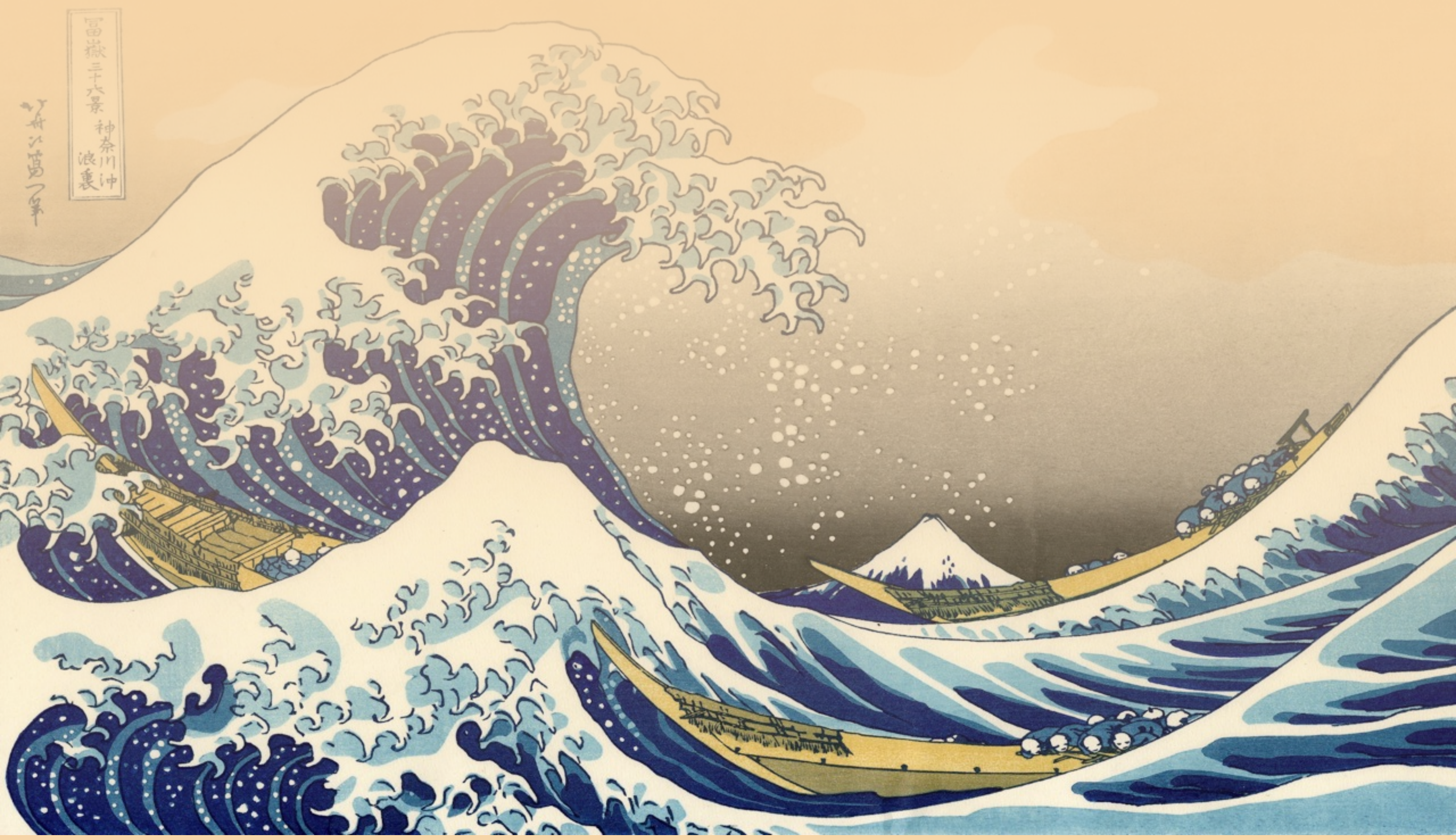
# Algebraic Geometry

## UT Austin, Spring 2016



**M390C NOTES: ALGEBRAIC GEOMETRY**

ARUN DEBRAY
FEBRUARY 2, 2016

These notes were taken in UT Austin's Math 390c (Algebraic Geometry) class in Spring 2016, taught by David Ben-Zvi. I live-TEXed them using `vim`, and as such there may be typos; please send questions, comments, complaints, and corrections to a.debray@math.utexas.edu.

Contents



Episode I.

## The Course Awakens: 1/19/16

> *"There was a mistranslation in Grothendieck's quote, 'the rising sea:' he was actually talking about raising an X-wing fighter out of a swamp using the Force."*

There are a lot of things that go under the scheme of algebraic geometry, but in this class we're going to use the slogan "algebra = geometry;" we'll try to understand algebraic objects in terms of geometry and vice versa.

There are two main bridges between algebra and geometry: to a geometric object we can associate algebra via functions, and the reverse construction might be less familiar, the notion of a spectrum. This is very similar to the notion of the spectrum of an operator.

We will follow the textbook of Ravi Vakil, *The Rising Sea*. There's also a course website.[1] The prerequisites will include some commutative algebra, but not too much category theory; some people in the class might be bored. Though we're not going to assume much about algebraic sets, basic algebraic geometry, etc., it will be helpful to have seen it.

Let's start. Suppose $X$ is a space; then, there's generally a notion of $\mathbb{C}$-valued functions on it, and this space might be $F(X)$. For example, if $X$ is a smooth manifold, we have $C^\infty(X)$, and if $X$ is a complex manifold, we have the holomorphic functions $\mathrm{Hol}(X)$.[2] Another category of good examples is *algebraic sets*, $X \subset \mathbb{C}^n$ that is given by the common zero set of a bunch of polynomials: $X = \{f_1(x) = \cdots = f_k(x) = 0\}$ for some $f_1, \dots, f_k \in \mathbb{C}[x_1, \dots, x_n]$. These have a natural notion of function, *polynomial functions*, which are polynomials in $\mathbb{C}[x_1, \dots, x_n]$ restricted to $X$, If $I(X)$ is the functions vanishing on $X$, then these functions are given by $\mathbb{C}[x_1, \dots, x_n]/I$.

The point is, on all of our spaces, the functions have a natural ring structure.[3] In fact, there's more: the constant functions are a map $\mathbb{C} \to F(X)$, and since $\mathbb{C}$ is a field, this map is injective. This means $F(X)$ is a $\mathbb{C}$-algebra, i.e. it is a $\mathbb{C}$-vector space with a commutative, $\mathbb{C}$-linear multiplication.

[1] https://www.ma.utexas.edu/users/benzvi/teaching/alggeom/syllabus.html.

[2] The best examples here are Riemann surfaces; when the professor imagines a "typical" or example algebraic variety, he sees a Riemann surface.

[3] In this class, all rings will be commutative and have a 1. Ring homomorphisms will send 1 to 1.

One of the things Grothendieck emphasized is that one should never look at a space (or an anything) on its own, but consider it along with maps between spaces. For example, given a map $\pi : X \to Y$ of spaces, we always have a *pullback* homomorphism $\pi^* : F(Y) \to F(X)$: if $f : Y \to \mathbb{C}$, then its pullback is $\pi^* y(x) = y(\pi(x))$. This tells us that we have a *functor* from spaces to commutative rings.

**Categories and Functors.** This is all done in Vakil's book, but in case you haven't encountered any categories in the streets, let's revisit them.

**Definition.** A *category* $\mathsf{C}$ consists of a set[4] of *objects* $\operatorname{Ob}\mathsf{C}$; if $X \in \operatorname{Ob}\mathsf{C}$, we just say $X \in \mathsf{C}$. We also have for every $X, Y \in \mathsf{C}$ the set $\operatorname{Hom}_{\mathsf{C}}(X, Y)$ of *morphisms*. For every $X, Y, Z \in \mathsf{C}$, there's a *composition map* $\operatorname{Hom}_{\mathsf{C}}(X, Y) \times \operatorname{Hom}_{\mathsf{C}}(Y, Z) \to \operatorname{Hom}_{\mathsf{C}}(Y, Z)$ and a unit $1_X \in \operatorname{Hom}_{\mathsf{C}}(X, X) = \operatorname{End}_{\mathsf{C}}(X)$ satisfying a bunch of axioms that make this behave like associative function composition.

To be precise, we want categories to behave like monoids, for which the product is associative and unital. In fact, a category with one object is a monoid. Thus, we want morphisms of categories to act like morphisms of monoids: they should send composition to composition.

**Definition.** A *functor* $F : \mathsf{C} \to \mathsf{D}$ is a function $F : \operatorname{Ob}\mathsf{C} \to \operatorname{Ob}\mathsf{D}$ with an induced map on the morphisms:

- If the map acts as $\operatorname{Hom}_{\mathsf{C}}(X, Y) \to \operatorname{Hom}_{\mathsf{D}}(F(X), F(Y))$, $F$ is called a *covariant* functor.
- If it sends $\operatorname{Hom}_{\mathsf{C}}(X, Y) \to \operatorname{Hom}_{\mathsf{D}}(F(Y), F(X))$, then $F$ is *contravariant*.

When we say "functor," we always mean a covariant functor, and here's the reason. Recall that for any monoid $A$ there's the *opposite monoid* $A^{\mathrm{op}}$ which has the same set, but reversed multiplication: $f \cdot_{\mathrm{op}} g = g \cdot f$. Similarly, given a category $\mathsf{C}$, there's an *opposite category* $\mathsf{C}^{\mathsf{op}}$ with the same objects, but $\operatorname{Hom}_{\mathsf{C}^{\mathsf{op}}}(X, Y) = \operatorname{Hom}_{\mathsf{C}}(Y, X)$. Then, a contravariant functor $\mathsf{C} \to \mathsf{D}$ is really a covariant functor $\mathsf{C}^{\mathsf{op}} \to \mathsf{D}$. Hence, in this class, we'll just refer to functors, with opposite categories where needed.

**Exercise.** Show that a functor $\mathsf{C}^{\mathsf{op}} \to \mathsf{D}$ induces a functor $\mathsf{C} \to \mathsf{D}^{\mathsf{op}}$.

When presented a category, you should always ask what the morphisms are; on the other hand, if someone tells you "the category of smooth manifolds," they probably mean that the morphisms are smooth functions.

Now, we see that pullback is a functor $F : \mathsf{Spaces} \to \mathsf{Ring}^{\mathsf{op}}$. One of the major goals of this class is to define a category of spaces on which this functor is an equivalence. This might not make sense, *yet*. This is the seed of "algebra = geometry."

**Definition.** Let $F, G : \mathsf{C} \rightrightarrows \mathsf{D}$ be functors. A *natural transformation* $\eta : F \Rightarrow G$ is a collection of maps: for every $X \in \mathsf{C}$, there's a map $\eta_X : F(X) \to G(X)$ satisfying a consistency condition: for every $f : X \to Y$ in $\mathsf{C}$, there's a commutative diagram

$$\begin{tikzcd} F(X) \arrow[r, "\eta_X"] \arrow[d, "F(f)"'] & G(X) \arrow[d, "G(f)"] \\ F(Y) \arrow[r, "\eta_Y"] & G(Y) \end{tikzcd}$$

That is, a natural transformation relates the objects and the morphisms, and reflects the structure of the category.

