My First Category Theory Textbook

Ralph Sarkis

Contents

0 Preliminaries 5
0.1 Abstract Algebra 5
0.2 Order Theory 11
0.3 Topology 14
1 Categories and Functors 19
1.1 Categories 19
1.2 Functors 27
1.3 Diagram Paving 34
2 Duality 37
2.1 Contravariant Functors 38
2.2 Opposite Category 40
2.3 Duality in Action 41
2.4 More Vocabulary 48
3 Limits and Colimits 53
3.1 Examples 53
3.2 Generalization 67
3.3 Diagram Chasing 74
4 Universal Properties 81
4.1 Examples 81
4.2 Generalization 93
4.3 Comma Categories 94
5 Natural Transformations 99
5.1 Functor Categories 99
5.2 The 2–category Cat 107
5.3 Equivalences 114

6 Yoneda Lemma 121

6.1 Representable Functors 121

6.2 Yoneda Lemma 126

6.3 Universality as Representability 130

7 Adjunctions 135

7.1 Equivalent Definitions 135

7.2 Results and Examples 141

8 Monads and Algebras 151

8.1 POV: Category Theory 151

8.2 POV: Universal Algebra 161

8.3 POV: Computer Programs 165

8.4 Exercises 169

9 Solutions to Exercises 171

9.1 Solutions to Chapter 1 171

9.2 Solutions to Chapter 2 172

9.3 Solutions to Chapter 3 174

9.4 Solutions to Chapter 4 177

9.5 Solutions to Chapter 5 179

9.6 Solutions to Chapter 6 181

9.7 Solutions to Chapter 7 181

9.8 Solutions to Chapter 8 182

0 Preliminaries

0.1 Abstract Algebra 5

0.2 Order Theory 11

0.3 Topology 14

Our main goal here is to introduce enough notation and terminology so that this book is self-contained.[0]

We assume you are familiar and comfortable with basic concepts about sets (e.g.: subsets, union, Cartesian product, cardinality, equivalence classes, quotients, etc.), functions (e.g.: injectivity, surjectivity, inverses, (pre)image, etc.), logic (e.g.: quantifiers, implication) and proofs (e.g.: you can write, read and understand proofs),[1] and we will not recall anything here. However, we need to have a little talk about foundations.

Several times in our coverage of category theory, we will use the term **collection** in order to avoid set-theoretical paradoxes. Collections are supposed to behave just like sets except that we will never consider collections containing other collections. We do not make it more formal because there are many ways to do it (dealing with so-called **size issues**),[2] and none of them are relevant to this course.

Still, you need to know why we cannot use sets as is usual in all other courses. In short, there exist collections of objects that cannot be sets.[3] In our case, we will need to talk about the collection of all sets and the collection of all groups (among others) and they cannot form sets. For the former, it is easy to see because if S is the set of all sets, then it contains all its subsets and hence $\mathcal{P}(S) \subseteq S$, this leads to the contradiction $|\mathcal{P}(S)| \leq |S| < |\mathcal{P}(S)|$.[4]

In the rest of this chapter, we cover the necessary background that we will use in the rest of the book. It is supposed to be a quick and (unfortunately) dry overview of stuff you may or may not have seen, so we will not dwell on explanations, intuitions and motivations.[5] You can safely skip these sections and come back whenever you click on a word or symbol that is defined here. We hope that this will save you from several trips to Wikipedia.

[0] Especially with the heavy use of the `knowledge` package, I felt it was necessary to cover enough background material in order to have the least amount of external links in the book.

[1] The very first things usually taught in early undergraduate mathematics courses.

[2] Most commonly, people use classes or Grothendieck universes. If this sticky point worries you, I suggest you keep it in the back of your mind and go read `https://arxiv.org/pdf/0810.1279.pdf` when you are a bit more comfortable with category theory.

[3] Famous examples include the collection of ordinal numbers which, by the Burali–Forti paradox, cannot be a set and the collection of all sets that do not contain themselves which, by the Russel paradox, cannot be a set.

[4] For a set X, $|X|$ denotes the **cardinal** of X and $\mathcal{P}(X)$ denotes the **powerset** of X, i.e. the set of all subsets of X. The strict inequality $|S| < |\mathcal{P}(S)|$ is due to Georg Cantor's famous diagonalization argument.

[5] Contrarily to the other chapters of this book.

0.1 Abstract Algebra

Here we recall definitions, examples and results you may have seen in classes on abstract algebra or linear algebra.[6]

[6] Monoids are not commonly covered, but they are simpler than groups and we need them at one point so we present them here.

Monoids

Definition 1 (Monoid)**.** A **monoid** is a set M equipped with a binary operation $\cdot : M \times M \to M$ (written infix) called **multiplication** and an **identity** element[7] 1_M satisfying for all $x, y, z \in M$

$$(x \cdot y) \cdot z = x \cdot (y \cdot z) \quad \text{and} \quad 1_M \cdot x = x = x \cdot 1_M.$$

If it satisfies $\forall x, y \in M, x \cdot y = y \cdot x$, M is a **commutative monoid**.

[7] Some authors call 1_M the **unit** or the **neutral** element.

Depending on the context, we will refer to a monoid either as M or $(M, \cdot)$ or $(M, \cdot, 1_M)$.

Remark 2. We will quickly drop the $\cdot$ symbol and denote multiplication with plain juxtaposition (i.e. $xy := x \cdot y$) for monoids and other algebraic structures with a multiplication.

Examples 3. 1. For any set S, the set of function from S to itself forms a monoid with the multiplication being composition of functions and the identity being the identity function $s \mapsto s$. We denote this monoid by S^S.

2. The sets $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ and $\mathbb{R}$[8] equipped with the operation of addition are all commutative monoids.

3. For any set S, the powerset $\mathcal{P}(S)$ has two simple monoid structures: one where the multiplication is $\cup$ and the identity is $\emptyset \subseteq S$, and the other where multiplication is $\cap$ and the identity is $S \subseteq S$.

[8] The symbols $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ and $\mathbb{R}$ denote respectively the sets of natural numbers, integers, rationals and real numbers.

Definition 4 (Submonoid)**.** Given a monoid M, a **submonoid** of M is a subset $N \subseteq M$ containing 1_M that is closed under multiplication (i.e. $\forall x, y \in N, x \cdot y \in N$).[9]

[9] This implies N is also a monoid with the multiplication and identity inherited from M.

Example 5. For any set S, the set of bijections from S to itself, denoted by Σ_S, is a submonoid of S^S because the composition of two bijections is bijective.

Definition 6 (Homomorphism)**.** Let M and N be two monoids, a **monoid homomorphism** from M to N is a function $f : M \to N$ satisfying the following property:

$$f(1_M) = 1_N \quad \text{and} \quad \forall x, y \in M, f(xy) = f(x)f(y).$$

When f is a bijection, we call it a **monoid isomorphism**, say that M and N are **isomorphic**, and write $M \cong N$.

Definition 7 (Kernel)**.** The **kernel** of a homomorphism $f : M \to N$ is the preimage of 1_N: $\ker(f) := f^{-1}(1_N)$. For any homomorphism f, $\ker(f)$ is a submonoid of M.[10]

[10] Similarly, the image of a homomorphism is also a submonoid.

Example 8. The inclusions $(\mathbb{N}, +) \to (\mathbb{Z}, +) \to (\mathbb{Q}, +) \to (\mathbb{R}, +)$ are all monoid homomorphisms with trivial kernel.[11] This implies this is also a chain of inclusions as submonoids.

[11] i.e. the kernel only contains the identity.

Definition 9 (Monoid action)**.** Let M be a monoid and S a set, an (left) **action** of M on S is an operation $\star : M \times S \to S$ satisfying for all $x, y \in M$ and $s \in S$

$$(x \cdot y) \star s = x \star (y \star s) \quad \text{and} \quad 1_M \star s = s.$$

The data $(M, S, \star)$ will also be called an M**–set** and we may refer to it abusively with S.

Any monoid action has a **permutation representation** defined to be the map

$$\sigma_\star : M \to S^S = x \mapsto (s \mapsto x \star s).$$

The properties of the action imply $\sigma_\star$ is a homomorphism. Conversely, given a homomorphism $\sigma : M \to S^S$ (i.e. $\sigma(1_M)$ is the identity function and $\sigma(xy) = \sigma(x) \circ \sigma(y)$ for any $x, y \in M$), there is a monoid action $\star_\sigma$ defined by $x \star_\sigma s = \sigma(x)(s)$.[12]

[12] These are inverse operations, i.e.

$$\sigma_{\star_\sigma} = \sigma \quad \text{and} \quad \star_{\sigma_\star} = \star.$$

Example 10. Any monoid M has a canonical left action on itself defined by $x \star m = xm$ for all $x, m \in M$.

Groups

Definition 11 (Group)**.** A **group** is set G equipped with a binary operation $\cdot : G \times G \to G$ called **multiplication**, an **inverse** operation $(-)^{-1} : G \to G$ and an **identity** element 1_G such that $(G, \cdot, 1_G)$ is a monoid and for all $x \in G$

$$x \cdot x^{-1} = 1_G = x^{-1} \cdot x.$$

If $(G, \cdot, 1_G)$ is a commutative monoid, we say that G is an **abelian group**.

Examples 12. 1. For any set S, we saw Σ_S was a submonoid of S^S, and it is in fact a group where the inverse of a function f is f^{-1} (it exists because f is bijective). We denote this group $\mathbf{\Sigma}_S$ and call it the group of **permutations** of S.[13]

[13] For $n \in \mathbb{N}$, Σ_n denotes the group of permutations of $\{1, \dots, n\}$.

2. The monoids on $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$ and $(\mathbb{R}, +)$ are also abelian groups with the inverse of x being $-x$.

3.

Definition 13 (Subgroup)**.** Given a group G, a **subgroup** of G is a submonoid H of G closed under taking inverses (i.e. $\forall x \in H, x^{-1} \in H$).[14]

[14] This implies H is also a group with the multiplication, inverse and identity inherited from G.

Example 14. For any group G and subset $S \subseteq G$, the subgroup **generated** by S inside G, denoted by $\langle S \rangle$ is the smallest subgroup containing S.[15]

[15] An explicit construction is

$$\langle S \rangle = \{x_1 \cdots x_n \mid n \in \mathbb{N}, x_1, \dots, x_n \in S \cup \{1_G\}\}.$$

Definition 15 (Homomorphism)**.** Let G and H be two groups, a **group homomorphism** from G to H is a monoid homomorphism $f : G \to H$. It follows that[16]

[16] For this, you need to show that inverses are unique.

$$\forall x \in G, f(x^{-1}) = f(x)^{-1}.$$

When f is a bijection, we call it a **group isomorphism**, say that G and H are **isomorphic**, and write $G \cong H$.

Example 16. For any group G and element $g \in G$, we call **conjugation** by g the homomorphism $\mathsf{c}_g : G \to G$ defined by $\mathsf{c}_g(x) = gxg^{-1}$.[17]

[17] It is a homomorphism as $g1_Gg^{-1} = gg^{-1} = 1_G$ and

$$gxyg^{-1} = gx1_Gyg^{-1} = gxg^{-1}gyg^{-1}.$$

Definition 17 (Kernel)**.** The **kernel** of a homomorphism $f : G \to H$ is the preimage of 1_H: $\ker(f) := f^{-1}(1_H)$. For any homomorphism f, $\ker(f)$ is a subgroup of G.[18]

[18] Similarly, the image of a homomorphism is also a subgroup.

Example 18. For any group G and element $g \in G$, $\ker(\mathsf{c}_g) = \{1_G\}$. Indeed, if $gxg^{-1} = 1_G$, conjugating by g^{-1} on both sides yields $x = 1_G$.

Definition 19 (Normal subgroup)**.** A subgroup N of G is called **normal** if for any $g \in G$ and $n \in N$, $gng^{-1} \in N$. In words, N is closed under conjugation by G. We write $N \lhd G$ when N is a normal subgroup of G.[19]

[19] The kernel of any homomorphism f is a normal subgroup as for any $h \in \ker f$ and any $g \in G$, we have

$$f(ghg^{-1}) = f(g)f(h)f(g)^{-1} = f(g)1f(g^{-1}) = 1.$$

Proposition 20. *For any subgroup H of G, the relation $\sim_H$ defined by*

$$g \sim_H g' \Leftrightarrow \exists h \in H, gh = g'$$

is an equivalence relation.

Proof. Any subgroup contains 1_G, so $g \sim_H g$ is witnessed by $g1_G = g$, hence $\sim_H$ is reflexive. If $gh = g'$, then $g = ghh^{-1} = g'h^{-1}$, thus $\sim_H$ is symmetric. If $gh = g'$ and $g'h' = g''$, then $ghh' = g''$ and since H is a subgroup $hh' \in H$, we conclude $\sim_H$ is transitive. □

Definition 21 (Quotient)**.** Let G be a group and N a normal subgroup of G, the multiplication of G is well-defined on equivalence classes of $\sim_N$, namely, if $g \sim_N g'$ and $h \sim_N h'$, then $gh \sim_N g'h'$.[20] The **quotient** G/N is the group whose elements are equivalence classes of $\sim_N$ with the multiplication $[g] \cdot [h] := [g \cdot h]$ and identity $1_{G/N} = [1_G]$ (where $[g]$ denotes the equivalence class of $\sim_N$ containing g).

[20] Suppose $gn = g'$ and $hn' = h'$ for $n, n' \in N$, then using the fact that $h^{-1}nh \in N$, we let $n'' := h^{-1}nhn' \in N$ and we find

$$g'h' = gnhn' = ghh^{-1}nhn' = ghn'',$$

thus $gh \sim_N g'h'$.

Definition 22 (Group action)**.** Let G be a group and S a set, an (left) **action** of G on S is a (left) monoid action of G on S. A set S equipped with action of G is called a G**–set**. It follows from the properties of an action that the function $s \mapsto g \star s$ is a bijection, hence the permutation representation $\sigma_\star$ is a homomorphism $G \to \Sigma_S$.

Example 23. Any group G has a canonical left action on itself defined by $x \star m = xm$ for all $x, m \in G$.

Definition 24 (Orbit)**.** Let S be a G–set, an **orbit** of S is a maximal subset of S closed under the action of G. Namely, it is a subset $A \subset S$ such that $g \star a \in A$ for any $g \in G$ and $a \in A$, and no subset strictly including A and strictly included in S ($A \subset A' \subset S$) has this property.

Rings

Definition 25 (Ring)**.** A **ring** is a set R equipped with a monoid structure $(R, \cdot, 1_R)$ and an abelian group structure $(R, +, 0_R)$[21] such that for all $x, y, z \in R$

[21] We call $\cdot$ the **multiplication** and $+$ the **addition** of the ring.

$$x \cdot (y + z) = (x \cdot y) + (x \cdot z).$$

If $(R, \cdot, 1_R)$ is a commutative monoid, we say that R is commutative.

Examples 26. 1. The abelian groups $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$ and $(\mathbb{R}, +)$ are also commutative rings with multiplication being the standard multiplication of numbers.

2. For any ring R and any $n \in \mathbb{N}$, the set of matrices $R^{n\times n}$ is a ring where addition is done pointwise, multiplication is the standard multiplication of matrices, $1_{R^{n\times n}}$ is the matrix with 1_R in each diagonal entry and 0_R everywhere else, and $0_{R^{n\times n}}$ is the matrix with 0_R everywhere.

Proposition 27. *Let R be a ring, for any $r \in R$, $0_R \cdot r = 0_R = r \cdot 0_R$.*

Proof. Here is the derivation for one equality (the other is symmetric):

$$0_R \cdot r = (1_R - 1_R) \cdot r = 1_R \cdot r - 1_R \cdot r = r - r = 0_R.$$

□

Definition 28 (Subring)**.** Given a ring R, a **subring** of R is a subset $S \subseteq R$ that is both a submonoid for $\cdot$ and a subgroup for $+$.[22]

[22] This implies S is also a ring with the multiplication and addition inherited from R.

Definition 29 (Homomorphism)**.** Let R and S be two rings, a **ring homomorphism** from R to S is a function $f : R \to S$ that is both a monoid homomorphism for the operation $\cdot$ and a group homomorphism for the operation $+$. Namely, it satisfies

$$\forall x, y \in R, f(x \cdot y) = f(x) \cdot f(y) \qquad f(1_R) = 1_S$$
$$\forall x, y \in R, f(x + y) = f(x) + f(y) \qquad f(0_R) = 0_S.$$

When f is a bijection, we call it a **ring isomorphism**, say that R and S are **isomorphic**, and write $R \cong S$.

Definition 30 (Kernel)**.** The **kernel** of a homomorphism $f : R \to S$ is the preimage of 0_S: $\ker f := f^{-1}(0_S)$. For any homomorphism, $\ker f$ is a subring of S.

As for monoids and groups, the image of a homomorphism is a subring, and as for groups the kernel satisfies an additional property: it is an ideal.

Definition 31 (Ideal)**.** Given a ring R, an **ideal** of R is a subring I such that for any $i \in I$ and $r, s \in R$, $ris \in I$.[23]

[23] An ideal is not only closed under multiplication but it is also preserved by multiplication by elements outside of the ideal.

Proposition 32. *For any subring S of R, the relation $\sim_S$ defined by*

$$r \sim_S r' \Leftrightarrow \exists s \in S, r + s = r'$$

is an equivalence relation.[24]

[24] Apply Proposition 20 to the group $(R, +)$ and its subgroup $(S, +)$.

Definition 33 (Quotient)**.** Let R be a ring and I be an ideal of R, the addition and multiplication of R are well-defined on equivalence classes of $\sim_I$, namely, if $r \sim_I r'$ and $s \sim_I s'$, then $r + s \sim_I r' + s'$ and $rs \sim_I r's'$.[25] The quotient R/I is the ring whos elements are equivalence classes of $\sim_I$ with the addition $[r] + [s] := [r + s]$, the multiplication $[r] \cdot [s] := [r \cdot s]$, $0_{R/I} := [0_R]$, and $1_{R/I} := [1_R]$.

[25] For addition, we can use the same proof as for quotient groups because I is a normal subgroup of $(R, +)$ (any subgroup of an abelian group is normal). For multiplication, suppose $r + i = r'$ and $s + j = s'$ for $i, j \in I$, then

$$r's' = (r + i)(s + j) = rs + rj + is + ij,$$

and since I is an ideal, $rj + is + ij \in I$. We conclude $rs \sim_I r's'$.

Definition 34 (Units)**.** An element of a ring is called a **unit** if it has a multiplicative inverse. Namely, $x \in R$ is a unit if there exists x^{-1} such that $xx^{-1} = 1_R = x^{-1}x$. We denote by $R^\times$ the set of units of R, it is a group with the multiplication inherited from R.

Example 35. The group of unit of $R^{n \times n}$ is called the **general linear group** over R and denoted by $\mathrm{GL}_n(R)$. It contains all the invertible[26] $n \times n$ matrices with entries in R.

[26] Sometimes called non-singular.

Proposition 36. *Any ring homomorphism $f : R \to S$ sends units of R to units of S.*[27]

Proof. If $x \in R$ has a multiplicative inverse x^{-1}, then the homomorphism properties imply

$$f(x)f(x^{-1}) = f(xx^{-1}) = f(1_R) = 1_S = f(1_R) = f(x^{-1}x) = f(x^{-1})f(x),$$

thus $f(x^{-1})$ is the multiplicative inverse of $f(x)$. □

[27] By restricting f to $R^\times$, we obtain a group homomorphism

$$f^\times : R^\times \to S^\times.$$

Fields

Definition 37 (Field)**.** A **field** is a commutative ring where every non-zero element is a unit.

Example 38. The rings $\mathbb{Q}$ and $\mathbb{R}$ are fields, but $\mathbb{Z}$ is not since the $\mathbb{Z}^\times = \{-1, 1\}$.

Definition 39 (Characteristic)**.** The **characteristic** of a field k is the minimum $n \in \mathbb{N}$ such that $1_k + \overset{n}{\cdots} + 1_k = 0_K$. If no such n exists, the characteristic of k is infinite.[28]

[28] One can show the characteristic of a field is never a composite number, it is either prime or infinite.

Examples 40. Fix a prime number p. The set $p\mathbb{Z}$ of multiples of p is an ideal of the ring $\mathbb{Z}$ and $\mathbb{Z}/p\mathbb{Z}$ is a field of characteristic p. The field $\mathbb{Q}$ has infinite characteristic.

Vector Spaces

Fix a field k.

Definition 41 (Vector space)**.** A **vector space** over k is a set an abelian group $(V, +, 0)$ along with an operation $\cdot : k \times V \to V$ called **scalar multiplication** such that the following holds for any $x, y \in k$ and $u, v \in V$:[29]

$$(xy) \cdot v = x \cdot (y \cdot v) \qquad 1 \cdot v = v$$
$$(x + y) \cdot v = x \cdot v + y \cdot v \qquad x \cdot (u + v) = x \cdot u + x \cdot v.$$

[29] We will not distinguish between the additions and zeros in k and V.

It follows that $0 \cdot v = 0$. We call elements of V **vectors**.

Example 42. For any $n \in \mathbb{N}$, the set k^n has a vector space structure, where addition and scalar multiplication are done pointwise, i.e.:

$$(u_1, \ldots, u_n) + (v_1, \ldots, v_n) = (u_1 + v_1, \ldots, u_n + v_n) \quad x \cdot (v_1, \ldots, v_n) = (xv_1, \ldots, xv_n).$$

Definition 43 (Subspace)**.** Given a vector space V, a **subspace** of V is a subset $W \subseteq V$ such that $0 \in W$, and for any $x \in k$ and $u, w \in W$, $x \cdot w \in W$ and $u + w \in W$.

Definition 44 (Linear map)**.** Let V and W be two vector spaces over k, a **linear map** from V to W is a function $T : V \to W$ satisfying

$$\forall x \in k, \forall u, v \in V, \quad T(x \cdot v) = x \cdot T(v) \qquad T(u + v) = T(u) + T(v).$$

When T is a bijection, we call it a **linear isomorphism**, say that R and S are **isomorphic**, and write $V \cong W$.

Definition 45 (Linear combination)**.** Let V be a vector space and $v_1, \ldots, v_n \in V$, a **linear combination** of these vectors is a sum

$$\sum_{i=1}^{n} a_i v_i = a_1 \cdot v_1 + \cdots + a_n v_n,$$

where $a_1, \ldots, a_n \in k$ are called the **coefficients**.

Definition 46 (Basis)**.** Let V be a vector space and $S \subseteq V$. We say that S is **linearly independent** if a linear combination of vectors in S is the zero vector if and only if all coefficients are zero. We say that S is **generating** if any $v \in V$ is a linear combination of vectors in S. We say that S is a **basis** of V if it is linearly independent and generating. The cardinality of a basis S of V is called the **dimension** of V.[30]

[30] Using the axiom of choice, one can show a basis always exists and all bases must have the same cardinality, hence the dimension of a vector space is well-defined.

Proposition 47. *A linear map $T : V \to W$ is completely determined by where it sends a basis of V.*

Proposition 48. *If a vector space V over k has dimension $n \in \mathbb{N}$, then $V \cong k^n$.*

Definition 49 (Dual)**.**

0.2 Order Theory

In this section, we briefly cover some early definitions and results from order theory. Since this subject is not usually taught in undergraduate courses, we spend a bit more time. In fact, we even introduce stuff we will not use later to make sure readers can get more familiar with the most important objects: posets and monotone functions.

Definition 50 (Poset)**.** A **poset** (short for partially ordered set) is a pair $(A, \leq)$ comprising a set A and a binary relation $\leq \subseteq A \times A$ that is

1. **reflexive** ($\forall x \in A, x \leq x$),
2. **transitive** ($\forall x, y, z \in A$ if $x \leq y$ and $y \leq z$ then $x \leq z$), and
3. **antisymmetric** ($\forall x, y \in A$ if $x \leq y$ and $y \leq x$ the $x = y$).

The relation is also called a **partial order**.[31]

[31] If antisymmetry is not satisfied, $\leq$ is called a **preorder**.

Examples 51. 1. The usual non-strict orders ($\leq$ and $\geq$) on $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ and $\mathbb{R}$ are all partial orders. The strict orders do not satisfy reflexivity.

2. The divisibility relation $\mid$ on $\mathbb{N}$ ($n \mid m$ if and only if n divides m) is a partial order.
3. For any set S, the powerset of S equipped with the subset relation ($\subseteq$) is a poset.
4. Any subset of a poset inherits a poset structure by restricting the partial order.

For any monoid M, there are three preorders defined by the so-called Green's relations:

$$\forall x, y \in M, x \leq_L y \Leftrightarrow \exists m \in M, x = my$$
$$\forall x, y \in M, x \leq_R y \Leftrightarrow \exists m \in M, x = ym$$
$$\forall x, y \in M, x \leq_J y \Leftrightarrow \exists m, m' \in M, x = mym'$$

Definition 52 (Monotone)**.** A function $f : (A, \leq_A) \to (B, \leq_B)$ between posets is **monotone** (or **order-preserving**) if for any $a, a' \in A$, $a \leq_A a' \implies f(a) \leq_B f(a')$.

Example 53. You probably already know lots of monotone functions, but let us give two less intuitive examples. Let $f : S \to T$ be a function, the **image map** of f[32] is the function $\mathcal{P}(S) \to \mathcal{P}(T)$ defined by $S \supseteq X \mapsto f(X) := \{f(x) \mid x \in X\}$. When both powersets are equipped with the inclusion partial order, the image map is monotone because $X \subseteq X' \subseteq S$ implies $f(X) \subseteq f(X')$.

The **preimage map** is

$$f^{-1} : \mathcal{P}(T) \to \mathcal{P}(S) = T \supseteq Y \mapsto f^{-1}(Y) := \{y \in S \mid f(y) \in Y\}.$$

It is also order-preserving because $Y \subseteq Y' \subseteq T$ implies $f^{-1}(Y) \subseteq f^{-1}(Y')$.

Proposition 54. *The composition of monotone functions between posets is monotone.*

Definition 55 (Dual)**.** The **dual order**[33] of a poset $(A, \leq)$, denoted by $(A, \leq)^{\text{op}}$, is the same set equipped with the converse relation $\geq$ defined by

$$\forall x, y \in A, x \geq y \Leftrightarrow y \leq x.$$

Definition 56 (Bounds)**.** Let $(A, \leq)$ be a poset and $S \subseteq A$, then $a \in A$ is an **upper bound** of S if $\forall s \in S, s \leq a$. Moreover, $a \in A$ is a **supremum** of S, if it is a least upper bound, that is, a is an upper bound of S and for any upper bound a' of S, $a \leq a'$. A supremum of S is denoted by $\vee S$, but when S contains only two elements, we use the infix notation $s_1 \vee s_2$ and call this a **join**.

A **lower bound** (resp. **infimum/meet**) of S is an upper bound (resp. supremum/join) of S in the dual order $(A, \leq)^{\text{op}}$.[34] An infimum of S is denoted by $\wedge S$ or $s_1 \wedge s_2$ in the binary case.

Proposition 57. *Infimums and supremums are unique when they exist.*[35]

Definition 58 (Complete lattice)**.** A **complete lattice** is a poset $(L, \leq)$ where every subset has a supremum and an infimum.[36] In particular, L has a smallest element $\vee\emptyset$ and a largest element $\wedge\emptyset$ (they are usually called **top** and **bottom** respectively).

Examples 59. 1. For any set S, $(\mathcal{P}(S), \subseteq)$ is a complete lattice. the supremum of a family of subsets is their union and the infimum is their intersection.

2. Defining supremums and infimums on the poset $(\mathbb{N}, |)$ is subtle. When $S \subseteq \mathbb{N}$ is non-empty, $\wedge S$ is the greatest common divisor of all elements in S and $\wedge\emptyset$ is 0 because any integer divides 0. For a finite and non-empty $S \subseteq \mathbb{N}$, $\vee S$ is the least common multiple of all elements in S. If S is infinite, then $\vee S$ is 0 and the supremum of the empty set is 1 because 1 divides any integer.

You might be wondering about possible posets where all infimums exist but not necessarily all supremums or vice-versa, it turns out that this is not possible as shown below.

Proposition 60. *Let $(L, \leq)$ be a poset, then the following are equivalent:*

(i) $(L, \leq)$ *is a complete lattice.*

[32] Which we abusively denote by f.

[33] This definition lets us avoid many symmetric arguments.

[34] Explicitly, $a \in A$ is a lower bound of S if $\forall s \in S, a \leq s$. It is an infimum of S if, in addition to being a lower bound of S, any lower bound a' of S satisfies $a' \leq a$.

[35] This holds by antisymmetry.

[36] Notice that, we can see $\vee$ and $\wedge$ as monotone maps from $(\mathcal{P}(L), \subseteq)$ to $(L, \leq)$.

(ii) *Any $S \subseteq L$ has a supremum.*

(iii) *Any $S \subseteq L$ has an infimum.*

Proof. (i) $\implies$ (ii), (i) $\implies$ (iii) and (ii) + (iii) $\implies$ (i) are all trivial. Also, by using duality, we only need to prove (ii) $\implies$ (iii).[37] For that, it suffices to note that, for any $S \subseteq L$, we can define $\wedge S$ to be the least upper bound for lower bounds of S. Formally,

$$\wedge S = \bigvee \{a \in L \mid \forall s \in S, a \leq s\}.$$

Defined that way, $\wedge S$ is a lower bound of S because if $s \in S$, then $s \geq a$ for every lower bound a of S, thus $\wedge S$, being the least upper bound of the lower bounds, is smaller than s. By definition, $\wedge S$ is greater than any other lower bound of S, hence it is indeed the infimum of S. □

[37] If this implication is true for any $(L, \leq)$, then it is true, in particular, for $(L, \geq)$. This implication for $(L, \geq)$ is equivalent to the converse implication for $(L, \leq)$.

Definition 61 (Fixpoints)**.** Let $f : (L, \leq) \to (L, \leq)$, a **pre-fixpoint** of L is an element $x \in L$ such that $f(x) \leq x$. A **post-fixpoint** is an element $x \in L$ such that $x \leq f(x)$. A **fixpoint** (or **fixed point**) of f is a pre- and post-fixpoint.

Theorem 62 (Knaester–Tarski[38])**.** *Let $(L, \leq)$ be a complete lattice and $f : L \to L$ be monotone.*

[38] This is actually a weaker version of the Knaester-Tarski theorem. The latter states that the fixpoints of a monotone function f form a complete lattice.

1. *The least fixpoint of f is the least pre-fixpoint $\mu f := \wedge\{a \in L \mid f(a) \leq a\}$.*

2. *The greatest fixpoint of f is the greatest post-fixpoint $\nu f := \vee\{a \in L \mid a \leq f(a)\}$.*

Proof. 1. Any fixpoint of f is in particular a pre-fixpoint, thus μf, being a lower bound of all pre-fixpoints, is smaller than all fixpoints. Moreover, because for any pre-fixpoint $a \in L$, $f(\mu f) \leq f(a) \leq a$, $f(\mu f)$ is also a lower bound of the pre-fixpoints, so $f(\mu f) \leq \mu f$. We infer that $f(f(\mu f)) \leq f(\mu f)$, so $f(\mu f)$ is a pre-fixpoint and $\mu f \leq f(\mu f)$. We conclude that μf is a fixpoint by antisymmetry.

The proof of the second item is the proof of the first item done in the dual order.

2. Any fixpoint of f is in particular a post-fixpoint, thus νf, being an upper bound of post-fixpoints, is bigger than all fixpoints. Moreover, because for any post-fixpoint $a \in L$, $a \leq f(a) \leq f(\nu f)$, $f(\nu f)$ is an upper bound of the post-fixpoints, so $\nu f \leq f(\nu f)$. We infer that $f(\nu f) \leq f(f(\nu f))$, so $f(\nu f)$ is a post-fixpoint and $f(\nu f) \leq \nu f$. We conclude that νf is a fixpoint by antisymmetry.

□

Definition 63 (Closure operator)**.** Let $(A, \leq)$ be a poset, a **closure operator** on A is a map $c : A \to A$ that is

1. monotone,

2. extensive ($\forall a \in A, a \leq c(a)$), and

3. idempotent ($\forall a \in A, c(a) = c(c(a))$).

Example 64. The floor ($\lfloor - \rfloor$) and ceiling ($\lceil - \rceil$) operations are closure operators on $(\mathbb{R}, \geq)$ and $(\mathbb{R}, \leq)$ respectively.

Definition 65 (Galois connection)**.** Given two posets $(A, \leq)$ and $(B, \sqsubseteq)$, a **Galois connection** is a pair of monotone functions $l : A \to B$ and $r : B \to A$ such that for any $a \in A$ and $b \in B$,

$$l(a) \sqsubseteq b \Leftrightarrow a \leq r(b).$$

For such a pair, we write $l \dashv r : A \to B$.

Proposition 66. *Let $l \dashv r : A \to B$ be a Galois connection, then l and r are monotone.*

Proof. Suppose $a \leq a'$, we will show $l(a) \sqsubseteq l(a')$. Since $l(a') \sqsubseteq l(a')$, using $\Rightarrow$ of the Galois connection yields $a' \leq r(l(a'))$, and, by transitivity, we have $a \leq r(l(a'))$. Then, using $\Leftarrow$ of the Galois connection, we find $l(a) \sqsubseteq l(a')$. We conclude that l is monotone.

A symmetric argument works to show r is monotone. □

Example 67.

Proposition 68. *Let $l \dashv r : A \to B$ be a Galois connection, then $r \circ l : A \to A$ is a closure operator.*

Proof. Since r and l are monotone, $r \circ l$ is monotone. Also, for any $a \in A$, $l(a) \sqsubseteq l(a)$ implies $a \leq r(l(a))$, so $r \circ l$ is extensive.

Now, in order to prove $r \circ l$ is idempotent, it is enough to show that[39]

$$r(l(a)) \geq r(l(r(l(a)))).$$

Observe that since $r(b) \leq r(b)$ for any $b \in B$, we have $l(r(b)) \leq b$, thus in particular, with $b = l(a)$, we have $l(r(l(a))) \leq l(a)$. Applying r which is monotone yields the desired inequality. □

[39] The $\leq$ inequality follows by extensiveness.

Proposition 69. *Let $l \dashv r : A \to B$ and $l' \dashv r : A \to B$ be Galois connections, then $l = l'$.*

Proposition 70. *Let $l \dashv r : A \to B$ and $l \dashv r' : A \to B$ be Galois connections, then $r = r'$.*

0.3 Topology

In this section, we introduce the basic terminology of topological spaces. Again we go a bit further than needed to help readers that first learn about topology here. We end this section by recalling some definitions about metric spaces.

Definition 71. A **topological space** is a pair (X, τ), where X is a set and $\tau \subseteq \mathcal{P}(X)$ is a family of subsets of X closed under arbitrary unions and finite intersections[40] whose elements are called **open sets** of X. We call τ a **topology** on X.

The **complement** of an open set U, denoted by U^c, is said to be **closed**.[41]

[40] For any family of open sets $\{U_i\}_{i \in I} \subseteq \tau$,

$$\bigcup_{i \in I} U_i \in \tau,$$

and if I is finite,

$$\bigcap_{i \in I} U_i \in \tau.$$

[41] Observe that both the empty set and the whole space are open and closed (sometimes referred to as **clopen**) because

$$\emptyset = \bigcup_{U \in \emptyset} U \text{ and } X = \bigcap_{U \in \emptyset} U \text{ and } \emptyset = X^c.$$

Example 72. On any set X, there are two trivial and extreme topologies.[42] The **discrete topology** $\tau_\top := \mathcal{P}X$ contains all the subsets of X. We can view $(X, \tau_\top)$ as a space where all points of X are separated from each other. The **codiscrete topology** $\tau_\bot := \{\emptyset, X\}$ contains only the subsets that must be open by definition of a topology. We can view $(X, \tau_\bot)$ as a space where all points of X are glued together with no space in-between.

[42] Trivial because

In the sequel, fix a topological space (X, τ).

Proposition 73. *Let $(C_i)_{i \in I}$ be a family of closed sets of X, then $\cap_{i \in I} C_i$ is closed and if I is finite, $\cup_{i \in I} C_i$ is also closed.*[43]

[43] This lemma gives an alternative to the axioms of Definition 71. Indeed, it is sometimes more convenient to define a topological space by giving its closed sets, and you can show the axioms about open sets still hold.

Proof. Both statements readily follow from DeMorgan's laws and the fact that the complement of a closed set is open and vice-versa. For the first one, DeMorgan's laws yield

$$\bigcap_{i \in I} C_i = \left(\bigcup_{i \in I} C_i^c \right)^c,$$

and the LHS is the complement of a union of opens, so it is closed. For the second one, DeMorgan's laws yield

$$\bigcup_{i \in I} C_i = \left(\bigcap_{i \in I} C_i^c \right)^c,$$

and the LHS is the complement of a finite intersection of opens, so it is closed. □

Proposition 74. *A subset $A \subseteq X$ is open if and only if for any $x \in A$, there exists an open $U \subseteq A$ such that $x \in U$.*

Proof. ($\Rightarrow$) For any $x \in A$, set $U = A$.

($\Leftarrow$) For each $x \in X$, pick an open $U_x \subseteq A$ such that $x \in A$, then we claim $A = \cup_{x \in A} U_x$ which is open[44]. The $\subseteq$ inclusion follows because each $x \in A$ has a set U_x in the union that contains x. The $\supseteq$ inclusion follows because each term of the union is a subset of A by assumption. □

[44] Arbitrary unions of opens are open.

Proposition 75. *A subset $A \subseteq X$ is closed if and only if for any $x \notin A$, there exists an open U such that, $x \in U$ and $U \cap A = \emptyset$.*[45]

[45] This result is simply a restatement of the last one by setting $A = A^c$.

Definition 76. Given $A \subseteq X$, the **closure** of A, denoted by $\overline{A}$ is the intersection of all closed sets containing A. One can show that $\overline{A}$ is the smallest closed set containing A.[46] Then, it follows that A is closed if and only if $\overline{A} = A$.

[46] $\overline{A}$ is closed because it is an intersection of closed sets and any closed sets containing A also contains $\overline{A}$ by definition.

Here are more easy results on the closure of a subset.

Proposition 77. *Given $A, B \subseteq X$ then the following statements hold:*

1. $A \subseteq B \implies \overline{A} \subseteq \overline{B}$
2. $A \subseteq \overline{A}$
3. $\overline{\overline{A}} = \overline{A}$
4. $\overline{\emptyset} = \emptyset$
5. $\overline{(A \cup B)} = \overline{A} \cup \overline{B}$

Proof of Lemma 77. 1. By definition, $\overline{B}$ contains B, thus A, but $\overline{B}$ is closed, so it must contain $\overline{A}$.

2. By definition.
3. $\overline{A}$ is closed, so its closure is itself.
4. 3 applied to $\emptyset$.
5. $\subseteq$ follows because the LHS is the smallest closed set containing $A \cup B$ and the RHS is closed and contains $A \cup B$.
   $\supseteq$: Since the RHS is closed, we have $\overline{(\overline{A} \cup \overline{B})} = \overline{A} \cup \overline{B}$ implying that the RHS is the smallest closed set containing $\overline{A} \cup \overline{B}$. Then, since the LHS is a closed set containing A and B, it contains $\overline{A}$ and $\overline{B}$ and hence must contain the RHS.

□

Remark 78. If we view $\mathcal{P}(X)$ as partial order equipped with the inclusion relation, the previous lemma is about good properties of the function $\overline{(-)} : \mathcal{P}(X) \to \mathcal{P}(X)$. Namely, we showed in the first three points that it is a monotone, extensive and idempotent, and therefore it is a closure operator.[47]

[47] In fact, this is where the terminology comes from.

Definition 79 (Dense)**.** A subset $A \subseteq X$ is said to be **dense** (in X) if any non-empty open set intersects A non-trivially, that is, $\forall \emptyset \neq U \in \tau, A \cap U \neq \emptyset$.

Proposition 80 (Decomposition)**.** *Let $A \subseteq X$, then $A = \overline{A} \cap (A \cup \left(\overline{A}\right)^c)$, where $\overline{A}$ is closed and $A \cup \left(\overline{A}\right)^c$ is dense. This results says that any subset of X can be decomposed into a closed and a dense set.*

Proof. The equality is clear[48] and $\overline{A}$ is closed by definition. It is left to show that $A \cup \left(\overline{A}\right)^c$ is dense. Let $U \neq \emptyset$ be an open set. If U intersects A, we are done. Otherwise, we have the following equivalences:

$$U \cap A = \emptyset \Leftrightarrow A \subseteq U^c \Leftrightarrow \overline{A} \subseteq U^c \Leftrightarrow U \subseteq \left(\overline{A}\right)^c,$$

where the second $\Rightarrow$ holds because U^c is closed. We conclude $U \cap \left(\overline{A}\right)^c \neq \emptyset$. □

[48] We use (in this order) distributivity of $\cap$ over $\cup$, the fact that a set and its complement intersect trivially and the inclusion $A \subseteq \overline{A}$:

$$\begin{aligned} \overline{A} \cap (A \cup (\overline{A})^c) &= (\overline{A} \cap A) \cup (\overline{A} \cap (\overline{A})^c) \\ &= A \cup \emptyset \\ &= A \end{aligned}$$

Proposition 81. *A subset $A \subseteq X$ is dense if and only if $\overline{A} = X$.*

Proof. ($\Rightarrow$) Since $\left(\overline{A}\right)^c$ is open but it intersects trivially the dense set A, it must be empty, thus $\overline{A}$ is the whole space.

($\Leftarrow$) Let U be an open set such that $U \cap A = \emptyset$, then A is contained in the closed set U^c, but this implies $\overline{A} \subseteq U^c$,[49] thus U is empty. □

[49] Recall that the closure of A is the smallest closed set containing A.

Definition 82 (Interior)**.** Let $A \subseteq X$, the **interior** of A, denoted by A^o is the union of all open sets contained in A. Similarly to the closure, we can check that that A^o is the largest open subset of A and thus that A is open if and only if $A = A^o$.[50]

[50] It also follows that $A \subseteq B \implies A^o \subseteq B^o$ and that $A^{oo} = A^o$.

We end this section by presenting a largely preferred way of defining a topology that avoid describing all open sets.

Definition 83 (Base)**.** Let X be a set, a **base** B is a set $B \subseteq \mathcal{P}(X)$ such that $X = \cup_{U \in B} U$ and any finite intersection of sets in B can be written as a union of sets in B.

Proposition 84. *Let X and $B \subseteq \mathcal{P}(X)$. If τ is the set of all unions of sets in B, then it is a topology on X. We say that τ is the topology **generated** by B.*

Proof. By assumption, we know that unions of opens are open and finite intersections of sets in B are open. It remains to show that finite intersections of unions of sets in B are also open. Let $U = \cup_{i \in I} U_i$ and $V = \cup_{j \in J} V_j$ with $U_i \in B$ and $V_j \in B$, then by distributivity, we obtain

$$U \cap V = \cup_{i \in I} U_i \bigcap \cup_{j \in J} V_j = \bigcup_{i \in I, j \in J} U_i \cap V_j,$$

so $U \cap V$ is open.[51] The lemma then follows by induction. □

[51] It is a union of opens.

In practice, instead of generating a topology from a base B, we start with any family $B_0 \subseteq \mathcal{P}(X)$ and let B be its closure under finite intersections, which satisfies the axioms of a base. Such a B_0 is often called a **subbase** for the topology generated by B.

Another very useful way to define topological spaces is to consider the topology induced by a metric.

Definition 85 (Metrics space). A **metric space** (X, d) is a set X together with a function $d : X \times X \to \mathbb{R}$ called a **metric** with the following properties for $x, y, z \in X$:

1. $d(x, y) \geq 0$
2. $d(x, y) = 0 \Leftrightarrow x = y$
3. $d(x, y) = d(y, x)$
4. $d(x, y) \leq d(x, z) + d(z, y)$

Definition 86 (Non-expansive). A function between metric spaces $f : (X, d_X) \to (Y, d_Y)$ is said to be **non-expansive**[52] if for all $x, x' \in X$,

$$d_Y(f(x), f(x')) \leq d_X(x, x').$$

[52] Also called 1–**Lipschitz** or **short**.

Proposition 87. *The composition of any two non-expansive maps is non-expansive.*

Definition 88 (Open ball). Let (X, d) be a metric space. Given a point $x \in X$ and a non-negative radius $r \in [0, \infty)$, the **open ball** of radius r centered at x is

$$B_r(x) := \{y \in X \mid d(x, y) < r.$$

Definition 89 (Induced topology). Any metric space (X, d) has an *induced topology* generated by the set of all open balls of X.[53]

In this topology, a set $S \subseteq X$ is open if and only if every point $x \in S$ is contained in an open ball which is contained in S.[54]

[53] This topology is sometimes called the open ball topology.

[54] Equivalently, $\forall x \in S, \exists r > 0, B_r(x) \subseteq S$.

Definition 90 (Convergence). Let (X, d) be a metric space, a sequence $\{p_n\}_{n \in \mathbb{N}} \subseteq X$ **converges** to $p \in X$ if

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall n \geq N, d(p_n, p) < \varepsilon.$$

Definition 91 (Cauchy sequence). Let (X, d) be a metric space, a sequence $\{p_n\}_{n \in \mathbb{N}} \subseteq X$ is called **Cauchy** if

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall m, n \geq N \implies d(p_n, p_m) < \varepsilon.$$

Definition 92 (Completeness). A metric space in which every Cauchy sequence converges is called **complete**.

1 Categories and Functors

1.1 Categories 19

1.2 Functors 27

1.3 Diagram Paving 34

As you will soon realize, many common mathematical objects can be viewed as categories or parts of a category, and often in several ways. Hence, there can be many starting points to motivate category theory even after restricting ourselves to the background of an undergraduate student in mathematics (see Chapter 0). I do not want to spend much time in the realm of informal explanations, so we will start from the notion of directed graphs, quickly get to the definition of a category and begin an enumeration of examples which will carry on (implicitly) for the rest of the book. We will also define functors which are to categories what homomorphisms are to groups (or rings, etc.), and list a bunch of examples.

1.1 Categories

Definition 93 (Directed graph)**.** A **directed graph** G consists of a collection of **nodes** or **objects** denoted G_0 and a collection of **arrows** or **morphisms** denoted G_1 along with two maps $s, t : G_1 \to G_0$, so that each arrow $f \in G_1$ has a **source** $s(f)$ and a **target** $t(f)$.

We draw a morphism as an arrow, the source being its tail and target being its head:

$$s(f) \xrightarrow{f} t(f)$$

Definition 94 (Paths)**.** A **path** in a directed graph G is a sequence of arrows $(f_1, \dots, f_k)$ that are **composable** in the sense that $t(f_i) = s(f_{i-1})$ for $i = 2, \dots, k$ as drawn below in (0). The collection of paths of length k in G will be denoted G_k.[55]

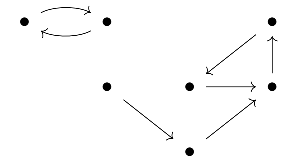
$$\bullet \xrightarrow{f_k} \bullet \xrightarrow{f_{k-1}} \bullet \cdots \bullet \xrightarrow{f_2} \bullet \xrightarrow{f_1} \bullet \qquad (0)$$

[55] The **length** of a path is the number of arrows in it. It is fitting that G_1 denotes the arrows of G and the paths of length 1 in G as they are the same thing.

Observe that when referring to a path as $(f_1, \dots, f_k)$ or drawing it as in (0), there is a mismatch in the ordering of the arrows. The order as drawn — also called the diagrammatic order — agrees with the usual notation in graph theory (the branch of mathematics concerned with studying graphs), and it is arguably a more intuitive representation of the word "path". The other order will be motivated when we will define the composition of arrows in a category. The main idea is that, conceptually, arrows coincide more closely with functions between mathematical objects, and if we see the arrows in (0) as functions, their composition is most of the time denoted by $f_1 \circ \cdots \circ f_k$.

Examples 95. It is very simple to give an example of a directed graph by drawing a bunch of nodes and arrows between them as in (1), G_0 is the collection of nodes,

G_1 is the collection of arrows and s and t can be inferred from looking at the head and tail of each arrow. Let us give more examples to motivate the next definition.



(1)

1. For any set X, there is a trivial directed graph with X as its collection of nodes and no arrows. The source and target maps are the unique functions $\varnothing \to X$. You can represent it by drawing a node for each element of X.[56]

   There is a slightly more complex directed graph whose nodes are the elements of X. For each pair $(x, x') \in X \times X$, we can add an arrow with source x and target x'. Drawing it is still fairly simple[57]: you draw a node for each element of X and an arrow from x to x' for each pair (x, x').[58]

2. Starting from a set X, we can define another directed graph by letting X be its only node and the collection of arrows be the set of functions from X to itself. The source and target maps are uniquely determined again, this time by their codomain that contains only the node X. This graph is already more interesting since the collection of arrows has a monoid structure. Indeed, the operation of composition of functions is associative, and the identity function is the identity for this operation.

3. Taking inspiration from the previous examples, we define a directed graph **Set**. It contains one node for every set, i.e., $\mathbf{Set}_0$ is the collection of all sets, [59] and one arrow with source X and target Y for every function $f : X \to Y$.

   Similarly to the last example, we recognize that the collection of arrows has a novel kind of structure induced by composition of functions and identity functions. It is not a monoid because you can only compose functions when one's source is the target of the other. In other words, composition of functions is not a binary operation $\circ : \mathbf{Set}_1 \times \mathbf{Set}_1 \to \mathbf{Set}_1$, it is of type $\mathbf{Set}_2 \to \mathbf{Set}_1$. Nonetheless, we still have associativity and identities which are at the core of the definition of a monoid. Since the theory of monoids is extremely rich and ubiquitous in mathematics, it is daring to study this seemingly more complex variant. We first need to make this structure abstract in the definition of a category.

[56] This is a very uninteresting directed graph.

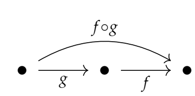
[57] Provided the set X is finite

[58] Note that there are so-called **loops** which are arrows from a node to itself because (x, x) is in $X \times X$.

[59] Notice how we could not have defined this graph if we required G_0 to be a set.

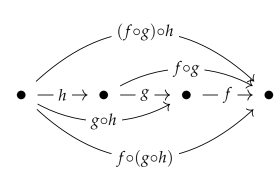
Definition 96 (Category)**.** A directed graph $\mathbf{C}$ along with a **composition** map $\circ : \mathbf{C}_2 \to \mathbf{C}_1$ is a **category** if it satisfies the following properties:

1. For any $(f, g) \in \mathbf{C}_2$, $s(f \circ g) = s(g)$ and $t(f \circ g) = t(f)$. This is more naturally understood visually in (2).



(2)

2. For any $(f, g, h) \in \mathbf{C}_3$, $f \circ (g \circ h) = (f \circ g) \circ h$, namely, composition is associative. Again, the graphic representation in (3) may be more revealing.



(3)

3. For any object $A \in \mathbf{C}_0$, there exists an **identity** morphism $u_{\mathbf{C}}(A) \in \mathbf{C}_1$ with A as its source and target that satisfies $u_{\mathbf{C}}(A) \circ f = f$ and $g \circ u_{\mathbf{C}}(A) = g$, for any $f, g \in \mathbf{C}_1$ where $t(f) = A$ and $s(g) = A$.

If the third property of Definition 96 is not satisfied, $\mathbf{C}$ is referred to as a **semicategory**. In rare occasions, authors choose to explicit when a category *does* satisfy this property, qualifying it as unital.

Remark 97 (Notation). In general, we will refer to categories with bold uppercase letters typeset with `\mathbf` ($\mathbf{C}$, $\mathbf{D}$, $\mathbf{E}$, etc.), their objects with uppercase letters (A, B, X, Y, Z, etc.) and their morphisms with lowercase letters (f, g, h, etc.). When the category is clear from the context, we denote the identity morphisms id_A instead of $u_{\mathbf{C}}(A)$. We say that two morphisms are **parallel** if they have the same source and target. Given morphisms f and g in a category $\mathbf{C}$, we say that f **factors through** g if there exists $h \in \mathbf{C}_1$ such that $f = g \circ h$ or $f = h \circ g$.

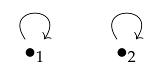
Observe that since $\circ$ is associative, it induces a unique composition map on paths of any finite lengths, which we abusively denote $\circ : \mathbf{C}_k \to \mathbf{C}_1$.[60] This lets us write $f_1 \circ f_2 \circ \cdots \circ f_k$ with no parentheses. Occasionally, we will refer to the image of the path under this map as the **composition of the path** or the **morphism that a path composes to**.

[60] Another abuse we make is to define $\circ : \mathbf{C}_0 \to \mathbf{C}_1$ by $X \mapsto \mathrm{id}_X$. That is, we identify objects of $\mathbf{C}$ with empty paths (of length 0) starting and ending at that object, and we consider the composition of an empty path to be the identity.

Examples 98 (Boring examples)**.** It can be really easy to construct a category by drawing its underlying directed graph and inferring the definition of the composition from it. Starting from the very simple graph depicted in (4), we can infer the definition of a category with a single object and its identity morphism. This category is denoted $\mathbf{1}$, the composition is trivial since $\mathrm{id}_\bullet \circ \mathrm{id}_\bullet = \mathrm{id}_\bullet$.


(4)

Similarly, we construct from the graph in (5) a category with two objects, their identity morphisms and nothing else. The composition is again trivial. This category will be denoted $\mathbf{1} + \mathbf{1}$.[61] More generally, for any collection $\mathbf{C}_0$, there is a category $\mathbf{C}$ whose collection of objects is $\mathbf{C}_0$ and whose collection of morphisms is $\mathbf{C}_1 := \{\mathrm{id}_X \mid X \in \mathbf{C}_0\}$. The composition map is completely determined by the third property in Definition 96.[62] A category without non-identity morphisms is called a **discrete category**.



(5)

[61] This notation is cleared up in Definition 216.

[62] i.e., for any $X \in \mathbf{C}_0$, $\mathrm{id}_X \circ \mathrm{id}_X = \mathrm{id}_X$.

The graph in (6) corresponds to the category with objects $\{A, B\}$ and morphisms $\{\mathrm{id}_A, \mathrm{id}_B, f\}$.

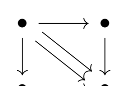
$$\mathrm{id}_A \circlearrowright A \xrightarrow{f} B \circlearrowleft \mathrm{id}_B \qquad (6)$$

The composition map is then completely determined by the properties of identity morphisms.[63] This category is called the interval category or the **walking arrow**, and it is denoted $\mathbf{2}$. Note however that $\mathbf{1} + \mathbf{1} \neq \mathbf{2}$.

[63] i.e., $f \circ \mathrm{id}_A = f$, $\mathrm{id}_B \circ f = f$, $\mathrm{id}_A \circ \mathrm{id}_A = \mathrm{id}_A$ and $\mathrm{id}_B \circ \mathrm{id}_B = \mathrm{id}_B$

Starting now, we will omit the identity morphisms from the diagrams (as is usual in the literature) for clarity reasons: they would hinder readability without adding information.

It is not always as straightforward to construct a category from a directed graph. For instance, if two distinct arrows have the same source and target, they must be explicitly drawn and the ambiguity in the composition must be dealt with. The graph in (7) is problematic in this way: it has two distinct paths of length two starting at the top-left corner and ending at the bottom-right corner. Since the composition of these paths can be equal to any of the two distinct morphisms between these corners, there is no category obviously corresponding to this graph.


(7)

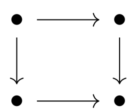
Since categories can be quite huge, it is rare that we draw all of a category at once. We will often draw diagrams with (labelled) nodes and arrows to represent the objects and morphisms within a category that we are focusing on. We also omit from our diagrams morphisms that can be inferred from the categorical structure. For instance, if we draw two composable morphisms as in (8), we do not draw the identity morphisms nor the composition $g \circ f$.

$$A \xrightarrow{f} B \xrightarrow{g} C \tag{8}$$

In many cases, not drawing all morphisms can lead to ambiguities like for (7). We have to be careful to avoid these, but sometimes we can resolve the problem by stating that the diagram is **commutative**.

Definition 99 (Commutativity)**.** Given a diagram representing objects and morphisms in a category, we say that it is **commutative** if the composition of any path of length greater than one is equal to the composition of any other path with the same source and target. The morphism resulting from the composition may or may not be depicted.

Examples 100. Arguably the most frequently used commutative diagram is the commutative square drawn in (9).

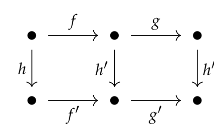


(9)

We say the square commutes when the bottom and top paths compose to the same (omitted in the diagram) morphism. The commutative square can also be seen as a category by inferring the missing morphism and the composition from commutativity. We can denote it $\mathbf{2} \times \mathbf{2}$.[64]

[64] This notation is explained in Definition 132.

Assuming that (10) commutes, we can infer that $f' \circ h = h' \circ f$, $g' \circ h' = h'' \circ g$, and $g' \circ f' \circ h = h'' \circ g \circ f$. Observe that the last equation can be derived from the first two which are equivalent to the commutativity of the two squares in (10). More generally, combining commutative diagrams in this way yields commutative diagrams, and this is the core of a powerful proof method called diagram paving that we introduce at the end of this chapter.



(10)

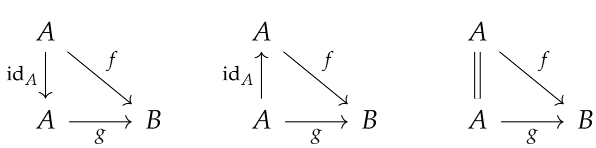
$$A \underset{g}{\overset{f}{\rightleftarrows}} B \tag{11}$$

Stating that (11) commutes is equivalent to stating that $f \circ g \circ f = f$ and $g \circ f \circ g = g$. We can also infer that $f \circ g \circ f \circ g = f \circ g$ and $g \circ f \circ g \circ f = g \circ f$, but this follows from the first two equality.

It would be odd to require that (7) commutes. It would imply that the two parallel morphisms are equal because they are both equal to the composition of the bottom and top paths. We will never draw parallel morphisms when they are supposed to be equal.

To assert that two morphisms $f, g : A \to B$ are equal using a diagram, we can say that either of the following is commutative, with a preference for the third one.[65]

[65] The equal sign in the third one can be read as id_A going in either direction.



(12)

Remark 101 (Convention). Reasoning with commutative diagrams is an acquired skill we will practice quite a lot in the following chapters. Yet there is no standard definition that everyone systematically uses.[66] For this reason, I decided to pick my favorite definition of commutativity which is uncommon.[67] In most cases, a diagram is called commutative when any two paths compose to the same morphism, but in practice, there are two exceptions handled by Definition 99:

1. Two parallel morphisms are not always equal in a commutative diagram. In fact, when parallel morphisms are drawn, it is usually to emphasize that they are distinct.

2. Unless otherwise stated, an endomorphism[68] drawn in a commutative diagram is not equal to the identity morphism (the composition of the empty path).

[66] This does not really lead to many misunderstandings anyway because what is meant by a diagram is usually made clear by the context.

[67] I have not seen the constraint on the length anywhere else.

[68] An **endomorphism** is a morphism whose source and target coincide.

Warning 102. Diagrams are not commutative by default. We will always specify when a diagram commutes. As our usage of commutative diagrams ramps up in the following chapters, you have to try to remember that.

Before moving on to more interesting categories, we introduce the Hom notation.

Definition 103 (Hom)**.** Let $\mathbf{C}$ be a category and $A, B \in \mathbf{C}_0$ be objects, the collection of all morphisms going from A to B is

$$\mathrm{Hom}_{\mathbf{C}}(A, B) := \{f \in \mathbf{C}_1 \mid s(f) = A \text{ and } t(f) = B\}.$$

This leads to an alternative way of defining the morphisms of $\mathbf{C}$, namely, one can describe $\mathrm{Hom}_{\mathbf{C}}(A, B)$ for all $A, B \in \mathbf{C}_0$ instead of describing $\mathbf{C}_1$ all at once. Defining the morphisms this way also takes care of the source and target functions implicitly.

Remark 104 (Notation). Some authors choose to denote the collection of morphisms between A and B with $\mathbf{C}(A, B)$. I prefer to use the latter notation when working with 2–categories[69] to highlight the fact that $\mathbf{C}(A, B)$ has more structure. Other authors use hom with a lowercase "h", my choice here is arbitrary.

[69] see Definition 370.

Definition 105 (Smallness)**.** A category $\mathbf{C}$ is called **small** if the collections of objects and morphisms are sets. If for all objects $A, B \in \mathbf{C}_0$, $\mathrm{Hom}_{\mathbf{C}}(A, B)$ is a set, $\mathbf{C}$ is said to be **locally small** and $\mathrm{Hom}_{\mathbf{C}}(A, B)$ is called a **hom-set**. A category that is not small can be referred to as **large**.

The following three examples will follow us throughout the book.

Example 106 (**Set**)**.** The category **Set** has the collection of sets as its objects and for any sets X and Y, $\mathrm{Hom}_{\mathbf{Set}}(X, Y)$ is the set of all the functions from X to Y.[70] The composition map is given by composition of functions (which is associative) and the identity maps serve as the identity morphisms. This category is locally small but not small.[71]

We will carry out many examples using **Set** because it is elaborate enough to be interesting, yet it is easy to understand because we are (assumed to be) very familiar with sets and functions.

[70] We already saw this directed graph in Example 95.3.

[71] By our argument at the start of Chapter 0: the collection of all sets cannot be a set.

Example 107. Let $(X, \leq)$ be a partially ordered set, it can be viewed as a category with elements of X as its objects. For any $x, y \in X$, the hom-set $\mathrm{Hom}_X(x, y)$ contains a single morphism if $x \leq y$ and is empty otherwise. The identity morphisms arise from the reflexivity of $\leq$. Since every hom-set contains at most one element and $\leq$ is transitive, the composition map is completely determined. Detailing this out, if $f : x \to y$ and $g : y \to z$ are morphisms, then we know that $x \leq y$ and $y \leq z$. Thus, transitivity implies that $x \leq z$ and there is a unique morphism $x \to z$, so it must be $g \circ f$.[72]

[72] Note that antisymmetry was not used in this argument, so one can more generally construct a category starting from a preorder. Such categories are called **thin** because each hom-set contains at most one morphism. It is straightforward to show the identities and composition ensure that any thin category $\mathbf{C}$ is constructed from the preorder $(\mathbf{C}_0, \leq)$ with

$$X \leq Y \Leftrightarrow \mathrm{Hom}_{\mathbf{C}}(X, Y) \neq \varnothing.$$

If a category corresponds to this construction for some poset, it is called **posetal**. In (13), we depict the posetal category associated to $(\mathbb{N}, \leq)$. The arrows between numbers n and $n + k$ are omitted for $k > 1$ as they can be inferred by the composition $n \leq n + 1 \leq n + 2 \leq \cdots \leq n + k$.

$$\overset{0}{\bullet} \longrightarrow \overset{1}{\bullet} \longrightarrow \overset{2}{\bullet} \longrightarrow \cdots \tag{13}$$

As a particular case of posetal categories, let (X, τ) be a topological space and note that the inclusion relation on open sets is a partial order on τ. Thus, X has a corresponding posetal category. More explicitly, the objects are open sets and for any $U, V \in \tau$, the hom-set $\mathrm{Hom}_X(U, V)$ contains the inclusion map i_{UV} if $U \subseteq V$ and is empty otherwise. This category will be denoted $\mathcal{O}(X, \tau)$ or $\mathcal{O}(X)$.

We will carry out many examples using posetal categories because it avoids difficulties arising from having different parallel morphisms.[73] In particular, every diagram drawn with objects and morphisms from a posetal category is commutative because the composition of any path is equal to the unique morphism between the source and target of that path. This also means some important aspects of a concept can be trivial when instantiating it for a posetal category.

[73] For the same reason, thin categories are also simple cases to carry out examples with.

Example 108 (Single object categories)**.** If a category $\mathbf{C}$ has a single object $*$, then all morphisms go from $*$ to $*$. In particular, $\mathbf{C}_1 = \mathrm{Hom}_{\mathbf{C}}(*, *)$ and $\mathbf{C}_2 = \mathbf{C}_1 \times \mathbf{C}_1$. Then, the associativity of $\circ$ and existence of id_* make $(\mathbf{C}_1, \circ)$ into a monoid.

Conversely, a monoid $(M, \cdot)$ can be represented by a single object category M, where $\mathrm{Hom}_M(*, *) = M$ and the composition map is the monoid operation.

Since many algebraic structures have an associative operation with an identity element, this yields a fairly general construction. The single object category associated to a monoid or group G will be denoted by $\mathbf{B}G$ and referred to as the **delooping** of G.

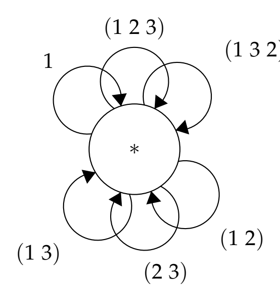


Figure 1.1: The delooping of the symmetric group S_3, a.k.a. $\mathbf{B}S_3$.

The natural numbers can also be endowed with the monoid structure of addition, hence a particular instance of a single object category is the delooping of $(\mathbb{N}, +)$. Notice that this category is very different from the posetal category $(\mathbb{N}, \leq)$. In the former, $\mathbb{N}$ is in correspondence with the morphisms while in the latter, it is in correspondence with the objects.

We will carry out many examples using deloopings of monoids or groups because it avoids difficulties arising from having two different objects.

Several simple examples of large categories arise as subcategories of **Set**.

Definition 109 (Subcategory)**.** Let $\mathbf{C}$ be a category, a category $\mathbf{C}'$ is a **subcategory** of $\mathbf{C}$ if, the following properties are satisfied.

1. The objects and morphisms of $\mathbf{C}'$ are objects and morphisms of $\mathbf{C}$ (i.e., $\mathbf{C}'_0 \subseteq \mathbf{C}_0$ and $\mathbf{C}'_1 \subseteq \mathbf{C}_1$).
2. The source and target maps of $\mathbf{C}'$ are the restrictions of the source and target maps of $\mathbf{C}$ on $\mathbf{C}'_1$ and for every morphism $f \in \mathbf{C}'_1$, $s(f), t(f) \in \mathbf{C}'_0$.
3. The composition map of $\mathbf{C}'$ is the restriction of the composition map of $\mathbf{C}$ on $\mathbf{C}'_2$ and for any $(f, g) \in \mathbf{C}'_2$, $f \circ_{\mathbf{C}'} g = f \circ_{\mathbf{C}} g \in \mathbf{C}'_1$.
4. The identity morphisms of objects in $\mathbf{C}'_0$ are the identity morphisms of objects in $\mathbf{C}_0$, i.e., $u_{\mathbf{C}}(A) = u_{\mathbf{C}'}(A)$ when $A \in \mathbf{C}'_0$.

Intuitively, one can see $\mathbf{C}'$ as being obtained from $\mathbf{C}$ by removing some objects and morphisms, but making sure that no morphism is left with no source or no target and that no path is left without its composition.

SOL **Exercise 110** (NOW!)**.** Find an example of a category $\mathbf{C}$ and a category $\mathbf{C}'$ that satisfy the first three conditions but not the fourth.

Definition 111 (Full and wide)**.** A subcategory $\mathbf{C}'$ of $\mathbf{C}$ is called **full** if for any objects $A, B \in \mathbf{C}'_0$, $\mathrm{Hom}_{\mathbf{C}'}(A, B) = \mathrm{Hom}_{\mathbf{C}}(A, B)$. It is called **wide** if $\mathbf{C}'_0 = \mathbf{C}_0$.[74]

[74] In words, a subcategory is full if the morphisms that were removed had their source or target removed as well, and it is wide if no objects were removed.

Examples 112 (Subcategories of **Set**)**.** We can selectively remove some objects and morphisms in **Set** to obtain the following categories.

1. Since the composition of injective functions is again injective, the restriction of morphisms in **Set** to injective functions yields a wide subcategory of **Set**, denoted by **SetInj**. Unsurprisingly, **SetSurj** can be constructed similarly.
2. Removing all infinite sets from **Set** yields the full subcategory of finite sets denoted **FinSet**.[75]
3. Further removing sets from **FinSet** and keeping only $\emptyset$, $\{1\}$, $\{1,2\}$, $\{1,2,3\}$, etc., we obtain the category **FinOrd** which is a small full subcategory of **Set**.[76]
4. Since the composition of monotone maps is monotone and the identity function is monotone, we can view each set $\{1, \ldots, n\}$ as ordered with $\leq$ and remove all morphisms that are not monotone from **FinOrd**. The resulting category is called the **simplex category** and denoted by $\mathbf{\Delta}$.

[75] This category is not small because there is no set of all finite sets.

[76] The name **FinOrd** is an abbreviation of finite ordinals, because we can also define **FinOrd** as the category of finite ordinals and functions between them.

Examples 113 (Concrete categories)**.** This second list of examples contains so-called concrete categories. Informally, they are categories of sets with extra structure, where morphisms are functions that preserve that extra structure.[77]

[77] Formally, see Definition 127.

1. The category $\mathbf{Set}_*$ is the category of **pointed** sets. Its objects are sets with a distinguished element, and its morphisms are functions that map distinguished elements to distinguished elements. The distinguished element of a pointed set

is the extra structure on top of the set, and morphisms between pointed set must preserve that structure. In more details, $(\mathbf{Set}_*)_0$ is the collection of pairs (X, x) where X is a set and $x \in X$, and for any two pointed sets (X, x) and (Y, y),

$$\mathrm{Hom}_{\mathbf{Set}_*}((X, x), (Y, y)) = \{f : X \to Y \mid f(x) = y\}.$$

The identity morphisms and composition are defined as in **Set**, so the axioms of a category clearly hold after checking that if $f : (X, x) \to (Y, y)$ satisfies $f(x) = y$ and $g : (Y, y) \to (Z, z)$ satisfies $g(y) = z$, then $(g \circ f)(x) = z$.

2. The category **Mon** is the category of monoids and their homomorphisms, let us be more explicit.[78] The objects are monoids, so $\mathbf{Mon}_0$ is the collection of all monoids, and the morphisms are monoid homomorphisms, so for any $M, N \in \mathbf{Mon}_0$, $\mathrm{Hom}_{\mathbf{Mon}}(M, N)$ is the set of homomorphisms from M to N. The composition in **Mon** is given by the composition of homomorphisms, we know it is well-defined because the composition of two homomorphisms is a homomorphism. Also, the composition is associative and the identity functions are homomorphisms, so we can define $u_{\mathbf{Mon}}(M) = \mathrm{id}_M$.

3. Similarly, the category of groups (resp. rings or fields) where the morphisms are group (resp. ring or field) homomorphisms is **Grp** (resp. **Ring** or **Field**). The category of abelian groups (resp. commutative monoids or rings) is a full subcategory of **Grp** (resp. **Mon** or **Ring**) denoted by **Ab** (resp. **CMon** or **CRing**).[79]

4. Let k be a fixed field, the category of vector spaces over k where the morphisms are linear maps is $\mathbf{Vect}_k$. The full subcategory of $\mathbf{Vect}_k$ consisting only of finite dimensional vector spaces is $\mathbf{FDVect}_k$.

5. The category of partially ordered sets where morphisms are order-preserving functions is denoted by **Poset**. It is a full subcategory of **Pre**, the category of preorders.

   A poset is a set A equipped with a binary relation $\leq \subseteq A \times A$ (the extra structure) that satisfies some axioms (reflexivity, transitivity and antisymmetry). In some sense, we can see the axioms as structure on top of the extra structure that is $\leq$. For example, we can consider the category **2Rel** of sets equipped with a binary relation (we do not require the axioms of posets to hold). An object of **2Rel** is a pair (A, R) where A is a set and $R \subseteq A \times A$ is a binary relation on A.[80] A morphism $(A, R_A) \to (B, R_B)$ is defined like order-preserving functions: it is a function $f : A \to B$ satisfying $\forall x, y \in A, (x, y) \in R_A \implies (f(x), f(y)) \in R_B$.

   The categories **Poset** and **Pre** are both full subcategories of **2Rel** where we only keep the relations satisfying the appropriate axioms.

6. The category of topological spaces where morphisms are continuous functions is denoted by **Top**.

7. The category of metric spaces where morphisms are nonexpansive functions is denoted by **Met**.

[78] These technicalities are essentially the same for the categories in the remainder of Example 113.

[79] Defining a category by saying it is a full subcategory of another one is a compact way of saying that we remove all the objects we do not want (e.g., the non-abelian groups) and nothing else.

[80] We use a nondescript letter for the relation instead of a symbol like $\leq$ to avoid being misled by the intuitions we have for partial orders.

In these last two examples, the choice of morphisms to take between spaces is not as clear cut as for the previous examples. For instance, one could ask the morphism between metric spaces to be continuous also, or for morphisms between topological spaces to map open sets to open sets (those are called open maps). In the end, the choice made depends on the context where the category is used.

SOL **Exercise 114.** An n–ary relation on a set A is a subset of A^n. Define the category n**Rel**.

Our next example is a large category that is neither a subcategory of **Set** nor a concrete category.

Example 115 (Rel). The category of sets and relations, denoted by **Rel**,[81] has as objects the collection of all sets, and for any sets X and Y, $\mathrm{Hom}_{\mathbf{Rel}}(X, Y)$ is the set of relations between X and Y, that is, the powerset of $X \times Y$. The composition of two relations $R \subseteq X \times Y$ and $S \subseteq Y \times Z$ is defined by

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

One can check that this composition is associative and that, for any set X, the **diagonal relation** $\Delta_X = \{(x,x) : x \in X\} \subseteq X \times X$ is the identity with respect to this composition.

[81] The notations for **Rel** and n**Rel** look close, but these categories see relations from very different points of view.

If you are not familiar with composition of relations, try to understand it visually. Draw the sets X, Y and Z as regions with dots inside, the relation R as wires connecting some dots in X and Y, and the relation S as wires connecting some dots in Y and Z. The relation $R\,;S$ relates a dot $x \in X$ to a dot $z \in Z$ if you can follow a wire in R and a wire in S to go from x to z.

Examples can also be helpful. Let $X = Y = Z$ be the set of all humans, R be the "cousin" relation (i.e., $(x,y) \in R$ whenever x and y are cousins) and S be the "sibling" relation. You can verify that $R;S = R$, $S;S = S$, but $R;R \neq R$.

Remark 116. You can view **Set** as a wide subcategory of **Rel** where you only take the relations $R \subseteq X \times Y$ satisfying for any $x \in X$,

$$|\,\{y \in Y \mid (x,y) \in R\}\,| = 1.$$

1.2 Functors

The list above is far from exhaustive; there are many more mathematical objects that can fit in a category and this is a main reason for studying this subject. Indeed, categories encapsulate a natural structure that accurately represents the heart of several mathematical theories from a global and abstract perspective.

If we were to develop category theory by mirroring the curriculum of most textbooks introducing abstract algebra, the rest of this chapter would be dedicated to exploring the insides of a category. We could talk about monomorphisms, epimorphisms, initial and terminal objects, subobjects, and even (co)limits inside a category. All these words will be defined in due time,[82] but not before explaining a guiding principle in category theory and setting an example by following it.

[82] Without relying on the rest of this chapter.

If we spend some more time studying Definition 96, we realize that the objects of a category carry little to no structure, and they are way less important than the morphisms. For example, the categories **Set**, **SetInj**, **SetSurj**, and **Rel** all have the same collection of objects, but they are very dissimilar.[83] As a matter of fact, there are alternative (albeit more messy) definitions of categories that do not refer to objects.

[83] We do not have enough tools yet to formally point out their differences.

Furthermore, a category only has superficial information about what its objects and morphisms are. For example, the category **Grp** is only a bunch of nodes and arrows, identities and a composition map. We cannot recover the definition of a group or a group homomorphism from that information. At first, this might seem detrimental: how can we prove things about groups if we do not know what they are? A good chunk of category theorists' mindset is contained in this snarky response.

We do not need to know what they are, only how they interact with each other.

As we advance through this book, we will get more sense of how true and powerful this idea can be.[84] We quickly start this journey by defining functors which are how categories interact with each other.

[84] One could argue the culminating point of this book (and any introduction to category theory) is the Yoneda lemma (see Chapter 6) which beautifully formalizes this idea.

Informally, a functor is a morphism of categories. Thus, to motivate the definition, we can look at other morphisms we have encountered. A clear similarity between categories like **Mon**, **Grp**, **Ring** or **Poset** is that all the objects are sets with some sort of structure that the morphisms preserve. In the first three categories, the structure on an object is the operations and identity elements that are preserved under homomorphisms, and in the last one, the structure on a poset is a relation that is preserved by order-preserving maps.[85] Hence, we go back to Definition 96, and we see that the structure of a category consists of the source and target maps, the composition map and the identities.

[85] Not all morphisms are functions that preserve structure, see e.g. morphisms in posetal categories.

Definition 117 (Functor)**.** Let $\mathbf{C}$ and $\mathbf{D}$ be categories, a **functor** $F : \mathbf{C} \rightsquigarrow \mathbf{D}$ is a pair of maps $F_0 : \mathbf{C}_0 \to \mathbf{D}_0$ and $F_1 : \mathbf{C}_1 \to \mathbf{D}_1$ such that diagrams (14), (15) and (16) commute where F_2 is induced by the definition of F_1 with $F_2 = (f, g) \mapsto (F_1(f), F_1(g))$.[86]

$$\begin{array}{ccccc} \mathbf{C}_0 & \xleftarrow{s} & \mathbf{C}_1 & \xrightarrow{t} & \mathbf{C}_0 \\ {\scriptstyle F_0}\downarrow & & \downarrow{\scriptstyle F_1} & & \downarrow{\scriptstyle F_0} \\ \mathbf{D}_0 & \xleftarrow[s]{} & \mathbf{D}_1 & \xrightarrow[t]{} & \mathbf{D}_0 \end{array} \tag{14}$$

$$\begin{array}{ccc} \mathbf{C}_2 & \xrightarrow{F_2} & \mathbf{D}_2 \\ {\scriptstyle \circ_{\mathbf{C}}}\downarrow & & \downarrow{\scriptstyle \circ_{\mathbf{D}}} \\ \mathbf{C}_1 & \xrightarrow[F_1]{} & \mathbf{D}_1 \end{array} \tag{15}$$

$$\begin{array}{ccc} \mathbf{C}_0 & \xrightarrow{F_0} & \mathbf{D}_0 \\ {\scriptstyle u_{\mathbf{C}}}\downarrow & & \downarrow{\scriptstyle u_{\mathbf{D}}} \\ \mathbf{C}_1 & \xrightarrow[F_1]{} & \mathbf{D}_1 \end{array} \tag{16}$$

[86] It is the first time we use commutative diagrams and we are already cheating a bit. Indeed, these diagrams do not represent objects and morphisms of a category we know. They could live in the category **Set** if $\mathbf{C}$ and $\mathbf{D}$ were small, but in the general case, we would need a category of collections and functions. It does not exist because there is no collection of all collections. Fortunately, this does not impact how we read these commutative diagrams.

Remark 118 (Digesting diagrams)*.* Once again, we emphasize that commutative diagrams will be heavily employed to make clearer and more compact arguments,[87] and that it will take time to get used to them. For now, let us unpack the definition above to ease its comprehension.

[87] This is especially true when using a blackboard or pen and paper because it makes it easier to point at things. Sadly, I cannot point at things on this PDF you are reading.

Commutativity of these diagrams is equivalent to having the following equalities:

$$s \circ F_1 = F_0 \circ s \qquad t \circ F_1 = F_0 \circ t \qquad F_1 \circ \circ_{\mathbf{C}} = \circ_{\mathbf{D}} \circ F_2 \qquad F_1 \circ u_{\mathbf{C}} = u_{\mathbf{D}} \circ F_0$$

Unrolling further, a functor $F : \mathbf{C} \rightsquigarrow \mathbf{D}$[88] must satisfy the following properties.

[88] The $\rightsquigarrow$ (`\rightsquigarrow`) notation for functors is not that common, they are usually denoted with plain arrows because they are morphisms. Nonetheless, I feel it is useful to have a special treatment for functors until you get accustomed to them. The squiggly arrow notation is sometimes used for Kleisli morphisms which we cover in Chapter 8.

i. For any $A, B \in \mathbf{C}_0$ and $f \in \mathrm{Hom}_{\mathbf{C}}(A, B)$, $F(f) \in \mathrm{Hom}_{\mathbf{D}}(F(A), F(B))$. This is equivalent to the commutativity of (14) which says $F_0(s(f)) = s(F_1(f))$ and $F_0(t(f)) = t(F_1(f))$.

ii. If $f, g \in \mathbf{C}_1$ are composable, then $F(f)$ and $F(g)$ are composable by i and $F(f \circ_{\mathbf{C}} g) = F(f) \circ_{\mathbf{D}} F(g)$ by commutativity of (15).

iii. If $A \in \mathbf{C}_0$, then $u_{\mathbf{D}}(F(A)) = F(u_{\mathbf{C}}(A))$ by commutativity of (16).[89]

[89] Alternatively, $\mathrm{id}_{F(A)} = F(\mathrm{id}_A)$.

The subscript on F is often omitted, as is common in the literature, when it is clear whether F is applied to an object or a morphism. We will also denote application of F with juxtaposition instead of parentheses, i.e., we can write FA and Ff instead of $F(A)$ and $F(f)$.

Examples 119 (Boring examples)**.** As usual, a few trivial constructions arise.

1. For any category $\mathbf{C}$, the **identity functor** $\mathrm{id}_{\mathbf{C}} : \mathbf{C} \rightsquigarrow \mathbf{C}$ is defined by letting $(\mathrm{id}_{\mathbf{C}})_0$ and $(\mathrm{id}_{\mathbf{C}})_1$ be identity maps on $\mathbf{C}_0$ and $\mathbf{C}_1$ respectively.

When the source and target of a functor coincide, we may refer to it as an **endofunctor**.

2. Let $\mathbf{C}$ be a category and $\mathbf{C}'$ a subcategory of $\mathbf{C}$, the **inclusion functor** $\mathcal{I} : \mathbf{C}' \rightsquigarrow \mathbf{C}$ is defined by letting $\mathcal{I}_0$ be the inclusion map $\mathbf{C}'_0 \hookrightarrow \mathbf{C}_0$ and $\mathcal{I}_1$ be the inclusion map $\mathbf{C}'_1 \hookrightarrow \mathbf{C}_1$.

3. Let $\mathbf{C}$ and $\mathbf{D}$ be categories and X be an object in $\mathbf{D}$, the **constant functor** $\Delta(X) : \mathbf{C} \rightsquigarrow \mathbf{D}$ sends every object to X and every morphism to id_X, i.e., $\Delta(X)_0(A) = X$ for any $A \in \mathbf{C}_0$ and $\Delta(X)_1(f) = \mathrm{id}_X$ for any $f \in \mathbf{C}_1$.

Examples 120 (Less boring)**.** Functors with the source being one of $\mathbf{1}$, $\mathbf{2}$ or $\mathbf{2} \times \mathbf{2}$[90] are a bit less boring. Let the target be a category $\mathbf{C}$ and let us analyze these functors.

[90] $\mathbf{2} \times \mathbf{2}$ is the commutative square in (9)

- Let $F : \mathbf{1} \rightsquigarrow \mathbf{C}$, F_0 assigns to the single object $\bullet \in \mathbf{1}_0$ an object $F(\bullet) \in \mathbf{C}_0$. Then, by commutativity of (16), F_1 is completely determined by $\mathrm{id}_\bullet \mapsto \mathrm{id}_{F(\bullet)}$. We conclude that functors of this type are in correspondence with objects of $\mathbf{C}$.
- Let $F : \mathbf{2} \rightsquigarrow \mathbf{C}$, F_0 assigns to A and B, two objects $FA, FB \in \mathbf{C}_0$ and F_1's action on identities is fixed. Still, there is one choice to make for $F_1(f)$ which must be a morphism in $\mathrm{Hom}_{\mathbf{C}}(FA, FB)$. Therefore, F sums up to a choice of two objects in $\mathbf{C}$ and a morphism between them. In other words, functors of this type are in correspondence with morphisms in $\mathbf{C}$.[91]
- Similarly (we leave the details as an exercise), functors of type $F : \mathbf{2} \times \mathbf{2} \rightsquigarrow \mathbf{C}$ are in correspondence with commutative squares inside the category $\mathbf{C}$.[92]

[91] After picking a morphism, the source and target are determined.

[92] i.e., pairs of pairs of composable morphisms $((f, g), (f', g')) \in \mathbf{C}_2 \times \mathbf{C}_2$ satisfying $f \circ g = f' \circ g'$.

Remark 121 (Functoriality)*.* We will use the term **functorial** as an adjective to qualify transformations that behave like functors and **functoriality** to refer to the property of behaving like a functor.

Throughout the rest of this book, the goal will essentially be to grow our list of categories and functors with more and more examples and perhaps exploit their properties wisely. Before pursuing this objective, we give important definitions analogous to injectivity and surjectivity of functions.

Definition 122 (Full and faithful)**.** Let $F : \mathbf{C} \rightsquigarrow \mathbf{D}$ be a functor. For $A, B \in \mathbf{C}_0$, denote the restriction of F_1 to $\mathrm{Hom}_{\mathbf{C}}(A, B)$ with

$$F_{A,B} : \mathrm{Hom}_{\mathbf{C}}(A, B) \to \mathrm{Hom}_{\mathbf{D}}(F(A), F(B)).$$

- If $F_{A,B}$ is injective for any $A, B \in \mathbf{C}_0$, then F is **faithful**.
- If $F_{A,B}$ is surjective for any $A, B \in \mathbf{C}_0$, then F is **full**.

