CHAPTER 1

Just enough category theory to be dangerous

Was mich nicht umbringt, macht mich stärker.

That which does not kill me, makes me stronger.
— *F. Nietzsche* [**N**, aphorism number 8]

Before we get to any interesting geometry, we need to develop a language to discuss things cleanly and effectively. This is best done in the language of categories. There is not much to know about categories to get started; it is just a very useful language. Like all mathematical languages, category theory comes with an embedded logic, which allows us to abstract intuitions in settings we know well to far more general situations.

Our motivation is as follows. We will be creating some new mathematical objects (such as schemes, and certain kinds of sheaves), and we expect them to act like objects we have seen before. We could try to nail down precisely what we mean by "act like", and what minimal set of things we have to check in order to verify that they act the way we expect. Fortunately, we don't have to — other people have done this before us, by defining key notions, such as *abelian categories*, which behave like modules over a ring.

Our general approach will be as follows. I will try to tell you what you need to know, and no more. (This I promise: if I use the word "topoi", you can shoot me.) I will begin by telling you things you already know, and describing what is essential about the examples, in a way that we can abstract a more general definition. We will then see this definition in less familiar settings, and get comfortable with using it to solve problems and prove theorems.

1.0.1. *Example: product.* For example, we will define the notion of *product* of schemes. We could just give a definition of product, but then you should want to know why this precise definition deserves the name of "product". As a motivation, we revisit the notion of product in a situation we know well: (the category of) sets. One way to define the product of sets U and V is as the set of ordered pairs $\{(u, v) \ : \ u \in U, v \in V\}$. But someone from a different mathematical culture might reasonably define it as the set of symbols $\{ {u \atop v} : \ u \in U, v \in V\}$. These notions are "obviously the same". Better: there is "an obvious bijection between the two".

This can be made precise by giving a better definition of product, in terms of a *universal property*. Given two sets M and N, a product is a set P, along with maps $\mu \colon P \to M$ and $\nu \colon P \to N$, such that for *any set* P' *with maps* $\mu' \colon P' \to M$ *and*

$\nu' \colon P' \to N$, these maps must factor *uniquely* through P:

(1.0.1.1)
$$\begin{tikzcd}
P' \arrow[dr, dotted, "\exists!"] \arrow[drr, "\nu'"] \arrow[ddr, "\mu'"'] & & \\
& P \arrow[r, "\nu"'] \arrow[d, "\mu"] & N \\
& M &
\end{tikzcd}$$

(The symbol $\exists$ means "there exists", and the symbol ! means "unique".) Thus a **product** is a *diagram*

$$\begin{tikzcd}
P \arrow[r, "\nu"] \arrow[d, "\mu"'] & N \\
M &
\end{tikzcd}$$

and not just a set P, although the maps μ and ν are often left implicit.

This definition agrees with the traditional definition, with one twist: there isn't just a single product; but any two products come with a *unique* isomorphism between them. In other words, the product is unique up to unique isomorphism. Here is why: if you have a product

$$\begin{tikzcd}
P_1 \arrow[r, "\nu_1"] \arrow[d, "\mu_1"'] & N \\
M &
\end{tikzcd}$$

and I have a product

$$\begin{tikzcd}
P_2 \arrow[r, "\nu_2"] \arrow[d, "\mu_2"'] & N \\
M &
\end{tikzcd}$$

then by the universal property of my product (letting (P_2, μ_2, ν_2) play the role of (P, μ, ν), and (P_1, μ_1, ν_1) play the role of (P', μ', ν') in (1.0.1.1)), there is a unique map $f \colon P_1 \to P_2$ making the appropriate diagram commute (i.e., $\mu_1 = \mu_2 \circ f$ and $\nu_1 = \nu_2 \circ f$). Similarly by the universal property of your product, there is a unique map $g \colon P_2 \to P_1$ making the appropriate diagram commute. Now consider the universal property of my product, this time letting (P_2, μ_2, ν_2) play the role of both (P, μ, ν) and (P', μ', ν') in (1.0.1.1). There is a unique map $h \colon P_2 \to P_2$ such that

$$\begin{tikzcd}
P_2 \arrow[dr, "h"] \arrow[drr, "\nu_2"] \arrow[ddr, "\mu_2"'] & & \\
& P_2 \arrow[r, "\nu_2"'] \arrow[d, "\mu_2"] & N \\
& M &
\end{tikzcd}$$

commutes. However, I can name two such maps: the identity map id_{P_2}, and $f \circ g$. Thus $f \circ g = \mathrm{id}_{P_2}$. Similarly, $g \circ f = \mathrm{id}_{P_1}$. Thus the maps f and g arising from the universal property are bijections. In short, there is a unique bijection between P_1 and P_2 preserving the "product structure" (the maps to M and N). This gives us the right to name any such product $M \times N$, since any two such products are uniquely identified.

This definition has the advantage that it works in many circumstances, and once we define categories, we will soon see that the above argument applies verbatim in any category to show that products, if they exist, are unique up to unique isomorphism. Even if you haven't seen the definition of category before, you can verify that this agrees with your notion of product in some category that you have seen before (such as the category of vector spaces, or the category of manifolds).

This is handy even in cases that you understand. For example, one way of defining the product of two manifolds M and N is to cut them both up into charts, then take products of charts, then glue them together. But if I cut up the manifolds in one way, and you cut them up in another, how do we know our resulting manifolds are the "same"? We could wave our hands, or make an annoying argument about refining covers, but instead, we should just show that they are "categorical products" and hence canonically the "same" (i.e., isomorphic). We will formalize this argument in §1.2.

Another set of notions we will abstract are categories that "behave like modules". We will want to define kernels and cokernels for new notions, and we should make sure that these notions behave the way we expect them to. This leads us to the definition of *abelian categories*, first defined by Grothendieck in his Tôhoku paper [**Gr1**].

In this chapter, we will give an informal introduction to these and related notions, in the hope of giving just enough familiarity to comfortably use them in practice.

1.1 Categories and functors

The introduction of the digit 0 *or the group concept was general nonsense too, and mathematics was more or less stagnating for thousands of years because nobody was around to take such childish steps...*
— *A. Grothendieck*, [**BP**, p. 4–5]

Before functoriality, people lived in caves.
— *B. Conrad*

We begin with an informal definition of categories and functors.

1.1.1. Categories.

A **category** consists of a collection of **objects**, and for each pair of objects, a set of **morphisms** (or **arrows**) between them. (For experts: technically, this is the definition of a *locally small category*. In the correct definition, the morphisms need only form a class, not necessarily a set, but see Caution 0.3.1.) Morphisms are often informally called **maps**. The collection of objects of a category $\mathscr{C}$ is often denoted $\mathrm{obj}(\mathscr{C})$, but we will usually denote the collection also by $\mathscr{C}$. If $A, B \in \mathscr{C}$, then the set of morphisms from A to B is denoted $\mathrm{Mor}(A, B)$. A morphism is often written

$f\colon A \to B$, and A is said to be the **source** of f, and B the **target** of f. (Of course, $\mathrm{Mor}(A, B)$ is taken to be disjoint from $\mathrm{Mor}(A', B')$ unless $A = A'$ and $B = B'$.)

Morphisms compose as expected: there is a composition $\mathrm{Mor}(B, C) \times \mathrm{Mor}(A, B) \to \mathrm{Mor}(A, C)$, and if $f \in \mathrm{Mor}(A, B)$ and $g \in \mathrm{Mor}(B, C)$, then their composition is denoted $g \circ f$. Composition is associative: if $f \in \mathrm{Mor}(A, B)$, $g \in \mathrm{Mor}(B, C)$, and $h \in \mathrm{Mor}(C, D)$, then $h \circ (g \circ f) = (h \circ g) \circ f$. For each object $A \in \mathscr{C}$, there is always an **identity morphism** $\mathrm{id}_A\colon A \to A$, such that when you (left- or right-)compose a morphism with the identity, you get the same morphism. More precisely, for any morphisms $f\colon A \to B$ and $g\colon B \to C$, $\mathrm{id}_B \circ f = f$ and $g \circ \mathrm{id}_B = g$. (If you wish, you may check that "identity morphisms are unique": there is only one morphism deserving the name id_A.) This ends the definition of a category.

We have a notion of **isomorphism** between two objects of a category (a morphism $f\colon A \to B$ such that there exists some — necessarily unique — morphism $g\colon B \to A$, where $f \circ g$ and $g \circ f$ are the identity on B and A respectively).

1.1.2. *Example.* The prototypical example to keep in mind is the category of sets, denoted *Sets*. The objects are sets, and the morphisms are maps of sets. (Because Russell's paradox shows that there is no set of all sets, we did not say earlier that there is a set of all objects. But as stated in §0.3, we are deliberately omitting all set-theoretic issues.)

1.1.3. *Example.* Another good example is the category Vec_k of vector spaces over a given field k. The objects are k-vector spaces, and the morphisms are linear transformations. (What are the isomorphisms?)

1.1.A. UNIMPORTANT EXERCISE. A category in which each morphism is an isomorphism is called a **groupoid**. (This notion is not important in what we will discuss. The point of this exercise is to give you some practice with categories, by relating them to an object you know well.)
(a) A perverse definition of a *group* is: a groupoid with one object. Make sense of this. (Similarly, in case you care: a perverse definition of a **monoid** is: a category with one object.)
(b) Describe a groupoid that is not a group.

1.1.B. EXERCISE. If A is an object in a category $\mathscr{C}$, show that the invertible elements of $\mathrm{Mor}(A, A)$ form a group (called the **automorphism group of** A, denoted $\mathrm{Aut}(A)$). What are the automorphism groups of the objects in Examples 1.1.2 and 1.1.3? Show that two isomorphic objects have isomorphic automorphism groups. (For readers with a topological background: if X is a topological space, then the fundamental groupoid is the category where the objects are points of X, and the morphisms $x \to y$ are paths from x to y, up to homotopy. Then the automorphism group of x_0 is the (pointed) fundamental group $\pi_1(X, x_0)$. In the case where X is connected, and $\pi_1(X)$ is not abelian, this illustrates the fact that for a connected groupoid — whose definition you can guess — the automorphism groups of the objects are all isomorphic, but not canonically isomorphic.)

1.1.4. *Example: abelian groups.* The abelian groups, along with group homomorphisms, form a category *Ab*.

1.1.5. *Important Example: Modules over a ring.* If A is a ring, then the A-modules form a category Mod_A. (This category has additional structure; it will be the prototypical example of an *abelian category*, see §1.5.) Taking $A = k$, we obtain Example 1.1.3; taking $A = \mathbb{Z}$, we obtain Example 1.1.4.

1.1.6. *Example: rings.* There is a category *Rings*, where the objects are rings, and the morphisms are maps of rings in the usual sense (maps of sets which respect addition and multiplication, and which send 1 to 1 by our conventions, §0.3).

1.1.7. *Example: topological spaces.* The topological spaces, along with continuous maps, form a category *Top*. The isomorphisms are homeomorphisms.

In all of the above examples, the objects of the categories were in obvious ways sets with additional structure (a **concrete category**, although we won't use this terminology). This needn't be the case, as the next example shows.

1.1.8. *Example: partially ordered sets.* A **partially ordered set**, (or **poset**), is a set S along with a binary relation $\geq$ on S satisfying:

(i) $x \geq x$ (reflexivity),
(ii) $x \geq y$ and $y \geq z$ imply $x \geq z$ (transitivity), and
(iii) if $x \geq y$ and $y \geq x$ then $x = y$ (antisymmetry).

A partially ordered set $(S, \geq)$ can be interpreted as a category whose objects are the elements of S, and with a single morphism from x to y if and only if $x \geq y$ (and no morphism otherwise).

A trivial example is $(S, \geq)$ where $x \geq y$ if and only if $x = y$. Another example is

(1.1.8.1)
$$\begin{tikzcd} & \bullet \arrow[d] \\ \bullet \arrow[r] & \bullet \end{tikzcd}$$

Here there are three objects. The identity morphisms are omitted for convenience, and the two non-identity morphisms are depicted. A third example is

(1.1.8.2)
$$\begin{tikzcd} \bullet \arrow[r] \arrow[d] & \bullet \arrow[d] \\ \bullet \arrow[r] & \bullet \end{tikzcd}$$

Here the "obvious" morphisms are again omitted: the identity morphisms, and the morphism from the upper left to the lower right. Similarly,

$$\cdots \longrightarrow \bullet \longrightarrow \bullet \longrightarrow \bullet$$

depicts a partially ordered set, where again, only the "generating morphisms" are depicted.

1.1.9. *Example: the category of subsets of a set, and the category of open subsets of a topological space.* If X is a set, then the subsets form a partially ordered set, where arrows are given by inclusion. (Be careful: you may be expecting the arrows to go the other way, because of Example 1.1.8.) Informally, if $U \subset V$, then we have exactly one morphism $U \to V$ in the category (and otherwise none). Similarly, if

X is a topological space, then the *open* sets form a partially ordered set, where the maps are given by inclusions.

1.1.10. *Definition.* A **subcategory** $\mathscr{A}$ of a category $\mathscr{B}$ has as its objects some of the objects of $\mathscr{B}$, and some of the morphisms of $\mathscr{B}$, such that the objects of $\mathscr{A}$ include the sources and targets of the morphisms of $\mathscr{A}$, and the morphisms of $\mathscr{A}$ include the identity morphisms of the objects of $\mathscr{A}$, and are preserved by composition. (For example, (1.1.8.1) is in an obvious way a subcategory of (1.1.8.2). Also, we have an obvious "inclusion" $i\colon \mathscr{A} \to \mathscr{B}$, which will soon be an example of a functor.)

1.1.11. Functors.

A **covariant functor** F from a category $\mathscr{A}$ to a category $\mathscr{B}$, denoted $F\colon \mathscr{A} \to \mathscr{B}$, is the following data. It is a map of objects $F\colon \operatorname{obj}(\mathscr{A}) \to \operatorname{obj}(\mathscr{B})$, and for each A_1, $A_2 \in \mathscr{A}$, and morphism $m\colon A_1 \to A_2$, a morphism $F(m)\colon F(A_1) \to F(A_2)$ in $\mathscr{B}$. We require that F preserves identity morphisms (for $A \in \mathscr{A}$, $F(\mathrm{id}_A) = \mathrm{id}_{F(A)}$), and that F preserves composition ($F(m_2 \circ m_1) = F(m_2) \circ F(m_1)$). (You may wish to verify that covariant functors send isomorphisms to isomorphisms.) A trivial example is the **identity functor** $\mathrm{id}\colon \mathscr{A} \to \mathscr{A}$, whose definition you can guess. Here are some less trivial examples.

1.1.12. *Example: a forgetful functor.* Consider the functor from the category of vector spaces (over a field k) Vec_k to *Sets*, that associates to each vector space its underlying set. The functor sends a linear transformation to its underlying map of sets. This is an example of a **forgetful functor**, where some additional structure is forgotten. Another example of a forgetful functor is $Mod_A \to Ab$ from A-modules to abelian groups, remembering only the abelian group structure of the A-module.

1.1.13. *Topological examples.* Examples of covariant functors include the fundamental group functor π_1, which sends a topological space X with choice of a point $x_0 \in X$ to a group $\pi_1(X, x_0)$ (what are the objects and morphisms of the source category?), and the ith homology functor $Top \to Ab$, which sends a topological space X to its ith homology group $H_i(X, \mathbb{Z})$. The covariance corresponds to the fact that a (continuous) morphism of pointed topological spaces $\phi\colon X \to Y$ with $\phi(x_0) = y_0$ induces a map of fundamental groups $\pi_1(X, x_0) \to \pi_1(Y, y_0)$, and similarly for homology groups.

1.1.14. *Example.* Suppose A is an object in a category $\mathscr{C}$. Then there is a functor $h^A\colon \mathscr{C} \to Sets$ sending $B \in \mathscr{C}$ to $\operatorname{Mor}(A, B)$, and sending $f\colon B_1 \to B_2$ to $\operatorname{Mor}(A, B_1) \to \operatorname{Mor}(A, B_2)$ described by

$$[g\colon A \to B_1] \longmapsto [f \circ g\colon A \to B_1 \to B_2].$$

This seemingly silly functor ends up surprisingly being an important concept.

1.1.15. Definitions. If $F\colon \mathscr{A} \to \mathscr{B}$ and $G\colon \mathscr{B} \to \mathscr{C}$ are covariant functors, then we define a functor $G \circ F\colon \mathscr{A} \to \mathscr{C}$ (the **composition** of G and F) in the obvious way. Composition of functors is associative in an evident sense.