**Definition.** A natural transformation $\eta$ is a *natural isomorphism* if for every $X \in \mathsf{C}$, the induced $\eta_X \in \operatorname{Hom}_{\mathsf{D}}(F(X), G(X))$ is an isomorphism.

This is equivalent to having a natural inverse to $\eta$.

So one might ask, what is the notion for which two categories are "the same?" One might naïvely suggest two functors whose composition is the identity functor, but this is bad. The set of objects isn't very useful: it doesn't capture the structure of the category. In general, asking for equality of objects is worse than asking for isomorphism of objects. Here's the right notion of sameness.

[4]This is wrong. But if you already know that, you know that worrying about set-theoretic difficulties is a major distraction here, and not necessary for what we're doing, so we're not going to worry about it.

**Definition.** Let $\mathsf{C}$ and $\mathsf{D}$ be categories. Then, a functor $F : \mathsf{C} \to \mathsf{D}$ is an *equivalence of categories* if there's a functor $G : \mathsf{D} \to \mathsf{C}$ such that there are natural isomorphisms $FG \to \mathrm{Id}_{\mathsf{D}}$ and $GF \to \mathrm{Id}_{\mathsf{C}}$.

This is a very useful notion, and as such it will be useful to see an equivalence that is not an isomorphism.

**Exercise.** Let $k$ be a field, and let $\mathsf{D} = \mathsf{fdVect}_k$, the category of finite-dimensional vector spaces and linear maps, and let $\mathsf{C}$ be the category whose objects are $\mathbb{Z}_{\geq 0}$, the natural numbers, with an object denoted $\langle n \rangle$, and with $\mathrm{Hom}(\langle n \rangle, \langle m \rangle) = \mathrm{Mat}_{m \times n}$. This is a category with composition given by matrix multiplication.

Let $F : \mathsf{C} \to \mathsf{D}$ send $\langle n \rangle \mapsto k^n$, and with the standard realization of matrices as linear maps. Show that $F$ is an equivalence of categories.

This category $\mathsf{C}$ has only some vector spaces, but for those spaces, it has all of the morphisms.

**Definition.** Let $F : \mathsf{C} \to \mathsf{D}$ be a functor.

- $F$ is *faithful* if all of the maps $\mathrm{Hom}_{\mathsf{C}}(X, Y) \hookrightarrow \mathrm{Hom}_{\mathsf{D}}(F(X), F(Y))$ are injective.
- $F$ is *fully faithful* if all of these maps are isomorphsism.
- $F$ is *essentially surjective* if every $X \in \mathsf{D}$ is isomorphic to $F(Z)$ for some $Z \in \mathsf{C}$.

The following theorem will also be a useful tool.

**Theorem 1.1.** *A functor $F : \mathsf{C} \to \mathsf{D}$ is an equivalence iff it is fully faithful and essentially surjective.*

So, to restate, we want a category of spaces that is the opposite category to the category of rings; this is what Grothendieck had in mind. In fact, let's peek a few weeks ahead and make a curious definition:

**Definition.** The *category of affine schemes* is $\mathsf{Rings}^{\mathsf{op}}$.

Of course, we'll make these into actual geometric objects, but categorically, this is all that we need.

Recall that if $f : M \to N$ is a set-theoretic map of manifolds, then $f$ is smooth iff its pullback sends $C^\infty$ functions on $N$ to $C^\infty$ functions on $M$. The first step in this direction is the following theorem, sometimes called *Gelfand duality*.

**Theorem 1.2** (Gelfand-Naimark)**.** *The functor $X \mapsto C^0(X)$ (the ring of continuous functions) defines an equivalence between the category of compact Hausdorff spaces and the (opposite) category of commutative $C^*$-algebras.*

This is an algebro-geometric result: it identifies a category of spaces with the opposite category of a category of algebraic objects.

However, we need to think harder than Gelfand duality in terms of compact, complex manifolds or in terms of algebraic spaces: for example, for $X = \mathbb{CP}^1$, $\mathrm{Hol}(X) = \mathbb{C}$: the only holomorphic functions are constant. The issue is that there are no partitions of unity in the holomorphic or algebraic world. This means we'll need to keep track of local data too, which will lead into the next few lectures' discussions on *sheaf theory*.

Returning to the example of algebraic sets, suppose $X$ and $Y$ are algebraic sets. What is the set of their morphisms? We decided the ring of functions was the polynomial functions $Y \to \mathbb{C}$, so we want maps $X \to Y$ to be those whose pullbacks send polynomial functions to polynomial functions. To be precise, the *ideal of* $X$ is $I(X) = \{f \in \mathbb{C}[x_1, \dots, x_n] \mid f|_X = 0\}$, defining a map $I$ from algebraic subsets of $\mathbb{C}^n$) to ideals in $\mathbb{C}[x_1, \dots, x_n]$. There's also a reverse map $V$,[5] sending an ideal $I$ to $V(I) = \{x \in \mathbb{C}^n \mid f(x) = 0 \text{ for all } f \in I\}$. From classical commutative algebra, it's a fact that this is finitely generated, so it's the vanishing locus of a finite number of polynomials, and therefore in fact an algebraic set.

The dictionary between algebraic sets and ideals of $\mathbb{C}[x_1, \dots, x_n]$ is one of many versions of the Nullstellensatz (more or less German for the "zero locus theorem"): if $J$ is an ideal, $I(V(J)) = \sqrt{J}$, its radical.

**Definition.** Let $R$ be a ring and $J \subset R$ be an ideal. Then, the *radical* of $J$ is $\sqrt{J} = \{r \in R \mid r^n \in J \text{ for some } n > 0\}$. One says that $J$ is *radical* if $J = \sqrt{J}$.

What this says is that $J$ is radical iff $R/J$ has no nonzero nilpotents.[6] Why are these kinds of ideals relevant? If $X \subset \mathbb{C}^n$ and $f$ vanishes on $X$, then so does $f^n$ for all $n$. That is, radicals encode the geometric property of vanishing, which is why $I(X)$ is a radical ideal.

[5] $V$ stands for "vanishing," "variety," or maybe "vendetta."

[6] Recall that if $R$ is a ring, an $r \in R$ is *nilpotent* if $r^n = 0$ for some $n$.

This is an outline of what classical algebraic geometry studies: it starts by defining algebraic subsets, and establishing a bijection between algebraic subsets of $\mathbb{C}^n$ and radical ideals of $\mathbb{C}[x_1, \dots, x_n]$. This isn't yet an equivalence of categories. Radical ideals correspond to finitely generated $\mathbb{C}$-algebras with no (nonzero) nilpotents: an ideal $I$ corresponds to the $\mathbb{C}$-algebra $\mathbb{C}[x_1, \dots, x_n]/I$.

This is all what the course is *not* about; we're going to replace the category of finitely generated, nilpotent-free $\mathbb{C}$-algebras with the category of *all* rings, but we want to keep some of the same intuition. This involves generalizing in a few directions at once, but we'll try to write down a dictionary; the defining principle is to identify spaces $X$ with rings $R = F(X)$, their ring of functions.

A point $x \in X$ is a map $i_x : x \to X$, so we get a pullback $i_x^* : F(X) \to \mathbb{C}$ given by evaluation at $x$. Let $\mathfrak{m}_x = \ker(i_x^*)$; since $\mathbb{C}$ is a field, this is a maximal ideal.[7] If $k$ is a field and $R$ is a $k$-algebra, then $R/I$ is also a $k$-algebra, so in particular if $I$ is maximal, then $k \hookrightarrow R/I$ is a map of fields, and therefore a field extension. Thus, if $k$ is algebraically closed (e.g. we're studying $\mathbb{C}$) and $R$ is a finitely generated $k$-algebra, then maximal ideals of $R$ are in bijection with homomorphisms $R \to k$.

Thus, given a ring $R$, we'll associate a set $\operatorname{MSpec}(R)$, the set of maximal ideals of $R$, such that $R$ should be its ring of functions. To do this, we'll say that an $r \in R$ is a "function" on $\operatorname{MSpec}(R)$ by acting on an $\mathfrak{m}_x \subset R$ as $r \bmod \mathfrak{m}_x$. This is a "number," since it's in a field, but the notion may be different at every point in $\operatorname{MSpec}(R)$! For example, if $R = \mathbb{Z}$, then $\operatorname{MSpec}(\mathbb{Z})$ is the set of primes, and $n \in \mathbb{Z}$ is a function which at 2 is $n \bmod 2$, at 3 is $n \bmod 3$, and so on.

A perhaps nicer example is when $R = \mathbb{R}[x]$, which has maximal ideals $(x - t)$ for all $t \in \mathbb{R}$. Here, evaluation sends $f(x) \mapsto f(x) \bmod (x - t) = f(t)$. That is, this is really evaluation, and here the quotient field is $\mathbb{R}$. So these look like good old real-valued functions, but these aren't all the maximal ideals: $(x^2 + 1)$ is also a maximal ideal, and $\mathbb{R}[x]/(x^2 + 1) \cong \mathbb{C}$. Then, we do get a kind of evaluation again, but we have to identify points and their complex conjugates.

So we'll have to find a good notion of geometry which generalizes from $\mathbb{C}$-algebras to $k$-algebras for any field $k$, to any commutative rings. We'll also have to think about nilpotents: we threw them away by thinking about zero sets, but they play a huge role in ring theory.

Episode II.

## Attack of the Cones: 1/21/16

*"To this end, we're going to give a crash course in category theory over the next few lectures; the door is over there."*

Remember that our general agenda is to match algebra and geometry; one way to express this idea is to take the category of rings and identify it with some category of geometric objects. However, we're going to reverse the arrows, and we'll get the category of affine schemes. These are some geometric spaces, with a contravariant functor from affine schemes to rings given by taking the ring of functions and a functor in the opposite direction called Spec.

One potential issue is that spaces may not have enough functions, e.g. $\mathbb{CP}^1$ as a complex manifold only has constant functions; as such, we'll enlarge our category to a whole category of schemes, which will also have an algebraic interpretation. Another weird aspect is that functions may take values in varying fields.

Schemes generalize geometry in three different directions: gluing spaces together to ensure we have enough functions is topology, like making manifolds; functions having varying codomains is useful for arithmetic and number theory; and allowing for rings with nilpotents feels a little like analysis.

Last time, we defined $\operatorname{MSpec}(R)$ for a ring $R$, the set of maximal ideals. It turns out that topology is not sufficient to understand these spaces; for example, the class of *local rings* are those with only one maximal ideal. There are many such rings, e.g. $\mathbb{C}[x]/(x^n)$, whose maximal ideal is $(x)$. In short, MSpec doesn't see nilpotents.

To any ring $R$, one can attach the category $\mathsf{Mod}_R$, whose objects are $R$-modules and morphisms are $R$-linear maps (those commuting with the action of $R$). This category is one of the more important things one studies in algebra, and we also want to express them in terms of geometric objects that are related somehow to $\operatorname{Spec} R$. This should also help us understand the algebraic properties of $R$-modules too.

[7] Recall that an ideal $I \subset R$ is *maximal* iff $R/I$ is a field. This is about the level of commutative algebra that we'll be assuming.

