From Set-Categories to Ab-Categories

If $\mathcal{C}$ is a category, recall that for each ordered pair of objects $x, y \in \mathcal{C}$ we have a set of arrows

$$\mathsf{Hom}_{\mathcal{C}}(x, y)$$

and for each ordered triple of objectx $x, y, z \in \mathcal{C}$ we have a composition function

$$\circ : \mathsf{Hom}_{\mathcal{C}}(y, z) \times \mathsf{Hom}_{\mathcal{C}}(x, y) \to \mathsf{Hom}_{\mathcal{C}}(x, z).$$

In this chapter we will consider categories in which the Hom sets and the composition function have extra structure. The prototype for this behavior is the category of abelian groups.

First consider the category Grp of all groups and homomorphisms. Given $G, H \in \mathsf{Grp}$ we define their *direct product* $G \times H \in \mathsf{Grp}$ as the Cartesian product set together with the componentwise group operation:

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

The group $G \times H$ together with the canonical projections $\pi_G(g, h) := g$ and $\pi_H(g, h) := h$ satisfies the universal property of the *categorical product* in Grp:

$$\begin{tikzcd}
 & & G \\
K \arrow[r, dashed, "\exists!"] \arrow[urr, bend left, "\varphi_G"] \arrow[drr, bend right, "\varphi_H"'] & G \times H \arrow[ur, "\pi_G"] \arrow[dr, "\pi_H"'] & \\
 & & H
\end{tikzcd}$$

That is, given a group $K \in \mathsf{Grp}$ and two homomorphisms $\varphi_G : K \to G$ and $\varphi_H : K \to H$, **there exists a unique homomorphism** $K \to G \times H$ making the diagram commute.

Exercise: Check this.

Now let $\mathsf{Ab} \subseteq \mathsf{Grp}$ denote the subcategory of abelian groups. This is a **full** subcategory in the sense that for all abelian groups $A, B \in \mathsf{Ab}$ we have

$$\mathsf{Hom}_{\mathsf{Ab}}(A, B) = \mathsf{Hom}_{\mathsf{Grp}}(A, B).$$

[That is, there is no such thing as an "abelian homomorphism" between abelian groups; we just use the regular group homomorphisms.] Since the direct product of abelian groups is again abelian, we conclude that the direct product also satisfies the universal property of the categorical product in Ab. But now we give it a new name.

Definition of Direct Sum. Given abelian groups $A, B \in \mathsf{Ab}$ we define their *direct sum* $A \oplus B \in \mathsf{Ab}$ as the Cartesian product set together with the componentwise group operation, which we denote as addition:

$$(a_1, b_1) + (a_2, b_2) := (a_1 + a_2, b_1 + b_2).$$

///

This is the "abstract" or the "external" direct sum. It is related to the usual "internal" direct sum as follows. Consider an abelian group $A \in \mathsf{Ab}$ and two subgroups $B_1, B_2 \subseteq A$. Then we have a group homomorphism

$$\sigma : B_1 \oplus B_2 \to A$$

defined by $\sigma(b_1, b_2) := b_1 + b_2$. This homomorphism is an isomorphism if and only if it is surjective (i.e., $A = B_1 + B_2$) and injective (i.e., $B_1 \cap B_2 = \{0_A\}$).

Normally it would be sacrilegious to use an additive notation for the categorical product, but the following theorem says that this is okay in the category of abelian groups.

The Coproduct in Ab. Consider two abelian groups $A, B \in \mathsf{Ab}$. Then the direct sum $A \oplus B \in \mathsf{Ab}$ together with the canonical injections $\iota_A(a) := (a, 0_B)$ and $\iota_B := (0_A, b)$ satisfies the universal property of the *categorical coproduct* in Ab. That is, given an abelian group $C \in \mathsf{Ab}$ and two homomorphisms $\varphi_A : A \to A \oplus B$ and $\varphi_B : B \to A \oplus B$, **there exists a unique homomorphism** $A \oplus B \to C$ making the following diagram commute:

$$\begin{tikzcd}
A \arrow[dr, "\iota_A"] \arrow[drr, bend left, "\varphi_A"] & & \\
& A \oplus B \arrow[r, dashed, "\exists!"] & C \\
B \arrow[ur, "\iota_B"'] \arrow[urr, bend right, "\varphi_B"'] & &
\end{tikzcd}$$

Proof: Let $\bar{\varphi} : A \oplus B \to C$ be some homomorphism making the diagram commute. Then for all $a \in A$ and $b \in B$ we must have

$$\begin{aligned}
\bar{\varphi}(a, b) &= \bar{\varphi}((a, 0_B) + (0_A, b)) \\
&= \bar{\varphi}(\iota_A(a) + \iota_B(b)) \\
&= \bar{\varphi}(\iota_A(a)) + \bar{\varphi}(\iota_B(b)) \\
&= \varphi_A(a) + \varphi_B(b).
\end{aligned}$$

