Algebraic Geometry

UT Austin, Spring 2016

M390C NOTES: ALGEBRAIC GEOMETRY

ARUN DEBRAY
FEBRUARY 23, 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 $f_1, \ldots, f_k \in \mathbb{C}[x_1, \ldots, x_n]$. These have a natural notion of function, *polynomial functions*, which are

[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.

polynomials in $\mathbb{C}[x_1, \ldots, x_n]$ restricted to X, If $I(X)$ is the functions vanishing on X, then these functions are given by $\mathbb{C}[x_1, \ldots, x_n]/I$.

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.

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* 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^{op} which has the same set, but reversed multiplication: $f \cdot_{\mathrm{op}} g = g \cdot f$. Similarly, given a category C, there's an *opposite category* C^{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 1.1. 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 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 η 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.

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

[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.

This is equivalent to having a natural inverse to η.

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.

Definition. Let C and 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 1.2. Let k be a field, and let $\mathsf{D} = \mathsf{fdVect}_k$, the category of finite-dimensional vector spaces and linear maps, and let 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 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 1.3. *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^∞ functions on N to C^∞ functions on M. The first step in this direction is the following theorem, sometimes called *Gelfand duality*.

Theorem 1.4 (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.

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

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.

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

---

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

[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.

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 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.

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 S and a relation $\leq$ on S that is

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

S has the structure of a category: the objects are the elements of S, and $\operatorname{Hom}(x, y)$ is $\{\mathrm{pt}\}$ if $x \leq y$ 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 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 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 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 C be a category.

- A *final* (or *terminal*) object in 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 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 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 Ab or 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 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 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 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 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 2.6. Show that $S^{-1}R$ is the initial object in 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 φ_T factors through $S^{-1}R$; this means that $S^{-1}R$ is the element of 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.

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

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* C.

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

Definition. Let A be an I-shaped diagram in 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 Cones_A,

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

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 2.7 (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 φ_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 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.

[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.

In linear categories, e.g. Ab, Mod_R, or 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.

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 2.8. 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 $\mathsf{C}^{\mathsf{initial}}$, 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.

[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 $* \to \mathsf{C}^{\mathsf{initial}}$, and in the other direction by projecting down onto the point.

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

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

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 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 $\begin{tikzcd} A \arrow[r, shift left, "f"] \arrow[r, shift right, "\varphi"'] & B \end{tikzcd}$. 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.

[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*.

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.

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 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 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]

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 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 C and 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.

[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.

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

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 φ. 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 id_X. The natural transformation φ 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 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 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 φ to 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 C, which is made much nicer since 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 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 C? There's an abstract categorical definition.

Definition. Let $f : X \to Y$ in 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 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 C.

To understand this, let's talk about Hom. For any category 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 φ!

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.

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

Definition. Let C and 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.

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 η_X be the map corresponding to 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 η and ε 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 η, and the second map collapses FG on the left by ε.

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

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.

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

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 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 4.6. $S \otimes_R T$ 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 $\operatorname{Spec} R \times \operatorname{Spec} S = \operatorname{Spec}(R \otimes S$: 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.

The Spectrum of a Ring: 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 C (the functors $I \to \mathsf{C}$) are also a category, the functor category $\mathsf{Fun}(I, \mathsf{C})$. This is also denoted 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 Δ^r to Δ 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 C and 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] $\mathfrak{p} \subset R$.

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

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 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 $\operatorname{Spec} k[x] = \mathbb{A}^1_k$.

[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.

Example 5.4 (Spec $\mathbb{Z}$)**.** Since $\mathbb{Z}$ is initial in the category of rings, then 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 Spec $\mathbb{Z}$, so Spec $\mathbb{Z}$ is connected, which is nice. The intuition is that every rational number is a function at 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 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 $\operatorname{Gal}(\mathbb{R}/\mathbb{C})$ 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.

Episode VI.

Functoriality of Spec: 2/4/16

Last time, we talked about Spec R for a ring R, the set of prime ideals. R acts like "functions" on this set: for an $r \in R$, "evaluating" it at a $\mathfrak{p} \in \operatorname{Spec} R$ returns $r \bmod \mathfrak{p}$ in $R/\mathfrak{p} \hookrightarrow \operatorname{Frac}(R/\mathfrak{p})$. Spec R has a lot of interesting structure, which we'll talk some more about today.

Recall that if k is a field, then Spec k is a point,[22] and if R is an integral domain, then (0) is prime, so there's a special point called the generic point. We also talked about a PID (actually a Euclidean domain), $k[x]$, where k is a field.

We also have affine n-space over a ring R, $\mathbb{A}^n_R = \operatorname{Spec} R[x_1, \dots, x_n]$. If $R = k$ is a field, then the affine line over k is $\mathbb{A}^1_k = \operatorname{Spec} k[x]$. This ring is a PID, and in particular primes correspond to irreducible polynomials, which correspond to Galois orbits of points in $\overline{k}$, along with one generic point.

The following theorem comes from commutative algebra.

Theorem 6.1.

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

[22] An easy way to remember this is that Spec k is a speck!

- *If R is a PID, then it's also a UFD, i.e. every $r \in R$ can be factored as $r = uf_1 \cdots f_k$, where $u \in R^\times$ and the f_j are irreducibles, unique up to units and scaling.*
- *If R is a UFD, then $R[x]$ is also a UFD.*

For a general UFD, there may exist prime ideals which are not principal.

The takeaway is that affine n-space comes from Spec of a PID. We can use this to better understand $\mathbb{A}^2_{\mathbb{C}}$ (or $\mathbb{A}^2_k$ when k is algebraically closed): there's a generic point (0), and maximal ideals $(s,t) \in k^2$ (given by the ideal $(x_1 - s, x_2 - t) \subset \mathbb{C}[x_1, x_2]$, and is clearly maximal, because the quotient is $\mathbb{C}$). However, we also have other prime ideals: if f is any irreducible polynomial, then (f) is prime (and vice versa, since we're in a UFD). In particular, there are lots of irreducibles, and therefore lots of prime ideals.

Thus, your picture could consist of all the points in k^2, and a generic point (which is dense, so draw it everywhere, maybe?), and then lots of points which are curves: for example, because $x_1^2 + x_2^2$ is irreducible in $\mathbb{C}[x,y]$, there's a point in $\mathbb{A}^2_{\mathbb{C}}$ that is the unit circle. And all lines exist, and other algebraic curves.

Exercise 6.2. These are all of the points in $\mathbb{A}^2_{\mathbb{C}}$: curves corresponding to irreducible polynomials, the maximal ideals, and the generic point.

We can learn more about $\mathbb{A}^n_k$ from the following theorem.

Theorem 6.3 (Hilbert's Nullstellensatz)**.** *If k is a field, then the maximal ideals of $k[x_1, \dots, x_n]$ have residue field k' a finite extension of k.*

That is, if $\mathfrak{m} \subset k[x_1, \dots, x_n]$ is maximal, then $k \hookrightarrow k[x_1, \dots, x_n]/\mathfrak{m}$ is a finite field extension. This is nice, because we don't have bizarre transcendental extensions. Additionally, if k is algebraically closed, then maximal ideals are what you think of as points: their residue fields have to be just k, and in fact they correspond to evaluation functions $k[x_1, \dots, x_n] \to k$, which are in bijection with $(t_1, \dots, t_n) \in k^n$.