**Crash Course in Categories.** There's a lot of categorical notions in algebraic geometry; it does strike one as a painful way to start a course, but hopefully we can get it out of our systems and move on to geometry knowing what we need. This corresponds to chapters 1 and 2 in the book.

We've seen several examples of categories: sets, groups, rings, etc. The next example is a useful class of categories.

**Definition.** A *poset* is a set $\mathsf{S}$ and a relation $\le$ on $\mathsf{S}$ that is

- *reflexive*, so $x \le x$ for all $x \in \mathsf{S}$;
- *transitive*, so if $x \le y$ and $y \le z$, then $x \le z$; and
- *antisymmetric*, so if $x \le y$ and $y \le x$, then $x = y$.

$\mathsf{S}$ has the structure of a category: the objects are the elements of $\mathsf{S}$, and $\mathrm{Hom}(x, y)$ is $\{\mathrm{pt}\}$ if $x \le y$ and is empty otherwise.

Transitivity means that we have composition, and reflexivity gives us identity maps.

This is an unusual example compared to things like "the category of all (somethings)," but is quite useful: a functor from the poset $\bullet \to \bullet$ to another category $\mathsf{C}$ is a choice of $A, B \in \mathsf{C}$ and a map $A \to B$; a functor from the poset $\mathbb{N}$ is the same as an infinite sequence in $\mathsf{C}$, and a commutative diagram is the same as a functor out of the category

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

into $\mathsf{C}$.

**Example 2.1.** A particularly important example of this: if $X$ is a topological space, then its open subsets form a poset under inclusion. Hence, they form a category, called $\mathsf{Top}(X)$. This category is important for sheaf theory, which we will say more about later. For example, if $A$ is an abelian group and $U \subset X$ is open, then let $\mathcal{O}_A(U)$ denote the abelian group of $A$-valued functions on $U$ (for example, $A$ might be $\mathbb{C}$, so $\mathcal{O}_A(U) = C^\infty(U)$). If $V \subset U$, then restriction of functions defines a map $\mathrm{res}_U^V : \mathcal{O}_A(U) \to \mathcal{O}_A(V)$. Since restriction obeys composition, then we've defined a functor $\mathcal{O}_A : \mathsf{Top}(X)^{\mathsf{op}} \to \mathsf{Ab}$ (or perhaps to $\mathbb{C}$-algebras, or another category); this is a *presheaf of abelian groups* (or $\mathbb{C}$-algebras, etc.).

To be precise, a *presheaf* on $X$ is a functor out of $\mathsf{Top}(X)^{\mathsf{op}}$. This is a way of organizing functions in a way that captures restriction; it will be very useful throughout this class.

Returning to category theory, one of its greatest uses is to capture structure through universal properties, rather than using explicit details of a given category. We'll give a few universal properties here.

**Definition.** Let $\mathsf{C}$ be a category.

- A *final* (or *terminal*) object in $\mathsf{C}$ is a $* \in \mathsf{C}$ such that for all $X \in \mathsf{C}$, there's a unique map $X \to *$.
- An *initial* object is a $* \in \mathsf{C}$ such that for all $X \in \mathsf{C}$, there's a unique map $* \to X$.

This is not the last time we'll have dual constructions produced by reversing the arrows.

**Example 2.2.** If $\mathsf{C}$ is a poset, then a terminal object is exactly a maximum element, and an initial object is a minimum element. Thus, in particular, they do not necessarily exist.

Nonetheless, if a final (or initial) object exists, it's necessarily unique.

**Proposition 2.3.** *Let $*$ and $*'$ be terminal objects in $\mathsf{C}$; then, there's a unique isomorphism $*$ to $*'$.*

*Proof.* There's a unique map $* \to *$, which therefore must be the identity, and there are unique maps $* \to *'$ and $*' \to *$, so composing these, we must get the identity, so such an isomorphism exists, and it must be unique, since there's only one map $* \to *'$. ⊠

By reversing the arrows, the same thing is true for initial objects. Thus, if such an object exists, it's unique, so one often hears "the" initial or final object. These will be useful for constructing other universal properties.

**Example 2.4.**

(1) In the category of sets, or in the category of topological spaces, the final object is a single point: everything maps to the point. The initial object is the empty set, since there's a unique (empty) map to any set or space.
(2) In $\mathsf{Ab}$ or $\mathsf{Vect}_k$ (abelian groups and vector spaces, respectively), 0 is both initial and terminal: the unique map is the zero map. An object that is initial and final is called a *zero object*; as in the case of sets, it may not exist.
(3) In the category of rings, 0 is terminal, but not initial (since a map out of 0 must send $0 = 1$ to 0 and 1). $\mathbb{Z}$ is initial, with the unique map determined by $1 \mapsto 1$.[8]
(4) Even though we don't really understand what an affine scheme is yet, we know that $\operatorname{Spec}\mathbb{Z}$ has to be a terminal object, and $\operatorname{Spec} 0$ has to be the initial object. Since we want this to be geometric, then $\operatorname{Spec}\mathbb{Z}$ will play the role of a point. It might not look like a point, but categorically it behaves like one.
(5) The category of fields is also interesting: setting $1 = 0$ isn't allowed, so there are neither initial nor terminal objects! If we specialize to fields of a given characteristic, then we get a unique map out of $\mathbb{Q}$ or $\mathbb{F}_p$, so the category of fields of a given characteristic is initial.
(6) The poset $\mathsf{Top}(X)$ has $\varnothing$ initial and $X$ terminal: it has top and bottom objects.

The fact that initial and terminal objects are unique means that if you characterize an object in terms of initial or terminal objects, then you know they're unique as soon as they exist.

**Definition.** If $R$ is a ring, we have the category $\mathsf{Alg}_R$ of $R$-algebras (rings $T$ with the extra structure of a map $R \to T$; morphisms must commute with this map). This is an example of something more general, called an *undercategory*: if $\mathsf{C}$ is a category and $X \in \mathsf{C}$, then the undercategory $X \downarrow \mathsf{C}$ is the category whose objects are data of $Y \in \mathsf{C}$ with $\mathsf{C}$-morphisms $a_Y : X \to Y$ and whose morphisms are commutative diagrams

$$\begin{tikzcd} Y_1 \arrow[rr, "\varphi"] & & Y_2 \\ & X. \arrow[ul, "a_{Y_1}"] \arrow[ur, "a_{Y_2}"'] & \end{tikzcd}$$

In the same way, the *overcategory* $X \uparrow \mathsf{C}$ is the same idea, but with maps to $X$ rather than from $X$ (e.g. spaces over a given space $X$).

Thus, it's possible to concisely define $\mathsf{Alg}_R = R \downarrow \mathsf{Ring}$. We will see other examples of this.

**Example 2.5** (Localization)**.** Let $R$ be a ring and $S \subset R$ be a multiplicative subset. Then, the *localization at $S$* is $S^{-1}R = \{r/s \mid r \in R, s \in S : r/s = r/s' \text{ when } s''(rs' - r's) = 0 \text{ for some } s'' \in S\}$. This is a construction we'll use a lot, so it will be useful to have a canonical characterization of them.

Now, let $\mathsf{C}$ be the category of $R$-algebras $T$ with maps $(\varphi_T : R \to T$ such that (and this is a property, not structure) $\varphi_T(s)$ is invertible in $T$ for all $s \in S$.

**Exercise.** Show that $S^{-1}R$ is the initial object in $\mathsf{C}$.

Note that the naïve idea that localization is "fractions in $S$" is true if $R$ is an integral domain, but if we have zero divisors, the $R$-algebra structure map $R \to S^{-1}R$ need not be injective. But the point is that if $T$ is an $R$-algebra where the elements of $S$ become invertible, the map $\varphi_T$ factors through $S^{-1}R$; this means that $S^{-1}R$ is the element of $\mathsf{C}$ that's "closest to $R$." However, you still have to concretely build it to show that it exists; however, we know already that it's determined up to unique isomorphism, so we say "the" localization.

Another very fundamental language for making constructions is that of limits and colimits. It may seem a little strange, but it's quite important.

**Definition.** Let $I$ be a *small category* (so its objects form a set); in the context of limits, we will refer to it as an *index category*. Then, a functor $A : I \to \mathsf{C}$ is called a *$I$-shaped* (or *$I$-indexed*) *diagram in* $\mathsf{C}$.

That is, if $m : i \to j$ is a morphism in $I$, then this diagram contains an arrow $A(m0 : A_i \to A_j$.

**Definition.** Let $A$ be an $I$-shaped diagram in $\mathsf{C}$. Then, a *cone* on $A$ is the data of an object $B \in \mathsf{C}$ and maps $A_i \to B$ for every $i \in I$ commuting with the morphisms in $I$. The cones on $A$ form a category $\mathsf{Cones}_A$,

[8] That rings and ring homomorphisms are unital is important for this to be true.

$$\begin{tikzcd}
A_1 \arrow[r] \arrow[d] \arrow[drr, bend left] & A_3 \arrow[dr] & \\
A_2 \arrow[ur] \arrow[rr] & & B
\end{tikzcd}$$

Figure 1. A cone on a diagram $A$.

where the morphisms are maps $B \to B'$ commuting with all the maps in the cone.

We can also take the category of "co-cones," which are data of maps from $B$ *into* the diagram. This is not quite the opposite category (since we want maps $B \to B'$ commuting with the maps into the diagram).[9]

**Definition.**

- The *colimit* $\varinjlim_I A$ is the initial object in the category of cones of $A$.
- The *limit* $\varprojlim_I A$ is the terminal object in the category of co-cones of $A$.

As before, colimits and limits may or may not exist, but if they do, they're unique up to unique isomorphism.

Colimits act like a quotient, and it's easier to map out of them. Correspondingly, limits behave like a subobject, and it's easier to map into them.

**Example 2.6** (Products and Coproducts)**.** Let $I = \bullet\ \bullet$ be a two-element discrete set (no non-identity arrows). Thus, an $I$-shaped diagram is just a choice of two spaces $A_1$ and $A_2$, so a colimit $C_1$ is the data of a unique map $\varphi_B$ for each $B \in \mathsf{C}$ fitting into the following diagram.

$$\begin{tikzcd}
A_1 \arrow[dr] \arrow[drr, bend left] & & \\
& C_1 \arrow[r, "\varphi_B"] & B \\
A_2 \arrow[ur] \arrow[urr, bend right] & &
\end{tikzcd}$$

This is called the *coproduct* of $A_1$ and $A_2$, denoted $A_1 \sqcup A_2$ or $A_1 \amalg A_2$.

Similarly, the limit of $A$ is called the *product* of $A_1$ and $A_2$, is denoted $A_1 \times A_2$, and fits into the diagram

$$\begin{tikzcd}
& & A_1 \\
B \arrow[r] \arrow[urr] \arrow[drr] & A_1 \times A_2 \arrow[ur] \arrow[dr] & \\
& & A_2
\end{tikzcd}$$

In the same way, if $I$ is a larger discrete set, we get coproducts and products of objects in $\mathsf{C}$ indexed by $I$, denoted $\coprod_I A_i$ and $\prod_I A_i$, respectively.

In the category of sets, the product is Cartesian product, and the coproduct is disjoint union. The same is true in topological spaces.

In the category of groups, the product is once again Cartesian product, but the coproduct is the free product (mapping out of it is the same as mapping out of the individual components, which is not true of the direct product). Note that this is distinct as underlying sets from the coproduct of sets.