A covariant functor $F\colon \mathscr{A} \to \mathscr{B}$ is **faithful** if for all $A, A' \in \mathscr{A}$, the map $\operatorname{Mor}_{\mathscr{A}}(A, A') \to \operatorname{Mor}_{\mathscr{B}}(F(A), F(A'))$ is injective, and **full** if it is surjective. A functor that is full and faithful is **fully faithful**. (For various philosophical reasons, the notion of "full" functor on its own is unimportant; "fully faithful" is the useful notion.) A subcategory $i\colon \mathscr{A} \to \mathscr{B}$ is a **full subcategory** if i is full. (Inclusions are always faithful, so there is no need for the phrase "faithful subcategory".) Thus a subcategory $\mathscr{A}'$ of $\mathscr{A}$ is full if and only if for all $A, B \in \operatorname{obj}(\mathscr{A}')$, $\operatorname{Mor}_{\mathscr{A}'}(A, B) = \operatorname{Mor}_{\mathscr{A}}(A, B)$. For example, the forgetful functor $\mathit{Vec}_k \to \mathit{Sets}$ is faithful, but not full; and if A is a ring, the category of finitely generated A-modules is a full subcategory of the category Mod_A of A-modules.

1.1.16. Definition. A **contravariant functor** is defined in the same way as a covariant functor, except the arrows switch directions: in the above language, $F(A_1 \to A_2)$ is now an arrow from $F(A_2)$ to $F(A_1)$. (Thus $F(m_2 \circ m_1) = F(m_1) \circ F(m_2)$, not $F(m_2) \circ F(m_1)$.)

It is wise to state whether a functor is covariant or contravariant, unless the context makes it very clear. If it is not stated (and the context does not make it clear), the functor is often assumed to be covariant.

Sometimes people describe a contravariant functor $\mathscr{C} \to \mathscr{D}$ as a covariant functor $\mathscr{C}^{\text{opp}} \to \mathscr{D}$, where $\mathscr{C}^{\text{opp}}$ is the same category as $\mathscr{C}$ except that the arrows go in the opposite direction. Here $\mathscr{C}^{\text{opp}}$ is said to be the **opposite category** to $\mathscr{C}$.

One can define fullness, etc. for contravariant functors, and you should do so.

1.1.17. *Linear algebra example.* If Vec_k is the category of k-vector spaces (introduced in Example 1.1.3), then taking duals gives a contravariant functor $(\cdot)^\vee\colon \mathit{Vec}_k \to \mathit{Vec}_k$. Indeed, to each linear transformation $f\colon V \to W$, we have a dual transformation $f^\vee\colon W^\vee \to V^\vee$, and $(f \circ g)^\vee = g^\vee \circ f^\vee$.

1.1.18. *Topological example (cf. Example 1.1.13) for those who have seen cohomology.* The ith cohomology functor $H^i(\cdot, \mathbb{Z})\colon \mathit{Top} \to \mathit{Ab}$ is a contravariant functor.

1.1.19. *Example.* There is a contravariant functor $\mathit{Top} \to \mathit{Rings}$ taking a topological space X to the ring of real-valued continuous functions on X. A morphism of topological spaces $X \to Y$ (a continuous map) induces the pullback map from functions on Y to functions on X.

1.1.20. *Example (the functor of points, cf. Example 1.1.14).* Suppose A is an object of a category $\mathscr{C}$. Then there is a contravariant functor $h_A\colon \mathscr{C} \to \mathit{Sets}$ sending $B \in \mathscr{C}$ to $\operatorname{Mor}(B, A)$, and sending the morphism $f\colon B_1 \to B_2$ to the morphism $\operatorname{Mor}(B_2, A) \to \operatorname{Mor}(B_1, A)$ via

$$[g\colon B_2 \to A] \longmapsto [g \circ f\colon B_1 \to B_2 \to A].$$

This example initially looks weird and different, but Examples 1.1.17 and 1.1.19 may be interpreted as special cases; do you see how? What is A in each case? This functor might reasonably be called the *functor of maps* (to A), but is actually known as the **functor of points**. We will meet this functor again in §1.2.11 and (in the category of schemes) in Definition 7.3.10.

1.1.21. ⋆ **Natural transformations (and natural isomorphisms) of covariant functors, and equivalences of categories.**

(This notion won't come up in an essential way until at least Chapter 7, so you shouldn't read this section until then.) Suppose F and G are two covariant functors from $\mathscr{A}$ to $\mathscr{B}$. A **natural transformation of covariant functors** $F \to G$ is the data of a morphism $m_A\colon F(A) \to G(A)$ for each $A \in \mathscr{A}$ such that for each $f\colon A \to A'$ in $\mathscr{A}$, the diagram

$$\begin{tikzcd} F(A) \arrow[r, "F(f)"] \arrow[d, "m_A"'] & F(A') \arrow[d, "m_{A'}"] \\ G(A) \arrow[r, "G(f)"'] & G(A') \end{tikzcd}$$

commutes. A **natural isomorphism** of functors is a natural transformation such that each m_A is an isomorphism. (We make analogous definitions when F and G are both contravariant.)

The data of functors $F\colon \mathscr{A} \to \mathscr{B}$ and $F'\colon \mathscr{B} \to \mathscr{A}$ such that $F \circ F'$ is naturally isomorphic to the identity functor $\mathrm{id}_{\mathscr{B}}$ on $\mathscr{B}$ and $F' \circ F$ is naturally isomorphic to $\mathrm{id}_{\mathscr{A}}$ is said to be an **equivalence of categories**. The right notion of when two categories are "essentially the same" is not *isomorphism* (a functor giving bijections of objects and morphisms) but *equivalence*. Exercises 1.1.C and 1.1.D might give you some vague sense of this. Later exercises (for example, that "rings" and "affine schemes" are essentially the same, once arrows are reversed, Exercise 7.3.E) may help too.

Two examples might make this strange concept more comprehensible. The double dual of a finite-dimensional vector space V is *not* V, but we learn early to say that it is canonically isomorphic to V. We can make that precise as follows. Let $f.d.Vec_k$ be the category of finite-dimensional vector spaces over k. Note that this category contains oodles of vector spaces of each dimension.

1.1.C. EXERCISE. Let $(\cdot)^{\vee\vee}\colon f.d.Vec_k \to f.d.Vec_k$ be the double dual functor from the category of finite-dimensional vector spaces over k to itself. Show that $(\cdot)^{\vee\vee}$ is naturally isomorphic to the identity functor on $f.d.Vec_k$. (Without the finite-dimensionality hypothesis, we only get a natural transformation of functors from id to $(\cdot)^{\vee\vee}$.)

Let $\mathscr{V}$ be the category whose objects are the k-vector spaces k^n for each $n \geq 0$ (there is one vector space for each n), and whose morphisms are linear transformations. The objects of $\mathscr{V}$ can be thought of as vector spaces with bases, and the morphisms as matrices. There is an obvious functor $\mathscr{V} \to f.d.Vec_k$, as each k^n *is* a finite-dimensional vector space.

1.1.D. EXERCISE. Show that $\mathscr{V} \to f.d.Vec_k$ gives an equivalence of categories, by describing an "inverse" functor. (Recall that we are being cavalier about set-theoretic assumptions, see Caution 0.3.1, so feel free to simultaneously choose bases for each vector space in $f.d.Vec_k$. To make this precise, you will need to use Gödel-Bernays set theory or else replace $f.d.Vec_k$ with a very similar small category, but we won't worry about this.)

1.1.22. $\star\star$ *Aside for experts.* Your argument for Exercise 1.1.D will show that (modulo set-theoretic issues) this definition of equivalence of categories is the same as another one commonly given: a covariant functor $F\colon \mathscr{A} \to \mathscr{B}$ is an equivalence of categories if it is fully faithful and every object of $\mathscr{B}$ is isomorphic to an object of the form $F(A)$ for some $A \in \mathscr{A}$ (F is **essentially surjective**, a term we will not need).

1.2 Universal properties determine an object up to unique isomorphism

Given some category that we come up with, we often will have ways of producing new objects from old. In good circumstances, such a definition can be made using the notion of a *universal property*. Informally, we wish that there were an object with some property. We first show that if it exists, then it is essentially unique, or more precisely, is unique up to unique isomorphism. Then we go about constructing an example of such an object to show existence.

Explicit constructions are sometimes easier to work with than universal properties, but with a little practice, universal properties are useful in proving things quickly and slickly. Indeed, when learning the subject, people often find explicit constructions more appealing, and use them more often in proofs, but as they become more experienced, they find universal property arguments more elegant and insightful.

1.2.1. Products were defined by a universal property. We have seen one important example of a universal property argument already in §1.0.1: products. You should go back and verify that our discussion there gives a notion of product in any category, and shows that products, *if they exist*, are unique up to unique isomorphism.

1.2.2. Initial, final, and zero objects. Here are some simple but useful concepts that will give you practice with universal property arguments. An object of a category $\mathscr{C}$ is an **initial object** if it has precisely one map to every object. It is a **final object** if it has precisely one map from every object. It is a **zero object** if it is both an initial object and a final object.

1.2.A. EXERCISE. Show that any two initial objects are uniquely isomorphic. Show that any two final objects are uniquely isomorphic.

In other words, *if* an initial object exists, it is unique up to unique isomorphism, and similarly for final objects. This (partially) justifies the phrase "*the* initial object" rather than "*an* initial object", and similarly for "*the* final object" and "*the* zero object". (Convention: we often say "the", not "a", for anything defined up to unique isomorphism.)

1.2.B. EXERCISE. What are the initial and final objects in *Sets*, *Rings*, and *Top* (if they exist)? How about in the two examples of §1.1.9?

1.2.3. Localization of rings and modules. Another important example of a definition by universal property is the notion of *localization* of a ring. We first review a constructive definition, and then reinterpret the notion in terms of universal property. A **multiplicative subset** S of a ring A is a subset closed under multiplication containing 1. We define a ring $S^{-1}A$. The elements of $S^{-1}A$ are of the form a/s where $a \in A$ and $s \in S$, and where $a_1/s_1 = a_2/s_2$ if (and only if) *for some* $s \in S$, $s(s_2a_1 - s_1a_2) = 0$. We define $(a_1/s_1) + (a_2/s_2) = (s_2a_1 + s_1a_2)/(s_1s_2)$, and $(a_1/s_1) \times (a_2/s_2) = (a_1a_2)/(s_1s_2)$. (If you wish, you may check that this equality of fractions really is an equivalence relation and the two binary operations on fractions are well-defined on equivalence classes and make $S^{-1}A$ into a ring.)

We have a canonical ring map

$$A \longrightarrow S^{-1}A \tag{1.2.3.1}$$

given by $a \mapsto a/1$. Note that if $0 \in S$, $S^{-1}A$ is the 0-ring.

There are two particularly important flavors of multiplicative subsets. The first is $\{1, f, f^2, \dots\}$, where $f \in A$. This localization is denoted A_f. (Can you describe an isomorphism $A_f \overset{\sim}{\longleftrightarrow} A[t]/(tf-1)$?) The second is $A \setminus \mathfrak{p}$, where $\mathfrak{p}$ is a prime ideal. This localization $S^{-1}A$ is denoted $A_\mathfrak{p}$. (Notational warning: If $\mathfrak{p}$ is a prime ideal, then $A_\mathfrak{p}$ means you're allowed to divide by elements not in $\mathfrak{p}$. However, if $f \in A$, A_f means you're allowed to divide by f. This can be confusing. For example, if (f) is a prime ideal, then $A_f \neq A_{(f)}$.)

Warning: sometimes localization is first introduced in the special case where A is an integral domain and $0 \notin S$. In that case, $A \hookrightarrow S^{-1}A$, but this isn't always true, as shown by the following exercise. (But we will see that noninjective localizations needn't be pathological, and we can sometimes understand them geometrically, see Exercise 3.2.L.)

1.2.C. EXERCISE. Show that $A \to S^{-1}A$ is injective if and only if S contains no zerodivisors. (A **zerodivisor** of a ring A is an element a such that there is a nonzero element b with $ab = 0$. The other elements of A are called **non-zerodivisors**. For example, an invertible element is never a zerodivisor. Counter-intuitively, 0 is a zerodivisor in every ring but the 0-ring. More generally, if M is an A-module, then $a \in A$ is a **zerodivisor for** M if there is a nonzero $m \in M$ with $am = 0$. The other elements of A are called **non-zerodivisors for** M. Equivalently, and *very* usefully, $a \in A$ is a non-zerodivisor for M if and only if $\times a : M \to M$ is an injection, or equivalently in the language of §1.5, if

$$0 \longrightarrow M \xrightarrow{\times a} M$$

is exact.)

If A is an integral domain and $S = A \setminus \{0\}$, then $S^{-1}A$ is called the **fraction field** of A, which we denote $K(A)$. The previous exercise shows that A is a subring of its fraction field $K(A)$. We now return to the case where A is a general (commutative) ring.

1.2.D. EXERCISE. Verify that $A \to S^{-1}A$ satisfies the following universal property: $S^{-1}A$ is initial among A-algebras B where every element of S is sent to an invertible element in B. (Recall: the data of "an A-algebra B" and "a ring map $A \to B$" are the same.) Translation: any map $A \to B$ where every element of S is sent to an invertible element must factor uniquely through $A \to S^{-1}A$. Another translation: a ring map out of $S^{-1}A$ is the same thing as a ring map from A that sends every element of S to an invertible element. Furthermore, an $S^{-1}A$-module is the same thing as an A-module for which $s \times \cdot : M \to M$ is an A-module isomorphism for all $s \in S$.

In fact, it is cleaner to *define* $A \to S^{-1}A$ by the universal property, and to show that it exists, and to use the universal property to check various properties $S^{-1}A$ has. Let's get some practice with this by *defining* localizations of modules by universal property. Suppose M is an A-module. We define the A-module map

$\phi\colon M \to S^{-1}M$ as being initial among A-module maps $M \to N$ such that elements of S are invertible in N ($s \times \cdot\colon N \to N$ is an isomorphism for all $s \in S$). More precisely, any such map $\alpha\colon M \to N$ factors uniquely through ϕ:

$$\begin{tikzcd} M \arrow[r, "\phi"] \arrow[dr, "\alpha"'] & S^{-1}M \arrow[d, dotted, "\exists!"] \\ & N \end{tikzcd}$$

(Translation: $M \to S^{-1}M$ is universal (initial) among A-module maps from M to modules that are actually $S^{-1}A$-modules. Can you make this precise by defining clearly the objects and morphisms in this category?)

Notice: (i) this determines $\phi\colon M \to S^{-1}M$ up to unique isomorphism (you should think through what this means); (ii) we are defining not only $S^{-1}M$, but also the map ϕ at the same time; and (iii) essentially by definition the A-module structure on $S^{-1}M$ extends to an $S^{-1}A$-module structure.

1.2.E. EXERCISE. Show that $\phi\colon M \to S^{-1}M$ exists, by constructing something satisfying the universal property. Hint: define elements of $S^{-1}M$ to be of the form m/s where $m \in M$ and $s \in S$, and $m_1/s_1 = m_2/s_2$ if and only if for some $s \in S$, $s(s_2m_1 - s_1m_2) = 0$. Define the additive structure by $(m_1/s_1) + (m_2/s_2) = (s_2m_1 + s_1m_2)/(s_1s_2)$, and the $S^{-1}A$-module structure (and hence the A-module structure) is given by $(a_1/s_1) \cdot (m_2/s_2) = (a_1m_2)/(s_1s_2)$.

1.2.F. EXERCISE.
(a) Show that localization commutes with finite products, or equivalently, with finite direct sums. In other words, if $M_1, \ldots, M_n$ are A-modules, describe an isomorphism (of A-modules, and of $S^{-1}A$-modules) $S^{-1}(M_1 \times \cdots \times M_n) \to S^{-1}M_1 \times \cdots \times S^{-1}M_n$.
(b) Show that localization commutes with *arbitrary* direct sums.
(c) Show that "localization does not necessarily commute with infinite products": the obvious map $S^{-1}(\prod_i M_i) \to \prod_i S^{-1}M_i$ induced by the universal property of localization is not always an isomorphism. (Hint: $(1, 1/2, 1/3, 1/4, \ldots) \in \mathbb{Q} \times \mathbb{Q} \times \cdots$.)

1.2.4. *Remark.* Localization does not always commute with Hom, see Example 1.5.10. But Exercise 1.5.H will show that in good situations (if the first argument of Hom is *finitely presented*), localization *does* commute with Hom.

1.2.5. Tensor products. Another important example of a universal property construction is the notion of a **tensor product** of A-modules

$$\begin{aligned} \otimes_A\colon \qquad \mathrm{obj}(\mathit{Mod}_A) \times \mathrm{obj}(\mathit{Mod}_A) &\longrightarrow \mathrm{obj}(\mathit{Mod}_A) \\ (M, N) &\longmapsto M \otimes_A N \end{aligned}$$

The subscript A is often suppressed when it is clear from context. The tensor product is often defined as follows. Suppose you have two A-modules M and N. Then elements of the tensor product $M \otimes_A N$ are finite A-linear combinations of symbols $m \otimes n$ ($m \in M$, $n \in N$), subject to relations $(m_1 + m_2) \otimes n = m_1 \otimes n + m_2 \otimes n$,

$m \otimes (n_1 + n_2) = m \otimes n_1 + m \otimes n_2$, $a(m \otimes n) = (am) \otimes n = m \otimes (an)$ (where $a \in A$, $m_1, m_2 \in M$, $n_1, n_2 \in N$). More formally, $M \otimes_A N$ is the free A-module generated by $M \times N$, quotiented by the submodule generated by $(m_1 + m_2, n) - (m_1, n) - (m_2, n)$, $(m, n_1 + n_2) - (m, n_1) - (m, n_2)$, $a(m, n) - (am, n)$, and $a(m, n) - (m, an)$ for $a \in A$, $m, m_1, m_2 \in M$, $n, n_1, n_2 \in N$. The image of (m, n) in this quotient is $m \otimes n$.