Thus, $\bar{\varphi}(a, b) := \varphi_A(a) + \varphi_B(b)$ is the unique function making the diagram commute. The only questions is whether this function $\bar{\varphi} : A \oplus B \to C$ is a group homomorphism, and for this we need the fact that C is **abelian**. For all (a_1, b_1) and (a_2, b_2) in $A \oplus B$ we have

$$\begin{aligned}
\bar{\varphi}((a_1, b_1) + (a_2, b_2)) &= \bar{\varphi}(a_1 + a_2, b_1 + b_2) \\
&= \varphi_A(a_1 + a_2) + \varphi_B(b_1 + b_2) \\
&= (\varphi_A(a_1) + \varphi_A(a_2)) + (\varphi_B(b_1) + \varphi_B(b_2)) \\
&\overset{!}{=} (\varphi_A(a_1) + \varphi_B(b_1)) + (\varphi_A(a_2) + \varphi_B(b_2)) \\
&= \bar{\varphi}(a_1, b_1) + \bar{\varphi}(a_2, b_2).
\end{aligned}$$

The commutativity of C was used in step "!". ///

Now recall that the **trivial group** $0 \in \mathsf{Ab}$ is a *zero object* in the sense that for each group $A \in \mathsf{Ab}$ there exists a unique homomorphism $A \to 0$ and a unique homomorphism $0 \to A$. By composing these homomorphisms for each pair of groups $A, B \in \mathsf{Ab}$ we obtain a unique zero homomorphism:

$$\begin{tikzcd} A \arrow[r, "\exists!"'] \arrow[rr, bend left, "0_{AB}"] & 0 \arrow[r, "\exists!"'] & B \end{tikzcd}$$

Finally, one can check that the canonical projections $\pi_A : A \oplus B \to A$, $\pi_B : A \oplus B \to B$ and the canonical injections $\iota_A : A \to A \oplus B$, $\iota_B : B \to A \oplus B$ satisfy the following commutative diagram:

$$\begin{tikzcd} A \arrow[rr, "\mathsf{id}_A"] \arrow[dr, "\iota_A"] \arrow[dd, "0_{AB}"'] & & A \\ & A \oplus B \arrow[ur, "\pi_A"] \arrow[dl, "\pi_B"] & \\ B & & B \arrow[ll, "\mathsf{id}_B"] \arrow[ul, "\iota_B"'] \arrow[uu, "0_{BA}"'] \end{tikzcd}$$

We summarize all this by saying that the direct sum is a *categorical biproduct* in Ab.

Addition of Homomorphisms in Ab. Let $A, B \in \mathsf{Ab}$ and consider any two homomorphisms $\varphi_1, \varphi_2 : A \to B$. We define the function $\varphi_1 + \varphi_2 : A \to B$ by "pointwise addition": for all $a \in A$ we set

$$(\varphi_1 + \varphi_2)(a) := \varphi_1(a) + \varphi_2(a).$$

The fact that B is an **abelian** group guarantees that this function $\varphi_1 + \varphi_2 : A \to B$ is actually a group homomorphism. ///

So what? Well, this tells us that the "Hom set" $\mathsf{Hom}_{\mathsf{Ab}}(A, B)$ is more than just a set; it's an abelian group. The group operation is pointwise addition and the identity element is the zero homomorphism $0_{AB} \in \mathsf{Hom}_{\mathsf{Ab}}(A, B)$.

Exercise: Check this.

And not only are the Hom sets abelian groups, but the composition functions respect this structure.

Biadditivity of Composition in Ab. Let $A, B, C \in \mathsf{Ab}$ and consider four homomorphisms: $\varphi_1, \varphi_2 \in \mathsf{Hom}_{\mathsf{Ab}}(B, C)$ and $\varphi_3, \varphi_4 \in \mathsf{Hom}_{\mathsf{Ab}}(A, B)$. Then we have

- $\varphi_1 \circ (\varphi_3 + \varphi_4) = (\varphi_1 \circ \varphi_3) + (\varphi_1 \circ \varphi_4)$
- $(\varphi_1 + \varphi_2) \circ \varphi_3 = (\varphi_1 \circ \varphi_3) + (\varphi_2 \circ \varphi_3)$

Proof: We will just prove the first statement. For all $a \in A$ we have

$$\begin{aligned}(\varphi_1 \circ (\varphi_3 + \varphi_4))(a) &= \varphi_1((\varphi_3 + \varphi_4)(a)) \\ &= \varphi_1(\varphi_3(a) + \varphi_4(a)) \\ &= \varphi_1(\varphi_3(a)) + \varphi_1(\varphi_4(a)) \\ &= (\varphi_1 \circ \varphi_3)(a) + (\varphi_1 \circ \varphi_4)(a) \\ &= [(\varphi_1 \circ \varphi_3) + (\varphi_1 \circ \varphi_4)](a).\end{aligned}$$

///

We will capture these two theorems with a definition.

