MATH 216: FOUNDATIONS OF ALGEBRAIC GEOMETRY

August 26, 2010.

Contents

CHAPTER 1

Introduction

I can illustrate the .... approach with the ... image of a nut to be opened. The first analogy that came to my mind is of immersing the nut in some softening liquid, and why not simply water? From time to time you rub so the liquid penetrates better, and otherwise you let time pass. The shell becomes more flexible through weeks and months — when the time is ripe, hand pressure is enough, the shell opens like a perfectly ripened avocado!

A different image came to me a few weeks ago. The unknown thing to be known appeared to me as some stretch of earth or hard marl, resisting penetration ... the sea advances insensibly in silence, nothing seems to happen, nothing moves, the water is so far off you hardly hear it ... yet finally it surrounds the resistant substance.

— *Grothendieck,* Récoltes et Semailles *p. 552-3, translation by Colin McLarty*

1.1 Goals

These are an updated version of notes accompanying a hard year-long class taught at Stanford in 2009-2010. I am currently editing them and adding a few more sections, and I hope a reasonably complete (if somewhat rough) version over the 2010-11 academic year at the site `http://math216.wordpress.com/`.

In any class, choices must be made as to what the course is about, and who it is for — there is a finite amount of time, and any addition of material or explanation or philosophy requires a corresponding subtraction. So these notes are highly inappropriate for most people and most classes. Here are my goals. (I do not claim that these goals are achieved; but they motivate the choices made.)

These notes currently have a very particular audience in mind: Stanford Ph.D. students, postdocs and faculty in a variety of fields, who may want to use algebraic geometry in a sophisticated way. This includes algebraic and arithmetic geometers, but also topologists, number theorists, symplectic geometers, and others.

The notes deal purely with the algebraic side of the subject, and completely neglect analytic aspects.

They assume little prior background (see §1.2), and indeed most students have little prior background. Readers with less background will necessarily have to work harder. It would be great if the reader had seen varieties before, but many students haven't, and the course does not assume it — and similarly for category theory, homological algebra, more advanced commutative algebra, differential geometry, ... Surprisingly often, what we need can be developed quickly from scratch. The cost is that the course is much denser; the benefit is that more people can follow it; they don't reach a point where they get thrown. (On the other hand, people who already have some familiarity with algebraic geometry, but want to understand the foundations more completely should not be bored, and will focus on more subtle issues.)

The notes seek to cover everything that one should see in a first course in the subject, including theorems, proofs, and examples.

They seek to be complete, and not leave important results as black boxes pulled from other references.

There are lots of exercises. I have found that unless I have some problems I can think through, ideas don't get fixed in my mind. Some are trivial — that's okay, and even desirable. As few necessary ones as possible should be hard, but the reader should have the background to deal with them — they are not just an excuse to push material out of the text.

There are optional (starred $\star$) sections of topics worth knowing on a second or third (but not first) reading. You should not read double-starred sections ($\star\star$) unless you really really want to, but you should be aware of their existence.

The notes are intended to be readable, although certainly not easy reading.

In short, after a year of hard work, students should have a broad familiarity with the foundations of the subject, and be ready to attend seminars, and learn more advanced material. They should not just have a vague intuitive understanding of the ideas of the subject; they should know interesting examples, know why they are interesting, and be able to prove interesting facts about them.

I have greatly enjoyed thinking through these notes, and teaching the corresponding classes, in a way I did not expect. I have had the chance to think through the structure of algebraic geometry from scratch, not blindly accepting the choices made by others. (Why do we need this notion? Aha, this forces us to consider this other notion earlier, and now I see why this third notion is so relevant...) I have repeatedly realized that ideas developed in Paris in the 1960's are simpler than I initially believed, once they are suitably digested.

1.1.1. Implications. We will work with as much generality as we need for most readers, and no more. In particular, we try to have hypotheses that are as general as possible without making proofs harder. The right hypotheses can make a proof easier, not harder, because one can remember how they get used. As an inflammatory example, the notion of quasiseparated comes up early and often. The cost is that one extra word has to be remembered, on top of an overwhelming number of other words. But once that is done, it is not hard to remember that essentially every scheme anyone cares about is quasiseparated. Furthermore, whenever the hypotheses "quasicompact and quasiseparated" turn up, the reader will likely immediately see a key idea of the proof.

Similarly, there is no need to work over an algebraically closed field, or even a field. Geometers needn't be afraid of arithmetic examples or of algebraic examples; a central insight of algebraic geometry is that the same formalism applies without change.

1.1.2. Costs. Choosing these priorities requires that others be shortchanged, and it is best to be up front about these. Because of our goal is to be comprehensive, and to understand everything one should know after a first course, it will necessarily take longer to get to interesting sample applications. You may be misled into thinking that one has to work this hard to get to these applications — it is not true!

1.2 Background and conventions

All rings are assumed to be commutative unless explicitly stated otherwise. All rings are assumed to contain a unit, denoted 1. Maps of rings must send 1 to 1. We don't require that $0 \neq 1$; in other words, the "0-ring" (with one element) is a ring. (There is a ring map from any ring to the 0-ring; the 0-ring only maps to itself. The 0-ring is the final object in the category of rings.) We accept the axiom of choice. In particular, any proper ideal in a ring is contained in a maximal ideal. (The axiom of choice also arises in the argument that the category of A-modules has enough injectives, see Exercise 23.2.E.)

The reader should be familiar with some basic notions in commutative ring theory, in particular the notion of ideals (including prime and maximal ideals) and localization. For example, the reader should be able to show that if S is a multiplicative set of a ring A (which we assume to contain 1), then the primes of $S^{-1}A$ are in natural bijection with those primes of A not meeting S (§4.2.6). Tensor products and exact sequences of A-modules will be important. We will use the notation $(A, \mathfrak{m})$ or $(A, \mathfrak{m}, k)$ for local rings — A is the ring, $\mathfrak{m}$ its maximal ideal, and $k = A/\mathfrak{m}$ its residue field. We will use (in Proposition 14.7.1) the structure theorem for finitely generated modules over a principal ideal domain A: any such module can be written as the direct sum of principal modules $A/(a)$.

We will not concern ourselves with subtle foundational issues (set-theoretic issues involving universes, etc.). It is true that some people should be careful about these issues. But is that really how you want to spend your life? (If so, a good start is [**KS**, §1.1].)

1.2.1. *Further background.* It may be helpful to have books on other subjects handy that you can dip into for specific facts, rather than reading them in advance. In commutative algebra, Eisenbud [**E**] is good for this. Other popular choices are Atiyah-Macdonald [**AM**] and Matsumura [**M-CRT**]. For homological algebra, Weibel [**W**] is simultaneously detailed and readable.

Background from other parts of mathematics (topology, geometry, complex analysis) will of course be helpful for developing intuition.

Finally, it may help to keep the following quote in mind.

Algebraic geometry seems to have acquired the reputation of being esoteric, exclusive, and very abstract, with adherents who are secretly plotting to take over all the rest of mathematics. In one respect this last point is accurate.

— David Mumford

Part I

Preliminaries

CHAPTER 2

Some category theory

That which does not kill me, makes me stronger. — Nietzsche

2.1 Motivation

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 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.

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 $\{{}^{u}_{v} : \ u \in U, v \in V\}$. 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 $\mu : P \to M$ and $\nu : P \to N$, such that for *any set* P' *with maps* $\mu' : P' \to M$ *and*

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

(2.1.0.1)
$$\begin{tikzcd} P' \arrow[dr, dotted, "\exists!"] \arrow[drr, bend left, "\nu'"] \arrow[ddr, bend right, "\mu'"'] & & \\ & P \arrow[r, "\nu"] \arrow[d, "\mu"] & N \\ & M & \end{tikzcd}$$

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 (2.1.0.1)), there is a unique map $f : P_1 \to P_2$ 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 $g : P_2 \to P_1$ 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 (2.1.0.1). There is a unique map $h : P_2 \to P_2$ such that

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

commutes. However, I can name two such maps: the identity map id_{P_2}, and $g \circ f$. Thus $g \circ f = \mathrm{id}_{P_2}$. Similarly, $f \circ g = \mathrm{id}_{P_1}$. 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, where the maps are taken to be linear maps; or the category of smooth manifolds, where the maps are taken to be smooth maps).

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 §2.3.

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 **[Gr]**.

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

2.2 Categories and functors

We begin with an informal definition of categories and functors.

2.2.1. Categories.

A **category** consists of a collection of **objects**, and for each pair of objects, a set of maps, or **morphisms** (or **arrows** or **maps**), between them. The collection of objects of a category $\mathcal{C}$ are often denoted $\mathrm{obj}(\mathcal{C})$, but we will usually denote the collection also by $\mathcal{C}$. If $A, B \in \mathcal{C}$, then the morphisms from A to B are denoted $\mathrm{Mor}(A, B)$. A morphism is often written $f : A \to B$, 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 $\mathrm{Mor}(A, B) \times \mathrm{Mor}(B, C) \to \mathrm{Mor}(A, C)$, 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 $A \in \mathcal{C}$, there is always an **identity morphism** $\mathrm{id}_A : A \to A$, such that when you (left- or right-)compose a morphism with the identity, you get the same morphism. More precisely, if $f : A \to B$ is a morphism, then $f \circ \mathrm{id}_A = f = \mathrm{id}_B \circ f$. (If you wish, you may check that "identity morphisms are unique": there is only on morphism deserving the name id_A.)

If we have a category, then we have a notion of **isomorphism** between two objects (a morphism $f : A \to B$ such that there exists some — necessarily unique — morphism $g : B \to A$, where $f \circ g$ and $g \circ f$ are the identity on B and A respectively), and a notion of **automorphism** of an object (an isomorphism of the object with itself).

2.2.2. *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 §1.2, we are deliberately omitting all set-theoretic issues.)

2.2.3. *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?)

2.2.A. UNIMPORTANT EXERCISE. A category in which each morphism is an isomorphism is called a **groupoid**. (This notion is not important in these notes. 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.
(b) Describe a groupoid that is not a group.

2.2.B. EXERCISE. If A is an object in a category $\mathcal{C}$, 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 2.2.2 and 2.2.3? 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.)

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

2.2.5. *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 §2.6.) Taking $A = k$, we obtain Example 2.2.3; taking $A = \mathbb{Z}$, we obtain Example 2.2.4.