If A is a field k, we recover the tensor product of vector spaces.

1.2.G. EXERCISE (IF YOU HAVEN'T SEEN TENSOR PRODUCTS BEFORE). Show that $\mathbb{Z}/(10) \otimes_{\mathbb{Z}} \mathbb{Z}/(12) \cong \mathbb{Z}/(2)$. (This exercise is intended to give some hands-on practice with tensor products.)

1.2.H. IMPORTANT EXERCISE: RIGHT-EXACTNESS OF $(\cdot) \otimes_A N$. Show that $(\cdot) \otimes_A N$ gives a covariant functor $Mod_A \to Mod_A$. Show that $(\cdot) \otimes_A N$ is a **right-exact functor**, i.e., if

$$M' \to M \to M'' \to 0$$

is an exact sequence of A-modules (which means $f \colon M \to M''$ is surjective, and M' surjects onto the kernel of f; see §1.5), then the induced sequence

$$M' \otimes_A N \to M \otimes_A N \to M'' \otimes_A N \to 0$$

is also exact. This exercise is repeated in Exercise 1.5.G, but you may get a lot out of doing it now. (You will be reminded of the definition of right-exactness in §1.5.6.)

In contrast, you can quickly check that tensor product is not left-exact: tensor the exact sequence of $\mathbb{Z}$-modules

$$0 \longrightarrow \mathbb{Z} \xrightarrow{\times 2} \mathbb{Z} \longrightarrow \mathbb{Z}/(2) \longrightarrow 0$$

with $\mathbb{Z}/(2)$.

The constructive definition of $\otimes$ is a weird definition, and really the "wrong" definition. To motivate a better one: notice that there is a natural A-bilinear map $M \times N \to M \otimes_A N$. (If $M, N, P \in Mod_A$, a map $f \colon M \times N \to P$ is A-**bilinear** if $f(m_1 + m_2, n) = f(m_1, n) + f(m_2, n)$, $f(m, n_1 + n_2) = f(m, n_1) + f(m, n_2)$, and $f(am, n) = f(m, an) = af(m, n)$.) *Any* A-bilinear map $M \times N \to P$ factors through the tensor product uniquely: $M \times N \to M \otimes_A N \to P$. (Think this through!)

We can take this as the *definition* of the tensor product as follows. It is an A-module T along with an A-bilinear map $t \colon M \times N \to T$, such that given any A-bilinear map $t' \colon M \times N \to T'$, there is a unique A-linear map $f \colon T \to T'$ such that $t' = f \circ t$.

$$\begin{tikzcd} M \times N \arrow[rr, "t"] \arrow[dr, "t'"'] & & T \arrow[dl, dotted, "\exists ! f"] \\ & T' & \end{tikzcd}$$

1.2.I. EXERCISE. Show that $(T, t \colon M \times N \to T)$ is unique up to unique isomorphism. Hint: first figure out what "unique up to unique isomorphism" means for such pairs, using a category of pairs (T, t). Then follow the analogous argument for the product.

In short: given M and N, there is an A-bilinear map $t \colon M \times N \to M \otimes_A N$, unique up to unique isomorphism, defined by the following universal property: for any A-bilinear map $t' \colon M \times N \to T'$ there is a unique A-linear map $f \colon M \otimes_A N \to T'$ such that $t' = f \circ t$.

As with all universal property arguments, this argument shows uniqueness *assuming existence*. To show existence, we need an explicit construction.

1.2.J. EXERCISE. Show that the construction of §1.2.5 satisfies the universal property of tensor product.

The three exercises below are useful facts about tensor products with which you should be familiar.

1.2.K. IMPORTANT EXERCISE.
(a) If M is an A-module and $A \to B$ is a morphism of rings, give $B \otimes_A M$ the structure of a B-module (this is part of the exercise). Show that this describes a functor $Mod_A \to Mod_B$.
(b) **(tensor product of rings)** If further $A \to C$ is another morphism of rings, show that $B \otimes_A C$ has a natural structure of a ring. Hint: multiplication will be given by $(b_1 \otimes c_1)(b_2 \otimes c_2) = (b_1 b_2) \otimes (c_1 c_2)$. (Exercise 1.2.U will interpret this construction as a fibered coproduct.)

1.2.L. IMPORTANT EXERCISE. If S is a multiplicative subset of A and M is an A-module, describe a natural isomorphism $(S^{-1}A) \otimes_A M \xrightarrow{\sim} S^{-1}M$ (as $S^{-1}A$-modules *and* as A-modules).

1.2.M. EXERCISE ($\otimes$ COMMUTES WITH $\oplus$). Show that tensor products commute with arbitrary direct sums: if M and $\{N_i\}_{i \in I}$ are all A-modules, describe an isomorphism

$$M \otimes (\oplus_{i \in I} N_i) \xrightarrow{\sim} \oplus_{i \in I} (M \otimes N_i).$$

1.2.6. Essential Example: Fibered products. Suppose we have morphisms $\alpha \colon X \to Z$ and $\beta \colon Y \to Z$ (in *any* category). Then the **fibered product** (or *fibred product*) is an object $X \times_Z Y$ along with morphisms $pr_X \colon X \times_Z Y \to X$ and $pr_Y \colon X \times_Z Y \to Y$, where the two compositions $\alpha \circ pr_X, \beta \circ pr_Y \colon X \times_Z Y \to Z$ agree, such that given any object W with maps to X and Y (whose compositions to Z agree), these maps factor through some unique $W \to X \times_Z Y$:

$$\begin{tikzcd}
W \arrow[dr, dotted, "\exists!"] \arrow[drr, bend left] \arrow[ddr, bend right] & & \\
& X \times_Z Y \arrow[r, "pr_Y"'] \arrow[d, "pr_X"] & Y \arrow[d, "\beta"] \\
& X \arrow[r, "\alpha"] & Z
\end{tikzcd}$$

(Warning: the definition of the fibered product depends on α and β, even though they are omitted from the notation $X \times_Z Y$.)

By the usual universal property argument, if it exists, it is unique up to unique isomorphism. (You should think this through until it is clear to you.) Thus the use of the phrase "the fibered product" (rather than "a fibered product") is reasonable, and we should reasonably be allowed to give it the name $X \times_Z Y$. We know what maps to it are: they are precisely maps to X and maps to Y that agree as maps to Z.

1.2.7. *Definition.* As an example, if $\pi \colon X \to Y$ is a morphism, and the fibered product $X \times_Y X$ exists, then this determines a **diagonal morphism** $\delta_\pi \colon X \to X \times_Y X$. The diagonal morphism will turn out to be a very useful notion.

Depending on your religion, the diagram

$$\begin{tikzcd} X \times_Z Y \arrow[r, "\mathrm{pr}_Y"] \arrow[d, "\mathrm{pr}_X"'] & Y \arrow[d, "\beta"] \\ X \arrow[r, "\alpha"] & Z \end{tikzcd}$$

is called a **fibered/pullback/Cartesian diagram/square** (six possibilities — and even more are possible if you prefer "fibred" to "fibered").

The right way to interpret the notion of fibered product is first to think about what it means in the category of sets.

1.2.N. EXERCISE (FIBERED PRODUCTS OF SETS). Show that in *Sets*,

$$X \times_Z Y = \{(x, y) \in X \times Y \ : \ \alpha(x) = \beta(y)\}.$$

More precisely, show that the right side, equipped with its evident maps to X and Y, satisfies the universal property of the fibered product. (This will help you build intuition for fibered products.)

1.2.O. EXERCISE. If X is a topological space, show that fibered products always exist in the category of open sets of X, by describing what a fibered product is. (Hint: it has a one-word description.)

1.2.P. EXERCISE. If Z is the final object in a category $\mathscr{C}$, and $X, Y \in \mathscr{C}$, show that "$X \times_Z Y = X \times Y$": "the" fibered product over Z is uniquely isomorphic to "the" product. Assume all relevant (fibered) products exist. (This is an exercise about unwinding the definition.)

1.2.Q. USEFUL EXERCISE: TOWERS OF CARTESIAN DIAGRAMS ARE CARTESIAN DIAGRAMS. If the two squares in the following commutative diagram are Cartesian diagrams, show that the "outside rectangle" (involving U, V, Y, and Z) is also a Cartesian diagram.

$$\begin{tikzcd} U \arrow[r] \arrow[d] & V \arrow[d] \\ W \arrow[r] \arrow[d] & X \arrow[d] \\ Y \arrow[r] & Z \end{tikzcd}$$

1.2.R. EXERCISE. Given morphisms $X_1 \to Y$, $X_2 \to Y$, and $Y \to Z$, show that there is a natural morphism $X_1 \times_Y X_2 \to X_1 \times_Z X_2$, assuming that both fibered products exist. (This is trivial once you figure out what it is saying. The point of this exercise is to see why it is trivial.)

1.2.S. IMPORTANT EXERCISE: THE DIAGONAL-BASE-CHANGE DIAGRAM. Suppose we are given morphisms $X_1, X_2 \to Y$ and $Y \to Z$. Show that the following diagram is a Cartesian square.

$$\begin{tikzcd} X_1 \times_Y X_2 \arrow[r] \arrow[d] & X_1 \times_Z X_2 \arrow[d] \\ Y \arrow[r] & Y \times_Z Y \end{tikzcd}$$

Assume all relevant (fibered) products exist. (If this exercise is too hard now, you can try it again at Exercise 1.3.B.) You will appreciate how useful this diagram is when you repeatedly use the diagonal morphism in proofs and constructions.

If you liked this problem, you may enjoy Exercise 11.1.C.

1.2.8. *Coproducts.* Define **coproduct** in a category by reversing all the arrows in the definition of product. Define **fibered coproduct** in a category by reversing all the arrows in the definition of fibered product. Coproduct is denoted by $\coprod$.

1.2.T. EXERCISE. Show that coproduct for *Sets* is disjoint union. This is why we use the notation $\coprod$ for disjoint union.

1.2.U. EXERCISE. Suppose $A \to B$ and $A \to C$ are two ring morphisms, so in particular B and C are A-modules. Recall (Exercise 1.2.K) that $B \otimes_A C$ has a ring structure. Show that there is a natural morphism $B \to B \otimes_A C$ given by $b \mapsto b \otimes 1$. (This is not necessarily an inclusion; see Exercise 1.2.G.) Similarly, there is a natural morphism $C \to B \otimes_A C$. Show that this gives a fibered coproduct on rings, i.e., that

$$\begin{tikzcd} B \otimes_A C & C \arrow[l] \\ B \arrow[u] & A \arrow[l] \arrow[u] \end{tikzcd}$$

satisfies the universal property of fibered coproduct.

1.2.9. Monomorphisms and epimorphisms.

1.2.10. Definition. A morphism $\pi\colon X \to Y$ is a **monomorphism** if any two morphisms $\mu_1\colon Z \to X$ and $\mu_2\colon Z \to X$ such that $\pi \circ \mu_1 = \pi \circ \mu_2$ must satisfy $\mu_1 = \mu_2$. In other words, there is at most one way of filling in the dotted arrow so that the diagram

$$\begin{tikzcd} Z \arrow[d, dotted, "\leq 1"'] \arrow[dr] & \\ X \arrow[r, "\pi"'] & Y \end{tikzcd}$$

commutes — for any object Z, the natural map $\mathrm{Mor}(Z, X) \to \mathrm{Mor}(Z, Y)$ is an injection. Intuitively, it is the categorical version of an injective map, and indeed this notion generalizes the familiar notion of injective maps of sets. (The reason we don't use the word "injective" is that in some contexts, "injective" will have an intuitive meaning which may not agree with "monomorphism". One example: in the category of divisible groups, the map $\mathbb{Q} \to \mathbb{Q}/\mathbb{Z}$ is a monomorphism but not injective. This is also the case with "epimorphism" (to be defined shortly) vs. "surjective".)

1.2.V. EXERCISE. Show that the composition of two monomorphisms is a monomorphism.

1.2.W. EXERCISE. Prove that a morphism $\pi\colon X \to Y$ is a monomorphism if and only if the fibered product $X \times_Y X$ exists, and the induced diagonal morphism $\delta_\pi\colon X \to X \times_Y X$ (Definition 1.2.7) is an isomorphism. We may then take this as the definition of monomorphism. (Monomorphisms aren't central to future discussions, although they will come up again. This exercise is just good practice.)

1.2.X. EASY EXERCISE. We use the notation of Exercise 1.2.R. Show that if $Y \to Z$ is a monomorphism, then the morphism $X_1 \times_Y X_2 \to X_1 \times_Z X_2$ you described in Exercise 1.2.R is an isomorphism. (Hint: for any object V, give a natural bijection between maps from V to the first and maps from V to the second. It is also possible to use the Diagonal-Base-Change diagram, Exercise 1.2.S.)

The notion of an **epimorphism** is "dual" to the definition of monomorphism, where all the arrows are reversed. This concept will not be central for us, although it turns up in the definition of an abelian category. Intuitively, it is the categorical version of a surjective map. (But be careful when working with categories of objects that are sets with additional structure, as epimorphisms need not be surjective. Example: in the category *Rings*, $\mathbb{Z} \to \mathbb{Q}$ is an epimorphism, but obviously not surjective.)

1.2.11. Representable functors and Yoneda's Lemma. Much of our discussion about universal properties can be cleanly expressed in terms of representable functors, under the rubric of "Yoneda's Lemma". Yoneda's lemma is an easy fact stated in a complicated way. Informally speaking, you can essentially recover an object in a category by knowing the maps into it. For example, we have seen that the data of maps to $X \times Y$ are naturally (canonically) the data of maps to X and to Y. Indeed, we have now taken this as the *definition* of $X \times Y$.

Recall Example 1.1.20. Suppose A is an object of category $\mathscr{C}$. For any object $C \in \mathscr{C}$, we have a set of morphisms $\operatorname{Mor}(C, A)$. If we have a morphism $f\colon B \to C$, we get a map of sets

$$\operatorname{Mor}(C, A) \longrightarrow \operatorname{Mor}(B, A), \tag{1.2.11.1}$$

by composition: given a map from C to A, we get a map from B to A by precomposing with $f\colon B \to C$. Hence this gives a contravariant functor $h_A\colon \mathscr{C} \to$ *Sets*. Yoneda's Lemma states that the functor h_A determines A up to unique isomorphism. More precisely:

1.2.Y. IMPORTANT EXERCISE THAT YOU SHOULD DO ONCE IN YOUR LIFE (YONEDA'S LEMMA).
(a) Suppose you have two objects A and A' in a category $\mathscr{C}$, and maps

$$i_C\colon \operatorname{Mor}(C, A) \longrightarrow \operatorname{Mor}(C, A') \tag{1.2.11.2}$$

that commute with the maps (1.2.11.1). Show that the i_C (as C ranges over the objects of $\mathscr{C}$) are induced from a unique morphism $g\colon A \to A'$. More precisely,

show that there is a unique morphism $g\colon A \to A'$ such that for all $C \in \mathscr{C}$, i_C is $u \mapsto g \circ u$.
(b) If furthermore the i_C are all bijections, show that the resulting g is an isomorphism. (Hint for both: This is much easier than it looks. This statement is so general that there are really only a couple of things that you could possibly try. For example, if you're hoping to find a morphism $A \to A'$, where will you find it? Well, you are looking for an element $\mathrm{Mor}(A, A')$. So just plug in $C = A$ to (1.2.11.2), and see where the identity goes.)

There is an analogous statement with the arrows reversed, where instead of maps into A, you think of maps *from* A. The role of the contravariant functor h_A of Example 1.1.20 is played by the covariant functor h^A of Example 1.1.14. Because the proof is the same (with the arrows reversed), you needn't think it through.

The phrase "Yoneda's Lemma" properly refers to a more general statement. Although it looks more complicated, it is no harder to prove.