Theorem 6.3 is equivalent to the following statement.

Corollary 6.4. *If $k \to K$ is a field extension, then K is finitely generated as a k-algebra iff it's a finite-dimensional k-vector space.*

The idea is that finite generation corresponds to a surjection $k[x_1, \dots, x_n] \twoheadrightarrow K$, which corresponds to a maximal ideal in $k[x_1, \dots, x_n]$.

We'll later use theorems like these to put finiteness conditions on different kinds of ring morphisms.

So looking at something like $\mathbb{A}^3$, there's the generic point (0), and "two-dimensional" points (f) for irreduicble f, and "zero-dimensional" points corresponding to maximal ideals.[23] Why do we have so many strange points, rather than just k^n? The answer is functoriality.

Theorem 6.5. Spec *is a functor* $\mathsf{Ring}^{\mathrm{op}} \to \mathsf{Set}$, *i.e. given a ring homomorphism $\phi : R \to T$, there's a set map $\Phi : \operatorname{Spec} T \to \operatorname{Spec} R$.*

Proof. The points of $\operatorname{Spec} T$ are maps $T \to K$ that generate K, for a field K, so composing $R \to T \to K$ gives us a homomorphism. It might not generate K, but it does generate a subfield, so this map corresponds to a prime ideal in R. ⊠

Less abstract proof. Let $\mathfrak{p} \subset T$ be prime. Then, $\phi^{-1}(\mathfrak{p}) \subset R$ is also a prime ideal: if $rs \in \phi^{-1}(\mathfrak{p})$, then $\phi(rs) \in \mathfrak{p}$, so one of $\phi(r)$ or $\phi(s)$ is in $\mathfrak{p}$, so one of r or s is in $\phi^{-1}(\mathfrak{p})$. ⊠

The preimage of a maximal idea is not necessarily maximal, which is one of the big reasons we look at more than just maximal ideals.

So prime ideals pull back, which is nice. But since R acts as functions on $\operatorname{Spec} R$, this should really be thought of as pullback of functions.

Quotients. The functoriality has some interesting consequences for our favorite ring operations. First, suppose $I \subset R$ is an ideal, so there's a quotient map $\phi : R \twoheadrightarrow R/I$. Thus, we get a map in the opposite direction, $\Phi : \operatorname{Spec} R/I \to \operatorname{Spec} R$.

Exercise 6.6. Show that an ideal $\mathfrak{p} \subset R/I$ is prime iff. $\phi^{-1}(\mathfrak{p})$ is a prime ideal in R containing I.

[23] There is a sense in which this can be made rigorous, and defines dimensions of schemes.

That is, it's an inclusion-preserving bijection, or considering only the prime ideals containing I, this is an isomorphism of posets. In any case, quotients give rise to injections $\operatorname{Spec} R/I \hookrightarrow \operatorname{Spec} R$.

For example, (xy) is an ideal in $\mathbb{C}[x,y]$, so $\mathbb{A}^2_{\mathbb{C}} \supset \operatorname{Spec} \mathbb{C}[x,y]/(xy)$. Geometrically, this is the inclusion of the coordinate axes into $\mathbb{A}^2_{\mathbb{C}}$: $\operatorname{Spec} \mathbb{C}[x,y]/I$ will be the zero locus of I in $\mathbb{A}^2_{\mathbb{C}}$. We'll let $V(I) = \operatorname{Spec} \mathbb{C}[x,y]/I \hookrightarrow \mathbb{A}^2_{\mathbb{C}}$. This is because if $I \subset \mathfrak{p} \subset \mathbb{C}[x,y]$, then $f \equiv 0 \bmod \mathfrak{p}$ for all $f \in I$.

For example, if $I = (xy)$ again, then the zero ideal doesn't contain (x,y), so the generic point isn't in this subset. And since xy isn't irreducible, then $\mathbb{C}[x,y]/(xy)$ isn't an integral domain; it has no generic point! However, there are two components with generic points: we can quotient to $\operatorname{Spec} \mathbb{C}[x,y]/(x)$ or $\operatorname{Spec} \mathbb{C}[x,y]/(y)$; each of these is a copy of $\mathbb{A}^1$ embedded (set-theoretically for now) in $V(xy)$, which are the x- and y-axis, respectively. These do have generic points, so $V(xy)$ has two "generic-like" points, which correspond to two particularly interesting prime ideals in $\mathbb{C}[x,y]/(xy)$.

To reiterate, if $\mathfrak{p}$ is a prime ideal, then we have a point $\mathfrak{p} \in \operatorname{Spec} R$, but also the map $j : \operatorname{Spec} R/\mathfrak{p} \to \operatorname{Spec} R$. How do these relate? Inside $R/\mathfrak{p}$, there's the zero ideal, which is the generic point, but this ideal corresponds to the ideal $\mathfrak{p} \subset R$. That is, j takes the generic point of $\operatorname{Spec} R/\mathfrak{p}$ to $\mathfrak{p}$. So these funny generic-like points are just the generic points of subschemes, which may be a little nicer perspective. Generic points are still weird, but once we have a topology, they just correspond to points that aren't closed (which exist in some topological spaces, yet might be less familiar). But to make Spec functorial, we have to accept these strange generic points. However, it can be a surprisingly convenient language: there's a nice representative point for any subscheme.

We can also draw $\mathbb{A}^1_{\mathbb{Z}} = \operatorname{Spec} \mathbb{Z}[x]$. Whatever it is, it comes with a map to $\operatorname{Spec} \mathbb{Z}$ (since $\mathbb{Z}$ is the initial object in the category of rings), which had a point for every prime in $\mathbb{Z}$, and a generic point. This in some sense keeps tract of the characteristics of your residue fields.

For any prime of $\mathbb{Z}$, such as 7, there's a quotient $\mathbb{Z}[x] \twoheadrightarrow \mathbb{F}_7[x]$, and this commutes with the inclusion maps $\phi : \mathbb{Z} \to \mathbb{Z}[x]$ (and the same for $\mathbb{F}_7$). Hence, we get a commutative diagram of sets

$$\begin{tikzcd}
\mathbb{A}^1_{\mathbb{F}_7} \arrow[r] \arrow[d] & \mathbb{A}^1_{\mathbb{Z}} \arrow[d, "\Phi"] \\
\operatorname{Spec} \mathbb{F}_7 \arrow[r] & \operatorname{Spec} \mathbb{Z}.
\end{tikzcd}$$

So $\Phi^{-1}(7)$ is the set of prime ideals of $\mathbb{Z}[x]$ whose intersection with $\mathbb{Z}$ contains 7, i.e. (7), and these correspond to the prime ideals of $\mathbb{F}_7[x]$. That is, we have a copy of $\mathbb{A}^1_{\mathbb{F}_7}$ as fibers of the map $\mathbb{A}^1_{\mathbb{Z}} \to \operatorname{Spec} \mathbb{Z}$, meaning $\mathbb{A}^1_{\mathbb{Z}}$ is a kind of surface, fibered over $\operatorname{Spec} \mathbb{Z}$.[24] Finally, what happens to the generic point? The fiber over (0) should correspond to maps to fields of characteristic 0, but these maps factor through $\mathbb{Q} \otimes_{\mathbb{Z}} (\mathbb{Z}[x]/\mathfrak{p})$, so (after a little work) one gets that $\Phi^{-1}(0) = \mathbb{A}^1_{\mathbb{Q}}$. See Figure 2 for one depiction of this.