In linear categories, e.g. $\mathsf{Ab}$, $\mathsf{Mod}_R$, or $\mathsf{Vect}_k$, $V \oplus W$ is the product and coproduct, and the same is true over all finite $I$. However, this is *not* true when $I$ is infinite: the coproduct is the direct sum, which takes finite sums of elements, and the product is the Cartesian product, which takes arbitrary sums of elements. It's worth working out why this is, and how it works.

[9] Some people switch the definitions of cones and co-cones, but since we're not going to use these words very much, it doesn't matter all that much.

Many of these categories are "sets with structure," e.g. groups, vector spaces, topological spaces, and so on. In these cases, there is a *forgetful functor* which forgets this structure: indeed, a group homomorphisms (continuous map, linear map) is a map of sets too.[10]

There's a useful principle here: *forgetful functors preserve limits*: if $F$ is a forgetful functor, then there is a canonical isomorphism $F(\varprojlim A) \cong \varprojlim F(A)$. This is something that can be defined more rigorously and proven. But one important corollary is that if you know what the limit looks like for sets, it's the same in groups, rings, vector spaces, topological spaces, and so on. However, this is very false for coproducts, e.g. the coproduct on groups is not the same as the one on sets.

This becomes a little cooler once we see limits that aren't just products.

**Example 2.7.** Consider the diagram of rings

$$\cdots \longrightarrow \mathbb{Z}/p^n \longrightarrow \cdots \longrightarrow \mathbb{Z}/p^3 \longrightarrow \mathbb{Z}/p^2 \longrightarrow \mathbb{Z}/p,$$

where each map is given by modding out by $p$. One can show that the limit exists, and it'll be the same as the limit of the underlying sets, a sequence of compatible elements; this limit is called the *p-adic integers*, denoted $\mathbb{Z}_p$. More generally, the same thing works for $\varprojlim R/I^n$ for an ideal $I \subset R$, and defines the *I-adic completion* $\widehat{R}_I$, which we'll revisit, since it has useful geometric meaning.

Episode III.

## The Yoneda Chronicles: 1/26/16

*"There's probably lots of notations [for this], so let me choose a bad one."*

Last time, we were talking about universal propeties, which tend to correspond to terminal or initial objects. This tends to characterize an object up to unique isomorphism, so there's in a sense only one solution.

There is *not* only one object. There might be a billion! Or infinitely many. But any two are uniquely isomorphic: if we take $\mathsf{C}^{\text{initial}}$, the subcategory of initial objects, it's equivalent to $*$, the category with one object and the identity morphism.[11] And we never look at categories up to isomorphism, only equivalence, so this is a better viewpoint.

We also started talking about limits and colimits last time; these are very important examples of universal properties. These are initial (resp. terminal) objects in the category of cones (resp. co-cones) of $I$-shaped diagrams, which are functors $I \to \mathsf{C}$. In other words, a colimit of a diagram is mapped to by every object in a diagram in a way compatible with the diagram maps, and such that any other mapped-to object factors through the limit; a limit maps to the diagram and factors through any other such map. Since these are initial or terminal objects, they are unique up to unique isomorphism, so one hears "the" (co)limit. Limits are analogous to subobjects, and colimits are more like quotients; as such, colimits tend to be more poorly behaved.

We also defined products and coproducts, which are limits and colimits, respectively, over a discrete set (made into a category by adding only the identity maps). For example, in the category of modules over a ring, the coproduct is direct sum, and the product is the Cartesian product; the difference between these is only felt at the infinite level, and the direct sum is more subtle. In the category of groups, the product of groups is the Cartesian product again (a group structure on the product as a set); on the other hand, the coproduct is *not* the coproduct of sets (disjoint union): it's the free product of groups, because maps out of $G$ and $H$ correspond to maps out of $G * H$. And this is different than the coproduct of abelian groups: it's direct sum (since abelian groups are $\mathbb{Z}$-modules). The patterns are: coproducts and products are quite different in general, and products are easier to understand than coproducts.

**Example 3.1** (Fiber products and coproducts)**.** Let

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

[10] If this seems vague, that's all right; it's possible to define and find forgetful functors more formally.

[11] The equivalence is given by any inclusion functor $* \to \mathsf{C}^{\text{initial}}$, and in the other direction by projecting down onto the point.

Limits across $I$ are called *fiber products*, and are terminal of objects fitting into the diagram

$$\begin{tikzcd} \varprojlim A_i \arrow[r] \arrow[d] & A_1 \arrow[d, "f"] \\ A_2 \arrow[r, "g"] & A_3. \end{tikzcd} \tag{3.1}$$

The fiber product is denoted $A_1 \times_{A_3} A_2$. In Set, these exist,[12] and for (3.1), is given by $A_2 \times_{A_3} A_2 = \{a_1, a_2 \mid f(a_1) = g(a_2)\}$.

The colimit of $I$ is $A_3$, since everything maps through $A_3$. This can be made more general; if a poset $P$ has a maximal element $m$, then $\varinjlim_P A_i = A_m$, and an analogous statement holds for minimal elements and limits. In fact, a cocone on a diagram is the addition of a maximal object; a colimit is trying to be the maximum of your diagram (which might not exist, but often does), and a limit is trying to be the minimum of your diagram.

The proper way to dualize this is to take colimits across

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

In this case, the colimit is called the *pushout*, and fits into the following diagram.

$$\begin{tikzcd} A_1 \arrow[r, "f"] \arrow[d, "g"'] & A_2 \arrow[d] \\ A_3 \arrow[r] & \varinjlim_I A_i. \end{tikzcd}$$

This is denoted $A_2 \amalg_{A_1} A_3$; in the category of sets, this is $A_2 \amalg A_3/\sim$, where $a_2 \sim a_3$ if there's an $a_1 \in A_1$ such that $f(a_1) = a_2$ and $g(a_1) = a_3$. Equivalence relations are a little harder to understand. And this isn't the pushout in other categories: in groups, the pushout is $G *_K H$, called the *free product with amalgamation*.

**Example 3.2** (Kernels and cokernels)**.** Suppose $\mathsf{C}$ is a category with a zero object 0 (so 0 is both initial and terminal). For any $A, B \in \mathsf{C}$, there's a unique map $0 : A \to B$ called the *zero map*, since there's a unique map $A \to 0$ and a unique map $0 \to B$, so composing them gives us the zero map.

Given any other $\varphi : A \to B$, we want to compare it with 0, so we're taking the (co)limit of the diagram $A \underset{\varphi}{\overset{f}{\rightrightarrows}} B$. The limit is called the *kernel*, denoted $\ker \varphi$, and the colimit is called the *cokernel*, denoted $\operatorname{coker} \varphi$.[13] Another way to think of (co)kernels is as fiber products and pushouts: they fit into the diagrams

$$\begin{tikzcd} \ker \varphi \arrow[r] \arrow[d] & A \arrow[d, "\varphi"] \\ 0 \arrow[r] & B \end{tikzcd} \qquad \begin{tikzcd} A \arrow[r] \arrow[d, "\varphi"] & 0 \arrow[d] \\ B \arrow[r] & \operatorname{coker} \varphi, \end{tikzcd}$$

and this may make their non-categorical constructions more clear.

These examples are useful in algebra, but now we also know that they're unique up to unique isomorphism, which can be quite useful. It's incredible how often these come up in algebra. It's also worth remembering that (co)limits tend to play well with (co)limits, in a way that can be made precise, but provides some useful intuition about what might be true.

[12] In fact, all limits exist in the category of sets. There are some set-theoretic issues involved in the proof, but we're not going to worry about that.

[13] This is an example of a more general construction, where one considers the diagram $f, g : A \rightrightarrows B$ for more general $f$ and $g$; the limit is called the *equalizer*, and the colimit is called the *coequalizer*.

**Example 3.3** (Completion)**.** We were also going to do algebraic geometry at some point, and one interesting algebraic construction that has a geometric analogue is *completion*: if $R$ is a ring and $I \subset R$ is an ideal, then the completion of $R$ at $I$, denoted $\widehat{R}_I$, is the limit of the diagram

$$\cdots \longrightarrow R/I^3 \longrightarrow R/I^2 \longrightarrow R/I.$$

When $R = \mathbb{Z}$ and $I = (p)$, this is the ring of *p-adic integers*, denoted $\mathbb{Z}_p$ or $\widehat{\mathbb{Z}}_{(p)}$.

In the category of sets, one can explicitly write down what the limit is, as a subset of the product:

$$\varprojlim_I A = \left\{ (a_i)_{i\in I} \in \prod_{i\in I} A_i \mid A(m)(a_i) = a_j \text{ for all } m : i \to j \in I \right\}.$$

This requires proof, but is on the homework.

Colimits in general are harder, but some colimits are easy, such as an increasing union of sets $U_1 \subset U_2 \subset U_3 \subset \cdots$. In this case, the colimit (in the category of sets) is just the union of all of these. A good example of these (albeit in a different category) is localization. $p^{-\infty}\mathbb{Z}$ is the localization $S^{-1}\mathbb{Z}$, where $S = \langle p \rangle = \{1, p, p^2, \dots\}$. Since we know this sits inside $\mathbb{Q}$, this is an increasing union of sets

$$\mathbb{Z} \cup p^{-1}\mathbb{Z} \cup p^{-2}\mathbb{Z} \cup \cdots .$$

This means we can write it as a colimit:

$$p^{-\infty}\mathbb{Z} = \varinjlim \left( \mathbb{Z} \longrightarrow p^{-1}\mathbb{Z} \longrightarrow p^{-2}\mathbb{Z} \longrightarrow \cdots \right). \tag{3.2}$$

This colimit takes place in the category $\mathsf{Ab}$ of abelian groups, also known as the category of $\mathbb{Z}$-modules. However, as $\mathbb{Z}$-modules, $p^{-1}\mathbb{Z} \cong \mathbb{Z}$, where $1/p \mapsto 1$. In other words, (3.2) is isomorphic to the diagram

$$p^{-\infty}\mathbb{Z} = \varinjlim \left( \mathbb{Z} \xrightarrow{\cdot p} \mathbb{Z} \xrightarrow{\cdot p} \mathbb{Z} \xrightarrow{\cdot p} \cdots \right),$$

and this makes sense in more generality, in particular when we don't have something like $\mathbb{Q}$ as a reference. In particular, for any ring $R$ and $r \in R$, we can take the limit as $R$-modules

$$r^{-\infty}R = \varinjlim \left( R \xrightarrow{r} R \xrightarrow{r} R \xrightarrow{r} \cdots \right).$$

If $R$ is a domain, then this sits inside its field of fractions, but otherwise we don't have a reference point. And we can start the construction with an arbitrary $R$-module $M$, defining $r^{-\infty}M$ as

$$r^{-\infty}M = \varinjlim \left( M \xrightarrow{r} M \xrightarrow{r} M \xrightarrow{r} \cdots \right).$$

In algebraic topology, there's a notion of a spectrum, which is an infinite sequence of topological spaces. People say this is a lot of machinery with the nebulous goal of inverting the suspension functor, but this is a very similar idea: we want to invert $r$ as many times as we can, so we have to string it out as an infinite sequence. Though this construction may look big, it has a simple purpose, which is useful to keep in mind. Localization is also given by a colimit, which we'll see in the exercise. It was already given by a universal property, but this nicer kind of universal property gives us some more information.