1.2.Z. ⋆ EXERCISE.
(a) Suppose A and B are objects in a category $\mathscr{C}$. Give a bijection between the natural transformations $h^A \to h^B$ of covariant functors $\mathscr{C} \to$ *Sets* (see Example 1.1.14 for the definition) and the morphisms $B \to A$.
(b) State and prove the corresponding fact for contravariant functors h_A (see Example 1.1.20). Remark: A contravariant functor F from $\mathscr{C}$ to *Sets* is said to be **representable** if there is a natural isomorphism

$$\xi\colon F \xrightarrow{\sim} h_A .$$

Thus the representing object A is determined up to unique isomorphism by the pair (F, ξ). There is a similar definition for covariant functors. (We will revisit this in §7.6, and this problem will appear again as Exercise 7.6.C. The element $\xi^{-1}(\mathrm{id}_A) \in F(A)$ is often called the "universal object"; do you see why?)
(c) **Yoneda's Lemma.** Suppose F is a covariant functor $\mathscr{C} \to$ *Sets*, and $A \in \mathscr{C}$. Give a bijection between the natural transformations $h^A \to F$ and $F(A)$. (The corresponding fact for contravariant functors is essentially Exercise 10.1.B.)

In fancy terms, Yoneda's lemma states the following. Given a category $\mathscr{C}$, we can produce a new category, called the **functor category** of $\mathscr{C}$, where the objects are contravariant functors $\mathscr{C} \to$ *Sets*, and the morphisms are natural transformations of such functors. We have a functor (which we can usefully call h) from $\mathscr{C}$ to its functor category, which sends A to h_A. Yoneda's Lemma states that this is a fully faithful functor, called the *Yoneda embedding*. (Fully faithful functors were defined in §1.1.15.)

1.2.12. *Joke.* The Yoda embedding, contravariant it is.

1.3 Limits and colimits

Limits and colimits are two important definitions determined by universal properties. They generalize a number of familiar constructions. I will give the definition first, and then show you why it is familiar. For example, fractions will be motivating examples of colimits (Exercise 1.3.D(a)), and the p-adic integers (Example 1.3.4) will be motivating examples of limits.

1.3.1. Limits. We say that a category is a **small category** if the objects form a set and the morphisms form a set. (This is a technical condition intended only for experts.) Suppose $\mathscr{I}$ is any small category, and $\mathscr{C}$ is any category. Then a functor $F\colon \mathscr{I} \to \mathscr{C}$ (i.e., with an object $A_i \in \mathscr{C}$ for each element $i \in \mathscr{I}$, and appropriate commuting morphisms dictated by $\mathscr{I}$) is said to be a **diagram indexed by** $\mathscr{I}$. We call $\mathscr{I}$ an **index category**. Our index categories will usually be partially ordered sets (Example 1.1.8), in which in particular there is at most one morphism between any two objects. (But other examples are sometimes useful.) For example, if $\square$ is the category

$$\begin{tikzcd} \bullet \arrow[r] \arrow[d] & \bullet \arrow[d] \\ \bullet \arrow[r] & \bullet \end{tikzcd}$$

and $\mathscr{A}$ is a category, then a functor $\square \to \mathscr{A}$ is precisely the data of a commuting square in $\mathscr{A}$.

Then the **limit of the diagram** is an object $\lim_{\mathscr{I}} A_i$ (or $\varprojlim_{\mathscr{I}} A_i$) of $\mathscr{C}$ along with morphisms $f_j\colon \lim_{\mathscr{I}} A_i \to A_j$ for each $j \in \mathscr{I}$, such that if $m\colon j \to k$ is a morphism in $\mathscr{I}$, then

$$\begin{tikzcd} \lim_{\mathscr{I}} A_i \arrow[d, "f_j"'] \arrow[dr, "f_k"] & \\ A_j \arrow[r, "F(m)"] & A_k \end{tikzcd} \tag{1.3.1.1}$$

commutes, and this object and maps to each A_i are universal (final) with respect to this property. More precisely, given any other object W along with maps $g_i\colon W \to A_i$ commuting with the $F(m)$ (if $m\colon j \to k$ is a morphism in $\mathscr{I}$, then $g_k = F(m) \circ g_j$), then there is a unique map

$$g\colon W \to \lim_{\mathscr{I}} A_i$$

so that $g_i = f_i \circ g$ for all i. (In some cases, the limit is sometimes called the **inverse limit** or **projective limit**. We won't use this language.) By the usual universal property argument, if the limit exists, it is unique up to unique isomorphism.

1.3.2. *Examples: products.* For example, if $\mathscr{I}$ is the partially ordered set

$$\begin{tikzcd} & \bullet \arrow[d] \\ \bullet \arrow[r] & \bullet \end{tikzcd}$$

we obtain the fibered product.

If $\mathscr{I}$ is

$$\bullet \qquad \bullet$$

we obtain the product.

If $\mathscr{I}$ is a set (i.e., the only morphisms are the identity maps), then the limit is called the **product** of the A_i, and is denoted $\prod_i A_i$. The special case where $\mathscr{I}$ has two elements is the example of the previous paragraph.

1.3.A. EXERCISE (REALITY CHECK). Suppose that the partially ordered set $\mathscr{I}$ has an initial object e. Show that the limit of any diagram indexed by $\mathscr{I}$ exists.

1.3.B. EXERCISE: THE DIAGONAL-BASE-CHANGE DIAGRAM, AGAIN. Solve 1.2.S again by identifying both $X_1 \times_Y X_2$ and $Y \times_{(Y \times_Z Y)} (X_1 \times_Z X_2)$ as the limit of the diagram

$$\begin{tikzcd}
X_1 \arrow[dr] & & \\
& Y \arrow[r] & Z \\
X_2 \arrow[ur] & &
\end{tikzcd}$$

1.3.3. *Example: formal power series.* For a ring A, the **(formal) power series**, $A[[x]]$, are often described informally (and somewhat unnaturally) as being the ring

$$A[[x]] = \{a_0 + a_1 x + a_2 x^2 + \cdots\}$$

(where $a_i \in A$, and the ring operations are the "obvious" ones). It is an example of a limit in the category of rings:

$$\begin{tikzcd}
& A[[x]] \arrow[d] \arrow[dr] \arrow[drr] & & \\
\cdots \arrow[r] & A[x]/(x^3) \arrow[r] & A[x]/(x^2) \arrow[r] & A[x]/(x).
\end{tikzcd}$$

The universal property of limits yields a natural ring morphism $A[x] \to A[[x]]$. If $A = \mathbb{R}$ or $\mathbb{C}$, this map factors through the ring of *convergent power series*.

1.3.4. *Example: the* p*-adic integers.* For a prime number p, the **p-adic integers** (or more informally, p**-adics**), $\mathbb{Z}_p$, are often described informally (and somewhat unnaturally) as being of the form

$$a_0 + a_1 p + a_2 p^2 + \cdots$$

(where $0 \leq a_i < p$). They are an example of a limit in the category of rings:

$$\begin{tikzcd}
& \mathbb{Z}_p \arrow[d] \arrow[dr] \arrow[drr] & & \\
\cdots \arrow[r] & \mathbb{Z}/(p^3) \arrow[r] & \mathbb{Z}/(p^2) \arrow[r] & \mathbb{Z}/(p).
\end{tikzcd}$$

(Warning: $\mathbb{Z}_p$ is sometimes used to denote the integers modulo p, but $\mathbb{Z}/(p)$ or $\mathbb{Z}/p\mathbb{Z}$ is better to use for this, to avoid confusion. Worse: by §1.2.3, $\mathbb{Z}_p$ also denotes those rationals whose denominators are a power of p. Hopefully the meaning of $\mathbb{Z}_p$ will be clear from the context.)

The similarity of Examples 1.3.3 and 1.3.4 is no coincidence. Formal power series and the p-adic integers are examples of *completions*, the topic of Chapter 28.

Limits do not always exist for any index category $\mathscr{I}$. However, you can often easily check that limits exist if the objects of your category can be interpreted as sets with additional structure, and arbitrary products exist (respecting the set-like structure).

1.3.C. IMPORTANT EXERCISE. Show that in the category *Sets*,

$$\left\{ (a_i)_{i \in \mathscr{I}} \in \prod_i A_i : F(m)(a_j) = a_k \text{ for all } m \in \operatorname{Mor}_{\mathscr{I}}(j,k) \subset \operatorname{Mor}(\mathscr{I}) \right\},$$

along with the obvious projection maps to each A_i, is the limit $\varprojlim_{\mathscr{I}} A_i$.

This clearly also works in the category Mod_A of A-modules (in particular Vec_k and *Ab*), as well as *Rings*.

From this point of view, $2 + 3p + 2p^2 + \cdots \in \mathbb{Z}_p$ can be understood as the sequence $(2, 2+3p, 2+3p+2p^2, \dots)$.

1.3.5. Colimits. More immediately relevant for us will be the dual (arrow-reversed version) of the notion of limit (or inverse limit). We just flip the arrows f_i in (1.3.1.1), and get the notion of a **colimit**, which is denoted $\operatorname{colim}_{\mathscr{I}} A_i$ (or $\varinjlim_{\mathscr{I}} A_i$). (You should draw the corresponding diagram.) Again, if it exists, it is unique up to unique isomorphism. (In some cases, the colimit is sometimes called the **direct limit**, **inductive limit**, or **injective limit**. We won't use this language. I prefer using limit/colimit in analogy with kernel/cokernel and product/coproduct. This is more than analogy, as kernels and products may be interpreted as limits, and similarly with cokernels and coproducts. Also, I remember that kernels "map to", and cokernels are "mapped to", which reminds me that a limit maps *to* all the objects in the big commutative diagram indexed by $\mathscr{I}$; and a colimit has a map *from* all the objects.)

1.3.6. *Joke.* A comathematician is a device for turning cotheorems into ffee.

Even though we have just flipped the arrows, colimits behave quite differently from limits.

1.3.7. *Example.* The abelian group $5^{-\infty}\mathbb{Z}$ of rational numbers whose denominators are powers of 5 is a colimit $\operatorname{colim}_{i \in \mathbb{Z}^+} 5^{-i}\mathbb{Z}$. More precisely, $5^{-\infty}\mathbb{Z}$ is the colimit of the diagram

$$\mathbb{Z} \longrightarrow 5^{-1}\mathbb{Z} \longrightarrow 5^{-2}\mathbb{Z} \longrightarrow \cdots$$

in the category of abelian groups.

The colimit over an index *set* I is called the **coproduct**, denoted $\coprod_i A_i$, and is the dual (arrow-reversed) notion to the product.

1.3.D. EXERCISE.
(a) Interpret the statement "$\mathbb{Q} = \operatorname{colim} \frac{1}{n}\mathbb{Z}$".
(b) Interpret the union of some subsets of a given set as a colimit. (Dually, the intersection can be interpreted as a limit.) The objects of the category in question are the subsets of the given set.

Colimits do not always exist, but there are two useful large classes of examples for which they do.

1.3.8. Definition. A nonempty partially ordered set $(S, \geq)$ is **filtered** (or is said to be a **filtered set**) if for each $x, y \in S$, there is a z such that $x \geq z$ and $y \geq z$. More generally (see Figure 1.1, a nonempty category $\mathscr{I}$ is **filtered** if:

(i) for each $x, y \in \mathscr{I}$, there is a $z \in \mathscr{I}$ and arrows $x \to z$ and $y \to z$, and
(ii) for every two arrows $u \colon x \to y$ and $v \colon x \to y$, there is an arrow $w \colon y \to z$ such that $w \circ u = w \circ v$.

(Other terminologies are also commonly used, such as "directed partially ordered set" and "filtered index category", respectively.)

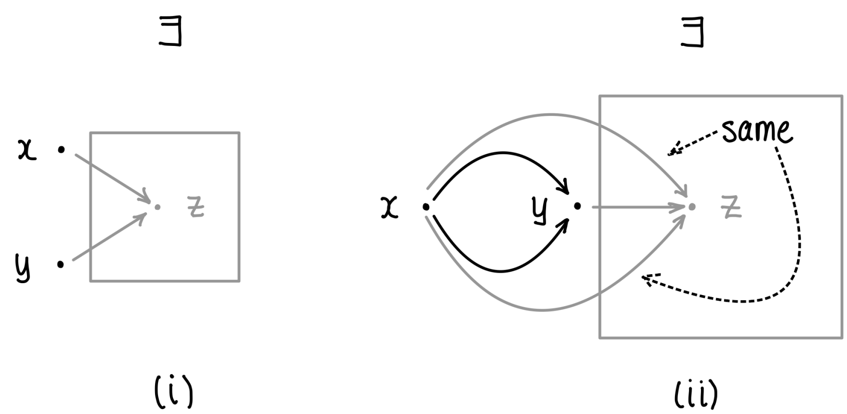


FIGURE 1.1. A filtered category (pictorial definition)

1.3.E. EXERCISE. Suppose $\mathscr{I}$ is filtered. (We will almost exclusively use the case where $\mathscr{I}$ is a filtered set.) Recall the symbol $\coprod$ for disjoint union of sets. Show that any diagram in *Sets* indexed by $\mathscr{I}$ has the following, with the obvious maps to it, as a colimit:

$$\left\{ (a_i, i) \in \coprod_{i \in \mathscr{I}} A_i \right\} \Big/ \left(\begin{array}{c} (a_i, i) \sim (a_j, j) \text{ if and only if there are } f \colon A_i \to A_k \text{ and} \\ g \colon A_j \to A_k \text{ in the diagram for which } f(a_i) = g(a_j) \text{ in } A_k \end{array} \right)$$

(You will see that the "filtered" hypothesis is there is to ensure that $\sim$ is an equivalence relation.)

For example, in Example 1.3.7, each element of the colimit is an element of something upstairs, but you can't say in advance what it is an element of. For instance, $17/125$ is an element of the $5^{-3}\mathbb{Z}$ (or $5^{-4}\mathbb{Z}$, or later ones), but not $5^{-2}\mathbb{Z}$.

This idea applies to many categories whose objects can be interpreted as sets with additional structure (such as abelian groups, A-modules, groups, etc.). For example, the colimit $\operatorname{colim} M_i$ in the category of A-modules Mod_A can be described as follows. The set underlying $\operatorname{colim} M_i$ is defined as in Exercise 1.3.E. To add the elements $m_i \in M_i$ and $m_j \in M_j$, choose an $\ell \in \mathscr{I}$ with arrows $u \colon i \to \ell$ and $v \colon j \to \ell$, and then define the sum of m_i and m_j to be $F(u)(m_i) + F(v)(m_j) \in M_\ell$.

The element $m_i \in M_i$ is 0 if and only if there is some arrow $u\colon i \to k$ for which $F(u)(m_i) = 0$, i.e., if it becomes 0 "later in the diagram". Last, multiplication by an element of A is defined in the obvious way.

1.3.F. EXERCISE. Verify that the A-module described above is indeed the colimit. (Make sure you verify that addition is well-defined, i.e., is independent of the choice of representatives m_i and m_j, the choice of ℓ, and the choice of arrows u and v. Similarly, make sure that scalar multiplication is well-defined.)

1.3.G. USEFUL EXERCISE (LOCALIZATION AS A COLIMIT). Generalize Exercise 1.3.D(a) to interpret localization of an integral domain as a colimit over a filtered set: suppose S is a multiplicative set of A, and interpret $S^{-1}A = \operatorname{colim} \frac{1}{s}A$ where the colimit is over $s \in S$, and in the category of A-modules. (Aside: Can you make some version of this work even if A isn't an integral domain, e.g., $S^{-1}A = \operatorname{colim} A_s$? This will work in the category of A-algebras.)

A variant of this construction works without the filtered condition, if you have another means of "connecting elements in different objects of your diagram". For example:

1.3.H. EXERCISE: COLIMITS OF A-MODULES WITHOUT THE FILTERED CONDITION. Suppose you are given a diagram of A-modules indexed by $\mathscr{I}$: $F\colon \mathscr{I} \to \mathit{Mod}_A$, where we let $M_i := F(i)$. Show that the colimit is $\oplus_{i \in \mathscr{I}} M_i$ modulo the relations $m_i - F(n)(m_i)$ for every $n\colon i \to j$ in $\mathscr{I}$ (i.e., for every arrow in the diagram). (Somewhat more precisely: "modulo" means "quotiented by the submodule generated by".)

1.3.9. Summary. One useful thing to informally keep in mind is the following. In a category where the objects are "set-like", an element of a limit can be thought of as a family of elements of each object in the diagram, that are "compatible" (Exercise 1.3.C). And an element of a colimit can be thought of as ("has a representative that is") an element of a single object in the diagram (Exercise 1.3.E). Even though the definitions of limit and colimit are the same, just with arrows reversed, these interpretations are quite different.

1.3.10. *Small remark.* In fact, colimits exist in the category of sets for all reasonable ("small") index categories (see for example [**E**, Thm. A6.1]), but that won't matter to us.

1.3.11. *Joke.* What do you call someone who reads a paper on category theory? Answer: A coauthor.