2.2.6. *Example: rings.* There is a category *Rings*, where the objects are rings, and the morphisms are morphisms of rings (which send 1 to 1 by our conventions, §1.2).

2.2.7. *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. This needn't be the case, as the next example shows.

2.2.8. *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 $x = y$.

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

(2.2.8.1)
$$\begin{tikzcd} & \bullet \arrow[d] \\ \bullet \arrow[r] & \bullet \end{tikzcd}$$

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

(2.2.8.2)
$$\begin{tikzcd} \bullet \arrow[r] \arrow[d] & \bullet \arrow[d] \\ \bullet \arrow[r] & \bullet \end{tikzcd}$$

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.

2.2.9. *Example: the category of subsets of a set, and the category of open sets in a topological space.* If X is a set, then the subsets form a partially ordered set, where the order is given by inclusion. Similarly, if X is a topological space, then the *open* sets form a partially ordered set, where the order is given by inclusion.

2.2.10. *Example.* A **subcategory** $\mathcal{A}$ of a category $\mathcal{B}$ has as its objects some of the objects of $\mathcal{B}$, and some of the morphisms, such that the morphisms of $\mathcal{A}$ include the identity morphisms of the objects of $\mathcal{A}$, and are closed under composition. (For example, (2.2.8.1) is in an obvious way a subcategory of (2.2.8.2).)

2.2.11. Functors.

A **covariant functor** F from a category $\mathcal{A}$ to a category $\mathcal{B}$, denoted $F : \mathcal{A} \to \mathcal{B}$, is the following data. It is a map of objects $F : \mathrm{obj}(\mathcal{A}) \to \mathrm{obj}(\mathcal{B})$, and for each A_1, $A_2 \in \mathcal{A}$, and morphism $m : A_1 \to A_2$, a morphism $F(m) : F(A_1) \to F(A_2)$ in $\mathcal{B}$. We require that F preserves identity morphisms (for $A \in \mathcal{A}$, $F(\mathrm{id}_A) = \mathrm{id}_{F(A)}$), and that F preserves composition ($F(m_1 \circ m_2) = F(m_1) \circ F(m_2)$). (You may wish to verify that covariant functors send isomorphisms to isomorphisms.)

If $F : \mathcal{A} \to \mathcal{B}$ and $G : \mathcal{B} \to \mathcal{C}$ are covariant functors, then we define a functor $G \circ F : \mathcal{A} \to \mathcal{C}$ in the obvious way. Composition of functors is associative in an evident sense.

2.2.12. *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.

2.2.13. *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 $f : X \to Y$ with $f(x_0) = y_0$ induces a map of fundamental groups $\pi_1(X, x_0) \to \pi_1(Y, y_0)$, and similarly for homology groups.

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

$$[g : A \to B_1] \mapsto [f \circ g : A \to B_1 \to B_2].$$

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

2.2.15. Full and faithful functors. A covariant functor $F : \mathcal{A} \to \mathcal{B}$ is **faithful** if for all $A, A' \in \mathcal{A}$, the map $\mathrm{Mor}_{\mathcal{A}}(A, A') \to \mathrm{Mor}_{\mathcal{B}}(F(A), F(A'))$ is injective, and **full** if it is surjective. A functor that is full and faithful is **fully faithful**. A subcategory $i : \mathcal{A} \to \mathcal{B}$ is a **full subcategory** if i is full. Thus a subcategory $\mathcal{A}'$ of $\mathcal{A}$ is full if and only if for all $A, B \in \mathrm{obj}(\mathcal{A}')$, $\mathrm{Mor}_{\mathcal{A}'}(A, B) = \mathrm{Mor}_{\mathcal{A}}(A, B)$.

2.2.16. 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 $\mathcal{F}(m_2 \circ m_1)\mathcal{F}(m_1) \circ \mathcal{F}(m_2)$, not the other way around.)

It is wise to always state whether a functor is covariant or contravariant. If it is not stated, the functor is often assumed to be covariant.

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

2.2.17. *Linear algebra example.* If *Vec*$_k$ is the category of k-vector spaces (introduced in Example 2.2.12), then taking duals gives a contravariant functor $\cdot^{\vee}$: *Vec*$_k \to$ *Vec*$_k$. Indeed, to each linear transformation $f : V \to W$, we have a dual transformation $f^{\vee} : W^{\vee} \to V^{\vee}$, and $(f \circ g)^{\vee} = g^{\vee} \circ f^{\vee}$.

2.2.18. *Topological example (cf. Example 2.2.13).* The the ith cohomology functor $H^i(\cdot, \mathbb{Z})$: *Top* $\to$ *Ab* is a contravariant functor.

2.2.19. *Example.* There is a contravariant functor *Top* $\to$ *Rings* taking a topological space X to the 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 maps on X.

2.2.20. *Example (cf. Example 2.2.14).* Suppose A is an object of a category $\mathcal{C}$. Then there is a contravariant functor $h_A : \mathcal{C} \to$ *Sets* sending $B \in \mathcal{C}$ to $\operatorname{Mor}(B, A)$, and sending the morphism $f : B_1 \to B_2$ to the morphism $\operatorname{Mor}(B_2, A) \to \operatorname{Mor}(B_1, A)$ via

$$[g : B_2 \to A] \mapsto [g \circ f : B_2 \to B_1 \to A].$$

This example initially looks weird and different, but Examples 2.2.17 and 2.2.19 may be interpreted as special cases; do you see how? What is A in each case?

2.2.21. $\star$ **Natural transformations (and natural isomorphisms) of functors, and equivalences of categories.**

(This notion won't come up in an essential away until at least Chapter 7, so you shouldn't read this section until then.) Suppose F and G are two functors from $\mathcal{A}$ to $\mathcal{B}$. A **natural transformation of functors** $F \to G$ is the data of a morphism $m_a : F(a) \to G(a)$ for each $a \in \mathcal{A}$ such that for each $f : a \to a'$ in $\mathcal{A}$, 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. The data of functors $F : \mathcal{A} \to \mathcal{B}$ and $F' : \mathcal{B} \to \mathcal{A}$ such that $F \circ F'$ is naturally isomorphic to the identity functor $I_{\mathcal{B}}$ on $\mathcal{B}$ and $F' \circ F$ is naturally isomorphic to $I_{\mathcal{A}}$ is said to be an **equivalence of categories**. This is the "right" notion of isomorphism of categories.

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 make can 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.

2.2.C. Exercise. Let $\cdot^{\vee\vee} : f.d.Vec_k \to f.d.Vec_k$ be the double dual functor from the category of 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-dimensional hypothesis, we only get a natural transformation of functors from id to $\cdot^{\vee\vee}$.)

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

2.2.D. Exercise. Show that $\mathcal{V} \to f.d.Vec_k$ gives an equivalence of categories, by describing an "inverse" functor. (We are assuming any needed version of the axiom of choice, §1.2, so feel free to simultaneously choose bases for each vector space in *f.d.Vec*$_k$.)

2.2.22. ⋆⋆ *Aside for experts.* One may show that this definition is equivalent to another one commonly given: a covariant functor $F : A \to B$ is an equivalence of categories if it is fully faithful and every object of B is isomorphic to an object of the form $F(a)$ (F is *essentially surjective*). One can show that such a functor has a *quasiinverse*, i.e., that there is a functor $G : B \to A$, which is also an equivalence, and for which there exist natural isomorphisms $G(F(A)) \cong A$ and $F(G(B)) \cong B$. Thus "equivalence of categories" is an equivalence relation. The notion of "equivalence of categories" is the right notion of what one thinks of as "isomorphism of categories" (I informally think of it as "essentially the same category"), but the reason for this would take too long to go into here.

2.3 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 construction more appealing, and use them more often in proofs, but as they become more experienced, find universal property arguments more elegant and insightful.

We have seen one important example of a universal property argument already in §2.1: 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.

2.3.1. Localization. A second example of 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 define $(a_1/s_1) \times (a_2/s_2) = (a_1a_2)/(s_1s_2)$, and $(a_1/s_1) + (a_2/s_2) = (s_2a_1 + s_1a_2)/(s_1s_2)$. We say that $a_1/s_1 = a_2/s_2$ if *for some* $s \in S$, $s(s_2a_1 - s_1a_2) = 0$. (This implies that $S^{-1}A$ is 0 if $0 \in S$.) We have a canonical map $A \to S^{-1}A$ given by $a \mapsto a/1$. If $S = \{f^n : n \in \mathbb{Z}^{\geq 0}\}$, where $f \in A$, we define $A_f := S^{-1}A$.

2.3.A. EXERCISE. Verify that $S^{-1}A$ satisfies the following universal property: $S^{-1}A$ is initial among A-algebras B where every element of S is sent to a unit in B. (Recall: the data of "an A-algebra B" and "a ring map $A \to B$" the the same.)

Warning: sometimes localization is first introduced in the special case where A is an integral domain. In that case, $A \hookrightarrow S^{-1}A$, but this isn't always true, as shown by the following result.

2.3.B. EXERCISE. Show that $A \to S^{-1}A$ is injective if and only if S contains no zero-divisors. (A **zero-divisor** of a ring A is an element a such that there is a non-zero element b with $ab = 0$. The other elements of A are called **non-zero-divisors**. For example, a unit is never a zero-divisor. Counter-intuitively, 0 is a zero-divisor in a ring A if and only if A is not the 0-ring.)

In fact, it is cleaner to *define* $S^{-1}A$ by this 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. Define $S^{-1}M$ as being initial among A-modules N for which $s \times \cdot : N \to N$ is an isomorphism for all $s \in S$.

2.3.C. EXERCISE. Show that if $S^{-1}M$ exists, then (a) there is a natural map $M \to S^{-1}M$ (here "natural" is meant informally as "obvious" — you know it if you see it), and (b) the A-module structure on $S^{-1}M$ extends to an $S^{-1}A$-module structure.

2.3.D. EXERCISE. Show that $S^{-1}M$ 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 $(a_1/s_1) \circ (m_2/s_2) = (a_1m_2)/(s_1s_2)$.

2.3.E. EXERCISE. Show that localization commutes with finite products. In other words, if $M_1, \dots, M_n$ are A-modules, describe an isomorphism $S^{-1}(M_1 \times \cdots \times M_n) \to S^{-1}M_1 \times \cdots \times S^{-1}M_n$.