All of these "nice" colimits are, more precisely, examples of a notion called filtered colimits. These are the analogues to Cauchy sequences: we know the limit exists if we have this condition, and it gives us nicer comparisons of elements later on in the sequence.

**Definition.**

- A poset $S$ is *filtered* if for all $x, y \in S$, there's a $z \in S$ majorizing $x$ and $y$, i.e. $z \geq x$ and $z \geq y$.
- A (small nonempty) category $\mathsf{C}$ is *filtered* if for all $x, y \in \mathsf{C}$, there's a $z \in \mathsf{C}$ and maps $x \to z$ and $y \to z$ and any two maps $f, g : x \rightrightarrows y$ have a coequalizer $h : y \to z$ (i.e. $f \circ h = g \circ h$).[14]

[14] Another way to think of this is the following: a poset is filtered if every finite subset has a maximum, and a category is filtered if every finite subcategory has a cone (a maximal element). Then, this guarantees nice things about infinite cones.

Notice that a finite filtered poset necessarily has a maximum, so this only becomes interesting in the infinite case.

The upshot is: filtered colimits exist, and they tend to have nice properties. For example, localizations are filtered, and (3.2) can be seen to match the definition explicitly, and increasing unions are filtered. Moreover, forgetful functors preserve filtered colimits. However, nontrivial finite colimits (such as pushouts) will not be filtered.

In the category of sets, one can give a construction for filtered colimits: if $I$ is a filtered category,

$$\varinjlim_I A_i = \coprod_I A_i / \sim,$$

where $a \sim b$ if they're eventually equivalent, i.e. if $a \in A_1$ and $b \in A_2$, then there's an $A_3$ in the diagram and maps $A_1 \to A_3$ and $A_2 \to A_3$ that map $a$ and $b$ to the same element.

In algebra, there are lots of statements like "localizations of direct sums are direct sums of localizations." This is true because both are colimits, and colimits play well with other colimits (though this does depend on the precise formulation of that principle). Similarly, completions of products are products of completions, because limits play well with limits. However, completions of direct sums might not do what you expect, nor localizations and products.

**Yoneda's lemma.**

*"The Force is everywhere; it surrounds us and binds us." – Yoda*

This is a slightly more mystical part of the class: we want to describe things not as they are, but as they are detected by things around them.

In fact, we get a surprising and powerful analogy from analysis: a category is much like an inner product space, where the objects of $\mathsf{C}$ are vectors, and the inner product is $A, B \mapsto \mathrm{Hom}(A, B)$. However, unlike inner products, this is not symmetric! This can be strange. The Yoneda lemma says, in this sense, that this pairing is nondegenerate: we can understand a "vector" completely by pairing it with other "vectors."

If $\mathsf{C}$ and $\mathsf{D}$ are categories, we can define the *functor category* $\mathsf{Fun}(\mathsf{C}, \mathsf{D})$, whose objects are (covariant) functors $\mathsf{C} \to \mathsf{D}$, and whose morphisms are natural transformations.

To a vector space $V$, we define the dual space $V^* = \mathrm{Hom}(V, k)$; the inner product structure defines a map $V \to V^*$, which is an isomorphism when the inner product is nondegenerate. This nondegeneracy is somewhat weak, and in fact feels more like the sense of distributions: if $V = C_c^\infty(\mathbb{R})$, its dual space $V^* = \mathrm{Dist}(\mathbb{R})$, the linear functionals on compactly supported, smooth functions. They're not isomorphic, but there is an embedding: any compactly supported smooth function defines a distribution.[15] Distributions are nice, because they're closed under lots of operations, so you can take your PDE or whatever and solve it in the distributional sense, and then try to get a regularity result showing it was in the original space the whole time.

In category theory, we're going to do something similar. For any $X \in \mathsf{C}$, let $h_X = \mathrm{Hom}_{\mathsf{C}}(-, X)$ and $h^X : \mathrm{Hom}_{\mathsf{C}}(X, -)$. These are functors $\mathsf{C}^{\mathsf{op}} \to \mathsf{Set}$ and $\mathsf{C} \to \mathsf{Set}$, respectively, e.g. $h_X : Y \mapsto \mathrm{Hom}_{\mathsf{C}}(Y, X)$. This is functorial because a map $Y \to Z$ induces a map $h_X(Z) \to h_X(Y)$ by pullback, which is contravariant, and composition is covariantly functorial for $h_X$. These are called the *functors (co)represented by* $X$.

Additionally, if $f : X \to X'$, then any map $Y \to X$ induces a map $Y \to X'$ by precomposing with $f$. In other words, $h_X$ is functorial in $X$! This defines a functor $h_- : \mathsf{C} \to \mathsf{Fun}(\mathsf{C}^{\mathsf{op}}, \mathsf{Set})$ sending $X \mapsto (Y \mapsto \mathrm{Hom}_{\mathsf{C}}(Y, X))$. This is weird and strange, but it's exactly like the embedding of a vector space into its dual.

We'll let $\widehat{\mathsf{C}} = \mathsf{Fun}(\mathsf{C}, \mathsf{Set})$.

**Lemma 3.4** (Yoneda)**.** *$h : \mathsf{C} \hookrightarrow \widehat{\mathsf{C}}$ is a full embedding.*

That is, for any $X, X' \in \mathsf{C}$, $\mathrm{Hom}_{\widehat{\mathsf{C}}}(h_X, h_{X'}) = \mathrm{Hom}_{\mathsf{C}}(X, X')$: we don't lose any information passing to $\widehat{\mathsf{C}}$. Or in other words, if you know all maps into $X$, then you know $X$.

For example, suppose we have a map of functors $\varphi : h_X \to h_{X'}$ in $\widehat{\mathsf{C}}$. This is a natural transformation, so for any $Y$, there's a map $h_X(Y) \to h_{X'}(Y)$ in a natural way. To prove the lemma, we want to construct a map $\psi : X \to X'$ which induces $\varphi$. So, how do we get such an element $\psi \in \mathrm{Hom}(X, X')$?

The only map we always have an any category is the identity, so let's look at $\mathrm{id}_X$. The natural transformation $\varphi$ induces a map $h_X(X) \to h_X(X')$, i.e. $\mathrm{Hom}(X, X) \to \mathrm{Hom}(X, X')$, so let $\psi = \varphi(\mathrm{id}_X)$. This

[15] If you haven't seen distributions, this is not really necessary to understand Yoneda's lemma.

assignment is a map $\mathrm{Hom}_{\widehat{\mathsf{C}}}(h_X, h_{X'}) \to \mathrm{Hom}_{\mathsf{C}}(X, X')$, and you can check this is the inverse to the map in the other direction. All this is doing is a little tautological, and as such it takes some time to sink in.

Episode IV.

## The Yoneda Chronicles, II: 1/28/16

*"I just like this stuff, sorry."*

Last time, we were talking about the Yoneda embedding; it's kind of strange, and you have to stare at it for a bit to get it. The analogy is that if $V$ is an inner product space, the map $v \mapsto \langle v, -\rangle$ defines an embedding $V \hookrightarrow V^*$ if the inner product is positive definite, so that $\langle v, -\rangle$ is nonzero (because $v \cdot v \neq 0$). The Yoneda embedding is sort of the same thing, but for a category $\mathsf{C}$ and its dual $\widehat{\mathsf{C}} = \mathsf{Fun}(\mathsf{C}^{\mathsf{op}}, \mathsf{Set})$. There's a contravariant functor $\mathsf{C} \to \widehat{\mathsf{C}}$ sending $X \mapsto h_X = \mathrm{Hom}(-, X)$, and the Yoneda lemma is that this is an embedding, or more precisely, a fully faithful functor: $\mathrm{Hom}_{\mathsf{C}}(X, Y) = \mathrm{Hom}_{\widehat{\mathsf{C}}}(h_X, h_Y)$. If you think of these as inner products, this is a "partial isometry:" there's an isometry onto the image.

The analogue of positive definiteness is that $\mathrm{id}_X \in \mathrm{Hom}_{\mathsf{C}}(X, X)$, so it must be nonempty. Then, we can transfer it around, enabling us to construct a map $X \to Y$ given a natural transformation $\varphi : h_X \to h_Y$, just by applying $\varphi$ to $\mathrm{id}_X$. Then, you can check that this is inverse to the map $\mathrm{Hom}_{\mathsf{C}}(X, Y) \to \mathrm{Hom}_{\widehat{\mathsf{C}}}(h_X, h_Y)$.

From the vector-spatial view, it's perhaps less surprising that you can understand the objects in a category in terms of the maps into them, but it's an extremely useful viewpoint: there are lots of operations you can perform in $\widehat{\mathsf{C}}$ (analogous to all the cool things you can do with distributions): for example, $\widehat{\mathsf{C}}$ has all limits and colimits. Then, you can try to understand how a construction in $\widehat{\mathsf{C}}$ relates to $\mathsf{C}$, which is made much nicer since $\mathsf{C}$ sits inside of $\widehat{\mathsf{C}}$.

One example is, if $X$ is a topological space, there's *functor of points* $h_X : \mathsf{Top}^{\mathsf{op}} \to \mathsf{Set}$ sending $Y \mapsto \mathrm{Hom}(Y, X)$. This captures a lot of the information of $X$: for example, the underlying set of $X$ is captured by $\mathrm{Hom}(*, X)$; paths are given by $\mathrm{Hom}(\mathbb{R}, X)$, and so on. In this setting, the Yoneda embedding tells us something that feels a little tautological: if you know all of the maps into $X$, you know $X$. This is not minimal by any means (and in practice, you end up using a less absurd amount of data), but it's a nice perspective, courtesy of abstract nonsense. Using it, we can translate questions about a category $\mathsf{C}$ into questions about the category of sets.

Given a functor $\mathsf{Top}^{\mathsf{op}} \to \mathsf{set}$, one might wonder whether it's $h_X$ for some $X$. This is the question of *representability*, one of the fundamental things in Grothendieck's worldview (a space is really a collection of maps into it), and we'll develop some ways to approach this question.

For example, what does it mean for a map $f : X \to Y$ to be injective in $\mathsf{C}$? There's an abstract categorical definition.

**Definition.** Let $f : X \to Y$ in $\mathsf{C}$. Then, $f$ is a *monomorphism* if whenever $g_1, g_2 : Z \rightrightarrows X$ and $f \circ g_1 = f \circ g_2$, then $g_1 = g_2$. A monomorphism is often denoted $f : X \hookrightarrow Y$.

The idea is: we care about $X$ as the maps into it, so if a map out of $X$ preserves all the information about maps into $X$, then it's analogous to injective.

**Definition.** Dually, an *epimorphism* $f : X \to Y$ in $\mathsf{C}$, written $f : X \twoheadrightarrow Y$, is a map such that whenever $g_1, g_2 : Y \rightrightarrows Z$ and $g_1 \circ f = g_2 \circ g$, then $g_1 = g_2$.

The Yoneda embedding shows up as follows.