- If $F_{A,B}$ is bijective for any $A, B \in \mathbf{C}_0$, then F is **fully faithful**.

SOL **Exercise 123** (NOW!). Show that the inclusion functor $\mathcal{I} : \mathbf{C}' \rightsquigarrow \mathbf{C}$ is faithful. Show it is full if and only if $\mathbf{C}'$ is a full subcategory.

SOL **Exercise 124.** Let $F : \mathbf{C} \rightsquigarrow \mathbf{D}$ and $G : \mathbf{D} \rightsquigarrow \mathbf{E}$. Show that

- if $G \circ F$ is faithful, then F is faithful, and
- if $G \circ F$ is full, then G is full.

As a generalization of the previous exercise, we note that a functor is full if and only if its image is a full subcategory of the target category.[93]

[93] The **image** of a functor $F : \mathbf{C} \rightsquigarrow \mathbf{D}$ is the subcategory of $\mathbf{D}$ containing all objects and morphisms in the image of F_0 and F_1.

Remark 125. While bijectivity is very strong to compare sets — it morally says that the elements of one set can be identified with the elements of another set — fully faithful functors are not as powerful. For instance, all functors between thin categories are fully faithful (because all the hom-sets are singletons). It should not be surprising that some fully faithful functors can be between two wildly unrelated categories because this property does not restrict the action on objects. We will see later what properties ensure that a functor strongly links the source and target category.

Examples 126. For all but the first example, we leave you to prove functoriality.[94] In the literature, a lot of functors are given only with their action on objects and the reader is supposed to figure out the action on morphisms. Not everyone has the same innate ability to do this, but I hope this book can give you enough experience to overcome this difficulty.

[94] It is an elementary task that is mostly relevant to the field of mathematics the functor comes from.

1. The **powerset functor** $\mathcal{P} : \mathbf{Set} \rightsquigarrow \mathbf{Set}$ sends a set X to its powerset $\mathcal{P}(X)$[95] and a function $f : X \to Y$ to the image map $\mathcal{P}(f) : \mathcal{P}(X) \to \mathcal{P}(Y)$. The latter sends a subset $S \subseteq X$ to
$$\mathcal{P}(f)(S) = f(S) := \{f(s) \mid s \in S\} \subseteq Y.$$
In order to prove that $\mathcal{P}$ is a functor, we need to show it makes diagrams (14), (15), and (16) commute. Equivalently, we can show it satisfies the three conditions in Remark 118.
   i. For any function $f : X \to Y$, the source and target of the image map $\mathcal{P}f$ are $\mathcal{P}X$ and $\mathcal{P}Y$ respectively as required.
   ii. Given two functions $f : X \to Y$ and $g : Y \to Z$, we can verify that $\mathcal{P}g \circ \mathcal{P}f = \mathcal{P}(g \circ f)$ by looking at the action of both sides on a subset $S \subseteq X$.
$$\begin{aligned}\mathcal{P}g(\mathcal{P}f(S)) &= \{g(y) \mid y \in \mathcal{P}f(S)\} \\ &= \{g(y) \mid y \in \{f(x) \mid x \in S\}\} \\ &= \{g(f(x)) \mid x \in S\}\end{aligned} \qquad \begin{aligned}\mathcal{P}(g \circ f)(S) &= \{(g \circ f)(x) \mid x \in S\} \\ &= \{g(f(x)) \mid x \in S\}\end{aligned}$$
   iii. Finally, the image map of id_X is the identity on $\mathcal{P}X$ because
$$\mathcal{P}\mathrm{id}_X(S) = \{\mathrm{id}_X(x) \mid x \in S\} = \{x \mid x \in S\} = S.$$

[95] The powerset of X is the set of all subsets of X.

The powerset functor is faithful because the same image map cannot arise from two different functions[96], it is not full because lots of functions $\mathcal{P}(X) \to \mathcal{P}(Y)$ are not image maps. A cardinality argument suffices: when $|X|, |Y| \geq 2$,

$$|\mathrm{Hom}_{\mathbf{Set}}(X,Y)| = |Y|^{|X|} < |\mathcal{P}(Y)|^{|\mathcal{P}(X)|} = |\mathrm{Hom}_{\mathbf{Set}}(\mathcal{P}(X), \mathcal{P}(Y))|.$$

[96] Indeed, if $f(x) \neq g(x)$, then $f(\{x\}) \neq g(\{x\})$.

2. The concrete categories of Examples 113 are defined using a functor.

   Definition 127 (Concrete category)**.** We call a category **C concrete** if it is paired (generally implicitly) with a faithful functor $U : \mathbf{C} \rightsquigarrow \mathbf{Set}$. In most cases, U is called the **forgetful functor** because it sends objects and morphisms of **C** to sets and functions by *forgetting* additional structure.

   The forgetful functor $U : \mathbf{Grp} \rightsquigarrow \mathbf{Set}$ sends a group $(G, \cdot, 1_G)$ to its underlying set G, *forgetting about the operation and identity*. It sends a group homomorphism $f : G \to H$ to the underlying function, *forgetting about the homomorphism properties*. It is faithful since if two homomorphisms have the same underlying function, then they are equal.[97]

   [97] We leave you the repetitive task to describe the forgetful functor for every concrete category in Examples 113.

   Briefly, functoriality of U follows from the facts that the underlying function of a homomorphism $f : G \to H$ goes between the underlying sets of G and H, the underlying function of a composition of homomorphisms is the composition of the underlying functions, and the underlying function of the identity homomorphism is the identity map.

3. It is also sometimes useful to consider *intermediate* forgetful functors. For example, $U : \mathbf{Ring} \rightsquigarrow \mathbf{Ab}$ sends a ring $(R, +, \cdot, 1_R, 0_R)$ to the abelian group $(R, +, 0_R)$, *forgetting about multiplication and* 1_R. It sends a ring homomorphism $f : R \to S$ to the same underlying function seen as a group homomorphism.[98] Not any old functor $\mathbf{Ring} \rightsquigarrow \mathbf{Ab}$ can be considered an intermediate forgetful functor. The key property is that forgetting about multiplication and 1_R ($\mathbf{Ring} \rightsquigarrow \mathbf{Ab}$) and then forgetting about the addition and 0_R ($\mathbf{Ab} \rightsquigarrow \mathbf{Set}$) is the same thing as forgetting all the ring structure at once ($\mathbf{Ring} \rightsquigarrow \mathbf{Set}$).

   [98] It can do that because part of the requirements for ring homomorphisms is to preserve the underlying additive group structure.

   The inclusion functor of **Poset** into 2**Rel** is also an intermediate forgetful functor. It forgets about all the properties of the partial order, but it does not forget about the binary relation.

4. In some cases, there is a canonical way to go in the opposite direction to the forgetful functor, it is called the free functor. For **Mon**, the free functor $F : \mathbf{Set} \rightsquigarrow \mathbf{Mon}$ sends a set X to the free monoid generated by X and a function $f : X \to Y$ to the unique group homomorphism $F(X) \to F(Y)$ that restricts to f on the set of generators.[99]

   [99] More details about free monoids are in Chapter 4.

   In Chapter 7, when covering adjunctions, we will study a strong relation between the forgetful functor U and the free functor F that will generalize to other mathematical structures.

5. Let $(X, \leq)$ and $(Y, \sqsubseteq)$ be posets, and $F : X \rightsquigarrow Y$ be a functor between their posetal categories. For any $a, b \in X$, if $a \leq b$, then $\mathrm{Hom}_X(a,b)$ contains a single element, thus $\mathrm{Hom}_Y(F(a), F(b))$ must contain a morphism as well,[100] or equivalently $F(a) \sqsubseteq F(b)$. This shows that F_0 is an order-preserving function on the posets.

   Conversely, any order-preserving function between X and Y will correspond to a unique functor as there is only one morphism in all the hom-sets.[101]

[100] The image of the element in $\mathrm{Hom}_X(a,b)$ under F.

[101] Given $f : (X, \leq) \to (Y, \sqsubseteq)$ order-preserving, the corresponding functor between the posetal categories of X and Y acts like f of the objects and sends a morphism $a \to b$ to the unique morphism $f(a) \to f(b)$ which exists because $a \leq b \implies f(a) \sqsubseteq f(b)$.

SOL **Exercise 128.** Let A and B be two sets, their powersets can be seen as posets with the order $\subseteq$. Thus, we can view $\mathcal{P}(A)$ and $\mathcal{P}(B)$ as posetal categories.

- Draw (using points and arrows) the category corresponding to $\mathcal{P}(\{0,1,2\})$.
- Show that the image and preimage functions defined below are functors between these categories.[102]

$$f : \mathcal{P}(A) \to \mathcal{P}(B) = S \mapsto \{f(a) \mid a \in S\}$$
$$f^{-1} : \mathcal{P}(B) \to \mathcal{P}(A) = S \mapsto \{a \in A \mid f(a) \in S\}$$

[102] i.e., they are order-preserving functions.

6. Let G and H be groups and $\mathbf{B}G$ and $\mathbf{B}H$ be their respective deloopings, then the functors $F : \mathbf{B}G \rightsquigarrow \mathbf{B}H$ are exactly the group homomorphisms from G to H.[103] Let $F : \mathbf{B}G \rightsquigarrow \mathbf{B}H$ be a functor, the action of F on objects is trivial since there is only one object in both categories. On morphisms, F_1 is a function from G to H which preserves composition and the identity morphism which, by definition, are the group multiplication and identity respectively. Thus, F_1 is a group homomorphism.

   Given a homomorphism $f : G \to H$, the reverse reasoning shows we obtain a functor $\mathbf{B}G \rightsquigarrow \mathbf{B}H$ by acting trivially on objects and with f on morphisms.

[103] Similarly for the deloopings of monoids.

7. For any group G, the functors $F : \mathbf{B}G \rightsquigarrow \mathbf{Set}$ are in correspondence with left actions of G. Indeed, if $S = F(*)$, then

$$F_1 : G = \mathrm{Hom}_{\mathbf{B}G}(*,*) \to \mathrm{Hom}_{\mathbf{Set}}(S,S)$$

   is such that $F(gh) = F(g) \circ F(h)$ for any $g, h \in G$ and $F(1_G) = \mathrm{id}_S$.[104] Moreover, since for any $g \in G$,

$$F(g^{-1}) \circ F(g) = F(g^{-1}g) = F(1_G) = \mathrm{id}_S = F(1_G) = F(gg^{-1}) = F(g) \circ F(g^{-1}),$$

   the function $F(g)$ is a bijection (its inverse is $F(g^{-1})$) and we conclude F_1 is the permutation representation of the group action defined by $g \star s = F(g)(s)$ for all $g \in G$ and $s \in S$.

   Given a group action on a set S, we leave you to show that letting $F_0 = * \mapsto S$ and F_1 be the permutation representation of the action yields a functor $F : \mathbf{B}G \rightsquigarrow \mathbf{Set}$.

[104] This is because gh is the composite of g and h in $\mathbf{B}G$ and 1_G is the identity morphism in $\mathbf{B}G$.

8. In the previous example, replacing $\mathbf{Set}$ with $\mathbf{Vect}_k$, one obtains k-linear representations of G instead of actions of G.[105]

[105] You might not know about linear representations, we just mention them in passing.

Remark 129 (Non-examples). From this long (and yet hardly exhaustive) list, one might get the feeling that every important mathematical transformation is a functor. This is not the case, so I wanted to show where functoriality can fail and hopefully give you a bit of intuition about why they fail. Here are two instances showcasing the two most common ways (in my experience) you can decide that a mapping is not functorial.

Let us define $F : \mathbf{FDVect}_k \rightsquigarrow \mathbf{Set}$ which assigns to any vector space over k a choice of basis. There is no non-trivializing way to define an action of F on linear maps which make F into a functor. One informal reason for this failure is that we cannot choose bases globally, so F is defined locally and its parts cannot be glued together.[106]

Another non-example is given by the center[107] of a group in **Grp**. A homomorphism $H \to G$ does not necessarily send the center of H in the center of G (take for instance $S_2 \hookrightarrow S_3$), thus, we cannot easily define the function $Z(H) \to Z(G)$ induced by the homomorphism (unless we send everything to $1_G \in Z(G)$). This time, Z is not a functor because it does not interact well with the morphisms of the category. Actually, if you decided to only keep group isomorphisms in the category, you could define the functor Z because isomorphisms preserve the center of groups.

[106] If you feel like you are making a non-canonical choice for every object, there is a good chance you are not dealing with a functor.

[107] The **center** of a group G, often denoted $Z(G)$, is the subset of G containing elements that commute with all other elements, i.e.,

$$Z(G) = \{x \in G \mid \forall g \in G, xg = gx\}.$$

In this chapter, we introduced a novel structure, namely categories, that functors preserve.[108] Since we also introduced several categories where objects had some structure that morphisms preserve, it is reasonable to wonder whether categories and functors are also part of a category. In fact, the only missing ingredient is the composition of functors (we already know what the source and target of a functor is and every category has an identity functor). After proving the following proposition, we end up with the category **Cat** where objects are small categories and morphisms are functors.[109]

[108] We defined functors precisely so that they preserve the structure of categories.

[109] In order to avoid paradoxes of the Russel kind, it is essential to restrict **Cat** to contain only small categories.

Proposition 130. *Let $F : \mathbf{C} \rightsquigarrow \mathbf{D}$ and $G : \mathbf{D} \rightsquigarrow \mathbf{E}$ be functors and $G \circ F : \mathbf{C} \rightsquigarrow \mathbf{E}$ be their* ***composition*** *defined by $G_0 \circ F_0$ on objects and $G_1 \circ F_1$ on morphisms. Then, $G \circ F$ is a functor.*

Proof. One could proceed with a really hands-on proof and show that $G \circ F$ satisfies the three necessary properties in a manner not unlike when proving the group homomorphisms compose. This should not be too hard, but you will have to deal with notation for objects, morphisms and the composition from all three different categories. This can easily lead to confusion or worse: boredom!

Instead, we will use the diagrams we introduced in the first definition of a functor. From the functoriality of F and G, we get two sets of three diagrams and combining them yields the diagrams for $G \circ F$.[110]

[110] Since F is a functor, the top two squares of (17) and the left squares of (18) and (19) commute. Since G is a functor, the bottom two squares (17) and the right squares of (18) and (19) commute.

$$\begin{array}{ccccc} \mathbf{C}_0 & \xleftarrow{s} & \mathbf{C}_1 & \xrightarrow{t} & \mathbf{C}_0 \\ {\scriptstyle F_0}\downarrow & & \downarrow{\scriptstyle F_1} & & \downarrow{\scriptstyle F_0} \\ \mathbf{D}_0 & \xleftarrow{s} & \mathbf{D}_1 & \xrightarrow{t} & \mathbf{D}_0 \\ {\scriptstyle G_0}\downarrow & & \downarrow{\scriptstyle G_1} & & \downarrow{\scriptstyle G_0} \\ \mathbf{E}_0 & \xleftarrow{s} & \mathbf{E}_1 & \xrightarrow{t} & \mathbf{E}_0 \end{array} \tag{17}$$

$$\begin{array}{ccccc} \mathbf{C}_2 & \xrightarrow{F_2} & \mathbf{D}_2 & \xrightarrow{G_2} & \mathbf{E}_2 \\ {\scriptstyle \circ_{\mathbf{C}}}\downarrow & & \downarrow{\scriptstyle \circ_{\mathbf{D}}} & & \downarrow{\scriptstyle \circ_{\mathbf{E}}} \\ \mathbf{C}_1 & \xrightarrow[F_1]{} & \mathbf{D}_1 & \xrightarrow[G_1]{} & \mathbf{E}_1 \end{array} \tag{18}$$

$$\begin{array}{ccccc} \mathbf{C}_0 & \xrightarrow{F_0} & \mathbf{D}_0 & \xrightarrow{G_0} & \mathbf{E}_0 \\ {\scriptstyle u_{\mathbf{C}}}\downarrow & & \downarrow{\scriptstyle u_{\mathbf{D}}} & & \downarrow{\scriptstyle u_{\mathbf{E}}} \\ \mathbf{C}_1 & \xrightarrow[F_1]{} & \mathbf{D}_1 & \xrightarrow[G_1]{} & \mathbf{E}_1 \end{array} \tag{19}$$

To finish the proof, you need to convince yourself that combining commutative diagrams in this way yields commutative diagrams. We proceed with a proof by example. Take diagram (19), we know the left and right square are commutative because F and G are functors. To show that the rectangle also commutes, we need to show the top path and bottom path from $\mathbf{C}_0$ to $\mathbf{E}_1$ compose to the same function. Here is the derivation:[111]

[111] In this case, both the diagram and the derivation are fairly simple. This will not stay true in the rest of the book, but the complexity of diagrams will grow way slower than the complexity of derivations, and we will mostly omit the latter for this reason.

$$\begin{aligned} G_1 \circ F_1 \circ u_{\mathbf{C}} &= G_1 \circ u_{\mathbf{D}} \circ F_0 && \text{left square commutes} \\ &= u_{\mathbf{E}} \circ G_0 \circ F_0 && \text{right square commutes} \end{aligned}$$ □

The category **Cat** is a concrete category. Intuitively, it is because categories are sets with extra sturcture that functors preserve. Rigorously, there is a forgetful functor $\mathbf{Cat} \rightsquigarrow \mathbf{Set}$.

SOL **Exercise 131** (NOW!)**.** Show that both assignments $\mathbf{C} \mapsto \mathbf{C}_0$ and $\mathbf{C} \mapsto \mathbf{C}_1$ yield functors $\mathbf{Cat} \rightsquigarrow \mathbf{Set}$.[112] Their action on morphism of categories (functors) is straightforward: the first sends F to F_0 and the second sends F to F_1. Show that the functor $(-)_0$ is not faithful, but $(-)_1$ is.

[112] Recall that we assumed $\mathbf{C} \in \mathbf{Cat}$ is small, meaning both $\mathbf{C}_0$ and $\mathbf{C}_1$ are sets.

We will often use $-$ as a **placeholder** for an input so the latter remains nameless. For instance, $f(-,-)$ means f takes two inputs. The type of the inputs and outputs will be made clear in the context.

This last exercise suggests we should view a category as a set of morphisms with extra structure. However, Definition 96 reveals we can also see a category as a directed graph with extra structure. We can make this formal by first defining the category **DGph** whose objects are small directed graphs and morphisms are functors without the requirement of (15) and (16).[113] There is a functor $\mathbf{Cat} \rightsquigarrow \mathbf{DGph}$ that simply forgets about composition and identities.

[113] Explicitly, a morphism $G \to G'$ is a pair of functions $F_0 : G_0 \to G'_0$ and $F_1 : G_1 \to G'_1$ satisfying for any $f \in G_1$, $F_0(s(f)) = s(F_1(f))$ and $F_0(t(f)) = t(F_1(f))$. Less cryptically, it is a mapping from objects to objects and arrows to arrows such that an arrow $A \to B$ is mapped to an arrow $F_0A \to F_0B$.

Since functors are also a new structure, one might expect that there are transformations between functors that preserve it. It is indeed the case, they are called natural transformations and they are the main subject of Chapter **??**. Moreover, although we will not cover it, there is a whole tower of abstraction that one could build in this way, and it is the subject of study of higher category theory.

1.3 Diagram Paving

If you are in awe at how wonderful the diagrammatic proof of Proposition 130, this section is for you. We introduce the proof technique called **diagram paving**[114] and

[114] Usually, diagram paving refers to a more general version of what I will show you. That technique is used in higher category theory.

set up some exercises for practice.

The key idea in that proof is that combining commutative diagrams yields commutative diagrams.[115] In general, paving a diagram that we want to show commutes is the process of progressivelly adding more objects and morphisms to obtain multiple diagrams we know (by hypothesis or previous lemmas) commute that combine into the original one.

[115] The term "combining" is not precisely defined, our intuition of what it means should be enough.

Let us clarify by example. In the setting of Proposition 130, to show that $G \circ F$ is a functor, we need to prove (14) instantiated with $G \circ F$ is commutative.[116] It is drawn in (20).

[116] We only do the first diagram.

$$\begin{array}{ccccc} \mathbf{C}_0 & \xleftarrow{s} & \mathbf{C}_1 & \xrightarrow{t} & \mathbf{C}_0 \\ {\scriptstyle (G\circ F)_0}\downarrow & & \downarrow{\scriptstyle (G\circ F)_1} & & \downarrow{\scriptstyle (G\circ F)_0} \\ \mathbf{E}_0 & \xleftarrow[s]{} & \mathbf{E}_0 & \xrightarrow[t]{} & \mathbf{E}_0 \end{array} \tag{20}$$

We can factor the action of $G \circ F$ and draw (21). We indicated with $\circlearrowleft$ that some parts of the diagram are known to commute (by definition of $G \circ F$).[117]

[117] We did not leave the arrow $(G \circ F)_1$ because it would make the diagram messy.

$$\begin{array}{ccccc} \mathbf{C}_0 & \xleftarrow{s} & \mathbf{C}_1 & \xrightarrow{t} & \mathbf{C}_0 \\ {\scriptstyle F_0}\downarrow & & \downarrow{\scriptstyle F_1} & & \downarrow{\scriptstyle F_0} \\ (G\circ F)_0 \; \circlearrowleft \; \mathbf{D}_0 & & \mathbf{D}_1 & & \mathbf{D}_0 \; \circlearrowleft \; (G\circ F)_0 \\ {\scriptstyle G_0}\downarrow & & \downarrow{\scriptstyle G_1} & & \downarrow{\scriptstyle G_0} \\ \mathbf{E}_0 & \xleftarrow[s]{} & \mathbf{E}_0 & \xrightarrow[t]{} & \mathbf{E}_0 \end{array} \tag{21}$$

Then we can decompose the two rectangles into four squares that all commute by hypothesis that F and G are functors.

$$\begin{array}{ccccc} \mathbf{C}_0 & \xleftarrow{s} & \mathbf{C}_1 & \xrightarrow{t} & \mathbf{C}_0 \\ {\scriptstyle F_0}\downarrow & \circlearrowleft & \downarrow{\scriptstyle F_1} & \circlearrowleft & \downarrow{\scriptstyle F_0} \\ (G\circ F)_0 \; \circlearrowleft \; \mathbf{D}_0 & \xleftarrow{s} & \mathbf{D}_1 & \xrightarrow{t} & \mathbf{D}_0 \; \circlearrowleft \; (G\circ F)_0 \\ {\scriptstyle G_0}\downarrow & \circlearrowleft & \downarrow{\scriptstyle G_1} & \circlearrowleft & \downarrow{\scriptstyle G_0} \\ \mathbf{E}_0 & \xleftarrow[s]{} & \mathbf{E}_0 & \xrightarrow[t]{} & \mathbf{E}_0 \end{array} \tag{22}$$

Finally, we recognize that all the commutative diagrams in (22) combine into (20), so the latter is commutative.

From now on, when doing proofs by paving a diagram, we will only show the last paved diagram. Instead of $\circlearrowleft$, we will use letters to indicate regions that commute so we can refer to each region in the text and explain why they commute.

There is one last thing we want to mention to end this chapter. We gave two central definitions, categories and functors, and we presented several examples of each. By defining products, we give you access to an unlimited amount of new categories and functors you can construct from known ones.[118]

[118] This is akin to products of groups, direct sums of vector spaces, etc. In Chapter 3, we will see how all of these constructions are instances of a more general construction called (categorical) product.

Definition 132 (Product category)**.** Let $\mathbf{C}$ and $\mathbf{D}$ be two categories, the **product** of $\mathbf{C}$ and $\mathbf{D}$, denoted by $\mathbf{C} \times \mathbf{D}$, is the category whose objects are pairs of objects in $\mathbf{C}_0 \times \mathbf{D}_0$ and for any two pairs $(X, Y), (X', Y') \in (\mathbf{C} \times \mathbf{D})_0$,[119]

[119] Explicitly, a morphism $(X, Y) \to (X', Y')$ is a pair of morphisms $X \to X'$ and $Y \to Y'$.

$$\mathrm{Hom}_{\mathbf{C}\times\mathbf{D}}((X,Y),(X',Y')) := \mathrm{Hom}_{\mathbf{C}}(X,X') \times \mathrm{Hom}_{\mathbf{D}}(Y,Y').$$

The identity morphisms and the composition are defined componentwise. Explicitly, for all $X \in \mathbf{C}_0$ and $Y \in \mathbf{D}_0$, $\mathrm{id}_{(X,Y)} = (\mathrm{id}_X, \mathrm{id}_Y)$, and for all $(f,f') \in \mathbf{C}_2$ and $(g,g') \in \mathbf{D}_2$, $(f,g) \circ (f',g') = (f \circ f', g \circ g')$.[120]

[120] We leave you to check that this defines the composition for all of $(\mathbf{C} \times \mathbf{D})_2$. Namely, if (f,g) and (f',g') are composable, then (f,f') and (g,g') are composable.

SOL **Exercise 133** (NOW!)**.** Verify that the category depicted in (9) is appropriately denoted by $\mathbf{2}\times\mathbf{2}$, i.e., that it is the product category formed with $\mathbf{C} = \mathbf{D} = \mathbf{2}$.

SOL **Exercise 134.** Show that the assignment $\Delta_{\mathbf{C}} : \mathbf{C} \rightsquigarrow \mathbf{C} \times \mathbf{C} = X \mapsto (X,X)$ is functorial, i.e., give its action on morphisms and show it satisfies the relevant axioms. We call $\Delta_{\mathbf{C}}$ the **diagonal functor**.

Definition 135 (Product functor)**.** Let $F : \mathbf{C} \rightsquigarrow \mathbf{C}'$ and $G : \mathbf{D} \rightsquigarrow \mathbf{D}'$ be two functors, the **product** of F and G, denoted $F \times G : \mathbf{C} \times \mathbf{D} \rightsquigarrow \mathbf{C}' \times \mathbf{D}'$, is defined componentwise on objects and morphisms, i.e., for any $(X,Y) \in (\mathbf{C}\times\mathbf{D})_0$ and $(f,g) \in (\mathbf{C}\times\mathbf{D})_1$,

$$(F \times G)(X,Y) = (FX, GY) \text{ and } (F\times G)(f,g) = (Ff, Gg).$$

Let us check this defines a functor.

i. By definition of $\mathbf{C}' \times \mathbf{D}'$, (Ff, Gg) is a morphism from (FX, GY) to (FX', GY').

ii. For $(f,f') \in \mathbf{C}_2$ and $(g,g') \in \mathbf{D}_2$, we have

$$\begin{aligned}(F\times G)((f,g)\circ(f',g')) &= (F\times G)(f\circ f', g\circ g')\\ &= (F(f\circ f'), G(g\circ g'))\\ &= (Ff\circ Ff', Gg\circ Gg')\\ &= (Ff,Gg)\circ(Ff',Gg')\\ &= (F\times G)(f,g)\circ(F\times G)(f',g').\end{aligned}$$

iii. Since F and G preserve identity morphisms, we have

$$(F\times G)(\mathrm{id}_{(X,Y)}) = (F\times G)(\mathrm{id}_X, \mathrm{id}_Y) = (F\mathrm{id}_X, G\mathrm{id}_Y) = (\mathrm{id}_{FX}, \mathrm{id}_{GY}) = \mathrm{id}_{(FX,GY)}.$$

SOL **Exercise 136** (NOW!)**.** Let $F : \mathbf{C}\times\mathbf{C}' \rightsquigarrow \mathbf{D}$ be a functor. For $X \in \mathbf{C}_0$, we define $F(X,-) : \mathbf{C}' \rightsquigarrow \mathbf{D}$ on objects by $Y \mapsto F(X,Y)$ and on morphisms by $g \mapsto F(\mathrm{id}_X, g)$. Show that $F(X,-)$ is a functor. Define $F(-,Y)$ similarly.

SOL **Exercise 137.** Let $F : \mathbf{C}\times\mathbf{C}' \to \mathbf{D}$ be an action defined on objects and morphisms satisfying

$$F(f,g) = F(f, \mathrm{id}_{t(g)}) \circ F(\mathrm{id}_{s(f)}, g) = F(\mathrm{id}_{t(f)}, g) \circ F(f, \mathrm{id}_{s(g)}).$$

Show that if for any $X \in \mathbf{C}_0$ and $Y \in \mathbf{C}'_0$, $F(X,-)$ and $F(-,Y)$ as defined above are functors, then F is a functor. In other words, the functoriality of F can be proven componentwise.

In the next chapters, we will present other interesting constructions of categories, but we can stop here for now.

2 Duality

2.1 Contravariant Functors 38
2.2 Opposite Category 40
2.3 Duality in Action 41
2.4 More Vocabulary 48

The concept of duality is ubiquitous throughout mathematics. It can relate two perspectives of the same object as for dual vector spaces, two complementary optimization problems such as a maximization and a minimization linear program, and even two seemingly unrelated subjects like topology and logic (Stone duality). While this vague principle of duality is behind many groundbreaking results, the duality in question here is categorical duality and it is a bit more precise.

Informally, there is nothing more to say than "Take all the diagrams in a definition/theorem, reverse the arrows and reap the benefits of the dual concept/result."[121] The more formal version will follow after we first exhibit the principle in action.

[121] In my opinion, this is already a very good reason to learn category theory because we can basically get twice as much math as before by framing things with a categorical language.

Recall that, intuitively, a functor is a structure-preserving transformation between categories. A simple example we have seen is functors between posets that are order-preserving functions. However, as a consequence, one might conclude that order-reversing functions impair the structure of a poset, which feels arbitrary. The same happens between deloopings of groups because anti-homomorphisms[122] do not arise as functors between such categories.

[122] An **anti-homomorphism** $f : G \to H$ is a function satisfying $f(gg') = f(g')f(g)$ and $f(1_G) = f(1_H)$.

For a more concrete situation, recall the powerset functor $\mathcal{P}$ described in Example 126.1. It assigns to any set X the powerset $\mathcal{P}X$, and to any function $f : X \to Y$ the image function $\mathcal{P}(f) : \mathcal{P}X \to \mathcal{P}Y$. There is another important function associated to f between powersets: the inverse image f^{-1} that assigns to $S \subset Y$ the set of points in X whose images are in S. Unfortunately, f^{-1} goes in the "wrong" direction $\mathcal{P}Y \to \mathcal{P}X$.

This is quite unsatisfactory because the assignment $f \mapsto f^{-1}$ is well-behaved, e.g. we have $\mathrm{id}_X^{-1} = \mathrm{id}_{\mathcal{P}X}$ for any set X and, for any functions $f : X \to Y$ and $g : Y \to X$, $(f \circ g)^{-1} = g^{-1} \circ f^{-1}$. This second equation looks just like the second condition on functors reversed. In words, taking the inverse image *preserves* composition but in reverse.

It seems arbitrary to distinguish between both options. There are two ways to remedy this discrepancy between intuition and formalism; both have duality as an underlying principle. In this chapter, we will describe both ways, dismiss one of them, and showcase the strength of duality while exploring more basic category theory.

2.1 Contravariant Functors

By modifying Definition 117 to require that $F(f)$ goes in the opposite direction, we obtain a contravariant functor. Incidentally, what we defined as a functor before is also called a **covariant** functor.

Definition 138 (Contravariant functor)**.** Let **C** and **D** be categories, a **contravariant functor** $F : \mathbf{C} \rightsquigarrow \mathbf{D}$ is a pair of maps $F_0 : \mathbf{C}_0 \to \mathbf{D}_0$ and $F_1 : \mathbf{C}_1 \to \mathbf{D}_1$ making diagrams (23), (24) and (25) commute.[123]

[123] Where F_2' is now induced by the definition of F_1 with $(f,g) \mapsto (F_1(g), F_1(f))$.

$$\begin{array}{ccccc} \mathbf{C}_0 & \xleftarrow{s} & \mathbf{C}_1 & \xrightarrow{t} & \mathbf{C}_0 \\ {\scriptstyle F_0}\downarrow & & \downarrow{\scriptstyle F_1} & & \downarrow{\scriptstyle F_0} \\ \mathbf{D}_0 & \xleftarrow[t]{} & \mathbf{D}_1 & \xrightarrow[s]{} & \mathbf{D}_0 \end{array} \tag{23}$$

$$\begin{array}{ccc} \mathbf{C}_2 & \xrightarrow{F_2'} & \mathbf{D}_2 \\ {\scriptstyle \circ_{\mathbf{C}}}\downarrow & & \downarrow{\scriptstyle \circ_{\mathbf{D}}} \\ \mathbf{C}_1 & \xrightarrow[F_1]{} & \mathbf{D}_1 \end{array} \tag{24}$$

$$\begin{array}{ccc} \mathbf{C}_0 & \xrightarrow{F_0} & \mathbf{D}_0 \\ {\scriptstyle u_{\mathbf{C}}}\downarrow & & \downarrow{\scriptstyle u_{\mathbf{D}}} \\ \mathbf{C}_1 & \xrightarrow[F_1]{} & \mathbf{D}_1 \end{array} \tag{25}$$

In words, F must satisfy the following properties.

i. For any $A, B \in \mathbf{C}_0$, if $f \in \mathrm{Hom}_{\mathbf{C}}(A, B)$ then $F(f) \in \mathrm{Hom}_{\mathbf{D}}(F(B), F(A))$.

ii. If $f, g \in \mathbf{C}_1$ are composable, then $F(f \circ g) = F(g) \circ F(f)$.

iii. If $A \in \mathbf{C}_0$, then $u_{\mathbf{D}}(F(A)) = F(u_{\mathbf{C}}(A))$.

Examples 139. Just like their covariant counterparts, contravariant functors are quite numerous. Here are a couple of simple ones.

1. Contravariant functors $F : (X, \leq) \rightsquigarrow (Y, \sqsubseteq)$ correspond to order-reversing functions between the posets X and Y and contravariant functors $F : \mathbf{B}G \rightsquigarrow \mathbf{B}H$ correspond to anti-homomorphisms between the groups G and H.

2. The **contravariant powerset functor** $2^{-} : \mathbf{Set} \rightsquigarrow \mathbf{Set}$ sends a set X to its powerset 2^X,[124] and a function $f : X \to Y$ to the preimage map $2^f : 2^Y \to 2^X$, the latter sends a subset $S \subseteq Y$ to

$$2^f(S) = f^{-1}(S) := \{x \in X \mid f(x) \in S\} \subseteq X.$$

[124] We use a different notation for the powerset to emphasize the difference between $\mathcal{P}$ and 2^{-}.

Next, there is a couple of functors that are key to understand the philosophy put forward by category theory.[125]

[125] We will talk more about it when covering the Yoneda lemma in Chapter **??**.

Example 140 (Hom functors)**.** Let **C** be a locally small category and $A \in \mathbf{C}_0$ one of its objects.[126] We define the covariant and contravariant **Hom functors** from **C** to **Set**.

[126] We need local smallness so that each $\mathrm{Hom}_{\mathbf{C}}(A, B)$ is a set and the functors land in **Set**.

1. The covariant Hom functor $\mathrm{Hom}_{\mathbf{C}}(A, -) : \mathbf{C} \rightsquigarrow \mathbf{Set}$ sends an object $B \in \mathbf{C}_0$ to the hom-set $\mathrm{Hom}_{\mathbf{C}}(A, B)$ and a morphism $f : B \to B'$ to the function

$$\mathrm{Hom}_{\mathbf{C}}(A, f) : \mathrm{Hom}_{\mathbf{C}}(A, B) \to \mathrm{Hom}_{\mathbf{C}}(A, B') = g \mapsto f \circ g.$$

This function is called **post-composition by** f and is denoted $f \circ (-)$.[127] Let us show $\mathrm{Hom}_{\mathbf{C}}(A, -)$ is a covariant functor.

i. For any $f \in \mathbf{C}_1$, it is clear from the definition that

$$\mathrm{Hom}_{\mathbf{C}}(A, s(f)) = s(f \circ (-)) \text{ and } \mathrm{Hom}_{\mathbf{C}}(A, t(f)) = t(f \circ (-)).$$

ii. For any $(f_1, f_2) \in \mathbf{C}_2$, we claim that

$$\mathrm{Hom}_{\mathbf{C}}(A, f_1 \circ f_2) = \mathrm{Hom}_{\mathbf{C}}(A, f_1) \circ \mathrm{Hom}_{\mathbf{C}}(A, f_2).$$

In the L.H.S., an element $g \in \mathrm{Hom}_{\mathbf{C}}(A, s(f_1 \circ f_2))$ is mapped to $(f_1 \circ f_2) \circ g$ and in the R.H.S., an element $g \in \mathrm{Hom}_{\mathbf{C}}(A, s(f_2))$ is mapped to $f_1 \circ (f_2 \circ g)$. Since $s(f_1 \circ f_2) = s(f_2)$ and composition is associative, we conclude that the two maps are the same.

iii. For any $B \in \mathbf{C}_0$, the post-composition by $u_{\mathbf{C}}(B)$ is defined to be the identity,[128] hence (16) also commutes.

2. The contravariant Hom functor $\mathrm{Hom}_{\mathbf{C}}(-, A) : \mathbf{C} \rightsquigarrow \mathbf{Set}$ sends an object $B \in \mathbf{C}_0$ to the hom-set $\mathrm{Hom}_{\mathbf{C}}(B, A)$ and a morphism $f : B \to B'$ to the function

$$\mathrm{Hom}_{\mathbf{C}}(f, A) : \mathrm{Hom}_{\mathbf{C}}(B', A) \to \mathrm{Hom}_{\mathbf{C}}(B, A) = g \mapsto g \circ f.$$

This function is called **pre-composition by** f and is denoted $(-) \circ f$.[129] Let us show $\mathrm{Hom}_{\mathbf{C}}(-, A)$ is a contravariant functor.

i. For any $f \in \mathbf{C}_1$, it is clear from the definition that

$$\mathrm{Hom}_{\mathbf{C}}(s(f), A) = t((-) \circ f)) \text{ and } \mathrm{Hom}_{\mathbf{C}}(t(f), A) = s((-) \circ f).$$

ii. For any $(f_1, f_2) \in \mathbf{C}_2$, we claim that

$$\mathrm{Hom}_{\mathbf{C}}(f_1 \circ f_2, A) = \mathrm{Hom}_{\mathbf{C}}(f_2, A) \circ \mathrm{Hom}_{\mathbf{C}}(f_1, A).$$

In the L.H.S., an element $g \in \mathrm{Hom}_{\mathbf{C}}(t(f_1 \circ f_2), A)$ is mapped to $g \circ (f_1 \circ f_2)$ and in the R.H.S., an element $g \in \mathrm{Hom}_{\mathbf{C}}(t(f_1), A)$ is mapped to $(g \circ f_1) \circ f_2$. Since $t(f_1 \circ f_2) = t(f_1)$ and composition is associative, we conclude that the two maps are the same.

iii. For any $B \in \mathbf{C}_0$, pre-composition by $u_{\mathbf{C}}(B)$ is defined to be the identity,[130] hence (25) also commutes.

It can take a bit of time to get comfortable with Hom functors. For now, we will give only one example of each kind (covariant and contravariant), but we will take more time to play with them later in the book.

Example 141 (Ring of functions)**.**

[127] Some authors denote $f \circ (-)$ as f^*, we prefer to keep this notation for later (see pullbacks).

[128] Namely, for any $f : A \to B$, $u_{\mathbf{C}}(B) \circ f = f$.

[129] Some authors denote $(-) \circ f$ as f_*, we prefer to keep this notation for later (see pushouts).

[130] Namely, for any $f : B \to A$, $f \circ u_{\mathbf{C}}(B) = f$.

Example 142 (Dual vector space)**.** In the category $\mathbf{Vect}_k$, there is a special object k,[131] let us see what the contravariant functor $\mathrm{Hom}_{\mathbf{Vect}_k}(-,k)$ does. It assigns to any vector space V the set of linear maps $V \to k$, that is, the carrier set of the dual space V^*. It assigns to linear maps $T : V \to W$, the function

$$\mathrm{Hom}_{\mathbf{Vect}_k}(W,k) \to \mathrm{Hom}_{\mathbf{Vect}_k}(V,k) = \phi \mapsto \phi \circ T.$$

We know that $\mathrm{Hom}_{\mathbf{Vect}_k}(V,k) = V^*$ can be seen as a vector space and it is easy to check that pre-composition by T is a linear map $W^* \to V^*$. Therefore, we find that the assignment $V \mapsto V^* = \mathrm{Hom}_{\mathbf{Vect}_k}(-,k)$ is a contravariant functor $\mathbf{Vect}_k \rightsquigarrow \mathbf{Vect}_k$.

We will not dwell too long on contravariant functors as we will see right away how they can be avoided, but first, let us give a reason why we want to avoid them.

SOL **Exercise 143.** Let $F : \mathbf{C} \rightsquigarrow \mathbf{D}$, $G : \mathbf{D} \rightsquigarrow \mathbf{E}$ be contravariant functors, and $G \circ F : \mathbf{C} \rightsquigarrow \mathbf{E}$ be their composition defined by $G_0 \circ F_0$ on objects and $G_1 \circ F_1$ on morphisms. Show that $G \circ F$ is a *covariant* functor.[132] Using diagrams will be easier.

2.2 Opposite Category

Another way to deal with order-reversing maps $(X, \leq) \to (Y, \subseteq)$ is to consider the reverse order on X and a covariant functor $(X, \geq) \rightsquigarrow (Y, \subseteq)$. This also works for anti-homomorphisms by constructing the opposite group G^{op} in which the operation is reversed, namely $g \cdot^{\mathrm{op}} h = hg$. The opposite category is a generalization of these constructions.

Definition 144 (Opposite category)**.** Let $\mathbf{C}$ be a category, we denote the **opposite** category with $\mathbf{C}^{\mathrm{op}}$ and define it by[133]

$$\mathbf{C}^{\mathrm{op}}_0 = \mathbf{C}_0,\ \mathbf{C}^{\mathrm{op}}_1 = \mathbf{C}_1,\ s^{\mathrm{op}} = t,\ t^{\mathrm{op}} = s,\ u_{\mathbf{C}^{\mathrm{op}}} = u_{\mathbf{C}}$$

with the composition defined by $f^{\mathrm{op}} \circ^{\mathrm{op}} g^{\mathrm{op}} = (g \circ f)^{\mathrm{op}}$.[134] This naturally leads to the following contravariant functor $(-)^{\mathrm{op}}_{\mathbf{C}} : \mathbf{C} \rightsquigarrow \mathbf{C}^{\mathrm{op}}$ which sends an object A to A^{op} and a morphism f to f^{op}. It is called the **opposite functor**.

With this definition, one can see contravariant functors as covariant functors. Formally, let $F : \mathbf{C} \rightsquigarrow \mathbf{D}$ be a contravariant functor, we can view F as covariant functor from $\mathbf{C}^{\mathrm{op}}$ to $\mathbf{D}$ or from $\mathbf{C}$ to $\mathbf{D}^{\mathrm{op}}$ via the compositions $F \circ (-)^{\mathrm{op}}_{\mathbf{C}^{\mathrm{op}}}$ and $(-)^{\mathrm{op}}_{D} \circ F$ respectively.[135]

In the rest of this book, we choose to work with covariant functors of type $\mathbf{C}^{\mathrm{op}} \rightsquigarrow \mathbf{D}$ instead of contravariant functors $\mathbf{C} \rightsquigarrow \mathbf{D}$,[136] and functors will be covariant by default.

Examples 145. 1. As hinted at before, the category corresponding to $(X, \geq)$ is the opposite category of $(X, \leq)$ and $(\mathbf{B}G)^{\mathrm{op}}$ is the category corresponding to the opposite group of G, i.e.: $(\mathbf{B}G)^{\mathrm{op}} = \mathbf{B}(G^{\mathrm{op}})$.

2. We have seen that functors $\mathbf{B}G \rightsquigarrow \mathbf{Set}$ correspond to left actions of a group G. You can check that functors $\mathbf{B}G^{\mathrm{op}} \rightsquigarrow \mathbf{Set}$ correspond to right actions of G.

[131] We know it is special because we know some linear algebra, but k also has some interesting categorical properties (see Exercise 176).

[132] We conclude that we cannot straightforwardly compose contravariant functors. This alone makes the following alternative more desirable because we want functors to be morphisms in a category, hence they must be composable.

[133] Intuitively, we reverse the direction of all morphisms in $\mathbf{C}$ and reverse the order of composition as well.

[134] Note that the $-^{\mathrm{op}}$ notation here is just used to distinguish elements in $\mathbf{C}$ and $\mathbf{C}^{\mathrm{op}}$ but the collection of objects and morphisms are the same.

[135] Recall from Exercise 143 that these compositions are covariant.

[136] We still had to learn about contravariant functors because you might encounter them in the wild.

3. The two Hom functors defined in Example 140 are now written

$$\mathrm{Hom}_{\mathbf{C}}(A,-) : \mathbf{C} \rightsquigarrow \mathbf{Set} \text{ and } \mathrm{Hom}_{\mathbf{C}}(-,A) : \mathbf{C}^{\mathrm{op}} \rightsquigarrow \mathbf{Set}.$$

By Exercise 137, they can be combined into a functor

$$\mathrm{Hom}_{\mathbf{C}}(-,-) : \mathbf{C}^{\mathrm{op}} \times \mathbf{C} \rightsquigarrow \mathbf{Set}$$

acting on objects as $(A,B) \mapsto \mathrm{Hom}_{\mathbf{C}}(A,B)$ and on morphisms as $(f,g) \mapsto (g \circ - \circ f)$. The condition in Exercise 137 is satisfied because[137]

$$\begin{aligned} \mathrm{Hom}_{\mathbf{C}}(f,g) &= g \circ - \circ f \\ &= \mathrm{id}_{t(g)} \circ (g \circ - \circ \mathrm{id}_{t(f)}) \circ f = \mathrm{Hom}_{\mathbf{C}}(f, \mathrm{id}_{t(g)}) \circ \mathrm{Hom}_{\mathbf{C}}(\mathrm{id}_{t(f)}, g) \\ &= g \circ (\mathrm{id}_{s(g)} \circ - \circ f) \circ \mathrm{id}_{s(f)} = \mathrm{Hom}_{\mathbf{C}}(\mathrm{id}_{s(f)}, g) \circ \mathrm{Hom}_{\mathbf{C}}(f, \mathrm{id}_{s(g)}). \end{aligned}$$

This will be called the Hom **bifunctor**.

[137] Looking at where the source and target functions are applied, these equalities do not match exactly what is in Exercise 137 since $\mathrm{Hom}_{\mathbf{C}}(-,-)$ is contravariant in the first component.

SOL **Exercise 146.** Let $F : \mathbf{C} \rightsquigarrow \mathbf{D}$ be a functor, show that its dual F^{op} defined by $A^{\mathrm{op}} \mapsto (FA)^{\mathrm{op}}$ on objects and $f^{\mathrm{op}} \mapsto (Ff)^{\mathrm{op}}$ on morphisms is a functor $\mathbf{C}^{\mathrm{op}} \rightsquigarrow \mathbf{D}^{\mathrm{op}}$.

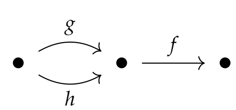
Remark 147. It is sometimes useful to compose the Hom bifunctor with other functors as follows. Given two functors $F, G : \mathbf{C} \rightsquigarrow \mathbf{D}$, there is a functor $\mathrm{Hom}_{\mathbf{D}}(F-, G-) : \mathbf{C}^{\mathrm{op}} \times \mathbf{C} \rightsquigarrow \mathbf{D}$ acting on objects by $(X,Y) \mapsto \mathrm{Hom}_{\mathbf{D}}(FX, GY)$ and on morphisms by $(f,g) \mapsto Gg \circ (-) \circ Ff$. One can check functoriality by showing

$$\mathrm{Hom}_{\mathbf{D}}(F-, G-) = \mathrm{Hom}_{\mathbf{D}}(-,-) \circ (F^{\mathrm{op}} \times G).$$

2.3 Duality in Action

Let us start to illustrate how duality can be useful while covering important definitions and results.

Definition 148 (Monomorphism)**.** Let $\mathbf{C}$ be a category, a morphism $f \in \mathbf{C}_1$ is said to be **monic** (or a **monomorphism**) if for any parallel morphisms g and h such that $(f,g), (f,h) \in \mathbf{C}_2$, $f \circ g = f \circ h$ implies $g = h$. Equivalently, f is monic if $g = h$ whenever the following diagram commutes.[138]



(26)

[138] According to Definition 99, this diagram commutes if and only if $f \circ g = f \circ h$ because the paths (f,g) and (f,h) are the only paths of length bigger than one.

Standard notation for a monomorphism is $\bullet \rightarrowtail \bullet$ (`\rightarrowtail`).[139]

[139] Another widespread notation is $\bullet \hookrightarrow \bullet$. I prefer to use the latter when we understand the morphism as an "inclusion" of the first object in the second. These are often monic.

Proposition 149. *Let* $\mathbf{C}$ *be a category and* $f : A \to B$ *a morphism, if there exists* $f' : B \to A$ *such that* $f' \circ f = \mathrm{id}_A$,[140] *then* f *is a monomorphism.*

[140] We say that f' is a **left inverse** of f.

Proof. If $f \circ g = f \circ h$, then $f' \circ f \circ g = f' \circ f \circ h$ implying $g = h$. □

Not all monomorphisms have a left inverse, those that do are called **split monomorphisms**.

Proposition 150. *Let* $\mathbf{C}$ *be a category and* $(f_1, f_2) \in \mathbf{C}_2$*, if* $f_1 \circ f_2$ *is a monomorphism, then* f_2 *is a monomorphism.*

Proof. Let $g, h \in \mathbf{C}_1$ be such that $f_2 \circ g = f_2 \circ h$, we readily get that $(f_1 \circ f_2) \circ g = (f_1 \circ f_2) \circ h$. Since $f_1 \circ f_2$ is a monomorphism, this implies $g = h$. □

The last two results hint at the fact that monomorphisms are analogues to injective functions and we will see that they are exactly the same in the category **Set**, but first let us introduce the dual concept after the formal definition of duality.

Definition 151 (Duality)**.** Given a definition or statement in an arbitrary category $\mathbf{C}$, one could view this concept inside the category $\mathbf{C}^{\mathrm{op}}$ and obtain a similar definition or statement where all morphisms and the order of composition are reversed, this is called the **dual** concept. Since $\mathbf{C}^{\mathrm{op}^{\mathrm{op}}} = \mathbf{C}$, taking the dual is an involution, namely, the dual of the dual of a thing is that thing. For a definition or result where multiple *arbitrary* categories are involved, the dual version is obtained by taking the opposite of all categories.[141] It is common but not systematic to refer to a dual notion with the prefix "co" (e.g.: presheaf and copresheaf).

[141] Note the emphasis on the word "arbitrary". For instance, a **presheaf** is a functor $F : \mathbf{C}^{\mathrm{op}} \rightsquigarrow \mathbf{Set}$ and the dual concept is a **copresheaf**, a functor $F : \mathbf{C} \rightsquigarrow \mathbf{Set}$; we did not take the opposite of **Set**.

Dualizing the definition of a monomorphism yields an epimorphism.

Definition 152 (Epimorphism)**.** Let $\mathbf{C}$ be a category, a morphism $f \in \mathbf{C}_1$ is said to be **epic** (or an **epimorphism**) if for any two parallel morphisms g and h such that $(g, f), (h, f) \in \mathbf{C}_2$, $g \circ f = h \circ f$ implies $g = h$. Equivalently, f is epic if $g = h$ whenever the following diagram commutes.[142]

$$\bullet \xrightarrow{f} \bullet \underset{h}{\overset{g}{\rightrightarrows}} \bullet \tag{27}$$

[142] Seeing the diagrams make it clearer that the concepts are dual. Reversing the arrows in (26) yields (27) and vice-versa.

Standard notation for an epimorphism is $\bullet \twoheadrightarrow \bullet$ (`\twoheadrightarrow`).

The dual versions of Propositions 149 and 150 also hold. Although translating our previous proofs to the dual case is straightforward, we will do the two next proofs relying on duality to convey the general sketch that works anytime a dual result needs to be proven.

Proposition 153. *Let* $\mathbf{C}$ *be a category and* $f : A \to B$ *a morphism, if there exists* $f' : B \to A$ *such that* $f \circ f' = \mathrm{id}_B$*, then* f *is epic.*[143]

[143] We say that f' is a **right inverse** of f.

Proof. Observe that f is epic in $\mathbf{C}$ if and only if f^{op} is monic in $\mathbf{C}^{\mathrm{op}}$ (reverse the arrows in the definition).[144] Moreover, by definition,

$$f'^{\mathrm{op}} \circ f^{\mathrm{op}} = (f \circ f')^{\mathrm{op}} = \mathrm{id}_B{}^{\mathrm{op}} = \mathrm{id}_{B^{\mathrm{op}}},$$

so by the result for monomorphisms, f^{op} is monic and hence f is epic. □

[144] This is another way to see that two concepts are dual.

Not all epimorphisms have a right inverse, those that do are called **split epimorphisms**.

Proposition 154. *Let* $\mathbf{C}$ *be a category and* $(f_1, f_2) \in \mathbf{C}_2$*, if* $f_1 \circ f_2$ *is epic, then* f_1 *is epic.*

Proof. Since ${f_2}^{\mathrm{op}} \circ {f_1}^{\mathrm{op}} = (f_1 \circ f_2)^{\mathrm{op}}$ is monic, the result for monomorphisms implies ${f_1}^{\mathrm{op}}$ is monic and hence f_1 is epic. □

Example 155 (Set). We mentioned that monomorphisms are like generalizations of injective functions, and you may have guessed that epimorphisms are, in the same sense, generalizations of surjective functions. Let us make this precise.

- A function $f : A \to B$ is a monomorphism in **Set** if and only if it is injective:[145]

  (⇐) Since f is injective, it has a left inverse, so it is monic by Proposition 149.

  (⇒) Given $a \in A$, let $g_a : \{*\} \to A$ be the function sending $*$ to a. For any $a_1 \neq a_2 \in A$, the functions g_{a_1} and g_{a_2} are different, hence $f \circ g_{a_1} \neq f \circ g_{a_2}$. Therefore, $f(a_1) \neq f(a_2)$ implying f is injective.

- A function $f : A \to B$ is an epimorphism if and only if it is surjective:[146]

  (⇐) Since f is surjective, it has a right inverse, so it is epic by Proposition 153.

  (⇒) Let $h : B \to \{0, 1\}$ be the constant function at 1 and $g : B \to \{0, 1\}$ be the indicator function of $\mathrm{Im}(f) \subseteq B$, namely,

$$g(x) = \begin{cases} 1 & \exists a \in A, x = f(a) \\ 0 & \text{otherwise} \end{cases}.$$

  We see that $g \circ f = h \circ f$ are both constant at 1, and f being epic implies $g = h$. Thus, any element of B is in the image of f, that is, f is surjective.

[145] As a consequence, since all injective functions have a left inverse, all the monomorphisms in **Set** are split.

[146] If you assume the axiom of choice, all surjective functions have a right inverse and thus all epimorphisms in **Set** are split.

Example 156 (Mon). Inside the category **Mon**, the monomorphisms are precisely the injective homomorphisms.

(⇒) Let $f : M \to M'$ be an injective homomorphisms and $g_1, g_2 : N \to M$ be two parallel homomorphisms. Suppose that $f \circ g_1 = f \circ g_2$, then for all $x \in N$, $f(g_1(x)) = f(g_2(x))$, so by injectivity of f, $g_1(x) = g_2(x)$. Therefore, $g_1 = g_2$ and since g_1 and g_2 were arbitrary, f is a monomorphism.

(⇐) Let $f : M \to M'$ be a monomorphism. Let $x, y \in M$ and define $p_x : (\mathbb{N}, +) \to M$ by $k \mapsto x^k$ and similarly for p_y. It is easy to show that p_x and p_y are homomorphisms.[147] If $f(x) = f(y)$, then, by the homomorphism property, for all $k \in \mathbb{N}$

$$f(p_x(k)) = f(x^k) = f(x)^k = f(y)^k = f(y^k) = f(p_y(k)).$$

In other words, we get $f \circ p_x = f \circ p_y$, so $p_x = p_y$ and $x = y$. This direction follows.

[147] It follows from the definition of x^k which is $x \overset{k}{\cdots} x$.

Conversely, an epimorphism is not necessarily surjective. For example, the inclusion homomorphism $i : (\mathbb{N}, +) \to (\mathbb{Z}, +)$ is clearly not surjective, but it is an epimorphism. Indeed, let $g, h : (\mathbb{Z}, +) \to M$ be two monoid homomorphisms satisfying $g \circ i = h \circ i$. In particular, $g(n) = h(n)$ for any $n \in \mathbb{N} \subset \mathbb{Z}$. It remains to show that also $g(-n) = h(-n)$: we have

$$\begin{aligned} h(n)g(-n) = g(n)g(-n) = g(n-n) = g(0) = 1_M \\ h(-n)h(n) = h(-n+n) = h(0) = 1_M, \end{aligned}$$

but then $g(-n) = h(-n)h(n)g(-n) = h(-n)$.

SOL **Exercise 157.** Show that a monomorphism in **Cat** is a functor that is faithful and injective on objects, it is called an **embedding**.[148]

SOL **Exercise 158.** Show that a morphism $f \in \mathbf{C}_1$ is monic if and only if the function $\mathrm{Hom}_{\mathbf{C}}(A, f) = f \circ -$ is injective for all $A \in \mathbf{C}_0$. Dually, show that f is epic if and only if the function $\mathrm{Hom}_{\mathbf{C}^{\mathrm{op}}}(A^{\mathrm{op}}, f^{\mathrm{op}}) = \mathrm{Hom}_{\mathbf{C}}(f, A) = - \circ f$ is injective for all $A \in \mathbf{C}_0$.

Remark 159. These alternative definitions of monomorphisms and epimorphisms are more categorical in nature. In fact, in the setting of enriched category theory they are preferable because they generalize easily.

Definition 160 (Isomorphism)**.** Let **C** be a category, a morphism $f : A \to B$ is said to be an **isomorphism** if there exists a morphism $f^{-1} : B \to A$ such that $f \circ f^{-1} = \mathrm{id}_B$ and $f^{-1} \circ f = \mathrm{id}_A$.[149]

SOL **Exercise 161.** Show that the property of being monic/epic/an isomorphism is invariant under composition, i.e., if f and g are composable monomorphisms, then $f \circ g$ is monic and similarly for epimorphisms and isomorphisms.

Remark 162. The results shown about monic and epic morphisms[150] imply that any isomorphism is monic and epic. However, the converse is not true as witnessed by the inclusion morphism i described in Example 156.[151] A category where all monic and epic morphisms are isomorphisms (e.g.: **Set**) is called **balanced**. If there exists an isomorphism between two objects A and B, then they are **isomorphic**, denoted $A \cong B$. Isomorphic objects are also isomorphic in the opposite category,[152] that is, the concept of **isomorphism** is *self-dual*.

For most intents and purposes, we will not distinguish between isomorphic objects in a category because all the properties we care about will hold for one if and only if they hold for the other. This attitude should be somewhat familiar if you have done a bit of abstract algebra because it is natural to substitute the group $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}$ for $\mathbb{Z}/6\mathbb{Z}$ or k^n for an n–dimensional vector space over k. It is less natural in **Set** because, for instance, it equates the sets $\{0, 1\}$ and $\{a, b\}$ which may be too coarse-grained for our intuition.

Example 163 (**Set**)**.** A function $f : X \to Y$ in $\mathbf{Set}_1$ has an inverse f^{-1} if and only if f is bijective, thus isomorphisms in **Set** are bijections. As a consequence, we have $A \cong B$ if and only if $|A| = |B|$.[153]

Example 164 (**Cat**)**.** An isomorphism in **Cat** is a functor $F : \mathbf{C} \rightsquigarrow \mathbf{D}$ with an inverse $F^{-1} : \mathbf{D} \rightsquigarrow \mathbf{C}$. This implies that F_0 and F_1 are bijections[154] because $(F^{-1})_0$ is the inverse of F_0 and $(F^{-1})_1$ is the inverse of F_1.

Conversely, if $F : \mathbf{C} \rightsquigarrow \mathbf{D}$ is a functor whose component on objects and morphisms are bijective, we check that defining $F^{-1} : \mathbf{D} \rightsquigarrow \mathbf{C}$ with $(F^{-1})_0 := (F_0)^{-1}$ and $(F^{-1})_1 := (F_1)^{-1}$ yields a functor.

i. Let $f \in \mathrm{Hom}_{\mathbf{D}}(A, B)$, by bijectivity of F_0 and F_1, there are $X, Y \in \mathbf{C}_0$ and $g : X \to Y$ such that $FX = A$, $FY = B$ and $Fg = f$. Then, by definition,

$$s(F^{-1}f) = s(g) = X = F^{-1}FX = F^{-1}A, \text{ and}$$

[148] Finding a nice characterization of epimorphisms in **Cat** is an open question as far as I know.

[149] Then f^{-1} is called the **inverse** of f. One can check that if f' is a left inverse of f and f'' is a right inverse, then $f' = f'' = f^{-1}$. Hence, the inverse is unique.

[150] Proposition 149 and 153.

[151] This is not akin to the situation in **Set** because, there, all monomorphisms and epimorphisms are split (assuming the axiom of choice).

[152] Because the left inverse becomes the right inverse and vice-versa.

[153] This is in fact the definition of cardinality.

[154] When F_0 is a bijection, F_1 is a bijection if and only if F is fully faithful. Indeed, ...

$$t(F^{-1}f) = t(g) = Y = F^{-1}FY = F^{-1}B.$$

ii. For any $(f, f') \in \mathbf{D}_2$ with $f = Fg$ and $f' = Fg'$, we find

$$F^{-1}(f \circ f') = F^{-1}(Fg \circ Fg') = F^{-1}F(g \circ g') = g \circ g' = F^{-1}Fg \circ F^{-1}Fg' = Ff \circ Ff'.$$

iii. For any $A \in \mathbf{D}_0$ with $A = FX$, we find

$$F^{-1}\mathrm{id}_A = F^{-1}\mathrm{id}_{FX} = F^{-1}F\mathrm{id}_X = \mathrm{id}_X = \mathrm{id}_{F^{-1}FX} = \mathrm{id}_{F^{-1}A}.$$

We can conclude that isomorphisms in **Cat** are precisely the functors which are bijective on objects and morphisms. Furthermore, Footnote 154 implies they are precisely fully faithful functors that are bijective on objects.

Examples 165 (Concrete categories)**.** In a concrete category **C** with forgetful functor U, the underlying function of an isomorphism f must bijective because $U(f^{-1})$ is the inverse of Uf. This condition may be sufficient or not.

1. It is a simple exercise in an algebra class to show that isomorphisms in the categories **Mon**, **Grp**, **Ring**, **Field** and $\mathbf{Vect}_k$ are simply bijective homomorphisms.[155]

2. In **Poset**, an isomorphism between $(A, \leq_A)$ and $(B, \leq_B)$ is a bijective function $f : A \to B$ satisfying $a \leq_A a' \Leftrightarrow f(a) \leq_B f(a')$. Such a function is clearly monotone, but its inverse is also monotone as for any $b \leq_B b'$, we have $ff^{-1}(b) \leq_B ff^{-1}(b') \implies f^{-1}(b) \leq_A f^{-1}(b')$.

3. In **Top**, it is not enough to have a bijective continuous function, we need to require that it has a continuous inverse.[156] Such functions are called **homeomorphisms**.

[155] In fact, isomorphisms are commonly defined as bijective homomorphisms in said algebra class.

[156] Consider $X = \{0, 1\}$ with the two extreme topologies $\tau_\perp = \{\emptyset, X\}$ and $\tau_\top = \mathcal{P}X$. The identity map $\mathrm{id}_X : (X, \tau_\top) \to (X, \tau_\perp)$ is clearly bijective and continuous, but its inverse is not continuous. A similar example shows that a bijective monotone function is not necessarily a poset isomorphism.

Definition 166 (Initial object)**.** Let **C** be a category, an object $A \in \mathbf{C}_0$ is said to be **initial** if for any $B \in \mathbf{C}_0$, $|\mathrm{Hom}_{\mathbf{C}}(A, B)| = 1$, namely there are no two parallel morphisms with source A and every object has a morphism coming from A. The[157] initial object of a category, if it exists, is denoted $\emptyset$ and the *unique* morphism from $\emptyset$ to $X \in \mathbf{C}_0$ is denoted $[] : \emptyset \to X$.

[157] We will soon see why we can use *the* instead of *an*.

Definition 167 (Terminal object)**.** Let **C** be a category, an object $A \in \mathbf{C}_0$ is said to be **terminal**[158] if for any $B \in \mathbf{C}_0$, $|\mathrm{Hom}_{\mathbf{C}}(B, A)| = 1$, namely there are no two parallel morphisms with target A and every object has a morphism going to A. The terminal object of a category, if it exists, is denoted **1** and the *unique* morphism from $X \in \mathbf{C}_0$ into **1** is denoted $\langle\rangle : X \to \mathbf{1}$.

[158] The terminology final is also common.

Remark 168 (Notation). The motivation behind the notations $\emptyset$ and **1** is given shortly, but the notations for the morphisms will be explained in Chapter 3.

An object is initial in a category **C** if and only if it is terminal in $\mathbf{C}^{\mathrm{op}}$, so these two concepts are dual. Also, if an object is initial and terminal, we say it is a **zero** object and usually denote it **0**.[159]

[159] Clearly, the concept of zero object is self-dual.

Example 169 (Set). Let X be a set, there is a unique function from the empty set into X, it is the empty function.[160] We deduce that the empty set is the initial object in **Set**, hence the notation $\varnothing$. For the terminal object, we observe that there is a unique function $X \to \{*\}$ sending all elements of X to $*$, thus $\mathbf{1} = \{*\}$ is terminal in **Set**.

[160] Recall (or learn here) that a function $f : A \to B$ is defined via subset of $f \subseteq A \times B$ that satisfies $\forall a \in A, \exists! b \in B, (a, b) \in f$. When A is empty, $A \times B$ is empty and the only subset $\varnothing \subseteq A \times B$ satisfies the condition vacuously. In passing, when B is empty but A is not, the unique subset of $A \times B$ does not satisfy the condition, so there is no function $A \to \varnothing$.

In this example, we could have chosen any singleton to show it is terminal. However, that choice is irrelevant to a good category theorist because just as any two singletons are isomorphic (because they have the same cardinality), any two terminal objects are isomorphic.

Proposition 170. *Let* $\mathbf{C}$ *be a category and* $A, B \in \mathbf{C}_0$ *be initial, then* $A \cong B$.

Proof. Let f be the single element in $\mathrm{Hom}_{\mathbf{C}}(A, B)$ and f' be the single element in $\mathrm{Hom}_{\mathbf{C}}(B, A)$. Both the identity morphism id_A and $f' \circ f$ belong to $\mathrm{Hom}_{\mathbf{C}}(A, A)$ which must have cardinality 1 because A is initial. Similarly id_B and $f \circ f'$ belong to $\mathrm{Hom}_{\mathbf{C}}(B, B)$ which has cardinalty 1 because B is initial. We conclude that $f' \circ f = \mathrm{id}_A$ and $f \circ f' = \mathrm{id}_B$. In words, f and f' are inverses, thus $A \cong B$. □

Corollary 171 (Dual)**.** *Let* $\mathbf{C}$ *be a category and* $A, B \in \mathbf{C}_0$ *be terminal, then* $A \cong B$.[161]

[161] From now on, I let you prove many dual results on your own — I will try to continue doing the complicated ones. They are not necessarily great exercises, but you can certainly do them if you want to follow this book at a slower pace.

Rewording the last two results, we can say that initial (resp. terminal) objects are unique up to isomorphisms. However, the situation is quite nicer. Initial (resp. terminal) objects are unique up to *unique* isomorphisms. Indeed, if there is an isomorphism $f : A \to B$ and A and B are initial (resp. terminal), then, by definition, f is the unique morphism in $\mathrm{Hom}_{\mathbf{C}}(A, B)$.

SOL **Exercise 172.** Show that in **Cat**, the initial object is the empty category (no objects and no morphisms) and the terminal object is the category with one object $\mathbf{1}$ (hence the agreeing notation).[162]

[162] **Hint**: the unique functor $\langle\rangle : \mathbf{C} \to \mathbf{1}$ is the constant functor at the object $\bullet \in \mathbf{1}_0$.

Example 173 (Grp). Similarly to **Set**, the trivial group with one element is terminal in **Grp**. Moreover, note that there are no empty group (because a group must contain an identity element), but any group homomorphism from the trivial group $\{1\}$ into a group G must send 1 to 1_G, which completely determines the homomorphism. Therefore, the trivial group is also initial in **Grp**, it is the zero object.

Example 174 (Met). The terminal object in **Met** is the space with only one point $*$. The distance is determined by the axioms on a metric, because $d_{\mathbf{1}}(*, *)$ must be equal to 0.[163] The initial object in **Met** is the empty space, for the same reason that $\varnothing$ is initial in **Set**.

[163] The function sending all of X to $*$ is nonexpansive whatever the distance d on X because $d(x, y) \geq 0 = d_{\mathbf{1}}(*, *)$.

SOL **Exercise 175.** Find the initial and terminal objects in $\mathbf{Set}_*$.

SOL **Exercise 176.** Find the initial and terminal objects in $\mathbf{Vect}_k$.

SOL **Exercise 177.** Find a category with only two objects X and Y such that

(a) X is initial but not terminal and Y is terminal but not initial.

(b) X is initial but not terminal and Y neither terminal nor initial.

(c) X is terminal but not initial and Y is neither terminal nor initial.

(d) X is initial and terminal and Y is neither terminal nor initial.

Examples 178. Here are more examples of categories where initial and terminal objects may or may not exist.

1. $\exists$ terminal, $\nexists$ initial: Consider the poset $(\mathbb{N}, \geq)$ represented by diagram (28). It is clear that 0 is terminal and no element can be initial because $0 \geq x$ implies $x = 0$.

$$\overset{0}{\bullet} \longleftarrow \overset{1}{\bullet} \longleftarrow \overset{2}{\bullet} \longleftarrow \cdots \qquad (28)$$

2. $\nexists$ terminal, $\exists$ initial:[164] Recall the category **SetInj** of finite sets and injective functions. The empty set is still initial but the singletons are not terminal because a function from a set S into $\{*\}$ is never injective when $|S| > 1$.

[164] Of course, you could take the opposite of $(\mathbb{N}, \geq)$, that is $(\mathbb{N}, \leq)$, but that is not fun.

3. $\nexists$ terminal, $\nexists$ initial: Let G be a non-trivial group, the delooping of G has no terminal and no initial objects. The category $\mathbf{B}G$ has a single object $*$ with $\mathrm{Hom}_{\mathbf{B}G}(*, *) = G$, so $*$ cannot be initial nor terminal when $|G| > 1$.

   For a more interesting example, consider the category **Field**. Its underlying directed graph is disconnected[165] because there are no field homomorphisms between fields of different characteristic. Therefore, **Field** has no initial nor terminal objects.

[165] There are objects with no morphisms between them.

4. $\exists$ terminal, $\exists$ initial: The empty set is both initial and terminal in the category **Rel** because a relation $\emptyset \to A$ (resp. $A \to \emptyset$) is a subset of $\emptyset \times A$ (resp. $A \times \emptyset$), and the latter has a unique subset for all sets A.

   For an example with no zero object, let X be a non-empty topological space where τ is the collection of open sets.[166] The category of open sets $\mathcal{O}(X)$ satisfies

$$\mathrm{Hom}_{\mathcal{O}(X)}(U, V) = \begin{cases} \{i_{U,V}\} & U \subseteq V \\ \emptyset & U \not\subseteq V \end{cases}$$

   Since the empty set is contained in every open set, it is an initial object. Since the full set X contains every open set, it is a terminal object. Any other set cannot be initial as it cannot be contained in $\emptyset$ nor terminal as it cannot contain X. Moreover, note that the two objects are not isomorphic because $X \not\subseteq \emptyset$.

[166] Recall that it must contain $\emptyset$ and X.

SOL **Exercise 179.** Let $\mathbf{C}$ be a category with a terminal object $\mathbf{1}$. Show that any morphism $f : \mathbf{1} \to X$ is monic. State and prove the dual statement.

SOL **Exercise 180.** Let $\mathbf{C}$ and $\mathbf{D}$ be categories, and $\mathbf{1_C}$ and $\mathbf{1_D}$ be terminal objects in $\mathbf{C}$ and $\mathbf{D}$ respectively. Show that $(\mathbf{1_C}, \mathbf{1_D})$ is terminal in the $\mathbf{C} \times \mathbf{D}$. State and prove the dual statement.

Example 181. For our last application of duality in this section,[167] let X be a set and consider the posetal category $(\mathcal{P}X, \subseteq)$. We would like to define the union of

[167] Do not worry, we will have plenty of opportunities to use duality later.

two subsets of X in this category. The usual definition $A \cup B = \{x \in X \mid x \in A \text{ or } x \in B\}$ is not suitable because the data in the posetal category $\mathcal{P}X$ never refers to elements of X. In particular, the subsets $A, B \subseteq X$ are simply objects in the category and it is not clear to us how we can determine what elements are in A and B with our categorical tools (objects and morphisms).

We propose another characterization of the union of A and B. First, what is obvious, $A \cup B$ contains A and it contains B. Second, $A \cup B$ is the smallest subset of X containing A and B. Indeed, if $Y \subseteq X$ contains all elements in A and B, then it also contains $A \cup B$. Using the order $\subseteq$ (or equivalently, the morphisms in the category $\mathcal{P}X$), we have[168]

[168] We leave it as an exercise to show that $A \cup B$ is the only subset of X satisfying this property.

$$A, B \subseteq A \cup B \text{ and } \forall Y \text{ s.t. } A, B \subseteq Y \text{ then } A \cup B \subseteq Y.$$

This yields a definition of $\cup$ within the category $\mathcal{P}X$, which means we can look in the opposite of $\mathcal{P}X$ and dualize $\cup$.

The dual of this property (reversing all inclusions) is as follows.[169]

[169] The symbol $\square$ is a placeholder for the operation which we will find to be dual to union.

$$A \square B \subseteq A, B \text{ and } \forall Y \text{ s.t. } Y \subseteq A, B \text{ then } Y \subseteq A \square B$$

Putting this in words, $A \square B$ is the largest subset of X which is contained in A and B. That is, of course, the intersection $A \cap B$. In this sense, union and intersection are dual operations. If you search your memory for properties about union and intersection that you proved when you first learned about sets, you will find that they usually come in pairs, the first property being the dual of the second.[170]

[170] e.g.:

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$
$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

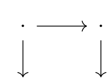
2.4 More Vocabulary

In the next chapter, we will start heavily using diagrams, and in order to generalize many concepts relying on diagrams, we will need a formal abstract definition of diagrams to work with. We introduce this definition here[171] and throw in a couple of new concepts and their duals to keep practicing with the central idea in this chapter.

[171] In the rest of the book, we use the term *diagram* to refer to both the informal pictures we draw and the formal mathematical object defined below. The context should disambiguate the two usages, but if you are not sure, rememeber that only the latter use will appear with a hyperlink on the word that links to Definition 182.

Definition 182 (Diagram)**.** A **diagram** in $\mathbf{C}$ is a functor $F : \mathbf{J} \rightsquigarrow \mathbf{C}$ where $\mathbf{J}$ is usually a small or even finite category. We say that J is the **shape** of the diagram F.

Remark 183. Diagrams are usually represented by (partially) drawing the image of F. While all the informal diagrams drawn up to this point can correspond to actual formal diagrams, it is not very pertinent to highlight this correspondence in a case-by-case basis. Indeed, the motivation behind Definition 182 is the need to abstract away from the drawings to work in more generality. For instance, when considering a commutative square in $\mathbf{C}$, it can be helpful to view it as the image of a functor with codomain $\mathbf{C}$ and domain the category $\mathbf{2} \times \mathbf{2}$ represented in (29).



(29)

Since diagrams are defined as functors, they interact well with other functors. For example, if $F : \mathbf{J} \rightsquigarrow \mathbf{C}$ is a diagram of shape $\mathbf{J}$ in $\mathbf{C}$ and $G : \mathbf{C} \rightsquigarrow \mathbf{D}$ is a functor, then $G \circ F$ is a diagram of shape $\mathbf{J}$ in $\mathbf{D}$. Some functors interact even more nicely with diagrams.

Definition 184. Let $F : \mathbf{C} \rightsquigarrow \mathbf{C}'$ be a functor and P a property[172] of diagrams.

- We say that F **preserves** diagrams with property P if for any diagram $D : \mathbf{J} \rightsquigarrow \mathbf{C}$, if D has property P, then $F \circ D$ has property P.

- We say that F **reflects** diagrams with property P if for any diagram $D : \mathbf{J} \rightsquigarrow \mathbf{C}$, if $F \circ D$ has property P, then D has property P.

[172] This is intentionally a vague term. In Chapter 3, we will have a more formal but less general definition of preserving and reflecting.

Warning 185. Preserving and reflecting a property P are not dual notions. The dual of preserving (resp. reflecting) P is preserving (resp. reflecting) the dual of P.

Example 186 (Commutativity)**.** By drawing the objects and morphisms in the image of a diagram $D : \mathbf{J} \rightsquigarrow \mathbf{C}$, we can still use Definition 99 to say whether D is commutative or not. Since functors preserve composition, if D is commutative and $F : \mathbf{C} \rightsquigarrow \mathbf{D}$ is any functor, $F \circ D$ is also commutative. Indeed, if two paths in $\mathbf{C}$ compose to the same morphism, then the composites of the paths after applying F are still equal. In other words, all functors preserve commutativity. We will use this fact many times in proofs[173] by drawing a commutative diagram and applying F to all objects and morphisms to get another commutative diagram.

[173] Without the rigor of defining the functors represented by the diagrams.

Commutativity is not reflected by all functors. Even if a diagram $D : \mathbf{J} \rightsquigarrow \mathbf{C}$ does not commute, composing D with the unique functor into the terminal category $\mathbf{1}$ yields a (trivially) commutative diagram $\langle\rangle \circ D : \mathbf{J} \rightsquigarrow \mathbf{1}$.

If $F : \mathbf{C} \rightsquigarrow \mathbf{D}$ is faithful, then F reflects commutativity. Let $D : \mathbf{J} \rightsquigarrow \mathbf{C}$ be a diagram and suppose $F \circ D$ is commutative. As in Definition 99, take a path in the image of D of length greater than one that composes to $p_1 : A \to B$ and another path that composes to $p_2 : A \to B$. After applying F, commutativity of $F \circ D$ ensures the two paths compose to the same morphism $p \in \mathrm{Hom}_{\mathbf{D}}(FA, FB)$. Moreover, p is the image of both p_1 and p_2, and since F is faithful, we conclude that $p_1 = p_2$.

The following two exercises are a quick investigation in preservation and reflection of simple properties we have seen in this chapter.

SOL Exercise 187. 1. Find an example of functor which does not preserve monomorphisms.[174]

2. Show that if $f \in \mathbf{C}_1$ is a split monomorphism, then $F(f)$ is also a split monomorphism, i.e.: any functor preserves split monomorphisms.

3. State and prove the dual statement.

4. Infer that all functors preserve isomorphisms, in particular functors send isomorphic objects to isomorphic objects.

[174] We can see a morphism as a diagram of shape $\mathbf{2}$ because a functor $\mathbf{2} \rightsquigarrow \mathbf{C}$ amounts to a choice of a morphism in $\mathbf{C}$. Thus, a functor F preserves monomorphisms if and only if whenever f is monic, $F(f)$ also is.

SOL Exercise 188. 1. Find an example of functor which does not reflect monomorphisms.[175]

2. Show that if F is faithful, then F reflects monomorphisms.

3. State and prove the dual statement.

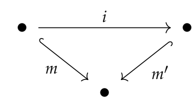
4. Show that if F is fully faithful, then F reflects isomorphisms.

[175] A functor reflects monomorphisms if whenever Ff is monic, f also is.

We have seen how to *categorify*[176] unions and intersections of subsets in Example 181. The next set-theoretical notion we categorify is subsets. A subset $I \subseteq S$ can be identified with the inclusion function $I \hookrightarrow S$, and since the latter is injective, we may want to consider monomorphisms with target S to be some kind of generalized subset. Observe however that an injection $I \rightarrowtail S$ is not necessarily an inclusion function. This does not matter because, in reality, we are interested in the image of this injection. We run into another obstacle because if two injections into S have the same image, they represent the same subset. We overcome this using the following exercise.

[176] Categorification is an imprecise term referring to the process of casting an idea in a more categorical language. Depending on the original idea and the context where it is used, there can be many ways to describe it with a categorical mind. In the following two chapters, we will spend some time categorifying several set-theoretical notions.

Two monomorphisms related by $\sim$.



SOL **Exercise 189.** Let $\mathbf{C}$ be a category and $X \in \mathbf{C}_0$, we define the relation $\sim$ on monomorphisms with target X by

$$m \sim m' \Leftrightarrow \exists \text{ isomorphism } i, m = m' \circ i.$$

Show that $\sim$ is an equivalence relation.

Definition 190 (Subobject)**.** Let $\mathbf{C}$ be a category, a **subobject** of $X \in \mathbf{C}_0$ is an equivalence class of the relation $\sim$ defined above. We will often abusively refer to a subobject simply with a monomorphism $Y \rightarrowtail X$ representing the class. The collection of subobjects of X is denoted $\mathrm{Sub}_{\mathbf{C}}(X)$. If for any $X \in \mathbf{C}_0$, $\mathrm{Sub}_{\mathbf{C}}(X)$ is a set, we say that $\mathbf{C}$ is **well-powered**.

Example 191 (**Set**)**.** Let $X \in \mathbf{Set}_0$, subobjects of X correspond to subsets of X.[177] Indeed, any subset $I \subseteq X$ has an inclusion function $i : I \hookrightarrow X$ which is injective, hence monic. For the other direction, we can show that $i : I \rightarrowtail X$ and $j : J \rightarrowtail X$ are in the same equivalence class in $\mathrm{Sub}_{\mathbf{Set}}(X)$ if and only if $\mathrm{Im}(i) = \mathrm{Im}(j)$.[178] We conclude that the correspondence between $\mathrm{Sub}_{\mathbf{Set}}(X)$ and $\mathcal{P}(X)$ sends $[i]$ to the image of i and $I \subseteq X$ to the equivalence class of the inclusion $i : I \hookrightarrow X$.

[177] The notation $\mathrm{Sub}_{\mathbf{Set}}(X)$ is perfect!

[178] ($\Rightarrow$) If $i \sim j$, then there exists a bijection f such that $i = j \circ f$. It follows that the image of j is the image of i.
($\Leftarrow$) Suppose $\mathrm{Im}(i) = \mathrm{Im}(j)$, we define $f : I \to J = x \mapsto j^{-1}(i(x))$, where j^{-1} is the left inverse of j. It is clear that $i = j \circ f$ and a quick computation shows f is an isomorphism with inverse $x \mapsto i^{-1}(j(x))$, where $i^{-1}(x)$ is the left inverse of i.

The next exercise generalizes the poset of subsets of X $(\mathcal{P}X, \subseteq)$.

SOL **Exercise 192.** Let $\mathbf{C}$ be a category and $X \in \mathbf{C}_0$, we define the relation $\leq$ on $\mathrm{Sub}_{\mathbf{C}}(X)$:

$$[m] \leq [m'] \Leftrightarrow \exists \text{ morphism } k, m = m' \circ k.$$

Show that $\leq$ is a well-defined partial order.

SOL **Exercise 193.** Show that the correspondence between $\mathcal{P}X$ and $\mathrm{Sub}_{\mathbf{Set}}(X)$ from Example 191 is an isomorphism of posets $(\mathcal{P}X, \subseteq) \cong (\mathrm{Sub}_{\mathbf{Set}}(X), \leq)$.[179]

[179] We saw what poset isomorphisms were in Example 165.2.

SOL **Exercise 194.** Show that a subobject in **Cat** is a subcategory.

We can use duality to obtain (for free) the notion of quotient objects.

Definition 195 (Quotients)**.** Let $\mathbf{C}$ be a category and $X \in \mathbf{C}_0$, there is an equivalence relation $\sim$ on epimorphisms with source X defined by

$$q \sim q' \Leftrightarrow \exists \text{ isomorphism } i, q = i \circ q'.$$

A **quotient object** (or simply quotient) of X is an equivalence class of the relation $\sim$ defined above.[180] The collection of quotients of X is denoted $\mathrm{Quot}_{\mathbf{C}}(X)$. If for any $X \in \mathbf{C}_0$, $\mathrm{Quot}_{\mathbf{C}}(X)$ is a set, we say that $\mathbf{C}$ is **co-well-powered**. There is a partial order $\leq$ on $\mathrm{Quot}_{\mathbf{C}}(X)$ defined by

[180] We will often abusively refer to a quotient simply with an epimorphism $X \twoheadrightarrow Y$.

$$[q] \leq [q'] \Leftrightarrow \exists \text{ morphism } k, q = k \circ q'.$$

The terminology for this dual notion is motivated by the following exercise.

SOL **Exercise 196.** Show that a quotient object of $G \in \mathbf{Grp}_0$ is a quotient group of G.

I love finding a categorical definition for something I am used to thikning of in classical terms.[181] It facilitates a better understanding of the essential components of the classical notion, and duality can open the gates to a parallel world where we can have just as much fun.

[181] This feeling led me to study more category theory.