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

$$\begin{aligned} \otimes_A : \quad & \operatorname{obj}(Mod_A) \times \operatorname{obj}(Mod_A) \longrightarrow \operatorname{obj}(Mod_A) \\ & (M, N) \longmapsto M \otimes_A N \end{aligned}$$

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

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

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

2.3.G. 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 $f : M \to M''$ is surjective, and M' surjects onto the kernel of f; see §2.6), 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 2.6.F, but you may get a lot out of doing it now.) (You will be reminded of the definition of right-exactness in §2.6.4.)

The constructive definition $\otimes$ 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 $f : M \times N \to P$ 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 $M \times N \to C$ factors through the tensor product uniquely: $M \times N \to M \otimes_A N \to C$. (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 $t : M \times N \to T$, such that given any A-bilinear map $t' : M \times N \to T'$, there is a unique A-linear map $f : T \to T'$ such that $t' = f \circ t$.

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

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

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

Note that this argument shows uniqueness *assuming existence*. We need to still show the existence of such a tensor product. This forces us to do something constructive. Fortunately, we already have:

2.3.I. EXERCISE. Show that the construction of §2.3.2 satisfies the universal property of tensor product.

The uniqueness of tensor product is our second example of the proof of uniqueness (up to unique isomorphism) by a *universal property*. If you have never seen this sort of argument before, then you might think you get it, but you should think over it some more. We will be using such arguments repeatedly in the future.

The two exercises below are some useful facts about tensor products with which you should be familiar. The first exercise deals with localization, so here is a brief introduction in case you haven't seen it before.

2.3.J. IMPORTANT EXERCISE. (a) If M is an A-module and $A \to B$ is a morphism of rings, show that $B \otimes_A M$ naturally has the structure of a B-module. (In fact this describes a functor $Mod_A \to Mod_B$, but you needn't show this unless you want to.)
(b) If further $A \to C$ is a morphism of rings, show that $B \otimes_A C$ has the 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 2.3.Y will interpret this construction as a coproduct.)

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

Here is another exercise involving a universal property.

2.3.3. Definition. An object of a category $\mathcal{C}$ 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.

2.3.L. EXERCISE. Show that any two initial objects are uniquely isomorphic. Show that any two final objects are uniquely isomorphic.

This (partially) justifies the phrase "*the* initial object" rather than "*an* initial object", and similarly for "*the* final object" and "*the* zero object".

2.3.M. EXERCISE. What are the initial and final objects are in *Sets*, *Rings*, and *Top* (if they exist)? How about the two examples of §2.2.9?

2.3.N. $\star$ EXERCISE. Prove Yoneda's Lemma.

2.3.4. Important Example: Fibered products. (This notion will be essential later.) Suppose we have morphisms $f : X \to z$ and $g : Y \to Z$ (in *any* category). Then the **fibered product** is an object $X \times_Z Y$ along with morphisms $\pi_X : X \times_Z Y \to X$ and $\pi_Y : X \times_Z Y \to Y$, where the two compositions $f \circ \pi_X, g \circ \pi_Y : X \times_Z Y \to Z$ agree, such that given any object W with maps to X and Y (whose compositions to Z agree), these maps factor through some unique $W \to X \times_Z Y$:

$$\begin{tikzcd}
W \arrow[dr, dotted, "\exists!"] \arrow[drr, bend left] \arrow[ddr, bend right] & & \\
& X \times_Z Y \arrow[r, "\pi_Y"'] \arrow[d, "\pi_X"] & Y \arrow[d, "g"] \\
& X \arrow[r, "f"] & Z
\end{tikzcd}$$

(Warning: the definition of the fibered product depends on f and g, 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.

Depending on your religion, the diagram

$$\begin{tikzcd} X \times_Z Y \arrow[r, "\pi_Y"'] \arrow[d, "\pi_X"] & Y \arrow[d, "g"] \\ X \arrow[r, "f"] & Z \end{tikzcd}$$

is called a **fibered/pullback/Cartesian diagram/square** (six possibilities).

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

2.3.O. EXERCISE. Show that in *Sets*,

$$X \times_Z Y = \{(x \in X, y \in Y) \ : \ f(x) = g(y)\}.$$

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.)

2.3.P. 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.)

2.3.Q. EXERCISE. If Z is the final object in a category $\mathcal{C}$, and $X, Y \in \mathcal{C}$, show that "$X \times_Z Y = X \times Y$": "the" fibered product over Z is uniquely isomorphic to "the" product. (This is an exercise about unwinding the definition.)

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

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

2.3.S. EXERCISE. Given $X \to Y \to Z$, show that there is a natural morphism $X \times_Y X \to X \times_Z X$, 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.)

2.3.T. USEFUL EXERCISE: THE MAGIC DIAGRAM. Suppose we are given morphisms $X_1, X_2 \to Y$ and $Y \to Z$. Describe the natural morphism $X_1 \times_Y X_2 \to$

$X_1 \times_Z X_2$. Show that the following diagram is a fibered square.

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

This diagram is surprisingly incredibly useful — so useful that we will call it the **magic diagram**.

2.3.5. Monomorphisms and epimorphisms.

2.3.6. Definition. A morphism $f : X \to Y$ is a **monomorphism** if any two morphisms $g_1, g_2 : Z \to X$ such that $f \circ g_1 = f \circ g_2$ must satisfy $g_1 = g_2$. In other words, for any other object Z, the natural map $\mathrm{Hom}(Z, X) \to \mathrm{Hom}(Z, Y)$ is an injection. This a generalization of an injection of sets. In other words, there is a unique way of filling in the dotted arrow so that the following diagram commutes.

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

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". This is also the case with "epimorphism" vs. "surjective".)

2.3.U. EXERCISE. Show that the composition of two monomorphisms is a monomorphism.

2.3.V. EXERCISE. Prove a morphism $X \to Y$ is a monomorphism if and only if the induced morphism $X \to X \times_Y X$ 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.)

2.3.W. EXERCISE. Suppose $Y \to Z$ is a monomorphism, and $X_1, X_2 \to Y$ are two morphisms. Show that $X_1 \times_Y X_2$ and $X_1 \times_Z X_2$ are canonically isomorphic. We will use this later when talking about fibered products. (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 magic diagram, Exercise 2.3.T)

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 is necessary for the definition of an abelian category. Intuitively, it is the categorical version of a surjective map.

2.3.7. Coproducts.

2.3.X. EXERCISE. Define **coproduct** in a category by reversing all the arrows in the definition of product. Show that coproduct for *Sets* is disjoint union. (This is why we use the notation $\coprod$ for disjoint union.)

2.3.Y. EXERCISE. Suppose $A \to B, C$ are two ring morphisms, so in particular B and C are A-modules. Recall (Exercise 2.3.J) 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 2.3.F.) 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.

2.3.Z. ⋆ EXERCISE (REPRESENTABLE FUNCTORS). Much of our discussion about universal properties can be cleanly expressed in terms of representable functors.
(a) Suppose A and B are objects in a category $\mathcal{C}$. Give a bijection between the natural transformations $h^A \to h^B$ of covariant functors $\mathcal{C} \to$ *Sets* (see Exercise 2.2.14 for the definition) and the morphisms $B \to A$.
(b) State the corresponding fact for contravariant functors h_A (see Exercise 2.2.20). Remark: a contravariant functor F from $\mathcal{C}$ to sets is said to be *representable* if there is a natural isomorphism $\xi : F \xrightarrow{\sim} h_A$. This exercise shows that 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.B.)
(c) **Yoneda's lemma.** Suppose F is a covariant functor $\mathcal{C} \to$ *Sets*, and $A \in \mathcal{C}$. Give a bijection between the natural transformations $h^A \to F$ and $F(A)$. State the corresponding fact for contravariant functors.

2.4 Limits and colimits

Limits and colimits provide two important examples defined by universal properties. They generalize a number of familiar constructions. I'll give the definition first, and then show you why it is familiar. For example, fractions will be motivating examples of colimits (Exercise 2.4.B(a)), and the p-adic numbers (Example 2.4.3) will be motivating examples of limits.

2.4.1. Limits. We say that a category is a **small category** if the objects and the morphisms are sets. (This is a technical condition intended only for experts.) Suppose $\mathcal{I}$ is any small category, and $\mathcal{C}$ is any category. Then a functor $F : \mathcal{I} \to \mathcal{C}$ (i.e. with an object $A_i \in \mathcal{C}$ for each element $i \in \mathcal{I}$, and appropriate commuting morphisms dictated by $\mathcal{I}$) is said to be a **diagram indexed by** $\mathcal{I}$. We call $\mathcal{I}$ an **index category**. Our index categories will be partially ordered sets (Example 2.2.8), 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 $\mathcal{A}$ is a category, then a functor $\square \to \mathcal{A}$ is precisely the data of a commuting square in $\mathcal{A}$.

Then the **limit** is an object $\varprojlim_{\mathcal{I}} A_i$ of $\mathcal{C}$ along with morphisms $f_j : \varprojlim_{\mathcal{I}} A_i \to A_j$ such that if $m : j \to k$ is a morphism in $\mathcal{I}$, then

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

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

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

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

we obtain the fibered product.

If $\mathcal{I}$ is

$$\bullet \qquad \bullet$$

we obtain the product.

If $\mathcal{I}$ 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 $\mathcal{I}$ has two elements is the example of the previous paragraph.

If $\mathcal{I}$ has an initial object e, then A_e is the limit, and in particular the limit always exists.

2.4.3. *Example: the* p*-adic numbers.* The p-adic numbers, $\mathbb{Z}_p$, are often described informally (and somewhat unnaturally) as being of the form $\mathbb{Z}_p = ? + ?p + ?p^2 + ?p^3 + \cdots$. They are an example of a limit in the category of rings:

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

Limits do not always exist for any index category $\mathcal{I}$. 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).

2.4.A. IMPORTANT EXERCISE. Show that in the category *Sets*,

$$\left\{ (a_i)_{i\in I} \in \prod_i A_i : F(m)(a_i) = a_j \text{ for all } m \in \mathrm{Mor}_{\mathcal{I}}(i,j) \in \mathrm{Mor}(\mathcal{I}) \right\},$$

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

This clearly also works in the category Mod_A of A-modules, and its specializations such as Vec_k and Ab.

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

2.4.4. Colimits. More immediately relevant for us will be the dual (arrow-reversed version) of the notion of limit (or inverse limit). We just flip all the arrows in that definition, and get the notion of a **colimit**. 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 $\mathcal{I}$; and a colimit has a map *from* all the objects.)