Definition of Ab-Category. Let $\mathcal{C}$ be a category. We will call this an Ab-*category* (also known as an Ab-*enriched category*) if for each pair of objects $x, y \in \mathcal{C}$ the Hom set $\mathsf{Hom}_{\mathcal{C}}(x, y)$ is an abelian group, and for each triple of objects $x, y, z \in \mathcal{C}$ the composition function

$$\circ : \mathsf{Hom}_{\mathcal{C}}(y, z) \times \mathsf{Hom}_{\mathcal{C}}(x, y) \to \mathsf{Hom}_{\mathcal{C}}(x, y)$$

is *biadditive*, i.e., if its restriction to each factor is a group homomorphism. ///

Technical Remark: Observe that **a biadditive function of abelian groups** $A \times B \to C$ **is not the same as a group homomorphism** $A \oplus B \to C$. For example, the multiplication function $\mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$ is biadditive because $(a + b)c = ac + bc$ and $a(b + c) = ab + ac$, but it is not a homomorphism because $(a + b)(c + d) \neq ac + bd$! However, we will see later that there **does exist** a (necessarily unique) abelian group $A \otimes B \in \mathsf{Ab}$ with the property that biadditive functions $A \times B \to C$ are "the same as" homomorphisms $A \otimes B \to C$. For this reason, most sources define Ab-categories by requiring that composition is a group homomorphim:

$$\circ : \mathsf{Hom}_{\mathcal{C}}(y, z) \otimes \mathsf{Hom}_{\mathcal{C}}(x, y) \to \mathsf{Hom}_{\mathcal{C}}(x, z).$$

More generally, one can define $\mathcal{V}$-*categories* (in which the Hom sets are objects in $\mathcal{V}$) as long as the category $\mathcal{V}$ has a suitable operation $\otimes : \mathcal{V} \times \mathcal{V} \to \mathcal{V}$. In this language, our usual "categories" become "Set-categories", with $\otimes$ given by the Cartesian product of sets. ///

Just as Set is the prototypical Set-category, Ab is the prototypical Ab-category.

A Ringoid With One Object

In this brief section we examine an important and surprising special case of Ab-categories.

Given any category $\mathcal{C}$ and any object $x \in \mathcal{C}$ we define the set of *endomorphisms* of x:

$$\mathsf{End}_{\mathcal{C}}(x) := \mathsf{Hom}_{\mathcal{C}}(x, x).$$

Note that we can think of $\mathsf{End}_{\mathcal{C}}(x) \subseteq \mathcal{C}$ as the full subcategory induced by the single object x. But recall that a category with one object is called a *monoid*. We conclude that $(\mathsf{End}_{\mathcal{C}}(x), \circ, \mathsf{id}_x)$ is a monoid with the associative operation $\circ$ and the identity element id_x.

Now if $A \in \mathsf{Ab}$ is an abelian group then the set of endomorphisms $\mathsf{End}_{\mathsf{Ab}}(A)$ has **two different algebraic structures**. As in any category, this set is a monoid:

$$(\mathsf{End}_{\mathsf{Ab}}(A), \circ, \mathsf{id}_A).$$

Since Ab is an Ab-category there is also an abelian group structure:

$$(\mathsf{End}_{\mathsf{Ab}}(A), +, 0_A).$$

Furthermore, the "biaddivity of composition" says precisely that the operation "$\circ$" distributes over the operation "$+$". In other words, we have a *ring* of endomorphisms:

$$(\mathsf{End}_{\mathsf{Ab}}(A), +, \circ, 0_A, \mathsf{id}_A).$$

But this is surprising because we never officially defined rings! Somehow the "ring" concept just emerged naturally from the category of abelian groups. Let me emphasize this:

The Endomorphisms of an Abelian Group Form a Ring.

In fact we will take this as the motivating example of a ring. And if you look carefully you will see that the abstract definition of rings is hiding above in plain sight.

Definition of Ringoid/Ring.

- A *ringoid* is a small Ab-category.
- A *ring* is a ringoid with one object.

///

Exercise: Verify that this definition of ring is equivalent to the one you know.

I'm not trying to be funny here; I actually think that this is the correct definition of rings. In the next section we will compare the abstract definition of a ring (a ringoid with one object) to the motivating example of a ring (endomorphisms of an abelian group).

From M-Sets to R-Modules

Now that we have defined rings, we want to define modules. The correct way to do this is to begin with the concept of a monoid acting on a group.

Recall that we can think of any abstract monoid $(M, \circ, \mathsf{id})$ as a category with one object. To be formal we will denote this category by $\mathsf{B}M$ and refer to its single object as "$*$". Thus we have $\mathsf{Obj}(\mathsf{B}M) = \{*\}$ and $\mathsf{Arr}(\mathsf{B}M) = \mathsf{End}_{\mathsf{B}M}(*) = M$. Now consider any functor from $\mathsf{B}M$ into the category of sets:

$$F : \mathsf{B}M \to \mathsf{Set}.$$

This functor consists of a family of functions:

- A function on objects, which assigns to the object $* \in \mathsf{B}M$ a set $F(*) \in \mathsf{Set}$,
- A function on arrows, which sends endomorphisms of $*$ to endomorphisms of $F(*)$:

$$F : \mathsf{End}_{\mathsf{B}M}(*) \to \mathsf{End}_{\mathsf{Set}}(F(*)).$$

If we define $X := F(*)$ then this is just a function $F : M \to \mathsf{End}_{\mathsf{Set}}(X)$, which must satisfy the two axioms for functors:

(i) Identity: $F(\mathsf{id}) = F(\mathsf{id}_*) = \mathsf{id}_{F(*)} = \mathsf{id}_X$,

(ii) Composition: For all $\alpha, \beta \in M$ we have $F(\alpha \circ \beta) = F(\alpha) \circ F(\beta)$.