For now, we only played with definitions without discovering anything deep. Some people maintain it is useless to take a categorical point of view if it does not lead to new results. Category theorists (I presume) believe that it helps organize our thoughts regardless of the mathematical outcomes. The rest of the book focuses on practicing categorical thinking without necessarily demonstrating its advantages other than its unifying/orgasitional power.

3 Limits and Colimits

3.1 Examples 53
3.2 Generalization 67
3.3 Diagram Chasing 74

The unifying power of categorical abstraction is arguably its biggest benefit. Indeed, it is often the case that many mathematical objects or results from different fields fit under the same categorical definition or fact. In my opinion, category theory is at its peak of elegance when a complex idea becomes close to trivial when viewed categorically, and when this same view helps link together the intuitions behind many ideas throughout mathematics.

The next two chapters concern one particular instance of this power: universal constructions. Along with Chapter 6, these three chapters constitute the heart of our investigation into a philosophical idea central to category theory:[182]

> A mathematical object is completely determined by its relations with other objects of the same kind.

[182] We already hinted at it in Chapter 1. I am not a good philosopher of mathematics, but I believe this statement is a fundamental belief in structuralism.

This chapter will cover limits and colimits which are special cases of universal constructions. We postpone the rigorous definition of the term "universal", so, for a while, I recommend you try to recognize *universality* as the thing that all definitions of (co)limits given below have in common.[183]

[183] This is also a good practice for reading more literature on category theory since "universal" can also be used informally.

The first section presents several examples. Each of its subsection is dedicated to one kind of limit or colimit of which a detailed example in **Set** is given along with a couple of interesting examples in other categories. It is not straightforward to build intuition about all kinds of (co)limits due to their innumerable applications. For now, I think it is fine if you are comfortable with the intuition in **Set** as it transposes well to concrete categories, but if you persist in learning category theory, you will get to see examples with other flavors. The second section gives a formal framework to talk about all the examples previously explored as well as a few general results. The third section is a training ground to practice a new proof technique called diagram chasing,[184] we will cover important results there too.

[184] It extends diagram paving using the tools seen in the chapter.

In the sequel, **C** denotes a category.

3.1 Examples

Before giving the definition of (co)limits which is very abstract, we present several examples of how they are used. These are very interesting on their own because they show you how a lot of things mathematicians care about in different contexts can be seen as the same abstract construction. Still, keep in mind that, after adding another level of abstraction, we will bring all these examples together as instances of (co)limits.

Products

Given two sets S and T, the most common construction of the Cartesian product $S \times T$ is conceptually easy: you take all pairs of elements S and T, that is,

$$S \times T := \{(s,t) \mid s \in S, t \in T\}.$$

This construction requires to pick out elements in S and T, form pairs of elements, and use the set-builder notation. While these steps are straightforward set-theoretically, it is not so clear how one would translate them into categorical language.[185] You can try to do it for the first step.

[185] Only working with the objects and morphisms of the category **Set**.

SOL **Exercise 197.** Inside the category **Set**, give a categorical definition of an element of a set. Your definition must only refer to objects and morphisms, so it can be generalized to other categories. Does your definition still correspond to an intuitive notion of elements inside **Poset**, **Grp**, **Cat**?

If one hopes to generalize products to other categories, the construction must only involve objects and morphisms.

Question 198. *What are essential functions (morphisms in* **Set***) to consider when studying $S \times T$?*

Answer. Projection maps. They are functions $\pi_1 : S \times T \to S$ and $\pi_2 : S \times T \to T$,[186] but that is not enough to define the product. Indeed, there are projection maps $\pi_1' : S \times T \times S \to S$ and $\pi_2' : S \times T \times S \to T$, but $S \times T \times S$ is not always isomorphic to $S \times T$. □

[186] The projections are defined by $\pi_1(s,t) = s$ and $\pi_2(s,t) = t$ for all $(s,t) \in S \times T$.

Question 199. *What is unique*[187] *about $S \times T$ with the projections π_1 and π_2?*

[187] Always up to isomorphism of course.

Answer. For one, π_1 and π_2 are surjective, and while they are not injective, they have an invertible-like property. Namely, given $s \in S$ and $t \in T$, the pair (s,t) is completely determined from $\pi_1^{-1}(s) \cap \pi_2^{-1}(t)$. □

Again, in order to get rid of the references to specific elements, another point of view is needed. Let X be a set of *choices* of pairs, an element $x \in X$ chooses elements in S and T via functions $c_1 : X \to S$ and $c_2 : X \to T$ (similar to the projections). Now, the *almost-inverse* defined above yields a function

$$! : X \to S \times T = x \mapsto \pi^{-1}(c_1(x)) \cap \pi^{-1}(c_2(x)).$$

This function maps $x \in X$ to an element in $S \times T$ that makes the same choice as x, and it is the only one that does so. Categorically, ! is the unique morphism in $\mathrm{Hom}_{\mathbf{C}}(X, S \times T)$ satisfying $\pi_i \circ ! = c_i$ for $i = 1,2$. Later, we will see that this property completely determines $S \times T$. For now, enjoy the power we gain from generalizing this idea.

Definition 200 (Binary product)**.** Let $A, B \in \mathbf{C}_0$. A (categorical) **binary product** of A and B is an object, denoted $A \times B$, along with two morphisms $\pi_A : A \times B \to A$ and $\pi_B : A \times B \to B$ called **projections** that satisfy the following universal property[188]: for every object $X \in \mathbf{C}_0$ with morphisms $f_A : X \to A$ and $f_B : X \to B$, there is a unique morphism $! : X \to A \times B$ making diagram (30) commute.

[188] Remember that the word universal is not yet defined, we are trying to get an idea of what it means with these examples.

$$\begin{array}{ccccc} & & X & & \\ & \swarrow^{f_A} & \downarrow ! & \searrow^{f_B} & \\ A & \xleftarrow[\pi_A]{} & A \times B & \xrightarrow[\pi_B]{} & B \end{array} \tag{30}$$

We will often denote $! = \langle f_A, f_B \rangle$ and call it the **pairing** of f_A and f_B.

Example 201 (**Set**)**.** Cleaning up the argument above, we show that the Cartesian product $A \times B$ with the usual projections is a binary product in **Set**. To show that it satisfies the universal property, let X, f_A and f_B be as in the definition. A function $! : X \to A \times B$ that makes (30) commute must satisfy

$$\forall x \in X, \pi_A(!(x)) = f_A(x) \text{ and } \pi_B(!(x)) = f_B(x).$$

Equivalently, $!(x) = (f_A(x), f_B(x))$. Since this uniquely determines $!$, $A \times B$ is indeed the binary product.

Example 202. Most of the constructions throughout mathematics with the name *product* can also be realized with a categorical product. Examples include the product of groups, rings or vector spaces, the product of topologies, etc. The fact that all these constructions are based on the Cartesian product of the underlying sets is a corollary of a deeper result about the forgetful functors that all these categories have in common.[189]

[189] We show in Chapter 7 that these forgetful functors are right adjoints and thus they preserve binary products (Proposition 436).

Let us give the details for **Mon**, they can be easily adapted for the other categories of algebraic objects (groups, rings, vector spaces) — this does not translate so readily for the product of two topological spaces.

Example 203. In another flavor, let X be a topological space and $\mathcal{O}(X)$ be the category of opens. If $A, B \subseteq X$ are open, what is their product? Following Definition 200, the existence of π_A and π_B imply that $A \times B$[190] is included in both sets, or equivalently $A \times B \subseteq A \cap B$.

[190] Recall that $\times$ denotes the categorical product, not the Cartesian product of sets.

Moreover, for any open set X included in A and B (via f_A and f_B), X should be included in $A \times B$ (via $!$).[191] In particular, X can be $A \cap B$ (it is open by definition of a topology), thus $A \cap B \subseteq A \times B$. In conclusion, the product of two open sets is their intersection. In an arbitrary poset, the same argument is used to show the product is the greatest lower bound/infimum/meet.

[191] Notice that uniqueness of $!$ is already given in a posetal category.

Remark 204. Given two objects in an arbitrary category, their product does not necessarily exist. Nevertheless, when it exists, one can (and we will) show that it is unique up to unique isomorphism.[192] Thus, in the sequel, we will speak of *the* product of two objects and similarly for other constructions presented in this chapter. Moreover, we will often refer to the object $A \times B$ alone (without the projections) as the product.

[192] The uniqueness of the isomorphism is under the condition that it preserves the structure of the product. We will clear up this subtlety in Remark 260.

SOL Exercise 205. Let (A, R) and (B, S) be two objects in **2Rel**.[193] We denote $R \wedge S$ the binary relation on $A \times B$ defined by (we write the relations infix like for orders)

$$(a, b)\, R \wedge S\, (a', b') \Leftrightarrow a\, R\, a' \text{ and } b\, S\, b'.$$

1. Show that $(A \times B, R \wedge S)$ is the product of (A, R) and (B, S) in **2Rel**.
2. Show that if R and S are reflexive/transitive/antisymmetric, then so is $R \wedge S$.
3. Conclude that, in both **Poset** and **Pre**, the product of any two objects exists.

[193] i.e. A and B are sets and $R \subseteq A \times A$ and $S \subseteq B \times B$.

SOL Exercise 206. Let A and B be two sets, find their product in the category **Rel**.

SOL Exercise 207. Let **C** and **D**, be two categories, we defined the product category $\mathbf{C} \times \mathbf{D}$ in Definition 132. Resolve the clash of notations by checking that $\mathbf{C} \times \mathbf{D}$ satisfies the universal property of the categorical product of **C** and **D**.

Before reaching even more generality, it is sane to check that we can prove some properties of the Cartesian product using the categorical definition. This would ensure that we are not venturing in useless abstract nonsense. We prove the harder one and leave you two easier ones as exercises.

Proposition 208. *Let $A, B, C \in \mathbf{C}_0$ be such that $A \times B$ and $B \times C$ exist. If $A \times (B \times C)$ exists, then $(A \times B) \times C$ exists and both products are isomorphic. In other words, the binary product is associative.*[194]

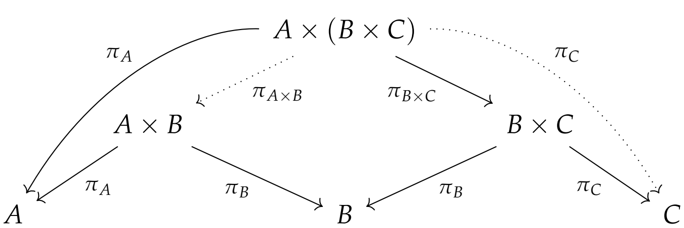
[194] Just like the Cartesian product is associative (up to isomorphism). The existence hypothesis is not necessary in **Set** because the Cartesian product of any two sets always exists.

Proof. We will show that $A \times (B \times C)$ satisfies the definition of the product $(A \times B) \times C$ with projections defined below. This means $(A \times B) \times C$ exists and the fact that $A \times (B \times C) \cong (A \times B) \times C$ follows trivially (we defined them to be the same object).[195]

First, we need two projections $\pi_{A \times B} : A \times (B \times C) \to A \times B$ and $\pi_C : A \times (B \times C) \to C$. In the diagram below, we show how to obtain them.[196]

[195] In any case, as we will prove in Proposition 259, if you had another construction for $(A \times B) \times C$, it would be isomorphic to ours.

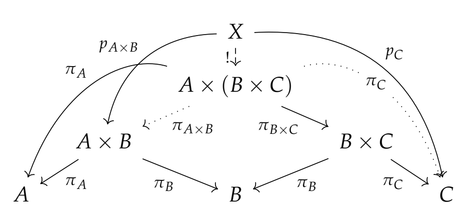
[196] We overload the notation and rely on the source and target of the morphisms to avoid confusion



(31)

The dotted arrow π_C is simply the composition $\pi_C \circ \pi_{B \times C}$. The dotted arrow $\pi_{A \times B}$ is obtained via the property of the product $A \times B$ and the morphisms $\pi_A : A \times (B \times C) \to A$ and $\pi_B \circ \pi_{B \times C} : A \times (B \times C) \to B$. It is the unique morphism making (31) commute, that is, $\pi_{A \times B} = \langle \pi_A, \pi_B \circ \pi_{B \times C} \rangle$.

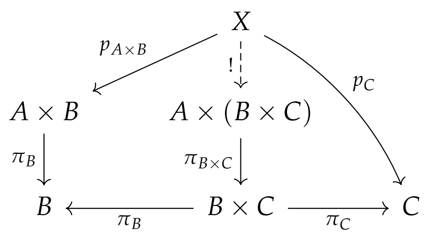
Suppose there is an object X and morphisms $p_{A \times B} : X \to A \times B$ and $p_C : X \to C$. We need to find $! : X \to A \times (B \times C)$ that makes (32) commute and is unique with that property. By post-composing with the appropriate projections, we can see how ! acts from the point of view of A, B and C:



(32)

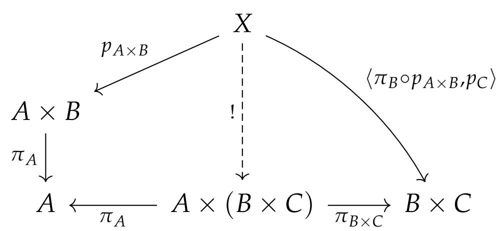
$$\begin{aligned}
\pi_A \circ ! &= \pi_A \circ \langle \pi_A, \pi_B \circ \pi_{B\times C} \rangle \circ ! = \pi_A \circ p_{A\times B} \\
\pi_B \circ \pi_{B\times C} \circ ! &= \pi_B \circ \langle \pi_A, \pi_B \circ \pi_{B\times C} \rangle \circ ! = \pi_B \circ p_{A\times B} \\
\pi_C \circ \pi_{B\times C} \circ ! &= p_C.
\end{aligned}$$

The last two equations tell us that $\pi_{B\times C} \circ !$ must make (33) commute.



(33)

Hence, by the universal property of $B \times C$, we must have $\pi_{B\times C} \circ ! = \langle \pi_B \circ p_{A\times B}, p_C \rangle$. This fact combined with the first equation tells us that ! makes (34) commute.



(34)

Hence, by the universal property of $A \times (B \times C)$, we must have $! = \langle \pi_A \circ p_{A\times B}, \langle \pi_B \circ p_{A\times B}, p_C \rangle \rangle$. Notice that the two uses of universal properties ensured that we found the unique possible choice for !. □

Remark 209. This has been our first proof using **diagram chasing**. It is different from diagram paving because the goal is to construct objects and morphisms that make some diagram commute (often with a proof of uniqueness of your construction).[197] Another unfortunate difference is that diagram chasing proofs are much harder to typeset. On the board, this proof can be done with one big diagram on which we point out the relevant parts at different moments in the proof. Here, we had to draw four diagrams for this proof in order to emphasize different parts of that huge diagram.

[197] In diagram paving, you only use objects and morphisms that are given. One can see diagram paving as part of diagram chasing because the commutativity proofs are done by combining smaller commutative diagrams.

Here are two simpler diagram chasing exercises for you to solve. It should help to highlight the important steps of the proof above. To show $A \times (B \times C)$ is the same thing as $(A \times B) \times C$, we showed the former satisfies the universal property of the latter. We built the appropriate projections, and given another object with maps to $A \times B$ and C, we showed how to construct the pairing of these maps, and finally we showed that pairing was unique.

SOL **Exercise 210.** Let $A, B \in \mathbf{C}_0$. If $A \times B$ exists, then $B \times A$ exists and both products are isomorphic. In other words, the binary product is commutative.[198]

[198] Just like the Cartesian product is commutative (up to isomorphism).

This statement is transparent in the definition of binary products because changing A for B in Definition 200 has no impact. Still, proving it is more rigorous.

SOL **Exercise 211.** Let $\mathbf{1}$ be the terminal object in $\mathbf{C}$. Show that for any $A \in \mathbf{C}_0$, the product of $\mathbf{1}$ and A is A.[199]

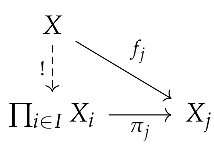
[199] This property is expected because in **Set**, $\mathbf{1} = \{*\}$ and
$$\{*\} \times A = \{(*, a) \mid a \in A\} \cong A.$$

To generalize the categorical product to more than two objects, one can, for instance, define the product of a finite family of sets recursively with the binary product.[200] This is well-defined thanks to the associativity and commutativity of $\times$, but this is not enough to get the infinite case. In contrast, generalizing the universal property illustrated in (30) yields a simpler definition that works even for arbitrary families. Instead of having only two objects and two projections, we will have a families of objects and projections indexed by an arbitrary set I.

[200] For a family $\{X_1, \dots, X_n\} \subseteq \mathbf{C}_0$:
$$\prod_{i=1}^{n} X_i = \begin{cases} X_1 & n = 1 \\ \left(\prod_{i=1}^{n-1} X_i\right) \times X_n & n > 1 \end{cases}$$

Definition 212 (Product)**.** Let $\{X\}_{i \in I}$ be an I–indexed family of objects of $\mathbf{C}$. The **product** of this family is an object $\prod_{i \in I} X_i$ along with projections $\pi_j : \prod_{i \in I} X_i \to X_j$ for all $j \in I$ satisfying the following universal property: for any object X with morphisms $\{f_j : X \to X_j\}_{j \in I}$, there is a unique morphism $! : X \to \prod_{i \in I} X_i$ making (35) commute for all $j \in I$.[201]

[201] Analogously to the binary case, we may write $! = \langle f_j \rangle_{j \in I}$ or, in the finite case, $! = \langle f_1, \dots, f_n \rangle$.



(35)

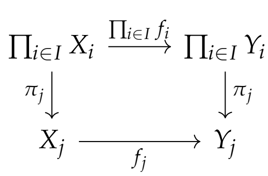
Warning 213. In a lot of cases, the arbitary product will be a straightforward generalization of the binary product,[202] but that is not true in all cases. For instance, in the category of open subsets of a topological space, the arbitrary product is not always the intersection. This is because arbitrary intersections of open sets are not necessarily open. To resolve this problem, it suffices to take the interior of the intersection which is open by definition.

[202] e.g. in **Set**, the Cartesian product of an arbitrary family of sets is still the set of ordered tuples (instead of pairs) of elements in the sets.

Commutativity and now associativity of categorical products are true by definition.[203] Here are three more properties of Cartesian products that generalize to categorical products.

[203] We mean the order of the X_is is not taken into account for the universal property. As we did for binary products, we will make this more rigorous in ...

SOL **Exercise 214** (NOW!)**.** Let $\{f_i : X_i \to Y_i\}_{i \in I}$ be a family of morphisms in $\mathbf{C}$, show that there is a unique morphism $\prod_{i \in I} f_i : \prod_{i \in I} X_i \to \prod_{i \in I} Y_i$ making the following square commute for all $j \in I$.



(36)

We call $\prod_{i \in I} f_i$ the **product** of the f_is. In the finite case, we write $f_1 \times \cdots \times f_n$.

In **Set**, the function $\prod_{i \in I} f_i$ acts on tuples in $\prod_{i \in I} X_i$ by applying f_i to the ith coordinate for every i.

SOL Exercise 215. Let X, Y and $\{X_i\}_{i\in I}$ be objects of $\mathbf{C}$ such that $\prod_{i\in I} X_i$ exists. For any family $f_i : X \to X_i$ and $g : Y \to X$ show that $\langle f_i\rangle_{i\in I} \circ g = \langle f_i \circ g\rangle_{i\in I}$. Conclude that for families $\{f_i : X_i \to Y_i\}_{i\in I}$ and $\{g_i : Z_i \to X_i\}_{i\in I}$, $(\prod f_i) \circ (\prod g_i) = \prod(f_i \circ g_i)$.[204]

[204] It may be useful to restate this in the binary case. For any $f : X \to Y$, $f' : X' \to Y'$, $g : Z \to X$ and $g' : Z' \to X'$, we have

$$(f \times f') \circ (g \times g') = (f \circ g) \times (g \circ g').$$

As a corollary, if $\mathbf{C}$ has all binary products, we get a functor $\mathbf{C} \times \mathbf{C} \rightsquigarrow \mathbf{C}$ sending (X, Y) to $X \times Y$ and (f, g) to $f \times g$.

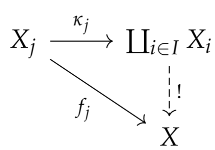
A family of objects in $\mathbf{C}$ is also called a **discrete diagram** because it corresponds to a functor from a discrete category (one with no non-identity morphisms) into $\mathbf{C}$.[205] The product of a family of objects is called the limit of the corresponding diagram. The big takeaway from last chapter is that each time we read a new definition, it is worth to dualize it. Thus, we ask: what is the colimit of a discrete diagram?

[205] Recall that a diagram is a functor into $\mathbf{C}$ (Definition 182).

Coproducts

Definition 216 (Coproduct)**.** Let $\{X\}_{i\in I}$ be an I–indexed family of objects in $\mathbf{C}$, its **coproduct** is an object, denoted $\coprod_{i\in I} X_i$ (or $X_1 + X_2$ in the binary case), along with morphisms $\kappa_j : X_j \to \coprod_{i\in I} X_i$ for all $j \in I$ called **coprojections** satisfying the following universal property: for any object X with morphisms $\{f_j : X_j \to X\}_{j\in I}$, there is a unique morphism $! : \coprod_{i\in I} X_i \to X$ making (37) commute for all $j \in I$.[206]

[206] We may denote $! = [f_j]_{j\in I}$ or, in the finite case, $! = [f_1, \ldots, f_n]$. We call it the **copairing** of $\{f_j\}_{j\in I}$.



(37)

Let us find out what coproducts of sets are.

Example 217 (**Set**)**.** Let $\{X_i\}_{i\in I}$ be a family of sets, first note that if $X_j = \emptyset$ for $j \in I$, then there is only one morphism $X_j \to X$ for any X.[207] In particular, (37) commutes no matter what $\coprod_{i\in I} X_i$ and X are. Therefore, removing X_j from this family does not change how the coproduct behaves, hence no generality is loss from assuming all X_is are non-empty.

[207] Because $\emptyset$ is initial.

Second, for any $j \in I$, let $X = X_j$, $f_j = \mathrm{id}_{X_j}$ and for any $j' \neq j$, let $f_{j'}$ be any function in $\mathrm{Hom}(X_{j'}, X_j)$.[208] Commutativity of (37) implies κ_j has a left inverse because $! \circ \kappa_j = f_j = \mathrm{id}_{X_j}$, so all coprojections are injective.

[208] One exists because X_j is non-empty.

Third, we claim that for any $j \neq j' \in I$, $\mathrm{Im}(\kappa_j) \cap \mathrm{Im}(\kappa_{j'}) = \emptyset$. Let $X = \{0, 1\}$, f_j and $f_{j'}$ be the constant functions sending everything to 0 and 1 respectively. The universal property implies that

$$\mathrm{Im}\left(! \circ \kappa_j\right) = \{0\} \neq \{1\} = \mathrm{Im}(! \circ \kappa_{j'}),$$

hence for any $x \in X_j$ and $x' \in X_{j'}$, we have $\kappa_j(x) \neq \kappa_{j'}(x')$.

In summary, the previous points say that $\coprod_{i\in I} X_i$ contains distinct copies of the images of all coprojections. Furthermore, the κ_js being injective, their image can be identified with the X_js to obtain[209]

[209] The symbol $\sqcup$ denotes the disjoint union of sets.

$$\bigsqcup_{i\in I} X_i \subseteq \coprod_{i\in I} X_i.$$

For the converse inclusion, in (37), let X be the disjoint union and the f_js be the inclusions. Assume there exists x in the R.H.S. that is not in the L.H.S., then we can define $!' : \coprod_{i \in I} X_i \to \bigsqcup_{i \in I} X_i$ that only differs from ! at x. Since x is not in the image of any coprojection, the diagrams still commute and this contradicts the uniqueness of !.

In conclusion, the coproduct in **Set** is the disjoint union and the coprojections are the inclusions.[210]

Remark 218. If this example looks more complicated than the product of sets, it is because we started knowing nothing concrete about coproducts of sets and gradually discovered what properties they had using specific objects and morphisms we know exist in **Set**. In contrast, we knew what products of sets were, and we just had to show they satisfied the universal property.[211]

In general, the hard part is to find what construction satisfies a universal property, proving it does is easier.

Examples 219. In the category of open sets of a space (X, τ), let $\{U_i\}_{i \in I}$ be a family of open sets and suppose $\coprod_i U_i$ exists. The coprojections yield inclusions $U_j \subseteq \coprod_i U_i$ for all $j \in I$, so $\coprod_i U_i$ must contain all U_js and thus $\cup_i U_i$. Moreover, in (37), letting f_j be the inclusion $U_j \hookrightarrow \cup_i U_i$ for all $j \in I$,[212] the existence of ! yields an inclusion $\coprod_i U_i \subseteq \cup_i U_i$. We conclude that the coproduct in this category is the union of open sets. In an arbitrary poset, the same argument is used to show the coproduct is the least upper bound/supremum/join.

In $\mathbf{Vect}_k$, the coproduct, also called the direct sum, is defined by[213]

$$\coprod_{i \in I} V_i = \bigoplus_{i \in I} V_i := \left\{ \vec{v} \in \prod_{i \in I} V_i \mid \vec{v}_i \neq 0 \text{ for finitely many } i\text{'s} \right\},$$

where $\kappa_j : V_j \hookrightarrow \coprod_i V_i$ sends v to $\kappa_j(v) \in \prod_i V_i$ satisfying $\kappa_j(v)_j = v$ and $\kappa_j(v)_{j'} = 0$ whenever $j \neq j'$. To verify this, let $\{f_j : V_j \to X\}_{j \in I}$ be a family of linear maps. We can construct ! by defining it on basis elements of the direct sum, which are just the basis elements of all V_js seen as elements of the sum (via the coprojections).[214] Indeed, if b is in the basis of V_j, we let $!(\kappa_j(b)) = f_j(b)$. Extending linearly yields a linear map $! : \coprod_i V_i \to X$. Uniqueness is clear because if $h : \coprod_i V_i \to X$ differs from ! on one of the basis elements, it does not make (37) commute.

SOL **Exercise 220.** Let A and B be two sets, show that their coproduct exists in the category **Rel** and find what it is.

SOL **Exercise 221.** Show that products are dual to coproducts, namely, if a product of a familiy $\{X_i\}_{i \in I}$ exists in $\mathbf{C}$, then this object and the projections are the coproduct of this family and the coprojections in $\mathbf{C}^{\mathrm{op}}$ and vice-versa. Conclude that you can define the **coproduct of morphisms** dually to Exercise 214, we denote them $\coprod_{i \in I} f_i$ or $f_1 + \cdots + f_n$ in the finite case.

Applying the duality between products and coproducts to Proposition 208 and Exercises 210 and 211, we get the following results.

[210] We recover the intuition for why empty sets can be ignored. A more general fact is proven in Exercise 221.

[211] One might argue that coming up with this universal property was the hard part in that case.

[212] These morphisms are in $\mathcal{O}(X)$ because $\cup_i U_i$ is open.

[213] Here, the symbol $\prod$ denotes the Cartesian product of the V_is as sets. The categorical product of vector spaces is also the direct sum, where the projections are the usual ones.

[214] It is necessary to require finitely many non-zero entries, otherwise the basis of the coproduct would not be the union of all bases of the V_js.

Corollary 222 (Dual)**.** *Taking binary coproducts is commutative and associative, and if* $\varnothing$ *is initial, then* $A + \varnothing \cong A$*.*[215]

[215] While in **Set**, we have $A \times \varnothing \cong \varnothing$, this does not generalize to all categories with binary products and an initial object, e.g. $\mathbf{Vect}_k$.

SOL **Exercise 223.** Dually to Exercise 215, show that if X, Y and $\{X_i\}_{i\in I}$ are objects of $\mathbf{C}$ such that $\coprod_{i\in I} X_i$ exists, then for any family $f_i : X_i \to X$ and $g : X \to Y$ show that $g \circ [f_i]_{i\in I} = [g \circ f_i]_{i\in I}$.

SOL **Exercise 224.** Let $\mathbf{C}$ have a terminal object $\mathbf{1}$. Show that the assignment $X \mapsto X + \mathbf{1}$ is functorial, i.e. define the action of $(- + \mathbf{1})$ on morphisms and show it satisfies the axioms of a functor.[216]

[216] We call $(- + \mathbf{1})$ the **maybe functor**.

In a very similar way to the product and coproduct, we will define various constructions in **Set**.[217]

[217] We will follow more closely the section on coproducts where we started with the definition of the (co)limit and then detailed an example in **Set**.

Equalizers

We briefly mentioned that a product (resp. coproduct) is a limit (resp. colimit) of a discrete diagram. The rest of the examples before generalizing will be (co)limits of small diagrams that contain non-identity morphisms.

Definition 225 (Fork)**.** A **fork** in $\mathbf{C}$ is a diagram of shape (38) or (39).

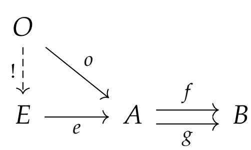
$$O \xrightarrow{o} A \underset{g}{\overset{f}{\rightrightarrows}} B \qquad (38) \qquad\qquad A \underset{g}{\overset{f}{\rightrightarrows}} B \xrightarrow{o} O \qquad (39)$$

These are dual notions, so we prefer to call (39) a **cofork**. If (38) commutes then $f \circ o = g \circ o$,[218] and we say that o **equalizes** f and g. If (39) commutes, then $o \circ f = o \circ g$, and we say that o **coequalizes** f and g.

[218] Recall that commutativity does not make parallel morphisms equal.

Definition 226 (Equalizer)**.** Let $A, B \in \mathbf{C}_0$ and $f, g : A \to B$ be parallel morphisms. The **equalizer** of f and g is an object E and a morphism $e : E \to A$ satisfying $f \circ e = g \circ e$ with the following universal property: for any morphism $o : O \to A$ equalizing f and g, there is a unique $! : O \to E$ making (40) commute.[219]

[219] Try to look for a common pattern in this definition and the definition of a product (both are instances of limits).



(40)

In other words, e is a morphism that equalizes f and g, and every other o that equalizes f and g factors through e uniquely. A common notation for e is $\mathrm{eq}(f, g)$. There is also a straightforward generalization to equalizers of more than two morphisms.[220]

[220] If $\{f_i\}_{i\in I}$ is a family of parallel morphisms, their equalizer is a morphism $e \in \mathbf{C}_1$ such that

$$\forall i, j \in I, f_i \circ e = f_j \circ e,$$

and every o with this property factors through e in a unique way.

Example 227 (**Set**)**.** Let $f, g : A \to B$ be two functions and suppose their equalizer exists and it is $e : E \to A$. By associativity, for any $h : O \to E$, the composite $e \circ h$ is a candidate for o in diagram (40) because $f \circ (e \circ h) = g \circ (e \circ h)$. What is more, if h' is such that $e \circ h = e \circ h'$, then $h = h'$ or it would contradict the uniqueness of !. We conclude that e is monic/injective.[221]

[221] This argument was independent of the category, hence we can conclude that an equalizer of parallel morphisms is always monic.

This implies E can be identified with its image under e. Since $f \circ e = g \circ e$, the image of e is contained in the subset $\{a \in A \mid f(a) = g(a)\}$. Now, by the universal property of the equalizer, letting O be this subset and o be the inclusion, there is an injection[222] $! : \{a \in A \mid f(a) = g(a)\} \hookrightarrow E$, thus both sets are equal. In conclusion, the equalizer of two parallel functions is the subset E in which they coincide and $e : E \hookrightarrow A$ is the inclusion.

[222] The fact that ! is an injection comes from the fact that the inclusion o is an injection and $e \circ ! = o$.

Examples 228. In a posetal category, hom-sets are singletons, so it must be the case that $f = g$ whenever f and g are parallel. Therefore, any $o : O \to A$ satisfies $f \circ o = g \circ o$. Written using the order notation, the universal property is then equivalent to the fact that $E \leq A$ and $O \leq A$ implies $O \leq E$. In particular, if $O = A$, then $A \leq E$, so $A = E$ by antisymmetry.

In **Ab**, **Ring** or $\mathbf{Vect}_k$, for the same reason that the Cartesian product of the underlying sets is the underlying set of the product,[223] the construction of equalizers is as in **Set**. However, since each of these categories have a notion of additive inverse for morphisms, the equalizer of f and g has a cooler name, that is, $\ker(f - g)$.[224]

[223] We explain this in Chapter 7.

[224] The equalizer of f and g is the subgroup/subring/subspace of A where f and g are equal, or equivalently, where $f - g$ is 0 (when $f - g$ and 0 are defined).

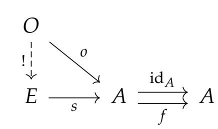
Definition 229 (Idempotents)**.** A morphism $f : A \to A \in \mathbf{C}_1$ is called **idempotent** when $f \circ f = f$. It is called **split idempotent** if there exist morphisms $s : E \to A$ and $r : A \to E$ such that $s \circ r = f$ and $r \circ s = \mathrm{id}_E$.[225]

[225] We can show that split idempotents are idempotent because

$$f \circ f = s \circ r \circ s \circ r = s \circ \mathrm{id}_E \circ r = f.$$

Proposition 230. *An idempotent morphism $f : A \to A \in \mathbf{C}_1$ is split idempotent if and only if the equalizer of id_A and f exists.*

Proof. ($\Rightarrow$) Let $f = s \circ r$ be such that $r \circ s = \mathrm{id}_E$, we claim that s is the equalizer. First, we can check that s equalizes id_A and f because $f \circ s = s \circ r \circ s = s \circ \mathrm{id}_E = s = \mathrm{id}_A \circ s$. Next, given $o : O \to A$ that also equalizes id_A and f, we need to find a morphism ! that makes (41) commute. Its uniqueness is given by s being monic (it has a left inverse). Noticing that $o = f \circ o = s \circ r \circ o$, we find $! = r \circ o$.



(41)

($\Leftarrow$) If $e : E \to A$ is the equalizer of f and id_A, then since f equalizes f and id_A, there exists $! : A \to E$ such that $e \circ ! = f$. By monicity of e, we find that $e \circ (! \circ e) = f \circ e = e$ implies $! \circ e = \mathrm{id}_A$, so f is a split idempotent (let $s = e$ and $r = !$). ☐

The first two examples had a relatively well-known instantiation in the category **Set**, namely, products are Cartesian products and coproducts are disjoint unions. The notion of equalizer of two functions, while just as intuitive as the others,[226] is less common in "classical" set theory. However, it still leads to a nice categorical definition of fiber.

[226] The equalizer of $f, g : A \to B$ is the subset of A where f and g are equal.

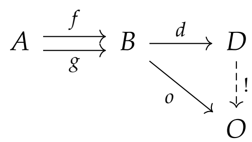
SOL **Exercise 231.** Let $f : A \to B$ be a function and $y \in B$, the *fiber* of y (under f) is $\{x \in A \mid f(x) = y\}$.[227] Give a categorical definition of fibers that does not rely on the special case of **Set**. Just like in Exercise 197, you should only refer to objects and morphisms. In particular, you can only use the categorical notion of elements (Definition 481). Does your definition still correspond to an intuitive notion of fibers inside **Poset**, **Grp**, **Cat**?

[227] Fiber is just a synonym for preimage (usually) taken at a single point.

The equalizer of f and g is the limit of the diagram containing only the two parallel morphisms, we define its colimit in the next section.

Coequalizers

Definition 232 (Coequalizer)**.** Let $A, B \in \mathbf{C}_0$ and $f, g : A \to B$ be parallel morphisms. The **coequalizer** of f and g is an object D and a morphism $d : B \to D$ satisfying $d \circ f = d \circ g$ with the following universal property: for any morphism $o : B \to O$ coequalizing f and g, there is a unique $! : D \to O$ making (42) commute.



(42)

In other words, d coequalizes f and g, and every other o that coequalizes f and g factors through d uniquely. A common notation for d is $\text{coeq}(f, g)$, and there is also a straightforward generalization to more than two morphisms.

Example 233 (Set). Let $f, g : A \to B$ be two functions and suppose $d : B \to D$ is their coequalizer. Similarly to the dual case, one can show that d is epic/surjective. Since $d \circ f = d \circ g$, for any $b, b' \in B$,

$$\big(\exists a \in A, f(a) = b \text{ and } g(a) = b'\big) \implies d(b) = d(b'). \qquad (*)$$

Denoting by $\sim$ the relation between two elements of B defined in the L.H.S. of $(*)$, the implication becomes $b \sim b' \implies d(b) = d(b')$. Note that $\sim$ is not necessarily an equivalence relation but $=$ is, thus, the converse implication does not always hold.[228]

Consequently, we consider the equivalence relation generated by $\sim$,[229] denoted by $\simeq$. As noted above, the forward implication $b \simeq b' \implies d(b) = d(b')$ still holds. For the converse, in (42), let $O := B/\simeq$ and $o : B \to B/\simeq$ be the quotient map. Post-composing with ! yields

$$d(b) = d(b') \implies o(b) = o(b') \implies b \simeq b'.$$

The equivalence $b \simeq b' \Leftrightarrow d(b) = d(b')$ and the fact that d is surjective means we can identify D with the quotient $B/\simeq$ and $d : B \to D$ with the quotient map.[230]

[228] For instance, when $b \sim b' \sim b''$, $d(b) = d(b'')$, but it might not be the case that $b \sim b''$.

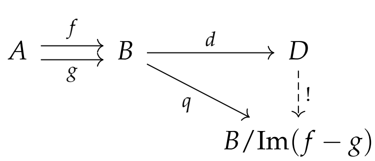
[229] In this case, it is simply the transitive closure.

[230] You can give the isomorphism $D \cong B/\simeq$.

Examples 234. In a posetal category, an argument dual to the one for equalizers shows the coequalizer of $f, g : A \to B$ is B.

In **Ab**, **Ring** or $\mathbf{Vect}_k$, let $f, g : A \to B$ be homomorphisms and suppose $d : B \to D$ is their coequalizers. Consider the homomorphism $f - g$, since d coequalizes f and g, $d \circ (f - g) = d \circ f - d \circ g = 0$, or equivalently, $\text{Im}(f - g) \subseteq \ker(d)$. Now, consider diagram (43) as an instance of (42), where q is the quotient map.[231]

[231] It is commutative because $q \circ (f - g) = 0$ by definition of q.



(43)

We claim that ! has an inverse, implying that $D \cong B/\text{Im}(f - g)$.[232] Indeed, for

[232] This is not enough to say that $B/\text{Im}(f - g)$ with the quotient map is the coequalizer, we leave you the task to complete the proof using this isomorphism that crucially satisfies $! \circ d = q$.

$[x] \in B/\mathrm{Im}(f-g)$, we must have

$$!^{-1}([x]) = !^{-1}(q(x)) = !^{-1}(!(d(x))) = d(x),$$

and it is only left to show $!^{-1}$ is well-defined because the inverse of a homomorphism is a homomorphism. This follows because if $[x] = [x']$, then there exists $y \in \mathrm{Im}(f-g)$ such that $x = x' + y$, so

$$!^{-1}(x) = d(x) = d(x'+y) = d(x') + d(y) = d(x') + 0 = !^{-1}(x').$$

In the special case that g is the constant 0 map, $B/\mathrm{Im}(f)$ is called the **cokernel** of f, denoted $\mathrm{coker}(f)$.

SOL **Exercise 235.** Show that an idempotent morphism $f : A \to A \in \mathbf{C}_1$ is split idempotent if and only if the coequalizer of f and id_A exists.

SOL **Exercise 236.** Try to dualize the definition of fibers from Exercise 231. What goes wrong?

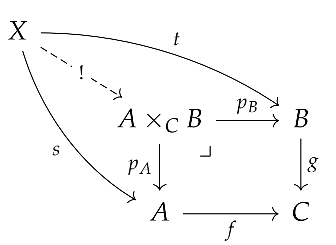
Pullbacks

Definition 237 (Cospan)**.** A **cospan** in $\mathbf{C}$ comprises three objects A, B, C and two morphisms f and g as in (44).[233]

$$A \xrightarrow{f} C \xleftarrow{g} B \tag{44}$$

[233] Just like forks, coforks and spans that we introduce later, cospan is simply a name that we give to a certain shape of diagram that occurs quite often.

Definition 238 (Pullback)**.** Let $A \xrightarrow{f} C \xleftarrow{g} B$ be a cospan in $\mathbf{C}$. Its **pullback** is an object $A \times_C B$ along with morphisms $p_A : A \times_C B \to A$ and $p_B : A \times_C B \to B$ such that $f \circ p_A = g \circ p_B$ and the following universal property holds: for any object X and morphisms $s : X \to A$ and $t : X \to B$ satisfying $f \circ s = g \circ t$, there is a unique morphism $! : X \to A \times_C B$ making (45) commute.[234]



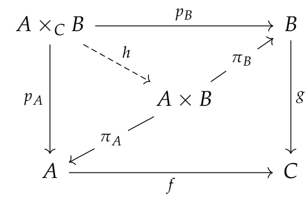
(45)

[234] The ⌟ symbol inside the square is a standard convention to specify that a square is not only commutative, but also a pullback square. Some authors call such a square *cartesian*, but this adjective has too many different meanings in category theory in my opinion, so we will not use it.

We call p_A the pullback of g **along** f and sometimes denote it $f^*(g)$. Symmetrically, p_B is the pullback of f along g, denoted $g^*(f)$.

A drawback of the notation $A \times_C B$ is that it does not refer to the morphisms f and g which are essential in the definition. An alternative notation is $f \times_C g$ (I learned about it here). An argument supporting this notation is in Exercise 333.

Example 239 (**Set**)**.** Let $A \xrightarrow{f} C \xleftarrow{g} B$ be a cospan in **Set** and suppose that its pullback is $A \xleftarrow{p_A} A \times_C B \xrightarrow{p_B} B$. Observe that p_A and p_B look like projections, and in fact, by the universality of the product $A \times B$, there is a map $h : A \times_C B \to A \times B$ such that $h(x) = (p_A(x), p_B(x))$ ((46) commutes). Consider the image of h, if



(46)

$(a,b) \in \mathrm{Im}(h)$, then there exists $x \in A \times_C B$ such that $p_A(x) = a$ and $p_B(x) = b$. Moreover, the commutativity of the square in (46) implies $f(a) = g(b)$, hence

$$\mathrm{Im}(h) \subseteq \{(a,b) \in A \times B \mid f(a) = g(b)\}.$$

Now, let X be the R.H.S., and $s = \pi_A|_X$ and $t = \pi_B|_X$ be the projections to A and B respectively restricted to $X \subseteq A \times B$. Our construction ensures $f \circ s = g \circ t$ hence there is a unique $! : X \to A \times_C B$ satisfying $p_A \circ ! = \pi_A|_X$ and $p_B \circ ! = \pi_B|_X$. Viewing h as going in the opposite direction to !,[235] we derive for any $(a,b) \in X$,[236]

[235] We just saw that the image of h is contained in X, so we can see h as a function $h : A \times_C B \to X$.

[236] We use the fact that $\pi_A \circ h \circ ! = p_A \circ !$ and similarly for B.

$$(h \circ !)(a,b) = (p_A(!(a,b)), p_B(a,b)) = (\pi_A(a,b), \pi_B(a,b)) = (a,b),$$

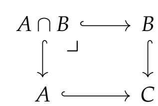
thus ! has a left inverse and is injective. Assume towards a contradiction that it is not surjective, then let $y \in A \times_C B$ not be in the image of ! and denote $x = !(p_A(y), p_B(y))$. Define $!'$ as acting exactly like ! except on $(p_A(y), p_B(y))$ where it goes to y instead of x. This ensure that $!'$ still makes the diagram commute, contradicting the uniqueness of !.

As a particular case, when one function in the cospan is an inclusion, say $g : B \hookrightarrow C$, the pullback is the preimage of B under f since[237]

[237] This can be seen as a generalization of the fibers defined in Exercise 231: seeing an element of C as a function $c : \mathbf{1} \to C$, the fiber $f^{-1}(c)$ is the pullback of c along f.

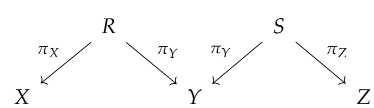
$$\{(a,b) \in A \times B \mid f(a) = g(b) = b\} \cong \{a \mid f(a) \in B\} = f^{-1}(B) \subseteq A.$$

You can also check that p_A is the inclusion $f^{-1}(B) \hookrightarrow A$ and p_B is f restricted to $f^{-1}(B)$. As a particular case of that, if the cospan consists of two inclusions $A \hookrightarrow C \hookleftarrow B$, then its pullback is the intersection $A \cap B$ with p_A and p_B being the inclusions.



Examples 240. In a posetal category, the commutativity of the square in (45) does not depend on the morphisms, thus the universal property is equivalent to the property of being a product.

The composition of relations R and S can be defined using pullbacks in **Set**. Given relations $R \subseteq X \times Y$ and $S \subseteq Y \times Z$, we can restrict the projections to R and S to obtain (47). Then, taking the pullback of the cospan in the middle and using the characterization of the pullback in **Set** from Example 239, we obtain



(47)

$$R \times_Y S = \left\{((x,y),(y',z)) \in R \times S \mid y = y'\right\}.$$

Observe in (48) that we have functions from $R \times_Y S$ to X and Z: $\pi_X \circ p_R$ and $\pi_Z \circ p_S$. Thus, by the universal property of the product $X \times Z$, there is a function $! : R \times_Y S \to X \times Z$. After a bit of computations, recalling that $p_R((x,y),(y',z)) = (x,y)$ and $p_S((x,y),(y',z)) = (y',z)$, we find that the image of ! is precisely the composite relation[238]

[238] Our argument here heavily relies on working with sets and functions, but there is a way to generalize relations in other nice enough categories using this idea.

$$S \circ R = \{(x,z) \mid \exists y, (x,y) \in R, (y,z) \in S\}.$$

$$\begin{array}{ccccccc} & & & R\times_Y S & & & \\ & & \swarrow^{p_R} & \lrcorner & \searrow^{p_S} & & \\ & R & & & & S & \\ \swarrow^{\pi_X} & & \searrow^{\pi_Y} & & \swarrow^{\pi_Y} & & \searrow^{\pi_Z} \\ X & & & Y & & & Z \end{array} \tag{48}$$

SOL Exercise 241. Let $f : X \to Y$ be a morphism in $\mathbf{C}$. Show f is monic if and only if the square in (49) is a pullback.[239]

$$\begin{array}{ccc} X & \xrightarrow{\mathrm{id}_X} & X \\ {\scriptstyle \mathrm{id}_X}\downarrow & \lrcorner & \downarrow{\scriptstyle f} \\ X & \xrightarrow[f]{} & Y \end{array} \tag{49}$$

[239] This result and its dual will sometimes be used to treat monomorphisms (resp. epimorphisms) as limits (resp. colimits). See e.g. Exercise 264 where you will show that monomorphisms are preserved by pullback preserving functors (see Definition 262).

SOL Exercise 242. Supposing (50) commutes, show that if the right square is a pullback and i and j are isomosphisms, then the rectangle is a pullback.[240]

$$\begin{array}{ccccc} X & \overset{i}{\longleftrightarrow} & A\times_C B & \xrightarrow{p_B} & B \\ {\scriptstyle h}\downarrow & & {\scriptstyle p_A}\downarrow & \lrcorner & \downarrow{\scriptstyle g} \\ Y & \underset{j}{\longleftrightarrow} & A & \xrightarrow[f]{} & C \end{array} \tag{50}$$

[240] i.e. X along with h and $p_B \circ i$ is a pullback of the cospan
$$Y \xrightarrow{f\circ j} C \xleftarrow{g} B.$$

Supposing (51) commutes, show that if the left square is a pullback and i and j are isomorphisms, then the rectangle is a pullback.

$$\begin{array}{ccccc} A\times_C B & \xrightarrow{p_B} & B & \overset{i}{\longleftrightarrow} & X \\ {\scriptstyle p_A}\downarrow & \lrcorner & \downarrow{\scriptstyle g} & & \downarrow{\scriptstyle h} \\ A & \xrightarrow[f]{} & C & \underset{j}{\longleftrightarrow} & Y \end{array} \tag{51}$$

When dualizing products and equalizers, the shape of the diagram did not change. Indeed, reversing all morphisms in a discrete diagram gives back a discrete diagram, and reversing two parallel morphisms yields two parallel morphisms. However, the opposite of a cospan is a span.

Pushouts

Definition 243 (Span)**.** A **span** in $\mathbf{C}$ comprises three objects A, B, C and two morphisms f and g as in (52).

$$A \xleftarrow{f} C \xrightarrow{g} B \tag{52}$$

Definition 244 (Pushout)**.** Let $A \xleftarrow{f} C \xrightarrow{g} B$ be a span in $\mathbf{C}$. Its **pushout** is an object, denoted $A +_C B$, along with morphisms $k_A : A \to A +_C B$ and $k_B : B \to A +_C B$ such that $k_A \circ f = k_B \circ g$ and the following universal property holds: for any object X and morphisms $s : A \to X$ and $t : B \to X$ satisfying $s \circ f = t \circ g$, there is a unique morphism $! : A +_C B \to X$ making (53) commute.[241]

[241] The $\ulcorner$ symbol is a standard convention to specify that the square is not only commutative, but also a pushout square.

$$\begin{array}{ccccc} C & \xrightarrow{g} & B & & \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle k_B} & \searrow{\scriptstyle t} & \\ A & \xrightarrow[k_A]{} & A +_C B & & \\ & \searrow{\scriptstyle s} & & \overset{!}{\dashrightarrow} & X \end{array} \tag{53}$$

We call k_A the pushout of g **along** f and sometimes denote it $f_*(g)$. Symmetrically, k_B is the pushout of f along g, denoted $g_*(f)$.

Example 245 (**Set**)**.** Let $A \xleftarrow{f} C \xrightarrow{g} B$ be a span in **Set** and suppose its pushout is $A \xrightarrow{k_A} A +_C B \xleftarrow{k_B} B$. Similarly to above, observe that k_A and k_B are like coprojections, so there is a unique map $! : A + B \to A +_C B$ such that $!(a) = k_A(a)$ and $!(b) = k_B(b)$. Furthermore, for any $c \in C$, $!(f(c)) = !(g(c))$, thus

$$\exists c \in C, f(c) = a \text{ and } g(c) = b \implies !(a) = !(b).$$

This is very similar to what happened for coequalizers and after working everything out, we obtain that $! : A + B \to A +_C B$ is the coequalizer of $\kappa_A \circ f$ and $\kappa_B \circ g$. This is a general fact that does not only apply in **Set** but in every category with binary coproducts and coequalizers.

As a particular case, if $C = A \cap B$ and f and g are simply inclusions, then $A +_C B = A \cup B$ (the *non-disjoint* union).

SOL **Exercise 246.** Show that if (54) is a pushout square, then d is the coequalizer of f and g. State and prove the dual statement.

$$\begin{array}{ccc} A & \xrightarrow{g} & B \\ {\scriptstyle f}\downarrow & \ulcorner & \downarrow{\scriptstyle d} \\ B & \xrightarrow[d]{} & D \end{array} \tag{54}$$

Example 247 (Rewriting)**.** The categorical approach to graph rewriting is full of uses of pushouts. In this example, we will try to give a flavor of a particular method called double-pushout rewriting (DPO) in an almost trivial setting using words instead of graphs. □.

Just as we defined products and coproducts for more than two objects, and equalizers and coequalizers for more than two morphisms (Footnote 220), we could define pullbacks (resp. pushouts) of multiple morphisms with the same target (resp. source). However, it starts to get messy at this point, so we will abstract away from specific examples of (co)limits.[242]

[242] There is a slick way of doing arbitrary pullbacks and pushouts (as opposed to the binary ones) that we explore in Exercise 333.

3.2 Generalization

There exists many other examples of (co)limits but these six examples give quite a good idea of what it is to be a limit or colimit. More precisely, we will see in Theorem 277 and Exercise 284 that any limit can be built out of products and equalizers or pullbacks and a terminal object. Dually, we can build colimits out of coproducts and coequalizers or pushouts and an initial object.

Let us try to informally spell out the general pattern in the definitions of each example.

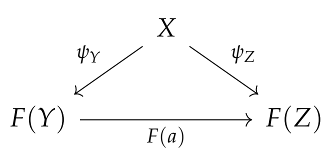
- We start with a shape for a diagram D (e.g. a discrete diagram, two parallel morphisms, a span, a cospan, etc.).
- The limit (resp. colimit) of D is an object L along with morphisms from L to every object in the diagram (resp. in the opposite direction) such that combining D with these morphisms yields a commutative diagram.
- These morphisms satisfy a universal property. For any object L' with morphisms from L' to every object in the diagram (resp. in the opposite direction) that commute with D, there is a unique $! : L' \to L$ (resp. $L \to L'$) such that combining all the morphisms with D yields a commutative diagram.

We have already formalized the first step when we defined diagrams in Definition 182. For the second and third step, notice that the morphisms given for L and L' have the same conditions, they form what we call a cone (resp. cocone).

Definitions

We start by formalizing limits.

Definition 248 (Cone)**.** Let $F : \mathbf{J} \rightsquigarrow \mathbf{C}$ be a diagram. A cone from X to F is an object $X \in \mathbf{C}_0$, called the **tip**, along with a family of morphisms $\{\psi_Y : X \to F(Y)\}$ indexed by objects $Y \in \mathbf{J}_0$ such that for any morphism $a : Y \to Z$ in $\mathbf{J}_1$, $F(a) \circ \psi_Y = \psi_Z$, i.e. diagram (55) commutes.

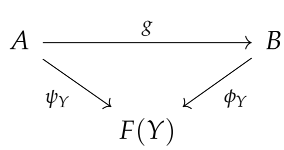


(55)

Often, the terminology cone over F is used.

Next, the fact that the morphism ! keeps everything commutative can be generalized. We say that ! is a morphism of cones.

Definition 249 (Morphism of cones)**.** Let $F : \mathbf{J} \rightsquigarrow \mathbf{C}$ be a diagram and $\{\psi_Y : A \to F(Y)\}_{Y \in \mathbf{J}_0}$ and $\{\phi_Y : B \to F(Y)\}_{Y \in \mathbf{J}_0}$ be two cones over F. A **morphism of cones** from A to B is a morphism $g : A \to B$ in $\mathbf{C}_1$ such that for any $Y \in \mathbf{J}_0$, $\phi_Y \circ g = \psi_Y$, i.e. (56) commutes.



(56)

After verifying that morphisms can be composed, the last two definitions give rise to the category of cones over a diagram F which we denote $\mathrm{Cone}(F)$. Finally, the universal property can be stated in terms of cones, thus giving the general definition of a limit. Indeed, the limit of a diagram F is a cone L over F such that for every cone L' over F, there is a unique cone morphism $! : L' \to L$ called the **mediating morphism**. Equivalently, L is the terminal object of $\mathrm{Cone}(F)$.

Definition 250 (Limit)**.** Let $F : \mathbf{J} \rightsquigarrow \mathbf{C}$ be a diagram, the **limit** of F, if it exists, is the terminal object of $\mathrm{Cone}(F)$. It is denoted $\lim_{\mathbf{J}} F$ or $\lim F$.

Remark 251*.* Often, $\lim F$ also designates the tip of the cone as an object in $\mathbf{C}$ rather than the whole cone.[243] We may also refer to the whole cone as the **limit cone**.

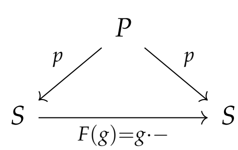
[243] This can sometimes be a source of confusion because many authors omit parts of the proof involving the rest of the cone, and the reader is expected to reconstruct the missing parts.

Examples 252. In the previous section, we gave three examples of limits: products are limits of discrete diagrams, equalizers are limits of diagrams with two parallel morphisms, and pullbacks are limits of cospans. We let you verify the details, and we add to this list three examples in increasing order of complexity.

1. Consider an empty diagram in $\mathbf{C}$, that is, the functor $\varnothing$ from the empty category to $\mathbf{C}$. A cone over $\varnothing$ is an object $X \in \mathbf{C}_0$, the tip, and nothing else as there are no objects in the diagram. Consequently, a morphism in $\mathrm{Cone}(\varnothing)$ is simply a morphism in $\mathbf{C}$ between the tips, so $\mathrm{Cone}(\varnothing)$ is the same as the original category $\mathbf{C}$ and $\lim \varnothing$ is the terminal object of $\mathbf{C}$ if it exists.[244]

[244] Equivalently, we can say that the terminal object is the product of an empty family.

2. Given a group G, recall from Example 126.7 that a G–set can be seen as a diagram in **Set**, i.e. a functor $\mathbf{B}G \rightsquigarrow \mathbf{Set}$. We claim that the limit of this diagram is the set $\mathsf{Fix}(S)$ of fixed points of the action (an element s of a G–set is a **fixed point** if $g \cdot s = s$).[245] Let $F : \mathbf{B}G \rightsquigarrow \mathbf{Set}$ be a G–set with $F(*) = S$, a cone over F is a set P along with a function $p : P \to S$ such that for any $g \in G$, (57) commutes.



(57)

[245] Recall that the limit of two parallel morphisms was called an equalizer. In this example, we are taking the limit of several parallel morphisms. Thus, one can also see the limit of F as the generalized equalizer of all the morphisms $g \cdot -$ with $g \in G$.

We infer from this diagram that the image of p is contained in the set of fixed points.[246] Therefore, p factors uniquely through the inclusion $\mathsf{Fix}(S) \hookrightarrow S$. We conclude that the cone formed by $\mathsf{Fix}(S) \hookrightarrow S$ is the limit cone.

[246] For any $x \in P$, we have $g \cdot p(x) = p(x)$.

3. Let x denote an indeterminate variable and k be a field, $k[x]$ denotes the ring of polynomials over x.[247] We will show that $k[\![x]\!]$, the ring of **formal power series** over x, can be defined as a limit.

   Let $I = \langle x \rangle$ be the ideal generated by x, it contains all the polynomials with no constant terms, and denote $I^n = \langle x^n \rangle$. In the sequel, we view elements of $k[x]/I^n$ as polynomials with degree at most $n-1$.[248] The following three key properties are satisfied (we leave the proofs to the interested readers).

   a) For any $n \leq m \in \mathbb{N}$ and $p \in k[x]/I^m$, forgetting about all terms in p of degree at least n yields a ring homomorphism $\pi_{m,n} : k[x]/I^m \to k[x]/I^n$.[249]

[247] In Chapter 6, we will describe a nice categorical definition of $k[x]$, but, for now, let us assume you know what polynomials are and how they can be added and multiplied together. You can skip this example if you are not familiar with rings.

[248] More accurately, $k[x]/I^n$ contains equivalence classes of polynomials, but their representatives are exactly the polynomials of degree at most $n-1$. Since $I^0 = k[x]$, the quotien $k[x]/I^0$ is the trivial ring, i.e. the zero object in **Ring**.

[249] Note that $\pi_{m,m}$ is the identity.

b) For any $n \in \mathbb{N}$, we can do the same thing for power series to obtain a homomorphism $\pi_{\infty,n} : k[\![x]\!] \to k[x]/I^n$.

c) Any composition of the homomorphisms above can be seen as a single homomorphism above. Namely, $\forall n \leq m \leq l \in \mathbb{N} \cup \infty$,

$$\pi_{m,n} \circ \pi_{l,m} = \pi_{l,n}.$$

Consider the posetal category $(\mathbb{N}, \geq)$, a) and c) imply that $F(n) := k[x]/I^n$ and $F(m \geq n) := \pi_{m,n}$ defines a functor $F : (\mathbb{N}, \geq) \to \mathbf{Ring}$. This is the diagram represented in (58).

$$\cdots \longrightarrow k[x]/I^n \xrightarrow{\pi_{n,n-1}} \cdots \longrightarrow k[x]/I^2 \xrightarrow{\pi_{2,1}} k[x]/I \xrightarrow{\pi_{1,0}} \mathbf{0} \tag{58}$$

Now, using b) and c), we see that $k[\![x]\!]$ along with $\{\pi_{\infty,n}\}_{n\in\mathbb{N}}$ is a cone over the diagram F. It is in fact the terminal cone. Let $\{p_n : R \to k[x]/I^n\}_{n\in\mathbb{N}}$ be another cone over F and $! : R \to k[\![x]\!]$ a morphism of cones. By commutativity, for any $m \leq n$, the coefficients for x^m of $!(r)$ and $p_n(r)$ must agree. Now, by commutativity of the cone $\{p_n\}_{n\in\mathbb{N}}$, $p_n(r)$ and $p_{n-1}(r)$ have the same coefficients except for x^n, thus we can compactly define $!$ by

$$!(r) := p_0(r) + \sum_{n>0} (p_n(r) - p_{n-1}(r)).$$

This completely determines $!$, so it is unique.[250]

[250] Existence follows from the same equation.

The construction of this diagram from quotienting different powers of the same ideal is used in different contexts, it is called the **ring completion** of $k[x]$ with respect to I. For instance, one can define the p–adic integers with base ring $\mathbb{Z}$ and the ideal generated by p for any prime p.

Codefinitions

Put simply, a colimit in $\mathbf{C}$ is a limit in $\mathbf{C}^{\mathrm{op}}$. I suggest you spend a bit of time trying to dualize all of the previous section on your own, but it is done below for completeness.

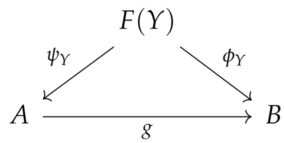
Definition 253 (Cocone)**.** Let $F : \mathbf{J} \rightsquigarrow \mathbf{C}$ be a diagram. A **cocone** from F to X is an object $X \in \mathbf{C}_0$ along with a family of morphisms $\{\psi_Y : F(Y) \to X\}$ indexed by objects of $Y \in \mathbf{J}_0$ such that for any morphism $a : Y \to Z$ in $\mathbf{J}$, $\psi_Z \circ F(a) = \psi_Y$, i.e. (59) commutes.

$$\begin{array}{ccc} F(Y) & \xrightarrow{F(a)} & F(Z) \\ & {\scriptstyle \psi_Y} \searrow \quad \swarrow {\scriptstyle \psi_Z} & \\ & X & \end{array} \tag{59}$$

Often, the terminology cocone under F is used.

Definition 254 (Morphism of cocones)**.** Let $F : \mathbf{J} \rightsquigarrow \mathbf{C}$ be a diagram and $\{\psi_Y : F(Y) \to A\}_{Y\in\mathbf{J}_0}$ and $\{\phi_Y : F(Y) \to B\}_{Y\in\mathbf{J}_0}$ be two cocones. A **morphism of cocones**

from A to B is a morphism $g : A \to B$ in $\mathbf{C}$ such that for any $Y \in \mathbf{J}_0$, $g \circ \psi_Y = \phi_Y$, i.e. (60) commutes.



(60)

The category of cocones under F is denoted $\mathrm{Cocone}(F)$.

Definition 255 (Colimit)**.** Let $F : \mathbf{J} \rightsquigarrow \mathbf{C}$ be a diagram, the **colimit** of F denoted $\mathrm{colim}F$, if it exists, is the initial object of $\mathrm{Cocone}(F)$.

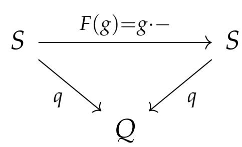
Examples 256. We dualize two examples from the previous section.

1. Dually to Example 252.1, $\mathrm{colim}\emptyset$ is the is the initial object of $\mathbf{C}$ if it exists.[251]

[251] Equivalently, the initial object is the coproduct of an empty family.

2. Dually to Example 252.2, we claim that the colimit of the diagram corresponding to a group action is the set of its orbits. Let $F : \mathbf{B}G \rightsquigarrow \mathbf{Set}$ be a G–set with $F(*) = S$, a cocone from F is a set Q along with a function $q : S \to Q$ such that for any $g \in G$, (61) commutes.

One can also see the colimit of F as the (generalized) coequalizer of all the morphisms $g \cdot -$ with $g \in G$.



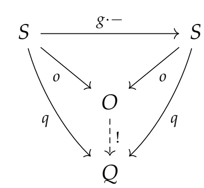
(61)

We infer that if there exists $g \in G$ such that $g \cdot s = s'$, then $q(s) = q(s')$. Denoting $o(s) := \{g \cdot s \mid g \in G\}$ to be the orbit of $s \in S$, the set of orbits of S

$$O := \{o(s) \mid s \in S\}$$

along with the map $o : S \to O$ forms a cocone from F since $o(g \cdot -) = o$.[252] This cocone is the colimit since for any $q : S \to Q$ as in (61), any $! : O \to Q$ making (62) commute is completely determined by $!(o(s)) = q(s)$ (which is well-defined since $o(s) = o(s') \implies \exists g \in G, g \cdot s = g \cdot s' \implies q(s) = q(s')$).

[252] Since the orbits are, by definition, stable under the action of G.



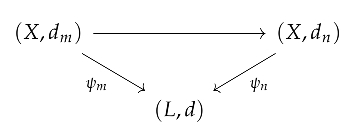
(62)

3. Let $X = \{x, y\}$, and for each nonzero $n \in \mathbb{N}$, let (X, d_n) denote the metric space where x and y have distance $\frac{1}{n}$ (all other distances must be 0). Since morphisms in $\mathbf{Met}$ are nonexpansive functions, for any $m \leq n$, the identity function $(X, d_m) \to (X, d_n)$ is a morphism in $\mathbf{Met}$.[253] We assemble all this data in a diagram of shape $(\mathbb{N}, \leq)$ (the opposite of (58)) depicted in (63).

[253] We have
$$d_m(x, y) = \frac{1}{m} \geq \frac{1}{n} = d_n(x, y).$$

$$(X, d_1) \longrightarrow (X, d_2) \longrightarrow \cdots \longrightarrow (X, d_n) \longrightarrow \cdots \qquad (63)$$

Recall the one point space $(\{*\}, d_1)$ is the terminal object $\mathbf{1}$ in $\mathbf{Met}$ (Example 174). The family $\{!_n : (X, d_n) \to \mathbf{1}\}$ comprising the unique morphisms to $\mathbf{1}$ is a cocone under (63), and we claim it is the colimit cocone.



(64)

Suppose $\psi_n : (X, d_n) \to (L, d)$ is a cocone under (63). Instantiating (59), we find that (64) commutes, hence $\psi_m(x) = \psi_n(x)$ and $\psi_m(y) = \psi_n(y)$ for every $m, n \in \mathbb{N}$. We can give one name ψ to the function $X \to L$ that underlies all ψ_n. For any $n \in \mathbb{N}$, the distance between $\psi(x)$ and $\psi(y)$ is bounded above by $\frac{1}{n}$, otherwise $\psi_n : (X, d_n) \to (L, d)$ would not be nonexpansive. Therefore, the distance can only be 0, and we conclude $\psi(x) = \psi(y)$.

A morphism of cocones f from $\{!_n\}$ to $\{\psi_n\}$ must satisfy $f(!_n(x)) = \psi_n(x) = \psi_n(y)$, so the only possible choice is the function sending $*$ to $\psi(x) = \psi(y)$.

SOL **Exercise 257** (Trivial (co)limits)**.** Show the following (co)limits always exist and find what they are.

1. The limit of a diagram with only one morphism.
2. The colimit of a diagram with only one morphism.
3. The limit of a span.
4. The colimit of a cospan.

Instantiating our examples (co)limits in posets was rather simple because they are thin categories, and every diagram in a thin category is commutative. This generalizes to all (co)limits.

SOL **Exercise 258.** Let $\mathbf{C}$ be a posetal category. Show that the limit (resp. colimit) of any diagram $F : \mathbf{J} \rightsquigarrow \mathbf{C}$ is the infimum (resp. supremum) of all points in the image of F.

Results

Proposition 259 (Uniqueness)**.** *Let $F : \mathbf{J} \rightsquigarrow \mathbf{C}$ be a diagram, the limit (resp. colimit) of F, if it exists, is unique up to unique isomorphism.*

Proof. This follows from the uniqueness of terminal (resp. initial) objects.[254] □

[254] Corollary 171 (resp. Proposition 170).

Remark 260. The isomorphism between two limits (also colimits) is unique when viewed as a morphism of cone. There might exists an isomorphism between the tips that is not a morphism of cone. For instance, let A, B and C be finite sets. One can check that both $A \times (B \times C)$ and $(A \times B) \times C$ are products of $\{A, B, C\}$ (with the usual projection maps). Thus, there is an isomorphism between them. One can check that, for it to be a morphism of cones, it must send $(a, (b, c))$ to $((a, b), c)$, but any other bijection between them is an isomorphism in **Set**.

For this reason, the limit really consists of the whole cone, and not just of the object at the tip. Unfortunately, this subtlety is not well cared for in the literature and it can and has led to errors.

Recall the definition of preserve and reflect we gave in Definition 184. With the framework of (co)limits, we can give more formal related definitions.

SOL **Exercise 261** (NOW!). Let $F : \mathbf{C} \rightsquigarrow \mathbf{C}'$ be a functor and $D : \mathbf{J} \rightsquigarrow \mathbf{C}$ be a diagram. The composition $F \circ D$ is a diagram of shape $\mathbf{J}$ in $\mathbf{C}'$. Show that sending a cone $\{\psi_X : A \to DX\}_{X \in \mathbf{J}_0}$ over F to $\{F\psi_X : FA \to FDX\}_{X \in \mathbf{J}_0}$ is a functor $F_D : \mathrm{Cone}(D) \rightsquigarrow \mathrm{Cone}(F \circ D)$. Dually, construct the functor $F^D : \mathrm{Cocone}(D) \rightsquigarrow \mathrm{Cocone}(F \circ D)$.

In words, $F \circ D$ is the diagram D where we applied F to all objects and morphisms. Then, F_D takes a cone over D and applies F to every object and morphism in it to obtain a cone over $F \circ D$.[255] This allows us to define preservation and reflection of (co)limits, as well as creation.

[255] Similarly for F^D.

Definition 262. Let $F : \mathbf{C} \rightsquigarrow \mathbf{C}'$ be a functor and $\mathbf{J}$ be a category.

- We say that F **preserves** limits of shape $\mathbf{J}$ if for any diagram $D : \mathbf{J} \rightsquigarrow \mathbf{C}$, if $\{\psi_X\}_{X \in \mathbf{J}_0}$ is the limit cone over D, then $\{F\psi_X\}_{X \in \mathbf{J}_0}$ is the limit cone over $F \circ D$. In other words, for any D, F_D preserves (in the sense of Definition 184) terminal objects.[256]
- We say that F **reflects** limits of shape $\mathbf{J}$ if for any diagram $D : \mathbf{J} \rightsquigarrow \mathbf{C}$, if $\{\psi_X\}_{X \in \mathbf{J}_0}$ is a cone over D and $\{F\psi_X\}_{X \in \mathbf{J}_0}$ is the limit cone over $F \circ D$, then $\{\psi_X\}_{X \in \mathbf{J}_0}$ is also the limit cone over D. In other words, for any D, F_D reflects (in the sense of Definition 184) terminal objects.
- We say that F **creates** limits of shape $\mathbf{J}$ if for any diagram $D : \mathbf{J} \rightsquigarrow \mathbf{C}$, if $\{\phi_X\}_{X \in \mathbf{J}_0}$ is a limit cone over $F \circ D$, then there exists a unique cone over D $\{\psi_X\}_{X \in \mathbf{J}_0}$ such that $F\psi_X = \phi_X$ and $\{\psi_X\}_{X \in \mathbf{J}_0}$ is a limit cone.

[256] We will often be less rigorous and write something like $\lim(F \circ D) = F(\lim_{\mathbf{J}} D)$. For instance, we will say that F preserves binary products if $FX \times FY = F(X \times Y)$ or $FX \times FY \cong F(X \times Y)$.

We leave to you the dualization of this definition.[257]

[257] Replace cone by cocone and limit by colimit.

These are more technical and rigorous than our previous notions of preservation and reflection of properties, but the intuition should stay the same. In practice, preservation is used way more often,[258] so let us practice a bit.

[258] In this book, we will not use the other two.

Example 263. Recall from Exercise 131 that we have two functors $(-)_0$ and $(-)_1$ from $\mathbf{Cat}$ to $\mathbf{Set}$. It follows from the definition of product categories that both preserve products. Indeed the objects of $\mathbf{C} \times \mathbf{D}$ are pairs of objects in $\mathbf{C}_0 \times \mathbf{D}_0$, and morphisms of $\mathbf{C} \times \mathbf{D}$ are pairs of morphisms in $\mathbf{C}_1 \times \mathbf{D}_1$, so

$$(\mathbf{C} \times \mathbf{D})_0 = \mathbf{C}_0 \times \mathbf{D}_0 \text{ and } (\mathbf{C} \times \mathbf{D})_1 = \mathbf{C}_1 \times \mathbf{D}_1.$$

SOL **Exercise 264.** Show that if F preserves pullbacks (i.e.: F preserves limits of cospans), then F preserves monomorphisms. State and prove the dual statement.

SOL **Exercise 265.** Show that if $F : \mathbf{C} \rightsquigarrow \mathbf{D}$ is an isomoprhism, then F preserves and reflects (co)limits of all shape.

As we already hinted at, oftentimes, forgetful functors preserve limits,[259] we let you prove a very specific instance of this.

[259] Due to results in Chapter 7.

SOL **Exercise 266.** Let $U : \mathbf{Set}_* \rightsquigarrow \mathbf{Set}$ be the forgetful functor from pointed sets to sets. Show that U preserves products, equalizers and pullbacks.

SOL Exercise 267. Fix $A \in \mathbf{C}_0$, show that the functor $\mathrm{Hom}_{\mathbf{C}}(A, -)$ preserves binary products. Namely, if $X, Y \in \mathbf{C}_0$ and $X \times Y$ exists, then

$$\mathrm{Hom}_{\mathbf{C}}(A, X \times Y) \cong \mathrm{Hom}_{\mathbf{C}}(A, X) \times \mathrm{Hom}_{\mathbf{C}}(A, Y).$$

Corollary 268 (Dual). *Fix* $A \in \mathbf{C}_0$, *show the functor* $\mathrm{Hom}_{\mathbf{C}}(-, A)$ *preserves binary coproducts when viewed as a functor* $\mathbf{C} \rightsquigarrow \mathbf{Set}^{\mathrm{op}}$, *i.e.:*

$$\mathrm{Hom}_{\mathbf{C}}(X + Y, A) \cong \mathrm{Hom}_{\mathbf{C}}(X, A) \times \mathrm{Hom}_{\mathbf{C}}(Y, A).$$

These last two results are strenghtened in Theorem 282 and Corollary 283. We are not done proving things about (co)limits, but we move on to the next section where we will do these proofs using diagram chasing.

3.3 Diagram Chasing

We show four results in increasing order of complexity to demonstrate diagram chasing through examples.

Proposition 269. *Let* $\{f_i, g_i : X_i \to Y_i\}_{i \in I}$ *be a familiy of parallel morphisms in* $\mathbf{C}$ *such that for any* $i \in I$, (65) *is an equalizer, then* (66) *is an equalizer.*

$$E_i \xrightarrow{e_i} X_i \underset{g_i}{\overset{f_i}{\rightrightarrows}} Y_i \qquad (65)$$

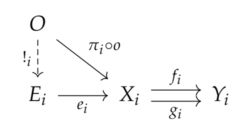
$$\prod_{i\in I} E_i \xrightarrow{\prod_{i\in I} e_i} \prod_{i\in I} X_i \underset{\prod_{i\in I} g_i}{\overset{\prod_{i\in I} f_i}{\rightrightarrows}} \prod_{i\in I} Y_i \qquad (66)$$

Proof. Suppose $o : O \to \prod_{i\in I} X_i$ also equalizes $\prod f_i$ and $\prod g_i$. We have the following implications.[260]

$$\begin{aligned} o \circ \prod f_i = o \circ \prod g_i &\implies \pi_i \circ \prod f_i \circ o = \pi_i \circ \prod g_i \circ o \\ &\implies f_i \circ \pi_i \circ o = g_i \circ \pi_i \circ o \end{aligned}$$

[260] The second implication uses (36).

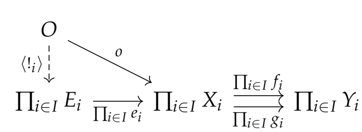
Consequently, for each $i \in I$, $\pi_i \circ o$ equalizes f_i and g_i, so it factors uniquely through e_i: $\pi_i \circ o = e_i \circ !_i$ as depicted in (**??**). The universal property of the product allows us to form the pairing $\langle !_i \rangle_{i\in I} : O \to \prod_{i\in I} E_i$, and we have the following derivation.



(67)

$$\begin{aligned} \pi_i \circ \prod e_i \circ \langle !_i \rangle &= e_i \circ \pi_i \circ \langle !_i \rangle \\ &= e_i \circ !_i \\ &= \pi_i \circ o \end{aligned}$$

We conclude from the universal property of $\prod X_i$ that $o = \prod e_i \circ \langle !_i \rangle$ as depicted in (68). It remains to show $\langle !_i \rangle$ is unique with this property.



(68)

If $m : O \to \prod E_i$ satisfies $\prod e_i \circ f = o$, then

$$e_i \circ \pi_i \circ f = \pi_i \circ \prod e_i \circ f = \pi_i \circ o,$$

but uniqueness of $!_i$ ensures $\pi_i \circ f = !_i$ (they both make (67) commute). This also means $f = \langle !_i \rangle_{i\in I}$, so we are done. □

Corollary 270 (Dual). *Let $\{f_i, g_i : X_i \to Y_i\}_{i \in I}$ be a familiy of parallel morphisms in* **C** *such that for any $i \in I$, $d_i : Y_i \to D_i$ is the coequalizer of f_i and g_i, then $\coprod d_i$ is the coequalizer of $\coprod f_i$ and $\coprod g_i$.*

One might summarize these results by saying that the product of equalizers is the equalizer of products,[261] and this is telling of a general fact about limits interacting with limits (dually colimits interacting with colimits), see Theorem **??** (Corollary **??**).

[261] Dually, the coproduct of coequalizers is the coequalizer of the coproducts.

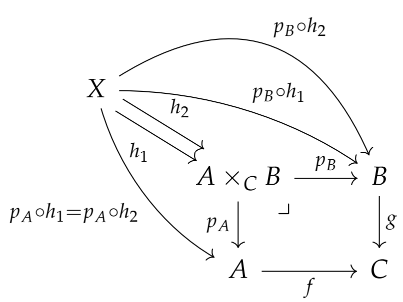
Theorem 271. *Consider the pullback square in* (69)*.*

$$\begin{array}{ccc} A \times_C B & \xrightarrow{p_B} & B \\ {\scriptstyle p_A}\downarrow & \lrcorner & \downarrow{\scriptstyle g} \\ A & \xrightarrow[f]{} & C \end{array} \tag{69}$$

If g is monic, then p_A also is. Symmetrically, if f is monic, then p_B also is.[262]

[262] This is commonly stated simply as: "The pullback of a monomorphism is a monomorphism."

Proof. Let $h_1, h_2 : X \to A \times_C B$ be such that $p_A \circ h_1 = p_A \circ h_2$, we need to show that $h_1 = h_2$. First, observe that h_1 and h_2 yield two cones over the cospan $A \xrightarrow{f} C \xleftarrow{g} B$ as depicted in (70).



(70)

The two cones are

$$\begin{array}{ccc} X & \xrightarrow{p_B \circ h_1} & B \\ {\scriptstyle p_A \circ h_1}\downarrow & & \\ A & & \end{array} \quad \text{and} \quad \begin{array}{ccc} X & \xrightarrow{p_B \circ h_2} & B \\ {\scriptstyle p_A \circ h_2}\downarrow & & \\ A & & \end{array}$$

They make the squares commute because the original pullback square commutes.

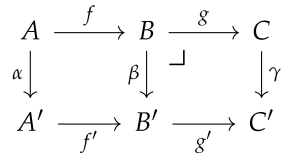
Furthermore, h_1 and h_2 are cone morphisms between X and $A \times_C B$ and since the pullback is the terminal cone over this cospan, they are unique. Now, we already have that the projections onto A is the same for both new cones, but we claim this is also true for the projections onto B. Indeed, because g is monic and the square commutes, we have the following implications.

$$\begin{aligned} p_A \circ h_1 = p_A \circ h_2 &\implies & f \circ p_A \circ h_1 = f \circ p_A \circ h_2 \\ &\implies & g \circ p_B \circ h_1 = g \circ p_B \circ h_2 \\ &\implies & p_B \circ h_1 = p_B \circ h_2 \end{aligned}$$

In other words, the two new cones are in fact the same cones, hence h_1 and h_2 are the same morphisms by uniqueness, which concludes our proof. □

Corollary 272 (Dual). *The pushout of an epimorphism is an epimorphism.*

Theorem 273 (Pasting Lemma)**.** *Consider* (71)*, where the right square is a pullback.*

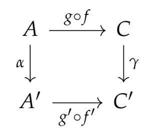


(71)

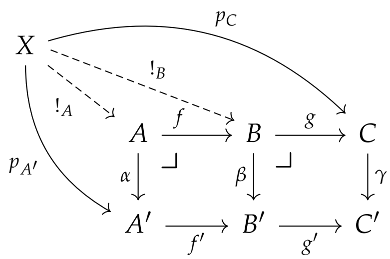
If (71) *commutes, the left square is a pullback if and only if the rectangle is.*[263]

[263] This result is called the **pasting lemma**.

Proof. (⇒) Explicitly, we have to show that $\alpha : A' \leftarrow A \rightarrow C : g \circ f$ is the pullback of $g' \circ f' : A' \rightarrow C' \leftarrow C : \gamma$, i.e., that (72) is a pullback square. The commutativity $g' \circ f' \circ \alpha = \gamma \circ g \circ f$ implies this is already a cone over the cospan we just described. Now, suppose there is another cone over this cospan, namely, there exist morphisms $p_{A'} : X \rightarrow A'$ and $p_C : X \rightarrow C$ satisfying $g' \circ f' \circ p_{A'} = \gamma \circ p_C$ as depicted in (73).



(72)



(73)

Notice that composing $p_{A'}$ with f', we obtain a cone over the cospan in the right square and by universality of B, this yields a unique morphism $!_B : X \rightarrow B$ satisfying $g \circ !_B = p_C$ and $\beta \circ !_B = f' \circ p_{A'}$. This second equality yields cone over the cospan in the left square, thus we get a unique morphism $!_A : X \rightarrow A$ satisfying $\alpha \circ !_A = p_{A'}$ and $f \circ !_A = !_B$. Composing the last equality with g, we get

$$g \circ f \circ !_A = g \circ !_B = p_C,$$

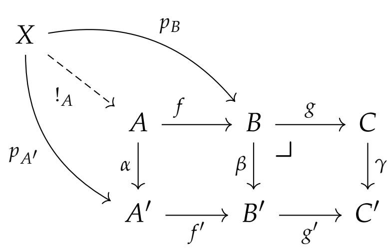
showing that $!_A$ is a morphism of cones over the rectangular cospan.

What is more, any other morphism $m : X \rightarrow A$ of cones over this cospan must satisfy

$$g \circ f \circ m = p_C \text{ and } \beta \circ f \circ m = f' \circ \alpha \circ m = f' \circ p_{A'},$$

and thus, $f \circ m$ is a morphism of cones over the cospan in the right rectangle. By uniqueness, $f \circ m = !_B$, so m is also a morphism of cones over the cospan in the left square, and by universality of A, $m = !_A$.

(⇐) Explicitly, we have to show that $\alpha : A' \leftarrow A \rightarrow B : f$ is the pullback of $f' : A' \rightarrow B \leftarrow B : \beta$.

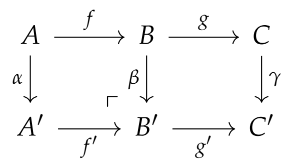


(74)

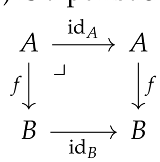
Let $p_{A'} : A' \leftarrow X \to B : p_B$ be a cone over the cospan of the left square (i.e. $\beta \circ p_B = f' \circ p_{A'}$). The commutativity of (71) implies $p_{A'} : A' \leftarrow X \to C : g \circ p_B$ is a cone over the rectangle cospan, then by universality, there exists a unique $!_A : X \to A$ such that $g \circ f \circ !_A = g \circ p_B$ and $\alpha \circ !_A = p_A$. Moreover, with the commutativity of the left square, we find that $f \circ !_A$ is a morphism of cones over the right cospan satisfying $\beta \circ f \circ !_A = f' \circ \alpha \circ !_A = f' \circ p_{A'} = \beta \circ p_B$ and $g \circ f \circ !_A = g \circ p_B$. But since our hypothesis on $p_{A'}$ and p_B implies p_B is a morphism of cones satisfying the same equations, by universality of B, $p_B = f \circ !_A$. Therefore, $!_A$ is a morphism of cone over the left cospan.

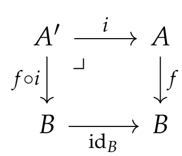
Finally, if $m : X \to A$ also satisfies $\alpha \circ m = p_{A'}$ and $f \circ m = p_B$. We find in particular that m is a morphism of cones over the rectangle cospan, hence by universality, $m = !_A$. □

Corollary 274 (Dual)**.** *If* (75) *commutes, the right square is a pushout if and only if the rectangle is.*


(75)

SOL Exercise 275. Show that (76) is a pullback square. Let $i : A' \to A$ be an isomorphism, show that (77) is a pullback square.[264]


(76)


(77)

[264] We can summarize the first square by saying that the pullback of any morphism along the identity gives back the original morphism. The second square is basically a converse to the statement "pullbacks are unique up to isomorphism" in this very special case.