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

2.4.5. *Example.* The ring $5^{-\infty}\mathbb{Z}$ of rational numbers whose denominators are powers of 5 is a colimit $\varinjlim 5^{-i}\mathbb{Z}$. More precisely, $5^{\infty}\mathbb{Z}$ is the colimit of

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

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.

2.4.B. EXERCISE. (a) Interpret the statement "$\mathbb{Q} = \varinjlim \frac{1}{n}\mathbb{Z}$". (b) Interpret the union of the 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 don't always exist, but there are two useful large classes of examples for which they do.

2.4.6. Definition. A 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, a category $\mathcal{I}$ is **filtered** if:

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

2.4.C. EXERCISE. Suppose $\mathcal{I}$ is filtered. (We will be almost exclusively using the case where $\mathcal{I}$ is a filtered set.) Show that any diagram in *Sets* indexed by $\mathcal{I}$ has the following as a colimit:

$$\left\{ a \in \coprod_{i \in \mathcal{I}} A_i \right\} \Big/ (a_i \in A_i) \sim f(a_i) \in A_j \text{ for every } f : A_i \to A_j \text{ in the diagram.}$$

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, in Example 2.4.5, 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 example, $17/125$ is an element of the $5^{-3}\mathbb{Z}$ (or $5^{-4}\mathbb{Z}$, or later ones), but not $5^{-2}\mathbb{Z}$. More generally, in the category of A-modules Mod_A, each element a of the colimit $\varinjlim A_i$ can be interpreted as an element of *some* $a \in A_i$. The element $a \in \varinjlim A_i$ is 0 if there is some $m : i \to j$ such that $F(m)(a) = 0$ (i.e. if it becomes 0 "later in the diagram"). Furthermore, two elements interpreted as $a_i \in A_i$ and $a_j \in A_j$ are the same if there are some arrows $m : i \to k$ and $n : j \to k$ such that $F(m)(a_i) = F(n)(a_j)$, i.e. if they become the same "later in the diagram". To add two elements interpreted as $a_i \in A_i$ and $a_j \in A_j$, we choose arrows $m : i \to k$ and $n : j \to k$, and then interpret their sum as $F(m)(a_i) + F(n)(a_j)$.

2.4.D. EXERCISE. Verify that the A-module described above is indeed the colimit.

2.4.E. USEFUL EXERCISE (LOCALIZATION AS COLIMIT). Generalize Exercise 2.4.B(a) to interpret localization of a ring as a colimit over a filtered set: suppose S is a multiplicative set of A, and interpret $S^{-1}A = \varinjlim \frac{1}{s}A$ where the limit is over $s \in S$.

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:

2.4.F. EXERCISE: COLIMITS OF A-MODULES WITHOUT THE FILTERED CONDITION. Suppose you are given a diagram of A-modules indexed by $\mathcal{I}$: $F : \mathcal{I} \to Mod_A$, where we let $A_i := F(i)$. Show that the colimit is $\oplus_{i \in \mathcal{I}} A_i$ modulo the relations $a_j - F(m)(a_i)$ for every $m : i \to j$ in $\mathcal{I}$ (i.e. for every arrow in the diagram).

The following exercise shows that you have to be careful to remember which category you are working in.

2.4.G. UNIMPORTANT EXERCISE. Consider the filtered set of abelian groups $p^{-n}\mathbb{Z}_p/\mathbb{Z}_p$. Show that this system has colimit $\mathbb{Q}_p/\mathbb{Z}_p$ in the category of abelian groups, and the colimit 0 in the category of finite abelian groups. Here $\mathbb{Q}_p$ is the fraction field of $\mathbb{Z}_p$, which can be interpreted as $\cup p^{-n}\mathbb{Z}_p$.

2.4.7. 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 an element in each object in the diagram, that are "compatible" (Exercise 2.4.A). And an element of a colimit can be thought of ("has a representative that is") an element of a single object in the diagram (Exercise 2.4.C). Even though the definitions of limit and colimit are the same, just with arrows reversed, these interpretations are quite different.

2.5 Adjoints

We next come to an very useful construction 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 $F : \mathcal{A} \to \mathcal{B}$ and $G : \mathcal{B} \to \mathcal{A}$ are **adjoint** if there is a natural bijection for all $A \in \mathcal{A}$ and $B \in \mathcal{B}$

$$\tau_{AB} : \mathrm{Mor}_{\mathcal{B}}(F(A), B) \to \mathrm{Mor}_{\mathcal{A}}(A, G(B)). \tag{2.5.0.1}$$

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

$$\begin{tikzcd} \mathrm{Mor}_{\mathcal{B}}(F(A'), B) \arrow[r, "Ff^*"] \arrow[d, "\tau_{A'B}"] & \mathrm{Mor}_{\mathcal{B}}(F(A), B) \arrow[d, "\tau_{AB}"] \\ \mathrm{Mor}_{\mathcal{A}}(A', G(B)) \arrow[r, "f^*"] & \mathrm{Mor}_{\mathcal{A}}(A, G(B)) \end{tikzcd} \tag{2.5.0.2}$$

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

2.5.A. EXERCISE. Write down what this diagram should be. (Hint: do it by extending diagram (2.5.0.2) above.)

2.5.B. EXERCISE. Show that the map τ_{AB} (2.5.0.1) is given as follows. For each A there is a map $\eta_A : A \to GF(A)$ so that for any $g : F(a) \to B$, the corresponding $f : A \to G(B)$ is given by the composition

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

Similarly, there is a map $\epsilon_B : B \to FG(B)$ for each B so that for any $f : A \to G(B)$, the corresponding map $g : F(A) \to B$ is given by the composition

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

Here is an example of an adjoint pair.

2.5.C. EXERCISE. Suppose M, N, and P are A-modules. Describe a bijection $\mathrm{Mor}_A(M \otimes_A N, P) \leftrightarrow \mathrm{Mor}_A(M, \mathrm{Mor}_A(N, P))$. (Hint: try to use the universal property.)

2.5.D. EXERCISE. Show that $\cdot \otimes_A N$ and $\mathrm{Mor}_A(N, \cdot)$ are adjoint functors.

2.5.1. ⋆ *Fancier remarks we won't use.* You can check that the left adjoint determines the right adjoint up to unique natural isomorphism, and vice versa, by a universal property argument. The maps η_A and ϵ_B of Exercise 2.5.B are called the **unit** and **counit** of the adjunction. This leads to a different characterization of adjunction. Suppose functors $F : \mathcal{A} \to \mathcal{B}$ and $G : \mathcal{B} \to \mathcal{A}$ are given, along with natural transformations $\epsilon : FG \to \mathrm{id}$ and $\eta : \mathrm{id} \to GF$ with the property that $G\epsilon \circ \eta G = \mathrm{id}_G$ (for each $B \in \mathcal{G}$, the composition of $\eta_{G(B)} : G(B) \to GFG(B)$ and $G(\eta_B) : GFG(B) \to G(B)$ is the identity) and $\eta F \circ F\epsilon = \mathrm{id}_F$. 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.

2.5.2. *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.

2.5.3. Example: groupification. Here is another motivating example: getting an abelian group from an abelian semigroup. An abelian semigroup is just like an abelian group, except you don't require an inverse. One example is the non-negative integers $0, 1, 2, \dots$ under addition. Another is the positive integers under multiplication $1, 2, \dots$. From an abelian semigroup, you can create an abelian group. Here is a formalization of that notion. If S is a semigroup, then its **groupification** is a map of semigroups $\pi : S \to G$ such that G is a group, and any other map of semigroups from S to a *group* G' factors *uniquely* through G.

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

2.5.E. EXERCISE. Construct groupification H from the category of abelian semigroups to the category of abelian groups. (One possibility of a construction: given an abelian semigroup S, the elements of its groupification $H(S)$ are (a, b), 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 semigroup map $S \to H(S)$.) Let F be the forgetful morphism 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; this is called a "reflective subcategory". 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.)

2.5.F. EXERCISE. Show that if a semigroup is *already* a group then groupification is the identity morphism, by the universal property.

2.5.G. EXERCISE. The purpose of this exercise is to give you some practice with "adjoints of forgetful functors", the means by which we get groups from semigroups, and sheaves from presheaves. Suppose A is a ring, and S is a multiplicative subset. Then $S^{-1}A$-modules are a fully faithful subcategory of the category of A-modules (meaning: the objects of the first category are a subset of the objects of the second; and the morphisms between any two objects of the second that are secretly objects of the first are just the morphisms from the first). Then $M \to S^{-1}M$ satisfies a universal property. Figure out what the universal property is, and check that it holds. In other words, describe the universal property enjoyed by $M \to S^{-1}M$, and prove that it holds.

(Here is the larger story. Every $S^{-1}A$-module is an A-module, and this is an injective map, so we have a covariant forgetful functor $F : Mod_{S^{-1}A} \to Mod_A$. 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 $G : Mod_A \to Mod_{S^{-1}A}$, 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).)

Here is a table of adjoints that will come up for us.

situation	category $\mathcal{A}$	category $\mathcal{B}$	left-adjoint $F : \mathcal{A} \to \mathcal{B}$	right-adjoint $G : \mathcal{B} \to \mathcal{A}$
A-modules			$\cdot \otimes_A N$	$\mathrm{Hom}_A(N, \cdot)$
ring maps $A \to B$	Mod_A	Mod_B	$\cdot_A B$ (extension of scalars)	forgetful (restriction of scalars)
(pre)sheaves on a topological space X	presheaves on X	sheaves on X	sheafification	forgetful
(semi)groups	semigroups	groups	groupification	forgetful
sheaves, $f : X \to Y$	sheaves on Y	sheaves on X	f^{-1}	f_*
sheaves of abelian groups or $\mathcal{O}$-modules, open immersions $f : U \hookrightarrow Y$	sheaves on U	sheaves on Y	$f_!$	f^{-1}
quasicoherent sheaves, $f : X \to Y$	quasicoherent sheaves on Y	quasicoherent sheaves on X	f^*	f_*

2.5.4. *Useful comment for experts.* One last comment 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 §2.6.10.

2.6 Kernels, cokernels, and exact sequences: A brief introduction to abelian categories

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. Each of these notions generalizes the previous one.

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 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 these notes 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.

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

- Ad1. For each $A, B \in \mathcal{C}$, $\mathrm{Mor}(A, B)$ 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. $\mathcal{C}$ has a zero object, denoted 0. (This is an object that is simultaneously an initial object and a final object, Defn. 2.3.3.)
- 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**.