**Lemma 4.1.** *$f : X \to Y$ is a monomorphism iff $h_f : h_X \to h_Y$ is pointwise injective, i.e. for every $Z \in \mathsf{C}$, $h_X(Z) \hookrightarrow h_Y(Z)$. Similarly, $f$ is an epimorphism iff $h_f : h_X \to h_Y$ is pointwise surjective.*

So we can take this strange notion of monomorphism (or epimorphism) and translate it into something nice. In functional analysis, functions have nice linear properties induced pointwise from $\mathbb{R}$, and similarly, here, morphisms can make use of the nice structure of Set.

For example, all limits and colimits exist in $\mathsf{Fun}(\mathsf{C}^{\mathsf{op}}, \mathsf{Set})$, like in the category of sets. What does this mean? (Yes, it's pretty crazily abstract.) A diagram $I \to \widehat{\mathsf{C}}$ is a diagram of functors with natural transformations between them. Then, we can define a new functor "$\varinjlim F_i$" sending an $X \in \mathsf{C}$ to $\varinjlim_I F_i(X)$ (which exists, because this limit is in Set). You should check that this is well-defined as a functor, and has the right universal property for the colimit, so we can remove the quotation marks; it's really the colimit. The point is: colimits in an abstract category might be weird or hard to define, but we know what they're like in sets, which is nice. And the same thing works for limits; the analogy is that addition or scalar multiplication of $\mathbb{R}$-valued functions on a space are done pointwise: for these functors, we're doing everything with the values of the functors. So $\mathsf{Fun}(\mathsf{C}^{\mathsf{op}}, \mathsf{Set})$ is like this nice promised land, but we need to know how to relate it to questions in $\mathsf{C}$.

To understand this, let's talk about Hom. For any category $\mathsf{C}$ and $Z \in \mathsf{C}$, we have a functor $\mathrm{Hom}_{\mathsf{C}}(Z, –) : \mathsf{C} \to \mathsf{Set}$. This functor preserves limits: the analogy is that maps into a subspace of a given vector space are a subset,[16] so $\mathrm{Hom}(V, U) \subset \mathrm{Hom}(V, W)$. That is, "maps into a subspace is a subspace of maps." And since limits are sort of like subspaces, this can be a mneomnic for $\mathrm{Hom}(Z, –)$ preserving limits.

Things here aren't hard, just unwinding notation. The maps $\mathrm{Hom}(Z, \varprojlim A_i)$ is a cone on the diagram of the $A_i$: it comes with maps $Z \to A_i$ compatible with the directed maps $A_i \to A_j$ — and we said that compatible collections are exactly what limits are in the category of sets, so this is $\varprojlim \mathrm{Hom}(Z, A_i)$. That $\mathrm{Hom}(Z, –)$ preserves limits is very important, and we will use it many times.

One might wonder about $\mathrm{Hom}_{\mathsf{C}}(–, Z)$, but this is just $\mathrm{Hom}_{\mathsf{C}^{\mathsf{op}}}(Z, –)$, so we see that $\mathrm{Hom}_{\mathsf{C}}(–, Z)$ sends colimits to limits, since it's a contravariant functor. Thus, $\mathrm{Hom}_{\mathsf{C}}(\varinjlim A_i, Z) = \varprojlim \mathrm{Hom}(A_i, Z)$. The mneomnic is that maps out of a quotient $V/U \to W$ are a subspace of maps $V \to W$ (those vanishing on $U$).

This may feel like symbol gymnastics, but we're almost done with the Yoneda embedding for a long time. Here's the final result.

**Corollary 4.2.** *The Yoneda embedding* $\mathsf{C} \hookrightarrow \mathsf{Fun}(\mathsf{C}^{\mathsf{op}}, \mathsf{Set})$ *preserves limits.*

This is, again, chasing symbols: $h_{\varprojlim A_i} = \mathrm{Hom}(–, \varprojlim A_i) = \varprojlim \mathrm{Hom}(–, A_i) = \varprojlim h_{A_i}$. This is another instance of the mantra that limits are easy: you can calculate limits in any category in terms of limits of sets.

For colimits, this is completely false; this might initially seem bad, but it's actually something good. We have some word of affine schemes, which we still don't get geometrically (we will, don't worry), but categorically is $\mathsf{Ring}^{\mathsf{op}}$. Using the Yoneda embedding, we get a functor of points

$$\begin{tikzcd}
\mathsf{AffSch} \arrow[r, equal] \arrow[d, hook] & \mathsf{Ring}^{\mathsf{op}} \arrow[d, hook] \\
\mathsf{Fun}(\mathsf{AffSch}^{\mathsf{op}}, \mathsf{Set}) \arrow[r, equal] & \mathsf{Fun}(\mathsf{Ring}, \mathsf{Set}).
\end{tikzcd}$$

We will be defining schemes by gluing together affine schemes, which is a kind of colimit. Hence, it's helpful that we don't preserve colimits, so we get nontrivial schemes. In other words, *a scheme is a functor* $\mathsf{Ring} \to \mathsf{Set}$ *with certain properties*. This is useful, because not all spaces have enough functions out of them, so they're not captured by $\mathsf{Ring}^{\mathsf{op}}$, and we need to pass to the functor category.

**Adjoint Functors.** Again, the analogy to vector spaces will be instructive: if $\varphi : V \to W$ is linear, then there's an adjoint map $\varphi^\dagger : W^* \to V^*$, corresponding to matrix transpose. But if $V$ and $W$ are inner product spaces, then the isomorphisms $V \cong V^*$ and $W \cong W^*$ allow us to realize $\varphi^\dagger$ as a map $W \to V$, and the key property is that for any $v \in V$ and $w \in W$, $\langle \varphi(v), w \rangle_W = \langle v, \varphi^\dagger(w) \rangle_V$; this is enough to completely characterize the adjoint. These are very useful, because they're in a way the closest thing to an inverse: a map $\varphi : V \to W$ factors through an isomorphism $(\ker V)^\perp \to \mathrm{Im}(\varphi)$, and the adjoint $\varphi^\dagger : \mathrm{Im}(\varphi) \to (\ker \varphi)^\perp$ *is* the inverse to $\varphi$!

Now, we're going to do the same thing with categories, with $\langle \cdot, \cdot \rangle$ replaced with Hom again. But since this isn't symmetric, we have left- and right-flavored adjoints.

**Definition.** Let $\mathsf{C}$ and $\mathsf{D}$ be categories and $F : \mathsf{C} \to \mathsf{D}$ and $G : \mathsf{D} \to \mathsf{C}$ be functors. Then, $(F, G)$ is an *adjoint pair* (order matters: $F$ is *left adjoint* to $G$ and $G$ is *right adjoint* to $F$) if there exists a natural isomorphism $\mathrm{Hom}_{\mathsf{D}}(F(–), –) \leftrightarrows \mathrm{Hom}_{\mathsf{C}}(–, G–)$. In other words, for every $X \in \mathsf{C}$ and $Y \in \mathsf{D}$, there's an isomorphism $\mathrm{Hom}_{\mathsf{D}}(FX, Y) = \mathrm{Hom}_{\mathsf{C}}(X, GY)$ that's functorial in both $X$ and $Y$.

[16] If this sounds dumb, remember that maps into a quotient are not a quotient.

There are other ways to rewrite this; the Wikipedia article is pretty good. For example, out of this structure there's a canonical map $\eta_X \in \mathrm{Hom}(X, GFX)$ (which doesn't have an obvious analogue in the world of vector spaces): this is the same as $\mathrm{Hom}(FX, FX)$, so let $\eta_X$ be the map corresponding to $\mathrm{id}_{FX}$. More precisely, there's a natural transformation $\eta : \mathrm{id}_{\mathsf{C}} \to GF$. In the same way, there's a natural transformation $\varepsilon : FG \to \mathrm{id}_{\mathsf{D}}$ given by pulling back the identity map.

Sometimes, having $\eta$ and $\varepsilon$ is more convenient than the standard definition of adjointness, so one can start with natural transformations $\eta : \mathrm{id}_{\mathsf{C}} \to FG$ and $\varepsilon : GF \to \mathrm{id}_{\mathsf{D}}$. Then, it's a theorem that if they satisfy the "mark of Zorro" axiom, that the following diagram commutes, where the first map adds $GF$ on the right by $\eta$, and the second map collapses $FG$ on the left by $\varepsilon$.

$$\begin{tikzcd} F \arrow[r] \arrow[rr, bend right, "\mathrm{id}"'] & FGF \arrow[r] & F \end{tikzcd}$$

What are adjoints used for? Everything, everywhere.

**Example 4.3** (Free and forgetful functors)**.** There's a pair of functors $\mathrm{Free} : \mathsf{Set} \to \mathsf{Grp}$ and $\mathrm{For} : \mathsf{Grp} \to \mathsf{Set}$. This is an adjunction, because a map out of a free group is determined exactly by where its generators go, so if $G$ is a group and $S$ is a set, then $\mathrm{Hom}_{\mathsf{Grp}}(\mathrm{Free}(S), G) = \mathrm{Hom}_{\mathsf{Set}}(S, \mathrm{For}(G))$.

We can generalize this: there are lots of forgetful functors, and we can define free functors as their left adjoints; in this way one realizes the usual definition of free abelian groups, for example.

Another example: there's a forgetful functor $\mathrm{For} : \mathsf{Mod}_R \to \mathsf{Set}$, and the notion of a free $R$-module is a left adjoint $\mathrm{Free} : \mathsf{Set} \to \mathsf{Mod}_R$, because $\mathrm{Hom}_{\mathsf{Mod}_R}(\mathrm{Free}(S), M) = \mathrm{Hom}_{\mathsf{Set}}(S, \mathrm{For}(M))$ for any set $S$ and $R$-module $M$. This is because a free$R$-module on a set $S$ is $R^S$ (the direct sum), and so the images of the generators are exactly what determine a map out of it.

But we've talked about functors to sets before: is For representable? A map $R \to M$ is determined by where it sends $M$: $\mathrm{For}(M) = \mathrm{Hom}_{\mathsf{Mod}_R}(R, M)$, so For is represented by $R$!

This is a special case of the most important adjunction.

**Example 4.4.** Let $R$ be a ring and $\mathsf{C} = \mathsf{Mod}_R$. We know that $\mathrm{Hom}_R(M, N)$ isn't just a set, but is naturally an $R$-module (you can add and multiply maps pointwise). Since we've been using Hom to denote sets, then we'll let *inner Hom* $\underline{\mathrm{Hom}}_R(M, N)$ denote the Hom as an $R$-module.

Thus, we've defined a functor $\underline{\mathrm{Hom}}_R(M, –) : \mathsf{Mod}_R \to \mathsf{Mod}_R$. Does it have a left adjoint?[17] That is, we need to look at $\mathrm{Hom}(N, \underline{\mathrm{Hom}}_R(M, P))$, whose elements send $n \in N$ to an $R$-linear map $M \to P$. We can recast these as maps $M \times N \to P$, which must be $R$-linear in both $M$ and $N$.

This may be looking familiar: we're looking for $R$-bilinear maps $M \times N \to P$ (that is, $\varphi(rm, n) = \varphi(m, rn) = r\varphi(m, n)$). And there is a universal object through which these factor through, the tensor product $M \otimes_R N$. By definition, bilinear maps $M \times N \to P$ correspond to linear maps $M \otimes_R N \to P$. You do have to construct it to show that it exists: it's the span of symbols $m \otimes n$, modded out by the equivalence relation $rm \otimes n = m \otimes rn$ (you can move scalars across the middle). There are a bunch of things to implicitly check here, some of which will be exercises for us.