1.4 Adjoints

We next come to a very useful notion closely related to universal properties. Just as a universal property "essentially" (up to unique isomorphism) determines an object in a category (assuming such an object exists), "adjoints" essentially determine a functor (again, assuming it exists). Two *covariant* functors $F\colon \mathscr{A} \to \mathscr{B}$

and $G\colon \mathscr{B} \to \mathscr{A}$ are **adjoint** if there is a natural bijection for all $A \in \mathscr{A}$ and $B \in \mathscr{B}$

$$\tau_{AB}\colon \operatorname{Mor}_{\mathscr{B}}(F(A), B) \to \operatorname{Mor}_{\mathscr{A}}(A, G(B)). \tag{1.4.0.1}$$

We say that (F, G) form an **adjoint pair**, and that F is **left-adjoint** to G (and G is **right-adjoint** to F). We say F is a **left adjoint** (and G is a **right adjoint**). By "natural" we mean the following. For all $f\colon A \to A'$ in $\mathscr{A}$, we require

$$\begin{tikzcd} \operatorname{Mor}_{\mathscr{B}}(F(A'), B) \arrow[r, "Ff^*"] \arrow[d, "\tau_{A'B}"] & \operatorname{Mor}_{\mathscr{B}}(F(A), B) \arrow[d, "\tau_{AB}"] \\ \operatorname{Mor}_{\mathscr{A}}(A', G(B)) \arrow[r, "f^*"] & \operatorname{Mor}_{\mathscr{A}}(A, G(B)) \end{tikzcd} \tag{1.4.0.2}$$

to commute, and for all $g\colon B \to B'$ in $\mathscr{B}$ we want a similar commutative diagram to commute. (Here f^* is the map induced by $f\colon A \to A'$, and Ff^* is the map induced by $Ff\colon F(A) \to F(A')$.)

1.4.A. EXERCISE. Write down what this diagram should be.

1.4.B. EXERCISE. Show that the map τ_{AB} (1.4.0.1) has the following properties. For each A there is a map $\eta_A\colon A \to GF(A)$ so that for any $g\colon F(A) \to B$, the corresponding $\tau_{AB}(g)\colon A \to G(B)$ is given by the composition

$$A \xrightarrow{\eta_A} GF(A) \xrightarrow{Gg} G(B).$$

Similarly, there is a map $\epsilon_B\colon FG(B) \to B$ for each B so that for any $f\colon A \to G(B)$, the corresponding map $\tau^{-1}_{AB}(f)\colon F(A) \to B$ is given by the composition

$$F(A) \xrightarrow{Ff} FG(B) \xrightarrow{\epsilon_B} B.$$

Here is a key example of an adjoint pair.

1.4.C. EXERCISE. Suppose M, N, and P are A-modules (where A is a ring). Describe a bijection $\operatorname{Hom}_A(M \otimes_A N, P) \leftrightarrow \operatorname{Hom}_A(M, \operatorname{Hom}_A(N, P))$. (Hint: try to use the universal property of $\otimes$.)

1.4.D. EXERCISE (TENSOR-HOM ADJUNCTION). Show that $(\cdot)\otimes_A N$ and $\operatorname{Hom}_A(N, \cdot)$ are adjoint functors.

1.4.E. EXERCISE. Suppose $B \to A$ is a morphism of rings. If M is an A-module, you can create a B-module M_B by considering it as a B-module. This gives a functor $\cdot_B\colon \mathit{Mod}_A \to \mathit{Mod}_B$. Show that this functor is right-adjoint to $\cdot \otimes_B A$. In other words, describe a bijection

$$\operatorname{Hom}_A(N \otimes_B A, M) \cong \operatorname{Hom}_B(N, M_B)$$

functorial in both arguments. (This adjoint pair is very important.)

1.4.1. ⋆ *Fancier remarks we won't use.* You can check that the left adjoint determines the right adjoint up to natural isomorphism, and vice versa. The maps η_A and ϵ_B of Exercise 1.4.B are called the *unit* and *counit* of the adjunction. This leads to a different characterization of adjunction. Suppose functors $F\colon \mathscr{A} \to \mathscr{B}$ and $G\colon \mathscr{B} \to \mathscr{A}$ are given, along with natural transformations $\eta\colon \mathrm{id}_{\mathscr{A}} \to GF$ and $\epsilon\colon FG \to \mathrm{id}_{\mathscr{B}}$ with the property that $G\epsilon \circ \eta G = \mathrm{id}_G$ (for each $B \in \mathscr{B}$, the composition of $\eta_{G(B)}\colon G(B) \to$

$GFG(B)$ and $G(\epsilon_B)\colon GFG(B) \to G(B)$ is the identity) and $\epsilon F \circ F\eta = \mathrm{id}_F$. Then you can check that F is left-adjoint to G. These facts aren't hard to check, so if you want to use them, you should verify everything for yourself.

1.4.2. *Examples from other fields.* For those familiar with representation theory: Frobenius reciprocity may be understood in terms of adjoints. Suppose V is a finite-dimensional representation of a finite group G, and W is a representation of a subgroup $H < G$. Then induction and restriction are an adjoint pair $(\mathrm{Ind}_H^G, \mathrm{Res}_H^G)$ between the category of G-modules and the category of H-modules.

Topologists' favorite adjoint pair may be the suspension functor and the loop space functor.

1.4.3. Example: groupification of abelian semigroups. Here is another motivating example: getting an abelian group from an abelian semigroup. (An **abelian semigroup** is just like an abelian group, except we don't require an identity or an inverse. Morphisms of abelian semigroups are maps of sets preserving the binary operation. One example is the non-negative integers $\mathbb{Z}^{\geq 0} = \{0, 1, 2, \dots\}$ under addition. Another is the positive integers $\mathbb{Z}^+ = \{1, 2, 3, \dots\}$ under addition. Yet another is the positive integers $\mathbb{Z}^+$ under multiplication. You may enjoy groupifying all three.) From an abelian semigroup, you can create an abelian group. In our examples, from the nonnegative integers under addition $(\mathbb{Z}^{\geq 0}, +)$, we create the integers $(\mathbb{Z}, +)$, and from the positive integers under multiplication $(\mathbb{Z}^+, \times)$, we create the positive rationals $(\mathbb{Q}^+, \times)$. Here is a formalization of that notion. A **groupification** of an abelian semigroup S is a map of abelian semigroups $\pi\colon S \to G$ such that G is an abelian group, and any map of abelian semigroups from S to an abelian *group* G' factors *uniquely* through G:

$$\begin{tikzcd} S \arrow[r, "\pi"] \arrow[dr] & G \arrow[d, dotted, "\exists!"] \\ & G' \end{tikzcd}$$

(Perhaps "abelian groupification" would be more precise than "groupification".)

1.4.F. Exercise (an abelian group is groupified by itself). Show that if an abelian semigroup is *already* a group then the identity morphism is the groupification. (More correct: the identity morphism is *a* groupification.) Note that you don't need to construct groupification (or even know that it exists in general) to solve this exercise.

1.4.G. Exercise. Construct the "groupification functor" H from the category of nonempty abelian semigroups to the category of abelian groups. (One possible construction: given an abelian semigroup S, the elements of its groupification $H(S)$ are ordered pairs $(a, b) \in S \times S$, which you may think of as $a - b$, with the equivalence that $(a, b) \sim (c, d)$ if $a + d + e = b + c + e$ for some $e \in S$. Describe addition in this group, and show that it satisfies the properties of an abelian group. Describe the abelian semigroup map $S \to H(S)$.) Let F be the forgetful functor from the category of abelian groups *Ab* to the category of abelian semigroups. Show that H is left-adjoint to F.

(Here is the general idea for experts: We have a full subcategory of a category. We want to "project" from the category to the subcategory. We have

$$\mathrm{Mor}_{\text{category}}(S, H) = \mathrm{Mor}_{\text{subcategory}}(G, H)$$

automatically; thus we are describing the left adjoint to the forgetful functor. How the argument worked: we constructed something which was in the smaller category, which automatically satisfies the universal property.)

1.4.H. EXERCISE (CF. EXERCISE 1.4.E). The purpose of this exercise is to give you more practice with "left adjoints of forgetful functors", the means by which we get abelian groups from abelian semigroups, and sheaves from presheaves. Suppose A is a ring, and S is a multiplicative subset. Then $S^{-1}A$-modules are a full subcategory (§1.1.15) of the category of A-modules (via the obvious inclusion $Mod_{S^{-1}A} \hookrightarrow Mod_A$). Then $Mod_A \to Mod_{S^{-1}A}$ can be interpreted as an adjoint to the forgetful functor $Mod_{S^{-1}A} \to Mod_A$. State and prove the correct statements.

(Here is the larger story. Every $S^{-1}A$-module is an A-module, and this is an injective map, so we have a forgetful functor $F\colon Mod_{S^{-1}A} \to Mod_A$. In fact this is a fully faithful functor: it is injective on objects, and the morphisms between any two $S^{-1}A$-modules *as A-modules* are just the same when they are considered as $S^{-1}A$-modules. Then there is a functor $G\colon Mod_A \to Mod_{S^{-1}A}$, which might reasonably be called "localization with respect to S", which is left-adjoint to the forgetful functor. Translation: If M is an A-module, and N is an $S^{-1}A$-module, then $\mathrm{Mor}(GM, N)$ (morphisms as $S^{-1}A$-modules, which are the same as morphisms as A-modules) are in natural bijection with $\mathrm{Mor}(M, FN)$ (morphisms as A-modules).)

Table 1 gives most of the adjoints that will come up for us. Other examples will also come up, such as the adjoint pair $(\sim, \Gamma_\bullet)$ between graded modules over a graded ring, and quasicoherent sheaves on the corresponding projective scheme (§15.7).

1.4.4. *Last comments only for people who have seen adjoints before.* If (F, G) is an adjoint pair of functors, then F commutes with colimits, and G commutes with limits. Also, limits commute with limits and colimits commute with colimits. We will prove these facts (and a little more) in §1.5.14.

1.5 An introduction to abelian categories

Ton papier sur l'Algèbre homologique a été lu soigneusement, et a converti tout le monde (même Dieudonné, qui semble complètement fonctorisé!) à ton point de vue.

Your paper on homological algebra was read carefully and converted everyone (even Dieudonné, who seems to be completely functorised!) to your point of view.

— *J.-P. Serre, letter to A. Grothendieck, Jul 13, 1955* [**GrS**, p. 17-18]

Since learning linear algebra, you have been familiar with the notions and behaviors of kernels, cokernels, etc. Later in your life you saw them in the category of abelian groups, and later still in the category of A-modules.

We will soon define some new categories (certain sheaves) that will have familiar-looking behavior, reminiscent of that of modules over a ring. The notions

situation	category $\mathscr{A}$	category $\mathscr{B}$	left adjoint $F\colon \mathscr{A} \to \mathscr{B}$	right adjoint $G\colon \mathscr{B} \to \mathscr{A}$
A-modules (Ex. 1.4.D)	Mod_A	Mod_A	$(\cdot) \otimes_A N$	$\operatorname{Hom}_A(N, \cdot)$
ring maps $B \to A$ (Ex. 1.4.E)	Mod_B	Mod_A	$(\cdot) \otimes_B A$ (extension of scalars)	$M \mapsto M_B$ (restriction of scalars)
(pre)sheaves on a topological space X (Ex. 2.4.K)	presheaves on X	sheaves on X	sheafification	forgetful
(semi)groups (§1.4.3)	semigroups	groups	groupification	forgetful
sheaves, $\pi\colon X \to Y$ (Ex. 2.7.B)	sheaves on Y	sheaves on X	π^{-1}	π_*
sheaves of abelian groups or $\mathscr{O}$-modules, open embeddings $\pi\colon U \hookrightarrow Y$ (Ex. 23.4.G)	sheaves on U	sheaves on Y	$\pi_!$	π^{-1}
quasicoherent sheaves, $\pi\colon X \to Y$ (Prop. 14.5.7)	$QCoh_Y$	$QCoh_X$	π^*	π_*
ring maps $B \to A$ (Ex. 17.1.J)	Mod_A	Mod_B	$M \mapsto M_B$ (restriction of scalars)	$N \mapsto \operatorname{Hom}_B(A, N)$
quasicoherent sheaves, affine $\pi\colon X \to Y$ (Ex. 17.1.K(b))	$QCoh_X$	$QCoh_Y$	π_*	$\pi^{!?}$

TABLE 1. Some important adjoint pairs

of kernels, cokernels, images, and more will make sense, and they will behave "the way we expect" from our experience with modules. This can be made precise through the notion of an *abelian category*. Abelian categories are the right general setting in which one can do "homological algebra", in which notions of kernel, cokernel, and so on are used, and one can work with complexes and exact sequences.

We will see enough to motivate the definitions that we will see in general: monomorphism (and subobject), epimorphism, kernel, cokernel, and image. But in this book we will avoid having to show that they behave "the way we expect" in a general abelian category because the examples we will see are directly interpretable in terms of modules over rings. In particular, it is not worth memorizing the definition of abelian category.

Two central examples of an abelian category are the category *Ab* of abelian groups, and the category Mod_A of A-modules. The first is a special case of the second (just take $A = \mathbb{Z}$). As we give the definitions, you should verify that Mod_A is an abelian category.

We first define the notion of *additive category*. We will use it only as a stepping stone to the notion of an abelian category. Two examples you can keep in mind while reading the definition: the category of free A-modules (where A is a ring), and real (or complex) Banach spaces.

1.5.1. Definition. A category $\mathscr{C}$ is said to be **additive** if it satisfies the following properties.

Ad1. For each $A, B \in \mathscr{C}$, $\mathrm{Mor}(A, B)$ is an abelian group, such that composition of morphisms distributes over addition. (You should think about what this means — it translates to two distinct statements.)

Ad2. $\mathscr{C}$ has a zero object, denoted 0. (This is an object that is simultaneously an initial object and a final object, Definition 1.2.2.)

Ad3. It has products of two objects (a product $A \times B$ for any pair of objects), and hence by induction, products of any finite number of objects.

In an additive category, the morphisms are often called **homomorphisms**, and Mor is denoted by Hom. In fact, this notation Hom is a good indication that you're working in an additive category. A functor between additive categories preserving the additive structure of Hom, is called an **additive functor**.

1.5.2. *Remarks.* It is a consequence of the definition of additive category that finite products are also finite coproducts (i.e., sums) — the details don't matter to us. The symbol $\oplus$ is used for this notion. Also, it is quick to show that additive functors send zero objects to zero objects (show that Z is a 0-object if and only if $\mathrm{id}_Z = 0_Z$; additive functors preserve both id and 0), and preserve products.

One motivation for the name 0-object is that the 0-morphism in the abelian group $\mathrm{Hom}(A, B)$ is the composition $A \to 0 \to B$. (We also remark that the notion of 0-morphism thus makes sense in any category with a 0-object.)

(A cleaner axiomatization of additive categories that makes clear that the abelian group structure of $\mathrm{Mor}(A, B)$ is intrinsic to the category itself is the following, [**Lur**, p. 21–22]. *A0. $\mathscr{C}$ has a zero object. A1. $\mathscr{C}$ has products of any two objects, and coproducts of any two objects.* By the universal property of product and coproduct, we have natural morphisms $\phi_{AB} : A \coprod B \to A \times B$. *A2.* ϕ_{AB} *is an isomorphism.* This allows us to define a binary operation on $\mathrm{Mor}(A, B)$, with $f + g$ (for $f, g \in \mathrm{Mor}(A, B)$) defined by the composition

$$A \xrightarrow{(f,g)} B \times B \xrightarrow{\phi_{BB}^{-1}} B \coprod B \longrightarrow B$$

where the last map is the "codiagonal" defined by universal property of coproduct. A little work shows that this endows $\mathrm{Mor}(A, B)$ with the structure of a **commutative monoid**, i.e., an abelian semigroup with identity. The identity is the composition $A \to 0 \to B$. *A3. This commutative monoid* $\mathrm{Mor}(A, B)$ *is an abelian group.*)

1.5.3. The category of A-modules Mod_A is clearly an additive category, but it has even more structure, which we now formalize as an example of an abelian category.

1.5.4. Definition. Let $\mathscr{C}$ be a category with a 0-object (and thus 0-morphisms). A **kernel** of a morphism $f \colon B \to C$ is *defined to be* a map $i \colon A \to B$ such that $f \circ i = 0$, and that is universal with respect to this property. Diagramatically:

$$\begin{tikzcd}
Z \arrow[d, dotted, "\exists!"'] \arrow[dr] \arrow[drr, "0"] & & \\
A \arrow[r, "i"] \arrow[rr, bend right, "0"'] & B \arrow[r, "f"] & C
\end{tikzcd}$$

(Note that the kernel is not just an object; it is a morphism of an object to B. In practice the term is often applied to just the object, and the intended interpretation is clear from the context.) Hence it is unique up to unique isomorphism by universal property nonsense. The kernel is written $\ker f \to B$. A **cokernel** (denoted $\operatorname{coker} f$) is defined dually by reversing the arrows — do this yourself. The kernel of $f\colon B \to C$ is the limit (§1.3) of the diagram

(1.5.4.1)
$$\begin{tikzcd} & 0 \arrow[d] \\ B \arrow[r, "f"] & C \end{tikzcd}$$

and similarly the cokernel is a colimit (see (2.6.0.1)).

If $i\colon A \to B$ is a monomorphism, then we say that A is a **subobject** of B, where the map i is implicit. There is also the notion of **quotient object**, defined dually to subobject.