Definition 276 ((Co)completeness)**.** A category is said to be **(co)complete** (resp. **finitely** (co)complete) if any small (resp. finite) diagram has a (co)limit.

Theorem 277. *Suppose that a category* **C** *has all products and equalizers then* **C** *has all limits, i.e.* **C** *is complete.*

Proof. Let $F : \mathbf{J} \rightsquigarrow \mathbf{C}$ be a diagram, we will show that the limit of F is obtained from the equalizer of two morphisms[265]

$$u_1, u_2 : \prod_{X \in \mathbf{J}_0} F(X) \to \prod_{a \in \mathbf{J}_1} F(t(a)),$$

[265] Recall that s and t denote the sources and targets of morphisms.

which are defined below. The equalizer and the products it involves exist by hypothesis.

First, let us try to explain the intuition behind this construction. The limit of F is the terminal cone over F. In particular, it is a cone over F, namely, a family of morphisms $\psi_X : \lim F \to FX$ indexed by $X \in \mathbf{J}_0$ such that for any $a : X \to Y \in \mathbf{J}_1$, $Fa \circ \psi_X = \psi_Y$. Since **C** has products, we can also specify the morphisms in the cone by a single morphism $\psi : \lim F \to \prod_{X \in \mathbf{J}_0} FX$.[266]

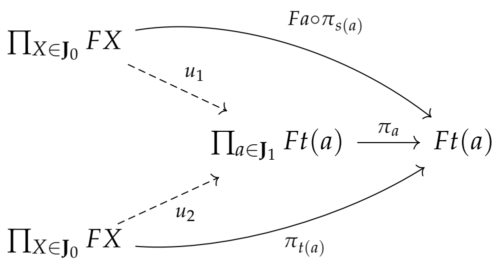
[266] The family $\{\psi_X\}$ gives rise to ψ by the universal property of the product and ψ gives rise to the family by post-composing with the projections $\pi_X : \prod_{X \in \mathbf{J}_0} FX \to FX$.

$$\psi_X = \pi_X \circ \psi$$

The additional property of the cone is now $\forall a : X \to Y \in \mathbf{J}_1, Fa \circ \pi_X \circ \psi = \pi_Y \circ \psi$. Replacing the objects X and Y with $s(a)$ and $t(a)$ respectively, we obtain two families of morphisms

$$\{Fa \circ \pi_{s(a)} : \prod_{X \in \mathbf{J}_0} FX \to Ft(a) \mid a \in \mathbf{J}_1\} \quad \text{and} \quad \{\pi_{t(a)} : \prod_{X \in \mathbf{J}_0} FX \to Ft(a) \mid a \in \mathbf{J}_1\}.$$

The universal property of products yields two parallel morphisms $u_1, u_2 : \prod_{X \in \mathbf{J}_0} FX \to \prod_{a \in \mathbf{J}_1} Ft(a)$ making (78) commute.



(78)

We find that ψ equalizes u_1 and u_2.[267] Since we did not use the fact that ψ is terminal yet, any cone over F yields a morphism from the tip to the product $\prod_{X \in \mathbf{J}_0} FX$ that equalizes u_1 and u_2. Moreover, this process can be reversed, hence any morphism that equalizes u_1 and u_2 corresponds to a cone over F.

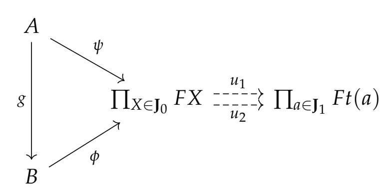
We are on a good track because we have shown that cones over F are in correspondence with cones over the parallel morphisms u_1 and u_2. If we can show there is also a correspondence between the morphisms of such cones, we will be able to conclude that the terminal cone over u_1 and u_2 (i.e. their equalizer) is the terminal cone over F (i.e. the limit of F).[268]

[267] We check that $u_1 \circ \psi = u_2 \circ \psi$ by post-composing with π_a for every $a \in \mathbf{J}_1$. Indeed, we have

$$\begin{aligned} \pi_a \circ u_1 \circ \psi &= Fa \circ \pi_{s(a)} \circ \psi \\ &= \pi_{t(a)} \circ \psi \qquad \text{(def. of } \psi) \\ &= \pi_a \circ u_2 \circ \psi, \end{aligned}$$

and the universal property of $\prod_{a \in \mathbf{J}_1} Ft(a)$ implies $u_1 \circ \psi = u_2 \circ \psi$.

[268] More abstractly, we show there is an isomorphism between the categories $\mathrm{Cone}(F)$ and $\mathrm{Cone}(U)$, where U is the diagram with only two parallel morphisms sent to u_1 and u_2. One can check that isomorphisms of categories preserve terminal objects (Exercise 265), so the equalizer of u_1 and u_2 is the limit of F.



(79)

Let $\{\psi_X, \phi_X : A \to FX\}_{X \in \mathbf{J}_0}$ be two cones over F, $g : A \to B$ be a morphism of cones, and ψ and ϕ be the corresponding morphism that equalize u_1 and u_2. We will show that (79) commutes. By definition of g, we have $\phi_X \circ g = \psi_X$ for any $X \in \mathbf{J}_0$, which we can rewrite as $\pi_X \circ \phi \circ g = \pi_X \circ \psi$. By the universal property of the product $\prod_{X \in \mathbf{J}_0} FX$, we conclude that $\phi \circ g = \psi$.

Conversely, given g that makes (79) commute, g is a morphism of cones over F because for any $X \in \mathbf{J}_0$, $\phi_X \circ g = \pi_X \circ \phi \circ g = \pi_X \circ \psi = \psi_X$.

In conclusion, let $\psi : L \to \prod_{X \in \mathbf{J}_0}$ be the equalizer of u_1 and u_2, the limit of F is the cone $\{\pi_X \circ \psi_X\}_{X \in \mathbf{J}_0}$.

□

Remark 278. The same proof yields a more general statement: For any cardinal κ, if a category $\mathbf{C}$ has all products of size less than κ and equalizers, then it has limits of any diagram with less than κ objects and morphisms.

Corollary 279 (Dual)**.** *If a category* $\mathbf{C}$ *has all coproducts of size less than* κ *and coequalizers, then it has colimits of any diagram with less than* κ *objects and morphisms.*

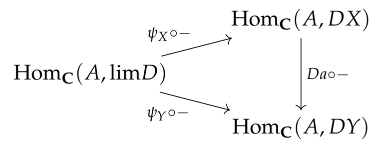
Definition 280. A functor $\mathbf{C} \rightsquigarrow \mathbf{D}$ is said to be (**finitely**) **(co)continuous** if it preserves all (finite) (co)limits.

SOL **Exercise 281.** Show that a functor is continuous if and only if it preserves products and equalizers. State and prove the dual statement.

Theorem 282. *Fix* $A \in \mathbf{C}_0$*, the functor* $\mathrm{Hom}_{\mathbf{C}}(A, -)$ *is continuous.*

Proof. We could show that $\mathrm{Hom}_{\mathbf{C}}(A, -)$ preserves equalizers and use Exercises 267 and 281, but the direct proof is not very long and it lets us get even more familiar with cones.

Let $D : \mathbf{J} \rightsquigarrow \mathbf{C}$ be a diagram and $\{\psi_X : \lim D \to DX\}_{X \in \mathbf{J}_0}$ be the limit cone, we need to show that $\{\psi_X \circ - : \mathrm{Hom}_{\mathbf{C}}(A, \lim D) \to \mathrm{Hom}_{\mathbf{C}}(A, DX)\}_{X \in \mathbf{J}_0}$ is a limit cone.

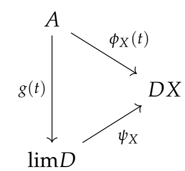


(80)

First, for any $a : X \to Y \in \mathbf{J}_1$, we have $Da \circ \psi_X = \psi_Y$, which implies (80) commutes. Hence, $\{\psi_X \circ -\}_{X \in \mathbf{J}_0}$ is a cone over $\mathrm{Hom}_{\mathbf{C}}(A, D-)$.

Next, if $\{\phi_X : T \to \mathrm{Hom}_{\mathbf{C}}(A, DX)\}_{X \in \mathbf{J}_0}$ is another cone over $\mathrm{Hom}_{\mathbf{C}}(A, D-)$, then observe that any $t \in T$ gives rise to a cone over D $\{\phi_X(t) : A \to DX\}_{X \in \mathbf{J}_0}$. Indeed, we have

$$Df \circ \phi_X(t) = ((Df \circ -) \circ \phi_X)(t) = \phi_Y(t).$$



(81)

We obtain a unique morphism of cones $g(t) : A \to \lim D$ making (81) commute for all $X \in \mathbf{J}_0$. This yields a function $g : T \to \mathrm{Hom}_{\mathbf{C}}(A, \lim D)$ that is a morphism of cones because combining (81) for every $t \in T$ yields $(\psi_X \circ -) \circ g = \phi_X$.

If $g' : T \to \mathrm{Hom}_{\mathbf{C}}(A, \lim D)$ is another morphism of cones, then we must have that $g'(t)$ also makes (81) for all $X \in \mathbf{J}_0$.[269] Therefore, $g'(t) : A \to \lim D$ is a morphism of cones and since $\lim D$ is terminal, we conclude $g'(t) = g(t)$ and $g' = g$. □

[269] We have
$$\psi_X \circ g'(t) = ((\psi \circ -) \circ g')(t) = \phi_X(t).$$

Corollary 283 (Dual). *Fix* $A \in \mathbf{C}_0$*, the functor* $\mathrm{Hom}_{\mathbf{C}}(-, A)$ *is continuous.*[270]

[270] More concisely, the Hom bifunctor is continuous in each argument.

SOL **Exercise 284.** Show that a category with all pullbacks and a terminal object is finitely complete.

Corollary 285 (Dual). *A category with all pushouts and an initial object is finitely cocomplete.*

Remark 286. We can conclude[271] that a functor is finitely continuous if and only if it preserves pullbacks and the terminal object and it is finitely coconituous if and only if it preserves pushouts and the initial object.

[271] Similarly to Exercise 281.

4 Universal Properties

4.1 Examples 81
4.2 Generalization 93
4.3 Comma Categories 94

We continue our exploration of universal constructions. This chapter is arranged like the previous one, we give lots of examples before abstracting away to define universal properties.[272] This abstracting step involves a new concept: comma categories, which are interesting in their own right.

[272] I estimate we have done enough diagram chasing, so we will not prove as much results as we did in Chapter 3.

4.1 Examples

Free Monoid

The construction of a *free* object is common to different fields of mathematics. Informally, when **C** is a category whose objects are objects of another category **D** equipped with extra structure (e.g. **C** is a concrete category and $\mathbf{D} = \mathbf{Set}$), the free **C**–object over a **D**–object X carries the least amount of structure possible to be considered a part of **C** while *containing* X.

The example we will carry out in **Mon** can be carried out in many other categories like **Grp**, **Ab**, **Ring**, etc. We choose **Mon** because the concrete characterization of a free monoid is simple.

Definition 287 (Classical)**.** The **free monoid** on a set A, denoted by A^*, is the set of finite words with symbols in A with the multiplication being concatenation of words and identity being the **empty word** ε.[273]

[273] Examples of finite words in $\{\mathtt{a},\mathtt{b},\mathtt{c}\}^*$ are `a`, `ab`, `abc`, `accabac`, etc. The concatenation of `abc` and `aacb` is `abcaacb`.

An intuitive way to see A^* is that it is the *smallest* monoid that contains A. We start from single-letter words which are just elements of A, and then generate the rest by concatenating bigger and bigger words together (before finally adding ε).

In order to give a categorical characterization, we need to look at homomorphisms from or into the free monoid. Notice that any homomorphism $h^* : A^* \to M$ is completely determined by where h^* sends single-letter words, i.e., elements of A. Indeed, in order to satisfy the homomorphism property, we must have for any $\mathtt{a}, \mathtt{b} \in A$,

$$h^*(\mathtt{ab}) = h^*(\mathtt{a}) \cdot h^*(\mathtt{b}) \text{ and } h^*(\varepsilon) = 1_M.$$

In general, the unique homomorphism sending $\mathtt{a} \in A$ to $h(\mathtt{a})$ can be defined recursively:

$$h^*(w) = \begin{cases} h(\mathtt{a}) \cdot h^*(w') & \mathtt{a} \in A, w \in A^*, w = \mathtt{a}w' \\ 1_M & w = \varepsilon \end{cases}.$$

Concisely, for any function $h : A \to M$, there is a unique homomoprhisms $h^* : A^* \to M$ that sends $\mathtt{a}$ to $h(\mathtt{a})$. We call this fact the universal property of the free monoid.

We repeated several times that universal properties should determine an object up to isomorphism, let us check this. Suppose that a monoid N contains A and satisfies the same property, that is for any (set-theoretic) function $h : A \to M$, there is a unique homomorphism $h^*_N : N \to M$ with $h^*_N(\mathtt{a}) = h(\mathtt{a})$. We claim that N and A^* are isomorphic.

If we take $M = A^*$, and $h : A \to A^* = a \mapsto a$, then we get a homomorphism $h^*_N : N \to A^*$ using the property for N. If we take $M = N$ and the inclusion $i : A \hookrightarrow N$, then the property of A^* yields a homomorphism $i^* : A^* \to N$. By construction, $h^*_N \circ i^* : A^* \to A^*$ and $i^* \circ h^*_N : N \to N$ are both homomorphisms that send $\mathtt{a}$ to $\mathtt{a}$.[274] Note that $\mathrm{id}_{A^*} : A^* \to A^*$ and $\mathrm{id}_N : N \to N$ are also homomorphisms sending $\mathtt{a}$ to $\mathtt{a}$. By the uniqueness in the universal property, we conclude

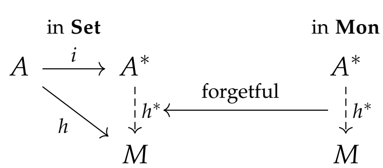
$$h^*_N \circ i^* = \mathrm{id}_{A^*} \text{ and } i^* \circ h^*_N = \mathrm{id}_N,$$

that is, A^* and N are isomorphic.

[274] Recall that both A^* and N contains all elements in A.

The universal property we gave above determined the free monoid up to isomorphism, so we are happy to make this into a definition. However, this definition cannot take place entirely in the category **Mon**. We had to implicitly rely on the fact that a monoid has an underlying set and homomorphisms are just functions satisfying additional properties. Our categorical definition thus relies on the forgetful functor $U : \mathbf{Mon} \rightsquigarrow \mathbf{Set}$.

Definition 288 (Categorical)**.** The free monoid of a set A is an object A^* in **Mon** along with a *canonical inclusion* $i : A \to U(A^*)$ that satisfies the following universal property: for any monoid M and function $h : A \to U(M)$, there exists a unique homomorphism $h^* : A^* \to M$ such that $U(h^*) \circ i = h$, namely, $h^*(i(a)) = h(a)$. This is summarized in (82).[275]



(82)

[275] We omit occurences of U as the underlying set (resp. function) of a monoid (resp. homomorphism) is often denoted with the same symbol as the monoid (resp. homomorphism).

We will see in Chapter 7 that the assignment $A \mapsto A^*$ can be assembled into a functor $-^* : \mathbf{Set} \rightsquigarrow \mathbf{Mon}$. It goes in the opposite direction to the forgetful functor, and in fact can be seen as a weak notion of inverse to U.

Abelianization

Our next example is very similar to the previous one. We add the least amount of structure to a group G to obtain an abelian group G^{ab}.[276]

[276] This assignment assembles into a weak inverse to the intermediate forgetful functor $\mathbf{Ab} \rightsquigarrow \mathbf{Grp}$.

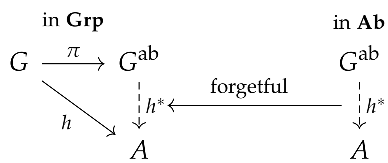
Definition 289 (Classical)**.** Let G be a group, the **abelianization** of G, denoted by G^{ab}, is the quotient of G by the **commutator subgroup** $G' := \{xyx^{-1}y^{-1} \mid x, y \in G\} \subseteq G$, that is $G^{\mathrm{ab}} := G/G'$.

Let us get more insight into this definition. The abelianization is supposed to be the *biggest* abelian quotient of G. To see why, note that if A is an abelian group, any homomorphism $h : G \to A$ must satisfy $h(xyx^{-1}y^{-1}) = 1_A$ for any $x, y \in G$.[277] Hence, G' is contained in the kernel of h. By the fundamental theorem of homomorphism (ref), there is a unique factorization $h = G \xrightarrow{\pi} G/G' \xrightarrow{h'} A$, where π is the canonical quotient map. We summarize this universal property as follows.

[277] The homomorphism property implies

$$\begin{aligned} h(xyx^{-1}y^{-1}) &= h(x)h(y)h(x)^{-1}h(y)^{-1} \\ &= h(x)h(x)^{-1}h(y)h(y)^{-1} \\ &= 1_A. \end{aligned}$$

Definition 290 (Categorical)**.** Let G be a group, the abelianization of G is an abelian group G^{ab} with a map $\pi : G \to G^{\text{ab}}$ satisfying the following universal property: for any homomorphism $h : G \to A$ where A is abelian, there is a unique homomorphism $h^* : G^{\text{ab}} \to A$ such that $h^* \circ \pi = h$. This is summarized in (83).



(83)

We can verify that this characterizes the abelianization of G up to isomorphism.[278]

[278] Compare with what we proved for free monoids.

SOL **Exercise 291.** Let $p : G \to H$ satisfy the universal property of $\pi : G \to G^{\text{ab}}$. Show that $G^{\text{ab}} \cong H$.

Vector Space Basis

This is the third and last example of the same flavor.[279]

[279] We now work with the forgetful functor $\mathbf{Vect}_k \rightsquigarrow \mathbf{Set}$.

Definition 292 (Classical)**.** Let V be a vector space over a field k, a **basis** for V is a subset $S \subseteq V$ that is linearly independent and generates V, namely, any $v \in V$ can be expressed as a linear combination of elements in S and any $s \in S$ cannot be expressed as a linear combination of elements in $S \setminus \{s\}$.

Once again, we would like to get rid of the content of this definition talking about elements, so we focus on what this means for linear maps coming out of V. Let S be a basis of V, W be another vector space over k and $T : V \to W$ be a linear map. By linearity, T is completely determined by where it sends the elements of S. Indeed, for any $v \in V$, write v as a linear combination $\sum_{s \in S} \lambda_s s$ with $\lambda_s \in k$ (only finitely many of the coefficients are non-zero), then $T(v) = \sum_{s \in S} \lambda_s T(s)$. We conclude that any (set-theoretic) function $t : S \to W$ extends to a unique linear map $T : V \to W$.[280]

[280] This is completely analogous to how any homomorphism from the free monoid A^* is determined by where it sends the generators (elements of A).

We claim that this property completely characterizes bases of V. Indeed, let $S \subseteq V$ be such that for any $t : S \to W$, there is a unique linear map $T : V \to W$ extending t. We will show that S is generating and linearly independent.

1. Let U be the subspace generated by S.[281] We claim that the quotient space V/U is $\{0\}$ implying $U = V$, i.e., S is generating. Let $t : S \to V/U$ be the function sending everything to 0, both the quotient map $\pi : V \to V/U$ and the 0 map $0 : V \to V/U$ extend t linearly.[282] By the uniqueness in the universal property, π and 0 must coincide, hence V/U must be trivial.

[281] It contains all linear combinations of elements in S.

[282] The former extends t because every linear combination of elements in S is in U which π sends to 0.

2. Fix $v \in S$, we will show that v is not a linear combination of elements in $S \setminus \{v\}$. First, we claim that v is not zero. If it were, then any function $t : S \to k$ sending v to a non-zero element could not be extended. Next, consider the function[283]

$$t : S \to V + V = \begin{cases} (s, 0) & s \neq v \\ (0, v) & s = v \end{cases}.$$

By the universal property, there exists a linear map $T : V \to V + V$ extending t. Notice that applying T to a linear combination of elements in S, we must obtain a vector of $V + V$ whose second coordinate is 0. However, the second coordinate of $T(v)$ is v, not 0. Hence, v is not a linear combination of elements in S. Our choice of v was arbitrary, so we can conclude that S is linearly independent.

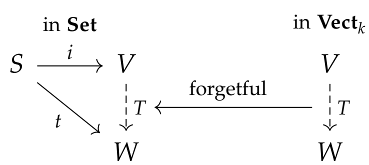
[283] Recall that the coproduct of vector spaces is their direct sum, i.e. $V + V = \{(u, w) \mid u, w \in V\}$ and operations are done coordinate-wise.

We have the following alternative definition of a vector space basis.[284]

[284] We are assuming a different point of view than we did for free monoids, but we are doing the same thing. One could start from a set S and say that V is the free vector space over S if there is the inclusion $i : S \to V$ satisfying (84).

This opposite point of view can be misleading. If we try to prove that this characterizes the basis up to isomorphism (i.e. if S and S' are bases of V, then $S \cong S'$), we will have a harder time than before. Comparing with the proofs for free monoids and abelianizations, we find we can easily prove that if V and W have S as a basis, then $V \cong W$.

Definition 293 (Categorical)**.** Let V be a vector space, a basis of V is a set S along with an inclusion $i : S \to V$ satisfying the following universal property: for any function $t : S \to W$ where W is a vector space, there is a unique linear map $T : V \to W$ such that $T \circ i = t$. This is summarized in (84).



(84)

The previous three examples of universal properties are all categorifications of a free construction. Here are two others we leave you to work out on your own.

SOL **Exercise 294.** What is the free partial order over a set S?

Recall that we can see a category as a directed graph with extra structure using the forgetful functor $U : \mathbf{Cat} \rightsquigarrow \mathbf{DGph}$ that forgets about composition and identities. From any directed graph G, we can construct a category of paths of G, denoted by $\mathbf{P}G$. The objects of $\mathbf{P}G$ are those of G, and the morphisms in $\mathrm{Hom}_{\mathbf{P}G}(A, B)$ are paths from A to B in G. The composition of two paths $A \xrightarrow{f_1} \cdots \xrightarrow{f_n} B$ and $B \xrightarrow{g_1} \cdots \xrightarrow{g_m} C$ is the concatenated path $A \xrightarrow{f_1} \cdots \xrightarrow{f_n} B \xrightarrow{g_1} \cdots \xrightarrow{g_m} C$, and the identity on A is the empty path going from A to A.[285]

[285] Of course, concatenating a path with the empty path does nothing.

SOL **Exercise 295.** Show that $\mathbf{P}G$ is the free category over the directed graph G. Moreover, show that when G has a single object, $\mathbf{P}G$ is the delooping of the free monoid ${G_1}^*$.

Exponential Objects

This section and the following two are motivated by important constructions in **Set** that we want to define categorically. Going further in this direction amounts to doing topos theory, namely, studying categories which look a lot like **Set**.

Remark 296. Let me repeat that there is a choice to make when doing such categorifications. Given a classical construction, we need to decide what is the core idea that we want to keep when we abstract away from concrete details. If this core idea allows you to recover the original construction when instantiating back in **Set**, then your abstraction is appropriate, but it might not be the only one.

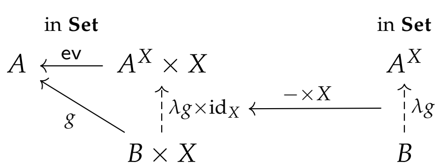
SOL Exercise 297. Let $\mathbf{C}$ be a category and $X \in \mathbf{C}_0$ be such that for any $Y \in \mathbf{C}_0$, $Y \times X$ exists. Show that $- \times X$ is a functor $\mathbf{C} \rightsquigarrow \mathbf{C}$.

Let A and X be sets, A^X commonly denotes the set of functions $X \to A$. In particular, **Set** is locally small and $\mathrm{Hom}_{\mathbf{Set}}(A, B)$ is a set, i.e., an object of **Set**. This is a somewhat exceptional situation, but there are other categories where hom-sets can actually be viewed as objects of the category.[286]

In hope to generalize this construction to other categories, let us study morphisms into A^X.[287] Given a set B and a morphism $f : B \to A^X$, there is a natural operation called **uncurrying** that takes f to $\lambda^{-1}f : B \times X \to A$ which basically evaluates both f and its output at the same time. Namely, $\lambda^{-1}f(b, x) = f(b)(x)$.

As a particular case, we consider the identity function $A^X \to A^X$. Uncurrying yields the **evaluation** function $\mathsf{ev} : A^X \times X \to A$ that evaluates the function in the first coordinate at the second coordinate: $\mathsf{ev}(f, x) = f(x)$.

Now, as the name suggests, uncurrying has an inverse operation called **currying**[288] which takes $g : B \times X \to A$ to $\lambda g : B \to A^X$ defined by $\lambda g(b) = x \mapsto g(b, x)$. Morally, λg delays the evaluation of g on the second input to later.[289] Moreover, notice that the currying of g satisfies $\mathsf{ev}(\lambda g(b), x) = g(b, x) \in A$ for any $b \in B$ and $x \in X$. Intuitively, $\lambda g(b)$ reads the first argument b and waits for the second argument, then $\mathsf{ev}(\lambda g(b), x)$ inputs x, so it is the same thing as doing $g(b, x)$. This along with the fact that currying and uncurrying are bijective operations[290] leads to a universal property that ev satisfies. It is summarized in (85).



(85)

This is entirely categorical, so we can define exponential objects as follows.

Definition 298 (Exponential)**.** Let $\mathbf{C}$ be a category and $X \in \mathbf{C}_0$ be such that $- \times X$ is a functor.[291] For $A \in \mathbf{C}_0$, the **exponential** A^X (if it exists) is an object A^X along with a morphism $\mathsf{ev} : A^X \times X \to A$ such that for all $g : B \times X \to A$, there is a unique $\lambda g : B \to A^X$ making (85) commute.

Informally, one can think of A^X as an object which behaves like $\mathrm{Hom}_{\mathbf{C}}(A, X)$. The terminology **internal hom** is often used (sometimes in more general contexts).

SOL Exercise 299. Let k be a field, and V and W be vector spaces over k. Show that the vector space $\mathrm{Hom}_{\mathbf{Vect}_k}(V, W)$ equipped with pointwise addition and scalar multiplication of linear maps is the exponential W^V.

[286] For instance, the set of linear maps $V \to W$ is a vector space where addition and scalar multiplication is done pointwise.

[287] A priori, there is no reason to prefer morphisms into A^X over morphisms out of A^X, but the intuition is cleaner with the former.

[288] Named in honor of Haskell Curry.

[289] For computer scientists, this is also related to the concept of *continuations*.

[290] Check that $\lambda\lambda^{-1}g = g$ and $\lambda^{-1}\lambda g = g$.

[291] i.e.: all binary products with $X \in \mathbf{C}_0$ exist.

SOL Exercise 300. Show that if $e : Y \times X \to A$ satisfies the same universal property as ev, then $Y \cong A^X$.[292]

Definition 301 (Cartesian closed)**.** When a category $\mathbf{C}$ has a terminal object and all exponentials A^X for all $A, X \in \mathbf{C}_0$ (in particular, it has all binary products[293]), we say it is **cartesian closed**.

The category of sets is cartesian closed. Here is an exercise calling back to when we showed many familiar properties of Cartesian products generalized to binary products.

SOL Exercise 302. Let $\mathbf{C}$ be a category with a terminal object $\mathbf{1}$, and let $X \in \mathbf{C}_0$. Show that X is the exponential $X^{\mathbf{1}}$ and $\mathbf{1}$ is the exponential $\mathbf{1}^X$,[294] i.e. find the evaluation morphisms and prove they satisfy the right universal property.

Subobject Classifier

SOL Exercise 303. Let $\mathbf{C}$ be a well-powered category with all pullbacks. We define $\mathrm{Sub}_{\mathbf{C}}$ on morphisms: it sends $f : X \to Y$ to $f^*(-) : \mathrm{Sub}_{\mathbf{C}}(Y) \to \mathrm{Sub}_{\mathbf{C}}(X)$ sending $m : I \rightarrowtail Y$ to $f^*(m)$, the pullback of m along f as depicted in (86). Show that this is well-defined (recall that a subobject of Y is an equivalence class of monomorphisms) and makes $\mathrm{Sub}_{\mathbf{C}}$ into a functor $\mathbf{C}^{\mathrm{op}} \rightsquigarrow \mathbf{Set}$.

$$\begin{array}{ccc} J & \longrightarrow & I \\ {\scriptstyle f^*(m)}\downarrow & \lrcorner & \downarrow{\scriptstyle m} \\ X & \xrightarrow[f]{} & Y \end{array} \tag{86}$$

In $\mathbf{Set}$, recall that subobjects are subsets. Hence, letting $\Omega = \{\bot, \top\}$ there is a correspondence between $\mathrm{Sub}_{\mathbf{Set}}(X)$ and $\mathrm{Hom}_{\mathbf{Set}}(X, \Omega)$, it sends $I \subseteq X$ to the characteristic function $\chi_I : X \to \Omega$,[295] and in the other direction $f : X \to \Omega$ is sent to $f^{-1}(\top) \subseteq X$. In particular, we have that $\chi_I^{-1}(\top) = I$, which we can write categorically as the following pullback.[296]

$$\begin{array}{ccc} I & \longrightarrow & \mathbf{1} \\ \big\downarrow & \lrcorner & \big\downarrow{\scriptstyle \top} \\ X & \xrightarrow[\chi_I]{} & \Omega \end{array} \tag{87}$$

Crucially, this pullback uniquely determines χ_I.[297] The role played by the two element set $\{\bot, \top\}$ can now be generalized to other categories.

Definition 304 (Subobject classifier)**.** Let $\mathbf{C}$ be a category with a terminal object $\mathbf{1}$. The **subobject classifier** (if it exists) is a morphism $\top : \mathbf{1} \to \Omega \in \mathbf{C}_1$ such that for any monomorphism $I \rightarrowtail X$ there is a unique morphism $\chi_m : X \to \Omega$ such that (87) is a pullback square. We call χ_I the **classifying morphism** of $I \rightarrowtail X$.

Example 305 ($\mathbf{Set}_*$)**.** We find the subobject classifier in $\mathbf{Set}_*$.

Let (X, x) be a pointed set, we first show that a subobject of (X, x) is a subset of X that contains x. An argument like the one in Example 155 shows that monomorphisms in $\mathbf{Set}_*$ are precisely the injective functions that preserve the point.[298] Hence, for a subset $I \subseteq X$ with $x \in I$, the inclusion $i : (I, x) \hookrightarrow (X, x)$ is a monomorphism. Moreover, we can show (as we did in Example 191) that two monomorphisms $(I, i) \rightarrowtail (X, x)$ and $(J, j) \rightarrowtail (X, x)$ are in the same equivalence class of

[292] We will stop proving that universal properties determine objects up to isomorphisms, the abstract result (stating that works for all universal properties) is Corollary **??**.

[293] It also follows that $\mathbf{C}$ has all finite products.

[294] Other properties about exponentials in $\mathbf{Set}$ can be generalized (e.g. $(X^Y)^Z \cong X^{Y \times Z}$), but we will wait until we see the Yoneda lemma to give more elegant proofs.

[295] The characteristic function χ_I is defined by

$$\chi_I(x) = \begin{cases} \top & x \in I \\ \bot & x \notin I \end{cases}.$$

[296] Recall our discussion on preimages in Example 239.

[297] If $f : X \to \Omega$ also makes (87) a pullback square, then $f^{-1}(\top) = I$, so f and χ_I must coincide. The preimage of f on $\top$ determines all of f because there is only one other value in the codomain of f.

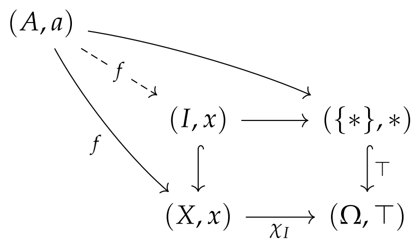
[298] We can also give a more abstract proof. The forgetful functor $\mathbf{Set}_* \rightsquigarrow \mathbf{Set}$ is faithful so it reflects monomorphisms by Exercise 188. Also, we saw in Exercise 266 that it preserves pullbacks, hence it preserves monomorphisms by Exercise 264.

$\mathrm{Sub}_{\mathbf{Set}_*}(X, x)$ if and only if their images coincide (and their image must contain x). We conclude that $\mathrm{Sub}_{\mathbf{Set}_*}(X, x)$ is in correspondence with $\{S \subseteq X \mid x \in S\}$.

The terminal object $\mathbf{1}$ in $\mathbf{Set}_*$ is the singleton $\{*\}$ with distinguished point $*$. Keeping the same notation $\Omega = \{\bot, \top\}$, we claim the subobject classifier is the unique morphism $\top : \mathbf{1} \to (\Omega, \top)$,[299] it sends $*$ to $\top$. For any subset $I \subseteq X$ that contains $x \in X$, we define the classifying morphism $\chi_I : (X, x) \to (\Omega, \top)$ as before (see Footnote 295), noting that it is a morphism in $\mathbf{Set}_*$ because x belongs to I so is mapped to $\top$. It clearly makes the square in (88) commute.[300]

[299] The terminal object $\mathbf{1}$ is also initial in $\mathbf{Set}_*$, see Exercise 175.

[300] Both paths send everything in I to $\top$.

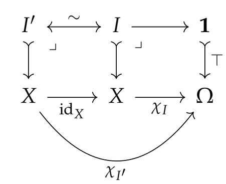


(88)

Now, for any morphism $f : (A, a) \to (X, x)$ making (88) commute, we find the image of f must be contained in I.[301] Therefore, we can factor f through the inclusion of I in X (necessarily uniquely). We conclude that the square in (88) is a pullback.

[301] Otherwise some $a \in A$ is mapped to $\bot$ in the bottom path but not the top path.

It remains to show χ_I is the only possible morphism making that possible. If another morphism χ' does, we apply the forgetful functor which preserves pullbacks (Exercise 266) to get a pullback in $\mathbf{Set}$. Because $\top : \mathbf{1} \to \Omega$ is the subobject classifier in $\mathbf{Set}$, χ' must be the classifying morphism which is the characteristic map χ_I.

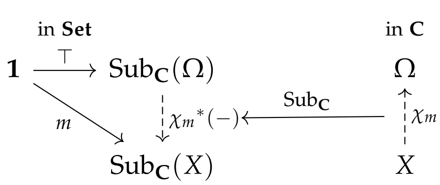


(89)

Before we can draw a diagram (akin to (82), (83), etc.) summarizing the universal property of the subobject classifier, we need to make sure that the classifying morphisms of two monomorphisms in the same equivalence class in $\mathrm{Sub}_{\mathbf{C}}(X)$ are equal. Let $I' \rightarrowtail X$ and $I \rightarrowtail X$ represent the same subobject, namely, there is an isomorphism $I' \cong I$ making the left square in (89) commute. The right square is a pullback by hypothesis and the left square is a pullback by Exercise 275. Therefore, the rectangle is a pullback by the pasting lemma, and we see that $\chi_{I'} = \chi_I \circ \mathrm{id}_X$ by uniqueness of the classifying morphism.

Now, in a well-powered category $\mathbf{C}$ that has a terminal object and all pullbacks,[302] the subobject classifier $\top : \mathbf{1} \to \Omega$ is such that for any subobject m of X, there is a unique morphism $\chi_m : X \to \Omega$ satisfying $\chi_m{}^*(\top) = m$. This is summarized in (90) where we identify $\top$ with the function $\mathbf{1} \to \mathrm{Sub}_{\mathbf{C}}(\Omega)$ picking out this equivalence class of $\top : \mathbf{1} \to \Omega$ in $\mathrm{Sub}_{\mathbf{C}}(\Omega)$ (recall that any morphism out of $\mathbf{1}$, in particular $\top : \mathbf{1} \rightarrowtail \Omega$, is monic by Exercise 179), and similarly for m.

[302] The definition of subobject classifier does not need the well-poweredness and the existence of all pullbacks, but they are necessary to have a universal property because it uses the functor $\mathrm{Sub}_{\mathbf{C}}$. In any case, subobject classifiers are usually used when these conditions are satisfied.



(90)

Notice that the dashed arrow gets reversed because $\mathrm{Sub}_{\mathbf{C}}$ is contravariant. We could also write "in $\mathbf{C}^{\mathrm{op}}$" and not reverse the arrow.

Power Objects

This is the third and last example that can motivate the study of topos theory.

Let X be a set, $\mathcal{P}X$ commonly denotes the set of all subsets of X. In particular, **Set** is well-powered and $\mathrm{Sub}_{\mathbf{Set}}(X)$ is a set, i.e., an object of **Set**. Again, this is an exceptional situation[303] that we would like to make abstract.

[303] This is even more exceptional than being cartesian closed. I do not have any simple examples, but we will see a couple of harder examples.

Let us study morphisms into $\mathcal{P}X$. A function $f : Y \to \mathcal{P}X$ assigns to each $y \in Y$ a (possibly empty) set $f(y)$ of values in X. We can also present the data of f as a subset Γ_f of $X \times Y$ containing the pair (x, y) whenever $x \in f(y)$. This yields a bijection between functions $f : Y \to \mathcal{P}X$ and subsets $\Gamma_f \subseteq X \times Y$[304]: given a subset $\Gamma \subseteq X \times Y$, we define $f_\Gamma : Y \to \mathcal{P}X$ by $f(y) = \{x \in X \mid (x, y) \in \Gamma\}$. The trick to rephrase this categorically is to note that Γ_f is the preimage of the "element of" subset $\in_X \subseteq X \times \mathcal{P}X$ under the function $\mathrm{id}_X \times f : X \times Y \to X \times \mathcal{P}X$.[305] Therefore, we have the following pullback (again, see Example 239).

[304] This generalizes the correspondence between elements of $\mathcal{P}X$ and $\mathrm{Sub}_{\mathbf{Set}}(X)$ because

$$\mathcal{P}X \cong \mathrm{Hom}(\mathbf{1}, \mathcal{P}X) \cong \mathrm{Sub}_{\mathbf{Set}}(X \times \mathbf{1}) \cong \mathrm{Sub}_{\mathbf{Set}}(X).$$

[305] We have that $(\mathrm{id}_X \times f)(x, y) = (x, f(y))$ is in $\in_X$ if and only if $x \in f(y)$ if and only if $(x, y) \in \Gamma_f$. Thus, $\Gamma_f = (\mathrm{id}_X \times f)^{-1}(\in_X)$.

$$\begin{array}{ccc} \Gamma_f & \longrightarrow & \in_X \\ \big\downarrow & \lrcorner & \big\downarrow \\ X \times Y & \xrightarrow[\mathrm{id}_X \times f]{} & X \times \mathcal{P}X \end{array} \tag{91}$$

We are ready to give the abstract definition.

Definition 306 (Power object)**.** Let $\mathbf{C}$ be a category and $X \in \mathbf{C}_0$ be such that $X \times -$ is a functor. The **power object** of X (if it exists) is an object $\mathfrak{P}X \in \mathbf{C}_0$ along with a monomorphism $\in_X \rightarrowtail X \times \mathfrak{P}X$ such that for any monomorphism $\gamma : \Gamma \rightarrowtail X \times Y$, there is a unique morphism $f_\gamma : Y \to \mathfrak{P}X$ making (92) a pullback square.

$$\begin{array}{ccc} \Gamma & \longrightarrow & \in_X \\ {\scriptstyle\gamma}\big\downarrow & \lrcorner & \big\downarrow \\ X \times Y & \xrightarrow[\mathrm{id}_X \times f_\gamma]{} & X \times \mathfrak{P}X \end{array} \tag{92}$$

Note that we obtain f_γ from γ instead of Γ_f from f (like we did in **Set**). In the end, it does not matter because the key property is that there is a correspondence between them. However, in the definition above, the fact that pullbacks are unique up to isomorphisms implies γ is uniquely determined by f_γ up to isomorphism,[306] hence we only need to require f_γ is uniquely determined by γ.

[306] More precisely, the subobject represented by γ is uniquely determined by γ.

Example 307 ($\mathbf{Set}_*$)**.** Recall that a subobject of (X, x) in $\mathbf{Set}_*$ is a subset of X that contains x. This suggests the power object of X may be the set of subsets of X containing x. However we still need to figure out what would be the distinguished point in that set. It turns out there is no point that works out. In fact, we can show that, in general, (X, x) does not have a power object.

We saw above that the power object $\mathfrak{P}(X, x)$ must satisfy

$$\mathrm{Hom}(\mathbf{1}, \mathfrak{P}(X, x)) \cong \mathrm{Sub}_{\mathbf{Set}_*}((X, x) \times \mathbf{1}).$$

Since $\mathbf{1}$ is initial in $\mathbf{Set}_*$, the L.H.S. is a singleton set. We recall that taking a product with the terminal object does nothing (Exercise 211), so the R.H.S. is the set of all subsets of X containing x. Hence, this isomorphism cannot be unless $(X, x) = \mathbf{1}$.[307]

[307] In that case, you can check $\mathbf{1}$ has a (uninteresting) power object.

Again, we want to draw a diagram that summarizes this universal property. Just like for subobject classifiers, we have to check f_γ is the same as $f_{\gamma'}$ when γ and γ' are representatives for the same subobject.

SOL Exercise 308. Let $\in_X \rightarrowtail X \times \mathfrak{P}X$ be the power object of $X \in \mathbf{C}_0$. Show that if γ and γ' are two monomorphisms equal in $\mathrm{Sub}_{\mathbf{C}}(X \times Y)$, then $f_\gamma = f_{\gamma'}$.

We can conclude that if $\mathbf{C}$ is well-powered and has a terminal object, the power object of $X \in \mathbf{C}_0$ is a monomorphism $\in_X \rightarrowtail X \times \mathfrak{P}X$ such that for any subobject γ of $X \times Y$, there is a unique morphism $f_\gamma : Y \to \mathfrak{P}X$ satisfying $(\mathrm{id}_X \times f_\gamma)^*(\in_X) = \gamma$. This is summarized in (93).

$$
\begin{array}{ccccc}
\text{in } \mathbf{Set} & & & & \text{in } \mathbf{C} \\
\mathbf{1} & \xrightarrow{\in_X} & \mathrm{Sub}_{\mathbf{C}}(X \times \mathfrak{P}X) & & \mathfrak{P}X \\
& \searrow^{\iota} & \Big\downarrow{\scriptstyle f_\gamma{}^*(\mathrm{id}_X \times -)} & \xleftarrow{\mathrm{Sub}_{\mathbf{C}}(X\times -)} & \Big\uparrow{\scriptstyle f_\gamma} \\
& & \mathrm{Sub}_{\mathbf{C}}(X \times Y) & & Y
\end{array}
\qquad (93)
$$

In the category **DGph**, any graph has power object.[308] Before proving this, we need to explain what are subobjects and how to take products and pullbacks in **DGph**.

[308] Recall that **DGph** contains only small directed graphs, those with a set of objects and a set of arrows. The morphisms in **DGph** are like functors, but without the requirements about preserving composition and identities (they are not defined in a directed graph).

Adapting the solution to Exercise 194, we find that the subobjects of $G \in \mathbf{DGph}_0$ are graphs H with $H_0 \subseteq G_0$ and $H_1 \subseteq G_1$ such that the source and target maps of H are restrictions of those of G. Similarly to subcategories, we can obtain H from G by deleting arrows and objects, and making sure the sources and targets of remaining arrows also remain.

Again taking inspiration from **Cat**, Definition 132 (see also Exercise 207) tells us how to define binary products of graphs if we forget about the composition and identities.[309]

[309] In other words, the forgetful functor $\mathbf{Cat} \rightsquigarrow \mathbf{DGph}$ preserves binary products.

We have not yet defined pullbacks in **Cat**, but we will do it only for **DGph** here because it is easier.

SOL Exercise 309. Given two morphisms $f : A \to C$ and $g : B \to C$ in **DGph**, find the pullback $A \times_C B$. Show that the functors $(-)_0 : \mathbf{DGph} \rightsquigarrow \mathbf{Set}$ and $(-)_1 : \mathbf{DGph} \rightsquigarrow \mathbf{Set}$[310] preserve pullbacks.

[310] These are defined like for **Cat** in Exercise 131.

The second part of this exercise is a hint for the first part, and it is what we will use shortly. Unrolling, it means the objects and arrows of $A \times_C B$ are defined as follows:

$$
\begin{aligned}
(A \times_C B)_0 &= \{(x, x') \in A_0 \times B_0 \mid f_0(x) = g_0(x')\} \\
(A \times_C B)_1 &= \{(e, e') \in A_1 \times B_1 \mid f_1(e) = g_1(e')\}.
\end{aligned}
$$

Example 310 (DGph). [311] Fix a graph X, we will find $\mathfrak{P}X$.

[311] I am writing this as if we are figuring it out together, but we will use a couple of clever tricks that come from higher-level arguments that we cannot give yet (seeing **DGph** as a functor category).

The universal property of $\mathfrak{P}X$ implies that there is a correspondence between morphisms $\mathbf{1} \to \mathfrak{P}X$ and subobjects of X (the terminal object in **DGph** is the graph with one object and one arrow). For **Cat**, we saw that a functor $\mathbf{1} \rightsquigarrow \mathbf{C}$ is just a choice of object in $\mathbf{C}_0$, but this is not the case in **DGph**. A morphism of graphs does not need to preserve identities, thus a morphism $\mathbf{1} \rightsquigarrow X$ is a choice of object plus a choice of loop on it. This means in $\mathfrak{P}X$, we should have one loop for each subgraph of X. Unfortunately, this does not tell us that much at this point.[312]

[312] We will come back to this later.

To give a complete (and enlightening) description of $\mathfrak{P}X$, we need to know what are its objects, its arrows and the source and target of its arrows. We will make use of a more general consequence of the universal property of $\mathfrak{P}X$: for any graph Y, $\mathrm{Hom}(Y, \mathfrak{P}X) \cong \mathrm{Sub}(X \times Y)$. We can find two graphs O and A such that $\mathrm{Hom}(O, \mathfrak{P}X)$ is in correspondence with the objects of $\mathfrak{P}X$ and $\mathrm{Hom}(A, X)$ with its arrows.

The graph O only contains one object o and no arrow. A morphism $O \to \mathfrak{P}X$ is then just a choice of an object that is the image of o. The product $X \times O$ has the same objects as X but no arrows.[313] Therefore, a subgraph of $X \times O$ is a subset of X_0, and we conclude that we can define $(\mathfrak{P}X)_0 = \mathcal{P}(X_0)$.

[313] By Definition 132, we have

$$(X \times O)_0 = X_0 \times O_0 = X_0 \times \{o\} \cong X, \text{ and}$$
$$(X \times O)_1 = X_1 \times O_1 = X_1 \times \varnothing \cong \varnothing.$$

The graph A contains two objects and one arrow a between them. It looks like the graph of **2**, but without the identity morphisms. A morphism $A \to \mathfrak{P}X$ is a choice of an arrow that is the image of a, and a redundant (determined by the first choice) choice for the image of the source and target of a. The product $X \times A$ can be viewed as two copies of the objects of X (one for each object of A), and for each arrow $f : x \to x'$ in X, there is an arrow from the first copy of x to the second copy of x'.[314] Here is a drawing of a small example.

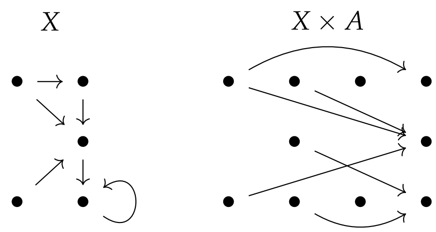
[314] By Definition 132, we have

$$(X \times A)_0 = X_0 \times A_0 = X_0 \times \{1, 2\} \cong X + X,$$

and all morphisms are of the form $(g, a) : (x, 1) \to (x', 2)$ where $g : x \to x'$. Thus,

$$\mathrm{Hom}_{X \times A}((x, 1), (x', 2)) = \mathrm{Hom}_X(x, x'),$$

and all other hom-sets are empty.



A subgraph $H \rightarrowtail X \times A$ can be seen as two subsets H^1 and H^2 of X_0[315] along with a set of arrows $H^a \subseteq X_1$ whose sources are in H^1 and targets are in H^2. We define $(\mathfrak{P}X)_1$ to be the set of all such triples (H^1, H^2, H^a) to ensure we have $\mathrm{Hom}(A, X) \cong (\mathfrak{P}X)_1 \cong \mathrm{Sub}(X \times A)$.

[315] H^1 contains the objects of H belonging to the first copy of X in $X \times A$ and H^2 contains the objects of H in the second copy.

It seems more than likely that H^1 and H^2, being objects of $\mathfrak{P}X$, are the source and target of the arrow (H^1, H^2, H^a). As a sanity check, let us verify that with this definition of source and target in $\mathfrak{P}X$, the loops are in correspondence with subgraphs of X; that is the first thing we discovered about $\mathfrak{P}X$. If $H^1 = H^2$, then the triple defines the subgraph of X containing all the objects in H^1 (or H^2) and all the arrows in H^a. Conversely, given a subgraph of $H \rightarrowtail X$, we let both H^1 and H^2 be the set of objects of H and H^a be the set of arrows of H.

We seem to be on the right track, and we need one last thing in the definition of power object,[316] the subgraph $\in_X$ of $X \times \mathfrak{P}X$. Since we are almost done, we will totally trust our intuition of what $\in_X$ should be without looking for more justifications. The objects of $\in_X$ are pairs (x, H) where $x \in X_0$ and $H \subseteq X_0$, it makes sense to require that $x \in H$. The arrows of $\in_X$ are pairs $(f, (H^1, H^2, H^a))$ where $f : x \to x'$, $x \in H^1$, $x' \in H^2$, it makes sense to require that $f \in H^a$.[317] We are ready to prove $\in_X \rightarrowtail \mathfrak{P}X$ satisfies the universal property of the power object of X.

[316] Before the proof of the universal property.

[317] Recall that every arrow in H^a has its source in H^1 and its target in H^2 just like f.

Let Γ be a subgraph of $X \times Y$ with inclusion $\gamma : \Gamma \to X \times Y$.[318] We need to define a morphism $f_\gamma : Y \to \mathfrak{P}X$ making (92) a pullback square, and we also need

[318] We assume without loss of generality that γ is an inclusion (not an arbitrary monomorphism) to avoid having different names for stuff in Γ and stuff in $X \times Y$.

to prove it is unique. Let us use Exercise 309 to compute the pullback for some yet undefined f_γ, and we will then figure out what constraints we obtain on f_γ when requiring that pullback to be Γ. Hopefully, these will uniquely define f_γ.

Call this pullback G. The objects of G are tuples[319]

$$((x,y),(x',S)) \subseteq (X \times Y)_0 \times (\in_X)_0$$

that satisfy, by commutativity of (92), $x = x'$ and $f_\gamma(y) = S \subseteq X_0$, and by definition of $\in_X$, $x' \in S$. Since the second pair is determined by the first, we can be equivalently write

$$G_0 = \{(x,y) \in X_0 \times Y_0 \mid x \in f_\gamma(y)\}.$$

Thus, to ensure G has the same objects as Γ, it is enough that f_γ satisfies $x \in f_\gamma(y) \Leftrightarrow (x,y) \in \Gamma$ which means $f_\gamma(y) = \{x \in X_0 \mid (x,y) \in \Gamma_0\}$.

The arrows of G are tuples

$$((g,h),(g',(H^1,H^2,H^a))) \subseteq (X \times Y)_1 \times (\in_X)_1$$

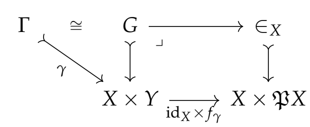
that satisfy, by commutativity of (92), $g = g'$ and $f_\gamma(h) = (H^1, H^2, H^a)$, and by definition of $\in_X$, $s(g) \in H^1$, $t(g) \in H^2$ and $g \in H^a$. Like above, we make things more concise:

$$G_1 = \{(g,h) \in X_1 \times Y_1 \mid s(g) \in f_\gamma(h)^1, t(g) \in f_\gamma(h)^2, g \in f_\gamma(h)^a\}.$$

To ensure G has the same arrows as Γ, we define $f_\gamma(h)$ be the arrow defined by the triple

$$(\{s(g) \mid (g,h) \in \Gamma_1\}, \{t(g) \mid (g,h) \in \Gamma_1\}, \{g \mid (g,h) \in \Gamma_1\}).$$

We leave you two final things to check. First, we only exhibited bijections between the objects and arrows of G and Γ, but in order for (92) to be a pullback, we have to make sure these bijections assemble into an isomorphism making (94) commute. Second, for any other f_γ, the pullback G is another subobject of $X \times Y$ (i.e. there is no isomorphism as in (94)).



(94)

Unlike for exponentials, there is no well-known terminology for a category with all power objects. This is because power objects are usually studied in categories with all finite limits, and when such a category has all power objects, it is called a topos.

Definition 311 (Topos)**.** A finitely complete category where every object has a power object is called an **(elementary) topos**.

Digression on Toposes

The goal of this section is to give an equivalent definition of a topos using exponentials and subobject classifiers. The proofs will be done in exercises, so it is your chance to do some more diagram chasing.

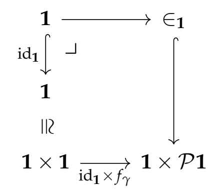
In **Set**,[320] the power object of the terminal set **1** is the set with two elements,

[319] Recall that $\in_X$ is a subgraph of $X \times \mathfrak{P}X$.

[320] Recall it is supposed to be the archetypal topos.

$\varnothing \subseteq \mathbf{1}$ and $\mathbf{1} \subseteq \mathbf{1}$. Now, $\gamma = \mathrm{id}_{\mathbf{1}} : \mathbf{1} \to \mathbf{1}$ is a monomorphism, so we can see it as a subobject in $\mathrm{Sub}_{\mathbf{Set}}(\mathbf{1})$ or $\mathrm{Sub}_{\mathbf{Set}}(\mathbf{1} \times \mathbf{1})$ via the isomorphism $\mathbf{1} \cong \mathbf{1} \times \mathbf{1}$. Using the universal property of $\mathcal{P}\mathbf{1}$, we find that $f_\gamma : \mathbf{1} \to \mathcal{P}\mathbf{1}$ sends the single element in $\mathbf{1}$ to $\mathbf{1} \in \mathcal{P}\mathbf{1}$.[321]

[321] More rigorously, it is the universal property of $\in_{\mathbf{1}} \subseteq \mathbf{1} \times \mathcal{P}\mathbf{1}$ which contains the only element in the R.H.S. You can check that this the only f_γ making (95) a pullback.



(95)

Notice that f_γ is (up to isomorphism) the same function as the subobject classifier $\top : \mathbf{1} \to \{\bot, \top\}$. In fact, in every topos, you can find the subobject classifier this way.

Example 312 (**DGph**).

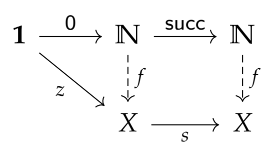
Natural Numbers Object

We end this section with a simpler example still related to toposes to some extent. Without going into the details, topos theory is a framework to study mathematical logic and set theory with a categorical point of view.[322] One of the fundamental building blocks of logic and set theory is the set of natural numbers $\mathbb{N} = \{0, 1, 2, \dots\}$ and the principle of induction tied to it. Let us restate the latter categorically.

[322] Ok, just a bit of informal details...
Grothendieck first defined a more constrained version of topos to help his research in algebraic geometry.
Lawvere and Tierney enlarged the notion of topos to the definition we gave, initiating a deep dive into the strong link between logic and toposes.
Later, Caramello launched a research programme on "toposes as bridges" that uses toposes to formally translate results and concepts between mathematical theories.

The set $\mathbb{N}$ comes with a distinguished element 0 that starts off inductive arguments. It corresponds to the function $0 : \mathbf{1} \to \mathbb{N}$ that picks out 0. For the inductive step, we rely on the function $\mathsf{succ} : \mathbb{N} \to \mathbb{N}$ that takes n to $n+1$.[323] The universal property of $\mathbb{N}$ is that for any pair of functions $z : \mathbf{1} \to X$, $s : X \to X$, there exists a unique $f : \mathbb{N} \to X$ making (96) commute.

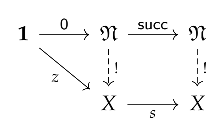
[323] The name succ refers to $n+1$ being the *successor* of n in $\mathbb{N}$.



(96)

The function f is defined inductively. We let $f(0)$ be the element of X in the image of z so that the triangle commutes, then we let $f(n+1) = s(f(n))$ to ensure the square commutes. This means for any $n \in \mathbb{N}$, $f(n) = s^n(z)$ where s^n denotes the composition $s \circ \overset{n}{\cdots} \circ s$ with $s^0 = \mathrm{id}_X$. We abstract away from **Set**.

Definition 313 (NNO)**.** In a category $\mathbf{C}$ with a terminal object $\mathbf{1}$, the **natural numbers object** or NNO (if it exists) is an object $\mathfrak{N} \in \mathbf{C}_0$ along with two morphisms $0 : \mathbf{1} \to \mathfrak{N}$ and $\mathsf{succ} : \mathfrak{N} \to \mathfrak{N}$ satisfying the following universal property: for any pair of morphisms $z : \mathbf{1} \to X$ and $s : X \to X$, there exists a unique morphism $! : \mathfrak{N} \to X$ making (97) commute.



(97)

SOL **Exercise 314.** Show that the NNO in **Poset** is $(\mathbb{N}, \leq)$ with the same zero and successor functions (now seen as morphisms in **Poset**).

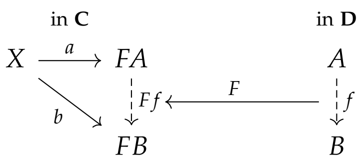
It is not evident how we could summarize the universal property of an NNO using a diagram exactly like the others. Still, the definition really feels like a universal property, so we should not forget this when generalizing what we have seen in all examples above.

4.2 Generalization

Diagrams (82), (83), (84), (85), (90) and (93) look so similar that you can try to infer the following definition unifying all these concepts under one roof.[324]

[324] Although, (85) looks like all arrows have been reversed, so, you guessed it, it will be an instance of the dual notion.

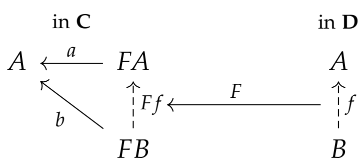
Definition 315 (Universal morphism)**.** If $F : \mathbf{D} \rightsquigarrow \mathbf{C}$ is a functor and $X \in \mathbf{C}_0$. A **universal morphism** from X to F is a morphism $a : X \to F(A)$ such that for any other morphism $b : X \to F(B)$, there is a unique morphism $f : A \to B$ in $\mathbf{D}$ such that $F(f) \circ a = b$, which is summarized in (98).



(98)

The dual notion is a universal morphism from F to X.[325] It is a morphism $a : F(A) \to X$ such that for any other morphism $F(B) \to X$, there is a unique morphism $f : B \to A$ in $\mathbf{D}$ satisfying $a \circ F(f) = b$. This is summarized below in (99).

[325] The duality is clear from how (99) is just (98) with all morphisms reversed. More abstractly, we can say that a universal morphism from F to X is a universal morphism from $X \in \mathbf{C}^{\mathrm{op}}$ to $F^{\mathrm{op}} : \mathbf{D}^{\mathrm{op}} \rightsquigarrow \mathbf{C}^{\mathrm{op}}$.



(99)

Examples 316. In practice and in the literrature, we often say that some construction satisfies a universal property without referring to the actual universal morphism. For example, we say that the free monoid satisfies a universal property, while the less ambiguous thing to say is that the inclusion of a set A into the free monoid A^* is a universal morphism from the set A to the fogetful functor $U : \mathbf{Mon} \rightsquigarrow \mathbf{Set}$.[326] Let us translate the other examples we gave above with this new terminology.

[326] You probably agree that the latter is a mouthful, but the former can feel very vague, especially when you are not familiar with the construction or universal properties in general.

1. The quotient map from a group G to its abelianization G^{ab} is the universal morphism from G to the forgetful functor $\mathbf{Ab} \rightsquigarrow \mathbf{Grp}$.
2. The set $S \subseteq V$ is a basis for the vector space V when the inclusion $S \hookrightarrow V$ is the universal morphism from S to the forgetful functor $\mathbf{Vect}_k \rightsquigarrow \mathbf{Set}$.
3. An exponential object is an object A^X along with the universal morphism ev from the functor $- \times X$ to A.[327]
4. A subobject classifier is a morphism $\top : \mathbf{1} \rightarrowtail \Omega$ such that the corresponding function $\top : \mathbf{1} \to \mathrm{Sub}_{\mathbf{C}}(\Omega)$ is the universal morphism from $\mathbf{1}$ to the functor $\mathrm{Sub}_{\mathbf{C}}$.
5. A power object of X is an object $\mathfrak{P}X$ along with the universal morphism $\in_X$ from $\mathbf{1}$ to $\mathrm{Sub}_{\mathbf{C}}(X \times -)$.

[327] This is an example of a universal morphism from a functor to an object, whereas all the other examples are universal morphisms from an object to a functor.

Another common practice is to use the word free in situations where we have a universal morphism to a forgetful functor (just like the free monoid). For instance, one could say that G^{ab} is the free abelian group over G, or that V is the free vector space over its basis. When you have two categories with an obvious forgetful functor between them, it can be useful to figure out if you can construct free objects. We will get back to this in Chapter 7.

A first approximation of the definition of universality is to say that a universal property is the property of being a universal morphism from X to F or from F to X. Unfortunately, this is too constrained. For instance, as we have said, the universal property of NNOs does not correspond to a universal morphism like that. Another example is subobject classifiers in categories that are not well-powered. In such categories, Sub is not a functor into **Set**,[328] so we cannot have a universal morphism from X to Sub.

[328] There might be another suitable codomain for that functor, but let us not think too hard about size issues.

In the next section, we will see that universal morphisms are initial or terminal objects in a comma category. It turns out that in the most general terms, being universal is best defined as being initial or terminal is some category. It may seem vague at first, but this perfectly describes all the universal properties we have used so far that fit the template "for all ... there exists a unique morphism ..."

Definition 317 (Universal property)**.** A **universal property** is the property of being initial or terminal in a category.[329]

[329] This rather underwhelming definition is also what led me to postpone it to this point, after we have seen many examples and uses of universal properties.

It readily follows (using Proposition 170 and Corollary 171) that universal properties determine things up to isomorphism.

SOL **Exercise 318.** Show that in any category $\mathbf{C}$ with a terminal object $\mathbf{1}$ (even if $\mathbf{C}$ is not well-powered), we can define a category whose objects are monomorphisms in $\mathbf{C}$ and $\top : X \rightarrowtail \Omega$ is terminal if and only if it is the subobject classifier in $\mathbf{C}$. In particular, if $\top$ is terminal in that category, then X is terminal in $\mathbf{C}$.

4.3 Comma Categories

Before moving on, we are going to have some fun with new definitions that let us construct new categories out of categories and functors. This section could have appeared in earlier chapters, but those were already dense, and this section ends with a more concise definition of universal morphisms as initial or terminal objects in comma categories.

Definition 319 (Comma category)**.** Given two functors $\mathbf{D} \overset{F}{\rightsquigarrow} \mathbf{C} \overset{G}{\leftsquigarrow} \mathbf{E}$, there is a category $F \downarrow G$,[330] called the **comma category**, whose objects are triples (X, Y, α) with $X \in \mathbf{D}_0$, $Y \in \mathbf{E}_0$ and $\alpha : F(X) \to G(Y)$ (in $\mathbf{C}_1$), and morphisms between (X_1, Y_1, α) and (X_2, Y_2, β) are pairs of morphisms $f : X_1 \to X_2$ in $\mathbf{D}_1$ and $g : Y_1 \to Y_2$

[330] Some authors denote this category F/G.

in $\mathbf{E}_1$ yielding a commutative square as in (100).

$$\begin{array}{ccc} F(X_1) & \xrightarrow{F(f)} & F(X_2) \\ {\scriptstyle\alpha}\downarrow & & \downarrow{\scriptstyle\beta} \\ G(Y_1) & \xrightarrow[G(g)]{} & G(Y_2) \end{array} \tag{100}$$

The identity morphism on (X, Y, α) is the pair $(\mathrm{id}_X, \mathrm{id}_Y)$ making (101) commute. The composition of (f, g) and (f', g') is $(f' \circ f, g' \circ f)$, it makes the following commute by paving with the commutative squares induced by (f, g) and (f', g').

$$\begin{array}{ccc} FX & \xrightarrow{F\mathrm{id}_X = \mathrm{id}_{FX}} & FX \\ {\scriptstyle\alpha}\downarrow & & \downarrow{\scriptstyle\alpha} \\ GY & \xrightarrow[G\mathrm{id}_Y = \mathrm{id}_{GY}]{} & GY \end{array} \tag{101}$$

$$\begin{array}{ccccc} & & \overset{F(f' \circ f)}{\frown} & & \\ FX_1 & \xrightarrow[Ff]{} & FX_2 & \xrightarrow[Ff']{} & FX_3 \\ {\scriptstyle\alpha}\downarrow & & & & \downarrow{\scriptstyle\gamma} \\ GY_1 & \xrightarrow{Gg} & GY_2 & \xrightarrow{Gg'} & GY_3 \\ & & \underset{G(g' \circ g)}{\smile} & & \end{array} \tag{102}$$

SOL Exercise 320. Given two functors $\mathbf{D} \xrightarrow{F} \mathbf{C} \xleftarrow{G} \mathbf{E}$ and their comma category $F \downarrow G$, show there are two forgetful functors $U_F : F \downarrow G \rightsquigarrow \mathbf{D}$ and $U_G : F \downarrow G \rightsquigarrow \mathbf{E}$ that send (X, Y, α) to X and to Y respectively.

Example 321 (NNO)**.** Let $\mathbf{C}$ be a category with a terminal object and a NNO, and let $\mathbf{1} + - : \mathbf{C} \rightsquigarrow \mathbf{C}$ be the maybe functor. The natural numbers object is the initial object in $(\mathbf{1} + -) \downarrow \mathrm{id}_{\mathbf{C}}$. The morphisms $0 : \mathbf{1} \to \mathfrak{N}$ and $\mathsf{succ} : \mathfrak{N} \to \mathfrak{N}$ can be copaired in $[0, \mathsf{succ}] : \mathbf{1} + \mathfrak{N} \to \mathfrak{N}$ that is an object of this comma category. An arbitrary object of $(\mathbf{1} + -) \downarrow \mathrm{id}_{\mathbf{C}}$ is a morphism $f : \mathbf{1} + X \to X$ which we can decompose as $[f \circ \kappa_{\mathbf{1}}, f \circ \kappa_X]$. Writing $z = f \circ \kappa_{\mathbf{1}}$ and $s = f \circ \kappa_X$, by the universal property of the NNO, there is a unique morphism making (97) commute. Equivalently, (103) commutes,[331] which means ! is the unique morphism from $[0, \mathsf{succ}]$ to f in the comma category $(\mathbf{1} + -) \downarrow \mathrm{id}_{\mathbf{C}}$.

$$\begin{array}{ccc} \mathbf{1} + \mathfrak{N} & \overset{\mathrm{id}_{\mathbf{1}} + !}{\dashrightarrow} & \mathbf{1} + X \\ {\scriptstyle[0,\mathsf{succ}]}\downarrow & & \downarrow{\scriptstyle f = [z,s]} \\ \mathfrak{N} & \underset{!}{\dashrightarrow} & X \end{array} \tag{103}$$

[331] If (97) commutes, we have $z = ! \circ 0$ and $s \circ ! = ! \circ \mathsf{succ}$. Thus, we have

$$\begin{aligned} [z, s] \circ (\mathrm{id}_{\mathbf{1}} + !) &= [z, s \circ !] \\ &= [! \circ 0, ! \circ \mathsf{succ}] \\ &= ! \circ [0, \mathsf{succ}]. \end{aligned}$$

Conversely, if (103) commutes, the same derivation shows $[z, s \circ !] = [! \circ 0, ! \circ \mathsf{succ}]$. By Corollary 268, we must have $z = ! \circ 0$ and $s \circ ! = ! \circ \mathsf{succ}$.

Definition 322 (Arrow category)**.** In the setting of Definition 319, if $F = G = \mathrm{id}_{\mathbf{C}}$, then $\mathrm{id}_{\mathbf{C}} \downarrow \mathrm{id}_{\mathbf{C}}$ is called the **arrow category** of $\mathbf{C}$ and denoted $\mathbf{C}^{\to}$. Its objects are morphisms in $\mathbf{C}$ and its morphisms are commutative squares in $\mathbf{C}$.[332] It may remind you of the category defined in Exercise 318.

[332] Less concisely, a morphism $\phi : f \to g$ between morphisms $f : X \to Y$ and $g : X' \to Y'$ is a pair of morphisms $\phi_X : X \to X'$ and $\phi_Y : Y \to Y'$ making (??) commute.

$$\begin{array}{ccc} X & \xrightarrow{f} & Y \\ {\scriptstyle\phi_X}\downarrow & & \downarrow{\scriptstyle\phi_Y} \\ X' & \xrightarrow[g]{} & Y' \end{array} \tag{104}$$

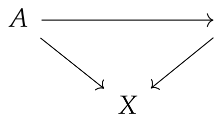
SOL Exercise 323. Let $\mathbf{C}$ be a category (note the change of font to distinguish the functors from their action).

1. Show that $\mathsf{id} : \mathbf{C} \rightsquigarrow \mathbf{C}^{\to}$ sending $X \in \mathbf{C}_0$ to id_X is functorial.

2. Show that $\mathrm{s} : \mathbf{C}^{\to} \rightsquigarrow \mathbf{C}$ sending $f \in \mathbf{C}_0^{\to}$ to $s(f)$ is functorial.
3. Show that $\mathrm{t} : \mathbf{C}^{\to} \rightsquigarrow \mathbf{C}$ sending $f \in \mathbf{C}_0^{\to}$ to $t(f)$ is functorial.

SOL **Exercise 324.** Show the assignment $\mathbf{C} \mapsto \mathbf{C}^{\to}$ yields a functor $\mathbf{Cat} \rightsquigarrow \mathbf{Cat}$.

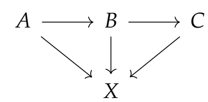
Definition 325 (Slice category)**.** In the setting of Definition 319, if $F = \mathrm{id}_{\mathbf{C}}$ and $G = \Delta(X) : \mathbf{1} \rightsquigarrow \mathbf{C}$ is a constant functor selecting one object $G(\bullet) = X \in \mathbf{C}_0$, then $\mathrm{id}_{\mathbf{C}} \downarrow \Delta(X)$ is called the **slice category** over X and denoted $\mathbf{C}/X$.[333] Its objects are morphisms in $\mathbf{C}$ with target X and its morphisms are commutative triangles with X as a tip as in (105).



(105)

[333] Some authors call this category $\mathbf{C}$ over X.

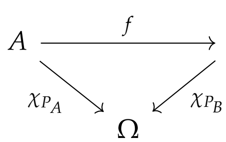
Identity morphisms are commutative triangles with the top morphism being identity and composition is done by combining triangles as in (106).



(106)

SOL **Exercise 326** (NOW!)**.** Suppose $\mathbf{C}$ has a terminal object $\mathbf{1}$, what is $\mathbf{C}/\mathbf{1}$?

Example 327. Recall that $\Omega = \{\bot, \top\}$ is the subobject classifier in $\mathbf{Set}$, that is, a function $A \to \Omega$ can be identified with the subset $f^{-1}(\top) \subseteq A$. Therefore, objects of $\mathbf{Set}/\Omega$ can be seen as sets A equipped with a distinguished subset $P \subseteq A$ that we will call a predicate.[334] Suppose (A, P_A) and (B, P_B) are sets equipped with predicates, what is a morphism $(A, P_A) \to (B, P_B)$ when we see these as objects in $\mathbf{Set}/\Omega$? It is a function $f : A \to B$ making (107) commute.[335]



(107)

[334] This terminology comes from the field of logic. You can think of predicates as things that might be satisfied or not by elements of a set. We say that $a \in A$ satisfies P if $a \in P$.

[335] Recall that $\chi_{P_A}(a) = \top \Leftrightarrow a \in P_A$ and similarly for P_B.

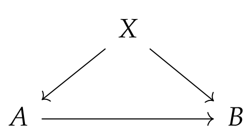
Equivalently, f must satisfy $a \in P_B \implies f(a) \in P_B$. Logically-minded people might call $\mathbf{Set}/\Omega$ the category of predicates and predicate-preserving functions. We can also view a predicate as a unary relation on A, and we recognize $\mathbf{Set}/\Omega$ is the category $1\mathbf{Rel}$.

SOL **Exercise 328.** Let $\mathbf{C}$ be a category with all finite products and fix $n \in \mathbb{N}$. Show the assignment $X \mapsto X^n = X \times \overset{n}{\cdots} \times X$ is functorial. Using this functor and intuition from the previous example, define $n\mathbf{Rel}$ as a comma category.

Definition 329 (Coslice category)**.** In the setting of Definition 319, if $G = \mathrm{id}_{\mathbf{C}}$ and $F = \Delta(X) : \mathbf{1} \rightsquigarrow \mathbf{C}$ is a constant functor selecting one object $F(\bullet) = X \in \mathbf{C}_0$, then $\Delta(X) \downarrow \mathrm{id}_{\mathbf{C}}$ is called the **coslice category** under X and denoted $X/\mathbf{C}$.[336] Its objects are morphisms in $\mathbf{C}$ with source X and its morphisms are commutative triangles with X as a tip as in (108).[337]

[336] Some authors call this category $\mathbf{C}$ under X.

[337] We leave you to dualize the definition of identities and composition from the definition of slice categories.



(108)

Example 330. In the solution to Exercise 197, we saw that a function $\mathbf{1} \to X$ in **Set** can be identified with the element of X it picks out. Therefore, objects of $\mathbf{1}/\mathbf{Set}$ can be seen as sets A equipped with a distinguished element $a \in A$. We already have a name for these things, they are pointed sets. Suppose (A, a) and (B, b) are pointed sets, what is a morphism $(A, a) \to (B, b)$ when we see these as objects of $\mathbf{1}/\mathbf{Set}$? It is a function $f : A \to B$ making (109) commute.

$$\begin{array}{ccc} & \mathbf{1} & \\ {\scriptstyle a}\swarrow & & \searrow{\scriptstyle b} \\ A & \xrightarrow[f]{} & B \end{array} \tag{109}$$

Equivalently, f must send a to b, i.e., $f(a) = b$. You might now recognize that $\mathbf{1}/\mathbf{Set}$ is really the category $\mathbf{Set}_*$ in disguise.

This example suggests we can define an abstract and general way of defining "pointed" things. However, recall that sometimes, $\mathbf{1}$ is not the right object to talk about elements. For instance, in **Grp**, $\mathbf{1}$ is also initial so, by the dual to Exercise 326, $\mathbf{1}/\mathbf{Grp}$ is the same thing as **Grp**. Still, we can easily define the category $\mathbf{Grp}_*$ of pointed groups: its objects are pairs (G, g) where G is a group and $g \in G$, and morphisms $(G, g) \to (H, h)$ are homomorphisms $f : G \to H$ satisfying $f(g) = h$.

SOL **Exercise 331.** Let $\mathbb{Z}$ be the group of integers equipped with addition. Show that one can define the category $\mathbf{Grp}_*$ as $\mathbb{Z}/\mathbf{Grp}$.

SOL **Exercise 332.** Show that for any category $\mathbf{C}$ and object $X \in \mathbf{C}_0$, the slice category $\mathbf{C}/X$ has a terminal object. State and prove the dual statement.

SOL **Exercise 333.** Show that the product of $f : A \to X$ and $g : B \to X$ in $\mathbf{C}/X$ exists if and only if the pullback of $A \xrightarrow{f} X \xleftarrow{g} B$ exists in $\mathbf{C}$. State and prove the dual statement.

These results can be summarized by saying that pullbacks are products in the slice category, and pushouts are coproducts in the coslice category. This allows us to define arbitrary (not binary) pullbacks and pushouts as arbitrary products and coproducts in the slice and coslice categories.[338]

[338] In the literature, these are called **fibered products** and **fibered sums** respectively.

SOL **Exercise 334.** Given two functors $\mathbf{D} \overset{F}{\rightsquigarrow} \mathbf{C} \overset{G}{\leftsquigarrow} \mathbf{E}$, show that an initial object in $F \downarrow G$ is a terminal object in $G^{\mathrm{op}} \downarrow F^{\mathrm{op}}$.

Back to universal properties. We give a more concise definition.

Proposition 335. *Let $F : \mathbf{D} \rightsquigarrow \mathbf{C}$ be a functor, $X \in \mathbf{C}_0$ and $\Delta(X) : \mathbf{1} \rightsquigarrow \mathbf{C}$ be the constant functor. A* ***universal morphism*** *from X to F is an initial object in $\Delta(X) \downarrow F$.*

Proof. Unrolling the definition of initial object in $\Delta(X) \downarrow F$, we find that it is a morphism $a : X \to F(A)$ such that for any other morphism $b : X \to F(B)$, there is unique morphism $(\bullet, A, a) \to (\bullet, B, b)$, that is, a unique morphism $f : A \to B$

making (110) commute.

$$\begin{array}{ccc} X & \xrightarrow{\mathrm{id}_X} & X \\ a\downarrow & & \downarrow b \\ FA & \xrightarrow[Ff]{} & FB \end{array} \tag{110}$$

This is exactly the situation depicted in (98). □

Corollary 336 (Dual)**.** *A universal morphism from F to X is a terminal object in $F \downarrow \Delta(X)$.*

Proof. We said that a universal morphism from F to X is a universal morphism from $X \in \mathbf{C}^{\mathrm{op}}$ to F^{op}. By the previous result, it is an initial object in $\Delta(X) \downarrow F^{\mathrm{op}}$. By Exercise 334, it is a terminal object in $F \downarrow \Delta(X)$. □

In case a universal property is realized by a universal morphism, we can formally prove that this property determines an object up to isomorphism.

SOL **Exercise 337** (NOW!)**.** Show that if there is a universal morphism from X to F and one from Y to F, then $X \cong Y$. State and prove the dual statement.

We have to postpone to Chapter 6 showing that, as we have claimed, any (co)limit satisfies a universal property. Still, you might have noticed that our definition of universal property also uses a special case of (co)limits, that is, initial and terminal objects. What is more, in the following chapters, we will introduce a couple more concepts which often coincide[339] with the concepts of (co)limits and universal properties.

[339] By *coincide*, we mean that one is a special case of the other or vice-versa or both directions.

5 Natural Transformations

5.1 Functor Categories 99
5.2 The 2–category Cat 107
5.3 Equivalences 114

In the previous chapters, we saw how to use the framework of categories to do mathematics. While fundamentally the same as "classical" mathematics,[340] doing mathematics with categories can feel different because we study mathematical structures from above rather than from the inside. Now, if we want to study group theory categorically, we have many options:

[340] We rely on rigorous logical arguments.

- We can study single-object categories where every morphism is invertible (deloopings of groups) and functors between them (group homomorphisms).[341]
- We can go one step higher and study the category **Grp** as a whole. We do not have access to what is inside a group, only how groups relate to each other.[342]
- We can climb another step and study **Grp** as an object of a category of categories.[343]
- In between the previous two items, we can study **Grp** as a subcategory of **Cat**. Taking the delooping is a fully faithful functor $\mathbf{B} : \mathbf{Grp} \rightsquigarrow \mathbf{Cat}$, so we identify **Grp** with its image in **Cat**. We still get to study how groups interact with each other, but also how they interact with other categories.

[341] This amounts to doing "classical" group theory.

[342] This has been our point of view until now.

[343] Recall that due to size issues, **Grp** is not an object of **Cat**, but we could carefully define a category of categories that contains **Grp**.

The first and last step are particular to groups, not all mathematical structures can be viewed as a categories. For instance, studying group theory requires to understand group homomorphisms which are functors, not categories. Taking the categorical mindset to the extreme,[344] we should only have to study how homomorphisms relate to each other, but what is a morphism between homomoprhisms? More generally, what is a morphism between functors?

[344] This might seem extreme at this point, but category theorists can go way further.

5.1 Functor Categories

Natural transformations are admittedly what made mathematicians want to study category theory in the first place. In short, they are morphisms between functors.

The abstract structure of a category is very familiar because it resembles what is found in algebraic structures such as groups, rings or vector spaces.[345] That is to say, it consists of the data of one or more sets with one or more operations satisfying one or more properties. The intuition for morphisms of algebraic structures ported well to categories: a functor comrpises functions between the carrier sets (object and morphisms) that preserve the operations (composition, source and target).

[345] In fact, it is technically called an essentially algebraic structure.

Unfortunately, the definition of a functor does not fit this pattern. It is hard to describe what is the "structure" of a functor. A first step towards defining morphisms between functors is to do it in some special cases.

Following the introduction, you can try to find a satisfying definition of morphism between group homomorphisms $f, g : G \to H$,[346] and then figure out its meaning when f and g are seen as functors $\mathbf{B}G \rightsquigarrow \mathbf{B}H$.

[346] Recall that morphisms should compose and there should be an identity morphism.

We will proceed with another special case. Given a functor $F : \mathbf{C} \rightsquigarrow \mathbf{Set}$, we would like to know what is a *subfunctor* of F.[347] To every object $X \in \mathbf{C}_0$, F assigns a set FX. It makes sense that a subfunctor F' sends X to a subset $F'X \subseteq FX$. To every $f \in \mathrm{Hom}_{\mathbf{C}}(X, Y)$, F assigns a function $Ff : FX \to FY$. It makes sense that a subfunctor F' sends f to a restriction of Ff on the domain $F'X$. Moreover, we need to require the image of $F'f$ (Ff restricted to $F'X$) lies in $F'Y$, otherwise the target of $F'f$ cannot be $F'Y$. We can summarize the constraints on F' with the following commutative square.[348]

[347] If we had a notion of morphisms between functors, we could define a subfunctor as a subobject, i.e. an equivalence class of monomorphisms.

$$\begin{array}{ccc} F'X & \hookrightarrow & FX \\ {\scriptstyle F'f}\downarrow & & \downarrow{\scriptstyle Ff} \\ F'Y & \hookrightarrow & FY \end{array} \tag{111}$$

[348] (111) commutes if and only if $F'f$ is the restriction of Ff to $F'X$.

It turns out this is enough to ensure that F' is a functor. Indeed, $F'(\mathrm{id}_X)$ is the identity map on FX restricted to $F'X$, which is the identity map on $F'X$. Also, for any $f : X \to Y$ and $g : Y \to X$, $F'f \circ F'g$ is the restriction of $F(g \circ f) = Fg \circ Ff$ to $F'X$.[349]

[349] You can check this manually, or pave the following diagram with the squares showing $F'f$ is Ff restricted to $F'X$ and $F'g$ is Fg restricted to $F'Y$.

$$\begin{array}{ccc} F'X & \hookrightarrow & FX \\ & & {\scriptstyle Ff}\downarrow \\ {\scriptstyle F'(g\circ f)}\downarrow & & FY \quad {\scriptstyle F(g\circ f)} \\ & & {\scriptstyle Fg}\downarrow \\ F'Z & \hookrightarrow & FZ \end{array} \tag{112}$$

Example 338. Let F be the maybe functor on **Set** and F' be the identity functor. One can verify that the family of inclusions of X inside $X + \mathbf{1}$ for all sets X yields commutative squares like (111).

We can generalize this to functors with arbitrary codomains.

SOL Exercise 339. Let $F : \mathbf{C} \rightsquigarrow \mathbf{D}$ be a functor. Suppose that for every $X \in \mathbf{C}_0$, there is a monomorphism $F'X \rightarrowtail FX$, and for every $f \in \mathrm{Hom}_{\mathbf{C}}(X, Y)$, there is a morphism $F'f$ making (111) commute. Show that F' is a functor $\mathbf{C} \rightsquigarrow \mathbf{D}$.

This does not strictly define a subfunctor because we still need to quotient by some equivalence saying when two functors represent the same subfunctor of F. Informally, if $F'X \rightarrowtail X$ and $F''X \rightarrowtail X$ always represent the same subobject in the same way, then F' and F'' represent the same subfunctor. To make this formal, we define morphisms of functors in full generality.

Definition 340 (Natural transformation)**.** Let $F, G : \mathbf{C} \rightsquigarrow \mathbf{D}$ be two (covariant) functors, a **natural transformation** $\phi : F \Rightarrow G$ is a map $\phi : \mathbf{C}_0 \to \mathbf{D}_1$ that satisfies $\phi(A) \in \mathrm{Hom}_{\mathbf{D}}(FA, GA)$ for all $A \in \mathbf{C}_0$ and makes (113) commute for any $f \in \mathrm{Hom}_{\mathbf{C}}(A, B)$.[350]

$$\begin{array}{ccc} F(A) & \xrightarrow{\phi(A)} & G(A) \\ {\scriptstyle F(f)}\downarrow & & \downarrow{\scriptstyle G(f)} \\ F(B) & \xrightarrow[\phi(B)]{} & G(B) \end{array} \tag{113}$$

[350] When doing proofs relying on naturality (i.e. the property of being natural), we will use (113) where we instantiate ϕ, F, G, A, B and f with the natural transformation, functors, objects and morphism that is needed in the proof. In order to make this instantiation less painful, we will use the shorthand $\mathsf{NAT}(\phi, A, B, f)$ and instantiate the parameters (we can omit F and G because they should be known from the context). I will try to be this precise whenever I use naturality, but it is very common to simply write "by naturality of ϕ" instead of $\mathsf{NAT}(\phi, A, B, f)$.