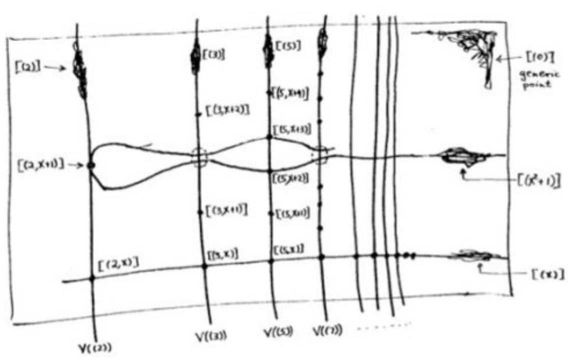

FIGURE 2. A drawing of $\mathbb{A}^1_{\mathbb{Z}} = \operatorname{Spec} \mathbb{Z}[x]$, from Mumford's red book on schemes.

[24] Eventually, when we define schemes, this will actually have geometric meaning, and will still be true for schemes.

We haven't used anything specific about $\mathbb{A}^1$, so we have a similar fibration over $\mathbb{A}^n_{\mathbb{Z}}$. If k is a field, then a map $\operatorname{Spec} k[x_1,\dots x_n]/I$ will be induced by a map $\mathbb{Z} \to k[x_1,\dots,x_n]/I$, which necessarily factors through the prime field of characteristic equal to k ($\mathbb{F}_p$ or $\mathbb{Q}$). Thus, it necessarily lives in the fiber for that prime ideal of $\operatorname{Spec}\mathbb{Z}$.

Localization. Suppose $S \subset R$ is multiplicative; then, the localization map $R \to S^{-1}R$ induces a map $\operatorname{Spec} S^{-1}R \to \operatorname{Spec} R$.

Exercise 6.7. Show that the prime ideals of $S^{-1}R$ are in bijection with the primes of R not meeting S.

Thus, localizations also give you subsets. One extreme example is $\operatorname{Spec}\mathbb{Q} \hookrightarrow \operatorname{Spec}\mathbb{Z}$ (a single point), but this is "infinitely generated" (i.e. S isn't finitely generated); we'll like finite localizations more. There are two particular examples we'll like.

- Given an $f \in R$, there's a multiplicative subset $S = \{1, f, f^2, \dots\}$, and the localization $S^{-1}R$ is denoted R_f or $f^{-1}R$. Then, $\operatorname{Spec} f^{-1}R = \{\mathfrak{p} \subset R : f \notin \mathfrak{p}\}$; that is, this is $\operatorname{Spec} R \setminus V(f)$ (since if $f \in \mathfrak{p}$, then $f \bmod \mathfrak{p} = 0$). This is a nice consequence of the weird definition of functions on our affine schemes: it's actually nice to know what it means for a function to be zero. You can also write $\operatorname{Spec} f^{-1}R = \operatorname{Spec} R \setminus \operatorname{Spec}(R/(f))$. This set is called $D(f)$, and when we have a topology, this will be referred to as a *distinguished open*.[25]

  For example, $\operatorname{Spec}(15)^{-1}\mathbb{Z}$ is the same as $\operatorname{Spec}\mathbb{Z}$, but with the points (3) and (5) removed. (The generic point is still in this set, since $15 \bmod (0) = 15$). Similarly, $\operatorname{Spec}((x^2+y^2)^{-1}\mathbb{C}[x,y])$ is the affine plane minus the circle.
- We also can define *localization at a prime*: if $\mathfrak{p} \subset R$ is prime, let $S = R \setminus \mathfrak{p}$, and we denote $S^{-1}R$ as $R_{\mathfrak{p}}$. This removes everything except the things that are inside $\mathfrak{p}$. That is, $\operatorname{Spec} R_{\mathfrak{p}}$ is the prime ideals contained in $\mathfrak{p}$.

  For example, look at $\operatorname{Spec}\mathbb{Z}_{(5)}$.[26] This contains (5) and (0), which can be thought of as a point with a little fuzziness around it. The rational numbers are also talking to it; it's not just $\mathbb{F}_5$. Yes, this may be a little weird.

  If we take $\operatorname{Spec}\mathbb{C}[x]_{(x-t)}$, we end up with the point $t \in \mathbb{A}^1_{\mathbb{C}}$ with a bit of fuzziness; it doesn't see any of the other points. But if we take $\operatorname{Spec}\mathbb{C}[x,y]_{(0,0)}$, there's the closed point (x,y) and the generic point (0), but for every irreducible polynomial f with $f(0,0) = 0$, its curve passes through the origin, so we get a nontrivial ideal of $\mathbb{C}[x,y]_{(0,0)}$. That is, geometrically, we also get a piece of each curve through the origin! This is the "local" of localization: ignore everything except what's happening arbitrarily close to 0.

If $\mathfrak{p}$ is prime, then $R_{\mathfrak{p}}$ is always a *local ring*, i.e. it always has a unique maximal ideal (since $\mathfrak{p}$ is the largest ideal contained in $\mathfrak{p}$).

Let's see what this looks like for a prime ideal that isn't maximal. $(x,z) \subset \mathbb{C}[x,y,z]$ is a point which is the y-axis (the generic point of $\operatorname{Spec}\mathbb{C}[y]$). Thus, $\operatorname{Spec}\mathbb{C}[x,y,z]_{(x,z)}$ has this point and the generic point of the plane. However, it will also contain the local data of surfaces intersecting this line. There's interesting ideal structure, but no other maximal ideals: there will be only one closed point.

Next time, we'll add topology to this picture; everything we did today is still true in the continuous world.

Episode VII.

The Zariski Topology: 2/9/16

Last time, we associated a set $\operatorname{Spec} R$ to a ring R, which was its set of prime ideals. An element $f \in R$ defines a "function" $f(\mathfrak{p}) = f \bmod \mathfrak{p} \in \operatorname{Frac}(R/\mathfrak{p})$. This is useful because we can attach two subsets $D(f) = \{\mathfrak{p} \mid f(\mathfrak{p}) \neq 0\} \subset \operatorname{Spec} R$,[27] and its complement $V(f) = \{\mathfrak{p} \mid f(\mathfrak{p}) = 0\} \subset \operatorname{Spec} R$. Thus, $V(f)$ is the set of prime ideals containing f.

[25] More generally, one can write $\operatorname{Spec} S^{-1}R = \bigcap_{f \in S} D(f)$, where S is any multiplicative subset.