2.6.2. *Remarks.* It is a consequence of the definition of additive category that finite direct products are also finite direct sums (coproducts) — 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 a is a 0-object if and only if $\mathrm{id}_a = 0_a$; additive functors preserve both id and 0), and preserves products.

One motivation for the name 0-object is that the 0-morphism in the abelian group $\mathrm{Hom}(A, B)$ is the composition $A \to 0 \to B$.

Real (or complex) Banach spaces are an example of an additive category. The category of free A-modules is another. The category Mod_A of A-modules is also an example, but it has even more structure, which we now formalize as an example of an abelian category.

2.6.3. Definition. Let $\mathcal{C}$ be an additive category. A **kernel** of a morphism $f : B \to C$ is a map $i : A \to B$ 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.) Hence it is unique up to unique isomorphism by universal property nonsense. A **cokernel** is defined dually by reversing the arrows — do this yourself. The kernel of $f : B \to C$ is the limit (§2.4) of the diagram

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

and similarly the cokernel is a colimit.

A morphism $i : A \to B$ in $\mathcal{C}$ is **monic** if for all $g : C \to A$ such that $i \circ g = 0$, we have $g = 0$:

$$\begin{tikzcd} C \arrow[d, "\therefore g=0"'] \arrow[dr, "0"] & \\ A \arrow[r, "i"] & B \end{tikzcd}$$

(Once we know what an abelian category is, you may check that a monic morphism in an abelian category is a monomorphism.) If $i : A \to B$ is monic, then we say that A is a **subobject** of B, where the map i is implicit. Dually, there is the notion of **epi** — reverse the arrows to find out what that is. The notion of **quotient object** is 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 monic morphism is the kernel of its cokernel.
(3) Every epi morphism is the cokernel of its kernel.

It is a non-obvious (and imprecisely stated) fact that every property you want to be true about kernels, cokernels, etc. follows from these three.

The **image** of a morphism $f : A \to B$ is defined as $\text{im}(f) = \ker(\operatorname{coker} f)$. It is the unique factorization

$$A \xrightarrow{\text{epi}} \text{im}(f) \xrightarrow{\text{monic}} B$$

It is the cokernel of the kernel, and the kernel of the cokernel. The reader may want to verify this as an exercise. It is unique up to unique isomorphism.

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 ring Mod_A, you can manipulate objects in any abelian category. This is made precise by Freyd-Mitchell Embedding Theorem. (The Freyd-Mitchell Embedding Theorem: If $\mathcal{A}$ is an abelian category such that

$\mathrm{Hom}(a, a')$ is a set for all $a, a' \in \mathcal{A}$, then there is a ring A and an exact, full faithful functor from $\mathcal{A}$ into Mod_A, which embeds $\mathcal{A}$ as a full subcategory. A proof is sketched in [**W**, §1.6], and references to a complete proof are given there. The moral is that to prove something about a diagram in some abelian category, we may pretend 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.) However, the abelian categories we'll 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'll show that sheaves of abelian groups on a topological space X form an abelian category (§3.5), 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.

2.6.4. Complexes, exactness, and homology.

We say a sequence

$$A \xrightarrow{f} B \xrightarrow{g} C \tag{2.6.4.1}$$

is a **complex** if $g \circ f = 0$, and is **exact** if $\ker g = \operatorname{im} f$. An exact sequence with five terms, the first and last of which are 0, is a **short exact sequence**. Note that $A \xrightarrow{f} B \longrightarrow C \longrightarrow 0$ being exact is equivalent to describing C as a cokernel of f (with a similar statement for $0 \longrightarrow A \longrightarrow B \xrightarrow{g} C$).

If (2.6.4.1) is a complex, then its **homology** (often denoted H) is $\ker g / \operatorname{im} f$. We say that the $\ker g$ are the **cycles**, and $\operatorname{im} f$ 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

$$A^\bullet : \qquad \cdots \longrightarrow A^{i-1} \xrightarrow{f^{i-1}} A^i \xrightarrow{f^i} A^{i+1} \xrightarrow{f^{i+1}} \cdots \tag{2.6.4.2}$$

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

$$0 \longrightarrow \ker f^i \longrightarrow A^i \longrightarrow \operatorname{im} f^i \longrightarrow 0 \tag{2.6.4.3}$$

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

2.6.A. EXERCISE. Describe exact sequences

$$0 \longrightarrow \operatorname{im} f^i \longrightarrow A^{i+1} \longrightarrow \operatorname{coker} f^i \longrightarrow 0 \tag{2.6.4.4}$$

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

(These are somehow dual to (2.6.4.3). In fact in some mirror universe this might have been given as the standard definition of homology.)

2.6.B. EXERCISE. 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). Show that $\sum(-1)^i \dim A^i = \sum(-1)^i h^i(A^\bullet)$. (Recall that $h^i(A^\bullet) = \dim \ker(d^i)/\operatorname{im}(d^{i-1})$.) 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.)

2.6.C. IMPORTANT EXERCISE. Suppose $\mathcal{C}$ is an abelian category. Define the category *Com*$_{\mathcal{C}}$ 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 $\mathcal{C}$, and the morphisms $A^\bullet \to B^\bullet$ are commuting diagrams

$$\begin{tikzcd} A^\bullet : \arrow[d] & \cdots \arrow[r] & A^{i-1} \arrow[r, "f^{i-1}"] \arrow[d] & A^i \arrow[r, "f^i"] \arrow[d] & A^{i+1} \arrow[r, "f^{i+1}"] \arrow[d] & \cdots \\ B^\bullet : & \cdots \arrow[r] & B^{i-1} \arrow[r, "g^{i-1}"] & B^i \arrow[r, "g^i"] & B^{i+1} \arrow[r, "g^{i+1}"] & \cdots \end{tikzcd} \tag{2.6.4.5}$$

Show that *Com*$_{\mathcal{C}}$ is an abelian category. (Feel free to deal with the special case *Mod*$_A$.)

2.6.D. IMPORTANT EXERCISE. Show that (2.6.4.5) induces a map of homology $H(A^i) \to H(B^i)$. (Again, feel free to deal with the special case *Mod*$_A$.)

We will later define when two maps of complexes are *homotopic* (§23.1), and show that homotopic maps induce isomorphisms on cohomology (Exercise 23.1.A), but we won't need that any time soon.

2.6.5. Theorem (Long exact sequence). — *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.) For a concise proof in the case of complexes of modules, and a discussion of how to show this in general, see [**W**, §1.3]. It will also come out of our discussion of spectral sequences as well (again, in the category of modules over a ring), see Exercise 2.7.E, but this is a somewhat perverse way of proving it.

2.6.6. *Exactness of functors.* If $F : \mathcal{A} \to \mathcal{B}$ is a covariant additive 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 $\mathcal{A}$ 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

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

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

A 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**.

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

2.6.E. 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 2.6.H(c).)

2.6.F. 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 $\cdot \otimes M$ is a right-exact covariant functor $Mod_A \to Mod_A$. (This is a repeat of Exercise 2.3.G.)
(c) Show that $\mathrm{Hom}(M, \cdot)$ is a left-exact covariant functor $Mod_A \to Mod_A$.
(d) Show that $\mathrm{Hom}(\cdot, M)$ is a left-exact contravariant functor $Mod_A \to Mod_A$.

2.6.G. 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

$$A^{\oplus q} \to A^{\oplus p} \to M \to 0 \tag{2.6.6.1}$$

Use (2.6.6.1) and the left-exactness of Hom to describe an isomorphism

$$S^{-1}\mathrm{Hom}_A(M, N) \cong \mathrm{Hom}_{S^{-1}A}(S^{-1}M, S^{-1}N).$$

(You might be able to interpret this in light of a variant of Exercise 2.6.H below, for left-exact contravariant functors rather than right-exact covariant functors.)

2.6.7. $\star$ **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 it is easy to prove in an ad hoc basis. Once they are second nature to you, many subtle facts will be come obvious to you as special cases. And you will see that they will get used (implicitly or explicitly) repeatedly.

2.6.8. $\star$ *Interaction of homology and (right/left-)exact functors.*

You might wait to prove this until you learn about cohomology in Chapter 20, when it will first be used in a serious way.

2.6.H. 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, [**N**, Ex. 2.6.H]. Suppose $F : \mathcal{A} \to \mathcal{B}$ is a covariant functor of abelian categories. Suppose $C^\bullet$ is a complex in $\mathcal{A}$.

(a) *(F right-exact yields* $FH^\bullet \longrightarrow H^\bullet$ *)* 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* $FH^\bullet \longleftarrow H^\bullet$ *)* 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 $FC^\bullet$.)

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

Hint for (a): use $C^p \xrightarrow{d^p} C^{p+1} \longrightarrow \operatorname{coker} d^p \longrightarrow 0$ to give an isomorphism $F \operatorname{coker} d^p \cong \operatorname{coker} F d^p$. Then use the first line of (2.6.4.4) to give a surjection $F \operatorname{im} d^p \twoheadrightarrow \operatorname{im} F d^p$. Then use the second line of (2.6.4.4) to give the desired map $FH^p C^\bullet \longrightarrow H^p F C^\bullet$. While you are at it, you may as well describe a map for the fourth member of the quartet $\{\ker, \operatorname{coker}, \operatorname{im}, H, \}$: $F \ker d^p \longrightarrow \ker F d^p$.

2.6.9. If this makes your head spin, you may prefer to think of it in the following specific case, where both $\mathcal{A}$ and $\mathcal{B}$ are the category of A-modules, and F is $\cdot \otimes N$ for some fixed N-module. Your argument in this case will translate without change to yield a solution to Exercise 2.6.H(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, 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.

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

A surprising number of arguments boil down the statement:

Limits (e.g. kernels) commute with limits and right-adjoints. In particular, both right-adjoints and limits are left exact.

as well as its dual:

Colimits (e.g. cokernels) commute with colimits and left-adjoints. In particular, both left-adjoints and colimits are left exact.

These statements were promised in §2.5.4. The latter has a useful extension:

In an abelian category, colimits over filtered index categories are exact.

("Filtered" was defined in §2.4.6.) If you want to use these statements (for example, later in these notes), you will have to prove them. Let's now make them precise.