An **abelian category** is an additive category satisfying three additional properties.

(1) Every map has a kernel and cokernel.
(2) Every monomorphism is the kernel of its cokernel.
(3) Every epimorphism is the cokernel of its kernel.

It is a nonobvious (and imprecisely stated) fact that every property you want to be true about kernels, cokernels, etc. follows from these three. (Warning: in part of the literature, additional hypotheses are imposed as part of the definition.)

The **image** of a morphism $f\colon A \to B$ is defined as $\operatorname{im}(f) = \ker(\operatorname{coker} f)$ whenever it exists (e.g., in every abelian category). The morphism $f\colon A \to B$ factors uniquely through $\operatorname{im} f \to B$ whenever $\operatorname{im} f$ exists, and $A \to \operatorname{im} f$ is an epimorphism and a cokernel of $\ker f \to A$ in every abelian category. The reader may want to verify this as a (hard!) exercise.

The cokernel of a monomorphism is called the **quotient**. The quotient of a monomorphism $A \to B$ is often denoted B/A (with the map from B implicit).

We will leave the foundations of abelian categories untouched. The key thing to remember is that if you understand kernels, cokernels, images and so on in the category of modules over a given ring, you can manipulate objects in any abelian category. This is made precise by the Freyd-Mitchell Embedding Theorem (Remark 1.5.5).

However, the abelian categories we will come across will obviously be related to modules, and our intuition will clearly carry over, so we needn't invoke a theorem whose proof we haven't read. For example, we will show that sheaves of abelian groups on a topological space X form an abelian category (§2.6), and the interpretation in terms of "compatible germs" will connect notions of kernels, cokernels etc. of sheaves of abelian groups to the corresponding notions of abelian groups.

1.5.5. *Small remark on chasing diagrams.* It is useful to prove facts (and solve exercises) about abelian categories by chasing elements. Unfortunately, some commonly used abelian categories, such as the category of complexes (to be defined in Exercise 1.5.C), do not have "elements" — they are not naturally "sets with additional structure" in any obvious way. Nonetheless, proof by element-chasing can be justified by the Freyd-Mitchell Embedding Theorem: If $\mathscr{C}$ is an abelian category whose objects form a set, then there is a ring A and an exact, fully faithful functor from $\mathscr{C}$ into Mod_A, which embeds $\mathscr{C}$ as a full subcategory. (Unfortunately, the ring A need not be commutative.) A proof is sketched in [**Weib**, §1.6], and references to a complete proof are given there. A proof is also given in [**KS1**, §9.7]. The upshot is that to prove something about a diagram in some abelian category, we may assume that it is a diagram of modules over some ring, and we may then "diagram-chase" elements. Moreover, any fact about kernels, cokernels, and so on that holds in Mod_A holds in any abelian category.

If invoking a theorem whose proof you haven't read bothers you, a short alternative is Mac Lane's "elementary rules for chasing diagrams", [**Mac**, Thm. 3, p. 204]; [**Mac**, Lem. 4, p. 205] gives a proof of the Five Lemma (Exercise 1.6.6) as an example.

But in any case, do what you need to do to put your mind at ease, so you can move forward. Do as little as your conscience will allow.

1.5.6. Complexes, exactness, and homology.

(In this entire discussion, we assume we are working in an abelian category.) We say a sequence

$$\cdots \longrightarrow A \xrightarrow{f} B \xrightarrow{g} C \longrightarrow \cdots \tag{1.5.6.1}$$

is a **complex at** B if $g \circ f = 0$, and is **exact at** B if $\ker g = \operatorname{im} f$. (More specifically, g has a kernel that is an image of f. Exactness at B implies being a complex at B — do you see why?) A sequence is a **complex** (resp., **exact**) if it is a complex (resp., exact) at each (internal) term. A **short exact sequence** is an exact sequence with five terms, the first and last of which are zeros — in other words, an exact sequence of the form

$$0 \longrightarrow A \longrightarrow B \longrightarrow C \longrightarrow 0.$$

For example, $0 \longrightarrow A \longrightarrow 0$ is exact if and only if $A = 0$;

$$0 \longrightarrow A \xrightarrow{f} B$$

is exact if and only if f is a monomorphism (with a similar statement for $A \xrightarrow{f} B \longrightarrow 0$);

$$0 \longrightarrow A \xrightarrow{f} B \longrightarrow 0$$

is exact if and only if f is an isomorphism; and

$$0 \longrightarrow A \xrightarrow{f} B \xrightarrow{g} C$$

is exact if and only if f is a kernel of g (with a similar statement for $A \xrightarrow{f} B \xrightarrow{g} C \longrightarrow 0$). To show some of these facts it may be helpful to prove that (1.5.6.1) is exact at B if and only if the cokernel of f is a cokernel of the kernel of g.

If you would like practice in playing with these notions before thinking about homology, you can prove the Snake Lemma (stated in Example 1.6.5, with a stronger version in Exercise 1.6.B), or the Five Lemma (stated in Example 1.6.6, with a stronger version in Exercise 1.6.C). (I would do this in the category of A-modules, but see [**KS1**, Lem. 12.1.1, Lem. 8.3.13] for proofs in general.)

If (1.5.6.1) is a complex at B, then its **homology at** B (often denoted by H) is $\ker g \,/ \operatorname{im} f$. (More precisely, there is some monomorphism $\operatorname{im} f \hookrightarrow \ker g$, and that H is the cokernel of this monomorphism.) Therefore, (1.5.6.1) is exact at B if and only if its homology at B is 0. We say that elements of $\ker g$ (assuming the objects of the category are sets with some additional structure) are the **cycles**, and elements of $\operatorname{im} f$ are the **boundaries** (so homology is "cycles mod boundaries"). If the complex is indexed in decreasing order, the indices are often written as subscripts, and H_i is the homology at $A_{i+1} \to A_i \to A_{i-1}$. If the complex is indexed in increasing order, the indices are often written as superscripts, and the homology H^i at $A^{i-1} \to A^i \to A^{i+1}$ is often called **cohomology**.

An exact sequence

$$(1.5.6.2) \qquad A^\bullet: \qquad \cdots \longrightarrow A^{i-1} \xrightarrow{f^{i-1}} A^i \xrightarrow{f^i} A^{i+1} \xrightarrow{f^{i+1}} \cdots$$

can be "factored" into short exact sequences

$$0 \longrightarrow \ker f^i \longrightarrow A^i \longrightarrow \ker f^{i+1} \longrightarrow 0$$

which is helpful in proving facts about long exact sequences by reducing them to facts about short exact sequences.

More generally, if (1.5.6.2) is assumed only to be a complex, then it can be "factored" into short exact sequences.

$$(1.5.6.3) \qquad 0 \longrightarrow \ker f^i \longrightarrow A^i \longrightarrow \operatorname{im} f^i \longrightarrow 0$$

$$0 \longrightarrow \operatorname{im} f^{i-1} \longrightarrow \ker f^i \longrightarrow H^i(A^\bullet) \longrightarrow 0$$

1.5.A. EXERCISE. Describe exact sequences

$$(1.5.6.4) \qquad 0 \longrightarrow \operatorname{im} f^i \longrightarrow A^{i+1} \longrightarrow \operatorname{coker} f^i \longrightarrow 0$$

$$0 \longrightarrow H^i(A^\bullet) \longrightarrow \operatorname{coker} f^{i-1} \longrightarrow \operatorname{im} f^i \longrightarrow 0$$

(These are somehow dual to (1.5.6.3). In fact in some mirror universe this might have been given as the standard definition of homology.) Assume the category is that of modules over a fixed ring for convenience, but be aware that the result is true for any abelian category.

1.5.B. EXERCISE AND IMPORTANT DEFINITION. Suppose

$$0 \xrightarrow{d^0} A^1 \xrightarrow{d^1} \cdots \xrightarrow{d^{n-1}} A^n \xrightarrow{d^n} 0$$

is a complex of finite-dimensional k-vector spaces (often called $A^\bullet$ for short). Define $h^i(A^\bullet) := \dim H^i(A^\bullet)$. Show that $\sum (-1)^i \dim A^i = \sum (-1)^i h^i(A^\bullet)$. In particular, if $A^\bullet$ is exact, then $\sum (-1)^i \dim A^i = 0$. (If you haven't dealt much with cohomology, this will give you some practice.)

1.5.C. IMPORTANT EXERCISE. Suppose $\mathscr{C}$ is an abelian category. Define the **category $Com_{\mathscr{C}}$ of complexes**) as follows. The objects are infinite complexes

$$A^\bullet: \qquad \cdots \longrightarrow A^{i-1} \xrightarrow{f^{i-1}} A^i \xrightarrow{f^i} A^{i+1} \xrightarrow{f^{i+1}} \cdots$$

in $\mathscr{C}$, and the morphisms $A^\bullet \to B^\bullet$ are commuting diagrams

$$\begin{tikzcd} \cdots \arrow[r] & A^{i-1} \arrow[r, "f^{i-1}"] \arrow[d] & A^i \arrow[r, "f^i"] \arrow[d] & A^{i+1} \arrow[r, "f^{i+1}"] \arrow[d] & \cdots \\ \cdots \arrow[r] & B^{i-1} \arrow[r, "g^{i-1}"] & B^i \arrow[r, "g^i"] & B^{i+1} \arrow[r, "g^{i+1}"] & \cdots \end{tikzcd} \tag{1.5.6.5}$$

Show that $Com_{\mathscr{C}}$ is an abelian category.

Feel free to deal with the special case of modules over a fixed ring. (Remark for experts: Essentially the same argument shows that $\mathscr{C}^{\mathscr{I}}$ is an abelian category for any small category $\mathscr{I}$ and any abelian category $\mathscr{C}$. This immediately implies that the category of presheaves on a topological space X with values in an abelian category $\mathscr{C}$ is automatically an abelian category, cf. §2.3.5.)

1.5.D. IMPORTANT EXERCISE. Show that (1.5.6.5) induces a map of homology $H^i(A^\bullet) \to H^i(B^\bullet)$. Show furthermore that H^i is a covariant functor $Com_{\mathscr{C}} \to \mathscr{C}$. (Again, feel free to deal with the special case Mod_A.)

1.5.7. *Homotopic maps induce the same maps on homology.* We say two maps of complexes $f\colon C^\bullet \to D^\bullet$ and $g\colon C^\bullet \to D^\bullet$ are **homotopic** if there is a sequence of maps $w\colon C^i \to D^{i-1}$ such that $f - g = dw + wd$.

1.5.E. EXERCISE. Show that two homotopic maps give the same map on homology.

1.5.8. Theorem (Long exact sequences). — *A short exact sequence of complexes*

$$\begin{tikzcd} 0^\bullet \arrow[d] & : & \cdots \arrow[r] & 0 \arrow[r] \arrow[d] & 0 \arrow[r] \arrow[d] & 0 \arrow[r] \arrow[d] & \cdots \\ A^\bullet \arrow[d] & : & \cdots \arrow[r] & A^{i-1} \arrow[r, "f^{i-1}"] \arrow[d] & A^i \arrow[r, "f^i"] \arrow[d] & A^{i+1} \arrow[r, "f^{i+1}"] \arrow[d] & \cdots \\ B^\bullet \arrow[d] & : & \cdots \arrow[r] & B^{i-1} \arrow[r, "g^{i-1}"] \arrow[d] & B^i \arrow[r, "g^i"] \arrow[d] & B^{i+1} \arrow[r, "g^{i+1}"] \arrow[d] & \cdots \\ C^\bullet \arrow[d] & : & \cdots \arrow[r] & C^{i-1} \arrow[r, "h^{i-1}"] \arrow[d] & C^i \arrow[r, "h^i"] \arrow[d] & C^{i+1} \arrow[r, "h^{i+1}"] \arrow[d] & \cdots \\ 0^\bullet & : & \cdots \arrow[r] & 0 \arrow[r] & 0 \arrow[r] & 0 \arrow[r] & \cdots \end{tikzcd}$$

induces a **long exact sequence in cohomology**

$$\begin{tikzcd} & & \cdots \arrow[r] & H^{i-1}(C^\bullet) \arrow[r] & {} \\ H^i(A^\bullet) \arrow[r] & H^i(B^\bullet) \arrow[r] & H^i(C^\bullet) \arrow[r] & {} & \\ H^{i+1}(A^\bullet) \arrow[r] & \cdots & & & \end{tikzcd}$$

(This requires a definition of the **connecting homomorphism** $H^{i-1}(C^\bullet) \to H^i(A^\bullet)$, which is "natural" in an appropriate sense.) In the category of modules over a ring, Theorem 1.5.8 will come out of our discussion of spectral sequences, see Exercise 1.6.F, but this is a somewhat perverse way of proving it. For a proof in general, see [**KS1**, Theorem 12.3.3]. You may want to prove it yourself, by first proving a weaker version of the Snake Lemma (Example 1.6.5), where in the hypotheses (1.6.5.1), the 0's in the bottom left and top right are removed, and in the conclusion (1.6.5.2), the first and last 0's are removed.

1.5.9. *Exactness of functors.* If $F \colon \mathscr{A} \to \mathscr{B}$ is an additive covariant functor from one abelian category to another, we say that F is **right-exact** if the exactness of

$$A' \longrightarrow A \longrightarrow A'' \longrightarrow 0,$$

in $\mathscr{A}$ implies that

$$F(A') \longrightarrow F(A) \longrightarrow F(A'') \longrightarrow 0$$

is also exact. Dually, we say that F is **left-exact** if the exactness of

$$0 \longrightarrow A' \longrightarrow A \longrightarrow A'' \qquad \text{implies}$$

$$0 \longrightarrow F(A') \longrightarrow F(A) \longrightarrow F(A'') \qquad \text{is exact.}$$

An additive contravariant functor is **left-exact** if the exactness of

$$A' \longrightarrow A \longrightarrow A'' \longrightarrow 0 \qquad \text{implies}$$

$$0 \longrightarrow F(A'') \longrightarrow F(A) \longrightarrow F(A') \qquad \text{is exact.}$$

The reader should be able to deduce what it means for a contravariant functor to be **right-exact**.

An additive covariant or contravariant functor is **exact** if it is both left-exact and right-exact.

1.5.F. EXERCISE. Suppose F is an exact functor. Show that applying F to an exact sequence preserves exactness. For example, if F is covariant, and $A' \to A \to A''$ is exact, then $FA' \to FA \to FA''$ is exact. (This will be generalized in Exercise 1.5.I(c).)

1.5.G. EXERCISE. Suppose A is a ring, $S \subset A$ is a multiplicative subset, and M is an A-module.
(a) Show that localization of A-modules $Mod_A \to Mod_{S^{-1}A}$ is an exact covariant functor.
(b) Show that $(\cdot) \otimes_A M$ is a right-exact covariant functor $Mod_A \to Mod_A$. (This is a repeat of Exercise 1.2.H.)
(c) Show that $\mathrm{Hom}(M, \cdot)$ is a left-exact covariant functor $Mod_A \to Mod_A$. If $\mathscr{C}$ is any abelian category, and $C \in \mathscr{C}$, show that $\mathrm{Hom}(C, \cdot)$ is a left-exact covariant functor $\mathscr{C} \to Ab$.
(d) Show that $\mathrm{Hom}(\cdot, M)$ is a left-exact contravariant functor $Mod_A \to Mod_A$. If $\mathscr{C}$ is any abelian category, and $C \in \mathscr{C}$, show that $\mathrm{Hom}(\cdot, C)$ is a left-exact contravariant functor $\mathscr{C} \to Ab$.

1.5.H. EXERCISE. Suppose M is a **finitely presented** A**-module**: M has a finite number of generators, and with these generators it has a finite number of relations; or usefully equivalently, fits in an exact sequence

$$A^{\oplus q} \longrightarrow A^{\oplus p} \longrightarrow M \longrightarrow 0 \tag{1.5.9.1}$$

Use (1.5.9.1) and the left-exactness of Hom to describe an isomorphism

$$S^{-1}\operatorname{Hom}_A(M,N) \xleftrightarrow{\sim} \operatorname{Hom}_{S^{-1}A}(S^{-1}M, S^{-1}N).$$

(You might be able to interpret this in light of a variant of Exercise 1.5.I below, for left-exact contravariant functors rather than right-exact covariant functors.)

1.5.10. *Example:* Hom *doesn't always commute with localization.* In the language of Exercise 1.5.H, take $A = N = \mathbb{Z}$, $M = \mathbb{Q}$, and $S = \mathbb{Z} \setminus \{0\}$.

1.5.11. ⋆ **Two useful facts in homological algebra.**

We now come to two (sets of) facts I wish I had learned as a child, as they would have saved me lots of grief. They encapsulate what is best and worst of abstract nonsense. The statements are so general as to be nonintuitive. The proofs are very short. They generalize some specific behavior that is easy to prove on an ad hoc basis. Once they are second nature to you, many subtle facts will become obvious to you as special cases. And you will see that they will get used (implicitly or explicitly) repeatedly.