Each $\phi(A)$ will be called a **component** of ϕ and may also be denoted with ϕ_A.

As usual, there is an **identity transformation** $\mathbb{1}_F : F \Rightarrow F$[351], it sends every object A to the identity map $\mathrm{id}_{F(A)}$. In the setting of Exercise 339, the monomorphisms $F'X \rightarrowtail FX$ are the components of a natural transformation $F' \Rightarrow F$.[352] Let us go back to our quest to define morphisms of group homomorphisms.

[351] The $\Rightarrow$ (`\Rightarrow`) notation is used more generally for morphisms between morphisms.

[352] To actually define subfunctors, we still need to tell you how to compose natural transformations, but we are not done with examples.

Example 341. Let $f, g : \mathbf{B}G \rightsquigarrow \mathbf{B}H$ be functors (i.e. group homomorphisms), both send the unique object $*$ in $\mathbf{B}G$ to $*$ in $\mathbf{B}H$. Thus, a natural transformation $\phi : f \Rightarrow g$ has a single component $\phi(*) : * \to *$ in H, which is simply an element $\phi \in H$. The commutativity condition is then exhibited by diagram (114) (which lives in $\mathbf{B}H$) for any $x \in G$.

$$\begin{array}{ccc} * & \xrightarrow{\phi} & * \\ {\scriptstyle f(x)}\downarrow & & \downarrow{\scriptstyle g(x)} \\ * & \xrightarrow[\phi]{} & * \end{array} \tag{114}$$

Recall that composition in $\mathbf{B}H$ is just multiplication in H, so naturality of ϕ says that for any $x \in G$, $\phi \cdot f(x) = g(x) \cdot \phi$. Equivalently, $\phi f(x) \phi^{-1} = g(x)$. Therefore, $g = c_\phi \circ f$ where c_ϕ denotes conjugation by ϕ.[353] In short, natural transformations between group homomorphisms correspond to factorizations through conjugations.

[353] In a group $(H, \cdot)$, **conjugation** by an element $h \in H$ is the homomorphism c_h defined $x \mapsto hxh^{-1}$.

Next, a concrete example closer to the general idea of a natural transformation.

Example 342. Fix some $n \in \mathbb{N}$ and define the functor $\mathrm{GL}_n : \mathbf{CRing} \rightsquigarrow \mathbf{Grp}$ by[354]

$$R \mapsto \mathrm{GL}_n(R) \text{ for any commutative ring } R \text{ and}$$
$$f \mapsto \mathrm{GL}_n(f) \text{ for any ring homomorphism } f.$$

[354] The map $\mathrm{GL}_n(f)$ is just the extension of f on $\mathrm{GL}_n(R)$ by applying f to every element of the matrices.

The second functor is $(-)^\times : \mathbf{CRing} \rightsquigarrow \mathbf{Grp}$ which sends a commutative ring R to its group of units $R^\times$ and a ring homomorphism f to $f^\times$, its restriction on $R^\times$. Checking these mappings define two (covariant) functors is left as an exercise, but one might expect these to be functors as they play nicely with the structure of the objects involved.

A natural transformation between these two functors is $\det : \mathrm{GL}_n \Rightarrow (-)^\times$ which maps a commutative ring R to $\det_R$, the function calculating the determinant of a matrix in $\mathrm{GL}_n(R)$. The first thing to check is that $\det_R \in \mathrm{Hom}_{\mathbf{Grp}}(\mathrm{GL}_n(R), R^\times)$ which is clear because the determinant of an invertible matrix is always a unit, $\det_R(I_n) = 1$ and $\det_R$ is a multiplicative map.[355] The second thing is to verify that diagram (115) commutes for any $f \in \mathrm{Hom}_{\mathbf{CRing}}(R, S)$:

[355] i.e. $\det_R(AB) = \det_R(A)\det_R(B)$.

$$\begin{array}{ccc} \mathrm{GL}_n(R) & \xrightarrow{\det_R} & R^\times \\ {\scriptstyle \mathrm{GL}_n(f)}\downarrow & & \downarrow{\scriptstyle f^\times = f|_{R^\times}} \\ \mathrm{GL}_n(S) & \xrightarrow[\det_S]{} & S^\times \end{array} \tag{115}$$

We will check the claim for $n = 2$, but the general proof should only involve more notation to write the bigger expressions, no novel idea. Let $a, b, c, d \in R$, we have

$$\begin{aligned}(\det_S \circ \mathrm{GL}_2(f))\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) &= \det_S\left(\begin{bmatrix} f(a) & f(b) \\ f(c) & f(d) \end{bmatrix}\right) \\ &= f(a)f(d) - f(b)f(c) \\ &= f(ad - bc) \\ &= f^\times(ad - bc) \\ &= (f^\times \circ \det_R)\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right).\end{aligned}$$

We conclude that the diagram commutes and that det is indeed a natural transformation.[356]

[356] Modulo the cases $n > 2$.

SOL **Exercise 343.** Recall the functors $\mathsf{s}, \mathsf{t} : \mathbf{C}^{\to} \rightsquigarrow \mathbf{C}$ defined in Exercise 323. Show that $\phi : \mathsf{s} \Rightarrow \mathsf{t}$ defined by $\phi(f) = f$ for any $f \in \mathbf{C}^{\to}_0 = \mathbf{C}_1$ is a natural transformation.

Because naturality is such a central idea to category theory (just as important as functoriality), we often use it post-rigorously. For instance, when studying a mathematical object X, we might follow some process to obtain another object $F(X)$, and another construction might yield $G(X)$, then we find a process ϕ to go from $F(X)$ to $G(X)$ and we say ϕ is **natural in** X. With these last three words, we implicitly mean a lot of things: that X is an object of some category, that F and G are functors from that category, and that ϕ is the component at X of a natural transformation $F \Rightarrow G$.

It is also possible that F and G take more than one parameter.

SOL **Exercise 344** (NOW!)**.** Let $F, G : \mathbf{C} \times \mathbf{C}' \rightsquigarrow \mathbf{D}$ be two functors. Show that a family

$$\{\phi_{X,Y} : F(X,Y) \to G(X,Y) \mid X \in \mathbf{C}_0, Y \in \mathbf{C}'_0\}$$

is a natural transformation if and only if for any $X \in \mathbf{C}_0$ and $Y \in \mathbf{C}'_0$, both[357]

$$\phi_{X,-} : F(X,-) \Rightarrow G(X,-) \text{ and } \phi_{-,Y} : F(-,Y) \Rightarrow G(-,Y)$$

are natural.

[357] Recall the definition of $F(X,-)$ and $F(-,Y)$ from Exercise 136. If only one of $\phi_{X,-}$ or $\phi_{-,Y}$ is natural, we say that ϕ is natural in X only, respectively Y only. In words, this exercise says that ϕ is natural in X and Y if and only if it is natural in X and natural in Y.

Examples 345 (Natural isomorphisms)**.** A **natural isomorphism** is a natural transformation whose components are all isomorphisms. We have already encountered several of them.

1. When defining exponentials, we saw that currying is a bijection $\mathrm{Hom}_{\mathbf{C}}(B \times X, A) \cong \mathrm{Hom}_{\mathbf{C}}(B, A^X)$. It turns out this is a natural isomorphism from the functor $\mathrm{Hom}_{\mathbf{C}}(- \times X, A) : \mathbf{C}^{\mathrm{op}} \rightsquigarrow \mathbf{Set}$ to $\mathrm{Hom}_{\mathbf{C}}(-, A^X) : \mathbf{C}^{\mathrm{op}} \rightsquigarrow \mathbf{Set}$. We simply need to check the square below commutes for any $f : B \to B'$.[358]

$$\begin{array}{ccc} \mathrm{Hom}_{\mathbf{C}}(B \times X, A) & \xrightarrow{g \mapsto \lambda g} & \mathrm{Hom}_{\mathbf{C}}(B, A^X) \\ {\scriptstyle -\circ(f \times \mathrm{id}_X)}\uparrow & & \uparrow{\scriptstyle -\circ f} \\ \mathrm{Hom}_{\mathbf{C}}(B' \times X, A) & \xrightarrow[g \mapsto \lambda g]{} & \mathrm{Hom}_{\mathbf{C}}(B', A^X) \end{array} \quad (116)$$

[358] Because these functors have $\mathbf{C}^{\mathrm{op}}$ as a source, note the reversal the arrows

Starting with g in the bottom left, we need to prove $\lambda g \circ f = \lambda(g \circ (f \times \mathrm{id}_X))$. The universal property of A^X tells us $\mathsf{ev} \circ (\lambda g \times \mathrm{id}_X) = g$. Pre-composing with $f \times \mathrm{id}_X$, we find

$$g \circ (f \times \mathrm{id}_X) = \mathsf{ev} \circ (\lambda g \times \mathrm{id}_X) \circ (f \times \mathrm{id}_X) = \mathsf{ev} \circ ((\lambda g \circ f) \times \mathrm{id}_X),$$

thus both $\lambda g \circ f$ and $\lambda(g \circ (f \times \mathrm{id}_X))$ make (117) commute, and they must be equal by uniqueness.

$$\begin{array}{ccc} A & \xleftarrow{\mathsf{ev}} & A^X \times X \\ & \nwarrow^{g \circ (f \times \mathrm{id}_X)} & \uparrow{\scriptstyle \lambda g \circ f = \lambda(g \circ (f \times \mathrm{id}_X))} \\ & & B \times X \end{array} \tag{117}$$

2. Without giving all the details, we note that the bijections

$$\begin{aligned} \mathrm{Hom}_{\mathbf{Set}}(A, M) &\cong \mathrm{Hom}_{\mathbf{Mon}}(A^*, M), \text{ and} \\ \mathrm{Hom}_{\mathbf{Grp}}(G, A) &\cong \mathrm{Hom}_{\mathbf{Ab}}(G^{\mathrm{ab}}, A) \end{aligned}$$

are also natural in A and M, and A and G respectively. They are the components of natural isomorphisms[359]

$$\begin{aligned} \mathrm{Hom}_{\mathbf{Set}}(-, U-) &\cong \mathrm{Hom}_{\mathbf{Mon}}(-^*, -), \text{ and} \\ \mathrm{Hom}_{\mathbf{Grp}}(-, U-) &\cong \mathrm{Hom}_{\mathbf{Ab}}(-^{\mathrm{ab}}, -). \end{aligned}$$

In particular, the assignments $A \mapsto A^*$ and $G \mapsto G^{\mathrm{ab}}$ are functorial, and these natural isomorphisms are witnesses to these functors being left adjoints to the corresponding forgetful functors.[360]

[359] Where U denotes the forgetful functors $\mathbf{Mon} \rightsquigarrow \mathbf{Set}$ and $\mathbf{Ab} \rightsquigarrow \mathbf{Grp}$ respectively.

[360] Adjoints are the topic of Chapter 7, where we will study more of these kind of natural isomorphisms.

Now, coming back to our idea that natural transformations are morphisms of functors, we shall explain how they compose.

Definition 346 (Vertical composition)**.** Let $F, G, H : \mathbf{C} \rightsquigarrow \mathbf{D}$ be parallel functors and $\phi : F \Rightarrow G$ and $\eta : G \Rightarrow H$ be two natural transformations. The **vertical composition** of ϕ and η, denoted $\eta \cdot \phi : F \Rightarrow H$ is defined by $(\eta \cdot \phi)(A) = \eta(A) \circ \phi(A)$ for all $A \in \mathbf{C}_0$. If $f : A \to B$ is a morphism in $\mathbf{C}$, then diagram (118) commutes by naturality of ϕ and η, showing that $\eta \cdot \phi$ is a natural transformation from F to H.

The notation $\cdot$ is not widespread, most authors use $\circ$ because vertical composition is the composition in a functor category. I believe the distinction is helpful as you learn this material.

$$\begin{array}{ccccc} F(A) & \xrightarrow{\phi(A)} & G(A) & \xrightarrow{\eta(A)} & H(A) \\ {\scriptstyle F(f)}\downarrow & & {\scriptstyle G(f)}\downarrow & & {\scriptstyle H(f)}\downarrow \\ F(B) & \xrightarrow[\phi(B)]{} & G(B) & \xrightarrow[\eta(B)]{} & H(B) \end{array} \tag{118}$$

The meaning of *vertical* will come to light when horizontal composition is introduced in a bit.

Definition 347 (Functor categories)**.** For any two categories $\mathbf{C}$ and $\mathbf{D}$, there is a **functor category** denoted $[\mathbf{C}, \mathbf{D}]$.[361] Its objects are functors from $\mathbf{C}$ to $\mathbf{D}$, its morphisms are natural transformations between such functors, and the composition is the vertical composition defined above. We leave you to check the associativity of $\cdot$ as it quickly follows from associativity of composition in $\mathbf{D}$. Similarly, you can verify the identity morphism for a functor F is $\mathbb{1}_F$.

[361] Some authors denote it $\mathbf{D}^{\mathbf{C}}$, analogously to the exponential of sets. In fact, $\mathbf{Cat}$ is cartesian closed and $[\mathbf{C}, \mathbf{D}]$ is the exponential. We give most of the proof in Example 373.5.

SOL Exercise 348 (NOW!)**.** Show that natural isomorphisms are precisely the isomorphisms in functor categories.

Functors that are naturally isomorphic are essentially the same functor; they send the same object to isomorphic objects and the same morphism to morphisms that are well-behaved under composition with isomorphisms between the source and targets. This suggests that a natural isomorphism between functors transfers all the properties, we check some of them in Exercise 349.

SOL Exercise 349. Let $F, G : \mathbf{C} \rightsquigarrow \mathbf{D}$ be two naturally isomorphic functors. Show that if F is full/faithful/(co)continuous, then so is G.

Example 350. Recall that a left action of a group G on a set S is just a functor $\mathbf{B}G \rightsquigarrow \mathbf{Set}$. Now, between two such functors $F, F' \in [\mathbf{B}G, \mathbf{Set}]$, a natural transformation is a single map $\sigma : F(*) \to F'(*)$ such that $\sigma \circ F(g) = F'(g) \circ \sigma$ for any $g \in G$. In other words, denoting $\cdot$ for both group actions on $F(*)$ and on $F'(*)$, σ satisfies $\sigma(g \cdot x) = g \cdot (\sigma(x))$ for any $g \in G$ and $x \in F(*)$. In group theory, such a map is called G–**equivariant**.

Therefore, the category $[\mathbf{B}G, \mathbf{Set}]$ can be identified as the category of G–sets (sets equipped with an action of G) with G–equivariant maps as the morphisms.

Examples 351. We can recover constructions we have seen before by studying categories of functors with a simple domain.

1. The terminal category $\mathbf{1}$ has a single object $\bullet$ and no morphism other than the identity. Recall that for any category $\mathbf{C}$, a functor $F : \mathbf{1} \rightsquigarrow \mathbf{C}$ is a simply a choice of object $F(\bullet) \in \mathbf{C}_0$ because $F(\mathrm{id}_\bullet)$ must be equal to $\mathrm{id}_{F(\bullet)}$. If $F, G \in [\mathbf{1}, \mathbf{C}]$, then a natural transformation $\phi : F \Rightarrow G$ is simply a choice of morphism $\phi : F(\bullet) \to G(\bullet)$ because the naturality square (119) for the only morphism $\mathrm{id}_\bullet$ is trivially commutative. Since vertical composition is just componentwise composition, $[\mathbf{1}, \mathbf{C}]$ can be identified with the category $\mathbf{C}$ itself.

$$\begin{array}{ccc} F(\bullet) & \xrightarrow{F(\mathrm{id}_\bullet)} & F(\bullet) \\ {\scriptstyle\phi}\downarrow & & \downarrow{\scriptstyle\phi} \\ G(\bullet) & \xrightarrow[G(\mathrm{id}_\bullet)]{} & G(\bullet) \end{array} \qquad (119)$$

2. Similarly, we can see a functor $F : \mathbf{1} + \mathbf{1} \rightsquigarrow \mathbf{C}$[362] as a choice of two objects $F(\bullet_1)$ and $F(\bullet_2)$ (not necessarily distinct), and a natural transformation $\phi : F \Rightarrow G$ between two such functors as a choice of two morphisms $\phi_1 : F(\bullet_1) \to G(\bullet_1)$ and $\phi_2 : F(\bullet_2) \to G(\bullet_2)$. Therefore, we infer that $[\mathbf{1} + \mathbf{1}, \mathbf{C}]$ can be identified with $\mathbf{C} \times \mathbf{C}$.

[362] Recall $\mathbf{1} + \mathbf{1}$ is the category depicted in (5).

3. Let us go one level harder. A functor $F : \mathbf{2} \rightsquigarrow \mathbf{C}$[363] is a choice of two objects FA and FB as well as a morphism $Ff : FA \to FB$. It can also be seen as a single choice of morphism Ff because FA and FB are determined to be the source and target of Ff respectively. A natural transformation $\phi : F \Rightarrow G$ between two such functors is *not* simply a choice of two morphisms $\phi_A : FA \to GA$ and $\phi_B : FB \to GB$ because, while the naturality squares for id_A and id_B trivially commute, the naturality square (120) for f is an additional constraint on ϕ. Namely, it says (ϕ_A, ϕ_B) makes a commutative square with Ff and Gf, hence we can identify $[\mathbf{2}, \mathbf{C}]$ with the arrow category $\mathbf{C}^{\to}$.

[363] Recall $\mathbf{2}$ is the category depicted in (6).

$$\begin{array}{ccc} FA & \xrightarrow{Ff} & FB \\ {\scriptstyle\phi_A}\downarrow & & \downarrow{\scriptstyle\phi_B} \\ GA & \xrightarrow[Gf]{} & GB \end{array} \qquad (120)$$

SOL Exercise 352. Show that the opposite of $[\mathbf{C}, \mathbf{D}]$ is $[\mathbf{C}^{\mathrm{op}}, \mathbf{D}^{\mathrm{op}}]$.

Viewing any category as a functor category as we did in the previous example has one major consequence formalized in the following results. In short, it says you can infer a lot of things from $[\mathbf{C}, \mathbf{D}]$ by studying $\mathbf{D}$. For instance, if $\mathbf{D}$ has all binary products, it follows that the product of functors F and G in $[\mathbf{C}, \mathbf{D}]$ is the functor sending $X \in \mathbf{C}_0$ to $FX \times GX$ and $f \in \mathbf{C}_1$ to $Ff \times Gf$.[364]

[364] Note that this is not the functor $F \times G$, the latter has type $\mathbf{C} \times \mathbf{C} \rightsquigarrow \mathbf{D} \times \mathbf{D}$.

Theorem 353. *Let* $\mathbf{C}$, $\mathbf{D}$ *and* $\mathbf{J}$ *be categories. If all limits of shape* $\mathbf{J}$ *exist in* $\mathbf{D}$, *then all such limits also exist in* $[\mathbf{C}, \mathbf{D}]$. *Moreover, for any diagram* $F : \mathbf{J} \rightsquigarrow [\mathbf{C}, \mathbf{D}]$ *and for all* $X \in \mathbf{C}_0$, *we have*[365]

$$(\lim_{\mathbf{J}} F)(X) = \lim_{\mathbf{J}}(F(-)(X)).$$

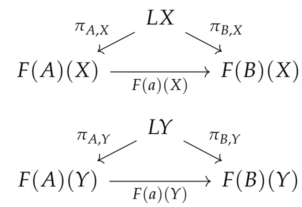
[365] This equation is commonly referred to as "limits in functor categories are computed pointwise".

Proof. Let us explain why the equation above makes sense (i.e. is well-typed).

On the L.H.S., since F is a diagram in $[\mathbf{C}, \mathbf{D}]$, its limit will be an object of $[\mathbf{C}, \mathbf{D}]$, namely a functor $\lim_{\mathbf{J}} F : \mathbf{C} \rightsquigarrow \mathbf{D}$. Thus if $X \in \mathbf{C}_0$, then $(\lim_{\mathbf{J}} F)(X)$ is an object in $\mathbf{D}$.

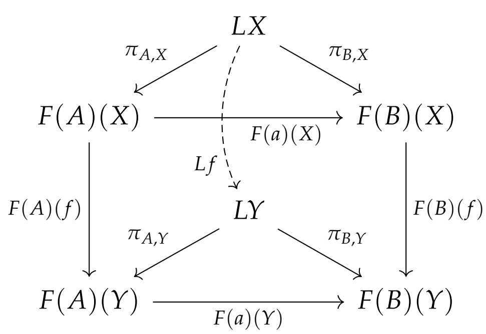
On the R.H.S., fix $X \in \mathbf{C}_0$ and observe that $F(-)(X)$ can be seen as a diagram $\mathbf{J} \rightsquigarrow \mathbf{D}$. Indeed, for $A \in \mathbf{J}_0$, $F(A)$ is a functor from $\mathbf{C}$ to $\mathbf{D}$, so $F(A)(X) \in \mathbf{D}_0$, and for $a : A \to B \in \mathbf{J}_1$, $F(a)$ is a natural transformation from $F(A)$ to $F(B)$, so $F(a)(X)$ (the component of $F(a)$ at X) is a morphism $F(A)(X) \to F(B)(X)$ in $\mathbf{D}$. Then, the limit of $F(-)(X)$ is an object in $\mathbf{D}$ (it exists by hypothesis).

We will define a functor L that sends X to $\lim_{\mathbf{J}}(F(-)(X))$, and we will show it is the limit of F, i.e. $L = \lim_{\mathbf{J}} F$.

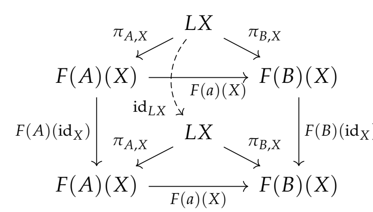


(121)

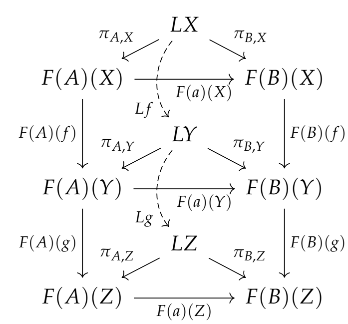
First, we need to define the action of L on morphisms. Let $f : X \to Y$, by definition, LX and LY are limits of $F(-)(X)$ and $F(-)(Y)$ respectively, the limit cones are depicted in (121). For any $a : A \to B$, the naturality of $F(a)$ means the front square in (122) commutes, so the family $\{F(A)(f) \circ \pi_{A,X} : LX \to F(A)(Y)\}_{A \in \mathbf{J}_0}$ forms a cone over $F(-)(Y)$, and the universal property of LY yields a unique morphism Lf making all of (122) commute.



(122)



(123)



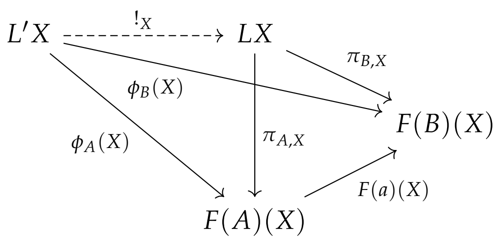
(124)

It follows from uniqueness that $L(\mathrm{id}_X) = \mathrm{id}_{LX}$ and $L(g \circ f) = Lg \circ Lf$ (check that these make (123) and (124) commute). Thus, we have our functor $L : \mathbf{C} \rightsquigarrow \mathbf{D}$.

Next, the back squares in (122) witness the fact that for any $A \in \mathbf{J}_0$, the morphisms $\pi_{A,X}$ are components of a natural transformation $\pi_A : L \Rightarrow F(A)$. Moreover, for any $a : A \to B \in \mathbf{J}_1$, $F(a) \cdot \pi_A = \pi_B$ holds because the commutativity of the triangles in (122) means for every $X \in \mathbf{C}_0$, $F(a)(X) \cdot \pi_{A,X} = \pi_{B,X}$. We conclude that the family $\{\pi_A : L \Rightarrow F(A)\}_{A \in \mathbf{J}_0}$ forms a cone over F. It remains to prove this is the limit cone.

Suppose $\{\phi_A : L' \Rightarrow F(A)\}_{A \in \mathbf{J}_0}$ is another cone over F, that is $F(a) \cdot \phi_A = \phi_B$ for any $a : A \to B \in \mathbf{J}_1$. Looking at the components at X, we find that $\{\phi_A(X) : L'X \to F(A)(X)\}_{A \in \mathbf{J}_0}$ forms a cone over $F(-)(X)$. Thus, the universal property of

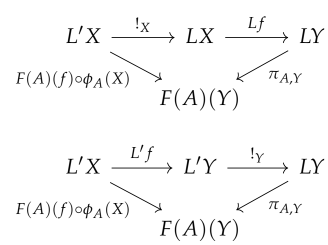
LX yields a unique morphism $!_X$ making (125) commute.



(125)

To show $!_X$ is natural in X, we need to show $Lf \circ !_X = !_Y \circ L'f$ for all $f : X \to Y$. Notice that the target of both sides is LY, so it might be possible to use the universal property of LY to conclude the equation holds. More precisely, we need to find a cone over $F(-)(Y)$ with tip $L'X$ and show $Lf \circ !_X$ and $!_Y \circ L'f$ are morphisms of cone, then by uniqueness they must be the same morphism.

The process we used to make the cone over $F(-)(Y)$ with tip LX in (122) still works for $L'X$. We get a cone $\{F(A)(f) \circ \phi_A(X) : L'X \to F(A)(Y)\}_{A \in \mathbf{J}_0}$. Now, the following derivations show that $Lf \circ !_X$ and $!_Y \circ L'f$ are morphisms of cone as depicted in (126). We conclude $!$ is natural, so we have a cone morphism $! : L' \Rightarrow L$.



(126)

$$\begin{aligned} \pi_{A,Y} \circ Lf \circ !_X &= F(A)(f) \circ \pi_{A,X} \circ !_X && (122) \\ &= F(A)(f) \circ \phi_A(X) && (125) \end{aligned}$$

$$\begin{aligned} \pi_{A,Y} \circ !_Y \circ L'f &= \phi_A(Y) \circ L'f && (125) \\ &= F(A)(f) \circ \phi_A(X) && \mathsf{NAT}(\phi, X, Y, f) \end{aligned}$$

Finally, for any other cone morphism $? : L' \Rightarrow L$, the component of $?$ at X make (125) commute, but $!_X$ is unique with this property. Hence $?_X = !_X$ for all $X \in \mathbf{C}_0$, and we conclude $?$ and $!$ coincide. We conclude that $\lim_{\mathbf{J}} F = L$. $\square$

Corollary 354 (Dual)**.** *Let* $\mathbf{C}$*,* $\mathbf{D}$ *and* $\mathbf{J}$ *be categories. If all colimits of shape* $\mathbf{J}$ *exist in* $\mathbf{D}$*, then all such colimits also exist in* $[\mathbf{C}, \mathbf{D}]$*, and they are computed pointwise.*[366]

[366] Uses Exercise 352.

If you are craving some more diagram chasing or you want to get more familiar with natural transformations and functor categories, you can try doing the following exercises without using Theorem 353 or Corollary 354.[367]

[367] You can essentially reproduce the same proof with the shape $\mathbf{J}$ fixed.

SOL **Exercise 355.** Suppose $\mathbf{D}$ has a terminal object $\mathbf{1}$. Show the constant functor $\Delta(\mathbf{1}) : \mathbf{C} \rightsquigarrow \mathbf{D}$ is terminal in $[\mathbf{C}, \mathbf{D}]$. State and prove the dual statement.

SOL **Exercise 356.** Suppose $\mathbf{D}$ has all binary products and let $F, G \in [\mathbf{C}, \mathbf{D}]_0$. Show that sending $X \in \mathbf{C}_0$ to $FX \times GX$ and $f \in \mathbf{C}_1$ to $Ff \times Gf$ is a functor and it is the product of F and G in $[\mathbf{C}, \mathbf{D}]$. State and prove the dual statement.

SOL **Exercise 357.** Suppose $\mathbf{D}$ has all equalizers and let $\phi, \psi : F \Rightarrow G$ be two parallel natural transformations. For $X \in \mathbf{C}_0$, let (127) be the equalizer in $\mathbf{D}$. Find the action of E on morphisms that make E into a functor $\mathbf{C} \rightsquigarrow \mathbf{D}$ and e into a natural transformation $e : E \Rightarrow F$. Finally, show that e is the equalizer of ϕ and ψ in $[\mathbf{C}, \mathbf{D}]$. State and prove the dual statement.

$$E(X) \xrightarrow{e_X} FX \underset{\psi_X}{\overset{\phi_X}{\rightrightarrows}} GX \qquad (127)$$

SOL **Exercise 358.** Suppose $\mathbf{D}$ has all pullbacks and let $\phi : F \Rightarrow G \Leftarrow H : \psi$ be a cospan of natural transformation. For $X \in \mathbf{C}_0$, let (128) be the pullback in $\mathbf{D}$. Find the action of P on morphisms that makes P into a functor $\mathbf{C} \rightsquigarrow \mathbf{D}$ and $\ell : P \Rightarrow F$ and $r : P \Rightarrow G$ into natural transformation. Finally, show that P with ℓ and r is the pullback of that cospan. State and prove the dual statement.

$$\begin{array}{ccc} P(X) & \xrightarrow{r_X} & HX \\ {\scriptstyle \ell_X}\downarrow & \lrcorner & \downarrow{\scriptstyle \psi_X} \\ FX & \xrightarrow[\phi_X]{} & GX \end{array} \qquad (128)$$

Example 359 ((Co)limits in **DGph**).

Another simple application of viewing a category as a functor category is to look at the evaluation functors.

SOL **Exercise 360.** For any object $X \in \mathbf{C}_0$, show that *evaluation* at X is a functor $-X : [\mathbf{C}, \mathbf{D}] \rightsquigarrow \mathbf{D}$. It sends F to FX and ϕ to ϕ_X.

We leave you to check that the source and target functors $\mathsf{s}, \mathsf{t} : \mathbf{C}^{\rightarrow} \rightsquigarrow \mathbf{C}$ are naturally isomorphic to the functors evaluating at $A \in \mathbf{2}_0$ and $B \in \mathbf{2}_0$ respectively.[368] Evaluating at the single object in $\mathbf{B}G$ yields a forgetful functor $[\mathbf{B}G, \mathbf{Set}] \rightsquigarrow \mathbf{Set}$. It sends a group action to the underlying set and an equivariant map to the underlying function.

[368] This offers an alternative way to show s and t are functors in one go.

Using Exercise 137, we can also conclude there is a functor $\mathsf{Ev} : \mathbf{C} \times [\mathbf{C}, \mathbf{D}] \rightsquigarrow \mathbf{D}$.[369] It sends (X, F) to $F(X)$ and $(f, \phi) : (X, F) \Rightarrow (Y, G)$ to $\phi_Y \circ F(f) = G(f) \circ \phi_X$.

[369] For a fixed $X \in \mathbf{C}_0$, we just saw $\mathsf{Ev}(X, -) = -X$ is a functor. For a fixed $F \in [\mathbf{C}, \mathbf{D}]_0$, $\mathsf{Ev}(-, F)$ is simply the functor F. The equation

$$\begin{aligned} \mathsf{Ev}(Y, \phi) \circ \mathsf{Ev}(f, F) &= \phi_Y \circ F(f) \\ &= G(f) \circ \phi_X \\ &= \mathsf{Ev}(f, G) \circ \mathsf{Ev}(X, \phi) \end{aligned}$$

holds by $\mathsf{NAT}(\phi, X, Y, f)$

We can now restate Theorem 353 and Corollary 354 by saying that when $\mathbf{D}$ has all (co)limits of shape $\mathbf{J}$, then Ev preserves (co)limits in its second component, i.e. for any $X \in \mathbf{C}_0$

$$\mathsf{Ev}(X, \lim_{\mathbf{J}} F) = \lim_{\mathbf{J}} \mathsf{Ev}(X, F-).$$

5.2 The 2–category Cat

It is now time to build intuition for the horizontal composition of natural transformations which will ultimately lead to the notion of a 2–category.

Definition 361 (The left action of functors)**.** Let $F, F' : \mathbf{C} \rightsquigarrow \mathbf{D}$, $G : \mathbf{D} \rightsquigarrow \mathbf{D}'$ be functors and $\phi : F \Rightarrow F'$ a natural transformation as summarized in (129).[370]

[370] Using squiggly arrows for functors in diagrams is very non-standard, but I believe it helps remember what kind of objects we are dealing with. Moreover, since these diagrams are not commutative, it makes a good contrast with the plain arrow notation which was mostly used for commutative diagrams.

$$\mathbf{C} \underset{F'}{\overset{F}{\rightrightarrows}} \mathbf{D} \xrightarrow{G} \mathbf{D}' \qquad (\phi : F \Rightarrow F') \qquad (129)$$

The functor G acts on ϕ by sending it to $G\phi := A \mapsto G(\phi(A)) : \mathbf{C}_0 \to \mathbf{D}'_1$. Showing that (130) commutes for any $f \in \mathrm{Hom}_{\mathbf{C}}(A, B)$ will imply that $G\phi$ is a natural transformation from $G \circ F$ to $G \circ F'$.

$$\begin{array}{ccc} (G \circ F)(A) & \xrightarrow{G\phi(A)} & (G \circ F')(A) \\ {\scriptstyle (G \circ F)(f)}\downarrow & & \downarrow{\scriptstyle (G \circ F')(f)} \\ (G \circ F)(B) & \xrightarrow[G\phi(B)]{} & (G \circ F')(B) \end{array} \qquad (130)$$

Consider this diagram after removing all applications of G, by naturality of ϕ, it is commutative. Since functors preserve commutativity, the diagram still commutes after applying G, hence $G\phi : G \circ F \Rightarrow G \circ F'$ is indeed natural.[371]

[371] More concisely, we apply G to $\mathsf{NAT}(\phi, A, B, f)$ to obtain (130).

We leave you to check this constitutes a left action, namely, for any $G : \mathbf{D} \rightsquigarrow \mathbf{D}'$, $G' : \mathbf{D}' \rightsquigarrow \mathbf{D}''$ and $\phi : F \Rightarrow F'$,

$$\mathrm{id}_{\mathbf{D}}\phi = \phi \text{ and } G'(G\phi) = (G' \circ G)\phi.$$

Definition 362 (The right action of functors)**.** Let $F, F' : \mathbf{C} \rightsquigarrow \mathbf{D}$, $H : \mathbf{C}' \rightsquigarrow \mathbf{C}$ be functors and $\phi : F \Rightarrow F'$ a natural transformation as summarized in (131).

$$\mathbf{C}' \xrightarrow{H} \mathbf{C} \underset{F'}{\overset{F}{\Downarrow \phi}} \mathbf{D} \tag{131}$$

The functor H acts on ϕ by sending it to $\phi H := A \mapsto \phi(H(A)) : \mathbf{C}'_0 \to \mathbf{D}_1$. Showing that (132) commutes for any $f \in \mathrm{Hom}_{\mathbf{C}'}(A, B)$ will imply that ϕH is a natural transformation from $F \circ H$ to $F' \circ H$.

$$\begin{array}{ccc} (F \circ H)(A) & \xrightarrow{\phi H(A)} & (F' \circ H)(A) \\ {\scriptstyle (F\circ H)(f)}\downarrow & & \downarrow{\scriptstyle (F'\circ H)(f)} \\ (F \circ H)(B) & \xrightarrow[\phi H(B)]{} & (F' \circ H)(B) \end{array} \tag{132}$$

Commutativity of (132) follows by naturality of ϕ: change f in diagram (113) with the morphism $H(f) : H(A) \to H(B)$, i.e. (132) is $\mathsf{NAT}(\phi, HA, HB, Hf)$.

We leave you to check this constitutes a right action, namely, for any $H : \mathbf{C}' \rightsquigarrow \mathbf{C}$, $H' : \mathbf{C}'' \rightsquigarrow \mathbf{C}'$ and $\phi : F \Rightarrow F'$,

$$\phi\mathrm{id}_{\mathbf{C}} = \phi \text{ and } (\phi H)H' = \phi(H \circ H').$$

Proposition 363. *The two actions commute, i.e. in the setting of* (133)*,* $G(\phi H) = (G\phi)H$.[372]

[372] For this reason, we will drop all the parentheses from such expressions. We will also drop the $\circ$ for composition of functors. All in all, expect to find expressions like $G'G\phi HH'$ and infer the natural transformation $A \mapsto G'(G(\phi(H(H'(A)))))$.

$$\mathbf{C}' \xrightarrow{H} \mathbf{C} \underset{F'}{\overset{F}{\Downarrow \phi}} \mathbf{D} \xrightarrow{G} \mathbf{D}' \tag{133}$$

Proof. In both the L.H.S. and the R.H.S., an object $A \in \mathbf{C}'_0$ is sent to $G(\phi(H(A)))$. □

SOL **Exercise 364** (NOW!)**.** In the setting of (133), show that the assignments $F \mapsto G \circ F \circ H$ and $\phi \mapsto G\phi H$ make a functor $G(-)H : [\mathbf{C}, \mathbf{D}] \rightsquigarrow [\mathbf{C}', \mathbf{D}']$.

A very useful consequence is that for any commutative diagram in $[\mathbf{C}, \mathbf{D}]$, we can pre-compose and post-compose with any functors and still obtain a commutative diagram. For instance, if (134) commutes in $[\mathbf{C}, \mathbf{D}]$, then for any functors $H : \mathbf{C}' \rightsquigarrow \mathbf{C}$ and $G : \mathbf{D} \rightsquigarrow \mathbf{D}'$ (135) commutes.[373]

[373] We will often use this property by writing things like "apply $G(-)H$ to (134)" to use the commutativity of (135) in a proof.

$$\begin{array}{ccc} X & \xrightarrow{\eta} & Y \\ {\scriptstyle \phi}\big\downarrow & & \big\downarrow{\scriptstyle \phi'} \\ X' & \xrightarrow[\eta']{} & Y' \end{array} \tag{134}$$

$$\begin{array}{ccc} G \circ X \circ H & \xrightarrow{G\eta H} & G \circ Y \circ H \\ {\scriptstyle G\phi H}\big\downarrow & & \big\downarrow{\scriptstyle G\phi' H} \\ G \circ X' \circ H & \xrightarrow[G\eta' H]{} & G \circ Y' \circ H \end{array} \tag{135}$$

We will refer to these two actions as the **biaction** of functors on natural transformations and they will motivate the definition of another way to compose natural transformations.

Let **C**, **D** and **E** be categories, $H, H' : \mathbf{C} \rightsquigarrow \mathbf{D}$ and $G, G' : \mathbf{D} \rightsquigarrow \mathbf{E}$ be functors and $\phi : H \Rightarrow H'$ and $\eta : G \Rightarrow G'$ be natural transformations. This is summarized in (136).

$$\mathbf{C} \underset{H'}{\overset{H}{\rightsquigarrow}} \!\Downarrow\!\phi\; \mathbf{D} \underset{G'}{\overset{G}{\rightsquigarrow}} \!\Downarrow\!\eta\; \mathbf{E} \tag{136}$$

The ultimate goal is to obtain a composition of ϕ and η that is a natural transformation $G \circ H \Rightarrow G' \circ H'$. Note that the biaction defined above yields four other natural transformations:

$$\begin{aligned} G\phi &: G \circ H \Rightarrow G \circ H' & \eta H &: G \circ H \Rightarrow G' \circ H \\ G'\phi &: G' \circ H \Rightarrow G' \circ H' & \eta H' &: G \circ H' \Rightarrow G' \circ H'. \end{aligned}$$

All of the functors involved go from **C** to **E**, so all four natural transformations fit in diagram (137) that lives in the functor category $[\mathbf{C}, \mathbf{E}]$.

$$\begin{array}{ccc} G \circ H & \xrightarrow{G\phi} & G \circ H' \\ {\scriptstyle \eta H}\big\downarrow & & \big\downarrow{\scriptstyle \eta H'} \\ G' \circ H & \xrightarrow[G'\phi]{} & G' \circ H' \end{array} \tag{137}$$

At first glance, this suggests two different definitions for the horizontal composition, that is, the composition of the top path $(\eta H' \cdot G\phi)$ or the composition of the bottom path $(G'\phi \cdot \eta H)$. Surprisingly, both definitions coincide.

Lemma 365. *Diagram* (137) *commutes, i.e.* $\eta H' \cdot G\phi = G'\phi \cdot \eta H$.[374]

[374] Similarly to NAT, we will refer to the commutativity of (137) with $\mathsf{HOR}(\phi, \eta)$. We use HOR because this lemma is crucial in the definition of HORizontal composition.

Proof. Fix an object $A \in \mathbf{C}_0$. Under $\eta H' \cdot G\phi$, it is sent to $\eta(H'(A)) \circ G(\phi(A))$ and under $G'\phi \cdot \eta H$, it is sent to $G'(\phi(A)) \circ \eta(H(A))$. Thus, the proposition is equivalent to saying diagram (138) is commutative (in **E**) for all $A \in \mathbf{C}_0$.

$$\begin{array}{ccc} (G \circ H)(A) & \xrightarrow{G(\phi(A))} & (G \circ H')(A) \\ {\scriptstyle \eta(H(A))}\big\downarrow & & \big\downarrow{\scriptstyle \eta(H'(A))} \\ (G' \circ H)(A) & \xrightarrow[G'(\phi(A))]{} & (G' \circ H')(A) \end{array} \tag{138}$$

This follows from $\mathsf{NAT}(\eta, HA, H'A, \phi(A))$. $\square$

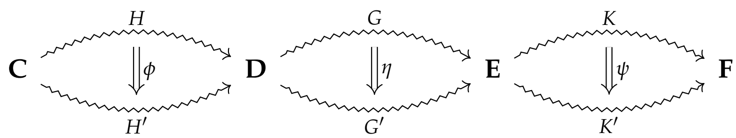
Definition 366 (Horizontal composition)**.** In the setting described in (136), we define the **horizontal composition** of η and ϕ by $\eta \diamond \phi = \eta H' \cdot G\phi = G'\phi \cdot \eta H$.[375]

[375] The $\diamond$ notation is not standard but there are no widespread symbol denoting horizontal composition. I have mostly seen $*$ or plain juxtaposition. Hopefully, you will encounter papers/books clear enough that you can typecheck to find what composition is being used.

One crucial point we have made in earlier chapters is that a notion of composition must satisfy associativity and have identities. We will show the former right after you show the latter.

SOL **Exercise 367.** Let $H : \mathbf{C}' \rightsquigarrow \mathbf{C}$, $F, F' : \mathbf{C} \rightsquigarrow \mathbf{D}$ and $G : \mathbf{D} \rightsquigarrow \mathbf{D}'$ be functors and $\phi : F \Rightarrow F'$ be a natural transformation (as in (133)). Show that $\phi \diamond \mathbb{1}_H = \phi H$ and $\mathbb{1}_G \diamond \phi = G\phi$. Infer that $\mathbb{1}_{\mathrm{id}_\mathbf{C}}$ is the identity at $\mathbf{C}$ for $\diamond$.

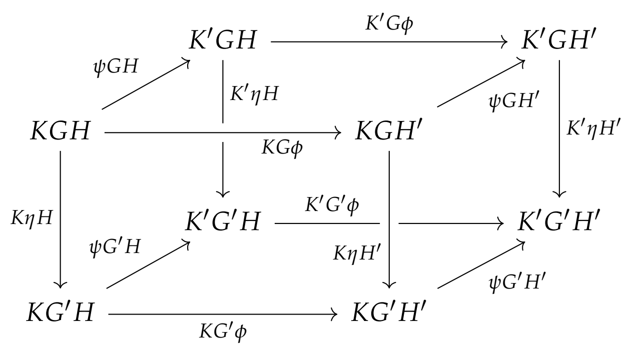
Proposition 368. *In the setting of* (139), $\psi \diamond (\eta \diamond \phi) = (\psi \diamond \eta) \diamond \phi$.



(139)

Proof. Similarly to how we constructed diagram (137) previously, we can use the biaction of functors and composition of functors to obtain the following diagram in $[\mathbf{C}, \mathbf{F}]$.[376]

[376] All $\circ$'s are left out for simplicity.



(140)

Here is how each face commutes.

Top: HOR($\psi, G\eta$)

Bottom: HOR($\psi, G'\eta$)

Left: HOR($\psi, \eta H$)

Right: HOR($\psi, \eta H'$)

Front: HOR($K\eta, \phi$)

Back: HOR($K'\eta, \phi$)

As detailed in the margin, this commutes because each face of the cube corresponds to a variant of diagram (137) (with some substitutions and application of a functor) and combining commutative diagrams yields commutative diagrams. Then, it follows that $\diamond$ is associative because[377] $\psi \diamond (\eta \diamond \phi)$ is the diagonal of the front face followed by the bottom right arrow, and $(\psi \diamond \eta) \diamond \phi$ is the top front arrow followed by the diagonal of the right face. □

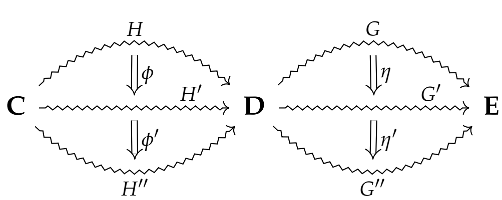
[377] We could have drawn only the front and right face, but the cube is cooler.

There is one last thing to conclude that **Cat** is a 2–category, namely, that the vertical and horizontal compositions interact nicely.

Proposition 369 (Interchange identity)**.** *In the setting of* (142)*, the **interchange identity** holds:*

$$(\eta' \cdot \eta) \diamond (\phi' \cdot \phi) = (\eta' \diamond \phi') \cdot (\eta \diamond \phi). \tag{141}$$

It is in the drawing of (142) that the intuition behind the terms vertical and horizontal is taken.



(142)

Proof. Akin to the other proofs, this is a matter of combining the right diagrams. After combining the diagrams in $[\mathbf{C}, \mathbf{E}]$ corresponding to $\eta \diamond \phi$ and $\eta' \diamond \phi'$, it is easy to see that the R.H.S. of (141) is the morphism going from $G \circ H$ to $G'' \circ H''$ in (143).

$$\begin{array}{ccccc}
G \circ H & \xrightarrow{G\phi} & G \circ H' & & \\
{\scriptstyle \eta H}\downarrow & & \downarrow{\scriptstyle \eta H'} & & \\
G' \circ H & \xrightarrow[G'\phi]{} & G' \circ H' & \xrightarrow{G'\phi'} & G' \circ H'' \\
& & {\scriptstyle \eta' H'}\downarrow & & \downarrow{\scriptstyle \eta' H''} \\
& & G'' \circ H' & \xrightarrow[G''\phi']{} & G'' \circ H''
\end{array} \tag{143}$$

Moreover, the diagram corresponding to the L.H.S. can be factored with the following equations (they follow from Exercise 364) yielding (144).

$$\begin{array}{ccccc}
G \circ H & \xrightarrow{G\phi} & G \circ H' & \xrightarrow{G\phi'} & G \circ H'' \\
{\scriptstyle \eta H}\downarrow & & & & \downarrow{\scriptstyle \eta H''} \\
G' \circ H & & & & G' \circ H'' \\
{\scriptstyle \eta' H}\downarrow & & & & \downarrow{\scriptstyle \eta' H''} \\
G'' \circ H & \xrightarrow[G''\phi]{} & G'' \circ H' & \xrightarrow[G''\phi']{} & G'' \circ H''
\end{array} \tag{144}$$

$$(\eta' \cdot \eta)H = \eta' H \cdot \eta H \qquad (\eta' \cdot \eta)H'' = \eta' H'' \cdot \eta H''$$
$$G(\phi' \cdot \phi) = G\phi' \cdot G\phi \qquad G''(\phi' \cdot \phi) = G''\phi' \cdot G''\phi$$

Combining (143) and (144), we obtain (145) from which the interchange identity readily follows.[378]

[378] The top right and bottom left square commute by $\mathsf{HOR}(\eta, \phi')$ and $\mathsf{HOR}(\eta', \phi)$ respectively. This implies all of (145) commutes and we have seen that the path from $G \circ H$ to $G'' \circ H''$ can be seen as the R.H.S. of (141) by looking at (143) or the L.H.S. by looking at (144). Thus, we infer the satisfaction of (141).

$$\begin{array}{ccccc}
G \circ H & \xrightarrow{G\phi} & G \circ H' & \xrightarrow{G\phi'} & G \circ H'' \\
{\scriptstyle \eta H}\downarrow & & \downarrow{\scriptstyle \eta H'} & & \downarrow{\scriptstyle \eta H''} \\
G' \circ H & \xrightarrow[G'\phi]{} & G' \circ H' & \xrightarrow{G'\phi'} & G' \circ H'' \\
{\scriptstyle \eta' H}\downarrow & & {\scriptstyle \eta' H'}\downarrow & & \downarrow{\scriptstyle \eta' H''} \\
G'' \circ H & \xrightarrow[G''\phi]{} & G'' \circ H' & \xrightarrow[G''\phi']{} & G'' \circ H''
\end{array} \tag{145}$$

□

All of the structure we have added on top of the category **Cat** can be abstracted away by saying that it is 2–category.

Definition 370 (Strict 2–cateory)**.** A **strict** 2**–category** consists of

- a category $\mathbf{C}$,
- for every $A, B \in \mathbf{C}_0$ a category $\mathbf{C}(A, B)$ with $\mathrm{Hom}_{\mathbf{C}}(A, B)$ as its objects and morphisms are called 2**–morphisms** (composition is denoted $\cdot$ and identities $\mathbb{1}$),
- a category with $\mathbf{C}_0$ as its objects, where the morphisms are pairs of parallel morphisms of $\mathbf{C}$ along with a 2–morphism between them. A morphism in this category is also called a 2**–cell**. The identity 2–cell at $A \in \mathbf{C}_0$ is the pair $(\mathrm{id}_A, \mathrm{id}_A)$ and the 2–morphism $\mathbb{1}_{\mathrm{id}_A}$ and composition of 2–cells is denoted $\diamond$),

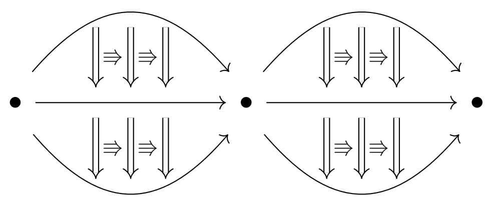
such that the interchange identity (141) holds.[379]

[379] The interchange identity does not come out of nowhere, it is equivalent to the composition $\circ$ being a functor $\mathbf{C}(B, C) \times \mathbf{C}(A, B) \rightsquigarrow \mathbf{C}(A, C)$ that acts on 2–morphisms by $\diamond$ for every $A, B, C \in \mathbf{C}_0$. We leave you to show this in the special case of the 2–category of categories in Exercise 372.

Digression on Higher/Enriched Categories

This book is not the place to further study 2–categories, but we can say a few interesting things about them. There are notions of morphisms between 2–categories (called 2–functors) and morphisms between them (called 2–natural transformations). The latter can be composed in three different ways (analog to vertical and horizontal composition for 2–morphisms) and all possible compositions interact well together. In particular,[380] there is a unique 2–natural transformation that is the composition of all 2–natural transformations in (146) (there are multiple ways to obtain it, depending on what compositions you do in what order, but as in the interchange identity, we require them to lead to the same 2–natural transformation).

[380] There are several so-called coherence axioms that describe how all compositions interact, but we state only one of them.



(146)

The category of 2–categories with 2–functors and 2–natural transformations is now an instance of a 3–category. The field of *higher category theory* studies the generalizations of this to n–categories for any n (even $n = \infty$!). However, most of higher category theory drops the *strict* part of our definition of 2–category because this condition is too strong. Very briefly, they allow the properties of composition, namely associativity, identities and interchange, to hold up to isomorphisms.

There is a relatively simple way to define strict n–categories using *enriched category theory*.[381] The definition of a locally small category can be seen as entirely taking place in the category **Set**. From this point of view, a locally small category is a collection $\mathbf{C}_0$ of objects equipped with

[381] I hope you can indulge this continued digression. While higher and enriched category theory are not as indispensible as basic category theory, they are quite powerful. We will not see how in this book, but I think these two little teasers might inspire some readers to find out by themselves.

- a set $\mathrm{Hom}_{\mathbf{C}}(A, B) \in \mathbf{Set}_0$ for every $A, B \in \mathbf{C}_0$,
- a function $\circ_{A,B,C} \in \mathrm{Hom}_{\mathbf{Set}}\left(\mathrm{Hom}_{\mathbf{C}}(B, C) \times \mathrm{Hom}_{\mathbf{C}}(A, B), \mathrm{Hom}_{\mathbf{C}}(A, C)\right)$ for every $A, B, C \in \mathbf{C}_0$,
- and a function $\mathrm{id}_A \in \mathrm{Hom}_{\mathbf{Set}}\left(\mathbf{1}, \mathrm{Hom}_{\mathbf{C}}(A, A)\right)$,

with conditions that can be stated as commutative diagrams in **Set**. Commutativity of (147) and (148) means that the identity morphisms are neutral with respect to composition and commutativity of (149) means composition is associative.

$$\begin{array}{ccc} \mathrm{Hom}_{\mathbf{C}}(B, C) \times \mathbf{1} & \xrightarrow{\mathrm{id} \times \mathrm{id}_B} & \mathrm{Hom}_{\mathbf{C}}(B, C) \times \mathrm{Hom}_{\mathbf{C}}(B, B) \\ & \searrow^{\pi_{\mathrm{Hom}_{\mathbf{C}}(B,C)}} & \downarrow \circ_{B,B,C} \\ & & \mathrm{Hom}_{\mathbf{C}}(B, C) \end{array} \tag{147}$$

$$\begin{array}{ccc} \mathrm{Hom}_{\mathbf{C}}(B,B) \times \mathrm{Hom}_{\mathbf{C}}(A,B) & \xleftarrow{\mathrm{id}_B \times \mathrm{id}} & \mathrm{Hom}_{\mathbf{C}}(A,B) \times \mathbf{1} \\ \downarrow{\circ_{A,B,B}} & \swarrow{\pi_{\mathrm{Hom}_{\mathbf{C}}(A,B)}} & \\ \mathrm{Hom}_{\mathbf{C}}(A,B) & & \end{array} \tag{148}$$

$$\begin{array}{ccc} \mathrm{Hom}_{\mathbf{C}}(C,D) \times \mathrm{Hom}_{\mathbf{C}}(B,C) \times \mathrm{Hom}_{\mathbf{C}}(A,B) & \xrightarrow{\circ_{B,C,D} \times \mathrm{id}} & \mathrm{Hom}_{\mathbf{C}}(B,D) \times \mathrm{Hom}_{\mathbf{C}}(A,B) \\ \downarrow{\mathrm{id} \times \circ_{A,B,C}} & & \downarrow{\circ_{A,B,D}} \\ \mathrm{Hom}_{\mathbf{C}}(C,D) \times \mathrm{Hom}_{\mathbf{C}}(A,C) & \xrightarrow[\circ_{A,C,D}]{} & \mathrm{Hom}_{\mathbf{C}}(A,D) \end{array} \tag{149}$$

It turns out we can abstract the properties of $\mathbf{1}$ and $\times$ that ensure we can do category theory: we say that $(\mathbf{Set}, \times, \mathbf{1})$ is a **monoidal category**.[382] Now, **enriched category theory** is done by replacing **Set** with another category that has a monoidal structure.

[382] The specific properties are not too relevant for us right now, but know that $\times$ and $\mathbf{1}$ are called the **tensor** and **unit** of the monoidal category.

Examples 371. 1. The category $\mathbf{1}$ is a monoidal category with the tensor and unit being trivial (there is only one object, so there is no choice). A category enriched in $\mathbf{1}$ is simply a collection $\mathbf{C}_0$ because there is no choice when defining $\mathrm{Hom}_{\mathbf{C}}(A,B) \in \mathbf{1}_0$, $\circ_{A,B,C} \in \mathbf{1}_1$ and $\mathrm{id}_A \in \mathbf{1}_1$.

2. Recall that categories can be seen as generalizations of monoids where elements have a source and target, and you can only multiply elements when they are composable. If we started from rings instead, we would have to say how morphisms can be added. For instance in **Ab**, given two parallel morphisms $f, f' : A \to B$, we can add them pointwise $(f + f')(a) = f(a) + f'(a)$.[383] This operation makes $\mathrm{Hom}_{\mathbf{Ab}}(A,B)$ an abelian group. Moreover, you can check that, just as multiplication commutes with addition in a ring, $g \circ (f + f') = (g \circ f) + (g \circ f')$ and $(f + f') \circ h = (f \circ h) + (f' \circ h)$.[384] This is equivalent to saying

$$\circ_{A,B,C} : \mathrm{Hom}_{\mathbf{Ab}}(B,C) \times \mathrm{Hom}_{\mathbf{Ab}}(A,B) \to \mathrm{Hom}_{\mathbf{Ab}}(A,C)$$

is a bilinear map, or equivalently,

$$\circ_{A,B,C} \in \mathrm{Hom}_{\mathbf{Ab}}\left(\mathrm{Hom}_{\mathbf{Ab}}(B,C) \otimes \mathrm{Hom}_{\mathbf{Ab}}(A,B), \mathrm{Hom}_{\mathbf{Ab}}(A,C)\right),$$

where $\otimes$ denotes the tensor product of abelian groups. Noting that $(\mathbf{Ab}, \otimes, \mathbb{Z})$ is a monoidal category, we simply say that **Ab** is enriched in **Ab**. You can check that $\mathbf{Vect}_k$ is also **Ab**–enriched.[385]

[383] The group operation in B is denoted by $+$ because it is commutative.

[384] However, in general,

$$(g + g') \circ (f + f') \neq (g \circ f) + (g' \circ f').$$

[385] You might encounter abelian categories in the wild, these are a special kind of **Ab**–enriched categories.

3. The category **Cat** of small categories is monoidal with the tensor being $\times$ and the unit being $\mathbf{1}$. A category enriched in **Cat** is a strict 2–category. For instance, the 2–category of categories is a collection $\mathbf{Cat}_0$ of objects, a category $\mathbf{Cat}(\mathbf{C}, \mathbf{D}) = [\mathbf{C}, \mathbf{D}]$ for every $\mathbf{C}, \mathbf{D} \in \mathbf{Cat}_0$, a functor $\mathrm{id}_{\mathbf{C}} : \mathbf{1} \rightsquigarrow [\mathbf{C}, \mathbf{C}]$ that picks the identity functor and, as you will show in Exercise 372, a morphism

$$\circ_{\mathbf{C},\mathbf{D},\mathbf{E}} \in \mathrm{Hom}_{\mathbf{Cat}}([\mathbf{D}, \mathbf{E}] \times [\mathbf{C}, \mathbf{D}], [\mathbf{C}, \mathbf{E}]).$$

The diagrams corresponding to (147), (148), and (149) (now they live in **Cat**) commute by results we have shown in this chapter.

4. Generalizing the previous item, a strict n–category is a category enriched in the category of strict $(n-1)$–categories.

5. The posetal category $([0,\infty],\geq)$ is monoidal with the tensor being $+$ (addition) and unit being 0.[386] A category enriched in $[0,\infty]$ is

   - a collection of objects X,
   - for every $x,y \in X$ an element $X(x,y) \in [0,\infty]$, and
   - for every $x,y,z \in X$, an element of $\mathrm{Hom}_{[0,\infty]}(X(y,z)+X(x,y), X(x,z))$.

   We can see the second point as a function $X \times X \to [0,\infty]$, and the third point says that $X(x,z) \leq X(x,y)+X(y,z)$.[387] This looks like a triangle inequality, and in fact all of X looks like a metric space, but where the distance can be infinite, the distance is not symmetric, and two distinct elements can be at distance 0.[388] A $[0,\infty]$–enriched category is also called a Lawvere metric space. If you are enjoying this introduction to enriched category theory, you can try to define *enriched functors*. You should find that for $[0,\infty]$, an enriched functor is a nonexpansive map between Lawvere metric spaces.[389]

[386] We define addition with ∞ in the intuitive way, $x+\infty = \infty + x = \infty$ for all $x \in [0,\infty]$.

[387] Recall there is an element in $\mathrm{Hom}_{[0,\infty]}(r,s)$ if and only if $r \geq s$.

[388] The fact that $X(x,x) = 0$ is witnessed by the identity morphism in $\mathrm{Hom}_{[0,\infty]}(0, X(x,x))$.

[389] This is one reason to define **Met** as we did.

SOL **Exercise 372** (NOW!)**.** Show that there is a functor $[\mathbf{D},\mathbf{E}] \times [\mathbf{C},\mathbf{D}] \rightsquigarrow [\mathbf{C},\mathbf{E}]$ whose action on objects is $(F,G) \mapsto F \circ G$.

5.3 Equivalences

Up to now, we supposedly have been doing everything up to isomorphism. However, in a 2–category and in particular in **Cat**, this can be too restrictive. Fortunately, the new "dimension" of natural transformations allows us to define a relaxed version of equality between categories called equivalence.

Recall that an isomorphism of categories is an isomorphism in the category **Cat**, namely, a functor $F : \mathbf{C} \rightsquigarrow \mathbf{D}$ with an inverse $G : \mathbf{D} \rightsquigarrow \mathbf{C}$ such that $F \circ G = \mathrm{id}_{\mathbf{D}}$ and $G \circ F = \mathrm{id}_{\mathbf{C}}$. As is typical in mathematics, one cannot distinguish between isomorphic categories as they only differ in notations and terminology.[390]

In many situations, we will describe an isomorphism between **C** and **D** by identifying the objects and morphisms in **C** with the objects and morphisms in **D**. That is, the functors are implicit in the discussion. For instance, in Example 330 we argued that $\mathbf{1}/\mathbf{Set}$ and $\mathbf{Set}_*$ are the same category. We really meant that they are isomorphic.[391] Only in rare cases (see Example 373.5 below) will we explicitly define the functor and its inverse.

[390] For example, the monoid isomorphism $\mathbb{N} \cong \{\mathsf{a}\}^*$ offers two ways to talk about the same mathematical object. In particular, it identifies 1 with a, 2 with aa, 3 with aaa, etc.

[391] The details of the construction of the isomorphisms are left to you.

Examples 373. Here are other examples of isomorphic categories that we have already seen and a couple of new ones.

1. It was already shown in Example 350 (the details were implicit) that for a group G, the category $[\mathbf{B}G, \mathbf{Set}]$ is isomorphic to the category of G–sets with G–equivariant maps as morphisms.

Another example for readers who know a bit of advanced algebra. Let k be a field and G a finite group, the categories of $k[G]$–modules ($k[G]$ is the group ring of k over G) and of k–linear representations of G are isomorphic.

2. In Example 351, three other isomorphisms were implicitly given:

$$[\mathbf{1}, \mathbf{C}] \cong \mathbf{C} \qquad [\mathbf{1}+\mathbf{1}, \mathbf{C}] \cong \mathbf{C} \times \mathbf{C} \qquad [\mathbf{2}, \mathbf{C}] \cong \mathbf{C}^{\rightarrow}.$$

3. The category **Rel** of sets with relations is isomorphic to $\mathbf{Rel}^{\mathrm{op}}$.[392] The functor $\mathbf{Rel} \rightsquigarrow \mathbf{Rel}^{\mathrm{op}}$ is the identity on objects and sends a relation $R \subseteq X \times Y$ to the opposite relation $Я \subseteq Y \times X$ (which is a morphism $X \to Y$ in $\mathbf{Rel}^{\mathrm{op}}$) defined by $(y, x) \in Я \Leftrightarrow (x, y) \in R$. The inverse is defined similarly.

4. Let $\mathbf{C}$ and $\mathbf{D}$ be categories the functor $\mathsf{swap} : \mathbf{C} \times \mathbf{D} \rightsquigarrow \mathbf{D} \times \mathbf{C}$ sends (A, B) to (B, A) and (f, g) to (g, f). It is easy to check that swap is a functor with inverse $\mathsf{swap}^{-1} : \mathbf{D} \times \mathbf{C} \rightsquigarrow \mathbf{C} \times \mathbf{D}$ defined in the obvious way.

5. Given three categories $\mathbf{C}$, $\mathbf{D}$ and $\mathbf{E}$, there is an isomorphism[393]

$$[\mathbf{C} \times \mathbf{D}, \mathbf{E}] \cong [\mathbf{C}, [\mathbf{D}, \mathbf{E}]].$$

The construction of the isomorphism follows the intuition of currying and uncurrying of functions, so the definitions are straightforward. Still, you will see that verifying the straightforward defintions are well-typed is cumbersome (but simple) because there are several levels of functors and natural transformations.

Let $F : \mathbf{C} \times \mathbf{D} \rightsquigarrow \mathbf{E}$, the currying of F is $\Lambda F : \mathbf{C} \rightsquigarrow [\mathbf{D}, \mathbf{E}]$ defined as follows. For $X \in \mathbf{C}_0$, the functor $\Lambda F(X)$ sends $Y \in \mathbf{D}_0$ to $F(X, Y)$ and $g \in \mathbf{D}_1$ to $F(\mathrm{id}_X, g)$. We showed in Exercise 136 that $\Lambda F(X) = F(X, -)$ is a functor. For $f \in \mathrm{Hom}_{\mathbf{C}}(X, X')$, we define the natural transformation $\Lambda F(f) : F(X, -) \Rightarrow F(X', -)$ by

$$\Lambda F(f)_Y = F(f, \mathrm{id}_Y) : F(X, Y) \to F(X', Y).$$

$$\begin{array}{ccc} F(X,Y) & \xrightarrow{F(\mathrm{id}_X, g)} & F(X,Y') \\ {\scriptstyle F(f,\mathrm{id}_Y)}\downarrow & & \downarrow{\scriptstyle F(f,\mathrm{id}_{Y'})} \\ F(X',Y) & \xrightarrow[F(\mathrm{id}_{X'}, g)]{} & F(X',Y') \end{array} \qquad (150)$$

The naturality square (150) is commutative because, by functoriality of F, the top and bottom path are equal to $F(f, g)$. We also have to show ΛF is a functor, namely $\Lambda F(\mathrm{id}_X) = \mathbb{1}_{F(X,-)}$ and $\Lambda F(f \circ f') = \Lambda F(f) \cdot \Lambda F(f')$. We can verify this componentwise using functoriality of F.

$$\Lambda F(\mathrm{id}_X)_Y = F(\mathrm{id}_X, \mathrm{id}_Y) = \mathrm{id}_{F(X,Y)}$$
$$\Lambda F(f \circ f')_Y = F(f \circ f', \mathrm{id}_Y) = F(f, \mathrm{id}_Y) \circ F(f', \mathrm{id}_Y) = \Lambda F(f)_Y \circ \Lambda F(f')_Y.$$

It remains to define $\Lambda-$ on morphisms. Given a natural transformation $\phi : F \Rightarrow F'$, we define $\Lambda\phi : \Lambda F \Rightarrow \Lambda F'$ at component $X \in \mathbf{C}_0$ by the natural transformation:

$$\Lambda\phi(X) = \phi_{X,-} : F(X, -) \Rightarrow F'(X, -).$$

We showed in Exercise 344 that $\phi_{X,-}$ is natural. Finally, we can check that $\Lambda-$ is a functor with the following derivations.[394]

$$\Lambda \mathbb{1}_F(X) = (\mathbb{1}_F)_{X,-} = \mathbb{1}_{F(X,-)}$$
$$\Lambda(\phi \cdot \eta)(X) = (\phi \cdot \eta)_{X,-} = \phi_{X,-} \cdot \eta_{X,-} = \Lambda\phi \cdot \Lambda\eta$$

[392] An arbitrary category $\mathbf{C}$ is not always isomorphic to its opposite. While the opposite functors $(-)^{\mathrm{op}}_{\mathbf{C}} : \mathbf{C} \rightsquigarrow \mathbf{C}^{\mathrm{op}}$ and $(-)^{\mathrm{op}}_{\mathbf{C}^{\mathrm{op}}} : \mathbf{C}^{\mathrm{op}} \rightsquigarrow \mathbf{C}$ are inverses of each other, they are contravariant functors, i.e. they are not morphisms in **Cat**.

[393] You might recognize a similarity with exponentials which rely on an isomorphism $\mathrm{Hom}_{\mathbf{C}}(B \times X, A) \cong \mathrm{Hom}_{\mathbf{C}}(B, A^X)$. The example here is more than an instance of exponentials of categories because the isomorphism is not only as sets but as categories.

[394] The second equation on the second line can be verified componentwise, i.e. for every $Y \in \mathbf{D}_0$

$$(\phi \cdot \eta)_{X,Y} = \phi_{X,Y} \circ \eta_{X,Y} = (\phi_{X,-} \cdot \eta_{X,-})_Y.$$

Conversely, let $F : \mathbf{C} \rightsquigarrow [\mathbf{D}, \mathbf{E}]$, the uncurrying of F is $\Lambda^{-1}F : \mathbf{C} \times \mathbf{D} \rightsquigarrow \mathbf{E}$ defined as follows. We use Exercise 137 to define $\Lambda^{-1}F$ componentwise. Fixing $X \in \mathbf{C}_0$, we know that $F(X)$ is a functor, so we set $\Lambda^{-1}F(X, -) = F(X)$. Fixing $Y \in \mathbf{D}_0$, we define $\Lambda^{-1}F(-, Y)$ on objects by sending $X \in \mathbf{C}_0$ to $F(X)(Y)$ and $f \in \mathbf{C}_1$ to $F(f)_Y$.[395] To show $\Lambda^{-1}F(-, Y)$ is a functor, we use the functoriality of F as follows.

[395] As a sanity check, if $f : X \to X'$, $F(f) : F(X) \Rightarrow F(X')$, thus the component at Y is $F(f)_Y : F(X)(Y) \to F(X')(Y)$ as desired.

$$\Lambda^{-1}F(\mathrm{id}_X, Y) = F(\mathrm{id}_X)_Y = \mathbb{1}_{F(X)_Y} = \mathrm{id}_{F(X)(Y)}$$
$$\Lambda^{-1}F(f \circ f', Y) = F(f \circ f')_Y = (F(f) \cdot F(f'))_Y = F(f)_Y \circ F(f')_Y.$$

Now, for every $f : X \to X'$ and $g : Y \to Y'$, the naturality of $F(f)$ implies the square in (151) commutes. This means we can let $\Lambda^{-1}F(f, g)$ be the diagonal, i.e.

$$\Lambda^{-1}F(f, g) := \Lambda^{-1}F(X', g) \circ \Lambda^{-1}F(f, Y) = \Lambda^{-1}F(f, Y') \circ \Lambda^{-1}F(X, g),$$

$$\begin{array}{ccc} F(X)(Y) & \xrightarrow{F(X)(g)} & F(X)(Y') \\ {\scriptstyle F(f)_Y}\downarrow & & \downarrow{\scriptstyle F(f)_{Y'}} \\ F(X')(Y) & \xrightarrow[F(X')(g)]{} & F(X')(Y') \end{array} \qquad (151)$$

and conclude by Exercise 137 that $\Lambda^{-1}F : \mathbf{C} \times \mathbf{D} \rightsquigarrow \mathbf{E}$ is a functor.

Given a natural transformation $\phi : F \Rightarrow F'$, we define $\Lambda^{-1}\phi : \Lambda^{-1}F \Rightarrow \Lambda^{-1}F'$ by $(\Lambda^{-1}\phi)_{X,Y} := (\phi_X)_Y$. By Exercise 344, it is enough to show naturality in one component at a time. Fix $X \in \mathbf{C}_0$, by hypothesis (ϕ_X is a morphism in $[\mathbf{D}, \mathbf{E}]$), $\phi_X : F(X) \Rightarrow F'(X)$ is natural in Y. Fix $Y \in \mathbf{D}_0$, we need to show the following square commutes.

$$\begin{array}{ccc} F(X)(Y) & \xrightarrow{\Lambda^{-1}F(f,Y)} & F(X')(Y) \\ {\scriptstyle (\phi_X)_Y}\downarrow & & \downarrow{\scriptstyle (\phi_{X'})_Y} \\ F'(X)(Y) & \xrightarrow[\Lambda^{-1}F'(f,Y)]{} & F'(X')(Y) \end{array} \qquad (152)$$

Recalling that $\Lambda^{-1}F(f, Y) = F(f)_Y$ and $\Lambda^{-1}F'(f, Y) = F'(f)_Y$, we recognize this square as $\mathsf{NAT}(\phi, X, X', f)$ evaluated at Y. Finally, we can check that $\Lambda^{-1}-$ is a functor with the following derivations.

$$(\Lambda^{-1}\mathbb{1}_F)_{X,Y} = ((\mathbb{1}_F)_X)_Y = \mathrm{id}_{F(X)(Y)} = (\mathbb{1}_{\Lambda^{-1}F})_{X,Y}$$
$$(\Lambda^{-1}\phi \cdot \eta)_{X,Y} = ((\phi \cdot \eta)_X)_Y = (\phi_X)_Y \circ (\eta_X)_Y = (\Lambda^{-1}\phi)_{X,Y} \cdot (\Lambda^{-1}\eta)_{X,Y}$$

The last step (I promise) of this proof is to show that $\Lambda-$ and $\Lambda^{-1}-$ are inverses of each other. The mindless computations below suffice.

$$\Lambda\Lambda^{-1}F(X)(Y) = \Lambda^{-1}F(X, Y) = F(X)(Y)$$
$$\Lambda\Lambda^{-1}F(f)_Y = \Lambda^{-1}F(f, Y) = F(f)_Y$$

$$\Lambda^{-1}\Lambda F(X, Y) = \Lambda F(X)(Y) = F(X, Y)$$
$$\Lambda^{-1}\Lambda F(f, g) = \Lambda F(X')(g) \circ \Lambda F(f)_Y = F(\mathrm{id}_{X'}, g) \circ F(f, \mathrm{id}_Y) = F(f, g)$$

Of course, the list above is not exhaustive, but it is time to introduce equivalences. Instead of requiring the round trips between $\mathbf{C}$ and $\mathbf{D}$ to be the identities, we merely require they are naturaly isomorphic to the identities.

Definition 374 (Equivalence)**.** A functor $F : \mathbf{C} \rightsquigarrow \mathbf{D}$ is an **equivalence** of categories if there exists a functor $G : \mathbf{D} \rightsquigarrow \mathbf{C}$ such that $F \circ G \cong \mathrm{id}_{\mathbf{D}}$ and $G \circ F \cong \mathrm{id}_{\mathbf{C}}$.[396] This is clearly symmetric, so we say two categories $\mathbf{C}$ and $\mathbf{D}$ are **equivalent**, denoted $\mathbf{C} \simeq \mathbf{D}$, if there is an equivalence between them. Moreover, we say that G is a **quasi-inverse** of F and vice-versa.

[396] Recall that $\cong$ between functors stands for natural isomorphisms.

This is certainly weaker than an isomorphism of categories, but it is still quite strong. In order to gain more intuition on how equivalences equate two categories, let us observe what properties this forces on the functor F. For all $f \in \mathrm{Hom}_{\mathbf{C}}(A, B)$, the following square commutes where ϕ_A and ϕ_B are isomorphisms.[397]

$$\begin{array}{ccc} A & \xrightarrow{\quad f \quad} & B \\ \phi_A^{-1}\uparrow\downarrow\phi_A & & \phi_B\downarrow\uparrow\phi_B^{-1} \\ GF(A) & \xrightarrow[GF(f)]{} & GF(B) \end{array} \tag{153}$$

[397] Naturality of ϕ only gives us $GF(f) \circ \phi_A = \phi_B \circ f$, but by composing with ϕ_A^{-1} or ϕ_B^{-1}, we obtain the commutativity of all of (153). In particular, we have $GF(f) = \phi_B \circ f \circ \phi_A^{-1}$.

This implies that the map $f \mapsto GF(f) : \mathrm{Hom}_{\mathbf{C}}(A, B) \to \mathrm{Hom}_{\mathbf{C}}(GF(A), GF(B))$ is a bijection. Indeed, pre-composition by ϕ_A^{-1} and post-composition by ϕ_B are both bijections,[398] so

$$f \mapsto \phi_B \circ f \circ \phi(A)^{-1} = GF(f)$$

is a bijection. Since A and B are arbitrary, we conclude $G \circ F$ is a fully faithful functor and a symmetric argument shows $F \circ G$ is also fully faithful. Then, it is easy to conclude that F and G must be fully faithful as well.[399]

[398] Recall the definitions of monomorphisms and epimorphisms and the fact that isomorphisms are monic and epic.

[399] Recall Exercise 124

What is more, the existence of an isomorphism $\eta_A : A \to FG(A)$ for any object A implies F (symmetrically G) has the following property.

Definition 375 (Essentially surjective)**.** A functor $F : \mathbf{C} \rightsquigarrow \mathbf{D}$ is **essentially surjective** if for any $X \in \mathbf{D}_0$, there exists $Y \in \mathbf{C}_0$ such that $X \cong F(Y)$.[400]

[400] Intuitively, this property means that while the image of F may not be all of $\mathbf{D}$, everything outside the image is at least isomorphic to somethig in the image.

We will show that these two properties (full faithfulness and essential surjectivity) are necessary and sufficient for F to be an equivalence.

Theorem 376. *A functor $F : \mathbf{C} \rightsquigarrow \mathbf{D}$ is an equivalence of categories if and only if F is fully faithful and essentially surjective.*

Proof. ($\Rightarrow$) Shown above.

($\Leftarrow$) We construct a functor $G : \mathbf{D} \rightsquigarrow \mathbf{C}$ such that $G \circ F \cong \mathrm{id}_{\mathbf{C}}$ and $F \circ G \cong \mathrm{id}_{\mathbf{D}}$.[401] Since F is essentially surjective, for any $A \in \mathbf{D}_0$, there exists an object $G(A) \in \mathbf{C}_0$ and an isomorphism $\phi_A : F(G(A)) \cong A$. Hence, $A \mapsto G(A)$ is a good candidate to describe the action of G on objects.

[401] The quasi-inverse of F. We can say *the* thanks to Exercise 377.

Next, similarly to the converse direction, note that for any $A, B \in \mathbf{D}_0$, the map

$$f \mapsto \phi_B \circ f \circ \phi_A^{-1}$$

is a bijection from $\mathrm{Hom}_{\mathbf{D}}(A, B)$ to $\mathrm{Hom}_{\mathbf{D}}(FG(A), FG(B))$. Moreover, since the functor F is fully faithful, it induces a bijection

$$F_{GA,GB} : \mathrm{Hom}_{\mathbf{C}}(G(A), G(B)) \to \mathrm{Hom}_{\mathbf{D}}(FG(A), FG(B))$$

which in turns yields a bijection

$$G_{A,B} : \mathrm{Hom}_{\mathbf{D}}(A,B) \to \mathrm{Hom}_{\mathbf{C}}(G(A),G(B)) = f \mapsto F^{-1}_{GA,GB}(\phi_B \circ f \circ \phi_A^{-1}).$$

This is the action of G on morphisms. Observe that the construction of G ensures that $F \circ G \cong \mathrm{id}_{\mathbf{D}}$ through the natural transformation ϕ. It remains to show that G is indeed a functor and find a natural isomorphism $\eta : G \circ F \cong \mathrm{id}_{\mathbf{C}}$.

For any composable morphisms $(f,g) \in \mathbf{D}_2$, it is easy to verify that

$$F(G(f) \circ G(g)) = FG(f) \circ FG(g) = FG(f \circ g),$$

so functoriality of G because F is faithful. To find η, recall that the definition of G yields commutativity of (154) for any $f \in \mathrm{Hom}_{\mathbf{C}}(A,B)$.

$$\begin{array}{ccc} F(A) & \xrightarrow{F(f)} & F(B) \\ \phi_{FA} \updownarrow & & \updownarrow \phi_{FB} \\ FGF(A) & \xrightarrow[FGF(f)]{} & FGF(B) \end{array} \tag{154}$$

Then, because F is fully faithful, (**??**) also commutes in $\mathbf{C}$ where $\eta_X = F^{-1}_{X,GFX}(\phi_{FX})$, and we conclude that η is a natural isomorphism $\mathrm{id}_{\mathbf{C}} \cong G \circ F$.[402]

$$\begin{array}{ccc} A & \xrightarrow{f} & B \\ \eta_A \updownarrow & & \updownarrow \eta_B \\ GF(A) & \xrightarrow[GF(f)]{} & GF(B) \end{array} \tag{155}$$

□

[402] You can manually derive that η_X is an isomorphism or use the fact that fully faithful functors reflect isomorphisms (Exercise 188).

The insight to extract from this argument is that two categories are equivalent if they describe the same objects and morphisms with the only relaxation that isomorphic objects can appear any number of times in either category. In contrast, categories can only be isomorphic if they have exactly the same objects and morphisms.

When constructing the quasi-inverse of F in Theorem 376, we had to pick one $G(A)$ for every A such that $A \cong FG(A)$ and one isomorphism $\phi_A : A \cong FG(A)$. These choices rely on the axiom of choice. There is some literature on doing category theory constructively and it relies on anafunctors. Those were defined precisely to avoid the axiom of choice in the proof above.

SOL **Exercise 377** (NOW!)**.** Let $F : \mathbf{C} \rightsquigarrow \mathbf{D}$ and $G, G' : \mathbf{D} \rightsquigarrow \mathbf{C}$ be two quasi-inverses to F. Show that $G \cong G'$.

SOL **Exercise 378.** Let $F : \mathbf{C} \rightsquigarrow \mathbf{D}$ be an equivalence. Show that if $F \cong F'$, then F' is an equivalence.

We will detail a couple of *easy* examples of equivalences and briefly mention a few *harder* ones. Of course, all the isomorphisms of categories we saw earlier are examples of equivalences where the natural isomorphisms are identities.

Examples 379 (Easy)**.** 1. Consider the full subcategory of **FinSet** consisting only of the sets $\emptyset, \{1\}, \{1,2\}, \dots, \{1,\dots,n\}, \dots$, we denoted it by **FinOrd**. The inclusion functor is fully faithful by definition and we claim it is essentially surjective. Indeed, any set $X \in \mathbf{FinSet}_0$ has a finite cardinality n, so $X \cong \{1,\dots,n\}$ and the latter belongs to **FinOrd**.

2. In a very similar fashion, an early result in linear algebra says that any finite dimensional vector space over a field k is isomorphic to k^n for some $n \in \mathbb{N}$. Thus, the category whose objects are k^n for all $n \in \mathbb{N}$ and morphisms are $m \times n$ matrices with entries in k,[403] which we denote $\mathbf{Mat}(k)$, is equivalent to the category of finite dimensional vector spaces.

3. A **partial** function $f : X \rightharpoonup Y$ is a function that may not be defined on all of X.[404] There is category $\mathbf{Par}$ of sets and partial functions where identity morphism and composition are defined straightforwardly.[405] We can view a partial function $f : X \rightharpoonup Y$ as a total function $f' : X \to Y + \mathbf{1}$ which sneds x to $f(x)$ when the latter is defined and to $* \in \mathbf{1}$ otherwise. Further extending f' to $[f', \mathrm{id}_{\mathbf{1}}] : X + \mathbf{1} \to Y + \mathbf{1}$, we can see any partial function as a function between pointed sets where the distinguished element corresponds to being undefined.

   This yields a fully faithful functor $F : \mathbf{Par} \rightsquigarrow \mathbf{Set}_*$ sending X to $(X + \mathbf{1}, *)$ and $f : X \rightharpoonup Y$ to $[f', \mathrm{id}_{\mathbf{1}}]$.[406] This functor is essentially surjective because for every pointed set (X, x), we find an isomorphism $(X \setminus \{x\} + \mathbf{1}, *) \to (X, x)$ that sends $y \in X \setminus \{x\}$ to y and $*$ to x. We infer the quasi-inverse to F sends a pointed set (X, x) to $X \setminus \{x\}$ and a function $f : (X, x) \to (Y, y)$ to the partial function $X \setminus \{x\} \to Y \setminus \{y\}$ that acts like f but is undefined whenever $f(a) = y$.

[403] After making a choice of basis for all k^n, an $m \times n$ matrix with entries in k corresponds to a linear map $k^n \to k^m$.

[404] In this context, a *normal* function defined on all of X is called **total**.

[405] You can view $\mathbf{Par}$ as the subcategory of $\mathbf{Rel}$ where you only take the relations $R \subseteq X \times Y$ satisfying for any $x \in X$ (cf. Remark 116),

$$|\{y \in Y \mid (x, y) \in R\}| \leq 1.$$

[406] We have already seen in Corollary 268 that $[f', \mathrm{id}_{\mathbf{1}}] = [g', \mathrm{id}_{\mathbf{1}}]$ if and only if $f' = g'$. It should be clear from the definition that $f' = g'$ if and only if $f = g$.

The first two examples and many other simple examples of equivalences are examples of skeletons. They are morally a subcategory where all the isomorphic copies are removed.

Definition 380 (Skeleton)**.** A category is called **skeletal** if there it contains no two isomorphic objects. A **skeleton** of a category is an equivalent skeletal category.

Examples 381. We have said that $\mathbf{FinOrd} \simeq \mathbf{FinSet}$ and $\mathbf{Mat}(k) \simeq \mathbf{FDVect}_k$ and we leave to you the easy task to check that these are examples of skeletons.[407]

Any posetal category is skeletal because whenever $x \leq y$ and $y \leq x$, we have $x = y$ which means no two distinct object can be isomorphic.

[407] Namely, you should show that no two sets in $\mathbf{FinOrd}$ are isomorphic and no two spaces in $\mathbf{Mat}(k)$ are isomorphic.

A category always has a skeleton if you assume the axiom of choice and the next result justifies us calling it *the* skeleton of a category.

SOL **Exercise 382.** Show that all skeletons of a category are isomorphic.