2.6.I. EXERCISE (KERNELS COMMUTE WITH LIMITS). Suppose $\mathcal{C}$ is an abelian category, and $a : \mathcal{I} \to \mathcal{C}$ and $b : \mathcal{I} \to \mathcal{C}$ are two diagrams in $\mathcal{C}$ indexed by $\mathcal{I}$. For convenience, let $A_i = a(i)$ and $B_i = b(i)$ be the objects in those two diagrams. Let

$h_i : A_i \to B_i$ be maps commuting with the maps in the diagram. (Translation: h is a natural transformation of functors $a \to b$, see §2.2.21.) Then the $\ker h_i$ form another diagram in $\mathcal{I}$ indexed by $\mathcal{I}$. Describe a natural isomorphism $\varprojlim \ker h_i \cong \ker(\varprojlim A_i \to \varprojlim B_i)$.

2.6.J. 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.)

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

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 $\mathcal{I}$. 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 $\mathcal{I}$. 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. $\square$

Of course, the dual statements to Exercise 2.6.J and Proposition 2.6.11 hold by the dual arguments.

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

2.6.K. EXERCISE. Show that in an abelian category, colimits over filtered index categories are exact. Right-exactness follows from the above discussion, so the issue is left-exactness. (Possible hint: After you show that localization is exact, Exercise 2.6.F(a), or sheafification is exact, Exercise 3.5.D, in a hands on way, you will be easily able to prove this. Conversely, this exercise will quickly imply those two.)

2.6.L. EXERCISE. Show that filtered colimits commute with homology. Hint: use the FHHF Theorem (Exercise 2.6.H), and the previous Exercise.

2.6.12. ⋆ **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

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

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

$$0 \to FM' \to FM \to FM''$$

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.

2.7 ⋆ Spectral sequences

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 spectres, 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 sequence 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 (which is 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 [**W**, Ch. 5] for more detailed information if you wish.

You should *not* read this section when you are reading the rest of Chapter 2. 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.

For concreteness, we work in the category Vec_k of vector spaces over a field k. However, everything we say will apply in any abelian category, such as the category Mod_A of A-modules.

2.7.1. Double complexes.

A **double complex** is a collection of vector spaces $E^{p,q}$ ($p, q \in \mathbb{Z}$), and "rightward" morphisms $d_{\rightarrow}^{p,q} : E^{p,q} \to E^{p,q+1}$ and "upward" morphisms $d_{\uparrow}^{p,q} : E^{p,q} \to E^{p+1,q}$. In the superscript, the first entry denotes the row number, and the second entry denotes the column number, in keeping with the convention for matrices, but opposite to how the (x, y)-plane is labeled. 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 $d_{\uparrow}$ satisfying (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 $(-1)^q d_{\uparrow}^{p,q}$. 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 $-f$.)

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

There are variations on this definition, where for example the vertical arrows go downwards, or some different subset of the $E^{p,q}$ are required to be zero, but I will leave these straightforward variations to you.

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 sufficent to prove lots of things.

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

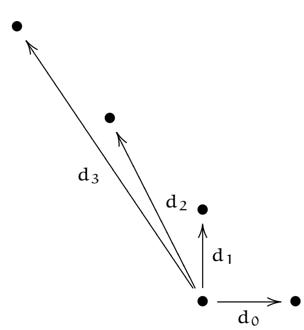
$${}_{\rightarrow}d_r^{p,q} : {}_{\rightarrow}E_r^{p,q} \rightarrow {}_{\rightarrow}E^{p+r,q-r+1}$$

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

(2.7.2.1)



(the morphisms each apply to different pages). Notice that the map always is "degree 1" in the grading of the single complex $E^{\bullet}$.

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 0th page $E_0^{p,q} = E^{p,q}$. In particular, if $E^{p,q} = 0$, then $E_r^{p,q} = 0$ for all r, so $E_r^{p,q} = 0$ unless $p, q \in \mathbb{Z}^{\geq 0}$.

Suppose now that $E^{\bullet,\bullet}$ is a **first quadrant double complex**, i.e. $E^{p,q} = 0$ for $p < 0$ or $q < 0$. 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

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

and thus we have canonical isomorphisms

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

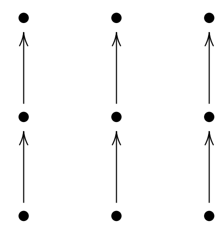
We denote this module $E_\infty^{p,q}$. The same idea works in other circumstances, for example if the double complex is only nonzero in a finite number of rows — $E^{p,q} = 0$ unless $p_0 < p < p_q$. This will come up for example in the long exact sequence and mapping cone discussion (Exercises 2.7.E and 2.7.F 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:

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

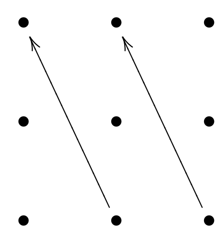
and so E_1 is just the table of cohomologies of the rows. You should check that there are now vertical maps $d_1^{p,q} : E_1^{p,q} \to E_1^{p+1,q}$ of the row cohomology groups, induced by $d_{\uparrow}$, and that these make the columns into complexes. (This is essentially the fact that a map of complexes 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 2.7.5 or in the long exact sequence in cohomology arising from a short exact sequence of complexes, Exercise 2.6.C. 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

$$E_\infty^{0,k} \xhookrightarrow{E_\infty^{1,k-1}} \ ? \ \xhookrightarrow{E_\infty^{2,k-2}} \cdots \xhookrightarrow{E^{0,k}} H^k(E^\bullet) \tag{2.7.2.2}$$

where the quotients are displayed above each inclusion. (I always forget which way the quotients are supposed to go, i.e. whether $E^{k,0}$ or $E^{0,k}$ is the subobject. One way of remembering it is by having some idea of how the result is proved.)

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 $E_r^{i,k-i}$ has precisely one non-zero 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_∞.

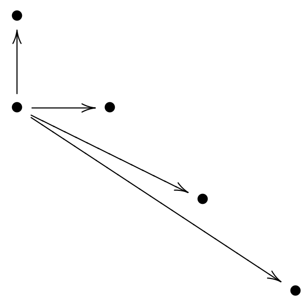
2.7.A. EXERCISE: INFORMATION FROM THE SECOND PAGE. Show that $H^0(E^\bullet) = E_\infty^{0,0} = E_2^{0,0}$ and

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

2.7.3. *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 (2.7.2.1)).

(2.7.3.1)



This spectral sequence is denoted ${}_{\uparrow}E_{\bullet}^{\bullet,\bullet}$ ("with the upwards 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 in the opposite order to that of (2.7.2.2) 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.

2.7.4. 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.

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

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

where the rows are exact in the middle (at B, C, D, G, H, I) 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

$$0 \to \ker\alpha \to \ker\beta \to \ker\gamma \to \operatorname{coker}\alpha \to \operatorname{coker}\beta \to \operatorname{coker}\gamma \to 0. \quad (2.7.5.1)$$

We plug this into our spectral sequence machinery. We first compute the cohomology using the rightwards orientation, i.e. using the order (2.7.2.1). 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 upwards 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-quesion-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 (2.7.5.1), 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?

2.7.B. UNIMPORTANT EXERCISE (GRAFTING EXACT SEQUENCES, A WEAKER VERSION 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

$$\cdots \longrightarrow W \longrightarrow \ker a \longrightarrow \ker b \longrightarrow \ker c$$

$$\longrightarrow \operatorname{coker} a \longrightarrow \operatorname{coker} b \longrightarrow \operatorname{coker} c \longrightarrow A' \longrightarrow \cdots .$$

2.7.6. *Example: the Five Lemma.* Suppose

(2.7.6.1)
$$\begin{tikzcd} F \arrow[r] & G \arrow[r] & H \arrow[r] & I \arrow[r] & J \\ A \arrow[r] \arrow[u, "\alpha"] & B \arrow[r] \arrow[u, "\beta"] & C \arrow[r] \arrow[u, "\gamma"] & D \arrow[r] \arrow[u, "\delta"] & E \arrow[u, "\epsilon"] \end{tikzcd}$$

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 rightwards orientation (2.7.2.1). 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 non-zero 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 upwards orientation (2.7.3.1). Then ${}_{\uparrow}E_1$ looks like this:

$$0 \longrightarrow 0 \longrightarrow ? \longrightarrow 0 \longrightarrow 0$$

$$0 \longrightarrow 0 \longrightarrow ? \longrightarrow 0 \longrightarrow 0$$

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're 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!

2.7.C. EXERCISE: THE SUBTLE 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.)

2.7.D. EXERCISE. If β and δ (in (2.7.6.1)) are injective, and α is surjective, show that γ is injective. Give the dual statement (whose proof is of course essentially the same).

2.7.E. EXERCISE. Use spectral sequences to show that a short exact sequence of complexes gives a long exact sequence in cohomology (Exercise 2.6.C).

2.7.F. EXERCISE (THE MAPPING CONE). Suppose $\mu : A^\bullet \to B^\bullet$ 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 20.2.4.)

The Grothendieck (or composition of functor) spectral sequence (Exercise 23.3.D) 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!

2.7.7. $\star\star$ **Complete definition of the spectral sequence, and proof.**

You should most definitely not read this section any time soon after reading the introduction to spectral sequences above. Instead, flip quickly through it to convince yourself that nothing fancy is involved.

We consider the rightwards orientation. The upwards orientation is of course a trivial variation of this.

2.7.8. *Goals.* We wish to describe the pages and differentials of the spectral sequence explicitly, and prove that they behave the way we said they did. More precisely, we wish to:

(a) describe $E_r^{p,q}$,
(b) verify that $H^k(E^\bullet)$ is filtered by $E_\infty^{p,k-p}$ as in (2.7.2.2),
(c) describe d_r and verify that $d_r^2 = 0$, and
(d) verify that $E_{r+1}^{p,q}$ is given by cohomology using d_r.

Before tacking these goals, you can impress your friends by giving this short description of the pages and differentials of the spectral sequence. We say that an element of $E^{\bullet,\bullet}$ is a (p,q)-*strip* if it is an element of $\oplus_{l\geq 0} E^{p+l,q-l}$ (see Fig. 2.1). Its non-zero entries lie on a semi-infinite antidiagonal starting with position (p,q). We say that the (p,q)-entry (the projection to $E^{p,q}$) is the *leading term* of the (p,q)-strip. Let $\boxed{S^{p,q}} \subset E^{\bullet,\bullet}$ be the submodule of all the (p,q)-strips. Clearly $S^{p,q} \subset E^{p+q}$, and $S^{0,k} = E^k$.