1.5.12. ⋆ *Interaction of homology and (right/left-)exact functors.*

You might wait to prove this until you learn about cohomology in Chapter 18, when it will first be used in a serious way.

1.5.I. IMPORTANT EXERCISE (THE FHHF THEOREM). This result can take you far, and perhaps for that reason it has sometimes been called the Fernbahnhof (FernbaHnHoF) Theorem, notably in [**Vak1**, Exer. 1.5.I]. Suppose $F\colon \mathscr{A} \to \mathscr{B}$ is a covariant functor of abelian categories, and $C^\bullet$ is a complex in $\mathscr{A}$.

(a) (F *right-exact yields* $FH^\bullet \longrightarrow H^\bullet F$) If F is right-exact, describe a natural morphism $FH^\bullet \to H^\bullet F$. (More precisely, for each i, the left side is F applied to the cohomology at piece i of $C^\bullet$, while the right side is the cohomology at piece i of $FC^\bullet$.)

(b) (F *left-exact yields* $FH^\bullet \longleftarrow H^\bullet F$) If F is left-exact, describe a natural morphism $H^\bullet F \to FH^\bullet$.

(c) (F *exact yields* $FH^\bullet \xleftrightarrow{\sim} H^\bullet F$) If F is exact, show that the morphisms of (a) and (b) are inverses and thus isomorphisms.

Hint for (a): use $C^i \xrightarrow{d^i} C^{i+1} \longrightarrow \operatorname{coker} d^i \longrightarrow 0$ to give an isomorphism $F\operatorname{coker} d^i \xleftrightarrow{\sim} \operatorname{coker} Fd^i$. Then use the first line of (1.5.6.4) to give an epimorphism $F\operatorname{im} d^i \twoheadrightarrow \operatorname{im} Fd^i$. Then use the second line of (1.5.6.4) to give the desired map $FH^iC^\bullet \to H^iFC^\bullet$. While you are at it, you may as well describe a map for the fourth member of the quartet $\{\operatorname{coker}, \operatorname{im}, H, \ker\}$: $F\ker d^i \to \ker Fd^i$.

1.5.13. If this makes your head spin, you may prefer to think of it in the following specific case, where both $\mathscr{A}$ and $\mathscr{B}$ are the category of A-modules, and F is $(\cdot) \otimes_A N$ for some fixed A-module N. Your argument in this case will translate without change to yield a solution to Exercise 1.5.I(a) and (c) in general. If $\otimes N$ is exact, then N is called a **flat** A-module. (The notion of flatness will turn out to be very important, and is discussed in detail in Chapter 24.)

For example, localization is exact (Exercise 1.5.G(a)), so $S^{-1}A$ is a *flat* A-algebra for all multiplicative sets S. Thus taking cohomology of a complex of A-modules commutes with localization — something you could verify directly.

1.5.14. *Interaction of adjoints, (co)limits, and (left- and right-) exactness.*

A surprising number of arguments boil down to the statement:

Limits commute with limits and right adjoints. In particular, in an abelian category, because kernels are limits, both limits and right adjoints are left-exact.

as well as its dual:

Colimits commute with colimits and left adjoints. In particular, because cokernels are colimits, both colimits and left adjoints are right-exact.

These statements were promised in §1.4.4, and will be proved below. The latter has a useful extension:

In Mod_A, colimits over filtered index categories are exact. "Filtered" was defined in §1.3.8.

1.5.15. $\star\star$ *Caution.* It is not true that in abelian categories in general, colimits over filtered index categories are exact. (Grothendieck realized the desirability of such colimits being exact, and formalized this as his "AB5" axiom, see for example [**Stacks**, tag 079A].) Here is a counterexample. Because the axioms of abelian categories are self-dual, it suffices to give an example in which a *cofiltered limit* fails to be exact (where **cofiltered** has the obvious dual definition to *filtered*), and we do this. Fix a prime p. In the category *Ab* of abelian groups, for each positive integer n, we have an exact sequence $\mathbb{Z} \to \mathbb{Z}/(p^n) \to 0$. Taking the limit over all n in the obvious way, we obtain $\mathbb{Z} \to \mathbb{Z}_p \to 0$, which is certainly not exact.)

Unimportant Remark 1.5.18 will dash another hope you may have.

1.5.16. If you want to use these statements (for example, later in this book), you will have to prove them. Let's now make them precise.

1.5.J. EXERCISE (KERNELS COMMUTE WITH LIMITS). Suppose $\mathscr{C}$ is an abelian category, and $a\colon \mathscr{I} \to \mathscr{C}$ and $b\colon \mathscr{I} \to \mathscr{C}$ are two diagrams in $\mathscr{C}$ indexed by $\mathscr{I}$. For convenience, let $A_i = a(i)$ and $B_i = b(i)$ be the objects in those two diagrams. Let $h_i\colon A_i \to B_i$ be maps commuting with the maps in the diagram. (Translation: h is a natural transformation of functors $a \to b$, see §1.1.21.) Then the $\ker h_i$ form another diagram in $\mathscr{C}$ indexed by $\mathscr{I}$. Describe a canonical isomorphism $\lim \ker h_i \overset{\sim}{\longleftrightarrow} \ker(\lim A_i \to \lim B_i)$, assuming the limits exist.

Implicit in the previous exercise is the idea that limits should somehow be understood as functors.

1.5.K. EXERCISE. Make sense of the statement that "limits commute with limits" in a general category, and prove it. (Hint: recall that kernels are limits. The previous exercise should be a corollary of this one.)

1.5.17. Proposition (right adjoints commute with limits). — *Suppose* $(F\colon \mathscr{C} \to \mathscr{D}, G\colon \mathscr{D} \to \mathscr{C})$ *is a pair of adjoint functors. If* $A = \lim A_i$ *is a limit in* $\mathscr{D}$ *of a diagram* *indexed by* $\mathscr{I}$*, then* $GA = \lim GA_i$ *(with the corresponding maps* $GA \to GA_i$*) is a limit in* $\mathscr{C}$*.*

Proof. We must show that $GA \to GA_i$ satisfies the universal property of limits. Suppose we have maps $W \to GA_i$ commuting with the maps of $\mathscr{I}$. We wish to show that there exists a unique $W \to GA$ extending the $W \to GA_i$. By adjointness of F and G, we can restate this as: Suppose we have maps $FW \to A_i$ commuting with the maps of $\mathscr{I}$. We wish to show that there exists a unique $FW \to A$ extending the $FW \to A_i$. But this is precisely the universal property of the limit. □

Of course, the dual statements to Exercise 1.5.K and Proposition 1.5.17 hold by the dual arguments.

If F and G are additive functors between abelian categories, and (F, G) is an adjoint pair, then (as kernels are limits and cokernels are colimits) G is left-exact and F is right-exact.

1.5.L. EXERCISE. Show that in Mod_A, colimits over filtered index categories are exact. (Your argument will apply without change to any abelian category whose objects can be interpreted as "sets with additional structure".) Right-exactness follows from the above discussion, so the issue is left-exactness. (Possible hint: After you show that localization is exact, Exercise 1.5.G(a), or stalkification is exact, Exercise 2.6.E, in a hands-on way, you will be easily able to prove this. Conversely, if you do this exercise, those two will be easy.)

1.5.M. EXERCISE. Show that filtered colimits commute with homology in Mod_A. Hint: use the FHHF Theorem (Exercise 1.5.I), and the previous Exercise.

In light of Exercise 1.5.M, you may want to think about how limits (and colimits) commute with homology in general, and which way maps go. The statement of the FHHF Theorem should suggest the answer. (Are limits analogous to left-exact functors, or right-exact functors?) We won't directly use this insight, but see §18.1 **(vii)** for an example.

Just as colimits are exact (not just right-exact) in especially good circumstances, limits are exact (not just left-exact) too. The following will be used twice in Chapter 28.

1.5.N. EXERCISE. Suppose

$$\begin{tikzcd}
 & \vdots \arrow[d] & \vdots \arrow[d] & \vdots \arrow[d] & \\
0 \arrow[r] & A_{n+1} \arrow[r] \arrow[d] & B_{n+1} \arrow[r] \arrow[d] & C_{n+1} \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & A_n \arrow[r] \arrow[d] & B_n \arrow[r] \arrow[d] & C_n \arrow[r] \arrow[d] & 0 \\
 & \vdots \arrow[d] & \vdots \arrow[d] & \vdots \arrow[d] & \\
0 \arrow[r] & A_0 \arrow[r] & B_0 \arrow[r] & C_0 \arrow[r] & 0
\end{tikzcd}$$

is an inverse system of exact sequences of modules over a ring, such that the maps $A_{n+1} \to A_n$ are surjective. (We say: "transition maps of the left term are surjective".) Show that the limit

$$0 \longrightarrow \lim A_n \longrightarrow \lim B_n \longrightarrow \lim C_n \longrightarrow 0 \tag{1.5.17.1}$$

is also exact. (You will need to define the maps in (1.5.17.1).)

1.5.18. *Unimportant Remark.* Based on these ideas, you may suspect that right-exact functors always commute with colimits. The fact that tensor product commutes with infinite direct sums (Exercise 1.2.M) may reinforce this idea. Unfortunately, it is not true — "double dual" $\cdot^{\vee\vee} : \mathit{Vec}_k \to \mathit{Vec}_k$ is covariant and right exact (in fact, exact), but does not commute with infinite direct sums, as $\oplus_{i=1}^{\infty}(k^{\vee\vee})$ is not isomorphic to $(\oplus_{i=1}^{\infty} k)^{\vee\vee}$.

1.5.19. ⋆ **Dreaming of derived functors.** When you see a left-exact functor, you should always dream that you are seeing the end of a long exact sequence. If

$$0 \longrightarrow M' \longrightarrow M \longrightarrow M'' \longrightarrow 0$$

is an exact sequence in abelian category $\mathscr{A}$, and $F \colon \mathscr{A} \to \mathscr{B}$ is a left-exact functor, then

$$0 \longrightarrow FM' \longrightarrow FM \longrightarrow FM''$$

is exact, and you should always dream that it should continue in some natural way. For example, the next term should depend only on M', call it R^1FM', and if it is zero, then $FM \to FM''$ is an epimorphism. This remark holds true for left-exact and contravariant functors too. In good cases, such a continuation exists, and is incredibly useful. We will discuss this in Chapter 23.

1.6 ⋆ Spectral sequences

Je suis quelque peu affolé par ce déluge de cohomologie, mais j'ai courageusement tenu le coup. Ta suite spectrale me paraît raisonnable (je croyais, sur un cas particulier, l'avoir mise en défaut, mais je m'étais trompé, et cela marche au contraire admirablement bien).

I am a bit panic-stricken by this flood of cohomology, but have borne up courageously. Your spectral sequence seems reasonable to me (I thought I had shown that it was wrong in a special case, but I was mistaken, on the contrary it works remarkably well).
— J.-P. Serre, letter to A. Grothendieck, March 14, 1956 [**GrS**, p. 38]

Spectral sequences are a powerful book-keeping tool for proving things involving complicated commutative diagrams. They were introduced by Leray in the 1940's at the same time as he introduced sheaves. They have a reputation for being abstruse and difficult. It has been suggested that the name 'spectral' was given because, like specters, spectral sequences are terrifying, evil, and dangerous. I have heard no one disagree with this interpretation, which is perhaps not surprising since I just made it up.

Nonetheless, the goal of this section is to tell you enough that you can use spectral sequences without hesitation or fear, and why you shouldn't be frightened when they come up in a seminar. What is perhaps different in this presentation is that we will use spectral sequences to prove things that you may have already seen, and that you can prove easily in other ways. This will allow you to get some hands-on experience for how to use them. We will also see them only in the special case of double complexes (the version by far the most often used in algebraic geometry), and not in the general form usually presented (filtered complexes, exact couples, etc.). See [**Weib**, Ch. 5] for more detailed information if you wish.

You should *not* read this section when you are reading the rest of Chapter 1. Instead, you should read it just before you need it for the first time. When you finally *do* read this section, you *must* do the exercises up to Exercise 1.6.F.

For concreteness, we work in the category Mod_A of module over a ring A. However, everything we say will apply in any abelian category. (And if it helps you feel secure, work instead in the category Vec_k of vector spaces over a field k.)

1.6.1. Double complexes.

A **double complex** is a collection of A-modules $E^{p,q}$ ($p, q \in \mathbb{Z}$), and "rightward" morphisms $d_{\rightarrow}^{p,q} : E^{p,q} \to E^{p+1,q}$ and "upward" morphisms $d_{\uparrow}^{p,q} : E^{p,q} \to E^{p,q+1}$. In the superscript, the first entry denotes the column number (the "x-coordinate"), and the second entry denotes the row number (the "y-coordinate"). (Warning: this is opposite to the convention for matrices.) The subscript is meant to suggest the direction of the arrows. We will always write these as $d_{\rightarrow}$ and $d_{\uparrow}$ and ignore the superscripts. We require that $d_{\rightarrow}$ and $d_{\uparrow}$ satisfy (a) $d_{\rightarrow}^2 = 0$, (b) $d_{\uparrow}^2 = 0$, and one more condition: (c) either $d_{\rightarrow} d_{\uparrow} = d_{\uparrow} d_{\rightarrow}$ (all the squares commute) or $d_{\rightarrow} d_{\uparrow} + d_{\uparrow} d_{\rightarrow} = 0$ (they all anticommute). Both come up in nature, and you can switch from one to the other by replacing $d_{\uparrow}^{p,q}$ with $(-1)^p d_{\uparrow}^{p,q}$. So I will assume that all the squares anticommute, but that you know how to turn the commuting case into this one. (You will see that there is no difference in the recipe, basically because the image and kernel of a homomorphism f equal the image and kernel respectively of $-f$.)

$$\begin{tikzcd} E^{p,q+1} \arrow[r, "d_{\rightarrow}^{p,q+1}"] & E^{p+1,q+1} \\ E^{p,q} \arrow[u, "d_{\uparrow}^{p,q}"] \arrow[r, "d_{\rightarrow}^{p,q}"'] \arrow[ur, phantom, "\text{anticommutes}"] & E^{p+1,q} \arrow[u, "d_{\uparrow}^{p+1,q}"'] \end{tikzcd}$$

There are variations on this definition, where for example the vertical arrows go downwards, or some subset of the $E^{p,q}$ is required to be zero.

From the double complex we construct a corresponding (single) complex $E^{\bullet}$ with $E^k = \oplus_i E^{i,k-i}$, with $d = d_{\rightarrow} + d_{\uparrow}$. In other words, when there is a *single* superscript k, we mean a sum of the kth antidiagonal of the double complex. The single complex is sometimes called the **total complex**. Note that $d^2 = (d_{\rightarrow} + d_{\uparrow})^2 = d_{\rightarrow}^2 + (d_{\rightarrow} d_{\uparrow} + d_{\uparrow} d_{\rightarrow}) + d_{\uparrow}^2 = 0$, so $E^{\bullet}$ is indeed a complex.

The cohomology of the single complex is sometimes called the **hypercohomology** of the double complex. We will instead use the phrase **cohomology of the double complex**.

Our initial goal will be to find the cohomology of the double complex. You will see later that we secretly also have other goals.

A spectral sequence is a recipe for computing some information about the cohomology of the double complex. I won't yet give the full recipe. Surprisingly, this fragmentary bit of information is sufficient to prove lots of things.

1.6.2. Approximate Definition. A **spectral sequence** with **rightward orientation** is a sequence of tables or **pages** ${}_{\rightarrow}E_0^{p,q}$, ${}_{\rightarrow}E_1^{p,q}$, ${}_{\rightarrow}E_2^{p,q}$, … ($p, q \in \mathbb{Z}$), where ${}_{\rightarrow}E_0^{p,q} = E^{p,q}$, along with a differential

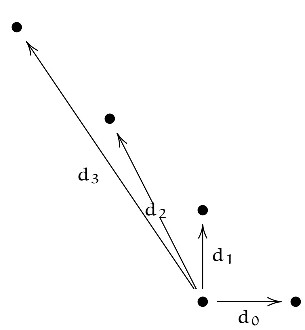
$$ {}_{\rightarrow}d_r^{p,q} \colon {}_{\rightarrow}E_r^{p,q} \longrightarrow {}_{\rightarrow}E_r^{p-r+1,q+r} $$

($r \in \mathbb{Z}^{\geq 0}$) with ${}_{\rightarrow}d_r^{p,q} \circ {}_{\rightarrow}d_r^{p+r-1,q-r} = 0$, and with an isomorphism of the cohomology of ${}_{\rightarrow}d_r$ at ${}_{\rightarrow}E_r^{p,q}$ (i.e., $\ker {}_{\rightarrow}d_r^{p,q} / \operatorname{im} {}_{\rightarrow}d_r^{p+r-1,q-r}$) with ${}_{\rightarrow}E_{r+1}^{p,q}$.

The orientation indicates that our 0th differential is the rightward one: $d_0 = d_{\rightarrow}$. The left subscript "$\rightarrow$" is usually omitted.