Here are other more interesting examples of equivalent categories.

Example 383 (Medium)**.** Let $\mathbf{C}$ be a category, the functor $\mathsf{id} : \mathbf{C} \rightsquigarrow \mathbf{C}^{\to}$ sends X to id_X and $f : X \to Y$ to the commutative square in (156). This functor is an equivalence if and only if all morphisms in $\mathbf{C}$ are isomorphisms.[408] It is clearly fully faithful, so it is left to show id is essentially surjective if and only if $\mathbf{C}$ is a groupoid.

[408] Such a category is called a **groupoid**.

$$\begin{array}{ccc} X & \xrightarrow{\mathrm{id}_X} & X \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle f} \\ Y & \xrightarrow[\mathrm{id}_Y]{} & Y \end{array} \qquad (156)$$

($\Rightarrow$) For any $f : X \to Y \in \mathbf{C}_1$, by hypothesis, there exists $A \in \mathbf{C}_0$ such that $\mathrm{id}_A \cong f$ in $\mathbf{C}^{\to}$. Let $(s : A \to X, t : A \to Y)$ be the isomorphism, its inverse must be (s^{-1}, t^{-1}). Looking at the chain of commutative squares in (157), we can infer that $s \circ t^{-1}$ is the inverse of f.[409]

[409] The composition $f \circ s \circ t^{-1}$ is the top path of the combined two leftmost squares, the bottom path is $t \circ t^{-1} \circ \mathrm{id}_Y = \mathrm{id}_Y$. The composition $s \circ t^{-1} \circ f$ is the bottom path of the combined two rightmost squares, the top path is $\mathrm{id}_X \circ s \circ s^{-1} = \mathrm{id}_X$.

$$\begin{array}{ccccccccc} Y & \xrightarrow{t^{-1}} & A & \xrightarrow{s} & X & \xrightarrow{s^{-1}} & A & \xrightarrow{s} & X \\ {\scriptstyle \mathrm{id}_Y}\downarrow & & {\scriptstyle \mathrm{id}_A}\downarrow & & {\scriptstyle f}\downarrow & & {\scriptstyle \mathrm{id}_A}\downarrow & & \downarrow{\scriptstyle \mathrm{id}_X} \\ Y & \xrightarrow[t^{-1}]{} & A & \xrightarrow[t]{} & Y & \xrightarrow[t^{-1}]{} & A & \xrightarrow[s]{} & X \end{array} \tag{157}$$

($\Leftarrow$) Let $f : X \to Y$ be an object of $\mathbf{C}^{\to}$, the inverse of f satisfies $f \circ f^{-1} = \mathrm{id}_Y$ and $f^{-1} \circ f = \mathrm{id}_X$, so the squares in (158) are isomorphisms in $\mathbf{C}^{\to}$ (they are inverses of each other). Thus, we find that f is isomorphic to id_X which is in the image of id.

$$\begin{array}{ccc} X & \xrightarrow{\mathrm{id}_X} & X \\ {\scriptstyle \mathrm{id}_X}\downarrow & & \downarrow{\scriptstyle f} \\ X & \xrightarrow[f]{} & Y \end{array} \qquad \begin{array}{ccc} X & \xrightarrow{f} & Y \\ {\scriptstyle \mathrm{id}_X}\downarrow & & \downarrow{\scriptstyle f^{-1}} \\ X & \xrightarrow[\mathrm{id}_X]{} & X \end{array} \tag{158}$$

SOL Exercise 384. The category **Setoid** is the full subcategory of 2**Rel** containing only (X, R) where R is an equivalence relation. Is **Set** equivalent to **Setoid**?

Examples 385 (Hard)**.** Examples of significant equivalences are all over the place in higher mathematics. However, they require a bit of work to describe them, thus let us only say a few words on a couple of them.

1. The equivalence between the category of affine schemes and the opposite of the category of commutative rings is a seminal result in scheme theory, a huge part of modern algebraic geometry.

2. The equivalence between Boolean lattices and Stone spaces is again seminal in the theory of Stone-type dualities. These can lead to deep connections between topology and logic. One application in particular is the study of the behavior of computer programs through formal semantics.

SOL Exercise 386. Show that equivalence of categories is an equivalence relation.

SOL Exercise 387. Show that $\mathbf{C} \simeq \mathbf{C}'$ and $\mathbf{D} \simeq \mathbf{D}'$ implies $[\mathbf{C}, \mathbf{D}] \simeq [\mathbf{C}', \mathbf{D}']$.

6 Yoneda Lemma

6.1 Representable Functors 121
6.2 Yoneda Lemma 126
6.3 Universality as Representability

We first defined an element of an object $X \in \mathbf{C}_0$ to be a morphism $\mathbf{1} \to X$. Our inspiration came from **Set** where $\mathrm{Hom}_{\mathbf{Set}}(\mathbf{1}, X) \cong X$. This is not a perfect categorification of the notion of element, because it works in some categories (e.g. **Poset**, **Top**, **Met**), but not in others (e.g. **Grp**, **Cat**[410], categories with no terminal object). In Exercise 331, we found a workaround for **Grp**, namely, elements of G correspond to morphisms $\mathbb{Z} \to G$.

[410] In **Cat**, a morphism $\mathbf{1} \to \mathbf{C}$ corresponds to an object of $\mathbf{C}_0$, but depending on the context, it may be more relevant to define an element of $\mathbf{C}$ to be a morphism of $\mathbf{C}_1$.

Armed with our new abstract tools from last chapter, in particular natural isomorphisms, we can rigorously explain why $\mathbf{1}$ seems to *represent* the choice of an element in **Set**, why $\mathbb{Z}$ plays that role in **Grp**, and go further to find other things that are *representable*.

This journey quickly leads to the Yoneda lemma which formalizes our conviction[411] that studying mathematical objects through their interactions with other objects is "enough".

[411] Hopefully, you have been convinced by earlier chapters.

6.1 Representable Functors

Throughout this chapter, let $\mathbf{C}$ be a locally small category. Recall that for an object $A \in \mathbf{C}_0$, there are two Hom functors from $\mathbf{C}$ to **Set**. The covariant one, $\mathrm{Hom}_{\mathbf{C}}(A, -)$, sends an object $B \in \mathbf{C}_0$ to $\mathrm{Hom}_{\mathbf{C}}(A, B)$ and a morphism $f : B \to B'$ to $f \circ (-)$. The contravariant one, $\mathrm{Hom}_{\mathbf{C}}(-, A)$, sends an object $B \in \mathbf{C}_0$ to $\mathrm{Hom}_{\mathbf{C}}(B, A)$ and a morphism $f : B \to B'$ to $(-) \circ f$. In order to lighten the notation, we denote these functors $H^A : \mathbf{C} \rightsquigarrow \mathbf{Set}$ and $H_A : \mathbf{C}^{\mathrm{op}} \rightsquigarrow \mathbf{Set}$ respectively.[412]

[412] It is somewhat standard to use sub- and superscript as an indication for the *variance* of a notation. Note however that, while H^A is covariant and H_A contravariant, we are not talking about this. Instead we are interested in their *variance* in the parameter A, and we will, given a morphism $f : A \to A'$, construct a natural transformation $H^{A'} \Rightarrow H^A$ which means H^A is contravariant in A, and similarly, H_A is covariant in A.

Although these functors are sometimes interesting on their own, their full power is unleashed when they are related to other functors through natural transformations. Before doing that, let us investigate how nice Hom functors are. For instance, many Hom functors can be described in simpler terms.

Examples 388. We are just revisiting things we already know.

1. Let $\mathbf{1} = \{*\}$ be the terminal object in **Set**, then what is the action of $H^{\mathbf{1}}$? For any object B,
$$H^{\mathbf{1}}(B) = \mathrm{Hom}_{\mathbf{Set}}(\mathbf{1}, B)$$
is easy to describe because for any element $b \in B$, there is a unique function $f : \mathbf{1} \to B = * \mapsto b$. Hence, there is an isomorphism from $H^{\mathbf{1}}(B)$ to B for any $B \in \mathbf{Set}_0$, it sends f to $f(*)$ and its inverse sends $b \in B$ to the map $* \mapsto b$.

Moreover, these isomorphisms are natural in B because (159) clearly commutes for any $f : B \to B'$, yielding a natural isomorphism $H^{\mathbf{1}} \cong \mathrm{id}_{\mathbf{Set}}$.

$$\begin{array}{ccc} H^{\mathbf{1}}(B) & \xrightarrow{f\circ(-)} & H^{\mathbf{1}}(B') \\ \updownarrow & & \updownarrow \\ B & \xrightarrow[f]{} & B' \end{array} \quad (159)$$

2. Consider again the terminal object but in the category **Grp**, namely, the group **1** only containing an identity element. Then, for any group G, the set $H^{\mathbf{1}}(G)$ is a singleton because any homomorphism $f : \mathbf{1} \to G$ must send the identity to the identity and no other choice can be made. Therefore, unlike in **Set**, $H^{\mathbf{1}}$ is very uninteresting and acts like the constant functor $\Delta(\mathbf{1}) : \mathbf{Grp} \rightsquigarrow \mathbf{Set}$, i.e. $H^{\mathbf{1}} \cong \Delta(\mathbf{1})$.

3. A better choice of object to mimic the behavior of $\mathrm{id}_{\mathbf{Grp}}$ is the additive group $\mathbb{Z}$. Indeed, for any $g \in G$, there is a unique homomorphism $f : \mathbb{Z} \to G$ sending 1 to g.[413] A very similar argument as above yields a natural isomorphism $H^{\mathbb{Z}} \cong \mathrm{id}_{\mathbf{Grp}}$.

4. The terminal object in **Cat** is the category **1** with a single object $\bullet$ and no morphism other than the identity. For any category $\mathbf{C} \in \mathbf{Cat}_0$, a functor $\mathbf{1} \rightsquigarrow \mathbf{C}$ is just a choice of object. Therefore, the same argument will show that $H^{\mathbf{1}} \cong (-)_0$, where $(-)_0$ sends a category to its set[414] of objects and a functor to its action restricted on objects.

   In order to obtain a similar way to extract morphisms, consider the category **2** with two objects and a single morphism between them. One obtains a natural isomorphism $H^{\mathbf{2}} \cong (-)_1$.[415]

[413] Note that f is completely determined by $f(1)$ because the homomorphism properties imply that $f(n) = f(1) + \overset{n}{\cdots} + f(1)$, $f(-n) = f(n)^{-1}$, and $f(0)$ must be the identity.

[414] Recall that **Cat** only contains small categories.

[415] You can prove this as we did for $H^{\mathbf{1}} \cong (-)_0$ or use Example 351.3.

Just like we benefitted from recognizing a category was isomorphic to a functor category (e.g. Theorem 353 and Corollary 354), we can benefit from finding a natural isomorphism between a functor and a Hom functor. For instance, we already know that the Hom functors are continuous,[416] and with Example 388.4 we can infer

$$\left(\prod_{i\in I} \mathbf{C}_i\right)_0 = \prod_{i\in I} (\mathbf{C}_i)_0 \text{ and } \left(\prod_{i\in I} \mathbf{C}_i\right)_1 = \prod_{i\in I} (\mathbf{C}_i)_1.$$

In words, the objects of a product of categories are tuples of objects of each category and similarly for morphisms.[417] This suggest carefully studying representable functors.

[416] Theorem 282 and Corollary 283. Also recall that a functor naturally isomorphic to a continuous functor is also continuous, see Exercise 349.

[417] We already knew that for the case of binary products, see Exercise 207.

Definition 389 (Representable functor)**.** A covariant functor $F : \mathbf{C} \rightsquigarrow \mathbf{Set}$ is **representable** if there is an object $X \in \mathbf{C}_0$ such that F is naturally isomorphic to $\mathrm{Hom}_{\mathbf{C}}(X, -)$. If F is contravariant, then it is representable if it is naturally isomorphic to $\mathrm{Hom}_{\mathbf{C}}(-, X)$.

Examples 390. Let us give examples of the contravariant kind.

1. Recall from Example 139.2 the contravariant powerset functor $2^{-} : \mathbf{Set} \rightsquigarrow \mathbf{Set}$. It sends a set X to its powerset $2^X = \mathcal{P}(X)$ and a function $f : X \to Y$ to the inverse image $2^f = f^{-1} : \mathcal{P}(Y) \to \mathcal{P}(X)$. We can identify subsets of a given set with functions from this set into $\Omega = \{\bot, \top\}$.[418] This yields a bijection $2^X \cong H_\Omega(X)$ that is natural in X. Indeed, for all $f : X \to Y$, you can check (160) commutes,[419]

[418] See our discussion of subobject classifers in **Set**.

[419] Starting with $p : Y \to \Omega$ in the bottom left. The top path yields

$$(p \circ f)^{-1}(\top) = \{x \in X \mid p(f(x)) = \top\}.$$

The bottom path yields

$$f^{-1}(p^{-1}(\top)) = \{x \in X \mid p(f(x)) = \top\}.$$

so $2^- \cong H_\Omega$.

$$\begin{array}{ccc} H_\Omega(X) & \underset{p \mapsto p^{-1}(\top)}{\overset{S \mapsto \chi_S}{\rightleftarrows}} & 2^X \\ {\scriptstyle -\circ f}\uparrow & & \uparrow{\scriptstyle 2^f = f^{-1}} \\ H_\Omega(Y) & \underset{p \mapsto p^{-1}(\top)}{\overset{S \mapsto \chi_S}{\rightleftarrows}} & 2^X \end{array} \tag{160}$$

2. Our first example of natural isomorphism (Example 345.1) was the currying of a morphism $\lambda : \mathrm{Hom}_{\mathbf{C}}(- \times X, A) \cong \mathrm{Hom}_{\mathbf{C}}(-, A^X)$, where A^X is an exponential object. It turns out exponential objects can be defined via this natural isomorphism. Namely, there is an isomorphism $\ell : \mathrm{Hom}_{\mathbf{C}}(- \times X, A) \cong \mathrm{Hom}_{\mathbf{C}}(-, E)$ if and only if E is the exponential object and $\ell_E^{-1}(\mathrm{id}_E)$ is the evaluation morphism.[420]

   ($\Leftarrow$) This was already shown in Example 345.1 modulo the fact that $\lambda^{-1}\mathrm{id}_{A^X} = \mathsf{ev}$. For the latter, it suffices to note that $\lambda\mathsf{ev}$ must be id_{A^X} to make (161) commute.

$$\begin{array}{ccc} A & \xleftarrow{\mathsf{ev}} & A^X \times X \\ & {\scriptstyle \mathsf{ev}}\nwarrow & \uparrow{\scriptstyle \lambda\mathsf{ev} \times \mathrm{id}_X} \\ & & A^X \times X \end{array} \tag{161}$$

   ($\Rightarrow$) Given ℓ, we show E is the exponential. For any $g : B \times X \to A$, we claim that $\ell_B(g)$ makes (162) commute. The naturality of ℓ^{-1} yields the following commutative square.

$$\begin{array}{ccc} \mathrm{Hom}_{\mathbf{C}}(B \times X, A) & \xleftarrow{\ell_B^{-1}} & \mathrm{Hom}_{\mathbf{C}}(B, E) \\ {\scriptstyle -\circ(\ell_B(g) \times \mathrm{id}_X)}\uparrow & & \uparrow{\scriptstyle -\circ \ell_B(g)} \\ \mathrm{Hom}_{\mathbf{C}}(E \times X, A) & \xleftarrow[\ell_E^{-1}]{} & \mathrm{Hom}_{\mathbf{C}}(E, E) \end{array} \tag{163}$$

   Starting in the bottom right with id_E, the bottom path sends it to $\ell_E^{-1}(\mathrm{id}_E) \circ (\ell_B(g) \times \mathrm{id}_X)$ and the top path sends to $\ell_B^{-1}(\ell_B(g)) = g$. Commutativity lets us conclude $\ell_E^{-1}(\mathrm{id}_E) \circ (\ell_B(g) \times \mathrm{id}_X) = g$, i.e. (162) commutes.

[420] The expression $\ell_E^{-1}(\mathrm{id}_E)$ might look like it comes out of nowhere, but it is not so mysterious. Given the natural isomorphism ℓ, if we are looking for a moprhism of type $E \times X \to A$, then we may as well look for a morphism of type $E \to E$ and use the bijection ℓ_E^{-1}. What morphism of type $E \to E$ do we have? Only one is guaranteed to exist, the identity id_E. This chapter contains several instances of this kind of forced choice.

$$\begin{array}{ccc} A & \xleftarrow{\ell_E^{-1}(\mathrm{id}_E)} & E \times X \\ & {\scriptstyle g}\nwarrow & \uparrow{\scriptstyle \ell_B(g) \times \mathrm{id}_X} \\ & & B \times X \end{array} \tag{162}$$

In the first items of Examples 388 and 390, we made an arbitrary choice of set. That is, we could have taken any singleton instead of $\mathbf{1}$ in the first case and any set with two elements instead of Ω in the second. More generally, one can show that if $A \cong B$, then $H_A \cong H_B$ and $H^A \cong H^B$.

SOL **Exercise 391** (NOW!). Let $A, B \in \mathbf{C}_0$ be isomorphic objects. Show that $H^A \cong H^B$. Dually, show that $H_A \cong H_B$.

In particular, for any object E isomorphic to the exponential A^X, we have

$$H_E \cong H_{A^X} \cong \mathrm{Hom}_{\mathbf{C}}(- \times X, A),$$

which means E is also the exponential. In Exercise 300, we also showed that if E satisfies the same universal property as A^X, then they must be isomorphic. In order to prove this using the natural isomorphism instead of the universal property, we would need a converse to Exercise 391.

Perhaps surprisingly, it is true and it will follow from the Yoneda lemma, but we prove it on its own first as a warm-up for the proof of the lemma.

Proposition 392. *Let $A, B \in \mathbf{C}_0$ be such that $H^A \cong H^B$, then $A \cong B$.*

Proof. The natural isomorphism gives two natural transformations $\phi : H^A \Rightarrow H^B$ and $\eta : H^B \Rightarrow H^A$ such that for any object $X \in \mathbf{C}_0$,

$$\eta_X \circ \phi_X : H^A(X) \to H^A(X) \quad \text{and} \quad \phi_X \circ \eta_X : H^B(X) \to H^B(X)$$

are identities. In order to show $A \cong B$, we will find two morphisms $f : B \to A$ and $g : A \to B$ such that $f \circ g = \mathrm{id}_A$ and $g \circ f = \mathrm{id}_B$. With the given data, there is no freedom to construct f and g. Since $\mathbf{C}$, A and B are arbitrary, there are only two morphisms that are required to exist, id_A and id_B. Next, we note that $\mathrm{id}_A \in H^A(A)$ and $\mathrm{id}_B \in H^B(B)$, hence, we can set $f := \phi_A(\mathrm{id}_A)$ and $g := \eta_B(\mathrm{id}_B)$.[421]

Now, $\phi_A(\mathrm{id}_A)$ is a morphism from B to A, so (164) commutes by naturality of η.

$$\begin{array}{ccc} H_B(A) & \xrightarrow{\eta_A} & H_A(A) \\ {\scriptstyle \phi_A(\mathrm{id}_A)\circ(-)}\uparrow & & \uparrow{\scriptstyle \phi_A(\mathrm{id}_A)\circ(-)} \\ H_B(B) & \xrightarrow[\eta_B]{} & H_A(B) \end{array} \tag{164}$$

We conclude, by starting with id_B in the bottom left, that

$$g \circ f = \phi_A(\mathrm{id}_A) \circ \eta_B(\mathrm{id}_B) = \eta_A(\phi_A(\mathrm{id}_A)) = \mathrm{id}_A.$$

A dual argument shows that

$$f \circ g = \eta_B(\mathrm{id}_B) \circ \phi_A(\mathrm{id}_A) = \phi_B(\eta_B(\mathrm{id}_B)) = \mathrm{id}_B,$$

and we have shown $A \cong B$. $\square$

Corollary 393 (Dual)**.** *Let $A, B \in \mathbf{C}_0$ be such that $H_A \cong H_B$, then $A \cong B$.*

Steve Awodey calls **Yoneda principle** the equivalences[422]

$$H^A \cong H^B \Leftrightarrow A \cong B \Leftrightarrow H_A \cong H_B.$$

This is the formalization of the philosophical point we mentionned a few times already: an object is determined up to isomorphism by all its relations with all other objects. The Hom functor H^A (or H_A) makes for an efficient description of all the relations between A and all other objects.

Let us give two more concrete examples of representable functors.

Example 394 (G acting on itself)**.** Any group G acts on itself by multiplication on the left. The corresponding functor, abusively denoted by $G : \mathbf{B}G \rightsquigarrow \mathbf{Set}$, sends $*$ to the set G and $g \in G$ to the bijection $h \mapsto gh$.[423] Fix another group action

[421] To emphasize the point about *no freedom*, try to convince yourself that any morphisms of type $B \to A$ and $A \to B$ that we can construct from id_A, id_B, ϕ and η (the only data we have) must be equal to f and g as we defined them.

[422] They are the combination of Exercise 391, Proposition 392 and Corollary 393.

[423] Its inverse is $h \mapsto g^{-1}h$.

$F \in [\mathbf{B}G, \mathbf{Set}]$, we showed a natural transformation $f : G \Rightarrow F$ is a G–equivariant map, it makes (165) commute for every $g \in G$.

$$\begin{array}{ccc} G & \xrightarrow{f_*} & F(*) \\ {\scriptstyle g-}\downarrow & & \downarrow{\scriptstyle g\star -} \\ G & \xrightarrow[f_*]{} & F(*) \end{array} \tag{165}$$

Starting with 1_G on the top left, we find that $f_*(g) = g \star f_*(1_G)$. Thus, the equivariant map f_* is completely determined by where it sends 1_G. Since there is no constraint on that choice, we get a bijection between natural transformations $G \Rightarrow F$ and elements of $F(*)$.

The assignment $F \mapsto F(*)$ is functorial as we have seen when defining Ev, and you can also see it as the forgetful functor $U : [\mathbf{B}G, \mathbf{Set}] \rightsquigarrow \mathbf{Set}$ that forgets about the action of G. Thus, we can ask whether the bijection above is natural in F, i.e. does (166) commute for every $h : F \Rightarrow F'$? It does commute as both paths send f to $\phi_*(f_*(1_G))$, hence we find that U is representable with $U \cong H^G$.

$$\begin{array}{ccc} \mathrm{Hom}_{[\mathbf{B}G,\mathbf{Set}]}(G,F) & \xrightarrow{f \mapsto f_*(1_G)} & F(*) \\ {\scriptstyle \phi\cdot -}\downarrow & & \downarrow{\scriptstyle \phi_*} \\ \mathrm{Hom}_{[\mathbf{B}G,\mathbf{Set}]}(G,F') & \xrightarrow[f \mapsto f_*(1_G)]{} & F'(*) \end{array} \tag{166}$$

Example 395 (Elements of a ring)**.** In **Ring** just like in **Grp**, the terminal object is the ring containing only one element that is the zero and identity at the same time. Thus, there can be no morphism $\mathbf{1} \to R$ unless $R = \mathbf{1}$.[424] We leave you to show $H^{\mathbf{1}}$ is naturally isomorphic to the constant functor $\Delta(\varnothing)$.

[424] A ring homomorphism must send 0 to 0 and 1 to 1, so if $0 = 1$ in the source then 0 must equal 1 in the target as well.

Let us try what we did for **Grp**: replace $\mathbf{1}$ with $\mathbb{Z}$. Unfortunately, a ring homomoprhism $f : \mathbb{Z} \to R$ is too constrained. We must have $f(0) = 0_R$ and $f(1) = 1_R$, and any other value is forced by the homomorphism properties:

$$f(n) = f(1) + \overset{n}{\cdots} + f(1) = 1_R + \overset{n}{\cdots} + 1_R \text{ and } f(-n) = -f(n).$$

This means $\mathbb{Z}$ is the initial ring, and we can prove $H^{\mathbb{Z}}$ is naturally isomorphic to the constant functor $\Delta(\mathbf{1})$ (see Exercise 405).

We need to add one element, say x, to $\mathbb{Z}$ so that f can map x anywhere, but no other choice can be made.[425] For the "map x anywhere" part, we must make sure that x is free of any constraint other than the properties of a ring. That is, it has an additive inverse $-x$, it satifies $x + x = (1x + 1x) = (1+1)x = 2x$ and other similar equations, it has powers like $x^2 = x \cdot x$ and $x^3 = x \cdot x \cdot x$, there are combinations like $5 + 2x + 4x^5$, and so on. The "no other choice" part is a consequence of the homomorphism properties. If the image of x is known, then the images of all the multiples and powers of x and combinations of them and other elements of $\mathbb{Z}$ are known too.

[425] This is essentially what we have done to go from $\mathbf{1}$ to $\mathbb{Z}$ in **Grp**. The integers can be seen as the group $\mathbf{1} = \{0\}$ where we add 1 (it is not the identity), its inverse -1 and letting the group operation do its thing. For instance $2 = 1 + 1$, $3 = 1 + 1 + 1$, etc.

In short, we are talking about the ring $\mathbb{Z}[x]$ of polynomials with one variable and coefficients in $\mathbb{Z}$. A ring homomorphism $\mathbb{Z}[x] \to R$ is completely determined by where it sends x, and we leave you to show $H^{\mathbb{Z}[x]}$ is naturally isomorphic to the forgetful functor **Ring** $\rightsquigarrow$ **Set**.[426]

[426] With the Yoneda principle, we now have the promised categorical definition of polynomials from Example 252.3. Exercise 396 generalizes this to multivariate polynomials with noninteger coefficients.

With a slight modification, we can show the units functor $(-)^\times :$ **Ring** $\rightsquigarrow$ **Set** (we also forget about the group structure) is representable. The ring $\mathbb{Z}[x, x^{-1}]$ is $\mathbb{Z}[x]$ where we add a multiplicative inverse to x. It satisfies all the expected equations (e.g. $x \cdot x^{-1} = 1$, $x^2 \cdot x^{-3} = x^{-1}$, etc.) and no other. A ring homomoprhism $f : \mathbb{Z}[x, x^{-1}] \to R$ must send x^{-1} to the inverse of $f(x)$. Therefore, $f(x)$ is now restricted to $R^\times$. We leave you to show $H^{\mathbb{Z}[x,x^{-1}]} \cong (-)^\times$.

SOL **Exercise 396.** Let $U :$ **Ring** $\rightsquigarrow$ **Set** be the forgetful functor and, for any $n \in \mathbb{N}$, $(-)^n :$ **Ring** $\rightsquigarrow$ **Ring** the n–wise product functor.

1. Show that $H^{\mathbb{Z}[x_1,\ldots,x_n]}$ is naturally isomorphic to the composition $U \circ (-)^n$.
2. For any ring R, show that $H^{R[x]} \cong H^R \times U$.[427]
3. Make up a categorical definition of $R[x_1, \ldots, x_n]$ using this characterization. Does item 1 make you more confident in your definition?

[427] For this to typecheck, the R.H.S. must be the product inside $[\mathbf{Ring}, \mathbf{Set}]$, i.e. $(H^R \times U)(S) = \mathrm{Hom}(R, S) \times S$.

6.2 Yoneda Lemma

Taking a closer look at our solution to Exercise 391, we find the assignments $A \mapsto H^A$ and $A \mapsto H_A$ are functorial.

Definition 397 (Yoneda embeddings)**.** The contravariant **Yoneda embedding**[428] $H^{(-)} : \mathbf{C}^{\mathrm{op}} \rightsquigarrow [\mathbf{C}, \mathbf{Set}]$ sends $A \in \mathbf{C}_0$ to the Hom functor H^A and a morphism $f : A' \to A$ to the natural transformation $H^f : H^A \Rightarrow H^{A'}$ defined by $H^f_B = \mathrm{Hom}_{\mathbf{C}}(f, B) = (-) \circ f$ for every $B \in \mathbf{C}_0$. The naturality of H^f follows from associativity[429]: for any $g : B \to B'$, (167) commutes.

$$\begin{array}{ccc} H^A(B) & \xrightarrow{(-)\circ f} & H^{A'}(B) \\ {\scriptstyle g\circ(-)}\downarrow & & \downarrow{\scriptstyle g\circ(-)} \\ H^A(B') & \xrightarrow[(-)\circ f]{} & H^{A'}(B') \end{array} \tag{167}$$

[428] Yoneda embeddings and the Yoneda lemma are named in honor of Nobuo Yoneda.

[429] Starting with h in the top left. The top path sends it to $g \circ (h \circ f)$ and the bottom path sends it to $(g \circ h) \circ f$. Since composition is associative, both paths are the same function.

The covariant embedding $H_{(-)} : \mathbf{C} \rightsquigarrow [\mathbf{C}^{\mathrm{op}}, \mathbf{Set}]$ sends $B \in \mathbf{C}_0$ to the Hom functor H_B and a morphism $f : B \to B'$ to the natural transformation $H_f : H_B \Rightarrow H_{B'}$ defined by $(H_f)_A = \mathrm{Hom}_{\mathbf{C}}(A, f) = f \circ (-)$ for any $A \in \mathbf{C}_0$.[430] In order to harmonize the notation, we write H^A_f instead of $(H_f)_A$. Now the subscript of H always goes in the target of the Hom, and the superscript alawys goes in the source.

[430] Naturality follows from associativity of composition again.

Another way to obtain these embeddings (incidentally proving they are functors) is to curry the Hom bifunctor. Indeed, you can verify that

$$H^- = \Lambda\mathrm{Hom}(-,-) \text{ and } H_- = \Lambda(\mathrm{Hom}(-,-) \circ \mathsf{swap}).$$

The embeddings are called like that (c.f. Exercise 157) because both functors are injective on objects[431] and fully faithful as will follow from the Yoneda lemma.

[431] If $A \neq B$, then $H^A(A)$ contains id_A but $H^B(A)$ does not, so $H^A \neq H^B$.

We now understand how an object $A \in \mathbf{C}_0$ can be understood by studying the representable H^A. In some sense, H^A tells us how A views the category it is in. Since the representable H^A is an object of the category $[\mathbf{C}, \mathbf{Set}]$, it is daring to try and understand it via the representable H^{H^A}. In other words, how does H^A see other functors in $[\mathbf{C}, \mathbf{Set}]$.

We have already got a problem. Even if $\mathbf{C}$ is locally small, there is no guarantee that $[\mathbf{C}, \mathbf{Set}]$ is locally small. Thus, $H^{H^A} = \mathrm{Hom}_{[\mathbf{C},\mathbf{Set}]}(H^A, -)$ might no be a well-defined functor.[432] To avoid confusing or cluttered notation, we write instead $\mathrm{Nat}(H^A, -)$ because, for a functor $F : \mathbf{C} \rightsquigarrow \mathbf{Set}$, $\mathrm{Nat}(H^A, F)$ is the collection of natural transformations from H^A to F.

[432] We do not know what category it lands in.

We already saw that for every morphism $f : B \to A$ in $\mathbf{C}$, there is an element $H^f \in \mathrm{Nat}(H^A, H^B)$. Does every natural transformation of this type arise like that?

Given a natural transformation $\alpha : H^A \Rightarrow H^B$ constructed from an unknown morphism $B \to A$, we can figure out what is that morphism by looking at $\alpha_A(\mathrm{id}_A)$.[433] Indeed, if $\alpha = H^f$ for some given $f : B \to A$, then

$$\alpha_A(\mathrm{id}_A) = H^f_A(\mathrm{id}_A) = \mathrm{id}_A \circ f = f.$$

Even if we do not know such an f, $\alpha_A(\mathrm{id}_A)$ is still a morphism $B \to A$. It turns out we can exploit naturality to show α must be the natural transformation $H^{\alpha_A(\mathrm{id}_A)}$.

What can we say when the target of α is not representable? i.e. $\alpha : H^A \Rightarrow F$ for some functor $F : \mathbf{C} \rightsquigarrow \mathbf{Set}$. Our trick from above tells us every such α yields an element $\alpha_A(\mathrm{id}_A) \in F(A)$. Again relying on naturality, we can show every element $a \in F(A)$ gives a transformation $\alpha : H^A \Rightarrow F$ satisfying $\alpha_A(\mathrm{id}_A) = a$.

In short, the surprising relation described by the Yoneda lemma is an isomorphism between $\mathrm{Nat}(H^A, F)$ and $F(A)$ that is natural in F and A. We first show the isomorphism and then the naturality.

[433] Once again, this choice is forced on us by the data we have. We are only given α and we need to find an element of $\mathrm{Hom}(B, A)$. It turns out α_A has type $\mathrm{Hom}(A, A) \to \mathrm{Hom}(B, A)$, so it remains to find an element of $\mathrm{Hom}(A, A)$. Since we know nothing else about $\mathbf{C}$, we can only pick id_A, because $\mathrm{Hom}(A, A)$ might contain no other morphism.

Lemma 398 (Yoneda lemma I). *For any $A \in \mathbf{C}_0$ and $F : \mathbf{C} \rightsquigarrow \mathbf{Set}$,*

$$\mathrm{Nat}(H^A, F) \cong F(A).$$

Proof. Let $\phi_{A,F} : \mathrm{Nat}(H^A, F) \to F(A)$ be defined by $\alpha \mapsto \alpha_A(\mathrm{id}_A)$.[434] In the opposite direction, let $\eta_{A,F} : F(A) \to \mathrm{Nat}(H^A, F)$ send an element $a \in F(A)$ to the natural transformation with components $(\eta_{A,F}(a))_B : \mathrm{Hom}_{\mathbf{C}}(A, B) \to F(B) = f \mapsto F(f)(a)$ for each $B \in \mathbf{C}_0$.[435] Checking (168) commutes for any $g : B \to B'$ shows that $\eta_{A,F}(a)$ is a natural transformation. Starting with f in the top left, the top path sends it to $F(g)(F(f)(a))$ and the bottom path sends it to $F(g \circ f)(a)$. These two are equal by functoriality, i.e. $F(g) \circ F(f) = F(g \circ f)$.

[434] As we said earlier, this is the only way to obtain an element of $F(A)$ from the given data.

[435] Again this definition is the only one that typechecks. With a functor F, an element of $F(A)$, and a morphism in $\mathrm{Hom}_{\mathbf{C}}(A, B)$, we can apply $F(f) : F(A) \to F(B)$ to get an element of $F(B)$.

$$\begin{array}{ccc} H^A(B) & \xrightarrow{F(-)(a)} & F(B) \\ {\scriptstyle g\circ(-)}\downarrow & & \downarrow{\scriptstyle F(g)} \\ H^A(B') & \xrightarrow[F(-)(a)]{} & F(B') \end{array} \tag{168}$$

We now check that $\phi_{A,F}$ and $\eta_{A,F}$ are inverses. First, $(\eta \circ \phi)_{A,F}$ sends $\alpha \in \mathrm{Nat}(H^A, F)$ to $\eta_{A,F}(\alpha_A(\mathrm{id}_A))$, and at any $B \in \mathbf{C}_0$, we have

$$\begin{aligned} (\eta_{A,F}(\alpha_A(\mathrm{id}_A)))_B(f) &= F(f)(\alpha_A(\mathrm{id}_A)) && \text{def of } \eta \\ &= \alpha_B(H^A(f)(\mathrm{id}_A)) && \mathsf{NAT}(\alpha, A, B, f) \\ &= \alpha_B(f \circ \mathrm{id}_A) && \text{def of } H^A \\ &= \alpha_B(f), \end{aligned}$$

thus $\alpha = (\eta \circ \phi)_{A,F}(\alpha)$.

Conversely, $(\phi \circ \eta)_{A,F}$ sends $a \in F(A)$ to $\eta_{A,F}(a)_A(\mathrm{id}_A) = F(\mathrm{id}_A)(a) = a$, and we can conclude that $\eta_{A,F}$ and $\phi_{A,F}$ are natural isomorphisms. □

Corollary 399 (Dual)**.** *For any $A \in \mathbf{C}_0$ and $F : \mathbf{C}^{\mathrm{op}} \rightsquigarrow \mathbf{Set}$, $\mathrm{Nat}(H_A, F) \cong F(A)$.*

We already mentionned a consequence of this result.

Corollary 400. *The Yoneda embeddings* $H^{(-)}$ *and* $H_{(-)}$ *are fully faithful.*[436]

436 Recall from Exercises 187 and 188 that when a functor F is fully faithful, $A \cong B$ if and only if $FA \cong FB$. Thus, Exercise 391, Proposition 392 and Corollary 393 are all corollaries of this.

Proof. Applying the lemma with $F = H^B$, we find an isomorphism

$$\mathrm{Nat}(H^A, H^B) \cong H^B(A) = \mathrm{Hom}_{\mathbf{C}}(B, A)$$

In the right to left direction, this isomorphism sends $f : B \to A$ to $H^f : H^A \Rightarrow H^B$.[437] This is the action of the functor $H^{(-)}$ on the homset $\mathrm{Hom}_{\mathbf{C}}(B, A)$. Therefore, for all $A, B \in \mathbf{C}_0$, $f \mapsto H^f$ is a bijection, which means $H^{(-)}$ is fully faithful.

The dual argument shows $H_{(-)}$ is fully faithful. □

437 By unrolling the definition of $\eta_{A,H^B}(f)$, we find its component at $A' \in \mathbf{C}_0$ sends $h \in \mathrm{Hom}_{\mathbf{C}}(A, A')$ to $h \circ f \in \mathrm{Hom}_{\mathbf{C}}(B, A')$. So $\eta_{A,H^B}(f) = H^f$.

Another consequence is that $\mathrm{Nat}(H^A, F)$ is a set (because it is isomorphic to $F(A)$ which is a set), and this allows us to formally state the second part of the Yoneda lemma.[438]

438 That $\phi_{A,F}$ and $\eta_{A,F}$ are natural in A and F.

The assignment $(A, F) \mapsto \mathrm{Nat}(H^A, F)$ is a functor $\mathbf{C} \times [\mathbf{C}, \mathbf{Set}] \rightsquigarrow \mathbf{Set}$ with the action on morphisms given by[439]

$$(g, \mu) : (A, F) \to (A', F') \mapsto \mu \cdot (-) \cdot H^g : \mathrm{Nat}(H^A, F) \to \mathrm{Nat}(H^{A'}, F').$$

439 If $g : A \to A'$, $\mu : F \Rightarrow F'$, and $\eta \in \mathrm{Nat}(H^A, F)$, we have the composite

$$H^{A'} \xRightarrow{H^g} H^A \xRightarrow{\eta} F \xRightarrow{\mu} F' \in \mathrm{Nat}(H^{A'}, F').$$

We can check this preserves identities and composition. The identity morphism on (A, F) is $(\mathrm{id}_A, \mathbb{1}_F)$, and it is sent to $\mathbb{1}_F \cdot (-) \cdot H^{\mathrm{id}_A}$, that is pre- and post-composition by the identities.[440] Given two morphisms $(g, \mu) : (A, F) \to (A', F')$ and $(g', \mu') : (A', F') \to (A'', F'')$, associativity of vertical composition implies

$$(\mu' \cdot (-) \cdot H^{g'}) \circ (\mu \cdot (-) \cdot H^g) = (\mu' \cdot \mu) \cdot (-) \cdot (H^g \cdot H^{g'}) = (\mu' \cdot \mu) \cdot (-) \cdot H^{g' \circ g}.$$

440 It follows from functoriality of $H^{(-)}$ that $H^{\mathrm{id}_A} = \mathbb{1}_{H^A}$.

The type of $\mathrm{Nat}(H^-, -)$ can be confusing. Just for a moment, think of $\mathrm{Nat}(-, -)$ as a Hom bifunctor.[441] Then, instead of seeing H^- as a functor $\mathbf{C}^{\mathrm{op}} \rightsquigarrow [\mathbf{C}, \mathbf{Set}]$, see it instead as $\mathbf{C} \rightsquigarrow [\mathbf{C}, \mathbf{Set}]^{\mathrm{op}}$. Then, $\mathrm{Nat}(H^-, -)$ is the composite

$$\mathbf{C} \times [\mathbf{C}, \mathbf{Set}] \xrightarrow{H^- \times \mathrm{id}} [\mathbf{C}, \mathbf{Set}]^{\mathrm{op}} \times [\mathbf{C}, \mathbf{Set}] \xrightarrow{\mathrm{Nat}(-,-)} \mathbf{Set}\ .$$

441 Strictly speaking $[\mathbf{C}, \mathbf{Set}]$ might not be locally small, so the functor $\mathrm{Nat}(-, -)$ is not well-defined.

The assignment $(A, F) \mapsto F(A)$ is another functor of the same type. We denoted it by Ev,[442] its action on morphisms is defined by

$$(g, \mu) : (A, F) \to (A', F') \mapsto F'(g) \circ \mu_A = \mu_{A'} \circ F(g) : F(A) \to F'(A').$$

442 See Example 373.5.

Lemma 401 (Yoneda lemma II)**.** *There is a natural isomorphism* $\mathrm{Nat}(H^-, -) \cong \mathsf{Ev}$.

Proof. The components of this isomorphism are the ones described in the first part. It remains to show that ϕ is natural in (A, F).[443] For any $(g, \mu) : (A, F) \to (A', F')$, we need to show the following square commutes.

443 By Exercise 344, it is enough to show it is natural in A and natural in F separately. We do both at the same time because it is not much harder.

$$\begin{array}{ccc} \mathrm{Nat}(H^A, F) & \xrightarrow{\phi_{A,F}} & F(A) \\ {\scriptstyle \mu \cdot (-) \cdot H^g} \downarrow & & \downarrow {\scriptstyle F'(g) \circ \mu_A} \\ \mathrm{Nat}(H^{A'}, F') & \xrightarrow[\phi_{A',F'}]{} & F'(A') \end{array} \tag{169}$$

Starting with a natural transformation $\alpha \in \mathrm{Nat}(H^A, F)$, the bottom path sends it to $(\mu \cdot \alpha \cdot H^g)_{A'}(\mathrm{id}_{A'})$ and the top path sends it to $(F'(g) \circ \mu_A)(\alpha_A(\mathrm{id}_A))$. The following derivation shows they are equal.

$$\begin{aligned}
(\mu \cdot \alpha \cdot H^g)_{A'}(\mathrm{id}_{A'}) &= (\mu_{A'} \circ \alpha_{A'})(H^g_{A'}(\mathrm{id}_{A'})) && \text{def of } \cdot \\
&= (\mu_{A'} \circ \alpha_{A'})(g) && \text{def of } H^g_{A'} \\
&= (\mu_{A'} \circ \alpha_{A'})(H^A_g(\mathrm{id}_A)) && \text{def of } H^A_g \\
&= (\mu_{A'} \circ \alpha_{A'} \circ H^A_g)(\mathrm{id}_A) && \\
&= (\mu_{A'} \circ F(g) \circ \alpha_A)(\mathrm{id}_A) && \mathsf{NAT}(\alpha, A, A', g) \\
&= (F'(g) \circ \mu_A)(\alpha_A(\mathrm{id}_A)) && \mathsf{NAT}(\mu, A, A', g) \qquad \square
\end{aligned}$$

Corollary 402 (Dual)**.** *There is a natural isomorphism* $\mathrm{Nat}(H_-, -) \cong \mathsf{Ev}$.[444]

[444] We can typecheck this as before. We see H_- as a functor $\mathbf{C}^{\mathrm{op}} \rightsquigarrow [\mathbf{C}^{\mathrm{op}}, \mathbf{Set}]^{\mathrm{op}}$ (c.f. Exercise 146). Then $\mathrm{Nat}(H_-, -) = \mathrm{Nat}(-,-) \circ H_- \times \mathrm{id}$.

While the Yoneda lemma is called a lemma, it is extremely important and powerful. We already said how it gives category theorists reasons to study an object through its relations to other objects (via the Yoneda principle). In a shallow exploration of category theory, this might seem like the only point[445] of the Yoneda lemma.

[445] I find it already quite grandiose.

Another result with a similar status in mathematics — it looks motivated only by philosophical and meta considerations — is Cayley's theorem. It states that any group is isomorphic to the subgroup of a permutation group.[446] Remarkably, the Yoneda lemma can be understood as a generalization of Cayley's theorem. This is our first application of Yoneda.

[446] It is important to group theorists because they are interested in studying symmetries of geometric shapes or other things, and these can easily be seen as subgroups of permutation groups. Thus, the abstract notion of group is made more concrete by Cayley's theorem.

Example 403 (Cayley's theorem with the Yoneda lemma)**.** Recall the first part of the Yoneda lemma which states that for a category $\mathbf{C}$, a functor $F : \mathbf{C} \rightsquigarrow \mathbf{Set}$ and an object $A \in \mathbf{C}_0$, we have

$$\mathrm{Nat}(\mathrm{Hom}(A, -), F) \cong F(A).$$

Moreover, we know the explicit maps, namely, a natural transformation ϕ in the L.H.S. is mapped to $\phi_A(\mathrm{id}_A)$ and an element $a \in F(A)$ is mapped to the natural transformation whose component at $B \in \mathbf{C}_0$ is $\phi_B = f \mapsto F(f)(a)$.

Let us apply this to $\mathbf{C}$ being the delooping of a group G. Recall that any functor $F : \mathbf{B}G \rightsquigarrow \mathbf{Set}$ sends $*$ to a set S and any $g \in G$ to a permutation of S, it corresponds to an action of G on S.

To use the Yoneda lemma, our only choice of object for A is $*$ and we will choose for F the Hom functor $F = \mathrm{Hom}_{\mathbf{B}G}(*, -)$. The Yoneda lemma yields

$$\mathrm{Nat}(\mathrm{Hom}_{\mathbf{B}G}(*, -), \mathrm{Hom}_{\mathbf{B}G}(*, -)) \cong \mathrm{Hom}_{\mathbf{B}G}(*, *).$$

We already know that the R.H.S. is G,[447] but we have to do a bit of work to understand the L.H.S. First, observe that a natural transformation $\phi : \mathrm{Hom}_{\mathbf{B}G}(*, -) \Rightarrow \mathrm{Hom}_{\mathbf{B}G}(*, -)$ is just one morphism $\phi_* : \mathrm{Hom}_{\mathbf{B}G}(*, *) \to \mathrm{Hom}_{\mathbf{B}G}(*, *)$. Namely, it is a map from G to G. Second, recalling that $\mathrm{Hom}_{\mathbf{B}G}(*, g) = g \circ (-)$ and that $*$ is

[447] By definition of $\mathbf{B}G$.

the only object in $\mathbf{C}_0$, we get that ϕ_* must only make (170) commute.

$$\begin{array}{ccc} G & \xrightarrow{\phi_*} & G \\ {\scriptstyle g\circ(-)}\downarrow & & \downarrow{\scriptstyle g\circ(-)} \\ G & \xrightarrow[\phi_*]{} & G \end{array} \tag{170}$$

This is equivalent to $\phi_*(g \cdot h) = g \cdot \phi_*(h)$, and we get that each ϕ_* is a G–equivariant map from G to itself.[448] Denote the set of such maps by $\mathrm{Hom}_G(G,G)$. We obtain that, as sets,

$$\mathrm{Hom}_G(G,G) \cong G.$$

Now, we can check that $\mathrm{Hom}_G(G,G)$ is a subgroup of Σ_G (the group of permutations of the set G) and that the bijection is in fact an group isomorphism. Cayley's theorem follows.

We have to show that id_G is in $\mathrm{Hom}_G(G,G)$, that maps in $\mathrm{Hom}_G(G,G)$ are bijective, and that they are stable under composition and taking inverses. First, we have $\mathrm{id}_G(g \cdot h) = g \cdot h = g \cdot \mathrm{id}_G(h)$, so $\mathrm{id}_G \in \mathrm{Hom}_G(G,G)$. Second, let f be a G–equivariant map. For any $g \in G$, we have $f(g) = f(g \cdot 1) = g \cdot f(1)$, that is f acts on G by right multiplication by $f(1)$. Thus, it is bijective with its inverse being right multiplication by $f(1)^{-1}$. Third, if f and f' are both G–equivariant map, then

$$(f \circ f')(g \cdot h) = f(f'(g \cdot h)) = f(g \cdot f'(h)) = g \cdot (f \circ f')(h),$$

hence $f \circ f'$ is G–equivariant. Finally, we saw f^{-1} is right multiplication by $f(1)^{-1}$, and it is G–equivariant as $f^{-1}(g \cdot h) = g \cdot h \cdot f(1)^{-1} = g \cdot f^{-1}(h)$. We conclude that $\mathrm{Hom}_G(G,G)$ is a subgroup of Σ_G.

The final check is that the Yoneda bijection $G \to \mathrm{Hom}_G(G,G)$ sending g to $(-) \cdot g$ is a group homomorphism.[449] It is clear that it sends the identity to the identity and for any $g, h \in G$

$$(-) \cdot gh = ((-) \cdot g) \cdot h = ((-) \cdot h) \circ ((-) \cdot g),$$

so this is a group homomorphism.

I would like to believe this book is not a "shallow exploration of category theory", so we will also see more concrete uses of Yoneda.

Example 404 (Exponentials in **DGph**)**.** We saw in Chapter 4 that **DGph** is a topos, so it has exponentials, but we did not write a nice description for them.[450] We will do this here relying on Yoneda and the isomorphism $\mathbf{DGph} \cong [V \rightrightarrows E, \mathbf{Set}]$ outlined in Example 359.

6.3 Universality as Representability

Representability is one of the two ways to describe universal constructions that we hinted at at the end of Chapter 4. In this section, we will explore how any universal property is equivalent to representability of some functor. Since (co)limits and

[448] We see G as a G–set with the action of left multiplication as in Example 394.

[449] isomorphism follows because it is a bijection.

[450] Theoretically, we know how to compute them because we have seen how to take power objects in Example 310 and (co)limits in Example 359, but we will take a more direct approach here.

universal morphisms are initial or terminal objects in some category, there is a first trivial way to express universality as representability.

SOL **Exercise 405** (NOW!)**.** Let $X \in \mathbf{C}_0$ and $\Delta(\mathbf{1}) : \mathbf{C} \rightsquigarrow \mathbf{Set}$ be the constant functor at the singleton $\mathbf{1} = \{\star\}$. Show that $\mathrm{Hom}_{\mathbf{C}}(X, -) \cong \Delta(\mathbf{1})$ if and only if X is initial. Dually, $\mathrm{Hom}_{\mathbf{C}}(-, X) \cong \Delta(\mathbf{1})$ if and only if X is terminal.[451]

[451] In the dual statement, the source of $\Delta(\mathbf{1})$ is $\mathbf{C}^{\mathrm{op}}$.

It turns out this result is not very useful.

Proposition 406. *Let $X, Y \in \mathbf{C}_0$. The product of X and Y exists if and only if there exists $P \in \mathbf{C}_0$ such that $\mathrm{Hom}_{\mathbf{C}\times\mathbf{C}}(\Delta_{\mathbf{C}}(-), (X, Y)) \cong \mathrm{Hom}_{\mathbf{C}}(-, P)$. The product is P.*

Proof. ($\Rightarrow$) Let $P = X \times Y$, for any $A \in \mathbf{C}_0$, there is an isomorphism

$$\mathrm{Hom}_{\mathbf{C}\times\mathbf{C}}((A, A), (X, Y)) \cong \mathrm{Hom}_{\mathbf{C}}(A, X \times Y)$$

which sends the pair $(f : A \to X, g : A \to Y)$ to $\langle f, g\rangle : A \to X \times Y$.[452] In the other direction, $p : A \to X \times Y$ is sent to the pair $(\pi_X \circ p, \pi_Y \circ p)$. Let us show it is natural in A. For any $m : A' \to A$, (171) commutes because the top path sends the pair (f, g) to the morphism $\langle f, g\rangle$ then to $\langle f, g\rangle \circ m = \langle f \circ m, g \circ m\rangle$ and the bottom path sends (f, g) to $(f, g) \circ (m, m) = (f \circ m, g \circ m)$ which is then sent to $\langle f \circ m, g \circ m\rangle$.

[452] Recall that $\langle f, g\rangle$ is the unique morphism satisfying $\pi_X \circ \langle f, g\rangle = f$ and $\pi_Y \circ \langle f, g\rangle = g$. Be careful not to confuse it with a pair of morphisms.

$$\begin{array}{ccc} \mathrm{Hom}_{\mathbf{C}\times\mathbf{C}}((A, A), (X, Y)) & \xrightarrow{\sim} & \mathrm{Hom}_{\mathbf{C}}(A, X \times Y) \\ {\scriptstyle -\circ(m,m)}\downarrow & & \downarrow{\scriptstyle -\circ m} \\ \mathrm{Hom}_{\mathbf{C}\times\mathbf{C}}((A', A'), (X, Y)) & \xrightarrow[\sim]{} & \mathrm{Hom}_{\mathbf{C}}(A', X \times Y) \end{array} \tag{171}$$

($\Leftarrow$) First, we define π_X and π_Y to be the pair of morphisms corresponding to id_P under the isomorphism $\mathrm{Hom}_{\mathbf{C}\times\mathbf{C}}((P, P), (X, Y)) \cong \mathrm{Hom}_{\mathbf{C}}(P, P)$.[453] Given two morphisms $f : A \to X$ and $g : A \to Y$, the isomorphism

[453] Once more, we are making a forced choice. To define the projections, we need two morphims $P \to X$ and $P \to Y$. By the natural isomorphism of the hypothesis, it is enough to find a morphism $P \to P$. We can only take id_P as we know nothing else about $\mathbf{C}$.

$$\mathrm{Hom}_{\mathbf{C}\times\mathbf{C}}((A, A), (X, Y)) \cong \mathrm{Hom}_{\mathbf{C}}(A, P)$$

yields a unique morphism $! : A \to P$. To see that $\pi_X \circ ! = f$ and $\pi_Y \circ ! = g$ we start with id_P in the top right of (172) which commutes by hypothesis.

$$\begin{array}{ccc} \mathrm{Hom}_{\mathbf{C}\times\mathbf{C}}((P, P), (X, Y)) & \xleftarrow{\sim} & \mathrm{Hom}_{\mathbf{C}}(P, P) \\ {\scriptstyle -\circ(!,!)}\downarrow & & \downarrow{\scriptstyle -\circ !} \\ \mathrm{Hom}_{\mathbf{C}\times\mathbf{C}}((A, A), (X, Y)) & \xleftarrow[\sim]{} & \mathrm{Hom}_{\mathbf{C}}(A, P) \end{array} \tag{172}$$

□

Corollary 407 (Dual)**.** *Let $X, Y \in \mathbf{C}_0$. The coproduct of X and Y exists if and only if there exists $S \in \mathbf{C}_0$ such that $\mathrm{Hom}_{\mathbf{C}\times\mathbf{C}}((X, Y), \Delta_{\mathbf{C}}(-)) \cong \mathrm{Hom}_{\mathbf{C}}(S, -)$. The coproduct is S.*[454]

[454] We implicitly use the fact that $(\mathbf{C} \times \mathbf{C})^{\mathrm{op}} \cong \mathbf{C}^{\mathrm{op}} \times \mathbf{C}^{\mathrm{op}}$.

In order to generalize these two results to arbitrary (co)limits, we define the generalized version of $\Delta_{\mathbf{C}}$.

Definition 408 (Generalized diagonal functor)**.** Let $\mathbf{J}$ be a small category, the **generalized diagonal functor** $\Delta^{\mathbf{J}}_{\mathbf{C}} : \mathbf{C} \rightsquigarrow [\mathbf{J}, \mathbf{C}]$ sends an object $X \in \mathbf{C}_0$ to the constant functor at X and a morphism $f : X \to Y \in \mathbf{C}_1$ to the natural transformation whose components are all $f : X \to Y$.

We have $\Delta^{\mathbf{J}}_{\mathbf{C}}(f) : X \Rightarrow Y$ because for any $a \in \mathbf{J}_1$, the square below commutes.

$$\begin{array}{ccc} X & \xrightarrow{X(a)=\mathrm{id}_X} & X \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle f} \\ Y & \xrightarrow[Y(a)=\mathrm{id}_Y]{} & Y \end{array}$$

Remark 409. This is a generalization of the diagonal functor $\Delta_{\mathbf{C}} : \mathbf{C} \rightsquigarrow \mathbf{C} \times \mathbf{C}$ because, with the isomorphism $[\mathbf{1}+\mathbf{1}, \mathbf{C}] \cong \mathbf{C} \times \mathbf{C}$ described in Example 351.2, we can identify $\Delta_{\mathbf{C}}$ with $\Delta^{\mathbf{1}+\mathbf{1}}_{\mathbf{C}}$.

Proposition 410. *Let $F : \mathbf{J} \rightsquigarrow \mathbf{C}$ be a diagram. The limit of F exists if and only if there is an object $L \in \mathbf{C}_0$ such that* $\mathrm{Nat}(\Delta^{\mathbf{J}}_{\mathbf{C}}(-), F) \cong \mathrm{Hom}_{\mathbf{C}}(-, L)$.[455] *The tip of the limit cone is L.*

[455] Recall that

$$\mathrm{Nat}(\Delta^{\mathbf{J}}_{\mathbf{C}}(-), F) = \mathrm{Nat}(-, F) \circ \Delta^{\mathbf{J}}_{\mathbf{C}}.$$

For this to be a functor $\mathbf{C}^{\mathrm{op}} \rightsquigarrow \mathbf{Set}$, it is important that $\mathbf{J}$ is small and $\mathbf{C}$ is locally small as it guarantees the functor category $[\mathbf{J}, \mathbf{C}]$ to be locally small too, hence $\mathrm{Nat}(\Delta^{\mathbf{J}}_{\mathbf{C}}(X), F)$ is a set for any $X \in \mathbf{C}_0$.

Proof. First, we note that for any $X \in \mathbf{C}_0$, a natural transformation $\psi : \Delta^{\mathbf{J}}_{\mathbf{C}}(X) \Rightarrow F$ is a cone over F with tip X. Indeed, for any $a : A \to B \in \mathbf{J}_1$, the naturality square in (173) is commutative.

$$\begin{array}{ccc} X & \xrightarrow{X(a)=\mathrm{id}_X} & X \\ {\scriptstyle \psi_A}\downarrow & & \downarrow{\scriptstyle \psi_B} \\ FA & \xrightarrow[F(a)]{} & FB \end{array} \tag{173}$$

This is equivalent to $\{\psi_A : X \to FA\}_{A \in \mathbf{J}_0}$ being a cone over F. Furthermore, a morphism of cones $\phi \to \psi$ is a morphism f between the tips such that $\forall A \in \mathbf{J}_0, \phi_A = \psi_A \circ f$. By looking at (174), we see this condition is equivalent to $\phi = \psi \cdot \Delta^{\mathbf{J}}_{\mathbf{C}}(f)$.

$$\begin{array}{ccccc} & & Y & \xrightarrow{\mathrm{id}_Y} & Y \\ & {\scriptstyle f}\swarrow & & {\scriptstyle f}\swarrow & \\ X & \xrightarrow{\mathrm{id}_X} & X & & \downarrow{\scriptstyle \phi_B} \\ & {\scriptstyle \psi_A}\searrow & {\scriptstyle \phi_A}\downarrow & {\scriptstyle \psi_B}\searrow & \\ & & FA & \xrightarrow[F(a)]{} & FB \end{array} \tag{174}$$

($\Rightarrow$) Let $\{\psi_A : L \to FA\}_{A \in \mathbf{J}_0}$ be the terminal cone over F (i.e. the limit) and see it as a natural transformation $\psi : \Delta^{\mathbf{J}}_{\mathbf{C}}(L) \Rightarrow F$. We need to define a natural isomorphism $\mathrm{Nat}(\Delta^{\mathbf{J}}_{\mathbf{C}}(-), F) \cong \mathrm{Hom}_{\mathbf{C}}(-, L)$. Similarly to the proofs of the previous section, we will see that we only need to see where id_L is sent to and the rest of the natural transformation will *construct itself*. Our only choice for the cone corresponding to id_L is ψ (it is the only cone we know exists).

Indeed, for any $f : X \to L$ the naturality square in (175) means the cone corresponding to $f : X \to L$ is $\{\psi_A \circ f : X \to FA\}_{A \in \mathbf{J}_0}$ by starting with id_L in the top right. Now, since ψ is the terminal cone, for any cone $\{\phi_A : X \to FA\}_{A \in \mathbf{J}_0}$, there is a unique morphism of cones $f : X \to L$ which satisfies $\forall A \in \mathbf{J}_0, \psi_A \circ f = \phi_A$. We conclude that $f \mapsto \psi \cdot \Delta^{\mathbf{J}}_{\mathbf{C}}(f)$ is a natural isomorphism.

$$\begin{array}{ccc} \mathrm{Nat}(\Delta^{\mathbf{J}}_{\mathbf{C}}(L), F) & \longleftrightarrow & \mathrm{Hom}_{\mathbf{C}}(L, L) \\ {\scriptstyle -\cdot\Delta^{\mathbf{J}}_{\mathbf{C}}(f)}\downarrow & & \downarrow{\scriptstyle -\circ f} \\ \mathrm{Nat}(\Delta^{\mathbf{J}}_{\mathbf{C}}(X), F) & \longleftrightarrow & \mathrm{Hom}_{\mathbf{C}}(X, L) \end{array} \tag{175}$$

($\Leftarrow$) Let $\psi : \Delta^{\mathbf{J}}_{\mathbf{C}}(L) \Rightarrow F$ be the cone corresponding to $\mathrm{id}_L \in \mathrm{Hom}_{\mathbf{C}}(L, L)$ under the natural isomorphism, we will show it is terminal. By the commutativity of (175) and bijectivity of the horizontal arrows, for any cone $\phi : \Delta^{\mathbf{J}}_{\mathbf{C}}(X) \Rightarrow F$, there is a unique morphism $f : X \to L$ such that $\phi = \psi \cdot \Delta^{\mathbf{J}}_{\mathbf{C}}(f)$. By the first paragraph of the proof, this is the unique morphism of cones showing ψ is terminal. □

Corollary 411 (Dual)**.** *Let $F : \mathbf{J} \rightsquigarrow \mathbf{C}$ be a diagram. The colimit of F exists if and only if there is an object $L \in \mathbf{C}_0$ such that* $\mathrm{Nat}(F, \Delta^{\mathbf{J}}_{\mathbf{C}}(-)) \cong \mathrm{Hom}_{\mathbf{C}}(L, -)$. *The tip of the colimit cone is L.*

Proposition 412. *Let $U : \mathbf{Mon} \rightsquigarrow \mathbf{Set}$ be the forgetful functor, A be a set and A^* be the free monoid on A, we have* $\mathrm{Hom}_{\mathbf{Set}}(A, U-) \cong \mathrm{Hom}_{\mathbf{Mon}}(A^*, -)$.

Proof. We have already shown before Definition 288 that sending $h : A \to M$ to $h^* : A^* \to M$ is a bijection of the desired type.[456] Now, we need to show it is natural in M. For any monoid homomorphism $f : M \to N$, (176) commutes (we omitted applications of U) because starting with $h : A \to M$, we have $(f \circ h)^* = f \circ h^*$.[457]

$$\begin{array}{ccc} \mathrm{Hom}_{\mathbf{Set}}(A, M) & \xrightarrow{\sim} & \mathrm{Hom}_{\mathbf{Mon}}(A^*, M) \\ {\scriptstyle f\circ -}\downarrow & & \downarrow{\scriptstyle f\circ -} \\ \mathrm{Hom}_{\mathbf{Set}}(A, N) & \xrightarrow[\sim]{} & \mathrm{Hom}_{\mathbf{Mon}}(A^*, N) \end{array} \tag{176}$$

□

[456] In the other direction, $h : A^* \to M$ is sent to $U(h) \circ i$ where $i : A \hookrightarrow A^*$ is the inclusion.

[457] To check this, let $w = a_1 \cdots a_n \in A^*$, we have

$$\begin{aligned} (f \circ h)^*(w) &= fh(a_1) \cdots fh(a_n) \\ &= f(h(a_1) \cdots h(a_n)) \\ &= f(h^*(w)). \end{aligned}$$

In the next Proposition, we will generalize this result to see how any universal morphism corresponds to some kind of representability and we will even give a converse direction. The generalizations of the proof is straightforward, so I suggest you try to get familiar with a specific case in the next exercise.

SOL **Exercise 413.** Let $\mathbf{C}$ be a category and $X \in \mathbf{C}_0$ be such that $- \times X$ is a functor. An object $A \in \mathbf{C}_0$ has an exponential $A^X \in \mathbf{C}_0$ if and only if $\mathrm{Hom}_{\mathbf{C}}(- \times X, A) \cong \mathrm{Hom}_{\mathbf{C}}(-, A^X)$.

Proposition 414. *Let $F : \mathbf{C} \rightsquigarrow \mathbf{D}$ be a functor and $X \in \mathbf{D}_0$. There is a universal morphism from X to F if and only if there exists $A \in \mathbf{C}_0$ such that* $\mathrm{Hom}_{\mathbf{D}}(X, F-) \cong \mathrm{Hom}_{\mathbf{C}}(A, -)$.

Proof. ($\Rightarrow$) Let $a : X \to FA$ be a universal morphism, by definition, for any $b : X \to FB$, there is a unique morphism $\phi_B(b) : A \to B$ such that $F(\phi_B(b)) \circ a = b$. In the other direction, ϕ_B^{-1} sending $f : A \to B$ to $Ff \circ a$ is the inverse of ϕ_B.[458] Let us now check that ϕ_B is natural. For any $m : B \to B'$, (177) commutes because when starting with $f : A \to B$ in the top right, the top path sends it to $Ff \circ a$ then to $Fm \circ Ff \circ a$ and the bottom path sends it to $m \circ f$ then to $F(m \circ f) \circ a$.

[458] We check they are inverses:

$$\begin{aligned} \phi_B^{-1}(\phi_B(b)) &= F(\phi_B(b)) \circ a = b \\ \phi_B(\phi_B^{-1}(f)) &= \phi_B(Ff \circ a) = f. \end{aligned}$$

$$\begin{array}{ccc} \mathrm{Hom}_{\mathbf{C}}(X, FB) & \xleftarrow{\sim} & \mathrm{Hom}_{\mathbf{D}}(A, B) \\ {\scriptstyle Fm\circ -}\downarrow & & \downarrow{\scriptstyle m\circ -} \\ \mathrm{Hom}_{\mathbf{C}}(X, FB') & \xleftarrow[\sim]{} & \mathrm{Hom}_{\mathbf{D}}(A, B') \end{array} \tag{177}$$

($\Leftarrow$) Let $a : X \to FA$ be the image of $\mathrm{id}_A : A \to A$ under the isomorphism $\mathrm{Hom}_{\mathbf{C}}(X, FA) \cong \mathrm{Hom}_{\mathbf{D}}(A, A)$, we claim that a is a universal morphism from X to F. Given $b : X \to FB$, let $\phi_B(b)$ be its image under the isomorphism $\mathrm{Hom}_{\mathbf{C}}(X, FB) \cong \mathrm{Hom}_{\mathbf{D}}(A, B)$, it satisfies $F(\phi_B(b)) \circ a = b$ because (178) commutes (start with id_A in the top right corner). The morphism $\phi_B(b)$ is unique with this property because any other $f : A \to B$ is the image of some $b' \neq b$ under ϕ_B yielding $Ff \circ a = b' \neq b$.

□

$$\begin{array}{ccc} \mathrm{Hom}_{\mathbf{C}}(X, FA) & \xleftarrow{\sim} & \mathrm{Hom}_{\mathbf{D}}(A, A) \\ {\scriptstyle F(\phi_B(b))\circ -}\downarrow & & \downarrow{\scriptstyle \phi_B(b)\circ -} \\ \mathrm{Hom}_{\mathbf{C}}(X, FB) & \xleftarrow[\sim]{} & \mathrm{Hom}_{\mathbf{D}}(A, B) \end{array} \tag{178}$$

Corollary 415 (Dual)**.** *Let $F : \mathbf{C} \rightsquigarrow \mathbf{D}$ be a functor and $X \in \mathbf{D}_0$. There is a universal morphism from F to X if and only if there exists $A \in \mathbf{C}_0$ such that* $\mathrm{Hom}_{\mathbf{D}}(F-, X) \cong \mathrm{Hom}_{\mathbf{C}}(-, A)$.

Comparing Propositions 410 and 414 and their duals, we infer that (co)limits satisfy universal properties.

Theorem 416. *Let $F \in [\mathbf{J}, \mathbf{C}]_0$ be a diagram.*

- *The limit of F exists if and only if there is a universal morphism from $\Delta^{\mathbf{J}}_{\mathbf{C}}$ to F.*
- *The colimit of F exists if and only if there is a universal morphism from F to $\Delta^{\mathbf{J}}_{\mathbf{C}}$.*

In the next chapter, we will lift these correspondence to a more global version. Namely, we will see how to assemble the universal morphisms for all diagrams of shape $\mathbf{J}$ (if they all exist) into something called a right adjoint to $\Delta^{\mathbf{J}}_{\mathbf{C}}$.

7 Adjunctions

7.1 Equivalent Definitions 135
7.2 Results and Examples 141

Remark 417. Adjunctions are very much everywhere in mathematics (once you learn to recognize them), and this inevitably means there are many angles to approach a first understanding. We will only get to see my favorite here, it can be roughly summarized in "adjunctions are global universal constructions", but of course I suggest you visit other resources to round out your intuitions.[459]

[459] I think feeling comfortable with adjunctions is a good signal that you are done with your journey in so-called basic category theory, and you are ready for the harder stuff (or you can apply basic category theory to other stuff).

In Chapter 4 on universal properties, we gave categorical descriptions of important constructions in mathematics. We defined the free monoid *on a set*, the abelianization *of a group*, and the exponential *of a set* by another one. The given set (resp. group) on which the constructions are applied is part of the definitions we gave, but we know that they can be applied to any set (resp. group). Therefore, one might ask if it is possible to define (categorically) the construction as a whole. For instance, the action of taking free monoids sends a set to a monoid, so it could be the action on objects of a functor from **Set** to **Mon**.

We start by explainig how this functor arises simply from the existence of free monoids on every set.[460] More abstractly, we show that having an object FX with a universal property based on X for every X means that F is a functor. Moreover, we will see that F is closely related to the functor used in the universal property. This relation is what we call an adjunction. The rest of the chapter will be dedicated to learning more about adjunctions through examples and properties.

[460] We spend a lot of time on this example, so you might want to revisit your understanding of free monoids before moving on.

7.1 Equivalent Definitions

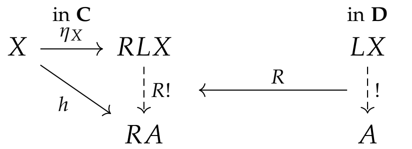
There are four very commonly used defintions of an adjunction.[461] We will start from the one that is most directly linked to the concrete setting of free monoids, and then develop the details (in the abstract setting) to get the other definitions. Finally, we will prove the equivalence between the definitions.

[461] Morally only three because one is dual to another.

Let us have two categories $\mathbf{C}$ and $\mathbf{D}$ and a functor $R : \mathbf{D} \rightsquigarrow \mathbf{C}$.[462] Suppose that for any $X \in \mathbf{C}_0$, we have a universal morphism from X to R, namely, we have an object $LX \in \mathbf{D}_0$ and a morphism $\eta_X : X \to RLX$ satisfying a universal property as in Definition 315 and summarized below.[463]

[462] In our concrete running example, $\mathbf{C} = \mathbf{Set}$, $\mathbf{D} = \mathbf{Mon}$ and R is the forgetful functor.

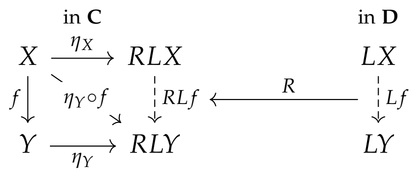
[463] For free monoids, LX is the free monoid on X, i.e. X^*, and η_X is the inclusion of X inside X^* (R only forgets the monoid structure).



(179)

We first show that the action $X \mapsto LX$ is functorial (yielding a functor $L : \mathbf{C} \rightsquigarrow \mathbf{D}$). For any $f : X \to Y$, the universality of η_X yields a unique morphism $Lf : LX \to LY$ satisfying $RLf \circ \eta_X = \eta_Y \circ f$ as summarized in (180).[464]

[464] For free monoids, $Lf : X^* \to Y^*$ is the homomorphism defined inductively by $Lf(\varepsilon) = \varepsilon$ and $Lf(w \cdot x) = Lf(w) \cdot f(x)$. Concretely, it applies f to every letter of the word.



(180)

The functoriality follows from the following equations showing that $L(\mathrm{id}_X) = \mathrm{id}_{LX}$ and $L(g \circ f) = Lg \circ Lf$ because these morphisms make the relevant diagrams commute:[465]

[465] The equations respectively show that id_{LX} makes (179) commute when h is replaced by id_X and $Lg \circ Lf$ does it when h is replaced by $g \circ f$.

$$R(\mathrm{id}_{LX}) \circ \eta_X = \mathrm{id}_{RLX} \circ \eta_X = \eta_X = \eta_X \circ \mathrm{id}_X$$
$$R(Lg \circ Lf) \circ \eta_X = RLg \circ RLf \circ \eta_X = RLg \circ \eta_Y \circ f = \eta_Z \circ (g \circ f).$$

Note that the definition of L on morphisms readily gives us that η is a natural transformation $\mathrm{id}_{\mathbf{C}} \Rightarrow RL$. The functor L constructed like that is called the left adjoint to R.[466]

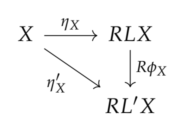
[466] For free monoids, L is the free monoid functor $\mathbf{Mon} \rightsquigarrow \mathbf{Set}$ sending X to X^* and it is the left adjoint to the forgetful functor $\mathbf{Mon} \rightsquigarrow \mathbf{Set}$.

Definition 418 (Left adjoint)**.** Let $R : \mathbf{D} \rightsquigarrow \mathbf{C}$ be a functor. A functor $L : \mathbf{C} \rightsquigarrow \mathbf{D}$ is called the **left adjoint** to R if there exists a natural transformation $\eta : \mathrm{id}_{\mathbf{C}} \Rightarrow RL$ such that for every X, $\eta_X : X \to RLX$ is a universal morphism from X to R, equivalently, η_X is initial in $\Delta(X) \downarrow R$.

Following the construction of L with another family of universal morphisms to R would yield another left adjoint. Thus, to justify the use of the definite article *the*, we can prove that the two left adjoints would be naturally isomorphic.

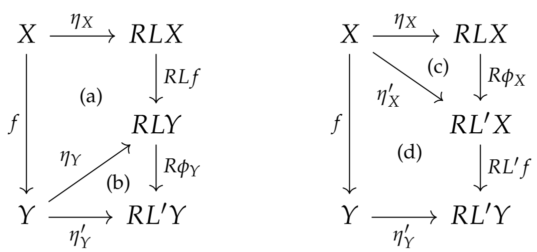
Proposition 419. *Let $R : \mathbf{D} \rightsquigarrow \mathbf{C}$ be a functor, and $L, L' : \mathbf{C} \rightsquigarrow \mathbf{D}$ be two left adjoints to R. Then, $L \cong L'$.*

Proof. Let $\eta : \mathrm{id}_{\mathbf{C}} \Rightarrow RL$ and $\eta' : \mathrm{id}_{\mathbf{C}} \Rightarrow RL'$ be the natural transformations witnessing L and L' respectively as left adjoints to R. For any X, since both $\eta_X : X \to RLX$ and $\eta'_X : X \to RL'X$ are initial in $\Delta(X) \downarrow R$, they must be isomorphic inside this comma category. This means there is an (unique) isomorphism $\phi_X : LX \to LX'$ making (181) commute. It is an isomorphism in $\Delta(X) \downarrow R$, but we find it is also an isomorphism in $\mathbf{D}$ by applying the forgetful functor $U_R : \Delta(X) \downarrow R \rightsquigarrow \mathbf{D}$ from Exercise 320 (recall Exercise 187.4).



(181)

It remains to show these components assemble into a natural transformation, i.e. that for any $f : X \to Y$, $L'f \circ \phi_X = \phi_Y \circ Lf$. We start by drawing the following two commutative diagrams.



(182)

Showing (182) commutes:

(a) $\mathsf{NAT}(\eta, X, Y, f)$.

(b) Definition of ϕ (181).

(c) Definition of ϕ (181).

(d) $\mathsf{NAT}(\eta', X, Y, f)$.

We find that both $\phi_Y \circ Lf$ and $L'f \circ \phi_X$ make (179) commute when h is replaced by $\eta'_Y \circ f$. Thus, by uniqueness, they must be equal. We conclude that ϕ is a natural isomorphism $L \Rightarrow L'$. □

The dual concept is called a right adjoint.

Definition 420 (Right adjoint)**.** Let $L : \mathbf{C} \rightsquigarrow \mathbf{D}$ be a functor. A functor $R : \mathbf{D} \rightsquigarrow \mathbf{C}$ is called the right adjoint to L is there exists a natural transformation $\varepsilon : LR \Rightarrow \mathrm{id}_{\mathbf{D}}$ such that for every X, $\varepsilon_X : LRX \to X$ is a universal morphism from L to X, equivalently, ε_X is terminal in $L \downarrow \Delta(X)$.

Corollary 421 (Dual)**.** *If $R, R' : \mathbf{D} \rightsquigarrow \mathbf{C}$ are two right adjoints to $L : \mathbf{C} \rightsquigarrow \mathbf{D}$, then $R \cong R'$.*

Example 422 (Cartesian closedness)**.** Let $\mathbf{C}$ be a category with all finite products (in particular, binary ones and a terminal object). Given two objects $A, X \in \mathbf{C}_0$, recall that their exponential exists if and only if there is a universal morphism $\mathsf{ev} : A^X \times X \to A$ from $- \times X$ to A.

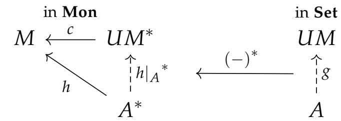
Fixing X, if this exponential exists for every $A \in \mathbf{C}_0$, then a dual argument to the one preceding Definition 418 shows that the assignment $A \mapsto A^X$ yields a functor $\mathbf{C} \rightsquigarrow \mathbf{C}$ that is right adjoint to $- \times X : \mathbf{C} \rightsquigarrow \mathbf{C}$ from Exercise 297, and moreover the evaluation morphisms are components of a natural transformation $(-)^X \times X \Rightarrow \mathrm{id}_{\mathbf{C}}$. By Definition 301, $\mathbf{C}$ is cartesian closed precisely when all functors $- \times X$ have a right adjoint.

Example 423 (Free monoids)**.** We saw that the free monoid functor $(-)^* : \mathbf{Set} \rightsquigarrow \mathbf{Mon}$ is left adjoint to the forgetful functor $U : \mathbf{Mon} \rightsquigarrow \mathbf{Set}$. We can also show that U is right adjoint to $(-)^*$. For any monoid $M \in \mathbf{Mon}_0$, we need to define a monoid homomorphism $UM^* \to M$. Since an element $w \in UM^*$ is a word whose letters are elements of M, we can multiply all the letters together with the monoid operation (the order does not matter thanks to associativity) to get one element of M. We call this function $c : UM^* \to M$, and the fact that it is a homomorphism also follows from associativity.

Now, for any set A and homomorphism $h : A^* \to M$, we know that the action of h is completely determined by where it sends the single-letter words.[467] More precisely, we know that if $w = \mathsf{a}_1 \cdots \mathsf{a}_n$ is a word in A^*, then $h(w) = h(\mathsf{a}_1) \cdots h(\mathsf{a}_n)$, where $\cdots$ denotes here the multiplciation in M. If we instead see $h(\mathsf{a}_1) \cdots h(\mathsf{a}_n)$ as a word in UM^*, i.e. $\cdots$ denotes concatenation of letters, it can be obtained

[467] You can see this as a consequence of either the classical Definition 287 or the categorical Definition 288 of free monoids.

by applying the restriction of h to A to every letter in w, i.e. $h(\mathsf{a}_1)\cdots h(\mathsf{a}_n) = h|_A{}^*(\mathsf{a}_1\cdots\mathsf{a}_n) = h|_A{}^*(w)$. This lets us see that $h|_A : A \to UM$ is the unique function satisfying $c(h|_A{}^*) = h$, and we conclude that c satisfies the appropriate universal property summarized in (183).



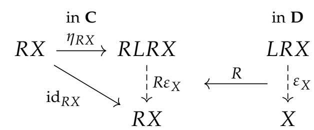
(183)

As for exponentials, we find that U is right adjoint to $(-)^*$.

In our running example, we now have a pair of functors ($(-)^*$ and U) adjoint to each other, one left adjoint and the other right adjoint. It turns out we can develop Example 423 abstractly and show that when L is left adjoint to R, then R is right adjoint to L, and vice-versa by duality.

Proposition 424. *Let $L : \mathbf{C} \rightsquigarrow \mathbf{D}$ and $R : \mathbf{D} \rightsquigarrow \mathbf{C}$ be two functors. If L is left adjoint to R, then R is right adjoint to L.*

Proof. Let $\eta : \mathrm{id}_{\mathbf{C}} \Rightarrow RL$ be the natural transformation witnessing L as left adjoint to R. We first define the components of a natural transformation $\varepsilon : LR \Rightarrow \mathrm{id}_{\mathbf{D}}$. For $X \in \mathbf{D}_0$, we need a morphism $LRX \to X$ in $\mathbf{D}$, and we know from the universal property of η_{RX} that it is enough to find a morphism $RX \to RX$. Of course we take the identity, and we let ε_X be the unique morphism given by the universality of η_{RX} such that $R(\varepsilon_X) \circ \eta_{RX} = \mathrm{id}_{RX}$ (see (184)).



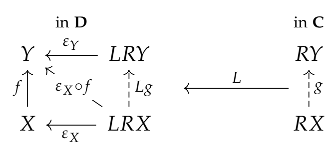
(184)

Next, we show that $\varepsilon_X : LRX \to X$ is a universal morphism from L to X. For any $f : LA \to X$, if $g : A \to RX \in \mathbf{C}_1$ is such that $f = \varepsilon_X \circ Lg$, then applying R and pre-composing with η_A, we obtain

$$\begin{aligned} Rf \circ \eta_A &= R\varepsilon_X \circ RLg \circ \eta_A \\ &= R\varepsilon_X \circ \eta_{RX} \circ g && \mathsf{NAT}(\eta, A, RX, g) \\ &= \mathrm{id}_{RX} \circ g && \text{definition of } \varepsilon_X \\ &= g. \end{aligned}$$

We conclude that $g := Rf \circ \eta_A$ is the unique morphism satisfying that $f = \varepsilon_X \circ Lg$, hence ε_X is universal.

Finally, we show that $\varepsilon : LR \Rightarrow \mathrm{id}_{\mathbf{D}}$ is natural. For any $f : X \to Y \in \mathbf{D}_1$, by universality, there is a unique morphism $g : RX \to RY$ such that $f \circ \varepsilon_X = \varepsilon_Y \circ Lg$ (see (185)) and by our derivation above, $g = Rf \circ R\varepsilon_X \circ \eta_{RX} \overset{(184)}{=} Rf$. Thus, we find that $f \circ \varepsilon_X = \varepsilon_Y \circ LRf$, namely ε is natural. ☐



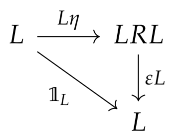
(185)

As a sanity check, notice that using the definition of ε_M in the case of free monoids, we get back the homomorphism c from Example 423. Indeed, instantiating (184), we find $\varepsilon_M : UM^* \to M$ is the unique homomorphism that acts like identity on single-letter words M (recall η_{UM} sends $x \in UM$ to the word $x \in UM^*$). It is easy to check c also acts like identity on single-letter words, so ε_M and c coincide by uniqueness.

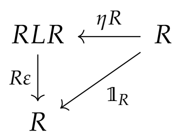
Corollary 425 (Dual)**.** *If R is right adjoint to L, then L is left adjoint to R.*

This makes Definitions 418 and 420 a bit unsatisfactory because they seem to focus on one side of relation between two functors. To resolve this, we bring up two important properties that arise from having a left and right adjoint, and we will see these also characterize adjoints.

First, we note that $\eta : \mathrm{id}_{\mathbf{C}} \Rightarrow RL$ and $\varepsilon : LR \Rightarrow \mathrm{id}_{\mathbf{D}}$ seem to have the right type to give rise to an equivalence between **C** and **D**. However, in general, nothing guarantees the components of η and ε are isomorphisms.[468] There is still some kind of invertibility property: η and ε satisfy the the **triangle identities** shown in (186) and (187) (they are commutative diagrams in $[\mathbf{C}, \mathbf{D}]$ and $[\mathbf{D}, \mathbf{C}]$ respectively).



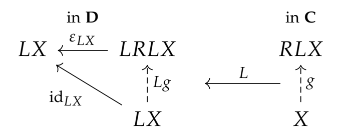
(186)



(187)

[468] It is clearly not the case in the free monoids example.

The second one holds by definition of ε_X (for any $X \in \mathbf{D}_0$, $R\varepsilon_X \circ \eta_{RX} = \mathrm{id}_{RX}$). For the first one, by universality of ε_X, there is a unique morphism $g : X \to RLX$ such that $\varepsilon_{LX} \circ Lg = \mathrm{id}_{LX}$ (see (188)), and by our derivation in the previous proof, $g = R(\mathrm{id}_{LX}) \circ \eta_X = \eta_X$. We find that $\varepsilon_{LX} \circ L\eta_X = \mathrm{id}_{LX}$ as desired.