Note that the differential $d = d_\uparrow + d_\rightarrow$ sends a (p,q)-strip x to a $(p,q+1)$-strip dx. If dx is furthermore a $(p+r, q+r+1)$-strip ($r \in \mathbb{Z}^{\geq 0}$), we say that x is an r-*closed* (p,q)-*strip*. We denote the set of such $\boxed{S_r^{p,q}}$ (so for example $S_0^{p,q} = S^{p,q}$,

$$\begin{array}{ccccc}
\ddots & 0 & 0 & 0 & 0 \\
0 & *^{p+2,q-2} & 0 & 0 & 0 \\
0 & 0 & *^{p+1,q-1} & 0 & 0 \\
0 & 0 & 0 & *^{p,q} & 0 \\
0 & 0 & 0 & 0 & 0^{p-1,q+1}
\end{array}$$

FIGURE 2.1. A (p,q)-strip (in $S^{p,q} \subset E^{p+q}$). Clearly $S^{0,k} = E^k$.

and $S_0^{0,k} = E^k$). An element of $S_r^{p,q}$ may be depicted as:

$$\begin{tikzcd}
\ddots \arrow[r] & ? & & & \\
& *^{p+2,q-2} \arrow[u] \arrow[r] & 0 & & \\
& & *^{p+1,q-1} \arrow[u] \arrow[r] & 0 & \\
& & & *^{p,q} \arrow[u] \arrow[r] & 0
\end{tikzcd}$$

2.7.9. *Preliminary definition of* $E_r^{p,q}$. We are now ready to give a first definition of $E_r^{p,q}$, which by construction should be a subquotient of $E^{p,q} = E_0^{p,q}$. We describe it as such by describing two submodules $Y_r^{p,q} \subset X_r^{p,q} \subset E^{p,q}$, and defining $E_r^{p,q} = X_r^{p,q}/Y_r^{p,q}$. Let $X_r^{p,q}$ be those elements of $E^{p,q}$ that are the leading terms of r-closed (p,q)-strips. Note that by definition, d sends $(r-1)$-closed $S^{p-(r-1),q+(r-1)-1}$-strips to (p,q)-strips. Let $Y_r^{p,q}$ be the leading $((p,q))$-terms of the differential d of $(r-1)$-closed $(p-(r-1), q+(r-1)-1)$-strips (where the differential is considered as a (p,q)-strip).

We next give the definition of the differential d_r of such an element $x \in X_r^{p,q}$. We take *any* r-closed (p,q)-strip with leading term x. Its differential d is a $(p+r, q-r+1)$-strip, and we take its leading term. The choice of the r-closed (p,q)-strip means that this is not a well-defined element of $E^{p,q}$. But it is well-defined modulo the $(r-1)$-closed $(p+1, r+1)$-strips, and hence gives a map $E_r^{p,q} \to E_r^{p+r,q-r+1}$.

This definition is fairly short, but not much fun to work with, so we will forget it, and instead dive into a snakes' nest of subscripts and superscripts.

We begin with making some quick but important observations about (p, q)-strips.

2.7.G. EXERCISE. Verify the following.

(a) $S^{p,q} = S^{p+1,q-1} \oplus E^{p,q}$.
(b) (Any closed (p, q)-strip is r-closed for all r.) Any element x of $S^{p,q} = S^{p,q}_0$ that is a cycle (i.e. $dx = 0$) is automatically in $S^{p,q}_r$ for all r. For example, this holds when x is a boundary (i.e. of the form dy).
(c) Show that for fixed p, q,

$$S^{p,q}_0 \supset S^{p,q}_1 \supset \cdots \supset S^{p,q}_r \supset \cdots$$

stabilizes for $r \gg 0$ (i.e. $S^{p,q}_r = S^{p,q}_{r+1} = \cdots$). Denote the stabilized module $S^{p,q}_\infty$. Show $S^{p,q}_\infty$ is the set of closed (p, q)-strips (those (p, q)-strips annihilated by d, i.e. the cycles). In particular, $S^{0,k}_r$ is the set of cycles in E^k.

2.7.10. *Defining* $E^{p,q}_r$.

Define $X^{p,q}_r := S^{p,q}_r / S^{p+1,q-1}_{r-1}$ and $Y := dS^{p-(r-1),q+(r-1)-1}_{r-1} / S^{p+1,q-1}_{r-1}$. Then $Y^{p,q}_r \subset X^{p,q}_r$ by Exercise 2.7.G(b). We define

$$\boxed{E^{p,q}_r = \frac{X^{p,q}_r}{Y^{p,q}_r} = \frac{S^{p,q}_r}{dS^{p-(r-1),q+(r-1)-1}_{r-1} + S^{p+1,q-1}_{r-1}}} \tag{2.7.10.1}$$

We have completed Goal 2.7.8(a).

You are welcome to verify that these definitions of $X^{p,q}_r$ and $Y^{p,q}_r$ and hence $E^{p,q}_r$ agree with the earlier ones of §2.7.9 (and in particular $X^{p,q}_r$ and $Y^{p,q}_r$ are both submodules of $E^{p,q}$), but we won't need this fact.

2.7.H. EXERCISE: $E^{p,k-p}_\infty$ GIVES SUBQUOTIENTS OF $H^k(E^\bullet)$. By Exercise 2.7.G(c), $E^{p,q}_r$ stabilizes as $r \to \infty$. For $r \gg 0$, interpret $S^{p,q}_r / dS^{p-(r-1),q+(r-1)-1}_{r-1}$ as the cycles in $S^{p,q}_\infty \subset E^{p+q}$ modulo those boundary elements of dE^{p+q-1} contained in $S^{p,q}_\infty$. Finally, show that $H^k(E^\bullet)$ is indeed filtered as described in (2.7.2.2).

We have completed Goal 2.7.8(b).

2.7.11. *Definition of* d_r.

We shall see that the map $d_r : E^{p,q}_r \to E^{p+r,q-r+1}$ is just induced by our differential d. Notice that d sends r-closed (p, q)-strips $S^{p,q}_r$ to $(p + r, q - r + 1)$-strips $S^{p+r,q-r+1}$, by the definition "r-closed". By Exercise 2.7.G(b), the image lies in $S^{p+r,q-r+1}_r$.

2.7.I. EXERCISE. Verify that d sends

$$dS^{p-(r-1),q+(r-1)-1}_{r-1} + S^{p+1,q-1}_{r-1} \to dS^{(p+r)-(r-1),(q-r+1)+(r-1)-1}_{r-1} + S^{(p+r)+1,(q-r+1)-1}_{r-1}.$$

(The first term on the left goes to 0 from $d^2 = 0$, and the second term on the left goes to the first term on the right.)

Thus we may define

$$\boxed{d_r : E_r^{p,q} = \frac{S_r^{p,q}}{dS_{r-1}^{p-(r-1),q+(r-1)-1} + S_{r-1}^{p+1,q-1}} \to}$$

$$\boxed{\frac{S_r^{p+r,q-r+1}}{dS_{r-1}^{p+1,q-1} + S_{r-1}^{p+r+1,q-r}} = E_r^{p+r,q-r+1}}$$

and clearly $d_r^2 = 0$ (as we may interpret it as taking an element of $S_r^{p,q}$ and applying d twice).

We have accomplished Goal 2.7.8(c).

2.7.12. *Verifying that the cohomology of* d_r *at* $E_r^{p,q}$ *is* $E_{r+1}^{p,q}$. We are left with the unpleasant job of verifying that the cohomology of

$$\frac{S_r^{p-r,q+r-1}}{dS_{r-1}^{p-2r+1,q-3}+S_{r-1}^{p-r+1,q+r-2}} \xrightarrow{d_r} \frac{S_r^{p,q}}{dS_{r-1}^{p-r+1,q+r-2}+S_{r-1}^{p+1,q-1}} \tag{2.7.12.1}$$

$$\xrightarrow{\qquad d_r \qquad} \frac{S_r^{p+r,q-r+1}}{dS_{r-1}^{p+1,q-1}+S_{r-1}^{p+r+1,q-r}}$$

is naturally identified with

$$\frac{S_{r+1}^{p,q}}{dS_r^{p-r,q+r-1} + S_r^{p+1,q-1}}$$

and this will conclude our final Goal 2.7.8(d).

We begin by understanding the kernel of the right map of (2.7.12.1). Suppose $a \in S_r^{p,q}$ is mapped to 0. This means that $da = db + c$, where $b \in S_{r-1}^{p+1,q-1}$. If $u = a - b$, then $u \in S^{p,q}$, while $du = c \in S_{r-1}^{p+r+1,q-r} \subset S^{p+r+1,q-r}$, from which u is r-closed, i.e. $u \in S_{r+1}^{p,q}$. Hence $a = b + u + x$ where $dx = 0$, from which $a - x = b + c \in S_{r-1}^{p+1,q-1} + S_{r+1}^{p,q}$. However, $x \in S^{p,q}$, so $x \in S_{r+1}^{p,q}$ by Exercise 2.7.G(b). Thus $a \in S_{r-1}^{p+1,q-1} + S_{r+1}^{p,q}$. Conversely, any $a \in S_{r-1}^{p+1,q-1} + S_{r+1}^{p,q}$ satisfies

$$da \in dS_{r-1}^{p+r,q-r+1} + dS_{r+1}^{p,q} \subset dS_{r-1}^{p+r,q-r+1} + S_{r-1}^{p+r+1,q-r}$$

(using $dS_{r+1}^{p,q} \subset S_0^{p+r+1,q-r}$ and Exercise 2.7.G(b)) so any such a is indeed in the kernel of

$$S_r^{p,q} \to \frac{S_r^{p+r,q-r+1}}{dS_{r-1}^{p+1,q-1} + S_{r-1}^{p+r+1,q-r}}.$$

Hence the kernel of the right map of (2.7.12.1) is

$$\ker = \frac{S_{r-1}^{p+1,q-1} + S_{r+1}^{p,q}}{dS_{r-1}^{p-r+1,q+r-2} + S_{r-1}^{p+1,q-1}}.$$

Next, the image of the left map of (2.7.12.1) is immediately

$$\mathrm{im} = \frac{dS_r^{p-r,q+r-1} + dS_{r-1}^{p-r+1,q+r-2} + S_{r-1}^{p+1,q-1}}{dS_{r-1}^{p-r+1,q+r-2} + S_{r-1}^{p+1,q-1}} = \frac{dS_r^{p-r,q+r-1} + S_{r-1}^{p+1,q-1}}{dS_{r-1}^{p-r+1,q+r-2} + S_{r-1}^{p+1,q-1}}$$

(as $S_r^{p-r,q-r+1}$ contains $S_{r-1}^{p-r+1,q+r-1}$).