If we abuse notation by writing the function $F(\alpha) : X \to X$ simply as $\alpha : X \to X$ then you will recognize these two axioms as the definition of a *monoid acting on a set* $M \circlearrowright X$:

(i) For $\mathsf{id} \in M$ and all $x \in X$ we have $\mathsf{id}(x) = x$,

(ii) For all $\alpha, \beta \in M$ and all $x \in X$ we have $(\alpha \circ \beta)(x) = \alpha(\beta(x))$.

In summary, if M is a monoid then a functor $F : \mathsf{B}M \to \mathsf{Set}$ is equivalent to a specific set $X \in \mathsf{Set}$ carrying a specific action $M \circlearrowright X$. We refer to the pair $M \circlearrowright X$ as an *M-set* and we refer to the functor category $M\text{-}\mathsf{Set} := \mathsf{Set}^{\mathsf{B}M}$ as the *category of M-sets*.

Recall that the arrows in $\mathsf{Set}^{\mathsf{B}M}$ are the natural transformations. Given two functors $F_1, F_2 : \mathsf{B}M \to \mathsf{Set}$ corresponding to monoid actions $M \circlearrowright X_1$ and $M \circlearrowright X_2$, one can check that a natural transformation $\Phi : F_1 \Rightarrow F_2$ is the same as set function $\Phi : X_1 \to X_2$ with the property that for all $\alpha \in M$ we have a commutative square:

$$\begin{tikzcd}
X_1 \arrow[r, "\Phi"] & X_2 \\
X_1 \arrow[u, "\alpha"] \arrow[r, "\Phi"'] & X_2 \arrow[u, "\alpha"']
\end{tikzcd}$$

Exercise: Check this.

In other words, a homomorphism of M-sets is a function Φ satisfying $\Phi\alpha = \alpha\Phi$ for all $\alpha \in M$. Physicists usually refer to such functions Φ as *intertwiners*.

Recall that a small category in which all arrows are isomorphisms is called a *groupoid*. If G is a group (thought of as a monoid) then the category $\mathsf{B}G$ is the corresponding groupoid with one object. At this point some people will joke that categories should really be called "monoid-oids".

Now let $(R, +, \circ, 0, \mathsf{id})$ be abstract ring. That is, suppose that $(R, \circ, \mathsf{id})$ is a monoid, $(R, +, 0)$ is an abelian group, and "$\circ$" distributes over "$+$". As above, we can encode the monoid $(R, \circ, \mathsf{id})$ as a category $\mathsf{B}R$ with one object. Then to incorporate the abelian group structure we must give $\mathsf{B}R$ the structure of Ab-category (or ringoid). The map $R \mapsto \mathsf{B}R$ thus assigns to each abstract ring R the corresponding ringoid $\mathsf{B}R$ with one object.

In this case, functors $F : \mathsf{B}R \to \mathsf{Set}$ correspond to actions of the underlying monoid $(R, \circ, \mathsf{id})$; they do not see the full ring structure of R. To capture the full structure of R we should instead consider "Ab-functors" from $\mathsf{B}R$ into some Ab-category.

Definition of Ab-Functors. Let $\mathcal{C}$ and $\mathcal{D}$ be Ab-categories and let $F : \mathcal{C} \to \mathcal{D}$ be a functor. We say that F is an Ab-*functor* if for each pair of objects $c_1, c_2 \in C$ the function

$$F : \mathsf{Hom}_{\mathcal{C}}(c_1, c_2) \to \mathsf{Hom}_{\mathcal{D}}(F(c_1), F(c_2))$$

is a group homomorphism. ///

[Remark: The old functors can be called Set-functors.]

Now we are ready to define modules.

Definition of R-Modules. Let R be an abstract ring and let $\mathsf{B}R$ be the corresponding Ab-category (ringoid) with one object. We think of the opposite category $\mathsf{B}R^{\mathsf{op}}$ as the ring R with the "order of multiplication" reversed. Now we define a *left R-module* as an Ab-functor

$$F : \mathsf{B}R \to \mathsf{Ab},$$

and we define a *right R-module* as an Ab-functor

$$F : \mathsf{B}R^{\mathsf{op}} \to \mathsf{Ab}.$$

We let $R\text{-}\mathsf{Mod} \subseteq \mathsf{Ab}^{\mathsf{B}R}$ and $\mathsf{Mod}\text{-}R \subseteq \mathsf{Ab}^{\mathsf{B}R^{\mathsf{op}}}$ denote the full subcategories whose objects are Ab-functors and whose arrows are (usual) natural transformations. ///

[Remark: We don't need to define "Ab-natural transformations" today because the categories $\mathsf{B}R$ and $\mathsf{B}R^{\mathsf{op}}$ have only one object.]

Exercise: Verify that this definition is equivalent to the one you know.