[26] The notation $\mathbb{Z}_{(5)}$ denotes localization at (5); this notation is distinct from the 5-adic numbers $\mathbb{Z}_5$, or the integers modulo 5, which are $\mathbb{Z}/5$.

[27] Here, D stands for "distinguished," though it also helps to think of it as "doesn't vanish."

Definition. If $f \in R$, f is called *nilpotent* if $f^n = 0$ for some $n > 0$. The set of nilpotent elements forms an ideal, called the *nilradical*, denoted $\mathfrak{N}(R)$ or $\operatorname{Nil}(R)$.

Exercise 7.1. Suppose $f \in R$. Then, $f(\mathfrak{p}) = 0$ for all $\mathfrak{p} \in \operatorname{Spec} R$ (also written $f \equiv 0$) iff f is nilpotent.

That is, functions aren't determined entirely by their values! But in integral domains, the nilradical is zero. In general, if $f \in \operatorname{Nil}(R)$, then $V(f) = \operatorname{Spec} R$ and $D(f) = \varnothing$.

We're going to define a topology on $\operatorname{Spec} R$, called the Zariski topology, based on these ideas. In topology, if a map $f : X \to \mathbb{R}$ is continuous, then $f^{-1}(0)$ is closed and $f^{-1}(\mathbb{R} \setminus 0)$ is open. We're going to do the same thing here, declaring $D(f)$ open and $V(f)$ closed. But that's not enough to describe a topology; we want to know what all the open (or closed) subsets are.

Last time, we defined an identification $V(f) = \operatorname{Spec}(R/(f))$, and therefore for any ideal $I \subset R$, $\operatorname{Spec} R/I = \bigcap_{f \in I} V(f)$. This is an intersection of closed subsets, so it should be closed. Thus, for any $S \subset R$, we can define

$$V(S) = \{x \in \operatorname{Spec} R \mid f(x) = 0 \text{ for all } f \in S\} = \bigcap_{f \in S} V(f).$$

We will declare these sets to be closed, so now we have lots of closed subsets, and they behave well under intersection.

We can try to do the same thing with opens, but it'll look a little different. For example, if S is multiplicative, then we could define "$D(S)$" to be $\operatorname{Spec} S^{-1}R = \bigcap_{f \in S} D(f)$, since we showed last time that $\{x \mid f(x) \neq 0 \text{ for all } f \in S\}$ is identified with $\operatorname{Spec} S^{-1}R$. The problem is, this is an arbitrary intersection of opens, so it might not be open. This doesn't seem like the right definition.

However, we can take advantage of de Morgan's laws: for any $S \subset R$, $\operatorname{Spec} R \setminus V(S) = \bigcup_{f \in S} D(f)$. This ought to be open, since arbitrary unions of open sets are. This breaks the symmetry between quotients and localization, but that's okay; all of these sets are unions of our basic distinguished open sets.

Definition. The *Zariski topology* on $\operatorname{Spec} R$ is the topology in which a $Z \subset \operatorname{Spec} R$ is closed if $Z = V(S)$ for some $S \subset R$.

Thus, the opens are $\operatorname{Spec} R \setminus V(S)$ for $S \subset R$. If this seems like a strange topology, the takeaway is that *a set is closed iff it's cut out by some equations*.

We can write a closed subset as the intersection of closed sets as the form $V(f)$: $Z \subset \operatorname{Spec} R$ is closed iff there's an $S \subset R$

$$Z = \bigcap_{f \in S} V(f) = \bigcap_{f : f|_Z = 0} V(f).$$

Thus, the *Zariski closure* of any $Z \subset \operatorname{Spec} R$ is just its closure in the Zariski topology:

$$\overline{Z} = \bigcap_{f \in R : f|_Z = 0} V(f).$$

Example 7.2. Suppose $R = \mathbb{C}[x]$, so $\operatorname{Spec} R = \mathbb{A}^1_{\mathbb{C}}$. Suppose $Z \subset \mathbb{A}^1_{\mathbb{C}}$ is any infinite subset; then, $\overline{\mathbb{Z}} = \mathbb{A}^1_{\mathbb{C}}$. Why is this? This is equivalent to saying "suppose I have a polynomial function vanishing on infinitely many points; then, it's equal to zero." Thus, $\overline{Z} = V(0) = \mathbb{A}^1_{\mathbb{C}}$; the only closed sets in this topology are finite. In other words, the open sets are huge!

This example works just as well for $\operatorname{Spec} \mathbb{Z}$: if $n \in \mathbb{Z}$ and $p \mid n$ for infinitely many primes, then $n = 0$. Thus, if $Z \subset \operatorname{Spec} \mathbb{Z}$ is an infinite set, then $\overline{Z} = \operatorname{Spec} \mathbb{Z}$.

The idea that the open sets are huge is true in general, and can be somewhat frustrating: this topology is quite coarse, and sometimes is hard to work with. The closed sets have formulas associated to them, and sometimes are easier to deal with.

Proposition 7.3. *The Zariski topology is indeed a topology.*

Proof. First, we need $\varnothing$ and $\operatorname{Spec} R$ to be both open and closed. They're both closed, as $\varnothing = V(1)$ and $\operatorname{Spec} R = V(0)$, and therefore both open as well.

Next, why are arbitrary unions of open sets open? This is equivalent to arbitrary intersections of closed sets being closed, but since intersections commute with each other,

$$\bigcap_{S\in\mathcal{S}} V(S) = \bigcap_{f\in\bigcup_{S\in\mathcal{S}} S} V(f).$$

Finally, we need finite intersections of opens to be open (or equivalently by induction, that the intersection of two opens is open). This is equivalent to finite unions of closed sets being closed. If $I_1, I_2 \subset R$ are sets, then $V(I_1) = V(\langle I_1\rangle)$ (we can just take the ideal generated by I_1), so we can assume I_1 and I_2 are ideals. Hence, $V(I_1) \cup V(I_2)$ is the set of $x \in \operatorname{Spec} R$ such that $i_1(x) = 0$ for all $i_1 \in I_1$ or $i_2(x) = 0$ for all $i_2 \in I_2$. This is equivalent to $x \in V(I_1 I_2)$, since these are linear combinations of products of ideals in I_1 and I_2. Thus, $V(I_1) \cup V(I_2) = V(I_1 I_2)$, which is closed, so we're happy. ⊠

Now, we can reinterpret the setwise constructions we made last week in terms of this topology. If $\mathfrak{p} \subset R$ is prime, then $\operatorname{Spec} R_\mathfrak{p} \subset \operatorname{Spec} R$, as we talked about, and the image is $\operatorname{Spec} R_\mathfrak{p} = \bigcap_{f(\mathfrak{p})\neq 0} D(f)$, i.e. this intersection is over $f \in R \setminus \mathfrak{p}$.

Zooming out a little, this topology realizes the next step in our dream: it plays nicely with the functor $\operatorname{Spec} : \mathsf{Ring}^{\mathsf{op}} \to \mathsf{Set}$, and thus provides it with more structure. More precisely:

Exercise 7.4. Spec is actually a functor $\mathsf{Ring}^{\mathsf{op}} \to \mathsf{Top}$, i.e. if $\phi : R \to T$ is a ring homomorphism, then the induced $T : \operatorname{Spec} T \to \operatorname{Spec} R$ is continuous.