The point is, this universal property is saying that $\mathrm{Hom}(N, \underline{\mathrm{Hom}}_R(M, P)) = \mathrm{Hom}(M \otimes_R N, P)$. Thus, $(M \otimes –, \mathrm{Hom}_R(M, –))$ is an adjoint pair! This is the real definition of the tensor product. If $R = \mathbb{Z}$, so we just have the category of abelian groups, then the tensor product will be written $M \otimes N$.

One useful thing to check is that if $S$ is an $R$-algebra, the map $R \to S$ defines an $S$-module structure on $S \otimes_R M$. That is, we have a functor $S \otimes_R – : \mathsf{Mod}_R \to \mathsf{Mod}_S$. As one example, if $R = \mathbb{Z}$, then this specializes to $S$ being any ring, and $A \mapsto A \otimes S$ sends $A$ to the free module $S \otimes A$.

Tensor products are always left adjoint, and this one also has a right adjoint: given an $S$-module $P$ and a map $R \to S$ (so $S$ is an $R$-algebra), then forgetting to the $R$-module structure is functorial, and $(S \otimes_R –, \mathrm{For})$ is another adjoint pair. This is why $S \otimes A$ is regarded as free.

**Example 4.5.** Now, suppose $S$ and $T$ are both $R$-algebras; then, $S \otimes_R T$ is more than just an $S$- or $T$-module; in fact, it's a ring (with $R$-, $S$-, and $T$-algebra structures). Over $\mathbb{Z}$, this specializes to the statement that the tensor product of two rings is still a ring. As a silly example, let $X$ be a set of $m$ points and $Y$ be a set of $n$ points. Then, $S = \mathbb{C}[X]$, the functions on $X$, is $\mathbb{C}^m$, a commutative ring with multiplication defined

[17] It turns out this does not have a right adjoint, which isn't too hard to convince yourself of.

pointwise. Similarly, $T$ is functions on $Y$, so $T \cong \mathbb{C}^n$. Thus, $S \otimes_{\mathbb{C}} T = \mathrm{Mat}_{m,n}$: there's a complex number for every pair $(m, n) \in X \times Y$. The multiplication is the silly one, pointwise multiplication (i.e. the one you told your linear algebra students to never, ever do), because this is the ring of functions on $X \times Y$, rather than usual matrix multiplication. This might be a little more motivation for this next statement.

**Exercise.** $S \otimes_R T$ is the coproduct $S \amalg T$ in the category of $R$-algebras, and the pushout $S \amalg_R T$ in the category of rings. In other words, it fits into the following diagram.

$$\begin{tikzcd}
R \arrow[r] \arrow[d] & S \arrow[d] \\
T \arrow[r] & S \otimes_R T
\end{tikzcd}$$

This may be confusing, because the coproduct of modules is direct sum. But the example with sets of points will be true more generally: in nice situations, $\mathrm{Fun}(X \times Y) \cong \mathrm{Fun}(X) \otimes \mathrm{Fun}(Y)$. Or in other words, $\mathrm{Hom}_{\mathsf{Ring}}(S \otimes T, U) = \mathrm{Hom}_{\mathsf{Ring}}(S, U) \times \mathrm{Hom}_{\mathsf{Ring}}(T, U)$, and there's a version with $R$-algebras as well.

Now, since affine schemes are the opposite category to the category of rings, then we know that $\mathrm{Spec}\, R \times \mathrm{Spec}\, S = \mathrm{Spec}(R \otimes S$: functions on $X \times Y$ are the tensor product of those on $X$ and those on $Y$. Strangely, though we know what products of affine schemes are, we don't know what affine schemes are yet.

Episode V.

## Schemes: 2/2/16

Before we delve into the world of schemes, we have just a little more to say about adjoints.

Recall that an adjoint pair $F : \mathsf{C} \rightleftarrows \mathsf{D} : G$ is the data of a natural isomorphism $\mathrm{Hom}_{\mathsf{C}}(M, GN) \cong \mathrm{Hom}_{\mathsf{D}}(FM, N)$ for all $M \in \mathsf{C}$ and $N \in \mathsf{D}$. One important example was the adjoint (Free, Forget): free functors lie on the left, because they're very easy to map out of (just specify where the generators go). The other important example may be an instance of the first example: $(\otimes_R, \underline{\mathrm{Hom}}_R)$, because $\mathrm{Hom}_{\mathsf{Mod}_R}(M, \underline{\mathrm{Hom}}_R(N, P))$ is the $R$-bilinear maps $M \times N \to P$, but this is $\mathrm{Hom}_R(M \otimes_R N, P)$, by the universal property for tensor product. So tensor products can be thought of as a free construction; another example of this is that given a map $R \to S$, the functor $S \otimes_R - : \mathsf{Mod}_R \to \mathsf{Mod}_S$ is left adjoint to the forgetful functor from $S$-modules to $R$-modules.

Finally, the last piece of abstract nonsense we'll discuss is the relation between adjoints and limits. If $I$ is an index category, then the $I$-shaped diagrams in a category $\mathsf{C}$ (the functors $I \to \mathsf{C}$) are also a category, the functor category $\mathsf{Fun}(I, \mathsf{C})$. This is also denoted $\mathsf{C}^I$.

There's a natural functor $\Delta : \mathsf{C} \to \mathsf{C}^I$ sending an $M \in \mathsf{C}$ to the diagram with $M$ at every index and the identity for every morphism, which of course commutes. Sometimes this is called the "stupid diagram," or more formally the *diagonal diagram* or *constant diagram*.

Every time you see a functor, your first question should be, *does it have an adjoint*? We can check on the left or on the right, so suppose we have a left adjoint $\Delta^\ell : \mathsf{C}^I \to \mathsf{C}$. Writing the meaning of this is less confusing than drawing pictures: if we have an $A_\bullet \in \mathsf{C}^I$ specified by

$$A_\bullet = \begin{tikzcd}
A_1 \arrow[d] \arrow[dr] & \\
A_2 \arrow[r] & A_3,
\end{tikzcd}$$

then the left adjoint has the property that for any $B \in \mathsf{C}$, a map $\Delta^\ell(A_\bullet) \to B$ is the data

$$\begin{tikzcd}
A_1 \arrow[d] \arrow[dr] \arrow[rr, bend left] & & B \arrow[d, "\mathrm{id}"] \arrow[dr, "\mathrm{id}"] & \\
A_2 \arrow[r] \arrow[rr, bend right] & A_3 \arrow[rr, bend right] & B_2 \arrow[r, "\mathrm{id}"] & B_3,
\end{tikzcd}$$

but if we collapse the diagonal $B$-diagram, this is exactly a cone! Thus, $\Delta^\ell = \varinjlim$ (which, recall, may not always exist), and similarly, a right adjoint $\Delta^r$ to $\Delta$ is $\varprojlim$. In fact, there's also a way to realize adjoints as certain kinds of limits; then, the following proposition is just a consequence of the principle that "(co)limits commute with (co)limits."

**Proposition 5.1.** *Right adjoints (resp. left adjoints) commutes with limits (resp. colimits). That is, if $(F, G)$ is an adjunction, then $F(\varinjlim A_i) = \varinjlim F(A_i)$ and $G(\varprojlim(B_i) = \varprojlim G(B_i)$.*

*Proof.* There's nothing particularly tricky here. If the adjoints are on the categories $\mathsf{C}$ and $\mathsf{D}$, then for any $A \in \mathsf{C}$, consider $\mathrm{Hom}_{\mathsf{C}}(A, \varprojlim(G(B_i)))$. If we can show this is the same as $\mathrm{Hom}_{\mathsf{C}}(A, G(\varprojlim(B_i)))$, then the Yoneda embedding says that $\varprojlim G(B_i) = G(\varprojlim B_i)$: we can show two things are the same by showing the maps into them are the same.

First, we said last time that Hom commutes with limits, so $\mathrm{Hom}_{\mathsf{C}}(A, \varprojlim G(B_i)) = \varprojlim \mathrm{Hom}_{\mathsf{C}}(A, G(B_i)) = \varprojlim \mathrm{Hom}_{\mathsf{D}}(F(A), B_i)$ by the adjunction. But since Hom commutes with limits, then this can be rewritten as $\mathrm{Hom}_{\mathsf{D}}(F(A), \varprojlim B_i) = \mathrm{Hom}_{\mathsf{C}}(A, G(\varprojlim B_i))$, again by the adjunction.

Then, the proof for left adjoints is the same, but in the opposite category. ⊠

To formalize this, you'd want to say why it's functorial in $A$, but this isn't the core content of the proof.

Proposition 5.1 is useful everywhere. For example, we said that "forgetful functors preserve limits;" since forgetful functors are right adjoint to free functors, then they must preserve limits. In particular, products, fiber products, and kernels are all preserved by forgetful functors.

Another application: since localization is a colimit and $S \otimes_R -$ is a right adjoint, then it should commute with localization. In particular, there's a natural isomorphism $S^{-1}M \otimes_R N \cong S^{-1}(M \otimes_R N)$. In particular, $S^{-1}R \otimes_R M \cong S^{-1}(R \otimes_R M) = S^{-1}M$, since there's a natural isomorphism $R \otimes_R M = M$. That is, localization of modules, as a functor, is $S^{-1}R \otimes_R -$. Another way to see this is that there's a forgetful functor $\mathsf{Mod}_{S^{-1}R} \to \mathsf{Mod}_R$, and the left adjoint functor is $S^{-1}R \otimes_R -$, the localization functor.

So localization is a tensor product, and therefore a localization. Thus, it commutes with arbitrary colimits: for example, since direct sums are colimits, then $S^{-1}(\bigoplus M_i) \cong \bigoplus S^{-1}M_i$ canonically, and tensor products commute with arbitrary direct sums. Moreover, pushouts, cokernels, and coequalizers all pass through tensor products.

On the other hand, completion cannot be written as a tensor product; it's a limit. Thus, it does not necessarily commute with direct sums, etc.

**Introduction to Schemes.** We can't really define a scheme yet (we're missing a key ingredient), but we can still talk a lot about them. Remember that our plan was to associate an affine scheme $\operatorname{Spec} R$ to a ring $R$, in a way that is a contravariant equivalence of categories. This is quite a strong desideratum, and so we want a strong construction. We'll find this has the following three ingredients: a set of points, a topology, and a structure sheaf of functions. We'll discuss the set today, and then move to the other two later.

Each ingredient is very necessary: for example, if $k$ is a field, then $\operatorname{Spec} k$ will be a point. There's only one topology here, but there are many nonisomorphic fields, so the structure sheaf will have to do something interesting. Why is this $\operatorname{Spec} k$? A point is the terminal object in Set, and fields have no interesting ideals: every map $k \to R$ for a ring $R$ is necessarily injective, hence a monomorphism. Hence, all maps $\operatorname{Spec} R \to \operatorname{Spec} k$ should be epimorphisms in Set, hence surjective. This is not a proof, just an *ansatz*.