Thus R-modules are just the "Ab-version" of groups/monoids acting on sets. Let's discuss some philosophy.

The Philosophy of Representation Theory. If G is a group/monoid, then we should study it by considering various Set-functors $F : \mathsf{B}G \to \mathcal{C}$ into various Set-categories $\mathcal{C}$. Such a functor is called a *representation of G in $\mathcal{C}$*. The prototype is the category of sets $\mathcal{C} = \mathsf{Set}$. Analogously, if R is a ring, then we should study it by considering various Ab-functors $F : \mathsf{B}R \to \mathcal{A}$ into various Ab-categories $\mathcal{A}$. Such a functor is called a *representation of R in $\mathcal{A}$*. The prototype is the category of abelian groups $\mathcal{A} = \mathsf{Ab}$.

Later we will define "representations of algebras", which simultaneously generalizes these two points of view.

Monic and Epic Arrows

If we want to study a ring R through its categories of representations then we need to know what kind of categories these are. It is a general phenomenon that the functor category $\mathcal{D}^{\mathcal{C}}$ shares many properties with its target category $\mathcal{D}$. Thus, the category of G-sets $\mathsf{Set}^{\mathsf{B}G}$ is similar to the category of sets Set (both are examples of "toposes", which we will not define in this class). Analogously, the category of R-modules $R\text{-}\mathsf{Mod} \subseteq \mathsf{Ab}^{\mathsf{B}R}$ is similar to the category of abelian groups Ab. Both $R\text{-}\mathsf{Mod}$ and Ab are examples of "abelian categories"; our next goal is to define this concept.

Of course, we already know that Ab is an Ab-category. But this is a weak structure that does not capture Ab very well. It turns out that the other key properties of Ab have to do with "monic" and "epic" arrows. These are the categorical generalizations of "injective" and "surjective" functions.

Definition of Monic and Epic Arrows. Let $\mathcal{C}$ be a category. We say that an arrow $\iota \in \mathsf{Arr}(\mathcal{C})$ is *monic* if it is **left cancellable**. That is, if for all $\alpha, \beta \in \mathsf{Arr}(\mathcal{C})$ such that $\iota \circ \alpha$ and $\iota \circ \beta$ are defined we have

$$\iota \circ \alpha = \iota \circ \beta \quad \Longleftrightarrow \quad \alpha = \beta.$$

More specifically we say that $\iota \in \mathsf{Arr}(\mathcal{C})$ is *split monic* if it has a **left inverse** $\sigma \in \mathsf{Arr}(\mathcal{C})$ such that $\sigma \circ \iota = \mathsf{id}$.

Dually, we say that $\pi \in \mathsf{Arr}(\mathcal{C})$ is *epic* if it is **right cancellable**. That is, if for all $\alpha, \beta \in \mathsf{Arr}(\mathcal{C})$ such that $\alpha \circ \pi$ and $\beta \circ \pi$ are defined we have

$$\alpha \circ \pi = \beta \circ \pi \quad \Longleftrightarrow \quad \alpha = \beta.$$

More specifically we say that $\pi \in \mathsf{Arr}(\mathcal{C})$ is *split epic* if it has a **right inverse** $\rho \in \mathsf{Arr}(\mathcal{C})$ such that $\pi \circ \rho = \mathsf{id}$. ///

[A guide to notation: The letters ι, σ, π and ρ stand for "injection", "section", "projection" and "retraction", respectively.]

Motivating Example. Let $\varphi : X \to Y$ be an arrow in the category of sets. Then

$$\begin{aligned} \varphi \text{ is monic} \quad &\Longleftrightarrow \quad \varphi \text{ is an injective function,} \\ \varphi \text{ is epic} \quad &\Longleftrightarrow \quad \varphi \text{ is a surjective function.} \end{aligned}$$

Proof: Consider two sets $X, Y \in \mathsf{Set}$ and let $\iota : X \to Y$ be an injective function. To show that ι is monic consider any two functions $\alpha, \beta : Z \to X$ such that $\iota \circ \alpha = \iota \circ \beta$. Then for all $x \in X$ we have $\iota(\alpha(x)) = \iota(\beta(x))$, which since ι is injective implies that $\alpha(x) = \beta(x)$. Since this is true for all $x \in X$ we conclude that $\alpha = \beta$. Conversely, let $\iota : X \to Y$ be a monic function. To show that ι is injective consider any elements $x_1, x_2 \in X$ such that $\iota(x_1) = \iota(x_2)$. Now let $\{*\}$ be a set with one element and define the functions $\alpha_1, \alpha_2 : \{*\} \to X$ by $\alpha_1(*) := x_1$ and $\alpha_2(*) := x_2$. Then we have $\iota \circ \alpha_1 = \iota \circ \alpha_2$, which since ι is monic implies that $\alpha_1 = \alpha_2$. We conclude that $x_1 = \alpha_1(*) = \alpha_2(*) = x_2$.