(188)

It is simple, but not very illuminating to see how these triangle identities hold in the free monoids example. Conversely, the next characterization of adjoints is in the spotlight of our running example. It abstractly states the slogan that it is the same thing to give a homomorphism out of the free monoid A^* or a function out of the set A.

Formally, we find a natural isomorphism[469]

$$\Phi : \mathrm{Hom}_{\mathbf{C}}(-, R-) \cong \mathrm{Hom}_{\mathbf{D}}(L-, -) : \Phi^{-1}.$$

[469] For free monoids, this gives

$$\mathrm{Hom}_{\mathbf{Set}}(A, M) \cong \mathrm{Hom}_{\mathbf{Mon}}(A^*, M),$$

which is inded what the slogan means.

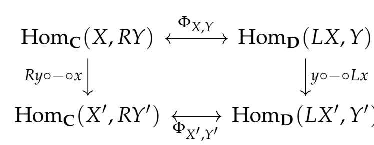
For $g : X \to RY$, we define $\Phi_{X,Y}(g) = \varepsilon_Y \circ Lg$ and for $f : LX \to Y$, we define $\Phi^{-1}_{X,Y}(f) = Rf \circ \eta_X$.[470] The derivations below show these are inverses (and it only relies on the triangle identities and naturality):

[470] You can certainly infer these definitions just by looking at the types. Also note because it will be useful that $\Phi_{X,Y}(\mathrm{id}_{RX}) = \varepsilon_X$ and $\Phi^{-1}_{X,Y}(\mathrm{id}_{LX}) = \eta_X$.

$$\Phi^{-1}_{X,Y}(\Phi_{X,Y}(g)) = R\varepsilon_Y \circ RLg \circ \eta_X = R\varepsilon_Y \circ \eta_{RY} \circ g = g \quad (189)$$

$$\Phi_{X,Y}(\Phi^{-1}_{X,Y}(f)) = \varepsilon_Y \circ LRf \circ L\eta_X = f \circ \varepsilon_{LX} \circ L\eta_X = f. \quad (190)$$

To show that Φ is natural, we need to show that (191) commutes for any $x : X' \to X$ and $y : Y \to Y'$. Starting with $g : X \to RY$ in the top left, the bottom path sends it to $Ry \circ g \circ x$ then to $\varepsilon_{Y'} \circ LRy \circ Lg \circ Lx$ and the top path sends g to $\varepsilon_Y \circ Lg$ then to $y \circ \varepsilon_Y \circ Lg \circ Lx$. The end results are equal by $\mathsf{NAT}(\varepsilon, Y, Y', y)$.



(191)

We can now give an unbiased definition (not focused on one side) of adjunction.

Definition 426 (Adjunction)**.** An **adjunction** between a functor $L : \mathbf{C} \rightsquigarrow \mathbf{D}$ and $R : \mathbf{D} \rightsquigarrow \mathbf{C}$ is the following data:

- A natural transformation $\eta : \mathrm{id}_{\mathbf{C}} \Rightarrow RL$ called the **unit** such that η_X is initial in $\Delta(X) \downarrow R$ for each $X \in \mathbf{C}_0$.
- A natural transformation $\varepsilon : LR \Rightarrow \mathrm{id}_{\mathbf{D}}$ called the **counit** such that ε_X is terminal in $L \downarrow \Delta(X)$ for each $X \in \mathbf{D}_0$.
- The unit η and counit ε satisfy the triangle identities.

- A natural isomorphism $\Phi : \mathrm{Hom}_{\mathbf{C}}(-, R-) \cong \mathrm{Hom}_{\mathbf{D}}(L-, -) : \Phi^{-1}$ such that $\Phi_{RX,X}(\mathrm{id}_{RX}) = \varepsilon_X$ and $\Phi^{-1}_{X,LX}(\mathrm{id}_{LX}) = \eta_X$.[471]

We denote $\mathbf{C} : L \dashv R : \mathbf{D}$ when there is an adjunction between $L : \mathbf{C} \rightsquigarrow \mathbf{D}$ and $R : \mathbf{D} \rightsquigarrow \mathbf{C}$ and we call L the left adjoint and R the right adjoint, and we say L and R are adjoints.[472]

Example 427 (Boring)**.** The identity functor on any category is self-adjoint: $\mathrm{id}_{\mathbf{C}} \dashv \mathrm{id}_{\mathbf{C}}$. Both the unit and counit are $\mathbb{1}_{\mathrm{id}_{\mathbf{C}}}$.[473]

While we resolved the bias in our definitions of adjoints, it cost us brevity. The culminating point of this section is the proof that all this data is not necessary to define an adjunction, giving only one of the fours points is enough. In other words, Definition 426 gives in fact four equivalent definitions of an adjunction.[474]

Theorem 428. *Two functors $L : \mathbf{C} \rightsquigarrow \mathbf{D}$ and $R : \mathbf{D} \rightsquigarrow \mathbf{C}$ are adjoints if at least one of the following holds.*

i. There is a natural transformation $\eta : \mathrm{id}_{\mathbf{C}} \Rightarrow RL$ such that η_X is initial in $\Delta(X) \downarrow R$ for each $X \in \mathbf{C}_0$.

ii. There is a natural transformation $\varepsilon : LR \Rightarrow \mathrm{id}_{\mathbf{D}}$ such that ε_X is terminal in $L \downarrow \Delta(X)$ for each $X \in \mathbf{D}_0$.

iii. There are two natural transformations $\eta : \mathrm{id}_{\mathbf{C}} \Rightarrow RL$ and $\varepsilon : LR \Rightarrow \mathrm{id}_{\mathbf{D}}$ that satisfy the triangle identities.[475]

iv. There is a natural isomorphism $\Phi : \mathrm{Hom}_{\mathbf{C}}(-, R-) \cong \mathrm{Hom}_{\mathbf{D}}(L-, -) : \Phi^{-1}$.

Proof. We have already shown that (i) gives rise to all the data of an adjunction at the start of the chapter.

For (ii), we can use duality. Indeed, taking the dual of Definition 426, we see that $\mathbf{C} : L \dashv R : \mathbf{D}$ if and only if $\mathbf{D}^{\mathrm{op}} : R^{\mathrm{op}} \dashv L^{\mathrm{op}} : \mathbf{C}^{\mathrm{op}}$ and η and ε swap their roles as unit and counit. Hence, from ε, we can derive an adjunction $R^{\mathrm{op}} \dashv L^{\mathrm{op}}$ as we did at the start of the chapter and duality yields $L \dashv R$.

For (iii), it is enough to show the components of the unit η_X are initial in $\Delta(X) \downarrow R$ and use (i).[476] Recall from (189) and (190) that for any $g : X \to RY \in \mathbf{C}_1$, there is a unique (because the components of Φ and Φ^{-1} are bijections) morphism $\Phi_{X,Y}(g) = \varepsilon_Y \circ Lg$ such that $R(\Phi_{X,Y}(g)) \circ \eta_X = \Phi^{-1}_{X,Y}(\Phi_{X,Y}(g)) = g$. Thus, η_X is a universal morphism as required.

For (iv), we will construct a unit satisfying (i). Fix $X \in \mathbf{C}_0$, we have a natural isomorphism $\Phi_{X,-} : \mathrm{Hom}_{\mathbf{C}}(X, R-) \cong \mathrm{Hom}_{\mathbf{D}}(LX, -)$. By Proposition 414, there is a universal morphism $\eta_X : X \to RLX$ from X to R.[477] This yields a natural transformation $\eta : \mathrm{id}_{\mathbf{C}} \Rightarrow RL$ because for any $f : X \to Y$, the commutativity of (192) implies (by starting with id_{LX} and id_{LY} in the top left and top right corners

[471] It follows by naturality that $\Phi_{X,Y}(g) = \varepsilon_Y \circ Lg$ and $\Phi^{-1}_{X,Y}(f) = Rf \circ \eta_X$, as we had above.

[472] When they are clear from the context or irrelevant, we omit the categories from the notation and write $L \dashv R$.

[473] You can prove this easily but it also follows from Proposition 435 and the fact that $\mathrm{id}_{\mathbf{C}}$ is its own inverse.

[474] There are still more equivalent definitions, but we have to limit ourselves to a finite list and we mentioned the parts of an adjunction that are most commonly used. One notable omission is that of adjunctions as Kan extensions.

[475] They satisfy

$$\varepsilon L \cdot L\eta = \mathbb{1}_L \qquad R\varepsilon \cdot \eta R = \mathbb{1}_R.$$

[476] You can check that the triangle identities ensure that the adjunction constructed from (i) will have ε as a counit.

[477] From the proof of Proposition 414, we recover $\eta_X = \Phi^{-1}_{X,LX}(\mathrm{id}_{LX})$.

respectively) $RLf \circ \eta_X = \Phi^{-1}_{X,LY}(Lf) = \eta_Y \circ f$.

$$\begin{array}{ccccc} \mathrm{Hom}_{\mathbf{D}}(LX, LX) & \xrightarrow{Lf\circ -} & \mathrm{Hom}_{\mathbf{D}}(LX, LY) & \xleftarrow{-\circ Lf} & \mathrm{Hom}_{\mathbf{D}}(LY, LY) \\ \Phi_{X,LX}\updownarrow & & \Phi_{X,LY}\updownarrow & & \updownarrow\Phi_{Y,LY} \\ \mathrm{Hom}_{\mathbf{C}}(X, RLX) & \xrightarrow[RLf\circ -]{} & \mathrm{Hom}_{\mathbf{C}}(X, RLY) & \xleftarrow[-\circ f]{} & \mathrm{Hom}_{\mathbf{C}}(Y, RLY) \end{array} \quad (192)$$

You can check the natural isomorphism constructed with (i) coincides with Φ. □

Each item of Theorem 428 can be seen as a definition of adjunctions.[478] We would like to spend a bit more time on point (iv) which is, in our opinion, the hardest definition to internalize and yet the easiest one to use in concrete contexts. The definition of an adjunction according to (iv) can be stated as follows.

[478] That is how most textbooks present it.

Two functors $L : \mathbf{C} \rightsquigarrow \mathbf{D}$ and $R : \mathbf{D} \rightsquigarrow \mathbf{C}$ are adjoint if there is a natural isomorphism[479]

$$\mathrm{Hom}_{\mathbf{C}}(-, R-) \cong \mathrm{Hom}_{\mathbf{D}}(L-, -).$$

[479] We use Remark 147 to define

$$\mathrm{Hom}_{\mathbf{C}}(-, R-) := \mathrm{Hom}_{\mathbf{C}}(-,-) \circ (\mathrm{id}_{\mathbf{C}^{\mathrm{op}}} \times R)$$
$$\mathrm{Hom}_{\mathbf{D}}(L-, -) := \mathrm{Hom}_{\mathbf{D}}(-,-) \circ (L^{\mathrm{op}} \times \mathrm{id}_{\mathbf{D}})$$

Less concisely, for any $X \in \mathbf{C}_0$ and $Y \in \mathbf{D}_0$, there is an isomorphism $\Phi_{X,Y} : \mathrm{Hom}_{\mathbf{C}}(X, RY) \cong \mathrm{Hom}_{\mathbf{D}}(LX, Y)$ such that for any $f : X \to X' \in \mathbf{C}_1$ and $g : Y \to Y' \in \mathbf{D}_1$, (193) commutes. We split the naturality in two squares because we will often use one square on its own[480] as we did on both sides of (192).

[480] This is possible by Exercise 344.

$$\begin{array}{ccccc} \mathrm{Hom}_{\mathbf{C}}(X', RY) & \xrightarrow{-\circ f} & \mathrm{Hom}_{\mathbf{C}}(X, RY) & \xrightarrow{Rg\circ -} & \mathrm{Hom}_{\mathbf{C}}(X, RY') \\ \Phi_{X',Y}\updownarrow & & \Phi_{X,Y}\updownarrow & & \updownarrow\Phi_{X,Y'} \\ \mathrm{Hom}_{\mathbf{D}}(LX', Y) & \xrightarrow[-\circ Lf]{} & \mathrm{Hom}_{\mathbf{D}}(LX, Y) & \xrightarrow[g\circ -]{} & \mathrm{Hom}_{\mathbf{D}}(LX, Y') \end{array} \quad (193)$$

In a very informal sense, the bijections $\Phi_{X,Y}$ let us embed $\mathbf{C}$ in $\mathbf{D}$ and vice-versa in a compatible way, that is, morphisms between $X \in \mathbf{C}_0$ and $Y \in \mathbf{D}_0$ can be seen either by viewing X in $\mathbf{D}$ via L or viewing Y in $\mathbf{C}$ via R.[481]

[481] For the adjunction $\mathbf{Set} : (-)^* \dashv U : \mathbf{Mon}$, any set can be viewed as the monoid of words over it, and any monoid can be viewed as a set by forgetting the operation.

To make proofs go smoother, we will often use the superscript notation $(-)^{\mathrm{t}}$ to denote an application of a component of Φ or Φ^{-1}. That is, for any $X \in \mathbf{C}_0$ and $Y \in \mathbf{D}_0$, we have

$$(-)^{\mathrm{t}} : \mathrm{Hom}_{\mathbf{C}}(X, RY) \cong \mathrm{Hom}_{\mathbf{D}}(LX, Y) : (-)^{\mathrm{t}}.$$

We call f^{t} the **transpose** of f.[482]

[482] Unfortunately, the term *transpose* is probably inspired by matrix transposition, but I do not know of a technical way to realize one as an instance of the other. Some authors also write f^* or $f^\sharp$ for the transpose of f.

7.2 Results and Examples

There are a couple of very important results in this section (Theorem 441 and Theorem 446), but we will start slow.

We already proved in Proposition 419 that two left adjoints to the same functor must be isomorphic.[483] That proof used the first definition of left adjoints we saw with a natural family of universal morphisms. Let us prove the same thing, but relying on our two new definitions instead.[484]

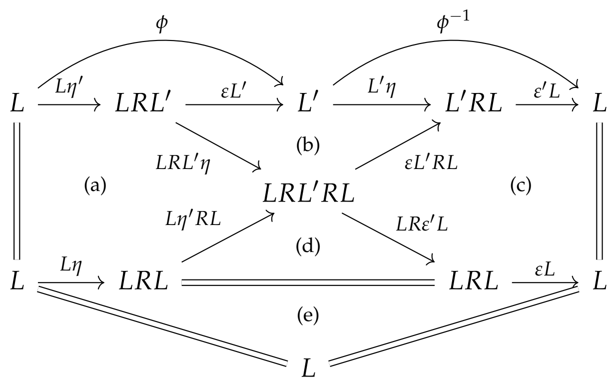
[483] With our new notation: if $L \dashv R$ and $L' \dashv R$, then $L \cong L'$, and dually if $L \dashv R$ and $L \dashv R'$, then $R \cong R'$.

[484] We omit the second item in Definition 426 because it is dual to the proof we already gave.

Proof of Proposition 419 via triangle identities. Let η and ε be the unit and counit of the adjunction $\mathbf{C} : L \dashv R : \mathbf{D}$, η' and ε' be those of $\mathbf{C} : L' \dashv R : \mathbf{D}$. Guided by the types, it is easy to compose the natural transformations we have to obtain two new natural transformations of type $L \Rightarrow L'$ and $L' \Rightarrow L$:

$$\phi = L \xRightarrow{L\eta'} LRL' \xRightarrow{\varepsilon L'} L' \quad \text{and} \quad \phi^{-1} = L' \xRightarrow{L'\eta} L'RL \xRightarrow{\varepsilon' L} L.$$

It remains to show ϕ^{-1} is the inverse of ϕ. We show $\phi^{-1} \circ \phi = \mathbb{1}_L$ by paving the following diagram (it lives in $[\mathbf{C}, \mathbf{D}]$).



(194)

Showing (194) commutes:

(a) Apply $L(-)'$ to $\mathsf{HOR}(\eta', \eta)$.

(b) By $\mathsf{HOR}(\varepsilon L', \eta)$ or $\mathsf{HOR}(\varepsilon, L'\eta)$.

(c) Apply $(-)L$ to $\mathsf{HOR}(\varepsilon, \varepsilon')$.

(d) Apply $L(-)L$ to the triangle identity (187) instantiated for η' and ε'.

(e) Apply the triangle identity (186) for η and ε.

We leave you to show $\phi \circ \phi^{-1}$ by paving a similar diagram (where L, η and ε swap roles with L', η' and ε'). $\square$

Proof of Proposition 419 via transposes. For any $X \in \mathbf{C}_0$, we define $\phi_X : LX \to L'X$ to be the image of $\mathrm{id}_{L'X} \in \mathrm{Hom}_{\mathbf{D}}(L'X, L'X)$ under the composition of the natural isomorphisms

$$\mathrm{Hom}_{\mathbf{D}}(L'X, L'X) \cong \mathrm{Hom}_{\mathbf{C}}(X, RL'X) \cong \mathrm{Hom}_{\mathbf{D}}(LX, L'X).$$

Then, for any $f : X \to Y$, the naturality squares in (195) imply $L'f \circ \phi_X = \phi_Y \circ Lf$.[485]

$$\begin{array}{ccccc} \mathrm{Hom}_{\mathbf{D}}(L'X, L'X) & \xrightarrow{L'f\circ-} & \mathrm{Hom}_{\mathbf{D}}(L'X, L'Y) & \xleftarrow{-\circ L'f} & \mathrm{Hom}_{\mathbf{D}}(L'Y, L'Y) \\ \updownarrow & & \updownarrow & & \updownarrow \\ \mathrm{Hom}_{\mathbf{C}}(X, RL'X) & \xrightarrow{RL'f\circ-} & \mathrm{Hom}_{\mathbf{C}}(X, RLY) & \xleftarrow{-\circ f} & \mathrm{Hom}_{\mathbf{C}}(Y, RLY) \\ \updownarrow & & \updownarrow & & \updownarrow \\ \mathrm{Hom}_{\mathbf{D}}(LX, L'X) & \xrightarrow[L'f\circ-]{} & \mathrm{Hom}_{\mathbf{D}}(LX, L'Y) & \xleftarrow[-\circ Lf]{} & \mathrm{Hom}_{\mathbf{D}}(LY, L'Y) \end{array} \tag{195}$$

[485] Start with $\mathrm{id}_{L'X}$ and $\mathrm{id}_{L'Y}$ at the top left and top right respectively and compare the results at the bottom middle.

We conclude that $\phi : L \Rightarrow L'$ is natural. With a symmetric argument, we construct $\phi^{-1} : L' \Rightarrow L$[486] and we check that they are inverses with (196) and (197).

[486] i.e.: ϕ_X^{-1} is the image of id_{LX} under

$$\mathrm{Hom}_{\mathbf{D}}(LX, LX) \cong \mathrm{Hom}_{\mathbf{C}}(X, RLX) \cong \mathrm{Hom}_{\mathbf{D}}(L'X, LX$$

$$\begin{array}{ccc} \mathrm{Hom}_{\mathbf{D}}(LX, LX) & \xrightarrow{\phi_X\circ-} & \mathrm{Hom}_{\mathbf{D}}(LX, L'X) \\ \updownarrow & & \updownarrow \\ \mathrm{Hom}_{\mathbf{D}}(L'X, LX) & \xrightarrow[\phi_X\circ-]{} & \mathrm{Hom}_{\mathbf{D}}(L'X, L'X) \end{array} \tag{196}$$

$$\begin{array}{ccc} \mathrm{Hom}_{\mathbf{D}}(L'X, L'X) & \xrightarrow{\phi_X^{-1}\circ -} & \mathrm{Hom}_{\mathbf{D}}(L'X, LX) \\ \updownarrow & & \updownarrow \\ \mathrm{Hom}_{\mathbf{D}}(LX, L'X) & \xrightarrow[\phi_X^{-1}\circ -]{} & \mathrm{Hom}_{\mathbf{D}}(LX, LX) \end{array} \tag{197}$$

Starting with id_{LX} in the top left of (196) and reaching the top right, we find that the image of $\phi_X \circ \phi_X^{-1}$ under the isomorphism is ϕ_X which is the image of $\mathrm{id}_{L'X}$, thus $\phi_X \circ \phi_X^{-1} = \mathrm{id}_{L'X}$. We proceed with a symmetric argument for (197). □

Of the three different proofs of Proposition 419, the second one using the triangle identities seems to be the quickest. You can judge for yourself which proof you prefer. In the rest of this chapter, we will see many examples of adjunctions and results about adjoint functors and try to have a balance between the different definitions we use.[487]

[487] We try to care about which definition is easiest to use.

We start with a converse to Proposition 419. When L has a right adjoint R and R' is isomorphic to R, then R' is also right adjoint to L.

SOL **Exercise 429.** Show that if $\mathbf{C} : L \dashv R : \mathbf{D}$ is an adjunction and $R \cong R'$, then $L \dashv R'$. State the dual statement and prove it.

Our main point in the introduction to this chapter was that grouping universal morphisms together as we did into an adjunction yields a notion of *global* universal construction. In particular, we can characterize when a category has *all* (co)limits of shape $\mathbf{J}$.

Theorem 430. *A category* $\mathbf{C}$ *has all limits of shape* $\mathbf{J}$ *if (and only if)*[488] *the functor* $\Delta^{\mathbf{J}}_{\mathbf{C}}$ *has a right adjoint.*

[488]

Proof. ($\Rightarrow$) For each diagram $F : \mathbf{J} \rightsquigarrow \mathbf{C}$, we pick (with the axiom of choice) a limit $\lim_{\mathbf{J}} F$ given by completeness and a universal morphism $\Delta^{\mathbf{J}}_{\mathbf{C}} \to F$ given by Theorem 416. By our argument at the start of the chapter, we get an adjunction $\Delta^{\mathbf{J}}_{\mathbf{C}} \dashv \lim_{\mathbf{J}}$.

($\Leftarrow$) Suppose $\mathbf{C} : \Delta^{\mathbf{J}}_{\mathbf{C}} \dashv L : [\mathbf{J}, \mathbf{C}]$ with unit η and let $F : \mathbf{J} \rightsquigarrow \mathbf{C}$ be a diagram. By definiton, $\eta_F : \Delta^{\mathbf{J}}_{\mathbf{C}} L(F) \to F$ is a universal morphism from $\Delta^{\mathbf{J}}_{\mathbf{C}}$ to F. Thus, by Theorem 416, $L(F)$ is the limit of F. □

Corollary 431 (Dual)**.** *A category* $\mathbf{C}$ *has all colimits of shape* $\mathbf{J}$ *if and only if the functor* $\Delta^{\mathbf{J}}_{\mathbf{C}}$ *has a left adjoint.*

We saw how families of universal morphisms give rise to an adjunction, so we could make our examples from Chapter 4 into adjunctions. Here, we carry out a similar but new example.

Example 432. Recall from Exercise 224 the maybe functor $- + \mathbf{1}$. Denote $\mathbf{1} = \{*\}$ for the terminal object of $\mathbf{Set}$. We consider a very similar functor $- + \mathbf{1} : \mathbf{Set} \rightsquigarrow \mathbf{Set}_*$ sending a set X to $(X + \mathbf{1}, *)$ and $f : X \to Y$ to $f + \mathrm{id}_{\mathbf{1}} : X + \mathbf{1} \to Y + \mathbf{1}$. In the other direction, we have the forgetful functor $U : \mathbf{Set}_* \rightsquigarrow \mathbf{Set}$ that forgets about the distinguished element of a pointed set. We claim that $- + \mathbf{1} \dashv U$.

First, for every set X, we need to define $\eta_X : X \to U((X + \mathbf{1}, *)) = X + \mathbf{1}$. The only obvious choice is to let η_X be the inclusion of X in $X + \mathbf{1}$ and one can check it makes η into a natural transformation $\mathrm{id}_{\mathbf{Set}} \Rightarrow U(- + \mathbf{1})$.

Check η and ε are natural:

$$\begin{array}{ccc} X & \xrightarrow{\eta_X} & X + \mathbf{1} \\ f\downarrow & & \downarrow f + \mathrm{id}_{\mathbf{1}} \\ Y & \xrightarrow[\eta_Y]{} & Y + \mathbf{1} \end{array} \qquad \begin{array}{ccc} (X, x) & \xrightarrow{\varepsilon_{(X,x)}} & (X + \mathbf{1}, *) \\ f\downarrow & & \downarrow f + \mathrm{id}_{\mathbf{1}} \\ (Y, y) & \xrightarrow[\varepsilon_{(Y,y)}]{} & (Y + \mathbf{1}, *) \end{array}$$

Second, for every pointed set (X, x), we need to define $\varepsilon_{(X,x)} : (X + \mathbf{1}, *) \to (X, x)$. Again, there is one clear choice, i.e.: acting like the identity on X and sending $*$ to x, we will denote $\varepsilon_{(X,x)} = [\mathrm{id}_X, * \mapsto x]$.

Finally, after checking the triangle identities which we instantiate below,[489] we conclude that $- + \mathbf{1} \dashv U$.

$$\begin{array}{ccc} (X+\mathbf{1},*) & \xrightarrow{\eta_X + \mathrm{id}_{\mathbf{1}}} & ((X+\mathbf{1})+\mathbf{1},\star) \\ & \searrow^{\mathrm{id}_{X+\mathbf{1}}} & \downarrow [\mathrm{id}_{X+\mathbf{1}}, \star \mapsto *] \\ & & (X+\mathbf{1},*) \end{array} \tag{198}$$

$$\begin{array}{ccc} X & \xrightarrow{\eta_X} & X+\mathbf{1} \\ & \searrow^{\mathrm{id}_X} & \downarrow [\mathrm{id}_X, * \mapsto x] \\ & & X \end{array} \tag{199}$$

[489] When dealing with a set $(X+\mathbf{1})+\mathbf{1}$, we will denote $*$ for the element of the inner $\mathbf{1}$ and $\star$ for the outer one.
In (199), $X = U(X,x)$.

A good exercise in categorical thinking is to generalize this example to an arbitrary category $\mathbf{C}$ with binary coproducts and a terminal object.[490]

[490] See ... for a solution.

Example 433 (Top). Let $U : \mathbf{Top} \rightsquigarrow \mathbf{Set}$ be the forgetful functor sending a topological space to its underlying set. We will find a left and a right adjoint to U.

Left adjoint: Fix a topological space (X,τ) and a set Y. We need to find a topological space (LY,λ) so that continuous functions $(LY,\lambda) \to (X,\tau)$ are in correspondence with functions $Y \to X$. It turns out there is a trivial topology that we can put on Y that makes any function $f : Y \to X$ continuous, it is called the **discrete topology** and contains all the subsets of Y.[491] We can check that any function $f : Y \to X$ is continuous relative to the discrete topology because for any open set $U \in \tau$, $f^{-1}(U)$ is a subset of Y and hence it is open in $(Y, \mathcal{P}(Y))$. After checking that sending Y to $(Y,\mathcal{P}(Y))$ and $f : Y \to Y'$ to $f : (Y,\mathcal{P}(Y)) \to (Y', \mathcal{P}(Y'))$ is a functor, we denote it disc, we find can conclude that disc $\dashv U$.

[491] It is clear that the set of all subsets of Y is a topology because any union or intersection of subsets is still a subset.

Right adjoint: Fix a topological space (X,τ) and a set Y. We need to find a topological space (LY,λ) so that continuous functions $(X,\tau) \to (LY,\lambda)$ are in correspondence with functions $X \to Y$. Again, there is a trivial topology that we can put on Y that makes any function $f : X \to Y$ continuous, it is called the **codiscrete topology** and contains only the empty set and the full space Y.[492] We can check that any function $f : X \to Y$ is continuous relative to the codiscrete topology because the $f^{-1}(\emptyset) = \emptyset$ and $f^{-1}(Y) = X$ must be open by the definition of a topology. After checking that sending Y to $(Y, \{\emptyset, Y\})$ and $f : Y \to Y'$ to $f : (Y,\{\emptyset,Y\}) \to (Y', \{\emptyset, Y'\})$ is a functor, we denote it codisc, we can conclude that $U \dashv$ codisc.

[492] Since $\emptyset \cap Y = \emptyset$ and $\emptyset \cup Y$, we conclude that $\{\emptyset, Y\}$ is closed under any union and intersection, hence it is a topology.

We found our first chain of adjunctions disc $\dashv U \dashv$ codisc. Another interesting one is $\mathrm{colim}_{\mathbf{J}} \dashv \Delta^{\mathbf{J}}_{\mathbf{C}} \dashv \lim_{\mathbf{J}}$ in a category $\mathbf{C}$ with all limits and colimits of shape $\mathbf{J}$. A less interesting one is $\cdots \dashv \mathrm{id}_{\mathbf{C}} \dashv \mathrm{id}_{\mathbf{C}} \dashv \mathrm{id}_{\mathbf{C}} \dashv \cdots$. Here is a chain of five adjunctions.

SOL Exercise 434. Let $\mathbf{C}$ be a category and $\mathsf{id}, \mathsf{s}, \mathsf{t}$ be the functors described in Exercise 323. Show they are related by the adjunctions $\mathsf{t} \dashv \mathsf{id} \dashv \mathsf{s}$. Suppose furthermore that $\mathbf{C}$ has an initial object $\emptyset$ and a terminal object $\mathbf{1}$. Show that the constant functor at $\mathrm{id}_{\emptyset}$ is left adjoint to t and the constant functor at $\mathrm{id}_{\mathbf{1}}$ is right adjoint to s.

As a final example, we show that any equivalence gives rise to two adjunctions. In this sense[493], one can see a left (resp. right) adjoint to a functor F as an approximation to a left (resp. right) inverse that is even coarser than a quasi-inverse.[494]

[493] And in another sense related to Kan extensions.

[494] Furthermore, it follows from Proposition 419 (resp. Corollary 421) that the left (resp. right) adjoint of F is the left (resp. right) inverse or quasi-inverse when the latter exists.

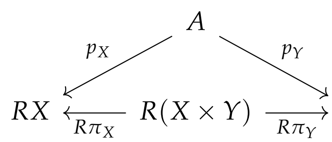
Proposition 435. *Let $L : \mathbf{C} \rightsquigarrow \mathbf{D}$ and $R : \mathbf{D} \rightsquigarrow \mathbf{C}$ be quasi-inverses, then $L \dashv R$ and $R \dashv L$.*

Proof. It is enough to show $L \dashv R$ as the definition of quasi-inverses is symmetric. □

Proposition 436. *Let $\mathbf{C} : L \dashv R : \mathbf{D}$ be adjoint functors and $X, Y \in \mathbf{D}_0$. If $X \times Y$ exists, then $R(X \times Y)$ with the projections $R(\pi_X)$ and $R(\pi_Y)$ is the product $R(X) \times R(Y)$.*[495]

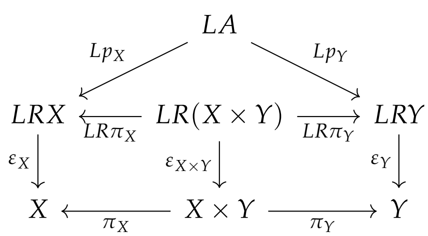
[495] In other words, right adjoints preserve binary products.

Proof. Let $p_X : A \to RX$ and $p_Y : A \to RY$ be such that (200) commutes.



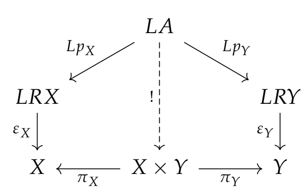
(200)

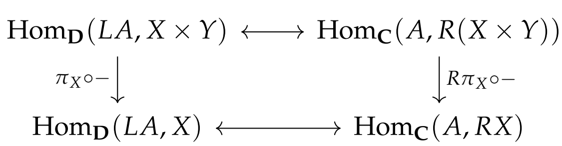
We need to show there is a unique mediating morphism $A \to R(X \times Y)$. First, we will get rid of the applications of R at the bottom, in order to use the universal property of the product $X \times Y$. To do this, we apply L to (200) and use the counit $\varepsilon : LR \Rightarrow \mathrm{id}_{\mathbf{D}}$ to obtain (201).



(201)

The universal property of $X \times Y$ tells us there is a unique $! : LA \to X \times Y$ such that $\pi_X \circ ! = \varepsilon_X \circ Lp_X$ and $\pi_Y \circ ! = \varepsilon_Y \circ Lp_Y$. We claim that $!^{\mathrm{t}}$ is the mediating morphism of (200), i.e.: $R\pi_X \circ !^{\mathrm{t}} = p_X$ and $R\pi_Y \circ !^{\mathrm{t}} = p_Y$. Using the adjunction $L \dashv R$, we obtain the following commutative square.



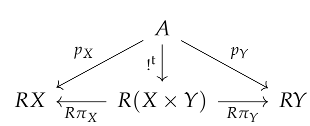


(202)

Now, starting with ! on the top left corner, we obtain the following derivation.

$$
\begin{aligned}
p_X &= {p_X}^{\mathrm{t}\mathrm{t}} \\
&= (\varepsilon_X \circ Lp_X)^{\mathrm{t}} \\
&= (\pi_X \circ !)^{\mathrm{t}} && \text{definition of } ! \\
&= R\pi_X \circ !^{\mathrm{t}} && \text{commutativity of (202)}
\end{aligned}
$$

Replacing X with Y in the previous argument shows $!^{\mathrm{t}}$ makes (203) commute. For the uniqueness, note that if $m : A \to R(X \times Y)$ can replace $!^{\mathrm{t}}$, then (204) commutes



(203)

which implies by uniqueness of ! that $m^{\mathsf{t}} = \varepsilon_{X\times Y} \circ Lm = \,!$. Transposing yields $!^{\mathsf{t}} = m$.

$$\begin{array}{ccccc} & & LA & & \\ & \swarrow^{Lp_X} & \downarrow^{Lm} & \searrow^{Lp_Y} & \\ LRX & \xleftarrow[LR\pi_X]{} & LR(X\times Y) & \xrightarrow[LR\pi_Y]{} & LRY \\ \downarrow^{\varepsilon_X} & & \downarrow^{\varepsilon_{X\times Y}} & & \downarrow^{\varepsilon_Y} \\ X & \xleftarrow[\pi_X]{} & X\times Y & \xrightarrow[\pi_Y]{} & Y \end{array} \tag{204}$$

□

Corollary 437 (Dual). *Let* $\mathbf{C} : L \dashv R : \mathbf{D}$ *be adjoint functors and* $A, B \in \mathbf{C}_0$. *If* $A + B$ *exists, then* $L(A+B)$ *with the coprojections* $L\kappa_A$ *and* $L\kappa_B$ *is the coproduct* $LA \times LB$.[496]

[496] In other words, left adjoints preserve binary coproducts.

Proposition 438. *Let* $\mathbf{C} : L \dashv R : \mathbf{D}$ *be adjoint functors. If* $g : X \to Y \in \mathbf{D}_1$ *is monic, then* $R(g)$ *is monic.*[497]

[497] In other words, right adjoints preserve monomorphisms.

Proof. Let $h_1, h_2 : Z \to R(X)$ be such that $R(g) \circ h_1 = R(g) \circ h_2$, we need to show that $h_1 = h_2$. Since $L \dashv R$, we have the following commutative square.

$$\begin{array}{ccc} \mathrm{Hom}_{\mathbf{C}}(Z, RX) & \longleftrightarrow & \mathrm{Hom}_{\mathbf{D}}(LZ, X) \\ {\scriptstyle Rg\circ -}\downarrow & & \downarrow{\scriptstyle g\circ -} \\ \mathrm{Hom}_{\mathbf{C}}(Z, RY) & \longleftrightarrow & \mathrm{Hom}_{\mathbf{D}}(LZ, Y) \end{array} \tag{205}$$

Starting with h_1 and h_2 in the top left corner, we find that[498]

$$g \circ {h_1}^{\mathsf{t}} = (Rg \circ h_1)^{\mathsf{t}} = (Rg \circ h_2)^{\mathsf{t}} = g \circ {h_2}^{\mathsf{t}},$$

[498] The first and last equality follow from commutativity of (205) and the middle equality is a hypothesis.

which, by monicity of g implies ${h_1}^{\mathsf{t}} = {h_2}^{\mathsf{t}}$. This in turn means that $h_1 = h_2$ because $(-)^{\mathsf{t}}$ is a bijection. □

Corollary 439 (Dual). *Let* $\mathbf{C} : L \dashv R : \mathbf{D}$ *be adjoint functors. If* $f : A \to B \in \mathbf{C}_1$ *is epic, then* $L(f)$ *is epic.*[499]

[499] In other words, left adjoints preserve epimorphisms.

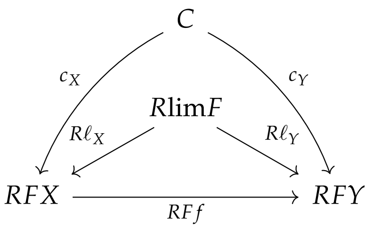
Remark 440. We want to put the emphasis on a crucial step in the proof above which was to derive $g \circ {h_1}^{\mathsf{t}} = (Rg \circ h_1)^{\mathsf{t}}$ from (205). By varying the arguments slightly (i.e.: going around the square in another direction or considering the naturality square involving pre-composition), we cook up four similar equations that can be helpful.[500]

[500] For instance, (207) was a crucial step in the proof of Proposition 436: we used (202) to derive $(\pi_X \circ !)^{\mathsf{t}} = R\pi_X \circ !^{\mathsf{t}}$.

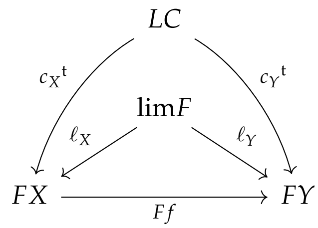
$$\forall g : X \to Y, f : Z \to RX, \qquad g \circ f^{\mathsf{t}} = (Rg \circ f)^{\mathsf{t}} \tag{206}$$
$$\forall g : X \to Y, f : LZ \to X, \qquad (g \circ f)^{\mathsf{t}} = Rg \circ f^{\mathsf{t}} \tag{207}$$
$$\forall g : LX \to Y, f : Z \to X, \qquad g^{\mathsf{t}} \circ f = (g \circ Lf)^{\mathsf{t}} \tag{208}$$
$$\forall g : X \to RY, f : Z \to X, \qquad (g \circ f)^{\mathsf{t}} = g^{\mathsf{t}} \circ Lf \tag{209}$$

Theorem 441. *Right adjoints are continuous.*

Proof. Let $\mathbf{C} : L \dashv R : \mathbf{D}$ be an adjunction and $F : \mathbf{J} \rightsquigarrow \mathbf{D}$ be a diagram in $\mathbf{D}$ whose limit cone is $\{\ell_X : \lim F \to FX\}_{X \in \mathbf{J}_0}$. We claim that $\{R\ell_X : R\lim F \to RFX\}_{\mathbf{J}_0}$ is the limit cone of $R \circ F$. For any other cone making (210) commute for any $f : X \to Y \in \mathbf{J}_1$, we can apply transposition to the c_X's to obtain (211) which commutes by (206).[501]

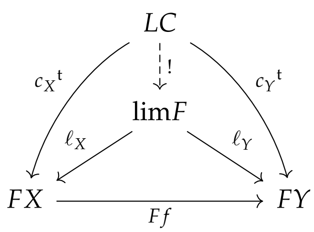


(210)

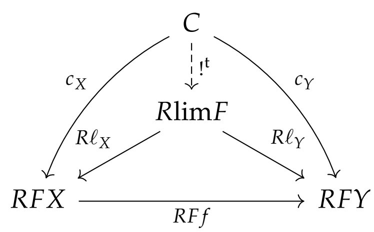


(211)

By the universal property of $\lim F$, there is a unique mediating morphism $! : LC \to \lim F$ making (212) commute. Transposing ! yields a mediating morphism making (213) commutes by (207).[502]



(212)



(213)

Finally, $!^{\mathrm{t}}$ is the only mediating morphism that fits in (213) because if $m : C \to R\lim F$ fits, then $m^{\mathrm{t}} : LC \to \lim F$ fits in (212)[503] and by uniqueness of !, $m^{\mathrm{t}} = !$ which further implies $m = !^{\mathrm{t}}$. □

[501] In (206), putting $g := Ff$ and $f := c_X$, we obtain

$$c_Y{}^{\mathrm{t}} = (RFf \circ c_X)^{\mathrm{t}} = Ff \circ c_X{}^{\mathrm{t}}.$$

[502] In (207), putting $g := \ell_X$ and $f := !$, we obtain

$$c_X = (c_X{}^{\mathrm{t}})^{\mathrm{t}} = (\ell_X \circ !)^{\mathrm{t}} = R\ell_X \circ !^{\mathrm{t}}.$$

Symmetrically, we have

$$c_Y = (c_Y{}^{\mathrm{t}})^{\mathrm{t}} = (\ell_Y \circ !)^{\mathrm{t}} = R\ell_Y \circ !^{\mathrm{t}}.$$

[503] Suppose $R\ell_X \circ m = c_X$, then we use (206) to conclude

$$c_X{}^{\mathrm{t}} = (R\ell_X \circ m)^{\mathrm{t}} = \ell_X \circ m^{\mathrm{t}},$$

and similarly for Y.

Corollary 442 (Dual)**.** *Left adjoints are cocontinuous.*

Theorem 443. *If* $\mathbf{C} : L \dashv R : \mathbf{D}$ *and* $\mathbf{D} : L' \dashv R' : \mathbf{E}$ *are two adjunctions, then* $\mathbf{C} : L'L \dashv RR' : \mathbf{E}$ *is an adjunction.*[504]

[504] This theorem is often referred to as *adjunctions can be composed.*

Proof. Let η and ε be the unit and counit of the first adjunction and η' and ε' be the unit and counit of the second one. We define the following unit and counit for the composite adjunction:

$$\widehat{\eta} = R\eta' L \cdot \eta : \mathrm{id}_{\mathbf{C}} \Rightarrow RR'L'L$$
$$\widehat{\varepsilon} = \varepsilon' \cdot L'\varepsilon R' : L'LRR' \Rightarrow \mathrm{id}_{\mathbf{E}}.$$

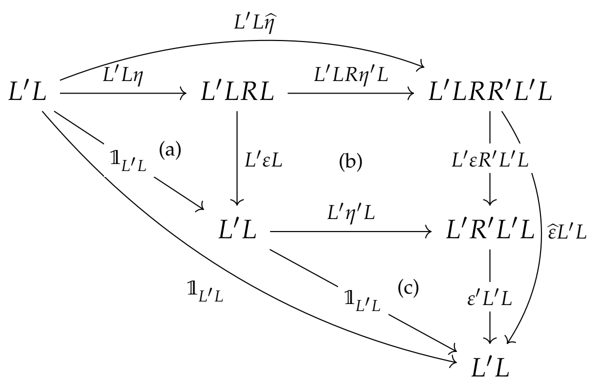
The following diagrams show the triangle identities.

Showing (214) commutes:

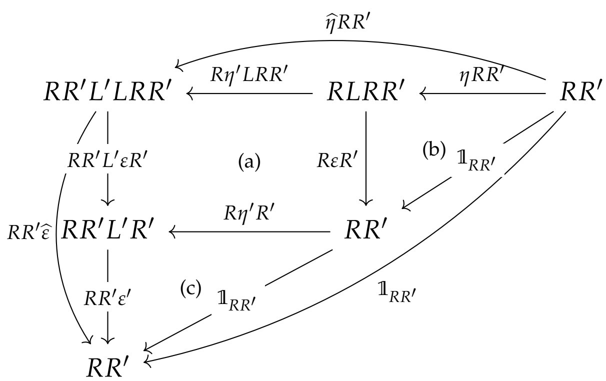
(a) Apply $L'(-)$ to the left triangle identity of η and ε.

(b) Apply $L'(-)L$ to $\mathsf{HOR}(\varepsilon, \eta')$.

(c) Apply $(-)L$ to the left triangle identity of



(214)



(215)

Showing (215) commutes:

(a) Apply $R(-)R'$ to $\mathsf{HOR}(\eta', \varepsilon)$.

(b) Apply $(-)R'$ to the right triangle identity of η and ε.

(c) Apply $R(-)$ to the right triangle identity of η' and ε'.

□

Proposition 444. *If* $\mathbf{D} : L \dashv R : \mathbf{E}$ *is an adjunction, then there is an adjunction* $[\mathbf{C}, \mathbf{D}] : (L \circ -) \dashv (R \circ -) : [\mathbf{C}, \mathbf{E}]$.

Proof. We simplify the notation a little bit by writing $L-$ and $R-$ instead of $L \circ -$ and $R \circ -$ respectively. First, we can see that $L-$ and $R-$ are functors by Exercise 372,[505] they send a natural transformation $\phi : F \Rightarrow G$ to $L\phi$ and $R\phi$ respectively. Composing them yields $RL- : [\mathbf{C}, \mathbf{D}] \rightsquigarrow [\mathbf{C}, \mathbf{D}]$ and $LR- : [\mathbf{C}, \mathbf{E}] \rightsquigarrow [\mathbf{C}, \mathbf{E}]$. Let $\eta : \mathrm{id}_{\mathbf{D}} \Rightarrow RL$ and $\varepsilon : LR \Rightarrow \mathrm{id}_{\mathbf{E}}$ be the unit and counit of $L \dashv R$. We claim that $\eta- = F \mapsto \eta F$ and $\varepsilon- = G \mapsto \varepsilon G$ are the unit and counit of an adjunction $L- \dashv R-$.

[505] They are compositions:

$$L- = (- \circ -) \circ (\Delta(L) \times \mathrm{id}_{[\mathbf{C},\mathbf{D}]})$$
$$R- = (- \circ -) \circ (\Delta(R) \times \mathrm{id}_{[\mathbf{C},\mathbf{E}]}).$$

Alternatively, we can use Example 373.5 where we described currying for functors. In that setting, we have

$$L- = \Lambda(- \circ -)(L)$$
$$R- = \Lambda(- \circ -)(R).$$

To see that $\eta-$ and $\varepsilon-$ are natural transformations of the right type, we can recognize them in the image of $\Lambda(- \circ -)$ (noting that $\mathrm{id}_{\mathbf{D}}- = \mathrm{id}_{[\mathbf{C},\mathbf{D}]}$ and $\mathrm{id}_{\mathbf{E}}- = \mathrm{id}_{[\mathbf{C},\mathbf{E}]}$):

$$\eta- = \Lambda(- \circ -)(\eta) : \mathrm{id}_{[\mathbf{C},\mathbf{D}]} \Rightarrow RL-$$
$$\varepsilon- = \Lambda(- \circ -)(\varepsilon) : LR- \Rightarrow \mathrm{id}_{[\mathbf{C},\mathbf{E}]}.$$

It is left to show the triangle identities hold assuming they hold for η and ε. In the following derivations, we use three simple facts:[506]

[506] They can be shown by proving the equality at each component.

- the biaction of $F-$ and $G-$ on $\phi-$ yields $(F\phi G)-$,
- $(\phi-) \cdot (\phi'-) = (\phi \cdot \phi')-$, and

- $(\mathbb{1}_F)- = \mathbb{1}_{F-}$.

Now, the triangle identities hold by:

$$(\varepsilon-)(L-)\cdot(L-)(\eta-) = (\varepsilon L-)\cdot(L\eta-) = (\varepsilon L\cdot L\eta)- = (\mathbb{1}_L)- = \mathbb{1}_{L-}$$
$$(R-)(\varepsilon-)\cdot(\eta-)(R-) = (R\varepsilon-)\cdot(\eta R-) = (R\varepsilon\cdot\eta R)- = (\mathbb{1}_R)- = \mathbb{1}_{R-}.$$

□

Corollary 445 (Dual)**.** *If* $\mathbf{D} : L \dashv R : \mathbf{E}$ *is an adjunction, then there is an adjunction* $[\mathbf{C},\mathbf{D}] : -L \dashv -R : [\mathbf{C},\mathbf{E}]$.

Theorem 446. *Let* $\mathbf{D}$ *be a category with all limits of shape* $\mathbf{J}$*. For any category* $\mathbf{C}$*, the functor category* $[\mathbf{C},\mathbf{D}]$ *has all limits of shape* $\mathbf{J}$ *and the limit of any diagram* $F : \mathbf{J} \rightsquigarrow [\mathbf{C},\mathbf{D}]$ *satisfies for any* $X \in \mathbf{C}_0$, $(\lim_{\mathbf{J}} F)(X) = \lim_{\mathbf{J}}(F(-)(X))$.[507]

Proof. From previous results, we have the following chain of adjunctions.

$$[\mathbf{C},\mathbf{D}] \underset{\lim_{\mathbf{J}}\circ -}{\overset{\Delta^{\mathbf{J}}_{\mathbf{D}}\circ -}{\underset{\longleftarrow}{\overset{\longrightarrow}{\perp}}}} [\mathbf{C},[\mathbf{J},\mathbf{D}]] \underset{\Lambda}{\overset{\Lambda^{-1}}{\underset{\longleftarrow}{\overset{\longrightarrow}{\perp}}}} [\mathbf{C}\times\mathbf{J},\mathbf{D}] \underset{-\circ\mathsf{swap}^{-1}}{\overset{-\circ\mathsf{swap}}{\underset{\longleftarrow}{\overset{\longrightarrow}{\perp}}}} [\mathbf{J}\times\mathbf{C},\mathbf{D}] \underset{\Lambda^{-1}}{\overset{\Lambda}{\underset{\longleftarrow}{\overset{\longrightarrow}{\perp}}}} [\mathbf{J},[\mathbf{C},\mathbf{D}]] \quad (216)$$

From left to right. The first adjunction is induced by Proposition 444 and the adjunction $\Delta^{\mathbf{J}}_{\mathbf{D}} \dashv \lim_{\mathbf{J}}$ given by completeness of $\mathbf{D}$. The second adjunction is obtained from Proposition 435 and the fact that Λ and Λ^{-1} are inverses. The third adjunction is induced by Corollary 445 and the canonical isomorphism $\mathsf{swap} : \mathbf{C}\times\mathbf{J} \rightsquigarrow \mathbf{J}\times\mathbf{C}$.[508] The fourth adjunction is similar to the second one.

There is a simpler way to describe the composition of the three rightmost adjunctions. If we view a functor $F : \mathbf{C} \rightsquigarrow [\mathbf{J},\mathbf{D}]$ as taking two arguments and write it $F(-_1)(-_2)$, the composition $\Lambda \circ (-\circ \mathsf{swap}) \circ \Lambda^{-1}$ (the top path) swaps the order of the arguments to yield the functor $F(-_2)(-_1) : \mathbf{J} \rightsquigarrow [\mathbf{C},\mathbf{D}]$. The bottom path swaps back the arguments.

Next, we show that the composition of the top path is $\Delta^{\mathbf{J}}_{[\mathbf{C},\mathbf{D}]}$. Starting with a functor $F : \mathbf{C} \rightsquigarrow \mathbf{D}$, the first left adjoint sends it to $\Delta^{\mathbf{J}}_{\mathbf{D}} \circ F$ which sends $X \in \mathbf{C}_0$ to the constant functor at FX and $f : X \to Y \in \mathbf{C}_1$ to the natural transformation whose components are all $Ff : FX \to FY$. Applying the three other left adjoints, we obtain a functor which sends any $j \in \mathbf{J}_0$ to the functor F and any $m : j \to j' \in \mathbf{J}_1$ to $\mathbb{1}_F$. We conclude that the top path sends F to the constant functor at F.

We obtain a right adjoint to $\Delta^{\mathbf{J}}_{[\mathbf{C},\mathbf{D}]}$ by composing all the right adjoins in (216) with Theorem 443 and thus $[\mathbf{C},\mathbf{D}]$ has all limits of shape $\mathbf{J}$. To compute them, we can compose the right adjoints in (216) to find $(\lim_{\mathbf{J}} F)(X) = \lim_{\mathbf{J}}(F(-)(X))$. □

Corollary 447 (Dual)**.** *Let* $\mathbf{D}$ *be a category with all colimits of shape* $\mathbf{J}$*. For any category* $\mathbf{C}$*, the functor category* $[\mathbf{C},\mathbf{D}]$ *has all colimits of shape* $\mathbf{J}$ *and the colimit of any diagram* $F : \mathbf{J} \rightsquigarrow [\mathbf{C},\mathbf{D}]$ *satisfies for any* $X \in \mathbf{C}_0$, $(\mathrm{colim}_{\mathbf{J}} F)(X) = \mathrm{colim}_{\mathbf{J}}(F(-)(X))$.[509]

Corollary 448. *If a category* $\mathbf{D}$ *is (finitely) complete or cocomplete, then so is* $[\mathbf{C},\mathbf{D}]$ *for any category* $\mathbf{C}$.

[507] This means limits in functor categories are taken pointwise, just like we proved in Theorem 353

[508] One could also see that $-\circ\mathsf{swap}$ and $-\circ\mathsf{swap}^{-1}$ are inverses.

[509] In other words, colimits are taken pointwise. You can use Exercise 352 or draw a similar chain of adjunctions as in (216).

SOL **Exercise 449.** Let $\mathbf{C}$ have all limits of shape $\mathbf{J}$ and $\mathbf{C} : L \dashv R : \mathbf{D}$ be an adjunction. Using Theorem 430, Corollary 421, Theorem 443 and Proposition 444, show that R preserves all limits of shape $\mathbf{J}$.

8 Monads and Algebras

8.1 POV: Category Theory 151
8.2 POV: Universal Algebra 161
8.3 POV: Computer Programs 165
8.4 Exercises 169

8.1 POV: Category Theory

We will start from the concept of an adjunction which, as we hope was made clear in the previous chapter, is ubiquitous and powerful throughout mathematics. However, we will start with a great oversimplification; we will assume the categories concerned are posetal.

An adjunction between posets $(P, \leq)$ and $(Q, \sqsubseteq)$ is a pair of order-preserving functions $L : P \to Q$ and $R : Q \to P$ satisfying for any $p \in P$ and $q \in Q$, $L(p) \sqsubseteq q \iff p \leq R(q)$. You might recognize this as a Galois connection from Chapter 0, this explains the notation $L \dashv R$ we introduced back then.

Let us derive again the properties of the composite $R \circ L$ using what we know about adjoints.[510]

It is of course a monotone function but we can derive a couple of additional properties. First, the existence of the unit $\eta : \mathrm{id}_P \Rightarrow RL$ means that for any $p \in P$, there is $\eta_p : p \to RL(p)$, so RL is extensive.[511] Second, the existence of the counit $\varepsilon : RL \Rightarrow \mathrm{id}_P$ means that for any $p \in P$, there is $R(\varepsilon_{L(p)}) : RLRL(p) \to RL(p)$ and $RL(\eta_p) : RL(p) \to RLRL(p)$, so RL is idempotent (i.e.: $\forall p \in P, RL(p) = RLRL(p)$). This means RL is a closure operator.

We will generalize this discussion to arbitrary categories now. Let $\mathbf{C} : L \dashv R : \mathbf{D}$ be an adjoint pair, we have two natural transformations $\eta : \mathrm{id}_{\mathbf{C}} \Rightarrow RL$ and $R\varepsilon L : RLRL \Rightarrow RL$ that interact well together due to the triangle identities. Applying $R(-)$ to (186) and $(-)L$ to (187) yields two diagrams that we combine into (217). We can add to the diagram coming from $\mathsf{HOR}(\varepsilon, \varepsilon)$ which act on by $R(-)L$ to obtain (218).

$$\begin{array}{ccccc} RL & \xrightarrow{RL\eta} & RLRL & \xleftarrow{\eta RL} & RL \\ & {\scriptstyle \mathbb{1}_{RL}}\searrow & \downarrow{\scriptstyle R\varepsilon L} & \swarrow{\scriptstyle \mathbb{1}_{RL}} & \\ & & RL & & \end{array} \qquad (217)$$

$$\begin{array}{ccc} RLRLRL & \xrightarrow{R\varepsilon LRL} & RLRL \\ {\scriptstyle RLR\varepsilon L}\downarrow & & \downarrow{\scriptstyle R\varepsilon L} \\ RLRL & \xrightarrow[R\varepsilon L]{} & RL \end{array} \qquad (218)$$

These diagrams are precisely what is required to define a monad.

Definition 450 (Monad)**.** A **monad** is a triple comprised of an endofunctor $M : \mathbf{C} \rightsquigarrow \mathbf{C}$ and two natural transformations $\eta : \mathrm{id}_{\mathbf{C}} \Rightarrow M$ and $\mu : M^2 \Rightarrow M$ called the **unit** and **multiplication** respectively that make (219) and (220) commute in $[\mathbf{C}, \mathbf{C}]$.

[510] Recall that we showed $R \circ L$ was a closure operator in Proposition 68.

[511] i.e.: $\forall p \in P, p \leq RL(p)$.

$$\begin{array}{ccccc} M & \xrightarrow{M\eta} & M^2 & \xleftarrow{\eta M} & M \\ & \searrow^{\mathbb{1}_M} & \downarrow \mu & {}^{\mathbb{1}_M}\swarrow & \\ & & M & & \end{array} \qquad (219) \qquad\qquad \begin{array}{ccc} M^3 & \xrightarrow{\mu M} & M^2 \\ {\scriptstyle M\mu}\downarrow & & \downarrow{\scriptstyle \mu} \\ M^2 & \xrightarrow[\mu]{} & M \end{array} \qquad (220)$$

Examples 451. Our discussion above tells us that any adjoint pair $L \dashv R$ corresponds to a monad $(RL, \eta, R\varepsilon L)$, so all the examples of adjunctions you have seen correspond to suitable examples of monads. For instance, all closure operators are monads. Here are more examples described from adjunctions in Chapter 7.

1. The adjunction $\mathbf{Set} : (-)^* \dashv U : \mathbf{Mon}$ yields the free monoid monad abusively denoted $(-)^* : \mathbf{Set} \rightsquigarrow \mathbf{Set}$ sending a set A to the underlying set of the free monoid on A. The unit sends $a \in A$ to the word $a \in A^*$ by inclusion and the multiplication sends a finite word over finite words over A to the concatenation of the words.[512]

2. Similarly to the previous example, there is monad $k[-]$ on $\mathbf{Set}$ sending A to the underlying set of the vector space $k[A]$.[513]

3.

4. Both adjunctions with the forgetful functor $\mathbf{Top} \rightsquigarrow \mathbf{Set}$ induce the identity monad.

[512] e.g.: it sends `(aa)(ab)(bb)` to `aaabbb`.

[513] We leave you to figure out the unit and multiplication depending on your preferred way to construct $k[A]$ (either as polynomials over variables in A or functions from A to k).

Examples 452. Here, we describe three simple yet very useful examples and let you ponder on the adjunctions they might or might not originate from.

1. Suppose $\mathbf{C}$ has (binary) coproducts and a terminal object $\mathbf{1}$, then $(- + \mathbf{1}) : \mathbf{C} \rightsquigarrow \mathbf{C}$ is a monad.[514] We write inl^{X+Y} (resp. inr^{X+Y}) for the coprojection of X (resp. Y) into $X+Y$.[515] First, note that for a morphism $f : X \to Y$,

$$f + \mathbf{1} = [\mathsf{inl}^{Y+\mathbf{1}} \circ f, \mathsf{inr}^{Y+\mathbf{1}}] : X + \mathbf{1} \to Y + \mathbf{1}.$$

The components of the unit are given by the coprojections, i.e.: $\eta_X = \mathsf{inl}^{X+\mathbf{1}} : X \to X + \mathbf{1}$, and the components of the multiplication are

$$\mu_X = [\mathsf{inl}^{X+\mathbf{1}}, \mathsf{inr}^{X+\mathbf{1}}, \mathsf{inr}^{X+\mathbf{1}}] : X + \mathbf{1} + \mathbf{1} \to X + \mathbf{1}.$$

Checking that (219) commutes, we have for any $X \in \mathbf{C}$:

$$\begin{aligned} \mu_X \circ (\eta_X + \mathbf{1}) &= [\mu_X \circ \mathsf{inl}^{(X+\mathbf{1})+\mathbf{1}} \circ \eta_X, \mu_X \circ \mathsf{inr}^{(X+\mathbf{1})+\mathbf{1}}] \\ &= [[\mathsf{inl}^{X+\mathbf{1}}, \mathsf{inr}^{X+\mathbf{1}}] \circ \mathsf{inl}^{X+\mathbf{1}}, \mathsf{inr}^{X+\mathbf{1}}] \\ &= [\mathsf{inl}^{X+\mathbf{1}}, \mathsf{inr}^{X+\mathbf{1}}] \\ &= \mathrm{id}_{X+\mathbf{1}} \\ &= [\mathsf{inl}^{X+\mathbf{1}}, \mathsf{inr}^{X+\mathbf{1}}] \\ &= \mu_X \circ \mathsf{inl}^{(X+\mathbf{1})+\mathbf{1}} \\ &= \mu_X \circ \eta_{X+\mathbf{1}} \end{aligned}$$

[514] It is called the **maybe monad**. It is a generalization of the maybe functor defined in Exercise 224 and you may want to generalize the adjunction described in Example 432 to this setting before going to the next section.

[515] These notations are very common in the community of programming language research, they stand for *injection left* (resp. *right*). We may omit the superscript in case it is too cumbersome.

For (220), we have for any $X \in \mathbf{C}$:

$$\begin{aligned}
\mu_X \circ (\mu_X + \mathbf{1}) &= [\mu_X \circ \mathsf{inl}^{(X+\mathbf{1})+\mathbf{1}} \circ \mu_X, \mu_X \circ \mathsf{inr}^{(X+\mathbf{1})+\mathbf{1}}] \\
&= [[\mathsf{inl}^{X+\mathbf{1}}, \mathsf{inr}^{X+\mathbf{1}}] \circ \mu_X, \mathsf{inr}^{X+\mathbf{1}}] \\
&= [[\mathsf{inl}^{X+\mathbf{1}}, \mathsf{inr}^{X+\mathbf{1}}, \mathsf{inr}^{X+\mathbf{1}}], \mathsf{inr}^{X+\mathbf{1}}] \\
&= [\mu_X, \mathsf{inr}^{X+\mathbf{1}}] \\
&= [[\mathsf{inl}^{X+\mathbf{1}}, \mathsf{inr}^{X+\mathbf{1}}], \mathsf{inr}^{X+\mathbf{1}}, \mathsf{inr}^{X+\mathbf{1}}] \\
&= [\mu_X \circ \mathsf{inl}^{(X+\mathbf{1})+\mathbf{1}}, \mu_X \circ \mathsf{inr}^{(X+\mathbf{1})+\mathbf{1}}, \mu_X \circ \mathsf{inr}^{(X+\mathbf{1})+\mathbf{1}}] \\
&= \mu_X \circ \mu_{X+\mathbf{1}}
\end{aligned}$$

2. The covariant powerset functor $\mathcal{P} : \mathbf{Set} \rightsquigarrow \mathbf{Set}$ is a monad with the following unit and multiplication:

$$\eta_X : X \to \mathcal{P}(X) = x \mapsto \{x\} \text{ and } \mu_X : \mathcal{P}(\mathcal{P}(X)) \to \mathcal{P}(X) = F \mapsto \bigcup_{s \in F} s.$$

Checking that (219) commutes, we have for any $S \subseteq \mathcal{P}(X)$:

$$\begin{aligned}
\mu_X(\mathcal{P}(\eta_X)(S)) &= \mu_X\left(\{\{x\} \mid x \in S\}\right) \\
&= \bigcup_{x \in S} \{x\} \\
&= S \\
&= \bigcup \{S\} \\
&= \mu_X(\{S\}) \\
&= \mu_X(\eta_{\mathcal{P}(X)}(S))
\end{aligned}$$

For (220), we have for any $\mathcal{F} \in \mathcal{P}(\mathcal{P}(\mathcal{P}(X))$:

$$\begin{aligned}
\mu_X(\mu_{\mathcal{P}(X)}(\mathcal{F})) &= \mu_X\left(\bigcup_{F \in \mathcal{F}} F\right) \\
&= \bigcup_{\substack{s \in \mathcal{P}(X) \\ \exists F \in \mathcal{F}, s \in F}} s \\
&= \{x \in X \mid \exists s \in \mathcal{P}(X), x \in s \text{ and } \exists F \in \mathcal{F}, s \in F\} \\
&= \bigcup_{F \in \mathcal{F}} \bigcup_{s \in F} s \\
&= \mu_X\left(\left\{\bigcup_{s \in F} s \mid F \in \mathcal{F}\right\}\right) \\
&= \mu_X(\mathcal{P}(\mu_X)(\mathcal{F}))
\end{aligned}$$

3. The functor $\mathcal{D} : \mathbf{Set} \to \mathbf{Set}$ sends a set X to the set of finitely supported distributions on X, i.e.:

$$\mathcal{D}(X) := \{\varphi \in [0,1]^X \mid \sum_{x \in X} \varphi(x) = 1 \text{ and } \varphi(x) \neq 0 \text{ for finitely many } x\text{'s}\}.$$

It sends a function $f : X \to Y$ to the function between distributions

$$\lambda\varphi^{\mathcal{D}(X)}.\lambda y^Y.\varphi(f^{-1}(y)).$$

More verbosely, the weight of $\mathcal{D}(f)(\varphi)$ at point y is equal to the total weight of φ on the preimage of y under f. It is a monad with unit $\eta_X = x \mapsto \delta_x$, where δ_x is the Dirac distribution at x (all the weight is at x), and multiplication

$$\mu_X = \Phi \mapsto \lambda x^X. \sum_{\phi \in \operatorname{supp}(\Phi)} \Phi(\phi) \cdot \phi(x),$$

where $\operatorname{supp}(\Phi)$ is the support of Φ, i.e.: $\operatorname{supp}(\Phi) := \{\varphi \mid \Phi(\varphi) \neq 0\}$.

After looking long enough for adjunctions giving rise to the monads in Examples 452, two questions dare to be asked. Does every monad arise from an adjunction in the same way as above? If yes, is that adjunction unique?

The second question might not be as natural to novices in category theory but it is almost as important as the first one. Indeed, uniqueness is a very strong property and if every monad had a unique corresponding adjunction, one might expect it to be fairly easy to find. This is part of the beauty of category theory. We are working with very little data M, η and μ so if it completely determined an adjunction $L \dashv R$ with its unit and counit and the natural isomorphism $\mathrm{Hom}(L-,-) \cong \mathrm{Hom}(-,R-)$, it could not do so in a very convoluted way merely because there is not that many ways to manipulate the original data.

In any case, we will respectively give a positive and negative answer to these questions. Fortunately, while we might not benefit from the power of uniqueness, there are two special adjunctions arising from a monad whose descriptions are fairly straightforward. In the order we present them, the first is due to Kleisli and the second to Eilenberg and Moore. In the rest of this section, (M, η, μ) will be a monad on a category $\mathbf{C}$.

Kleisli Category $\mathbf{C}_M$

An intuitive way to think about monads is through the idea of **generalized elements**.[516] Given an object $A \in \mathbf{C}_0$, we can view MA as extending A with more *general* or *structured* elements built from A.

[516] This is not a formal term.

In this picture, the morphisms $\eta_A : A \to MA$ give a way to understand anything inside A trivially as a general element of A. The morphisms $\mu_A : M^2A \to MA$ imply that higher order structures can be collapsed so that generalized elements over generalized elements of A are generalized elements of A. The functoriality of M implies that the new structures in MA are somewhat independent of A. Indeed, for every morphisms $f : A \to B$, there is a morphism $Mf : MA \to MB$ which, by naturality of η ($Mf(\eta_A) = \eta_B(f)$), acts just like f on the trivial generalization of elements in A. Commutativity of (219) says that the trivial generalization[517] of a generalized element is indeed trivial, namely, after collapsing via μ, we end up with what we started with. Finally, the associativity of μ (i.e.: commutativity of (220))

[517] There are two ways to do it corresponding to the L.H.S. and R.H.S. of (219).

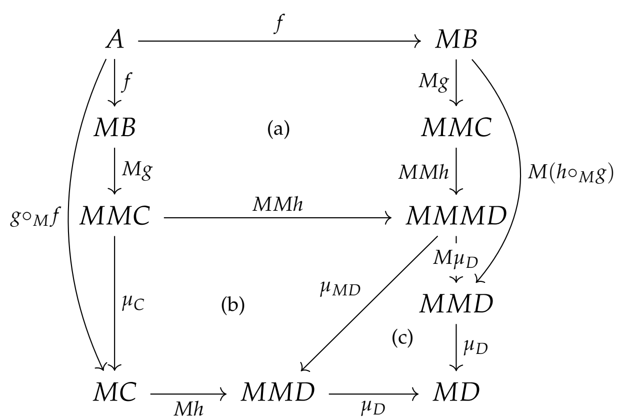
corresponds to the fact that in higher order of generalizations, one can collapse the structure at every level in any order and end up with the same thing.

Now, we can also consider **generalized morphisms**. Let us say we were given an ill-defined morphism $f : A \to B$ that sends some of the stuff in A outside of B. One way to fix this might be to consider general elements of B and see f as a morphism $A \to MB$. We will call such morphisms **Kleisli morphisms** and write $f : A \twoheadrightarrow B$ for $f : A \to MB$.[518]

[518] Another common notation for Kleisli morphisms is $f : A \rightsquigarrow B$ but this clashes with our notation for functors.

With an arbitrary functor F, you might have a hard time to come up with a way to compose two Kleisli morphisms $A \to FB$ and $B \to FC$ or even define the identity Kleisli morphism $A \to FA$, but the data of a monad lets you do just that. Indeed, given $f : A \twoheadrightarrow B$ and $g : B \twoheadrightarrow C$, while g is not composable with f, Mg is so we have $Mg \circ f : A \to MMC$ and it suffices to apply the multiplication μ_C to obtain $\mu_C \circ Mg \circ f : A \twoheadrightarrow C$. We denote $g \circ_M f := \mu_C \circ Mg \circ f$ and call it the **Kleisli composition**. Also, for any $A \in \mathbf{C}_0$, the component of the unit at A yields a Kleisli morphism $\eta_A : A \twoheadrightarrow A$. Let us check that $\circ_M$ is associative and that η_A behaves like the identity with respect to $\circ_M$.

Let $f : A \twoheadrightarrow B$, $g : B \twoheadrightarrow C$ and $h : C \twoheadrightarrow D$ be Kleisli morphisms, the compositions $h \circ_M (g \circ_M f)$ and $(h \circ_M g) \circ_M f$ are respectively the bottom and top path of the following commutative diagram, so we conclude that $\circ_M$ is associative.



(221)

Showing (221) commutes:
(a) Trivial.
(b) $\mathsf{NAT}(\mu, C, MD, h)$.
(c) Components of (220) at D.

We show that $\eta_B \circ_M f = f$ and $f \circ_M \eta_A = f$ with the following derivations.

$$\begin{aligned} \eta_B \circ_M f &= \mu_B \circ M\eta_B \circ f \\ \text{by L.H.S. of (219)} &= \mathrm{id}_{MB} \circ f \\ &= f \end{aligned} \qquad \begin{aligned} f \circ_M \eta_A &= \mu_B \circ Mf \circ \eta_A \\ \text{by } \mathsf{NAT}(\eta, A, MB, f) &= \mu_B \circ \eta_{MB} \circ f \\ \text{by R.H.S. of (219)} &= \mathrm{id}_{MB} \circ f \\ &= f \end{aligned}$$

This leads to the definition of the category $\mathbf{C}_M$.[519]

[519] Notice that we had to use all the data from the monad: the naturality of η and μ, the commutativity of both diagrams (219) and (220) as well as functoriality of M (the latter was used implicitly).

Definition 453 ($\mathbf{C}_M$). Let $\mathbf{C}$ be a category and (M, η, μ) a monad on $\mathbf{C}$. The **Kleisli category** of M, denoted $\mathbf{C}_M$[520], has the same objects as $\mathbf{C}$ and the morphisms in $\mathrm{Hom}_{\mathbf{C}_M}(A, B)$ are the elements of $\mathrm{Hom}_{\mathbf{C}}(A, MB)$. The identity for $A \in \mathbf{C}_0$ is $\eta_A : A \to MA$ and composition is $\circ_M$.

[520] Some authors denote it $\mathrm{Kl}(M)$.

Examples 454. We describe the Kleisli category for the monads in Examples 452.

1. By identifying a Kleisli morphism $f : A \nrightarrow B$ with a partial function $A \rightharpoonup B$ as we did in Example 379.3, we can show that $\mathbf{Set}_{-+1} \cong \mathbf{Par}$.

2. In $\mathbf{Set}_{\mathcal{P}}$, objects are sets and morphisms are functions $r : X \to \mathcal{P}(Y)$. Viewing the latter as a relation $R \subseteq X \times Y$ defined by $(x,y) \in R \Leftrightarrow y \in r(x)$, we can verify that composition of relations corresponds to Kleisli composition in $\mathbf{Set}_{\mathcal{P}}$.[521]

   Let $r : X \to \mathcal{P}(Y)$ and $s : Y \to \mathcal{P}(Z)$ be Kleisli morphisms, R, S and SR be the relations corresponding to r, s and $s \circ_{\mathcal{P}} r$. We need to show $SR = S \circ R$. Fix $x \in X$, we have

$$(s \circ_{\mathcal{P}} r)(x) = (\mu_Z^{\mathcal{P}} \circ \mathcal{P}(s) \circ r)(x) = \bigcup \mathcal{P}(s)(r(x)) = \{z \in Z \mid \exists y \in r(x), z \in s(y)\}.$$

   Since $y \in r(x) \Leftrightarrow (x,y) \in R$ and $z \in s(y) \Leftrightarrow (y,z) \in S$, we conclude that

$$(x,z) \in SR \Leftrightarrow z \in (s \circ_{\mathcal{P}} r)(x) \Leftrightarrow (x,z) \in S \circ R.$$

   After a bit more administrative arguments, one finds that $\mathbf{Set}_{\mathcal{P}} \cong \mathbf{Rel}$.

3.

[521] Composition of relations was defined in Example 115.

Since we can view any object of $\mathbf{C}$ as an object of $\mathbf{C}_M$, we may wonder if we can do the same with morphisms to obtain a functor $\mathbf{C} \rightsquigarrow \mathbf{C}_M$. The key idea is to view $f : A \to B$ as a generalized morphism by trivially generalizing its target, that is, by post-composing with η_B. We claim that $F_M : \mathbf{C} \rightsquigarrow \mathbf{C}_M$ acting as identity on objects and post-composing by components of η on morphisms is a functor.[522] Indeed, $F_M(\mathrm{id}_A) = \eta_A$ is the identity on A in $\mathbf{C}_M$ and

$$\begin{aligned}
F_M(g \circ f) &= \eta_C \circ g \circ f \\
&= Mg \circ \eta_B \circ f && \mathsf{NAT}(\eta, B, C, g) \\
&= Mg \circ \mu_B \circ M(\eta_B) \circ \eta_B \circ f && \text{by (219)} \\
&= \mu_C \circ MMg \circ M(\eta_B) \circ \eta_B \circ f && \mathsf{NAT}(\mu, B, C, g) \\
&= \mu_C \circ M(\eta_C) \circ Mg \circ \eta_B \circ f && M\mathsf{NAT}(\eta, B, C, g) \\
&= F_M(g) \circ_M F_M(f). && \text{def. of } \circ_M
\end{aligned}$$

[522] Explicitly, for any $A \in \mathbf{C}_0$, $F_M(A) = A$ and for any $f : A \to B$, $F_M(f) = \eta_B \circ f$.

We will now construct a right adjoint $U_M : \mathbf{C}_M \rightsquigarrow \mathbf{C}$ to F_M. Given A and B objects of both $\mathbf{C}$ and $\mathbf{C}_M$, the Kleisli morphisms from $F_M A$ to B are precisely the morphisms in $\mathbf{C}$ from A to MB, thus we infer that the identity function is an isomorphism $\mathrm{Hom}_{\mathbf{C}_M}(F_M A, B) \cong \mathrm{Hom}_{\mathbf{C}}(A, MB)$. This implies U_M sends B to MB and we can define U_M on morphisms by imposing the naturality of the aforementioned isomorphism. Given $g : A \nrightarrow B$, starting with η_A on the top left of (222), we find that $U_M g \circ \eta_A = g$ which implies $U_M g = \mu_B \circ Mg$.[523]

$$\begin{array}{ccc}
\mathrm{Hom}_{\mathbf{C}_M}(A, A) & \xleftrightarrow{\mathrm{id}} & \mathrm{Hom}_{\mathbf{C}}(A, MA) \\
{\scriptstyle g \circ_M (-)} \downarrow & & \downarrow {\scriptstyle U_M g \circ (-)} \\
\mathrm{Hom}_{\mathbf{C}_M}(A, B) & \xleftrightarrow[\mathrm{id}]{} & \mathrm{Hom}_{\mathbf{C}}(A, MB)
\end{array} \qquad (222)$$

[523] This implication is subtle. While it is true that we do not yet know if another f satisfies $f \circ \eta_A = g$. Once we know (in a few moments) defining $U_M g = \mu_B \circ Mg$ yields an adjunction $F_M \dashv U_M$ whose unit is η, we know that η_A is universal and uniqueness of $U_M g$ follows.

As a sanity check (and for a bit of practice), let us verify U_M is a functor. For any $A \in \mathbf{C}_{M0}$, $U_M(\eta_A) = \mu_A \circ M(\eta_A) = \mathrm{id}_A$ by the L.H.S. of (219) and for any for any $f : A \nrightarrow B$ and $g : B \nrightarrow C$,

$$\begin{aligned}
U_M(g \circ_M f) &= U_M(\mu_C \circ Mg \circ f) \\
&= \mu_C \circ M(\mu_C \circ Mg \circ f) \\
&= \mu_C \circ M(\mu_C) \circ MMg \circ Mf \\
&= \mu_C \circ \mu_{MC} \circ MMg \circ Mf && \text{by (220)} \\
&= \mu_C \circ Mg \circ \mu_B \circ Mf && \text{by naturality of } \mu \\
&= U_M(g) \circ U_M(f).
\end{aligned}$$

Let us now verify that $F_M \dashv U_M$. Let $A, B \in \mathbf{C}_0$ (we view B as an object of $\mathbf{C}_M$), we saw that the identity function is an isomorphism $\mathrm{Hom}_{\mathbf{C}_M}(F_M A, B) \cong \mathrm{Hom}_{\mathbf{C}}(A, U_M B)$ and we now check it is natural. We need to show (223) commutes for any $f : A' \to A$ and $g : B \nrightarrow B'$. It follows from this derivation starting with $k : A \nrightarrow B$ in the top left.

$$\begin{array}{ccc}
\mathrm{Hom}_{\mathbf{C}_M}(A, B) & \xleftrightarrow{\mathrm{id}} & \mathrm{Hom}_{\mathbf{C}}(A, MB) \\
{\scriptstyle g \circ (-) \circ_M F_M f} \downarrow & & \downarrow {\scriptstyle U_M g \circ (-) \circ f} \\
\mathrm{Hom}_{\mathbf{C}_M}(A', B') & \xleftrightarrow[\mathrm{id}]{} & \mathrm{Hom}_{\mathbf{C}}(A', MB')
\end{array} \quad (223)$$

$$\begin{aligned}
g \circ_M k \circ_M F_M f &= \mu_{B'} \circ M(g) \circ \mu_B \circ M(k) \circ \eta_A \circ f \\
&= \mu_{B'} \circ M(g) \circ \mu_B \circ \eta_{MB} \circ k \circ f && \text{by naturality of } \eta \\
&= \mu_{B'} \circ M(g) \circ \mathrm{id}_{MB} \circ k \circ f && \text{by (219)} \\
&= \mu_{B'} \circ M(g) \circ k \circ f \\
&= U_M g \circ k \circ f
\end{aligned}$$

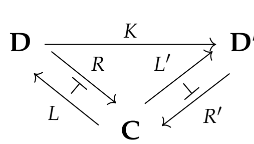
Finally, in order to achieve our initial goal of finding an adjunction that induces the original monad, we need to make sure the monad arising from $F_M \dashv U_M$ is (M, η, μ). First, we check that $U_M F_M = M$. On objects, it is clear. On a morphism $f : A \to B$, we have

$$U_M(F_M(f)) = U_M(\eta_B \circ f) = \mu_B \circ M(\eta_B) \circ Mf \overset{(219)}{=} Mf.$$

Next, as η_A is the image of the identity on A in $\mathbf{C}_M$ under the natural isomorphismcomponent, the unit of the adjunction is the unit of the monad. The counit of the adjunction at A is $\varepsilon_A = \mathrm{id}_{MA}$, thus $(U_M \varepsilon F_M)_A = U_M(\mathrm{id}_{F_M A}) = \mu_A \circ M(\mathrm{id}_{MA}) = \mu_A$.

Recall that we claimed $F_M \dashv U_M$ was special in some way and that this was the (informal) reason why it was relatively easy to find, the next proposition will make this precise.

Definition 455 (Adj_M)**.** Let $\mathbf{C}$ be a category and (M, η, μ) a monad on $\mathbf{C}$. The **category of adjunctions inducing** M is denoted Adj_M. Its objects are adjoint pairs $L \dashv R$ with unit η and counit ε sastisfying $R \circ L = M$ $R\varepsilon L = \mu$. Its morphisms $L \dashv R \to L' \dashv R'$) are functors K satisfying $K \circ L = L'$ and $R' \circ K = R$ as in (224).

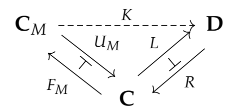


(224)

We can restate the end result of the discussion above as $F_M \dashv U_M$ being an object of Adj_M. It is special because it is initial.

Proposition 456. *The adjunction $F_M \dashv U_M$ is initial in* Adj_M.

Proof. Let $\mathbf{C} : L \dashv R : \mathbf{D} \in \mathrm{Adj}_M$ with unit η and counit ε, we claim there is a unique functor $K : \mathbf{C}_M \rightsquigarrow \mathbf{D}$ satisfying $K \circ F_M = L$ and $R \circ K = U_M$ as in (225).



(225)

On objects, K is determined by $KA = KF_MA = LA$. To a morphism $f : A \nrightarrow B$, we need to assign a morphism in $Kf \in \mathrm{Hom}_{\mathbf{D}}(LA, LB)$ such that $RKf = U_Mf = \mu_B \circ Mf = R\varepsilon_{LB} \circ RLf$. It is clear that $Kf = \varepsilon_{LB} \circ Lf$ is a candidate but to show it is unique, we consider the following naturality square coming from the adjunction $L \dashv R$.

$$\begin{array}{ccc} \mathrm{Hom}_{\mathbf{D}}(LA, LA) & \xrightarrow{R-\circ\eta_A} & \mathrm{Hom}_{\mathbf{C}}(A, RLA) \\ {\scriptstyle Kf\circ(-)}\downarrow & & \downarrow{\scriptstyle RKf\circ(-)} \\ \mathrm{Hom}_{\mathbf{D}}(LA, LB) & \xleftarrow[\varepsilon_{LB}\circ L-]{} & \mathrm{Hom}_{\mathbf{C}}(A, RLB) \end{array} \tag{226}$$

Starting with id_{LA} in the top left and reaching the bottom left, we find

$$\begin{aligned} Kf &= \varepsilon_{LB} \circ LRKf \circ L\eta_A \\ &= \varepsilon_{LB} \circ LR\varepsilon_{LB} \circ LRLf \circ L\eta_A && \text{hypothesis on } RKf \\ &= \varepsilon_{LB} \circ LR\varepsilon_{LB} \circ L\eta_{RLB} \circ Lf && \mathsf{NAT}(\eta, A, RLB, f) \\ &= \varepsilon_{LB} \circ \varepsilon_{LRLB} \circ L\eta_{RLB} \circ Lf && \mathsf{HOR}(\varepsilon, \varepsilon)L \\ &= \varepsilon_{LB} \circ \varepsilon_{LMB} \circ L\eta_{MB} \circ Lf && RL = M \\ &= \varepsilon_{LB} \circ \mathrm{id}_{MB} \circ Lf && \text{triangle identity} \\ &= \varepsilon_{LB} \circ Lf \end{aligned}$$

To finish the proof, let us verify K is functorial.

$$K(u_{\mathbf{C}_M}(A)) = K(\eta_A) = \varepsilon_{LB} \circ L(\eta_A) \overset{(186)}{=} \mathrm{id}_A$$

$$\begin{aligned} K(g \circ_M f) &= K(\mu_C \circ RLg \circ f) \\ &= \varepsilon_{LC} \circ L(\mu_C) \circ LRLg \circ Lf \\ &= \varepsilon_{LC} \circ LR\varepsilon_{LC} \circ LRLg \circ Lf && \text{by hypothesis on } \varepsilon \\ &= \varepsilon_{LC} \circ \varepsilon_{LRLC} \circ LRLg \circ Lf && \mathsf{HOR}(\varepsilon, \varepsilon)L \\ &= \varepsilon_{LC} \circ Lg \circ \varepsilon_{LB} \circ Lf && \mathsf{NAT}(\varepsilon, LB, LRLC, Lg) \\ &= Kg \circ Kf \end{aligned}$$

□

SOL Exercise 457. Let $K : L \dashv R \to L' \dashv R'$ be a morphism in Adj_M, ε and ε' be the counits of the source and target respectively. Show that $K\varepsilon = \varepsilon' K$.

Eilenberg–Moore Category $\mathbf{C}^M$

For the second solution to the problem of finding an adjunction inducing a given monad, we look at the more structural side of monads.

Definition 458 (M–algebra)**.** Let (M, η, μ) be a monad, an **Eilenberg–Moore algebra** for M or simply M–algebra is a pair (A, α) consisting of an object $A \in \mathbf{C}_0$ and a morphism $\alpha : MA \to A$ such that (227) and (228) commute.

$$\begin{array}{ccc} A & \xrightarrow{\eta_A} & MA \\ & \searrow^{\mathrm{id}_A} & \downarrow \alpha \\ & & A \end{array} \quad (227) \qquad \begin{array}{ccc} M^2A & \xrightarrow{\mu_A} & MA \\ M\alpha \downarrow & & \downarrow \alpha \\ MA & \xrightarrow[\alpha]{} & A \end{array} \quad (228)$$

We will often denote an M–algebra using only its underlying object or its underlying morphism.