Since we're defining Φ to be the same underlying map of sets, these two visions of the Spec functor commute with the forgetful functor $\mathsf{Top} \to \mathsf{Set}$.

The intuition behind why this map is continuous is that it acts as a pullback on "functions" (elements of R); the proof that shows pullbacks are continuous on manifolds provides intuition for the proof here.

Now, we have more structure, but not enough for Spec to be an equivalence (for example, $\operatorname{Spec}\mathbb{R}$ and $\operatorname{Spec}\mathbb{C}$ are both points, but $\mathbb{R} \not\cong \mathbb{C}$ as rings).

Now, let's talk about one of the weirdnesses of the Zariski topology.

Claim. Let R be an integral domain. Then, the generic point $(0) \in \operatorname{Spec} R$ is dense.

Proof. $\overline{(0)}$ is the intersection of all $r \in R$ with $r \bmod 0 = 0$, i.e. of only $V(0)$.But $V(0) = \operatorname{Spec} R$, so $\overline{(0)} = \operatorname{Spec} R$. ⊠

Corollary 7.5. *Let* $\mathfrak{p} \in \operatorname{Spec} R$. *Then,* $\overline{\mathfrak{p}} = \operatorname{Spec} R/\mathfrak{p}$ *as a subset of* $\operatorname{Spec} R$.

The idea is that $R/\mathfrak{p}$ is an integral domain, so (0) is dense in $\operatorname{Spec} R/\mathfrak{p}$, and $(0) \in \operatorname{Spec} R/\mathfrak{p}$ maps to $\mathfrak{p}$ in $\operatorname{Spec} R$; since the inclusion $\operatorname{Spec} R/\mathfrak{p} \hookrightarrow \operatorname{Spec} R$ is continuous (which remains to be checked), it sends closures to closures.

Thus, prime ideals $\mathfrak{p}$ correspond to points, but can also be thought of the subschemes $R/\mathfrak{p}$.

Corollary 7.6. *A point* $\mathfrak{p} \in \operatorname{Spec} R$ *is closed iff it's a maximal ideal.*

The maximal ideals are the "usual" points that we're used to; closed points behave more like our intuition. All the new points, non-maximal prime ideals, are generic points of subschemes.

Another way to say this is that there's a partial order relation on points defined by inclusion. This is again unlike our usual geometric intuition.

Definition. Let $x \leftrightarrow \mathfrak{p}$ and $y \leftrightarrow \mathfrak{q}$ be points in $\operatorname{Spec} R$. Then, x is a *specialization* of y if $x \in \overline{y}$, i.e. $\mathfrak{q} \subset \mathfrak{p}$, and y is called a *generization* of x.

The minimal elements of this poset are the closed points, which correspond to the maximal ideals.

For example, in $\operatorname{Spec}\mathbb{Z}$, the generic point (0) corresponds to the image of $\operatorname{Spec}\mathbb{Q}$ induced by $\mathbb{Z} \hookrightarrow \mathbb{Q}$. It's dense, but it's not open: $\operatorname{Spec}\mathbb{Z} \setminus (0)$ is an infinie subset, and we saw above that infinite subsets aren't closed. Nonetheless, it's still Spec of a localization (the field of fractions). The takeaway is that Spec of a localization is not open in general.

Another useful fact is that the distinguished opens $D(f)$ form a basis for the Zariski topology, i.e. any open $U \subset \operatorname{Spec} R$ is a union of these open sets. This is because $U = \operatorname{Spec} R \setminus V(I)$ for some ideal $I \subset R$. Hence $\operatorname{Spec} R \setminus U = \bigcap_{f\in I} V(f)$, and hence $U = \bigcup_{f\in I} D(f)$. This is more of a tongue-twister than a proof, but it all comes down to complements of intersections becoming unions of complements.

Claim. If $f, g \in R$, then $D(f) \subset D(g)$ iff $f^n \in (g)$ for some $n > 0$.

Proof. Again, we'll unwind definitions. If $D(f) \subset D(g)$, then $f|_{V(g)} = 0$, so $f \bmod (g)$ vanishes everywhere on $\operatorname{Spec} R/(g)$. Thus, $f \in \mathrm{Nil}(R/(g))$, so $f^n \equiv 0 \bmod g$ for some n. That is, $f^n \in (g)$. The converse is analogous. ⊠

This is a nice bridge between algebra and geometry.

So we have a basis of distinguished opens and an inclusion relation between them. We also have a nice property about coverings.

Exercise 7.7. Show that if $S \subset R$, then $\operatorname{Spec} R = \bigcup_{f \in S} D(f)$ iff $(S) = R$.

That is, covering corresponds to that set generating the whole ring. But $(S) = R$ has a nice concrete meaning: there are $f_i \in S$ and $a_i \in R$ such that $\sum a_i s_i = 1$. This is necessarily a finite sum, because generating an ideal means only taking finite linear combinations. Thus, we get a curious corollary.

Corollary 7.8. $\operatorname{Spec} R$ *is quasicompact.*[28]

The proof is to replace an arbitrary cover by the finite combination that sums to 1.

There's no reasonable sense in which a scheme is compact, and certainly we don't want something like $\mathbb{A}^n_{\mathbb{C}}$ to be compact. The idea is that open sets are huge: in $\mathbb{A}^1_{\mathbb{C}}$, for example, any nonempty open is dense, and therefore any two nonempty opens have an intersection that's also dense in $\mathbb{A}^1_{\mathbb{C}}$! In other words, the Zariski topology is very far from Hausdorff.

We'd like to have notions that are compactness and Hausdorfness but for schemes, but the usual ones don't work. We'll define analogous notions (e.g. "separatedness" for the Hausdorff property), but until then, we have to be careful with keeping all the words right in the dictionary between usual geometry and algebraic geometry.

Example 7.9. If k is a field, then *infinite-dimensional affine space* over k is $X = \operatorname{Spec} k[x_1, x_2, \dots] = \mathbb{A}^\infty_k$. Then, $\mathbf{0} = V(x_1, x_2, \dots)$, so $X \setminus \mathbf{0}$ is open, but not quasicompact: the cover $X \setminus \mathbf{0} = \bigcup_{i \in \mathbb{N}} D(x_i)$ has no finite subcover. Thus, $X \setminus \mathbf{0}$ is not $\operatorname{Spec} T$ for any ring T! We'll eventually see how this is a scheme, but it isn't an affine scheme.

There are easier examples of schemes which aren't affine: the functions on $\mathbb{A}^2_{\mathbb{C}} \setminus 0$ are just $\mathbb{C}[x, y]$, which are the functions on $\mathbb{A}^2$, but saying this rigorously requires more work.

We can also talk about connectedness: when can we write $\operatorname{Spec} R = X \amalg Y$ for open and closed X and Y? The intuition is that we will be able to work in the opposite category: $R = S \times T$ iff $\operatorname{Spec} TR = \operatorname{Spec} S \amalg \operatorname{Spec} T$.