Now consider two sets $X, Y \in \mathsf{Set}$ and let $\pi : X \to Y$ be a surjective function. To show that π is epic consider any functions $\alpha, \beta : Y \to Z$ such that $\alpha \circ \pi = \beta \circ \pi$. Then for any element $y \in Y$, surjectivity of π implies that there exists an element with $\pi(x) = y$ and hence

$$\alpha(y) = \alpha(\pi(x)) = (\alpha \circ \pi)(x) = (\beta \circ \pi)(x) = \beta(\pi(x)) = \beta(y).$$

Since this is true for all $y \in Y$ we conclude that $\alpha = \beta$. Conversely, let $\pi : X \to Y$ be an epic function. To prove that π is surjective consider a set $\Omega = \{0, 1\}$ with two elements. Now let $\alpha_1 : Y \to \Omega$ be the constant function defined by $\alpha_1(y) := 1$ for all $y \in Y$, and let $\alpha_2 : Y \to \Omega$ be the "characteristic function of the image":

$$\alpha_2(y) := \begin{cases} 1 & \text{if } y \in \mathsf{im}\,\pi \\ 0 & \text{if } y \notin \mathsf{im}\,\pi \end{cases}.$$

We clearly have $\alpha_1(\pi(x)) = 1 = \alpha_2(\pi(x))$ for all $x \in X$ and hence $\alpha_1 \circ \pi = \alpha_2 \circ \pi$. Then since π is epic this implies that $\alpha_1 = \alpha_2$ and hence $\mathsf{im}\,\pi = Y$. We conclude that π is surjective. ///

[Remark: In the category of sets, every monic arrow $\iota : X \to Y$ splits. To see this, define a section $\sigma : Y \to X$ by sending each $y \in \mathsf{im}\,\iota$ to its unique preimage and sending each $y \notin \mathsf{im}\,\iota$

to some arbitrary element of X (assume that $X \neq \varnothing$). The statement that "every epic arrow in Set splits" is equivalent to the Axiom of Choice.]

The situation is not so straightforward in other concrete categories.

Exercise: Let $\mathcal{C} \subseteq \mathsf{Set}$ be a concrete category and consider any arrow $\varphi : X \to Y$ in $\mathcal{C}$.

(a) Copy the proof from above to show that

$$\begin{aligned} \varphi \text{ is monic} \quad &\Longrightarrow \quad \varphi \text{ is injective,} \\ \varphi \text{ is epic} \quad &\Longrightarrow \quad \varphi \text{ is surjective.} \end{aligned}$$

(b) Let $U : \mathcal{C} \to \mathsf{Set}$ be the "forgetful functor" that assigns to each object $X \in \mathcal{C}$ its underlying set $U(X) \in \mathsf{Set}$, and assume that U has a left adjoint "free functor" $F : \mathsf{Set} \to \mathcal{C}$. Use the "free object" $F(\{*\}) \in \mathcal{C}$ to prove that

$$\varphi \text{ is injective} \quad \Longrightarrow \quad \varphi \text{ is monic.}$$

(c) It is quite common for epic arrows to be non-surjective. For example, prove that the inclusion ring homomorphism $\iota : \mathbb{Z} \to \mathbb{Q}$ is epic even though it is clearly not surjective. The proof above fails because the category of rings doesn't have a "subobject classifier" such as $\Omega = \{0, 1\}$.

///

In categories (such as the category of rings) where epic arrows are not just surjective functions, we are forced to conclude that the concept of "surjective function" is not so natural.

My reason for introducing monic and epic arrows at this time is to discuss their relationship to kernels and cokernels in the category of abelian groups. The concepts of kernel and cokernel require a zero object, but they have a slight generalization that can be defined in any category.

Definition of Equalizer/Coequalizer. Let $\mathcal{C}$ be any category and consider two parallel arrows $\alpha, \beta : x \to y$. Let $\mathcal{I}$ be a category with two objects $\mathsf{Obj}(\mathcal{I}) = \{1, 2\}$ and two non-identity arrows $\mathsf{Hom}_{\mathcal{I}}(1, 2) = \{\delta, \varepsilon\}$, and define a (non-commutative) diagram $D : \mathcal{I} \to \mathcal{C}$ by

$$D(1) = x, \quad D(2) = y, \quad , D(\delta) = \alpha, \quad D(\varepsilon) = \beta.$$

The limit of D, if it exists, is called the *equalizer of α and β* and the colimit of D, if it exists, is called the *coequalizer of α and β*.

More explicitly, the equalizer is a pair $(\mathsf{eq}(\alpha, \beta), \iota)$ consisting of an object $\mathsf{eq}(\alpha, \beta) \in \mathcal{C}$ and an arrow $\iota : \mathsf{eq}(\alpha, \beta) \to x$ satisfying $\alpha \circ \iota = \beta \circ \iota$. Furthermore, if $\mu : z \to x$ is any arrow satisfying $\alpha \circ \mu = \beta \circ \mu$ then there exists a unique arrow $z \to \mathsf{eq}(\alpha, \beta)$ making the following diagram commute:

$$\begin{tikzcd} & & x \arrow[dr, "\alpha"] & \\ z \arrow[r, dashed, "\exists!"] \arrow[urr, bend left, "\mu"] \arrow[drr, bend right, "\mu"'] & \mathsf{eq}(\alpha,\beta) \arrow[ur, "\iota"] \arrow[dr, "\iota"'] & & y \\ & & x \arrow[ur, "\beta"'] & \end{tikzcd}$$