Thus the cohomology of (2.7.12.1) is

$$\ker / \operatorname{im} = \frac{S_{r-1}^{p+1,q-1} + S_{r+1}^{p,q}}{dS_r^{p-r,q+r-1} + S_{r-1}^{p+1,q-1}} = \frac{S_{r+1}^{p,q}}{S_{r+1}^{p,q} \cap (dS_r^{p-r,q+r-1} + S_{r-1}^{p+1,q-1})}$$

where the equality on the right uses the fact that $dS_r^{p-r,q+r+1} \subset S_{r+1}^{p,q}$ and an isomorphism theorem. We thus must show

$$S_{r+1}^{p,q} \cap (dS_r^{p-r,q+r-1} + S_{r-1}^{p+1,q-1}) = dS_r^{p-r,q+r-1} + S_r^{p+1,q-1}.$$

However,

$$S_{r+1}^{p,q} \cap (dS_r^{p-r,q+r-1} + S_{r-1}^{p+1,q-1}) = dS_r^{p-r,q+r-1} + S_{r+1}^{p,q} \cap S_{r-1}^{p+1,q-1}$$

and $S_{r+1}^{p,q} \cap S_{r-1}^{p+1,q-1}$ consists of (p,q)-strips whose differential vanishes up to row $p+r$, from which $S_{r+1}^{p,q} \cap S_{r-1}^{p+1,q-1} = S_r^{p,q}$ as desired.

This completes the explanation of how spectral sequences work for a first-quadrant double complex. The argument applies without significant change to more general situations, including filtered complexes.

CHAPTER 3

Sheaves

It is perhaps suprising that geometric spaces are often best understood in terms of (nice) functions on them. For example, a differentiable manifold that is a subset of $\mathbb{R}^n$ can be studied in terms of its differentiable functions. Because "geometric spaces" can have few (everywhere-defined) functions, a more precise version of this insight is that the structure of the space can be well understood by considering all functions on all open subsets of the space. This information is encoded in something called a *sheaf*. Sheaves were introduced by Leray in the 1940's. (The reason for the name is will be somewhat explained in Remark 3.4.2.) We will define sheaves and describe useful facts about them. We will begin with a motivating example to convince you that the notion is not so foreign.

One reason sheaves are slippery to work with is that they keep track of a huge amount of information, and there are some subtle local-to-global issues. There are also three different ways of getting a hold of them.

- in terms of open sets (the definition §3.2) — intuitive but in some way the least helpful
- in terms of stalks (see §3.4)
- in terms of a base of a topology (§3.7).

Knowing which to use requires experience, so it is essential to do a number of exercises on different aspects of sheaves in order to truly understand the concept.

3.1 Motivating example: The sheaf of differentiable functions.

Consider differentiable functions on the topological space $X = \mathbb{R}^n$ (or more generally on a smooth manifold X). The sheaf of differentiable functions on X is the data of all differentiable functions on all open subsets on X. We will see how to manage this data, and observe some of its properties. On each open set $U \subset X$, we have a ring of differentiable functions. We denote this ring of functions $\mathcal{O}(U)$.

Given a differentiable function on an open set, you can restrict it to a smaller open set, obtaining a differentiable function there. In other words, if $U \subset V$ is an inclusion of open sets, we have a "restriction map" $\mathrm{res}_{V,U} : \mathcal{O}(V) \to \mathcal{O}(U)$.

Take a differentiable function on a big open set, and restrict it to a medium open set, and then restrict that to a small open set. The result is the same as if you restrict the differentiable function on the big open set directly to the small open set.

In other words, if $U \hookrightarrow V \hookrightarrow W$, then the following diagram commutes:

$$\begin{tikzcd} \mathcal{O}(W) \arrow[rr, "\mathrm{res}_{W,V}"] \arrow[dr, "\mathrm{res}_{W,U}"'] & & \mathcal{O}(V) \arrow[dl, "\mathrm{res}_{V,U}"] \\ & \mathcal{O}(U) & \end{tikzcd}$$

Next take two differentiable functions f_1 and f_2 on a big open set U, and an open cover of U by some $\{U_i\}$. Suppose that f_1 and f_2 agree on each of these U_i. Then they must have been the same function to begin with. In other words, if $\{U_i\}_{i \in I}$ is a cover of U, and $f_1, f_2 \in \mathcal{O}(U)$, and $\mathrm{res}_{U,U_i} f_1 = \mathrm{res}_{U,U_i} f_2$, then $f_1 = f_2$. Thus we can *identify* functions on an open set by looking at them on a covering by small open sets.

Finally, given the same U and cover $\{U_i\}$, take a differentiable function on each of the U_i — a function f_1 on U_1, a function f_2 on U_2, and so on — and they agree on the pairwise overlaps. Then they can be "glued together" to make one differentiable function on all of U. In other words, given $f_i \in \mathcal{O}(U_i)$ for all i, such that $\mathrm{res}_{U_i, U_i \cap U_j} f_i = \mathrm{res}_{U_j, U_i \cap U_j} f_j$ for all i and j, then there is some $f \in \mathcal{O}(U)$ such that $\mathrm{res}_{U,U_i} f = f_i$ for all i.

The entire example above would have worked just as well with continuous function, or smooth functions, or just plain functions. Thus all of these classes of "nice" functions share some common properties. We will soon formalize these properties in the notion of a sheaf.

3.1.1. The germ of a differentiable function. Before we do, we first give another definition, that of the germ of a differentiable function at a point $p \in X$. Intuitively, it is a "shred" of a differentiable function at p. Germs are objects of the form $\{(f, \text{open } U) : p \in U, f \in \mathcal{O}(U)\}$ modulo the relation that $(f, U) \sim (g, V)$ if there is some open set $W \subset U, V$ containing p where $f|_W = g|_W$ (i.e., $\mathrm{res}_{U,W} f = \mathrm{res}_{V,W} g$). In other words, two functions that are the same in a neighborhood of p (but may differ elsewhere) have the same germ. We call this set of germs the stalk at p, and denote it $\mathcal{O}_p$. Notice that the stalk is a ring: you can add two germs, and get another germ: if you have a function f defined on U, and a function g defined on V, then $f + g$ is defined on $U \cap V$. Moreover, $f + g$ is well-defined: if f' has the same germ as f, meaning that there is some open set W containing p on which they agree, and g' has the same germ as g, meaning they agree on some open W' containing p, then $f' + g'$ is the same function as $f + g$ on $U \cap V \cap W \cap W'$.

Notice also that if $p \in U$, you get a map $\mathcal{O}(U) \to \mathcal{O}_x$. Experts may already see that we are talking about germs as colimits.

We can see that $\mathcal{O}_p$ is a local ring as follows. Consider those germs vanishing at p, which we denote $\mathfrak{m}_p \subset \mathcal{O}_p$. They certainly form an ideal: $\mathfrak{m}_p$ is closed under addition, and when you multiply something vanishing at p by any other function, the result also vanishes at p. We check that this ideal is maximal by showing that the quotient map is a field:

$$0 \longrightarrow \mathfrak{m} := \text{ideal of germs vanishing at } p \longrightarrow \mathcal{O}_p \xrightarrow{f \mapsto f(p)} \mathbb{R} \longrightarrow 0 \tag{3.1.1.1}$$

3.1.A. EXERCISE. Show that this is the only maximal ideal of $\mathcal{O}_p$. (Hint: show that every element of $\mathcal{O}_p \setminus \mathfrak{m}$ is invertible.)

Note that we can interpret the value of a function at a point, or the value of a germ at a point, as an element of the local ring modulo the maximal ideal. (We will see that this doesn't work for more general sheaves, but *does* work for things behaving like sheaves of functions. This will be formalized in the notion of a *local-ringed space*, which we will see, briefly, in §7.3.

3.1.2. *Aside.* Notice that $\mathfrak{m}/\mathfrak{m}^2$ is a module over $\mathcal{O}_p/\mathfrak{m} \cong \mathbb{R}$, i.e. it is a real vector space. It turns out to be naturally (whatever that means) the cotangent space to the manifold at p. This insight will prove handy later, when we define tangent and cotangent spaces of schemes.

3.1.B. EXERCISE FOR THOSE WITH DIFFERENTIAL GEOMETRIC BACKGROUND. Prove this.

3.2 Definition of sheaf and presheaf

We now formalize these notions, by defining presheaves and sheaves. Presheaves are simpler to define, and notions such as kernel and cokernel are straightforward. Sheaves are more complicated to define, and some notions such as cokernel require more thought. But sheaves are more useful because they are in some vague sense more geometric; you can get information about a sheaf locally.

3.2.1. Definition of sheaf and presheaf on a topological space X.
To be concrete, we will define sheaves of sets. However, *Sets* can be replaced by any category, and other important examples are abelian groups *Ab*, k-vector spaces Vec_k, rings *Rings*, modules over a ring Mod_A, and more. (You may have to think more when dealing with a category of objects that aren't "sets with additional structure", but there aren't any new complications. In any case, this won't be relevant for us.) Sheaves (and presheaves) are often written in calligraphic font, or with an underline. The fact that $\mathcal{F}$ is a sheaf on a topological space X is often written as

$$\begin{array}{c}\mathcal{F}\\ \Big| \\ X\end{array}$$

3.2.2. Definition: Presheaf. A **presheaf** $\mathcal{F}$ on a topological space X is the following data.

• To each open set $U \subset X$, we have a set $\mathcal{F}(U)$ (e.g. the set of differentiable functions in our motivating example). (Notational warning: Several notations are in use, for various good reasons: $\mathcal{F}(U) = \Gamma(U, \mathcal{F}) = H^0(U, \mathcal{F})$. We will use them all.) The elements of $\mathcal{F}(U)$ are called **sections of $\mathcal{F}$ over** U.

• For each inclusion $U \hookrightarrow V$ of open sets, we have a **restriction map** $\mathrm{res}_{V,U} : \mathcal{F}(V) \to \mathcal{F}(U)$ (just as we did for differentiable functions).

The data is required to satisfy the following two conditions.

• The map $\mathrm{res}_{U,U}$ is the identity: $\mathrm{res}_{U,U} = \mathrm{id}_{\mathcal{F}(U)}$.

Index