But even before that, decomposing as $X \amalg Y$ must mean there are functions i_X, i_Y on $\operatorname{Spec} R$ such that $i_X|_X = 1$, $i_Y|_X = 0$, and vice versa for Y. Thus, $i_X^2 = i_X$ and $i_X i_Y = 0$, so they're *orthogonal idempotents*; hence, connectedness corresponds to (not) finding orthogonal idempotents in R.

Episode VIII.

Connectedness, Irreducibility, and the Noetherian Condition: 2/11/16

"I just can't read my own handwriting."

Today, we'll talk about properties of the Zariski topology and its relation to the structure of rings; next week, we'll cover a little general nonsense about sheaf theory, and finally get to define schemes.

Connectedness. One of the most basic questions one can ask about a topological space is whether it's connected. It turns out that for the Zariski topology, connectedness correlates very nicely with an algebraic property.

Definition. If R is a ring, then $i \in R$ is an *idempotent* if $i^2 = i$.

[28] This just means "every open cover has a finite subcover;" in this schema, compactness is reserved for Hausdorff spaces, and we use quasicompactness to make the distinction clearer.

These are akin to projectors, and indeed, we have a complementary projector: $(1-i)^2 = 1-2i+i = 1-i$, so $i^\perp = 1-i$ is also an idempotent. Since R is commutative, then i and $i^\perp$ commute, and $i+i^\perp = 1$. Moreover, i and $i^\perp$ are *orthogonal idempotents*, in the sense that $ii^\perp = i - i = 0$.

Since these add to 1, $r = ri + ri^\perp$ for any $r \in R$, and so $R \cong iR \times i^\perp R$. Thus, the prime ideals in R are the disjoint union of the ones in iR and $i^\perp R$. Thus, $\operatorname{Spec} R = \operatorname{Spec} iR \amalg \operatorname{Spec} i^\perp R$, and this is true as topological spaces. Since $\operatorname{Spec} iR = V(i^\perp)$ and $\operatorname{Spec} i^\perp R = V(i)$, these are both clopen sets, and therefore $\operatorname{Spec} R$ isn't connected.

Conversely, we will be able to prove that if $\operatorname{Spec} R = X \amalg Y$ as topological spaces, there's an idempotent i_X which is valued 1 on X and 0 on Y, and in fact this lies in R, and so a decomposition $R = i_X R \times i_X^\perp R$. However, we don't have the techniques to prove this yet. Hence, $\operatorname{Spec} R$ is connected iff R has no idempotents.

In representation theory, one often studies associative algebras, such as the group algebra $k[G]$ associated to a finite group G over a field k. Inside A, there's a commutative ring, its center $Z(A)$. Wedderburn's theorem states that the idempotents $i \in Z(A)$, called *central idempotents*, are in bijection with the irreducible representations of G over k. Thus, commutative algebra is useful even in non-commutative algebra.

Irreducibility. The notion of irreducibility is one that doesn't come up in ordinary geometry.

Definition. A topological space X is *irreducible* if you can't write $X = Z_1 \cup Z_2$ for proper closed subsets $Z_1, Z_2 \subset X$.

In Euclidean geometry, this is absurd: consider the upper half-plane (including the x-axis) and the lower half-plane (including the x-axis) for $\mathbb{R}^2$. But the Zariski topology encodes things differently: it's a way of encoding algebraic structure on a space.

Suppose R is an integral domain. Then, there's a generic point $(0) \in \operatorname{Spec} R$ which is dense. Suppose $\operatorname{Spec} R = Z_1 \cup Z_2$ for proper closed subsets Z_1 and Z_2; then, (0) is in one of them, and so its closure is too (since they're both closed). However, that's the whole space, so one of them isn't a proper subset. Thus, if R is an integral domain, $\operatorname{Spec} R$ is irreducible.

Conversely, suppose R is any ring, and $\operatorname{Spec} R = X = Z_1 \cup Z_2$, for proper, closed subsets Z_1 and Z_2. Then, there exist functions $f_1, f_2 \in R$ such that $f_i|_{Z_i} = 0$ and $f_i \neq 0$, meaning neither f_i is nilpotent. However, $f_1 f_2 \equiv 0$, so it is nilpotent, and so $(f_1 f_2)^N = 0$. In particular, R is not an integral domain.

We must be careful, because this is not an if and only if.

Example 8.1 (Dual numbers)**.** We're going to get quite acquainted with the *dual numbers*, the ring $k[\varepsilon]/(\varepsilon^2)$, for a field k. Thus, ε is a nilpotent, so this ring isn't an integral domain. Then, $\operatorname{Spec} k[\varepsilon]/\varepsilon^2$ is a point, as is $\operatorname{Spec} k$, and so this is certainly irreducible! So integral domain implies irreducible, but not vice versa.

This ring will show up as a useful example because it's a simple example of how nilpotents work.

Another corollary is that $\operatorname{Spec} R$ as a topological space is insensitive to nilpotents in R, since nilpotents as functions are identically 0, and they don't affect the space of prime ideals. So we'll have to put a stronger structure on $\operatorname{Spec} R$ to distinguish these two rings.

Definition. A ring R is *reduced* if it has no nonzero nilpotents, i.e. $\operatorname{Nil}(R) = 0$.

Now, if we have a reduced ring R and $f, g \in R$ with $fg = 0$, $f, g \neq 0$, then $\operatorname{Spec} R = V(f) \cup V(g)$, and this is a proper decomposition (there are no nonzero nilpotents, so there are places where f and g don't vanish). The point is, *an integral domain corresponds to being reduced and irreducible*.

Recall that we have a dictionary between algebra and geometry: given an ideal $I \subset R$, there's a closed subset $V(I) \subset \operatorname{Spec} R$. Correspondingly, if $S \subset \operatorname{Spec} R$ is any subset, then the set $I(S) = \{r : r|_S \neq 0\}$ is an ideal of R. Then, $V(I(S)) = \bigcap_{f|_S = 0} V(f)$, so this is just the Zariski closure of S in $\operatorname{Spec} R$.

Correspondingly, $I(R(J)) = \sqrt{J} = \{r \in R : r^N \in J, N \gg 0\}$, because if $r|_{V(J)} = 0$, then $r \bmod J$ is identically 0 on $\operatorname{Spec} R/J$, i.e. it's nilpotent in R/J (so $r^N \in J$ for some N). If $J = \sqrt{J}$, then J is called *radical*.

This correspondence thus isn't perfectly bijective, but it's a nice dictionary. In particular, we've proven the following theorem.

Theorem 8.2. *These functions I and V provide a bijection between the closed subsets of* $\operatorname{Spec} R$ *and the radical ideals of R.*

Under this bijection, prime ideals correspond to points, but also to irreducible subsets (the closure of its generic point). To be precise, $\operatorname{Spec} R/I$ is closed in $\operatorname{Spec} R$, and I is radical iff R/I is reduced. It's also irreducible iff I is prime, but we saw that both of these are equivalent to R/I being an integral domain. This can be useful: a single point can be used to understand an entire irreducible subset, which is quite precise.