Dually, the cokernel is a pair $(\mathsf{coker}(\alpha,\beta),\pi)$ making the following diagram commute:

$$
\begin{tikzcd}
 & y \arrow[dr, "\pi"] \arrow[drr, bend left, "\mu"] & & \\
x \arrow[ur, "\alpha"] \arrow[dr, "\beta"'] & & \mathsf{coeq}(\alpha,\beta) \arrow[r, dashed, "\exists!"] & z \\
 & y \arrow[ur, "\pi"'] \arrow[urr, bend right, "\mu"'] & &
\end{tikzcd}
$$

///

Motivating Example. Let $\alpha,\beta : X \to Y$ be functions between sets, and consider the subset of X on which α and β agree:

$$E := \{x \in X : \alpha(x) = \beta(x)\} \subseteq X.$$

I claim that the inclusion function $\iota : E \to X$ is the category-theoretic equalizer. Indeed, if $\mu : Z \to X$ is any function satisfying $\alpha \circ \mu = \beta \circ \mu$ then we want to show that there exists a unique function $\bar{\mu} : Z \to E$ making the following diagram commute:

$$
\begin{tikzcd}
 & & X \arrow[dr, "\alpha"] & \\
Z \arrow[urr, bend left, "\mu"] \arrow[r, "\bar{\mu}"] \arrow[drr, bend right, "\mu"'] & E \arrow[ur, "\iota"] \arrow[dr, "\iota"'] & & Y \\
 & & X \arrow[ur, "\beta"'] &
\end{tikzcd}
$$

That is, for each $z \in Z$ we want to show that there exists a unique element $\bar{\mu}(z) \in E$ with the property that $\iota(\bar{\mu}(z)) = \mu(z)$. And this is certainly true because ι is injective.

The set-theoretic coequalizer is harder to construct. Here's a sketch: Equivalence relations on the set Y are the same as partitions. Let $\Pi(Y)$ be the collection of all partitions of Y. One can show that this is a complete lattice with $\wedge$ given by the coarsest common refinement and $\vee$ given by the finest common coarsening. Now let $\sim \subseteq Y \times Y$ be the coarsest common refinement of all partitions on Y with the property that $\alpha(x) \sim \beta(x)$ for all $x \in X$. Then the category-theoretic coequalizer is given by the quotient function $\pi : Y \to (Y/\sim)$ sending each element $y \in Y$ to its equivalence class $[y] \in (Y/\sim)$. (Details omitted.) ///

In the above example we see that every equalizer in Set is injective and every coequalizer in Set is surjective. This phenomenon generalizes to arbitrary categories as follows.

Equalizers are Monic, Coequalizers are Epic. Consider two parallel arrows $\alpha,\beta : x \to y$ in a category $\mathcal{C}$. Then the equalizer $\iota : \mathsf{eq}(\alpha,\beta) \to x$, if it exists, is a monic arrow. Dually, the coequalizer $\pi : y \to \mathsf{coeq}(\alpha,\beta)$, if it exists, is an epic arrow.

Proof: Let $\iota : \mathsf{eq}(\alpha,\beta) \to x$ be the equalizer of the arrows $\alpha,\beta : x \to y$ in $\mathcal{C}$. To prove that ι is monic, suppose that we have $\iota \circ \varphi = \iota \circ \psi$ for some arrows $\varphi,\psi : z \to \mathsf{eq}(\alpha,\beta)$. In this case we want to show that $\varphi = \psi$.

So let $\mu : z \to x$ be defined as the common arrow $\mu := \iota \circ \varphi = \iota \circ \psi$ so that the following diagram commutes:

$$\begin{tikzcd}
 & & x \arrow[dr, "\alpha"] & \\
z \arrow[r, dashed] \arrow[r, bend left, "\varphi"] \arrow[r, bend right, "\psi"'] \arrow[urr, bend left, "\mu"] \arrow[drr, bend right, "\mu"'] & \mathsf{eq}(\alpha, \beta) \arrow[ur, "\iota"] \arrow[dr, "\iota"'] & & y \\
 & & x \arrow[ur, "\beta"'] &
\end{tikzcd}$$

Finally, from the uniqueness of the dotted arrow (which is the **definition** of the equalizer) we conclude that $\varphi = \psi$. The proof for coequalizers is similar. ///

Another name for the equalizer/coequalizer is the difference kernel/cokernel. Why?

Definition of Kernel/Cokernel. Let $\mathcal{C}$ be a category with a zero object $0 \in \mathcal{C}$, so that between any ordered pair of objects $x, y \in \mathcal{C}$ there is a zero arrow $0_{xy} : x \to y$. Let $\mathcal{I}$ be a category with two objects $\mathsf{Obj}(\mathcal{I}) = \{1, 2\}$ and two non-identity arrows $\mathsf{Hom}_{\mathcal{I}}(1, 2) = \{\delta, \varepsilon\}$. Finally, let $\varphi : x \to y$ be any arrow in $\mathcal{C}$ and consider the diagram $D : \mathcal{I} \to \mathcal{C}$ defined by

$$D(1) = x, \quad D(2) = y, \quad , D(\delta) = \varphi, \quad D(\varepsilon) = 0_{xy}.$$

The limit of D, if it exists, is called the *kernel of* φ and the colimit of D, if it exists, is called the *cokernel of* φ.