More generally, we'd like points in $\operatorname{Spec} R$ should correspond to maps $\mathrm{pt} \to \operatorname{Spec} R$, which will correspond to a ring homomorphism $R \to K$, for a field $k$. How do we organize these homomorphisms?

**Definition.** If $R$ is a ring, define the set $\operatorname{Spec} R$ to be the set of prime ideals[18] $\mathfrak{p} \subset R$.

One's first naïve idea of what you'd want is the set of maximal ideals, which is a subset (after all, a maximal ideal is a prime ideal), but if $\mathfrak{m} \subset R$ is maximal, that's the same as a surjection $R \twoheadrightarrow k$. But if $\mathfrak{p}$ is a prime ideal, then the surjection $R \twoheadrightarrow R/\mathfrak{p}$ is onto an integral domain. So prime ideals are surjections onto integral domains.

Wait, why are we talking about integral domains? An integral domain means exactly having a field of fractions: if $I$ is an integral domain, let $S = I \setminus 0$, which is multiplicative, so we get a field $S^{-1}I$, and an

[18]For the purposes of this definition, $R$ is not an ideal of itself; we're only looking at proper ideals.

injective map $I \hookrightarrow S^{-1}I$. And a subring of a field must be an integral domain (since fields have no zero divisors). Hence, integral domains are exactly the rings which are subrings of fields. Thus, prime ideals give maps to fields, even if they may not be injective: if $\mathfrak{p}$ is a prime ideal, then $R \twoheadrightarrow R/\mathfrak{p} \hookrightarrow \operatorname{Frac}(R/\mathfrak{p})$, and the composite map may not be surjective, but its image generates the field $\operatorname{Frac}(R/\mathfrak{p})$.

In other words, a prime ideal is the same as a homomorphism $R \to k$ which generates $k$ as a field. The field associated to a prime ideal is called its *residue field*. This is one reason why prime ideals are still somehow reasonable. One can also define an equivalence relation on maps from $R$ to fields (there are many of these, thanks to e.g. field extensions), and prime ideals represent equivalence classes. So one might think that "prime ideals of $R$ are the ways in which $R$ talks to fields."

Now, suppose $r \in R$ and $\mathfrak{p} \subset R$ is prime (we'll think of it as a point $x \in \operatorname{Spec} R$). Then, there's an evaluation map $r(x) = r \bmod \mathfrak{p} \in R/\mathfrak{p}$, or even inside $\operatorname{Frac}(R/\mathfrak{p})$. So we can think of $R$ as the set of "regular functions" on $\operatorname{Spec} R$. The codomain field of the function $r(x)$ depends on the point $x$, which is quite strange, but we'll eventually pin down precisely what such a function means; meanwhile, this issue is one of the main weirdnesses of schemes you'll have to work with at first.

Hence, if $k$ is a field, then $\operatorname{Spec} k = \{(0)\} = \mathrm{pt}$, and any $r \in k$ gives a $k$-valued function on the point $(0)$, which is $r(\mathrm{pt}) = r$. Moreover, if $R$ is the zero ring, then $\operatorname{Spec} R = \varnothing$; this makes sense, because 0 is terminal in the category of rings, and $\varnothing$ is initial in the category of sets.

**Example 5.2.** Let's have a more interesting example, $\mathbb{A}^1_{\mathbb{C}}$, the *affine line over* $\mathbb{C}$, defined to be $\operatorname{Spec} \mathbb{C}[x]$.[19]

The maximal ideals in $\mathbb{C}[x]$ are exactly the irreducible (nonconstant) polynomials $\langle (x-t) \rangle \subset \mathbb{C}[x]$, and a $t \in \mathbb{C}$ defines a function on them which is precisely evaluation at $t$. However, there's one more prime ideal, the zero ideal.[20]

**Lemma 5.3.** *0 and $\langle (x-t) \rangle$ for $t \in \mathbb{C}$ are all of the prime ideals of $\mathbb{C}[x]$.*

*Proof.* Suppose $\mathfrak{p} \subset \mathbb{C}[x]$ is prime and nonzero. Then, let $f \in \mathfrak{p}$ be a polynomial of minimal degree in $\mathfrak{p}$. Then, $f$ must be nonconstant (if it were constant, it would be invertible, so $\mathfrak{p} = \mathbb{C}[x]$, which isn't the case). However, the degree of $f$ must be 1: if $\deg f > 1$, then since $\mathbb{C}$ is algebraically closed, then $f$ has a root, so if $\deg f > 1$, then $f = gh$, with $\deg g, \deg h > 0$. Thus, $g \in \mathfrak{p}$ or $h \in \mathfrak{p}$, because $\mathfrak{p}$ is prime, but both of them have degrees less than that of $f$, which is a contradiction, so $\deg f = 1$.

Thus, $\mathfrak{p} \supset \langle f \rangle$: *a priori*, it could be bigger. We'll use the property that $\mathbb{C}[x]$ is a Euclidean domain, so we can do polynomial long division, so if $g \in \mathfrak{p}$, then $g = f \cdot m + r$, for some $m, r \in \mathbb{C}[x]$ with $\deg r < \deg f$. But since $f, g \in \mathfrak{p}$, then $r = g - fm \in \mathfrak{p}$ as well, but $\deg r < \deg f <$ so $\deg r = 0$, and therefore $f \mid g$, i.e. $g \in \langle f \rangle$. ⊠

So now we know what $\mathbb{A}^1_{\mathbb{C}}$ is as a set. One can draw a picture of it: for every $t \in \mathbb{C}$, there's a point $\langle (x-t) \rangle \in \mathbb{A}^1_{\mathbb{C}}$, so we have a bunch of point; then, we have the point corresponding to $(0)$, which is "bigger." In general, if $R$ is an integral domain, the point corresponding to $(0)$ in $\operatorname{Spec} R$ will be called the *generic point*. Then, the residue field associated to each point $t \in \mathbb{C}$ is $\mathbb{C}$ again, and for the zero ideal we get $\operatorname{Frac}(\mathbb{C}[x]/0) = \mathbb{C}(x)$, the rational functions in $\mathbb{C}$.

Since the proof of Lemma 5.3 only depended on $\mathbb{C}$ being an algebraically closed field, the above example works just as well for $\operatorname{Spec} k$, when $k$ is any algebraically closed field: for every $t \in k$ we have a point with residue field $k$, and then the generic point $(0)$ with residue field $k(x)$, the rational functions on $\operatorname{Spec} k[x] = \mathbb{A}^1_k$.

**Example 5.4** ($\operatorname{Spec} \mathbb{Z}$)**.** Since $\mathbb{Z}$ is initial in the category of rings, then $\operatorname{Spec} \mathbb{Z}$ will be final in the category of affine schemes. So it will behave as a point, even though it doesn't look at all like one. Having a good geometric object corresponding to $\mathbb{Z}$ was a major motivator for Grothendieck, and was a feature of the scheme-theoretic approach over others.

The picture is a point for every prime $p \in \mathbb{Z}$, with residue field $\mathbb{F}_p$, but also the zero ideal, corresponding to the generic point, whose residue field is $\mathbb{Q}$. This point ends up being dense once we define a topology on $\operatorname{Spec} \mathbb{Z}$, so $\operatorname{Spec} \mathbb{Z}$ is connected, which is nice. The intuition is that every rational number is a function at

[19] More generally, if $k$ is any field, the *affine line over* $k$ is $\mathbb{A}^1_k = \operatorname{Spec} k[x]$.

[20] The condition that 0 is a prime ideal is equivalent to a ring being an integral domain, so in these cases we do have a distinguished prime ideal.

all but finitely many points: $19/15 \in \mathbb{Q}$, so we can evaluate $(19/15)(7) = 5 \bmod 7$, and do this everywhere except 3 and 5, which are its "poles." (Its value at the generic point is $19/15$ again.)

Since we have a map $\mathbb{Z} \to \mathbb{Q}$, then we'd better have a nice map $\operatorname{Spec} \mathbb{Q} \to \operatorname{Spec} \mathbb{Z}$, corresponding to morphisms of residue fields. Since $\operatorname{Spec} \mathbb{Q}$ is a point, we can just send it to the generic point, whose residue field is $\mathbb{Q}$. This is why we need prime ideals (and generic points as a consequence); if we're trying to mimic ring theory, this is just necessary. Classical algebraic geometry tended to restrict itself to finitely generated algebras over an algebraically closed field, which means that we must miss out on some ring theory.

**Example 5.5.** We can also talk about $\mathbb{A}^1_{\mathbb{R}} = \operatorname{Spec} \mathbb{R}[x]$, or more generally $\mathbb{A}^1_k$ where $k$ is not algebraically closed. This means classifying the prime ideals of $\mathbb{R}[x]$; since $\mathbb{R}$ isn't algebraically closed, it's no longer true that every prime ideal contains a linear factor. We do have $(x-t)$ for $t \in \mathbb{R}$ and $(0)$ again, but since $\mathbb{R}[x]$ is again a Euclidean domain, then it's a PID. Thus, all of our ideals are $(f)$ for some $f \in \mathbb{R}[x]$, and $(f)$ is prime iff $f$ is an irreducible polynomial.[21] This means we have to classify monic irreducible polynomials.

Over a field $k$, a monic irreducible polynomial is given exactly by a Galois orbit in $\overline{k}$. For $\mathbb{R}$, $\overline{\mathbb{R}} = \mathbb{C}$, and $\mathrm{Gal}(\mathbb{R}/\mathbb{C})$ is a group of order 2, generated by complex conjugation. Thus, the orbits are two points $\{z, \overline{z}\}$, the complex conjugate roots of an irreducible quadratic in $\mathbb{R}[x]$.

Thus, the picture of $\mathbb{A}^1_{\mathbb{R}}$ has a copy of $\mathbb{R}$ as usual (points with residue field $\mathbb{R}$), with a generic point $(0)$ with residue field $\mathbb{R}(x)$, and an "upper half-plane" (which is not strictly true, since we're identifying points in $\mathbb{C}$, rather than taking a subset) of $\{z, \overline{z}\}$ with residue field $\mathbb{C}$ (since, for example, $\mathbb{R}[x](x^2+1) \cong \mathbb{C}$). The point is: for a field that's not algebraically closed, there are points in the affine line whose residue field is a nontrivial field extension.

Given any $f \in \mathbb{R}[x]$, we get a function on $\operatorname{Spec} \mathbb{R}[x]$: $f((x^2+1)) = f \bmod (x^2+1) \in \mathbb{C}$. This is a complex number, but evaluating at $(x-t)$, with $t \in \mathbb{R}$, gives you a real number. This is a little funny, but the takeaway is that we have these interesting new points, since $\mathbb{R}$ isn't algebraically closed.

A fun exercise is to draw $\mathbb{A}^1_{\mathbb{F}_p}$, because $\overline{\mathbb{F}}_p/\mathbb{F}_p$ isn't a finite field extension: each finite extension has a Galois group $\mathbb{Z}/p$, so there are $(p-1)$ Galois orbits at each stage. It's definitely a strange thing, and not what you would think of as a line: the point is that the construction of a scheme has extra points you might not expect, in order to make the connection between rings and schemes work.

[21] There's this nice set of inclusions fields $\subset$ Euclidean domains $\subset$ PIDs $\subset$ UFDs $\subset$ integral domains.