For example, in $\mathbb{A}^2_{\mathbb{C}}$, irreducible subsets are in bijection with prime ideals, and irreducible polynomials give us prime ideals. However, the union of the coordinate axes is not irreducible (it's the union of the x-axis and the y-axis).

Notice that an irreducible set is automatically connected (and this translates to an algebraic statement, too), so disconnected subsets are reducible. What is this good for? Well, let $Y = \operatorname{Spec} R/I$; I want to understand this better, and so want to write this as a union of irreducible components $Y_1 \cup \cdots \cup Y_n$, so $Y_i = \operatorname{Spec} R/\mathfrak{p}_i$. Algebraically, if R is reduced, this means writing $I = \bigcap_{i=1}^n \mathfrak{p}_i$. The takeaway is that we could understand any ideal in terms of prime ideals.

This is not true in general: first of all, Y may not have a finite number of irreducible components, nor I be a finite intersection of prime ideals. For example, in infinite-dimensional affine space $\mathbb{A}^\infty_k = \operatorname{Spec} k[x_1, x_2, \dots]$, consider the set Y that's the union of all of the coordinate axes. This does not satisfy this finiteness condition, and we'd like a word for the condition that does. That word is Noetherian.

The Noetherian Condition. This should be thought of as a "finite-dimensionality." Dimension is weird enough in ordinary topology, but the weirdness of the Zariski topology allows finite-dimensionality to be easily defined.

Definition. A topological space X is *Noetherian* if it satisfies the *descending chain condition* (DCC): if $X \supset Z_1 \supset Z_2 \supset \cdots$ is an infinite sequence of closed subsets of X, then it eventually stabilizes. In other words, there's some n such that $Z_n = Z_{n+1} = Z_{n+1} = \cdots$.

This is a funny condition when one first sees it; just like irreducibility, it doesn't arise in ordinary topology. Any shrinking sequence of neighborhoods on a manifold shows that it's not Noetherian (and in the same way, there are infinite ascending chains). And these two notions are related.

Proposition 8.3. *If X is a Noetherian topological space, then any closed $Y \subset X$ can be written as a finite union of irreducibles $Y = Y_1 \cup \cdots \cup Y_n$. If we additionally specify that no Y_i contains any other Y_j, then this decomposition is unique.*

Proof. We'll prove this using a technique called *Noetherian induction*, which we'll use again.

Let S denote the set of closed subsets of X not admitting such a description; we would like to show that S is empty. If S is nonempty, we'll show there's a minimal element of S with respect to inclusion, and use it to derive a contradiction.

Suppose $Y_1 \in S$ is not minimal, then pick a $Y_2 \subset Y_1$ in S, and repeat this argument: we get a chain $Y_1 \supset Y_2 \supset Y_3 \supset \cdots$, so the Noetherian condition guarantees this stabilizes at some $Y_n = Y$, and this must be a minimal element for S.

Since Y doesn't have a decomposition into a finite number of irreducibles, then it's not irreducible, and so $Y = W \cup Z$ for proper closed subsets $W, Z \subset Y$. But since Y is minimal, then $W, Z \notin S$, so $W = W_1, \dots, W_n$ and $Z = Z_1, \dots, Z_m$ are decompositions into irreducibles. Hence, taking the union of these two, we have a finite decomposition into irreducibles for Y, which is a contradiction.

The uniqueness is pretty easy: it's very much like the uniqueness of prime factorization, which is not a coincidence. Suppose $Y = Z_1 \cup \cdots \cup Z_m = W_1 \cup \cdots \cup W_n$ are two decompositions, where no Z_i contains a Z_j, and the same for the W_i. Since $Z_1 \subset W_1 \cup \cdots \cup W_n$, then $Z_1 = (Z_1 \cap W_1) \cup \cdots \cup (Z_1 \cap W_n)$. Thus, since Z_1 is irreducible, one of these, without loss of generality $Z_1 \cap W_1$, is equal to Z_1. Thus, $Z_1 \subset W_1$, and with the Z_i and W_i switched, we have $W_1 \subset Z_2$, so $Z_1 \subset Z_2$, which we assumed was not the case unless $Z_1 = W_1$ (and then induction takes care of the rest). ⊠

The Noetherian condition arose first on rings.

Definition.

- A ring R is *Noetherian* if it satisfies the *ascending chain condition* (ACC) on ideals: if $I_1 \subset I_2 \subset I_3 \subset \cdots$ are all ideals of R, then there's some n such that $I_n = I_{n+1} = \cdots$.
- R is *Artinian* if it satisfies the descending chain condition on ideals.

One could correspondingly define Artinian topological spaces, which would have to satisfy the ascending chain condition on closed subsets. But very few spaces are Artinian (e.g. $\operatorname{Spec}\mathbb{Z}$ isn't); it suggests that the space is finite.

Exercise 8.4. Show that if R is Noetherian, then $\operatorname{Spec} R$ is a Noetherian space.

The converse is very false, since Spec doesn't detect nilpotents, so come up with a ring with an infinite ascending chain of nilpotents and you'll have a counterexample.

Lots of nice rings (or spaces) are Noetherian: all fields are, as is $\mathbb{Z}$, and quotients and localizations preserve the Noetherian condition. And there's a cleaner way to check the definition.

Exercise 8.5. Show that the following are equivalent.

(1) R is a Noetherian ring.
(2) Every ideal of R is finitely generated.
(3) For any closed $Z \subset R$ given by $Z = \operatorname{Spec} R/I$, Z is an intersection of finitely many $V(f)$ for $f \in R$.

So Noetherianness is a quite strong finite condition.

Theorem 8.6 (Hilbert basis theorem)**.** *If R is Noetherian, then so is $R[x]$.*

Combined with quotients and localizations, this is equivalent to the statement that if R is a ring, then all finitely generated R-algebras are Noetherian. This is a quite powerful statement: there are lots and lots of Noetherian rings, including pretty much any ring that one usually thinks about.

And we also get back the statement we were looking for when we defined this: if R is a Noetherian ring and $I \subset R$ is a radical ideal, then there are ideals $I = \mathfrak{p}_1 \cap \cdots \cap \mathfrak{p}_n$ for $\mathfrak{p}_j$ prime, or every closed subset is a finite union of irreducibles.

Proof of Theorem 8.6. Let $I \subset R[x]$ be nonzero; we'll show I is finitely generated. Pick an $f_1 \in I$ of minimal degree; then, pick an $f_2 \in I \setminus (f_1)$ of minimal degree, and $f_3 \in I \setminus (f_1, f_2)$ of minimal degree, and so forth.

Thus, we have the chain $(f_1) \subset (f_1, f_2) \subset (f_1, f_2, f_3) \subset \cdots$ in I. We will show it terminates, which means I is finitely generated, but we don't yet know $R[x]$ is Noetherian. Since we do know this for R, let a_i be the leading coefficient of f_i. Thus, we have a chain of ideals in R: $(a_1) \subset (a_1, a_2) \subset (a_1, a_2, a_3) \subset \cdots$. Since R is Noetherian, this stabilizes, so there's an n such that $a_n = \sum_{i<n} r_i a_i$ for some $r_i \in R$.