The order of the morphisms is best understood visually:

(1.6.2.1)



(the morphisms each apply to different pages). Notice that the map always is "degree 1" in terms of the grading of the single complex $E^\bullet$. (You should figure out what this informal statement really means.)

The actual definition describes what $E_r^{\bullet,\bullet}$ and $d_r^{\bullet,\bullet}$ really are, in terms of $E^{\bullet,\bullet}$. We will describe d_0, d_1, and d_2 below, and you should for now take on faith that this sequence continues in some natural way.

Note that $E_r^{p,q}$ is always a subquotient of the corresponding term on the ith page $E_i^{p,q}$ for all $i < r$. In particular, if $E^{p,q} = 0$, then $E_r^{p,q} = 0$ for all r.

Suppose now that $E^{\bullet,\bullet}$ is a **first quadrant double complex**, i.e., $E^{p,q} = 0$ for $p < 0$ or $q < 0$ (so $E_r^{p,q} = 0$ for all r unless $p, q \in \mathbb{Z}^{\geq 0}$). Then for any fixed p, q, once r is sufficiently large, $E_{r+1}^{p,q}$ is computed from $(E_r^{\bullet,\bullet}, d_r)$ using the complex

$$0 \xrightarrow{d_r^{p+r-1,q-r}} E_r^{p,q} \xrightarrow{d_r^{p,q}} 0$$

and thus we have canonical isomorphisms

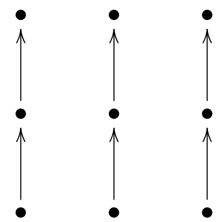
$$E_r^{p,q} \cong E_{r+1}^{p,q} \cong E_{r+2}^{p,q} \cong \cdots .$$

We denote this module $E_\infty^{p,q}$. The same idea works in other circumstances, for example if the double complex is only nonzero in a finite number of rows — $E^{p,q} = 0$ unless $q_0 < q < q_1$. This will come up for example in the mapping cone and long exact sequence discussion (Exercises 1.6.F and 1.6.E below).

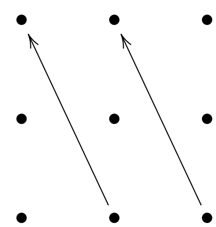
We now describe the first few pages of the spectral sequence explicitly. As stated above, the differential d_0 on $E_0^{\bullet,\bullet} = E^{\bullet,\bullet}$ is defined to be $d_\rightarrow$. The rows are complexes:

$$\text{The 0th page } E_0: \qquad \begin{array}{ccccc} \bullet & \longrightarrow & \bullet & \longrightarrow & \bullet \\ \bullet & \longrightarrow & \bullet & \longrightarrow & \bullet \\ \bullet & \longrightarrow & \bullet & \longrightarrow & \bullet \end{array}$$

and so E_1 is just the table of cohomologies of the rows. You should check that there are now vertical maps $d_1^{p,q} : E_1^{p,q} \to E_1^{p,q+1}$ of the row cohomology groups, induced by $d_\uparrow$, and that these make the columns into complexes. (This is essentially the fact that a map of complexes induces a map on homology.) We have "used up the horizontal morphisms", but "the vertical differentials live on".

The 1st page E_1: 

We take cohomology of d_1 on E_1, giving us a new table, $E_2^{p,q}$. It turns out that there are natural morphisms from each entry to the entry two above and one to the left, and that the composition of these two is 0. (It is a very worthwhile exercise to work out how this natural morphism d_2 should be defined. Your argument may be reminiscent of the connecting homomorphism in the Snake Lemma 1.6.5 or in the long exact sequence in cohomology arising from a short exact sequence of complexes, Theorem 1.5.8. This is no coincidence.)

The 2nd page E_2: 

This is the beginning of a pattern.

Then it is a theorem that there is a filtration of $H^k(E^\bullet)$ by $E_\infty^{p,q}$ where $p+q=k$. (We can't yet state it as an official **Theorem** because we haven't precisely defined the pages and differentials in the spectral sequence.) More precisely, there is a filtration

$$E_\infty^{0,k} \xhookrightarrow{E_\infty^{1,k-1}} \; ? \; \xhookrightarrow{E_\infty^{2,k-2}} \cdots \xhookrightarrow{E_\infty^{k,0}} H^k(E^\bullet) \tag{1.6.2.2}$$

where the quotients are displayed above each inclusion. (Here is a tip for remember which way the quotients are supposed to go. The differentials on later and later pages point deeper and deeper into the filtration. Thus the entries in the direction of the later arrowheads are the subobjects, and the entries in the direction of the later "arrowtails" are quotients. This tip has the advantage of being independent of the details of the spectral sequence, e.g., the "quadrant" or the orientation.)

We say that the spectral sequence ${}_{\rightarrow}E_\bullet^{\bullet,\bullet}$ **converges** to $H^\bullet(E^\bullet)$. We often say that ${}_{\rightarrow}E_2^{\bullet,\bullet}$ (or any other page) **abuts** to $H^\bullet(E^\bullet)$.

Although the filtration gives only partial information about $H^\bullet(E^\bullet)$, sometimes one can find $H^\bullet(E^\bullet)$ precisely. One example is if all $E_\infty^{i,k-i}$ are zero, or if all but one of them are zero (e.g., if $E_r^{\bullet,\bullet}$ has precisely one nonzero row or column, in which case one says that the spectral sequence **collapses** at the rth step, although we will not use this term). Another example is in the category of vector spaces over a field, in which case we can find the dimension of $H^k(E^\bullet)$. Also, in lucky circumstances, E_2 (or some other small page) already equals E_∞.

1.6.A. EXERCISE: INFORMATION FROM THE SECOND PAGE. Show that $H^0(E^\bullet) = E_\infty^{0,0} = E_2^{0,0}$ and

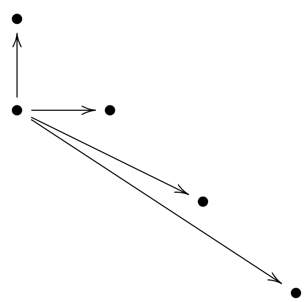
$$0 \longrightarrow E_2^{0,1} \longrightarrow H^1(E^\bullet) \longrightarrow E_2^{1,0} \xrightarrow{d_2^{1,0}} E_2^{0,2} \longrightarrow H^2(E^\bullet)$$

is exact.

1.6.3. *The other orientation.*

You may have observed that we could as well have done everything in the opposite direction, i.e., reversing the roles of horizontal and vertical morphisms. Then the sequences of arrows giving the spectral sequence would look like this (compare to (1.6.2.1)).

(1.6.3.1)



This spectral sequence is denoted ${}_\uparrow E_\bullet^{\bullet,\bullet}$ ("with the **upward orientation**"). Then we would again get pieces of a filtration of $H^\bullet(E^\bullet)$ (where we have to be a bit careful with the order with which ${}_\uparrow E_\infty^{p,q}$ corresponds to the subquotients — it is the opposite order to that of (1.6.2.2) for ${}_\rightarrow E_\infty^{p,q}$). Warning: in general there is no isomorphism between ${}_\rightarrow E_\infty^{p,q}$ and ${}_\uparrow E_\infty^{p,q}$.

In fact, this observation that we can start with either the horizontal or vertical maps was our secret goal all along. Both algorithms compute information about the same thing ($H^\bullet(E^\bullet)$), and usually we don't care about the final answer — we often care about the answer we get in one way, and we get at it by doing the spectral sequence in the *other* way.

1.6.4. Examples.

We are now ready to see how this is useful. The moral of these examples is the following. In the past, you may have proved various facts involving various sorts of diagrams, by chasing elements around. Now, you will just plug them into a spectral sequence, and let the spectral sequence machinery do your chasing for you.

1.6.5. *Example: Proving the Snake Lemma.* Consider the diagram

(1.6.5.1)
$$\begin{tikzcd}
0 \arrow[r] & D \arrow[r] & E \arrow[r] & F \arrow[r] & 0 \\
0 \arrow[r] & A \arrow[r] \arrow[u, "\alpha"] & B \arrow[r] \arrow[u, "\beta"] & C \arrow[r] \arrow[u, "\gamma"] & 0
\end{tikzcd}$$

where the rows are exact in the middle (at A, B, C, D, E, F) and the squares commute. (Normally the Snake Lemma is described with the vertical arrows pointing downwards, but I want to fit this into my spectral sequence conventions.)

We wish to show that there is an exact sequence

$$0 \to \ker \alpha \to \ker \beta \to \ker \gamma \to \operatorname{coker} \alpha \to \operatorname{coker} \beta \to \operatorname{coker} \gamma \to 0. \tag{1.6.5.2}$$

We plug this into our spectral sequence machinery. We first compute the cohomology using the rightward orientation, i.e., using the order (1.6.2.1). Then because the rows are exact, $E_1^{p,q} = 0$, so the spectral sequence has already converged: $E_\infty^{p,q} = 0$.

We next compute this "0" in another way, by computing the spectral sequence using the upward orientation. Then ${}_\uparrow E_1^{\bullet,\bullet}$ (with its differentials) is:

$$\begin{tikzcd}
0 \arrow[r] & \operatorname{coker} \alpha \arrow[r] & \operatorname{coker} \beta \arrow[r] & \operatorname{coker} \gamma \arrow[r] & 0 \\
0 \arrow[r] & \ker \alpha \arrow[r] & \ker \beta \arrow[r] & \ker \gamma \arrow[r] & 0.
\end{tikzcd}$$

Then ${}_\uparrow E_2^{\bullet,\bullet}$ is of the form:

$$\begin{tikzcd}
0 \arrow[drr] & 0 \arrow[drr] & & & & & \\
& 0 \arrow[drr] & ?? \arrow[drr] & ? \arrow[drr] & ? & 0 & \\
& 0 & ? & ? \arrow[drr] & ?? \arrow[drr] & 0 & \\
& & & & & 0 & 0
\end{tikzcd}$$

We see that after ${}_\uparrow E_2$, all the terms will stabilize except for the double question marks — all maps to and from the single question marks are to and from 0-entries. And after ${}_\uparrow E_3$, even these two double-question-mark terms will stabilize. But in the end our complex must be the 0 complex. This means that in ${}_\uparrow E_2$, all the entries must be zero, except for the two double question marks, and these two must be isomorphic. This means that $0 \to \ker \alpha \to \ker \beta \to \ker \gamma$ and $\operatorname{coker} \alpha \to \operatorname{coker} \beta \to \operatorname{coker} \gamma \to 0$ are both exact (that comes from the vanishing of the single question marks), and

$$\operatorname{coker}(\ker \beta \to \ker \gamma) \cong \ker(\operatorname{coker} \alpha \to \operatorname{coker} \beta)$$

is an isomorphism (that comes from the equality of the double question marks). Taken together, we have proved the exactness of (1.6.5.2), and hence the Snake Lemma! (Notice: in the end we didn't really care about the double complex. We just used it as a prop to prove the Snake Lemma.)

Spectral sequences make it easy to see how to generalize results further. For example, if $A \to B$ is no longer assumed to be injective, how would the conclusion change?

1.6.B. Unimportant Exercise (grafting exact sequences, a variant of the Snake Lemma). Extend the Snake Lemma as follows. Suppose we have a commuting diagram

$$\begin{tikzcd}
 & 0 \arrow[r] & X' \arrow[r] & Y' \arrow[r] & Z' \arrow[r] & A' \arrow[r] & \cdots \\
\cdots \arrow[r] & W \arrow[r] \arrow[u] & X \arrow[r] \arrow[u, "a"] & Y \arrow[r] \arrow[u, "b"] & Z \arrow[r] \arrow[u, "c"] & 0. \arrow[u] &
\end{tikzcd}$$

where the top and bottom rows are exact. Show that the top and bottom rows can be "grafted together" to an exact sequence

$$\begin{tikzcd}
\cdots \arrow[r] & W \arrow[r] & \ker a \arrow[r] & \ker b \arrow[r] & \ker c
\end{tikzcd}$$

$$\begin{tikzcd}
\arrow[r] & \operatorname{coker} a \arrow[r] & \operatorname{coker} b \arrow[r] & \operatorname{coker} c \arrow[r] & A' \arrow[r] & \cdots .
\end{tikzcd}$$

1.6.6. *Example: the Five Lemma.* Suppose

(1.6.6.1)
$$\begin{tikzcd}
F \arrow[r] & G \arrow[r] & H \arrow[r] & I \arrow[r] & J \\
A \arrow[r] \arrow[u, "\alpha"] & B \arrow[r] \arrow[u, "\beta"] & C \arrow[r] \arrow[u, "\gamma"] & D \arrow[r] \arrow[u, "\delta"] & E \arrow[u, "\epsilon"]
\end{tikzcd}$$

where the rows are exact and the squares commute.

Suppose α, β, δ, ϵ are isomorphisms. We will show that γ is an isomorphism.

We first compute the cohomology of the total complex using the rightward orientation (1.6.2.1). We choose this because we see that we will get lots of zeros. Then ${}_{\rightarrow}E_1^{\bullet,\bullet}$ looks like this:

$$\begin{tikzcd}
? & 0 & 0 & 0 & ? \\
? \arrow[u] & 0 \arrow[u] & 0 \arrow[u] & 0 \arrow[u] & ? \arrow[u]
\end{tikzcd}$$

Then ${}_{\rightarrow}E_2$ looks similar, and the sequence will converge by E_2, as we will never get any arrows between two nonzero entries in a table thereafter. We can't conclude that the cohomology of the total complex vanishes, but we can note that it vanishes in all but four degrees — and most important, it vanishes in the two degrees corresponding to the entries C and H (the source and target of γ).

We next compute this using the upward orientation (1.6.3.1). Then ${}_{\uparrow}E_1$ looks like this:

$$\begin{tikzcd}
0 \arrow[r] & 0 \arrow[r] & ? \arrow[r] & 0 \arrow[r] & 0 \\
0 \arrow[r] & 0 \arrow[r] & ? \arrow[r] & 0 \arrow[r] & 0
\end{tikzcd}$$

and the spectral sequence converges at this step. We wish to show that those two question marks are zero. But they are precisely the cohomology groups of the total complex that we just showed *were* zero — so we are done!

The best way to become comfortable with this sort of argument is to try it out yourself several times, and realize that it really is easy. So you should do the following exercises! Many can readily be done directly, but you should deliberately try to use this spectral sequence machinery in order to get practice and develop confidence.

1.6.C. EXERCISE: A SUBTLER FIVE LEMMA. By looking at the spectral sequence proof of the Five Lemma above, prove a subtler version of the Five Lemma, where one of the isomorphisms can instead just be required to be an injection, and another can instead just be required to be a surjection. (I am deliberately not telling you which ones, so you can see how the spectral sequence is telling you how to improve the result.)

1.6.D. EXERCISE: ANOTHER SUBTLE VERSION OF THE FIVE LEMMA. If β and δ (in (1.6.6.1)) are injective, and α is surjective, show that γ is injective. Give the dual statement (whose proof is of course essentially the same).

The next two exercises no longer involve first quadrant double complexes. You will have to think a little to realize why there is no reason for confusion or alarm.

1.6.E. IMPORTANT EXERCISE (THE MAPPING CONE). Suppose $\mu\colon A^\bullet \to B^\bullet$ is a morphism of complexes. Suppose $C^\bullet$ is the single complex associated to the double complex $A^\bullet \to B^\bullet$. ($C^\bullet$ is called the **mapping cone** of μ.) Show that there is a long exact sequence of complexes:

$$\cdots \to H^{i-1}(C^\bullet) \to H^i(A^\bullet) \to H^i(B^\bullet) \to H^i(C^\bullet) \to H^{i+1}(A^\bullet) \to \cdots .$$

(There is a slight notational ambiguity here; depending on how you index your double complex, your long exact sequence might look slightly different.) In particular, we will use the fact that μ induces an isomorphism on cohomology if and only if the mapping cone is exact. (We won't use it until the proof of Theorem 18.2.4.)

1.6.F. EXERCISE. Use spectral sequences to show that a short exact sequence of complexes gives a long exact sequence in cohomology (Theorem 1.5.8). (This is a generalization of Exercise 1.6.E.)

The Grothendieck composition-of-functors spectral sequence (Theorem 23.3.5) will be an important example of a spectral sequence that specializes in a number of useful ways.

You are now ready to go out into the world and use spectral sequences to your heart's content!

1.6.7. Complete definition of spectral sequences, and proof.

You should most definitely not read the precise definition of a spectral sequence, and the proof that they work as advertised, any time soon after reading the introduction to spectral sequences above. But after a suitable interval, you should at least flip through a construction and proof to convince yourself that nothing fancy is involved. The idea is not as bad as you might think, see [**Vak2**].

It is useful to notice that the proof implies that spectral sequences are functorial in the 0th page: the spectral sequence formalism has good functorial properties in the double complex. Unfortunately, Grothendieck's terminology "spectral functor" [**Gr1**, §2.4] did not catch on.