More explicitly, the kernel of φ is a pair $(\ker \varphi, \iota)$ where $\iota : \ker \varphi \to x$ is an arrow satisfying $\varphi \circ \iota = 0_{\ker \varphi, y}$, and if $\mu : z \to x$ is any arrow in $\mathcal{C}$ satisfying $\varphi \circ \mu = 0_{zy}$ then there exists a unique arrow $z \to \ker \varphi$ making the following diagram commute:

$$\begin{tikzcd}
z \arrow[dr, "\mu"] \arrow[drr, bend left, "0_{zy}"] \arrow[dd, dashed, "\exists!"'] & & \\
 & x \arrow[r, "\varphi"] & y \\
\ker \varphi \arrow[ur, "\iota"'] \arrow[urr, bend right, "0_{\ker \varphi, y}"'] & &
\end{tikzcd}$$

Dually, the cokernel of φ is a pair $(\mathsf{coker}\, \varphi, \pi)$ making the following diagram commute:

$$\begin{tikzcd}
 & & \mathsf{coker}\, \varphi \arrow[dd, dashed, "\exists!"] \\
x \arrow[r, "\varphi"] \arrow[urr, bend left, "0_{x, \mathsf{coker}\, \varphi}"] \arrow[drr, bend right, "0_{xz}"'] & y \arrow[ur, "\pi"] \arrow[dr, "\mu"'] & \\
 & & z
\end{tikzcd}$$

///

Consider two groups $G, H \in \mathsf{Grp}$ and a group homomorphism $\varphi : G \to H$. You verified on a previous exercise that the kernel of φ exists in Grp and it is given by the inclusion homomorphism from the set-theoretic kernel $\iota : \ker \varphi \to G$. If H is **abelian** then you verified that the cokernel of φ also exists and is given by the quotient homomorphism $\pi : H \to H/\mathsf{im}\, \varphi$, where $\mathsf{im}\,\varphi \subseteq H$ is the set-theoretic image. If H is not abelian then the image $\mathsf{im}\,\varphi \subseteq H$ might not be a **normal** subgroup. In this case the cokernel is the quotient by the "normal closure" of the image, $\pi : H \to H/\langle \mathsf{im}\,\varphi \rangle$ (i.e., the smallest normal subgroup containing $\mathsf{im}\,\varphi$).

Note that the inclusion homomorphism $\iota : \ker\varphi \to G$ is injective and the quotient homomorphism $\pi : H \to H/\langle \mathsf{im}\,\varphi\rangle$ is surjective. This phenomenon generalizes to arbitrary categories.

Kernels are Monic, Cokernels are Epic. Let $\mathcal{C}$ be any category with a zero object $0 \in \mathcal{C}$ and let $\varphi : x \to y$ be any arrow in $\mathcal{C}$. Then the kernel $\iota : \ker\varphi \to x$, if it exists, is a monic arrow and the cokernel $\pi : y \to \mathsf{coker}\,\varphi$, if it exists, is an epic arrow.

Proof: Let $\iota : \ker\varphi \to x$ be the kernel and consider any two arrows $\alpha, \beta : z \to \ker\varphi$ such that $\mu := \iota \circ \alpha = \iota \circ \beta$. Since $\varphi \circ \iota = 0_{\ker\varphi,y}$ we find that

$$\varphi \circ \mu = \varphi \circ (\iota \circ \alpha) = (\varphi \circ \iota) \circ \alpha = 0_{\ker\varphi,y} \circ \alpha.$$

Then it follows from the commutative diagram

$$\begin{tikzcd}
z \arrow[r, "\alpha"] \arrow[rr, bend left, "0_{\ker\varphi,y}\circ\alpha"] \arrow[dr, "\exists!"'] & \ker\varphi \arrow[r, "0_{\ker\varphi,y}"] \arrow[d, "\exists!"] & y \\
& 0 \arrow[ur, "\exists!"'] &
\end{tikzcd}$$

that $\varphi \circ \mu = 0_{\ker\varphi,y} \circ \alpha = 0_{zy}$. Putting all of this together produces a commutative diagram:

$$\begin{tikzcd}
z \arrow[dd, bend right, "\alpha"'] \arrow[dd, bend left, "\beta"] \arrow[dd, dashed] \arrow[dr, "\mu"] \arrow[drr, bend left, "0_{zy}"] & & \\
& x \arrow[r, "\varphi"] & y \\
\ker\varphi \arrow[ur, "\iota"'] \arrow[urr, bend right, "0_{\ker\varphi,y}"'] & &
\end{tikzcd}$$

Finally, from the uniqueness of the dotted arrow (which is the **definition** of the kernel) we conclude that $\alpha = \beta$, hence ι is monic. The proof for cokernels is similar. ///