Definition 459 (Homomorphism)**.** Let (M, η, μ) be a monad and (A, α) and (B, β) be two M–algebras. An M–algebra **homomorphism** or simply M–homomorphism from (A, α) to (B, β) is a morphism $h : A \to B$ making (229) commute.

$$\begin{array}{ccc} MA & \xrightarrow{Mh} & MB \\ \alpha \downarrow & & \downarrow \beta \\ A & \xrightarrow[h]{} & B \end{array} \quad (229)$$

After checking that the composition of two M–homomorphisms is an M–homomorphism and id_A is an M–homomorphism from (A, α) to itself whenever α is an M–algebra, we get a category of M–algebras and M–homomorphism called the **Eilenberg–Moore category** of M and denoted $\mathbf{C}^M$.

Since $\mathbf{C}^M$ was built from objects and morphisms in $\mathbf{C}$, there is an obvious forgetful functor $U^M : \mathbf{C}^M \rightsquigarrow \mathbf{C}$ sending an M–algebra (A, α) to its underlying object A and an M–homomorphism to its underlying morphism. We will now find a left adjoint $F^M : \mathbf{C} \rightsquigarrow \mathbf{C}^M$ to U^M. Since we want this adjunction to induce the monad M, we require that $U^M F^M = M$. It means F^M must send $A \in \mathbf{C}_0$ to an M–algebra on MA and $h \in \mathbf{C}_1$ to Mh. There is straightforward choice given to us by the data of M, that is, $F^M A = (MA, \mu_A : MMA \to MA)$ and it turns out naturality of μ yields commutativity of

$$\begin{array}{ccc} M^2A & \xrightarrow{M^2h} & M^2B \\ \mu_A \downarrow & & \downarrow \mu_B \\ MA & \xrightarrow[Mh]{} & MB \end{array}, \quad (230)$$

which implies Mh is indeed an M–homomorphism. Because M is a functor, we immediately obtain that F^M is a functor. We now show that $F^M \dashv U^M$ with unit η and counit ε satisfying $U^M \varepsilon F^M = \mu$.

Let us define the counit and verify the triangle identities. For an M–algebra $\alpha : MA \to A$, we want an M–homomorphism $\varepsilon_\alpha : F^M U^M A = (MA, \mu_A) \to (A, \alpha)$. Again, we have a straightforward choice since α, being an M–algebra, satisfies $\alpha \circ$

$\mu_A = \alpha \circ M\alpha$, hence we can set $\varepsilon_\alpha = \alpha$. The following derivations show the triangle identities hold.

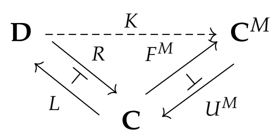
$$\varepsilon_{F^M A} \circ F^M \eta_A = \varepsilon_{\mu_A} \circ M\eta_A = \mu_A \circ M\eta_A = \mathrm{id}_{MA} = \mathrm{id}_{F^M A}$$
$$U^M \varepsilon_\alpha \circ \eta_{U^M(A,\alpha)} = \alpha \circ \eta_A = \mathrm{id}_A = \mathrm{id}_{U^M(A,\alpha)}$$

Lastly, we verify
$$U^M(\varepsilon_{F^M A}) = U^M(\varepsilon_{\mu_A}) = U^M(\mu_A) = \mu_A,$$
and we conclude $F^M \dashv U^M$ is an object of Adj_M.

Dually to Proposition 456, we show that this adjunction is special in a precise way.

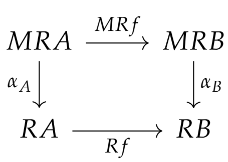
Proposition 460. *The adjunction* (F^M, U^M) *is terminal in* Adj_M.

Proof. Let $\mathbf{C} : L \dashv R : \mathbf{D} \in \mathrm{Adj}_M$ with unit η and counit ε, we claim there is a unique functor $K : \mathbf{D} \rightsquigarrow \mathbf{C}^M$ satisfying $K \circ L = F^M$ and $U^M \circ K = R$ as in (231).

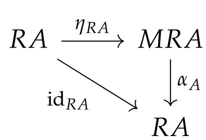


(231)

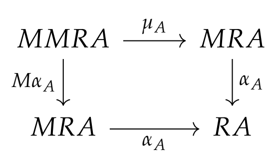
As before, we can determine K by the equation $U^M K = R$ which means it sends $A \in \mathbf{D}_0$ to an M–algebra on RA and $f : A \to B \in \mathbf{D}_1$ to an M–homomorphism $Rf : KA \to KB$. The only missing piece of this puzzle is the algebra structure on KA. We have two clues. First, Rf is an M–homomorphism, i.e.: denoting $KA = (RA, \alpha_A)$ and $KB = (RB, \alpha_B)$, we must ensure (232) commutes. Second, (KA, α_A) is an M–algebra, so (233) and (234) commute.



(232)



(233)



(234)

Replacing M with RL, we recognize the first diagram as a naturality square showing α is a natural transformation $RLR \Rightarrow R$ and the two other diagrams yield

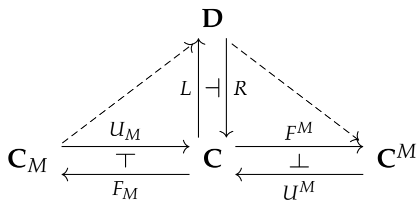
$$\alpha \cdot \eta R = \mathbb{1}_R \quad \text{and} \quad \alpha \cdot RL\alpha = \alpha \cdot \mu.$$

Moreover, we can see that $\alpha_A = R\varepsilon_A$ makes (233) commute by a triangle identity. This candidate also makes (232) commute because $R\varepsilon_A$ is a natural transformation and (234) commute because

$$\begin{aligned}
R\varepsilon_A \circ \mu_A &= R\varepsilon_A \circ R\varepsilon_{LA} && R\varepsilon L = \mu \\
&= R(\varepsilon_A \circ \varepsilon_{LA}) && \text{functoriality of } R \\
&= R(\varepsilon_A \circ LR(\varepsilon_A)) && \mathsf{HOR}(\varepsilon, \varepsilon) \\
&= R\varepsilon_A \circ MR\varepsilon_A && RL = M.
\end{aligned}$$

To verify uniqueness, recall that the counit of the adjunction $F^M \dashv U^M$ sends an M–algebra (X, x) to the M–homomorphism $x : (MX, \mu_X) \to (X, x)$. Thus, α_A is the result of applying the counit to KA and by Exercise **??**, we have $\alpha_A = K\varepsilon_A = R\varepsilon_A$. As K acts like R on morphisms, it is obviously functorial. $\square$

The following picture summarizes the last two sections.



(235)

With the following two results, one can see the Kleisli category inside the Eilenberg–Moore category as the full subcategory of free algebras.

SOL **Exercise 461.** Show that the unique morphism $F_M \dashv U_M \to F^M \dashv U^M$ is the functor $\mathbf{C}_M \rightsquigarrow \mathbf{C}^M$ sending $A \in \mathbf{C}_0$ to (MA, μ_A) and $f : A \twoheadrightarrow B$ to $\mu_B \circ Mf$.

Proposition 462. *The functor $\mathbf{C}_M \rightsquigarrow \mathbf{C}^M$ of Exercise 461 is fully faithful.*

Proof. **Full:** Suppose $g : MA \to MB$ is such that $g \circ \mu_A = \mu_B \circ Mg$, then

$$\mu_B \circ M(g \circ \eta_A) = \mu_B \circ Mg \circ M\eta_A = g \circ \mu_A \circ M\eta_A = g,$$

so g is the image of $g \circ \eta_A$ in $\mathbf{C}_M$.

Faithful: Suppose $\mu_B \circ Mg = \mu_B \circ Mf$, then pre-composing with η_A, we find that $f = f \circ_M \eta_A = g \circ_M \eta_A = g$. $\square$

8.2 POV: Universal Algebra

In this section, we will highlight the link between algebraic structures as you have encountered them in other classes with the Eilenberg–Moore algebras discussed above. We will only work over the category **Set**.[524] We start by developing an example.

[524] The ideas of universal algebra have be developed in other settings like enriched categories.

Example 463 ($\mathcal{P}_{\mathrm{ne}}$)**.** Consider the non-empty finite powerset functor $\mathcal{P}_{\mathrm{ne}}$ sending X to $\{S \in \mathcal{P}(X) \mid S \text{ is finite and non-empty}\}$. The same unit and multiplication as defined for $\mathcal{P}$ make $\mathcal{P}_{\mathrm{ne}}$ into a monad.[525] A $\mathcal{P}_{\mathrm{ne}}$–algebra is a function $\alpha : \mathcal{P}_{\mathrm{ne}}(A) \to A$ satisfying the equations $\alpha\{a\} = a$ and $\alpha(\mathcal{P}_{\mathrm{ne}}(\alpha)(S)) = \alpha(\bigcup S)$. From this, we can extract a binary operation $\oplus_\alpha : A \times A \to A$ by defining $x \oplus_\alpha y = \alpha\{x, y\}$. This operation is clearly commutative and idempotent,[526] but it is also associative by the following derivation.

[525] It is easy to see as the η and μ restrict to finite and non-empty.

[526] i.e.: $x \oplus_\alpha y = y \oplus_\alpha y$ and $x \oplus_\alpha x = x$.

$$\begin{aligned}(x \oplus_\alpha y) \oplus_\alpha z &= \alpha\{x, y\} \oplus_\alpha z \\ &= \alpha\{\alpha\{x, y\}, z\} \\ &= \alpha\{\alpha\{x, y\}, \alpha\{z\}\}\end{aligned}$$

$$\begin{aligned}&= \alpha\{\mathcal{P}_{\text{ne}}\alpha\{\{x,y\},\{z\}\}\}\\&= \alpha\{\mu_A\{\{x,y\},\{z\}\}\}\\&= \alpha\{x,y,z\}.\end{aligned}$$

Since a $\mathcal{P}_{\text{ne}}$–homomorphism $h : (A,\alpha) \to (B,\beta)$ commutes with α and β it also commutes with $\oplus_\alpha$ and $\oplus_\beta$.[527]

[527] i.e.: $h(a \oplus_\alpha a') = h(a) \oplus_\beta h(a')$.

Conversely, if $\oplus$ is an idempotent, associative and commutative binary operation on A, we can define $\alpha_\oplus$ on non-empty finite sets of A by iterating $\oplus$. Namely,

$$\alpha_\oplus\{x\} = x \oplus x \quad \text{and} \quad \alpha_\oplus\{x_1,\dots,x_n\} = x_1 \oplus x_2 \oplus \cdots \oplus x_n.$$

It is well-defined by associativity and commutativity and we can check that it is the inverse of the operation described in the previous paragraph. That is to say, we can check that $\alpha_{\oplus_\alpha} = \alpha$ and $\oplus_{\alpha_\oplus} = \oplus$. For the former, it is clear for singleton sets and for any $n > 1$, we have the following derivation.

$$\begin{aligned}\alpha_{\oplus_\alpha}\{x_1,\dots,x_n\} &= x_1 \oplus_\alpha \cdots \oplus_\alpha x_n\\&= \alpha\{x_1, x_2 \oplus_\alpha \cdots \oplus_\alpha x_n\}\\&= \vdots\\&= \alpha\{x_1,\alpha\{x_2,\alpha\{\cdots,\alpha\{x_n\}\}\}\}\\ \text{using } \alpha\circ\mathcal{P}_{\text{ne}}(\alpha) = \alpha\circ\mu_A &= \alpha\{x_1,x_2,\alpha\{\cdots,\alpha\{x_n\}\}\}\\&= \vdots\\&= \alpha\{x_1,\dots,x_n\}\end{aligned}$$

For the latter, we have

$$x \oplus_{\alpha_\oplus} y = \alpha_\oplus\{x,y\} = x \oplus y.$$

A set equipped with an idempotent, commutative and associative binary operation is called a **semilattice**[528] and we have shown above that $\mathcal{P}_{\text{ne}}$–algebras are in correspondence with semilattices. Through the introduction of basic notions in universal algebra, we will explain how this correspondence is functorial and generalize the core idea behind it.

[528] A semilattice can also be called a sup-semilattice, join-semilattice, inf-semilattice or meet-semilattice. This is because a semilattice can also be defined as a poset where all supremums/joins (resp. , infimums/meets) exist.

Definition 464 (Algebraic theory)**.** An **algebraic signature**[529] is a set Σ of operation symbols along with **arities** in $\mathbb{N}$, we denote $f : n \in \Sigma$ for an n–ary operation symbol f in Σ. Given a set X, one constructs the set of **Σ–terms** with variables in X, denoted $T_\Sigma(X)$ by iterating operations symbols:

[529] Also called algebraic similarity type.

$$\forall x \in X,\, x \in T_\Sigma(X)$$
$$\forall t_1,\dots,t_n \in T_\Sigma(X), f : n \in \Sigma,\, f(t_1,\dots,t_n) \in T_\Sigma(X).$$

An **equation**[530] E over Σ is a pair of Σ–terms over a set of dummy variables which we usually denote with an equality sign (e.g.: $s = t$ for $s,t \in T_\Sigma(X)$ and X is the set of dummy variables). We will call the tuple (Σ, E) an **algebraic theory**.

[530] Also called axiom.

Example 465. The algebraic theory of semilattices contains a single binary operation $\Sigma_S = \{\oplus : 2\}$ and the following equations in E_S:[531]

$$x \oplus x = x \qquad I\text{: idempotence}$$
$$x \oplus y = y \oplus x \qquad C\text{: commutativity}$$
$$(x \oplus y) \oplus z = x \oplus (y \oplus z). \qquad A\text{: associativity}$$

[531] It will be made clear why this is the theory of semilattices shortly.

Let $X = \{x, y, z\}$, the set of Σ–terms contains infinitely many terms, e.g.: $x \oplus y$, $x \oplus (y \oplus z)$, $(x \oplus x) \oplus (y \oplus z) \oplus (z \oplus x)$, etc.[532]

[532] The parentheses are here to denote the order in which the operation symbols was applied. While in semilattices, the operation $\oplus$ satisfies the equations making the parentheses and order irrelevant, when describing terms over the signature, we cannot remove them.

Definition 466 ((Σ, E)–algebras)**.** Given an algebraic theory (Σ, E), a (Σ, E)**–algebra** is a set A along with operations $f^A : A^n \to A$ for all $f : n \in \Sigma$ such that the pairs of terms in E are always equal when the operation symbols and dummy variables are instantiated in A.[533] We usually denote Σ^A for the set operations f^A.

[533] The operation symbol f is always instantiated by f^A and a dummy variable can be instantiated by any element of A. For instance, suppose (A, f^A, g^A) is a (Σ, E)–algebra and $f(x, g(y)) = g(y)$ is an equation in E, then for any $a, b \in A$, $f^A(a, g^A(b)) = g^A(b)$.

Examples 467. As is suggested by the terminology, the common algebraic structures can be defined with simple algebraic theories.

1. We can define a monoid as an algebra for the signature $\{\cdot : 2, 1 : 0\}$ and the equations $x \cdot (y \cdot z) = (x \cdot y) \cdot z$, $1 \cdot x = x$, $x \cdot 1 = x$. We will say that this is the algebraic theory of monoids.
2. Adding the unary operation $(-)^{-1}$ and the equations $x \cdot x^{-1} = 1$ and $x^{-1} \cdot x = 1$, we obtain the theory of groups.
3. Adding the equation $x \cdot y = y \cdot x$ yields the theory of abelian groups.
4. With the signature $\{+ : 2, \cdot : 2, 1 : 0, 0 : 0\}$, we can add the abelian group equations for the operation $+$ (identity is 0), the monoid equations for $\cdot$ (identity is 1) and the distributivity equation $x \cdot (y + z) = (x \cdot y) + (x \cdot z)$ and thus obtain the theory of rings.
5. The theory of semilattices has this named because a (Σ_S, E_S)–algebra is a semilattice.

We also have homomorphisms between (Σ, E)–algebras.

Definition 468 ((Σ, E)–algebra homomorphisms)**.** Given two (Σ, E)–algebras A and B, a **homomorphism** between them is a map $h : A \to B$ commuting with all operations in Σ, that is $\forall f : n \in \Sigma, h \circ f^A = f^B \circ h^n$.[534]

[534] We write h^n for componentwise application of the map h to vectors in A^n, i.e.: $h^n(a_1, \ldots, a_n) = (h(a_1), \ldots, h(a_n))$.

The category of (Σ, E)–algebras and their homomorphisms (with the obvious composition and identities) is denoted $\mathrm{Alg}(\Sigma, E)$.

Example 469 (Σ_S, E_S)**.** Recall from Example 463 that $\mathcal{P}_{\mathrm{ne}}$–algebras correspond to semilattices. Up to a couple of missing functoriality arguments, we have shown that the categories $\mathbf{Set}^{\mathcal{P}_{\mathrm{ne}}}$ and $\mathrm{Alg}(\Sigma_S, E_S)$ are isomorphic. We say that (Σ_S, E_S) is an **algebraic presentation** of the monad $\mathcal{P}_{\mathrm{ne}}$ or that the theory of semilattices presents the monad $\mathcal{P}_{\mathrm{ne}}$.

It turns out all algebraic theories present at least one monad.

Definition 470 (Term monad)**.** Let (Σ, E) be an algebraic theory, one can assign to any set X, the set $T_{\Sigma,E}(X)$ of terms in $T_\Sigma(X)$ modulo the equations in E.[535] This can be extended to functions $f : X \to Y$, by variable substitution, i.e.: $T_\Sigma(f)$ acts on a term t by replacing all occurrences of $x \in X$ with $f(x) \in Y$ and $T_{\Sigma,E}(f)$ acts on equivalence classes by $[t] \mapsto [T_\Sigma(f)(t)]$. We obtain a functor $T_{\Sigma,E}$ on which we can put a monad structure.

The unit is obvious because any element of X is a Σ–term, thus $\eta_X : X \to T_{\Sigma,E}(X)$ maps x to the equivalence class containing the term x. The multiplication is derived from the fact that applying operations in Σ to Σ–terms yields Σ–terms. More explicitly, μ_X is a *flattening* operation defined recursively by

$$\forall t \in T_\Sigma(X), \mu_X([[t]]) = [t]$$
$$\forall f : n \in \Sigma, t_1, \dots, t_n \in T_\Sigma T_{\Sigma,E}(X), \mu_X([f(t_1, \dots, t_n)]) = [f(\mu_X([t_1]), \dots, \mu_X([t_n]))]$$

One can show that $\mathbf{Set}^{T_{\Sigma,E}}$ is the category of (Σ, E)–algebras.

[535] Let us not waste time here to make this more formal as there is a lot to say that is not relevant to the rest of this story. We say that two terms s and t are equal modulo E if we can rewrite s using the equations in E and obtain t. The informal notion of *rewriting* is good enough for us (we hope you got a sense of what rewriting means when learning about high school algebra).

Unfortunately, the term monads are not very simple to work with[536] and it is often desirable to find other simpler monads which are presented by the same theory or conversely to find an algebraic presentation for a given monad.

[536] In fact, you might have realized we chose to not even bother.

Examples 471. 1. The algebraic theory presenting $\mathcal{D}$ is called the theory of **convex algebras** and is denoted $(\Sigma_{\mathsf{CA}}, E_{\mathsf{CA}})$, it consists of a binary operation $+_p : 2$ for any $p \in (0,1)$ which is meant to represent a choice between the two terms in the operation, the left one being chosen with probability p and the second one with probability $1-p$. There are three equations in the theory that morally ensure that terms representing the same probabilistic choice are equal.[537]

$$x +_p x = x \qquad I_p\text{: idempotence}$$
$$x +_p y = y +_{\overline{p}} x \qquad C_P\text{: skew-commutativity}$$
$$(x +_q y) +_p z = x +_{pq} (y +_{\frac{p\overline{q}}{\overline{pq}}} z) \qquad A_p\text{: skew-associativity}$$

These equations are necessary for every distribution in $\mathcal{D}X$ to correspond uniquely to an equivalence class in $T_{\Sigma_{\mathsf{CA}}, E_{\mathsf{CA}}}(X)$.

[537] For $x \in [0,1]$, we denote $\overline{x} := 1 - x$.

2. The monad $(- + \mathbf{1})$ is particular because it is really simple and combines very well with other monads.

Proposition 472. *For any monad M, there is a monad structure on the composition $M(- + \mathbf{1})$. Moreover, if M is presented by (Σ, E) the monad $M(- + \mathbf{1})$ is presented by $(\Sigma \cup \{* : 0\}, E)$, that is, the new theory only has an additional constant*[538] *which is neutral with respect to the operation symbols.*

Proof. Postponed to Exercise 478. □

[538] A 0–ary opeartion is more commonly called a constant.

We often qualify theories with an added constant as **pointed**. For instance, the theories presented by $\mathcal{P}_{\text{ne}}(- + \mathbf{1})$ and $\mathcal{D}(- + \mathbf{1})$ are those of **pointed semilattices** and **pointed convex algebras** respectively.

Remark 473 (Lawvere's way). There is another way to do universal algebra *more categorically* still very much linked to monads: Lawvere theories. Algebras over a Lawvere theory[539] are defined more abstractly using the categorical language and, on this account, they enjoy straightforward generalization through enrichment or lifting to higher order categories.

[539] They are called models of the theory.

8.3 POV: Computer Programs

In this section, we will develop on an original idea by Eugenio Moggi that monads are suitable models for a general notion of *computation*. In the sequel, we will use the terms *type* and *set* interchangeably.

Moggi gave a justification for using monads in computer science (particularly in programming semantics) via the informal intuition of *computational types*. For a type A, the computational type of A should contain all computations which return a value of type A. It is intended for the interpretation of *computation* to be made explicit by an instance of a monad. In most cases, it can be thought of as a piece of code which returns some value, but for now, we start by building the intuition in an abstract sense.

Let MA denote the computational type of A and MMA the computational type of MA, that is computations returning values which are themselves computations of type A. The following items should coincide with our intuition of computation.

1. For any $x \in A$, there is a trivial computation $\mathsf{return}\ x \in MA$.
2. For any $C \in MMA$, we can reduce C to $\mathsf{flatten}(C) \in MA$ which executes C and the computation returned by C to obtain a final return value of type A.
3. If $C \in MA$, then $\mathsf{flatten}(\mathsf{return}\ C) = C$.
4. If $C \in MA$ and $C' \in MMA$ does the same computation as C but instead of returning a value x, it returns the computation $\mathsf{return}\ x$, then $\mathsf{flatten}(C') = C$.
5. If $MMMA$ is the computational type of MMA and $C \in MMMA$, then there are two ways to flatten C. First, there is the computation C_1 which executes C and executes the returned computation (of type MMA) to obtain a final value of type MA, hence $C_1 \in MMA$ and $\mathsf{flatten}(C_1) \in MA$. Second, C_2 executes C and flattens the returned computation to obtain a final value of type MA, C_2 is also of type MMA and $\mathsf{flatten}(C_2) \in MA$. These two operations should yield the same result.

Now, a monad M is a description of computational types that is general, namely, for any type A, the monad M gives a type MA behaving as expected. You can check that $x \mapsto \mathsf{return}\ x$ is the unit of this monad and $\mathsf{flatten}$ is the multiplication.

Examples 474. Here, we list more examples commonly used in computer science.

List monad: For any set X, let $L(X)$ denote the set of all finite lists whose elements are chosen in X. This is a functor that sends a function $f : X \to Y$ to its extension on lists $L(f) : L(X) \to L(Y)$ which applies f to all elements on the list (in lots of programming languages, one writes $L(f) := \mathsf{map}(f, -)$).Then, we can put a monad structure on L. The unit maps send an element $x \in X$ to the list containing only that element: $\eta_X = x \mapsto [x]$. The multiplication maps concatenate all the lists in a lists of lists: $\mu_X = [\ell_1, \ldots, \ell_n] \mapsto \ell_1\ell_2 \cdots \ell_n$. It is easy to check diagrams (219) to (220) commute.

Termination: In order to model computations that might terminate with no output, the monad $(- + \mathbf{1})$ is often used. For any type X, the type $X + \mathbf{1}$ has all the values of type X and an additional termination value denoted $*$. The behavior of the unit and multiplication of the monad can be interpreted as the fact that the stage of the computation that leads to a termination is irrelevant. This monad is also known as the Maybe monad.

Non-deterministic choice: The model for nondeterministic choice is given by the monad $\mathcal{P}_{\mathrm{ne}}$. The elements of $S \in \mathcal{P}_{\mathrm{ne}}(X)$ are seen as the possible outcomes of a nondeterministic choice. The unit is basically viewing a deterministic choice as a nondeterministic choice. The multiplication reduces the number of choices without changing the behavior. For instance, consider a process that nondeterministically chooses between two boxes containing two coins each and then chooses a coin in the box. By simply observing the final choice, we would not be able to distinguish it from a process that nondeterministically chooses between the four coins from the start.

Probabilistic choice: In the same vein, probabilistic choice can be interpreted with the monad $\mathcal{D}$ of finitely supported distributions.

Exceptions: As a generalization of termination, we can put a monad structure on the functor $(\cdot + E)$ where E is a set of exceptions that the computation can raise.

This view sheds light on one important features of monads we have not yet explored. If M and $\widehat{M}$ are monads describing computational effects, it is natural to ask for a way to combine them. Indeed, it does not seem too ambitious to have a model for programs which, for instance, make nondeterministic choices and also might terminate with no output. It turns out there is a very useful tool to deal with this at the level of monads.

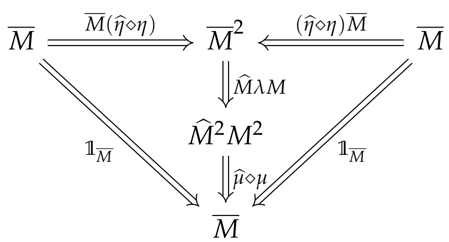
Definition 475 (Monad distributive law)**.** Let (M, η, μ) and $(\widehat{M}, \widehat{\eta}, \widehat{\mu})$ be two monads on $\mathbf{C}$, a natural transformation $\lambda : M\widehat{M} \Rightarrow \widehat{M}M$ is called a **monad distributive law of** M **over** $\widehat{M}$ if it makes (236), (237) commute.

$$\begin{array}{ccccc} M & \xRightarrow{M\widehat{\eta}} & M\widehat{M} & \xLeftarrow{\eta\widehat{M}} & \widehat{M} \\ & \searrow_{\widehat{\eta}M} & \Downarrow \lambda & \swarrow_{\widehat{M}\eta} & \\ & & \widehat{M}M & & \end{array} \tag{236}$$

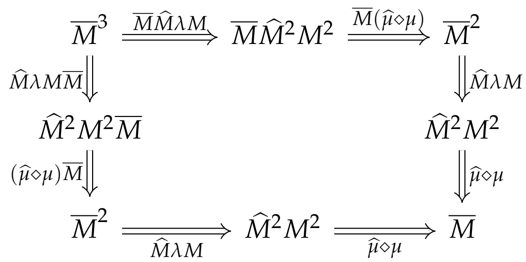
$$\begin{array}{ccccccccc} MM\widehat{M} & & \xRightarrow{\mu\widehat{M}} & & M\widehat{M} & & \xLeftarrow{M\widehat{\mu}} & & M\widehat{M}\widehat{M} \\ \Downarrow M\lambda & & & & \Downarrow \lambda & & & & \Downarrow \lambda\widehat{M} \\ M\widehat{M}M & \xRightarrow[\lambda M]{} & \widehat{M}MM & \xRightarrow[\widehat{M}\mu]{} & \widehat{M}M & \xLeftarrow[\widehat{\mu}M]{} & \widehat{M}\widehat{M}M & \xLeftarrow[\widehat{M}\lambda]{} & \widehat{M}M\widehat{M} \end{array} \tag{237}$$

Proposition 476. *If $\lambda : M\widehat{M} \Rightarrow \widehat{M}M$ is a monad distributive law, then the composite $\overline{M} = \widehat{M}M$ is a monad with unit $\overline{\eta} = \widehat{\eta} \diamond \eta$ and multiplication $\overline{\mu} = (\widehat{\mu} \diamond \mu) \cdot \widehat{M}\lambda M$.*

Proof. We have to show that the following instances of (219) and (220) commute.

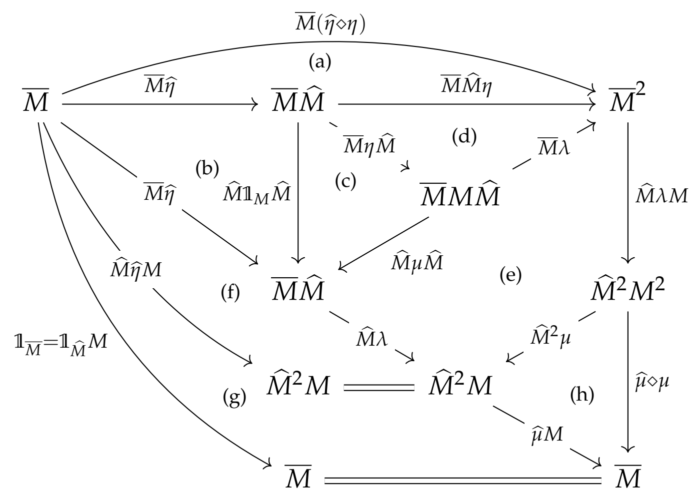


(238)



(239)

For the left part of (238), we have the following paving, the justifications of each part is given in the margin (the notation (219).L (resp. .R) means only the left (resp. right) part of the diagram is considered).



(240)

Showing (240) commutes:

(a) Definition of $\diamond$ and functoriality of $\overline{M}$.

(b) $\widehat{M}\mathbb{1}_M\widehat{M}$ is the identity transformation.

(c) Act on (219).L with $\widehat{M}$ on the left and right.

(d) Act on (236).R with $\overline{M}$ on the left.

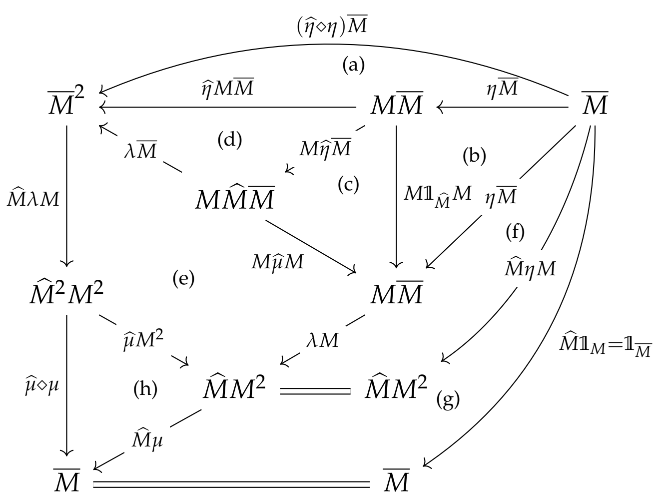
(e) Act on (237).L with $\widehat{M}$ on the left.

(f) Act on (236).L with $\widehat{M}$ on the left.

(g) Act on (219) with M on the right.

(h) Definition of $\diamond$.

For the right part of (238), we have the following paving.

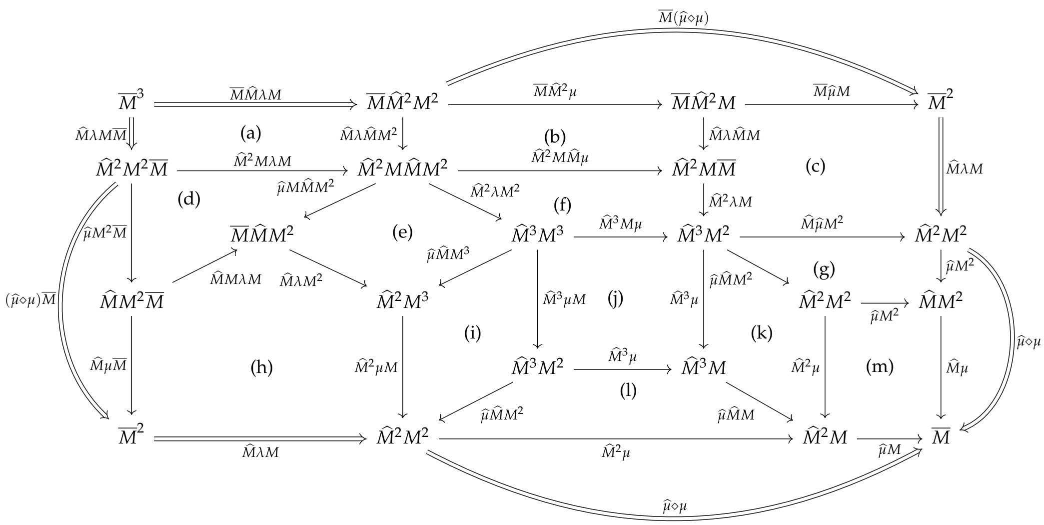


(241)

Showing (241) commutes:

(a)

For (239), we do the same thing.



(a) Def of $\widehat{M}\lambda \diamond \lambda M$.

(b) Def of $\widehat{M}\lambda\widehat{M} \diamond \mu$.

(c) Apply $\widehat{M}(\cdot)M$ to (237).R.

(d) Def of $\widehat{\mu} \diamond M\lambda M$.

(e) Def of $\widehat{\mu} \diamond \lambda M^2$.

(f) Def of $\widehat{M}^2\lambda \diamond \mu$.

(g) Apply $(\cdot)M^2$ to associativity of $\widehat{\mu}$ (220).

(h) Apply $\widehat{M}(\cdot)M$ to (237).L.

(i) Def of $\widehat{\mu}\widehat{M} \diamond \mu M$.

(j) Apply $\widehat{M}^3$ to associativity of μ (220).

(k) Def of $\widehat{\mu}\widehat{M} \diamond \mu$.

(l) Same as (k): Def of $\widehat{\mu}\widehat{M} \diamond \mu$.

(m) Def of $\widehat{\mu} \diamond \mu$.

□

Corollary 477. *If* **C** *has (binary) coproducts and a terminal object* **1** *and M is a monad, then $M(- + \mathbf{1})$ is also monad.*

Proof. We will exhibit a monad distributive law of M over $(- + \mathbf{1})$. We claim

$$\iota_X : MX + \mathbf{1} \to M(X+1) = [M(\mathsf{inl}^{X+\mathbf{1}}), \eta_{X+\mathbf{1}} \circ \mathsf{inr}^{X+\mathbf{1}}]$$

is a monad distributive law $\iota : (- + \mathbf{1})M \Rightarrow M(- + \mathbf{1})$. Then, it follows by Proposition 476. □

SOL Exercise 478. Show Proposition 472 with the monad structure on $M(- + \mathbf{1})$ given in Corollary 477.

Example 479 (Rings)**.** Consider the term monads for the theory of monoids and abelian groups $T_{\mathbf{Mon}}$ and $T_{\mathbf{Ab}}$. You can check that they are the monads induced by the free-forgetful adjunctions between **Mon** and **Set** and **Ab** and **Set**. Also, $T_{\mathbf{Mon}}$ is the same thing as the list monad. Call the binary operation of $T_{\mathbf{Mon}}$ and $T_{\mathbf{Ab}}$ the product and sum respectively.

Then, by identifying products of sums (elements of $T_{\mathbf{Mon}}T_{\mathbf{Ab}}X$) with sums of products (elements of $T_{\mathbf{Ab}}T_{\mathbf{Mon}}X$) by *distributing* the product over the sum as we are used to do with, say, real numbers, we obtain a monad distributive law of $T_{\mathbf{Mon}}$ over $T_{\mathbf{Ab}}$. The resulting composite monad $T_{\mathbf{Ab}}T_{\mathbf{Mon}}$ is the term monad for the theory of rings. The term distributive law comes from this example.

Remark 480. It is not always possible to combine monads in such a natural way. For instance, it was shown that no distributive law exist between $\mathcal{P}_{\text{ne}}$ and $\mathcal{D}$ and even that no monad structure can exist on $\mathcal{P}_{\text{ne}}\mathcal{D}$ or $\mathcal{D}\mathcal{P}_{\text{ne}}$. Thus, modelling combined probabilistic and nondeterministic effects has been quite a hard endeavor and is still an active area of research I discovered in an internship with Matteo Mio and Valeria Vignudelli at ENS de Lyon last summer.

If you are looking for more applications of this perspective on monads and especially if you enjoyed the assignment on Brzozowski's algorithm, I suggest you look into the paper *Generalizing Determinization From Automata to Coalgebras* available at `https://arxiv.org/abs/1302.1046`.

8.4 Exercises

1. Show that the triple $(\mathcal{D}, \eta, \mu)$ described in Example 452.3 is a monad.
2. Show that the Kleisli category of the powerset monad is the category **Rel** of relations.
3. Show that ι defined in the proof of Corollary 477 is a monad distributive law.
4. Show Proposition 472 with the monad structure on $M(- + \mathbf{1})$ given in Corollary 477.

9 Solutions to Exercises

9.1 Solutions to Chapter 1 171
9.2 Solutions to Chapter 2 172
9.3 Solutions to Chapter 3 174
9.4 Solutions to Chapter 4 177
9.5 Solutions to Chapter 5 179
9.6 Solutions to Chapter 6 181
9.7 Solutions to Chapter 7 181
9.8 Solutions to Chapter 8 182

9.1 Solutions to Chapter 1

Solution to Exercise 110. Take any monoid M with an idempotent element $x \neq 1_M$ (it satisfies $x \cdot x = x$). Letting $\mathbf{C}$ be $\mathbf{B}M$ and $\mathbf{C}'$ contain the object $*$ and only the morphism x yields a suitable example because the identity in $\mathbf{C}'$ is x. □

Solution to Exercise 134. On morphisms, we define $\Delta_{\mathbf{C}}(f) = (f, f)$. The functoriality properties hold because everything in $\mathbf{C} \times \mathbf{C}$ is done componentwise.

i. For $f : X \to Y$, we have $(f, f) : (X, X) \to (Y, Y)$.

ii. For $f : X \to Y$ and $g : Y \to Z$, we have $(g, g) \circ (f, f) = (g \circ f, g \circ f)$.

iii. For any $X \in \mathbf{C}_0$, we have $\Delta_{\mathbf{C}}(\mathrm{id}_X) = (\mathrm{id}_X, \mathrm{id}_X) = \mathrm{id}_{(X,X)}$.

□

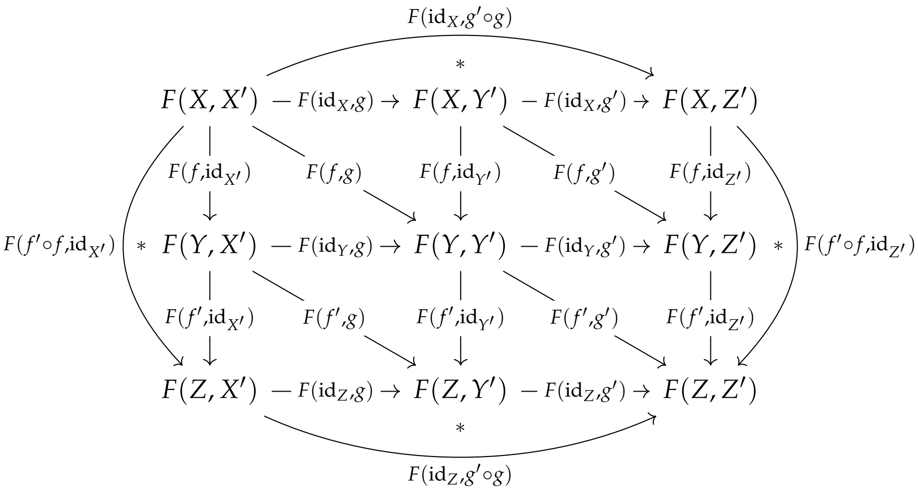
Solution to Exercise 136. A quick way to show $F(X, -)$ is a functor is to recognize it as the composition of F with $\Delta(X) \times \mathrm{id}_{\mathbf{C}'}$, where $\Delta(X)$ is the constant functor at X. Similarly, $F(-, Y) := F \circ (\mathrm{id}_{\mathbf{C}} \times Y)$. □

Solution to Exercise 137. Let us show the three properties of functoriality.

i. For any $(f, g) : (X, X') \to (Y, Y')$, by hypothesis, we have the following commutative square showing $F(f, g)$ has the right source and target.

$$\begin{array}{ccc} F(X,X') & \xrightarrow{F(\mathrm{id}_X, g)} & F(X,Y') \\ {\scriptstyle F(f,\mathrm{id}_{X'})}\downarrow & \searrow^{F(f,g)} & \downarrow{\scriptstyle F(f,\mathrm{id}_{Y'})} \\ F(Y,X') & \xrightarrow[F(\mathrm{id}_Y, g)]{} & F(Y,Y') \end{array}$$

ii. Let us have two morphisms $(f, g) : (X, X') \to (Y, Y')$ and $(f', g') : (Y, Y') \to (Z, Z')$ in $\mathbf{C} \times \mathbf{C}'$. The hypothesis on $F(-, -)$ gives the four commutative squares below and the functoriality of F in each component gives the commutativity of the parts denoted by $*$.



We conclude from the commutativity of the whole diagram that $F(f', g') \circ F(f, g) = F(f' \circ f, g' \circ g)$.

iii. For any $(A, B) \in (\mathbf{C} \times \mathbf{C}')_0$, the functoriality of either component yields

$$F(\mathrm{id}_{(A,B)}) = F(\mathrm{id}_A, \mathrm{id}_B) = \mathrm{id}_{F(A,B)}.$$

□

9.2 Solutions to Chapter 2

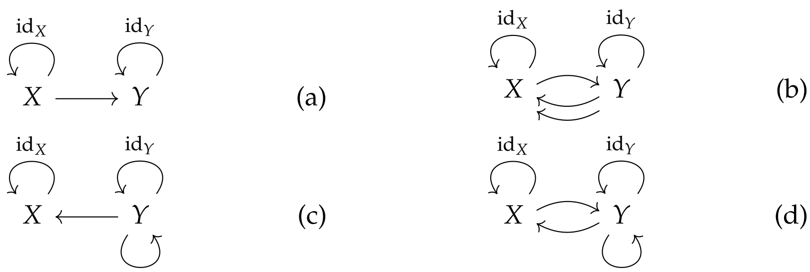
Solution to Exercise 161. Let us have two morphisms $f : X \to Y$ and $g : Y \to Z$.

- Suppose f and g are monic. For any $h_1, h_2 : Z \to Z'$ satisfying $h_1 \circ g \circ f = h_2 \circ g \circ f$, monicity of f implies $h_1 \circ g = h_2 \circ g$ which in turn, by monicity of g imply $h_1 = h_2$. Thus, $g \circ f$ is monic.
- We apply duality. Suppose f and g are epic, then f^{op} and g^{op} are monic so $(g \circ f)^{\mathrm{op}} = f^{\mathrm{op}} \circ g^{\mathrm{op}}$ is monic, thus $g \circ f$ is epic.
- If f and g are isomorphisms, then it is easy to check that $f^{-1} \circ g^{-1}$ is the inverse of $g \circ f$, implying $g \circ f$ is an isomorphism.

□

Solution to Exercise 177. We draw the categories with all the morphisms and we let you infer the composition[540] and show that they fit the requirement (by counting morphisms).

[540] The categories (a) and (b) have a uniquely determined composition. For (c) and (d), composing the non-identity endomorphism with itself can yield either itself or id_Y.



□

Solution to Exercise 180. Let (X, Y) be an object of $\mathbf{C} \times \mathbf{D}$, the pair consisting of $\langle\rangle_{\mathbf{C}} : X \to \mathbf{1_C}$ and $\langle\rangle_{\mathbf{D}} : Y \to \mathbf{1_D}$ is a morphism

$$(\langle\rangle_{\mathbf{C}}, \langle\rangle_{\mathbf{D}}) : (X, Y) \to (\mathbf{1_C}, \mathbf{1_D})$$

in $\mathbf{C} \times \mathbf{D}$. Any other morphism of this type is a pair (f, g) consisting of $f : X \to \mathbf{1_C}$ and $g : Y \to \mathbf{1_D}$, but by definition of terminal objects, we must have $f = \langle\rangle_{\mathbf{C}}$ and $g = \langle\rangle_{\mathbf{D}}$. Hence, $(\langle\rangle_{\mathbf{C}}, \langle\rangle_{\mathbf{D}})$ is the unique morphism in $\mathrm{Hom}_{\mathbf{C}\times\mathbf{D}}((X, Y), (\mathbf{1_C}, \mathbf{1_D}))$.

For the dual statement, we need to show that $(\emptyset_{\mathbf{C}}, \emptyset_{\mathbf{D}})$ is initial in $\mathbf{C} \times \mathbf{D}$ whenever $\emptyset_{\mathbf{C}}$ and $\emptyset_{\mathbf{D}}$ are initial in $\mathbf{C}$ and $\mathbf{D}$ respectively. Applying the opposite construction, we find that $\emptyset_{\mathbf{C}}$ and $\emptyset_{\mathbf{D}}$ are terminal in $\mathbf{C}^{\mathrm{op}}$ and $\mathbf{D}^{\mathrm{op}}$ respectively. Thus, the proof above shows $(\emptyset_{\mathbf{C}}, \emptyset_{\mathbf{D}})$ is terminal in $\mathbf{C}^{\mathrm{op}} \times \mathbf{D}^{\mathrm{op}}$. Now, a simple (tedious) unrolling of the definitions should convince you that $\mathbf{C}^{\mathrm{op}} \times \mathbf{D}^{\mathrm{op}} = (\mathbf{C} \times \mathbf{D})^{\mathrm{op}}$, so $(\emptyset_{\mathbf{C}}, \emptyset_{\mathbf{D}})$ is also the terminal object in $(\mathbf{C} \times \mathbf{D})^{\mathrm{op}}$. Therefore, $(\emptyset_{\mathbf{C}}, \emptyset_{\mathbf{D}})$ is initial in $\mathbf{C} \times \mathbf{D}$. □

Solution to Exercise 187. 1. Let $f : A \to B$ be the only non-identity morphism in $\mathbf{2}$, it is a monomorphism vacuously because there is only one morphism with target A (id_A). Now, for any morphism $m : X \to Y \in \mathbf{C}_1$, we can define $F : \mathbf{2} \rightsquigarrow \mathbf{C}$ by $FA = X$, $FB = Y$ and $Ff = m$ and it will be a functor. Thus, choosing m that is not monic yields the required example.

2. If f is split monic, it has a right inverse f'. This implies Ff' is the right inverse of Ff because $Ff \circ Ff' = F(f \circ f') = F(\mathrm{id}) = \mathrm{id}$. We conclude that Ff is split monic.

3. We need to show that functors preserve split epimorphisms. By duality, if f is split epic, then f^{op} is split monic, thus it is preserved by the functor F^{op}. And $Ff = (F^{\mathrm{op}}(f^{\mathrm{op}}))^{\mathrm{op}}$ is split epic.

4. Functors preserve isomorphisms because a morphism is an isomorphism if and only if it is split epic and split monic.[541] If $A \cong B$ and $i : A \to B$ is an isomorphism, then $Fi : FA \to FB$ is an isomorphism, so $FA \cong FB$.

□

[541] Because split epic is equivalent to having a left inverse and split monic is equivalent to having a right inverse.

Solution to Exercise 188. 1. Let $\mathbf{C}$ be a category with at least one morphism f that is not monic, the only functor $\langle\rangle : \mathbf{C} \rightsquigarrow \mathbf{1}$ sends f to $\mathrm{id}_\bullet$ which is monic.

2. Suppose that $F(f)$ is monic and let g and h be such that $f \circ g = f \circ h$. By monicity of $F(f)$, $F(f) \circ F(g) = F(f \circ g) = F(f \circ h) = F(f) \circ F(h)$ implies $F(g) = F(h)$. Since F is faithful, $g = h$.

3. We need to show faithful functors reflect epimorphisms.

□

Solution to Exercise 189. Let us have three monomorphisms $m : Y \rightarrowtail X$, $n : Z \rightarrowtail X$ and $o : W \rightarrowtail X$.

Reflexivity: We have $m \circ \mathrm{id}_Y = m$ thus $m \sim m$.

Symmetry: Suppose that $m \sim n$, namely, there is an isomorphism $i : Y \to X$ such that $m = n \circ i$. Then, pre-composing with the isomorphism i^{-1} yields $m \circ i^{-1} = n$ which implies $n \sim m$.

Transitivity: If $m \sim n$ and $n \sim o$, then there exist isomorphisms $i : Y \to Z$ and $i' : W \to Z$ satisfying $m = n \circ i$ and $n = o \circ i'$. Therefore, we have $m = o \circ i' \circ i$ which implies $m \sim o$.[542] □

[542] Recall that the composition of two isomorphisms is an isomorphism.

Solution to Exercise 192. Let us have five monomorphisms $m : Y \rightarrowtail X$, $m : Y' \rightarrowtail X$, $n : Z \rightarrowtail X$, $n' : Z' \rightarrowtail X$ and $o : W \rightarrowtail X$.[543]

[543] Recall that we often use m to refer to $[m]$.

Well-defined: Suppose that $m \leq n$, $m' \sim m$ and $n \sim n'$, namely, there is a morphism $k : Y \to Z$ and isomorphisms $i : Y \circ Y'$ and $i' : Z' \to Z$ such that $m = n \circ k$, $m' = m \circ i$ and $n = n' \circ i'$. Combining these equalities yields $m' = n' \circ i' \circ k \circ i$ which witnesses $m' \leq n'$.

Reflexivity: We have $m \circ \mathrm{id}_Y = m$ thus $m \leq m$.

Antisymmetry: If $m \leq n$ and $n \leq m$, then there exist morphisms $k : Y \to Z$ and $k' : Z \to Y$ satisfying $m = n \circ k$ and $n = m \circ k'$. Combining these two equalities yield $m = m \circ k' \circ k$ and $n = n \circ k \circ k'$. Therefore, since m and n are monic, we infer that $k' \circ k = \mathrm{id}_Y$ and $k \circ k' = \mathrm{id}_Z$. This means k is an isomorphism and $m \sim n$ (so $[m] = [n]$).

Transitivity: If $m \leq n$ and $n \leq o$, then there exist morphisms $k : Y \to Z$ and $k' : W \to Z$ satisfying $m = n \circ k$ and $n = o \circ k'$. Therefore, we have $m = o \circ k' \circ k$ which implies $m \leq o$. □

9.3 Solutions to Chapter 3

Solution to Exercise 197. There is a simple correspondence between a set S and the set of functions $\mathbf{1} \to S$.[544] An element $s \in S$ is sent to the function assigning s to $*$, and a function $f : \mathbf{1} \to S$ is sent to $f(*) \in S$. This suggests to define an element of a set S by a function from $\{*\}$ to S. This is indeed a categorical definition because we can abstract away from **Set**.

[544] Recall that the terminal object in **Set** is the singleton $\{*\}$, or any other singleton.

Definition 481 (Element)**.** In a category $\mathbf{C}$ with a terminal object $\mathbf{1}$,[545] an *element* of an object $X \in \mathbf{C}_0$ is a morphism in $\mathrm{Hom}_{\mathbf{C}}(\mathbf{1}, X)$.

[545] We need this requirement in the definition because some categories may not have a terminal object, and we would not know what object could replace it.

Unfortunately, this definition does not represent our intuition about elements faithfully in all categories with a terminal object.

- In **Poset**, the terminal object is the set $\{*\}$ with the only possible order $\leq_{\mathbf{1}}= \{(*,*)\}$. Any function $\mathbf{1} \to (X, \leq)$ is monotone because $\leq$ has to be reflexive. Thus, the same correspondence as for **Set** works, and an element of $(X, \leq)$ in the categorical sense can be seen as an "actual" element of the poset.

- In **Grp**, the terminal object is the trivial group with a single identity element. For any group G, there is only one homomorphism $\mathbf{1} \to G$ that must send the identity in $\mathbf{1}$ to the identity in G. Hence, there is only one categorical element of G no matter its size.

- In **Cat**, the terminal object is $\mathbf{1}$, the category with a single object $\bullet$ and a single morphism $\mathrm{id}_\bullet$. There is a simple correspondence between objects of a category $\mathbf{C}$ and functors $\mathbf{1} \rightsquigarrow \mathbf{C}$. In one direction, it sends $X \in \mathbf{C}_0$ to the functor sending $\bullet$ to X and $\mathrm{id}_\bullet$ to id_X. In the other direction, it sends $F : \mathbf{1} \rightsquigarrow \mathbf{C}$ to $F(\bullet) \in \mathbf{C}_0$. Therefore, a categorical element of $\mathbf{C}$ is an object of $\mathbf{C}$.[546]

[546] It is harder to decide whether the definition makes sense here. An other intuitive notion of element of **C** could be a morphism instead of a object.

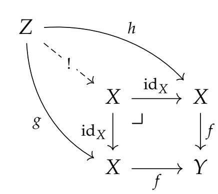
□

Solution to Exercise 210. As we have said that binary products are unique up to isomorphism, it is enough to show that $A \times B$ satisfies the same universal property as $B \times A$. Let π_A and π_B be the projections of $A \times B$, we claim that $B \xleftarrow{\pi_B} A \times B \xrightarrow{\pi_A} A$ is the product of B and A. Indeed, for any $B \xleftarrow{p_B} X \xrightarrow{p_A} A$, we use the original universal property of $A \times B$ to find a unique mediating morphism $! : X \to A \times B$ such that $\pi_B \circ ! = p_B$ and $\pi_A \circ ! = p_A$. □

Solution to Exercise 211. □

Solution to Exercise 214. The existence and uniqueness of $\prod_{i \in I} f_i$ is given by the universal property of the product $\prod_{i \in I} Y_i$ with for each $j \in I$, the morphism $f_j \circ \pi_j : \prod_{i \in I} X_i \to Y_j$. □

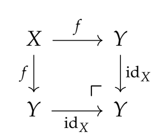
Solution to Exercise 241. ($\Rightarrow$) Suppose $f : X \to Y$ is monic, commutativity of (49) is trivial. For any $X \xleftarrow{g} Z \xrightarrow{h} X$ satisfying $f \circ g = f \circ h$, we have $g = h$. Thus $g = h$ is the mediating morphism ! of (242), it is unique because $\mathrm{id}_X \circ m = g$ implies $m = g$.



(242)

($\Leftarrow$) For any $g, h : Z \to X$ satisfying $f \circ g = f \circ h$, the universal property of the pullback tells us there is a unique $! : Z \to X$ making (242) commute. Since ! satisfies $g = \mathrm{id}_X \circ ! = h$, we conclude $g = ! = h$, thus f is a monomorphism.

The dual statement is that $f : X \to Y$ is epic if and only if (243) is a pushout. We leave the proof to you. □



(243)

Solution to Exercise 257. We recognize that 1 and 2 are dual statements and so are 3 and 4. We will prove something more general from which 1 and 3 follows, and use duality for 2 and 4.

Proposition 482. *Let* **J** *be a category with an initial object* $\emptyset$*. For any diagram* $F : \mathbf{J} \rightsquigarrow \mathbf{C}$*, we have* $\lim_{\mathbf{J}} F = F(\emptyset)$*.*

Proof. The limit cone comprises $F(\varnothing)$ as the tip, and for each $X \in \mathbf{D}_0$, $\phi_X = F(!_X) : F(\varnothing) \to FX$ is the image of the unique morphism $!_X : \varnothing \to X$ under F. We can verify this is a cone over F because for any $a : X \to X'$ in $\mathbf{J}_1$, $F(a) \circ F(!_X) = F(a \circ !_X) = F(!_{X'})$.[547]

[547] By initiality of $\varnothing$, $a \circ !_X = !_{X'}$.

Let $\{\psi_X : X \to FX\}$ be another cone over F. There is a morphism $\psi_\varnothing : X \to F(\varnothing)$, and it is a morphism of cones to the limit cone because for any $X \in \mathbf{J}_0$, $F(!_X) \circ \psi_\varnothing = \psi_X$ is a consequence of the ψ_Xs forming a cone. Any other morphism of cone $f : X \to F(\varnothing)$ must satisfy $F(!_\varnothing) \circ f = \psi_\varnothing$, but $!_\varnothing$ is the identity on $\varnothing$, hence $\mathrm{id}_{F(\varnothing)} \circ f = \psi_\varnothing$ implies $f = \psi_\varnothing$. □

Corollary 483 (Dual)**.** *Let* $\mathbf{J}$ *be a category with a terminal object* $\mathbf{1}$*. For any diagram* $F : \mathbf{J} \rightsquigarrow \mathbf{C}$*, we have* $\mathrm{colim}_{\mathbf{J}} F = F(\mathbf{1})$*.*

We can now apply these results to Exercise 257.

1. The limit of $A \to B$ is A.
2. The colimit of $A \to B$ is B.
3. The limit of $A \leftarrow B \to C$ is B.
4. The colimit of $A \to B \leftarrow C$ is B.

□

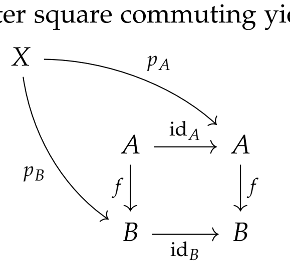
Solution to Exercise **??**. If $\{\psi_X : A \to DX\}_{X \in \mathbf{J}_0}$ is a cone over F, then the family $\{F\psi_X\}_{X \in \mathbf{J}_0}$ is a cone over $F \circ D$ since $Da \circ \psi_X = \psi_Y$ implies $FDa \circ F\psi_X = F\psi_Y$ for any $a : X \to Y \in \mathbf{J}_1$. On morphisms F_D sends $g : \{\psi_X\}_{X \in \mathbf{J}_0} \to \{\phi_X\}_{X \in \mathbf{J}_0}$ to $Fg : \{F\psi_X\}_{X \in \mathbf{J}_0} \to \{F\phi_X\}_{X \in \mathbf{J}_0}$. Again, the fact that Fg is a morphism of cones follows straightforwardly from

$$\phi_X \circ g = \psi_X \implies F\phi_X \circ Fg = F\psi_X.$$

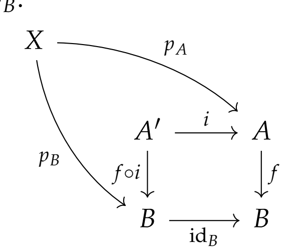
Observe that cones and cocones are dual in the sense that $\mathrm{Cone}(D)$ is the same as $\mathrm{Cocone}(D^{\mathrm{op}})$.[548] Therefore, $F^D : \mathrm{Cocone}(D) \rightsquigarrow \mathrm{Cocone}(F \circ D)$ can be defined as $F^{\mathrm{op}}_{D^{\mathrm{op}}} : \mathrm{Cone}(D^{\mathrm{op}}) \rightsquigarrow \mathrm{Cone}(F^{\mathrm{op}} \circ D^{\mathrm{op}}) = \mathrm{Cone}((F \circ D)^{\mathrm{op}})$. □

[548] Recall that D^{op} is the functor $\mathbf{J}^{\mathrm{op}} \rightsquigarrow \mathbf{C}^{\mathrm{op}}$ with the same action as D.

Solution to Exercise 275. Let $p_A : X \to A$ and $p_B : X \to B$ be such that (244) commutes. A mediating morphism $! : X \to A$ must satisfy $\mathrm{id}_A \circ ! = p_A$ and $f \circ ! = p_B$. The first equality ensures $! = p_A$ is unique and satisfies the second equality because the outer square commuting yields $f \circ p_A = p_B$.



(244)

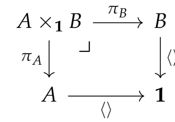


(245)

Let $p_A : X \to A$ and $p_B : X \to B$ be such that (245) commutes. A unique mediating morphism $! : X \to A$ must satisfy $i \circ ! = p_A$ and $f \circ i \circ ! = p_B$. Post-composing the first equality by i^{-1} implies $! = i^{-1} \circ p_A$ is unique and satisfies the second equality because $f \circ i \circ i^{-1} \circ p_A = f \circ p_A = p_B$. □

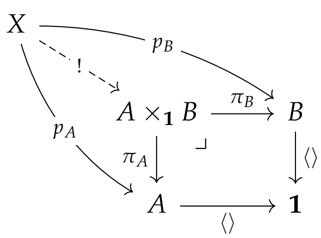
Solution to Exercise 284. We will show that if $\mathbf{C}$ has all pullbacks and a terminal object, then it has all finite products and equalizers. This implies, using Remark 278, that $\mathbf{C}$ is finitely complete.

For finite products, recall that it is enough to show that $\mathbf{C}$ has all binary products as it already has the empty product (the terminal object). We claim that the pullback of $A \xrightarrow{\langle\rangle} \mathbf{1} \xleftarrow{\langle\rangle} B$ is the binary product $A \times B$.



(246)

Indeed, for any $A \xleftarrow{p_A} X \xrightarrow{p_B} B$, we have $\langle\rangle \circ p_A = \langle\rangle \circ p_B$, thus, there is a unique morphism $! : X \to A \times_{\mathbf{1}} B$ making (247) commute. Since the commutativity of the squares always hold, this is equivalent to the unviersal property of the binary product. Hence $A \times B \cong A \times_{\mathbf{1}} B$.



(247)

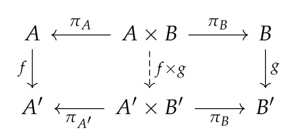
□

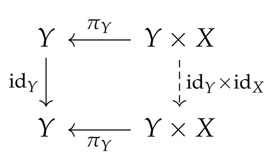
9.4 Solutions to Chapter 4

Solution to Exercise 297. If $\mathbf{C}$ has all binary products, we recall from Footnote 204 that sending (X, Y) to $X \times Y$ is a functor, and we use Exercise 136 to define $- \times X$. If $\mathbf{C}$ only has binary products with X, we do this manually.

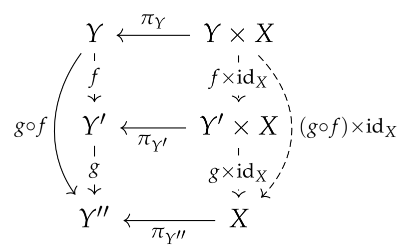
We define $- \times X$ on morphisms by sending $f : Y \to Y' \in \mathbf{C}_1$ to $f \times \mathrm{id}_X : Y \times X \to Y' \times X$. Functoriality follows from the definition of $\times$ on morphisms. Indeed, $\mathrm{id}_Y \times \mathrm{id}_X$ is the only morphism making (248) commute and $(g \circ f) \times \mathrm{id}_X$ is the only morphism making (249) commute.

Recall that if $f : A \to A'$ and $g : B \to B'$, $f \times g : A \times B \to A' \times B'$ is the unique morphism making the diagram below commute:





(248)



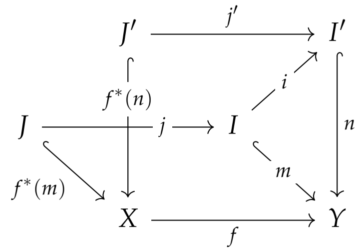
(249)

□

Solution to Exercise 303. First, we know that the pullback of the monomorphism m along f is monic by Theorem 271. Next, for $n : I' \rightarrowtail X \in \mathrm{Sub}_{\mathbf{C}}(Y)$, we need to show $[m] = [n]$ implies $[f^*(m)] = [f^*(n)]$.[549] In (250), we need to show there is an

549 Recall that $[m] = [n]$ when there is an isomorphism i satisfying $n = m \circ i$.

isomorphism $i' : J \to J'$ making everything commute.



(250)

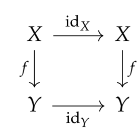
By the pullback property of J', there is a unique mediating morphism $i' : J \to J'$ commuting with (250).[550] Similarly, the pullback property of J, there is a unique mediating morphism $i'^{-1} : J' \to J$ commuting with (250).[551] The fact that i' and i'^{-1} are inverses follows from viewing $i'^{-1} \circ i'$ as a mediating morphism from the pullback J to itself which must be the identity by uniqueness. Similarly for $i' \circ i'^{-1}$.

[550] Use the fact that $n \circ i^{-1} \circ j = m \circ j = f \circ f^*(m)$.

[551] Use the fact that $m \circ i \circ j' = n \circ j' = f \circ f^*(n)$.

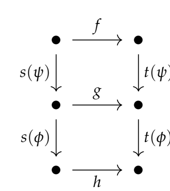
For functoriality of $\mathrm{Sub}_{\mathbf{C}}$, we need to show $\mathrm{id}^*(m) = m$ and $g^*(f^*(m)) = f \circ g^*(m)$. The first equality follows from Exercise 275 and the second from the pasting lemma. □

Solution to Exercise 323. 1. On morphisms, id sends $f : X \to Y$ to the commutative square $f : \mathrm{id}_X \to \mathrm{id}_Y$ depicted in (**??**). Since the identity of $\mathrm{id}_X \in \mathbf{C}_0^{\to}$ is $\mathrm{id}_X : \mathrm{id}_X \to \mathrm{id}_X$ and the composition of commutative squares is done by composing the left part and right part independently, we conclude that $\mathrm{id}(f \circ g) = f \circ g = \mathrm{id}(f) \circ \mathrm{id}(g)$. Thus, id is a functor.



(251)

2. On morphisms, s sends a commutative square $\phi : f \to g$ to the morphism $s(f) \to s(g)$ in the square, we denote it $s(\phi)$. In other words, we send a commutative square to its left part. Again, since the composition in $\mathbf{C}^{\to}$ is done independently on the left and right part, we find that $s(\phi \circ \psi) = s(\phi) \circ s(\psi)$, thus s is a functor (see (252) for a visual aid).



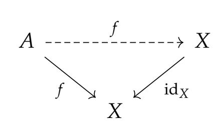
(252)

3. On morphisms, t sends a commutative square $\phi : f \to g$ to the morphism $t(f) \to t(g)$ in the square, we denote it $t(\phi)$. With a similar argument to the second point, we conclude that t is a functor.

□

Solution to Exercise **??**. □

Solution to Exercise 332. The terminal object of $\mathbf{C}/X$ is the identity morphism $\mathrm{id}_X : X \to X$. For any object of the slice category $f : A \to X$, we have the commutative triangle (253) with $! = f$. Uniqueness of ! follows from $\mathrm{id}_X \circ ! = f \implies ! = f$.



(253)

The dual statement is that id_X is the initial object of $X/\mathbf{C}$. □

9.5 Solutions to Chapter 5

Solution to Exercise 344. ($\Rightarrow$) For any $g : Y \to Y'$, the naturality of ϕ yields this commutative square.

$$\begin{array}{ccc} F(X,Y) & \xrightarrow{\phi_{X,Y}} & G(X,Y) \\ {\scriptstyle F(X,g)=F(\mathrm{id}_X,g)}\big\downarrow & & \big\downarrow{\scriptstyle G(\mathrm{id}_X,g)=G(X,g)} \\ F(X,Y') & \xrightarrow[\phi_{X,Y'}]{} & G(X,Y') \end{array} \tag{254}$$

We conclude that $\phi_{X,-}$ is a natural transformation $F(X,-)$. A symmetric argument works for $\phi_{-,Y}$ (see (255)).

$$\begin{array}{ccc} F(X,Y) & \xrightarrow{\phi_{X,Y}} & G(X,Y) \\ {\scriptstyle F(f,\mathrm{id}_Y)}\big\downarrow & & \big\downarrow{\scriptstyle G(f,\mathrm{id}_Y)} \\ F(X',Y) & \xrightarrow[\phi_{X',Y}]{} & G(X',Y) \end{array} \tag{255}$$

($\Leftarrow$) For any $(f,g) : (X,Y) \to (X',Y')$, we note that, by functoriality, $F(f,g) = F(f,\mathrm{id}_{Y'}) \circ F(\mathrm{id}_X,g)$ and similarly for G. Thus, we can combine the naturality of $\phi_{X,-}$ and $\phi_{-,Y}$ to obtain the commutativity of $\phi_{X,Y}$ as shown in (256).

$$\begin{array}{ccccc} & F(X,Y) & \xrightarrow{\phi_{X,Y}} & G(X,Y) & \\ & \big\downarrow{\scriptstyle F(\mathrm{id}_X,g)} & & {\scriptstyle G(\mathrm{id}_X,g)}\big\downarrow & \\ {\scriptstyle F(f,g)} & F(X,Y') & \xrightarrow{\phi_{X,Y'}} & G(X,Y') & {\scriptstyle G(f,g)} \\ & \big\downarrow{\scriptstyle F(f,\mathrm{id}_{Y'})} & & {\scriptstyle G(f,\mathrm{id}_{Y'})}\big\downarrow & \\ & F(X',Y') & \xrightarrow[\phi_{X',Y'}]{} & G(X',Y') & \end{array} \tag{256}$$

□

Solution to Exercise 348. Let $F, G : \mathbf{C} \rightsquigarrow \mathbf{D}$ be functors.

($\Rightarrow$) If $\phi : F \Rightarrow G$ is a natural isomorphism, then it has an inverse $\phi^{-1} : G \Rightarrow F$ which satisfies $\phi \cdot \phi^{-1} = \mathbb{1}_G$ and $\phi^{-1} \cdot \phi = \mathbb{1}_F$. Looking at each components, we find $\phi_X \circ (\phi^{-1})_X = \mathrm{id}_X$ and $(\phi^{-1})_X \circ \phi_X = \mathrm{id}_X$, hence they are isomorphisms.

($\Leftarrow$) Let $\phi : F \Rightarrow G$ be a natural transformation such that ϕ_X is an isomorphism for each $X \in \mathbf{C}_0$. We claim that the family ϕ_X^{-1} is the inverse of ϕ. After we show that this family is a natural transformation $G \Rightarrow F$, the construction implies it is the inverse of ϕ. For any $f : X \to Y \in \mathbf{C}_1$, the naturality of ϕ implies $\phi_Y \circ F(f) = G(f) \circ \phi_X$. Pre-composing with ϕ_X^{-1}, we have $G(f) = \phi_Y \circ F(f) \circ \phi_X^{-1}$ and therefore

$$\phi_Y^{-1} \circ G(f) = \phi_Y^{-1} \circ \phi_Y \circ F(f) \circ \phi_X^{-1} = F(f) \circ \phi_X^{-1}$$

yields the naturality of ϕ^{-1}. □

Solution to Exercise 352. We have already seen in Exercise 146 that we can take the dual of a functor $F : \mathbf{C} \rightsquigarrow \mathbf{D}$ to obtain a functor $F^{\mathrm{op}} : \mathbf{C}^{\mathrm{op}} \rightsquigarrow \mathbf{D}^{\mathrm{op}}$. It remains to check that a natural transformation $F \Rightarrow G$ can be identified with a natural transformation $G^{\mathrm{op}} \Rightarrow F^{\mathrm{op}}$. This follows from observing that the naturality square (257) in $\mathbf{D}$ corresponds to the naturality square (258) in $\mathbf{D}^{\mathrm{op}}$.[552]

[552] i.e.: (257) commutes if and only if (258) commutes.

$$\begin{array}{ccc} FX & \xrightarrow{\phi_X} & GX \\ {\scriptstyle Ff}\downarrow & & \downarrow{\scriptstyle Gf} \\ FY & \xrightarrow[\phi_Y]{} & GY \end{array} \quad (257) \qquad \begin{array}{ccc} G^{\text{op}}Y & \xrightarrow{\phi_Y} & F^{\text{op}}Y \\ {\scriptstyle G^{\text{op}}f}\downarrow & & \downarrow{\scriptstyle F^{\text{op}}f} \\ G^{\text{op}}X & \xrightarrow[\phi_X]{} & F^{\text{op}}X \end{array} \quad (258)$$ □

Solution to Exercise 372. On morphisms, this functor must send a pair of natural transformations $\eta : F \Rightarrow F'$ and $\phi : G \Rightarrow G'$ to a natural transformation $FG \Rightarrow F'G'$. This is exactly what horizontal composition does.

To see that horizontal composition is functorial, first note that $\mathbb{1}_F \diamond \mathbb{1}_G = \mathbb{1}_{FG}$. Next, the fact that horizontal composition commutes with composition of functors is exactly the interchange identity. □

Solution to Exercise 386. We need to show that $\simeq$ is reflexive, symmetric and transitive. Symmetry is trivial because the definition of $\mathbf{C} \simeq \mathbf{D}$ is symmetric. Reflexivity follows from the fact that the identity functor on any category is fully faithful and essentially surjective.

For transitivity, given the categories and functors represented in (259) with natural isomorphisms $\phi : FG \Rightarrow \text{id}_{\mathbf{D}}$, $\psi : GF \Rightarrow \text{id}_{\mathbf{C}}$, $\phi' : F'G' \Rightarrow \text{id}_{\mathbf{E}}$ and $\psi' : G'F' \Rightarrow \text{id}_{\mathbf{D}}$, we claim that the composition $G \circ G'$ is the quasi-inverse of $F' \circ F$.

$$\mathbf{C} \underset{G}{\overset{F}{\rightleftarrows}} \mathbf{D} \underset{G'}{\overset{F'}{\rightleftarrows}} \mathbf{E} \quad (259)$$

Since the biaction of functors preserves natural isomorphisms,[553] we have two natural isomorphisms

$$\phi' \cdot (F'\phi G') : F'FGG' \Rightarrow \text{id}_{\mathbf{E}} \text{ and } \psi \cdot (G\psi' F) : GG'F'F \Rightarrow \text{id}_{\mathbf{C}},$$

which shows $\mathbf{C} \simeq \mathbf{E}$. □

[553] This holds because acting on the left or right with a functor is a functor, part of this is shown in the next solution and it also follows from the previous exercise.

Solution to Exercise 387. We will show the following two implications

$$\begin{aligned} \forall \mathbf{D} \quad \mathbf{C} \simeq \mathbf{C}' &\implies [\mathbf{C}, \mathbf{D}] \simeq [\mathbf{C}', \mathbf{D}] \\ \forall \mathbf{C} \quad \mathbf{D} \simeq \mathbf{D}' &\implies [\mathbf{C}, \mathbf{D}] \simeq [\mathbf{C}, \mathbf{D}'] \end{aligned}$$

and infer that $\mathbf{C} \simeq \mathbf{C}'$ and $\mathbf{D} \simeq \mathbf{D}'$ implies

$$[\mathbf{C}, \mathbf{D}] \simeq [\mathbf{C}', \mathbf{D}] \simeq [\mathbf{C}', \mathbf{D}'].$$

For the first implication, let $F : \mathbf{C} \rightsquigarrow \mathbf{C}'$ and $G : \mathbf{C}' \rightsquigarrow \mathbf{C}$ be quasi-inverses. We define the functor $(-)F : [\mathbf{C}', \mathbf{D}] \rightsquigarrow [\mathbf{C}, \mathbf{D}]$ that acts on functors by pre-composition and on natural transformations by the right action in Definition 362.[554] Similarly, we define the functor $(-)G : [\mathbf{C}, \mathbf{D}] \rightsquigarrow [\mathbf{C}', \mathbf{D}]$. We claim that $(-)F$ and $(-)G$ are quasi-inverses.

[554] i.e.: $H : \mathbf{C} \rightsquigarrow \mathbf{D}$ is mapped to $HF = H \circ F$ and $\phi : H \Rightarrow H'$ is mapped to ϕF. Functoriality follows from the properties of the right action.
Another way to show functoriality is to recall that $\phi F = \phi \diamond \mathbb{1}_F$ and hence $(-)F$ is the composition of the functor

$$\text{id}_{[\mathbf{C}', \mathbf{D}]} \times F : [\mathbf{C}', \mathbf{D}] \times \mathbf{1} \rightsquigarrow [\mathbf{C}', \mathbf{D}] \times [\mathbf{C}, \mathbf{C}']$$

with the horizontal composition functor defined in Exercise 372.

Let $\Phi : GF \Rightarrow \text{id}_{\mathbf{C}}$ be a natural isomorphism witnessing F and G being quasi-inverses, then $(-)\Phi$ is a natural isomorphism from $(-)GF$ to $\text{id}_{[\mathbf{C}, \mathbf{D}]}$. Indeed, for any $\phi : H \Rightarrow H' \in [\mathbf{C}, \mathbf{D}]_1$, (260) commutes as the top path and bottom path are both equal to $\phi \diamond \Phi$ and $H\Phi$ is an isomorphism because Φ is and functors preserve isomorphisms.

$$\begin{array}{ccc} HGF & \xrightarrow{H\Phi} & H \\ {\scriptstyle \phi GF}\downarrow & & \downarrow{\scriptstyle \phi} \\ H'GF & \xrightarrow[H'\Phi]{} & H' \end{array} \quad (260)$$

We leave to you the symmetric argument showing $(-)FG \cong \mathrm{id}_{[\mathbf{C}',\mathbf{D}]}$ and the similar argument for the second implication. □

9.6 Solutions to Chapter 6

Solution to Exercise 391. Let $f : A \to B$ be an isomorphism. For each $X \in \mathbf{C}_0$, we know from our definition of the other Hom functor that $- \circ f : \mathrm{Hom}(B, X) \to \mathrm{Hom}(A, X)$ is an isomorphism.[555] It remains to show that this is natural in X. For all $g : X \to X'$ we have to show (261) commutes.

$$\begin{array}{ccc} H^A(X) & \xrightarrow{(-)\circ f} & H^B(X) \\ {\scriptstyle g\circ(-)}\downarrow & & \downarrow{\scriptstyle g\circ(-)} \\ H^A(X') & \xrightarrow[(-)\circ f]{} & H^B(X') \end{array} \qquad (261)$$

Starting with h in the top left, both paths send it to $g \circ h \circ f$. We conclude that $H^A \cong H^B$. We leave the proof of the dual to you. □

[555] All functors preserve isomorphisms, see Exercise 187.

Solution to Exercise 405. ($\Rightarrow$) Suppose there is a natural isomorphism $\phi : \mathrm{Hom}_{\mathbf{C}}(X, -) \Rightarrow \mathbf{1}$, then for any object $Y \in \mathbf{C}_0$, there is a bijection $\mathrm{Hom}_{\mathbf{C}}(X, Y) \cong \{\star\}$. Hence, there is a unique morphism $X \to Y$.

($\Leftarrow$) Suppose that X is initial, then for any $Y \in \mathbf{C}_0$, we have an isomorphism $\phi_Y : \mathrm{Hom}_{\mathbf{C}}(X, Y) \to \mathbf{1}(Y)$ which sends the unique morphism $X \to Y$ to $\star$. We need to show this family is natural in Y. Let $f : Y \to Y' \in \mathbf{C}_1$, (262) clearly commutes because all sets are singletons.

$$\begin{array}{ccc} \mathrm{Hom}_{\mathbf{C}}(X,Y) & \xrightarrow{\phi_Y} & \mathbf{1}(Y) \\ {\scriptstyle f\circ-}\downarrow & & \downarrow{\scriptstyle \mathrm{id}_{\mathbf{1}}} \\ \mathrm{Hom}_{\mathbf{C}}(X,Y') & \xrightarrow[\phi_{Y'}]{} & \mathbf{1}(Y') \end{array} \qquad (262)$$

□

9.7 Solutions to Chapter 7

Solution to Exercise 434. We will proceed by defining the units and counits because, as you will see, they are practically given and then we will verify they satisfy the triangle identities. We denote (ϕ_X, ϕ_Y) for a commutative square with $\mathsf{s}(\phi_X, \phi_Y) = \phi_X$ and $\mathsf{t}(\phi_X, \phi_Y) = \phi_Y$

($\mathsf{t} \dashv \mathsf{id}$) The component of the unit at $f \in \mathbf{C}_0^{\to}$ is a commutative square from f to $\mathsf{id}(\mathsf{t}(f)) = \mathrm{id}_{t(f)}$. You should convince yourself that (263) is the only such square that is guaranteed to exist no matter what $\mathbf{C}$ is, we have $\eta_f = (f, \mathrm{id}_{t(f)})$. The component of the counit at $X \in \mathbf{C}_0$ is a morphism from $t(\mathrm{id}_X) = X$ to X. Again, the only possible choice is $\varepsilon_X = \mathrm{id}_X$. We check in the following derivations that the triangle identities hold.

$$\begin{array}{ccc} s(f) & \xrightarrow{f} & t(f) \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle \mathrm{id}_{t(f)}} \\ t(f) & \xrightarrow[\mathrm{id}_{t(f)}]{} & t(f) \end{array} \qquad (263)$$

$$\varepsilon_{\mathsf{t}(f)} \circ \mathsf{t}(\eta_f) = \mathrm{id}_{t(f)} \circ \mathrm{id}_{t(f)} = \mathrm{id}_{\mathsf{t}(f)}$$
$$\mathsf{id}(\varepsilon_X) \circ \eta_{\mathsf{id}(X)} = (\mathrm{id}_X, \mathrm{id}_X) \circ (\mathrm{id}_X, \mathrm{id}_X) = (\mathrm{id}_X, \mathrm{id}_X) = \mathrm{id}_{\mathsf{id}(X)}.$$

($\mathsf{id} \dashv \mathsf{s}$) The component of the unit at $X \in \mathbf{C}_0$ is a morphism from X to $\mathsf{s}(\mathsf{id}(X)) = X$, thus $\eta_X = \mathrm{id}_X$. The component of the counit at $f \in \mathbf{C}_0^{\to}$ is a commutative square from $\mathsf{id}(\mathsf{s}(f)) = \mathrm{id}_{s(f)}$ to f. Again, there is only once choice: $\varepsilon_f = (\mathrm{id}_{s(f)}, f)$ depicted in (264). The following derivations show the triangle identities hold.

$$\begin{array}{ccc} s(f) & \xrightarrow{\mathrm{id}_{s(f)}} & s(f) \\ {\scriptstyle \mathrm{id}_{s(f)}}\downarrow & & \downarrow{\scriptstyle f} \\ s(f) & \xrightarrow[f]{} & t(f) \end{array} \qquad (264)$$

$$\varepsilon_{\mathsf{id}(X)} \circ \mathsf{id}(\eta_X) = (\mathrm{id}_X, \mathrm{id}_X) \circ (\mathrm{id}_X, \mathrm{id}_X) = (\mathrm{id}_X, \mathrm{id}_X) = \mathrm{id}_{\mathsf{id}(X)}$$
$$\mathsf{s}(\varepsilon_f) \circ \eta_{\mathsf{s}(f)} = \mathrm{id}_{s(f)} \circ \mathrm{id}_{s(f)} = \mathrm{id}_{\mathsf{s}(f)}.$$

($? \dashv \mathsf{t}$) If t has a left adjoint ?, then there is a isomorphism $\mathrm{Hom}_{\mathbf{C}^{\to}}(?X, f) \cong \mathrm{Hom}_{\mathbf{C}}(X, t(f))$ that is natural in X and f. □

Solution to Exercise 449. Using Theorem 430, Theorem 443 and Proposition 444, we can obtain two chains of adjunctions.

$$\mathbf{C} \underset{R}{\overset{L}{\underset{\longleftarrow}{\overset{\longrightarrow}{\bot}}}} \mathbf{D} \underset{\lim_{\mathbf{J}}}{\overset{\Delta^{\mathbf{J}}_{\mathbf{D}}}{\underset{\longleftarrow}{\overset{\longrightarrow}{\bot}}}} [\mathbf{J}, \mathbf{D}] \qquad\qquad \mathbf{C} \underset{\lim_{\mathbf{J}}}{\overset{\Delta^{\mathbf{J}}_{\mathbf{C}}}{\underset{\longleftarrow}{\overset{\longrightarrow}{\bot}}}} [\mathbf{J}, \mathbf{C}] \underset{R-}{\overset{L-}{\underset{\longleftarrow}{\overset{\longrightarrow}{\bot}}}} [\mathbf{J}, \mathbf{D}]$$

Then, observing that both composite left adjoints are equal,[556] we conclude by Corollary 421 that $R\lim_{\mathbf{J}} \cong \lim_{\mathbf{J}}(R-)$. □

[556] Both $\Delta^{\mathbf{J}}_{\mathbf{D}} \circ L$ and $L\Delta^{\mathbf{J}}_{\mathbf{C}}$ send $X \in \mathbf{C}_0$ to the constant functor at LX.

9.8 Solutions to Chapter 8

Solution to Exercise 457. By the universal property of η' and one of the triangle identities, ε'_{KA} is the unique morphism such that $R'\varepsilon'_{KA} \circ \eta'_{R'KA} = \mathrm{id}_{R'KA}$ (see (265)).

$$\begin{array}{ccccc} R'KA & \xrightarrow{\eta'_{R'KA}} & R'L'R'KA & & L'R'KA \\ & {\scriptstyle \mathrm{id}_{R'KA}}\searrow & \Big\downarrow{\scriptstyle R'\varepsilon'_{KA}} & \xleftarrow{\quad R \quad} & \Big\downarrow{\scriptstyle \varepsilon'_{KA}} \\ & & R'KA & & KA \end{array} \qquad (265)$$

We claim that $K\varepsilon_A$ also fits in the place of ε'_{KA} in (265) which means they are equal by uniqueness. We need to show $R'K\varepsilon_A \circ \eta'_{R'KA} = \mathrm{id}_{R'KA}$. Recalling that $\eta' = \eta$ and $R'K = R$, we rewrite the equality as $R\varepsilon_A \circ \eta_{RA} = \mathrm{id}_{RA}$ which holds by a triangle identity. □