Thus, $f_n = a_n x^{\deg f_n} + \cdots$, so we can peel off the highest-degree term:

$$g = f_n - \sum_{i<n} r_i f_i x^{\deg f_n - \deg f_i}.$$

This is in I, but it's not contained in $(f_1, \ldots, f_{n-1})$, since then f_n would be (and we assumed the chain doesn't stabilize). However, it has lower degree than f_n does, and we assumed it was minimal, giving us a contradiction. ⊠

This reduction of polynomials to their coefficient rings is probably the same trick used to prove that if R is a UFD, then so is $R[x]$.

Unless you care about infinite-dimensional things, you probably won't ever have to worry about non-Noetherian spaces or rings.

As we'll see, the Zariski topology is somewhat weird, and encodes the algebra of a ring, but it doesn't pick up the geometry. It's a pictorial summary of the algebra; once we introduce some geometry, the geometry we get is much more like complex geometry.

Episode IX.

Revenge of the Sheaf: 2/16/16

This week, we're going to provide the last ingredient in the definition of an affine scheme: its sheaf of functions. To do that, we'll have to define sheaves abstractly.

Sheaves formalize the idea that functions are local: if X is a topological space, we consider functions $f : X \to \mathbb{R}$ (or to another topological space Y). This "locality" means the following things:

- If $U \subset X$ is open, we can restrict f to $f|_U$, which is a function on U. If $V \subset U$ is another open, restriction composes: $(f|_U)|_V) = f|_V$. This will generalize to the notion of a presheaf.
- If $X = U \cap V$, where U and V are open subsets, then we can glue functions that agree on the overlaps. In other words, if $\mathcal{F}(X)$ is the functions on X, then restriction gives us an injective map $\mathcal{F}(X) \to \mathcal{F}(U) \times \mathcal{F}(V)$, and the image is exactly the functions agreeing on $U \cap V$. This will generalize to a sheaf.

To define these formally, fix a category C; if it helps to be concrete, C will almost always be Set, Ab, or Ring for this class. Recall that if X is a topological space, then $\mathsf{Top}(X)$ is the category of open subsets in X, interpreted as a poset under inclusion.

Definition. The *category of presheaves* on X, $\mathsf{C}_X^{\mathrm{pre}}$ is the category whose objects are functors $\mathsf{Top}(X)^{\mathsf{op}} \to \mathsf{C}$ and whose morphisms are their natural transformations.

What does this actually mean? If $\mathcal{F}$ is a presheaf, then to any open $U \subset X$, we have its *sections* on U, $\mathcal{F}(U) \in \mathsf{C}$, and composition of morphisms means that if $W \subset V \subset U \subset X$ are open sets, then we have *restrictions maps* that commute: if res_U^V denotes restriction from U to V, then $\mathrm{res}_U^W = \mathrm{res}_V^W \circ \mathrm{res}_U^V$.

A morphism of sheaves is a natural transformation $\Phi : \mathcal{F} \to \mathcal{G}$, i.e. for all $V \subset U$ as opens of X, there's a commutative diagram of maps in C:

$$\begin{tikzcd}
\mathcal{F}(U) \arrow[r, "\Phi(U)"] \arrow[d, "\mathrm{res}_U^V"'] & \mathcal{G}(U) \arrow[d, "\mathrm{res}_U^V"] \\
\mathcal{F}(V) \arrow[r, "\Phi(V)"] & \mathcal{G}(V).
\end{tikzcd}$$

For example, if X is a topological space, the continuous, real-valued functions $C(X;\mathbb{R})$ form a presheaf, by the assignment $U \mapsto C(U;\mathbb{R})$.

Now, we'd like to extract sheaves from this, by adding a *descent* (or *locality* or *gluing* or sheaf) axiom.

Definition. The category C_X of C-*valued sheaves* on X is the full subcategory[29] satisfying the *sheaf axiom*: let $U \subset X$ be open and $\mathfrak{U} = \{U_i\}_{i\in I}$ be an open cover of U (so U is the union of the U_i). For any $i, j \in I$, let $U_{ij} = U_i \cap U_j$; then, we require the diagram

$$\mathcal{F}(U) \xrightarrow{\prod \mathrm{res}_U^{U_i}} \prod_{i\in I} \mathcal{F}(U_i) \rightrightarrows \prod_{i,j\in I} \mathcal{F}(U_{ij})$$

is an equalizer diagram.

Well, that was compact; let's unpack it. Suppose $U = U_1 \cap U_2$; then, $U = U_1 \amalg_{U_{12}} U_2$ in $\mathsf{Top}(X)$, which means the diagram

$$U_{12} \rightrightarrows U_1 \amalg U_2 \longrightarrow U$$

is a *co*equalizer diagram.[30] Because a presheaf is a contravariant functor, we'd like it to turn the coequalizer diagram that encodes an open cover into an equalizer diagram.

What this actually means is that if I have objects $f_i \in \mathcal{F}(U_i)$ for each i, such that the restrictions all agree (so $\mathrm{res}_{U_i}^{U_{ij}} f_i = \mathrm{res}_{U_j}^{U_{ij}} f_j$ for all $i, j \in I$), then there exists a unique $f \in \mathcal{F}(X)$ such that $f_i = \mathrm{res}_U^{U_i} f$. That is, we can glue sheaves on an open cover, and can do so uniquely.

Example 9.1. Many kinds of functions on your space will form sheaves, e.g. C^∞ (smooth functions), C^ω (analytic functions), continuous functions. For any set Y, maps from X to Y form a sheaf in the category of sets.

In the future, when we say something is local, we will mean that it forms a sheaf: its values on an open cover determine it globally.

---

[29] "Sheaves are a full subcategory of presheaves" means that every sheaf is a presheaf, and the morphisms are the same.

[30] Something unusual is going on, because, strictly speaking, $U_1 \amalg U_2$ is not in $\mathsf{Top}(X)$. This diagram actually lives in Top. Gorthendieck reformulated a lot of this by recasting these as maps to your space, rather than subsets... but that's a story for another day.

Example 9.2 (Skyscraper sheaf)**.** We can also define sheaves which don't quite look like functions. Let C be a category with a terminal object $*$ (e.g. Set), and $S \in \mathsf{C}$. Pick an $x \in X$, and define the *skyscraper sheaf* $i_{x,*}S$ by

$$i_{x,*}S(U) = \begin{cases} S, & x \in U \\ *, & x \notin U. \end{cases}$$

This is a presheaf, because this definition plays well with restriction: if $V \subset U$, then either $x \notin U$ (so we restrict $* \to *$), $x \in V$ (so we restrict $S \to S$), or $x \in U \setminus V$ (so we restrict $S \to *$), which is fine. And it's a sheaf, because if we have an open cover for an open $U \subset X$, then U contains x iff its cover does, so everything works out. The name comes from a picture where S is stacked up at x, and there's nothing anywhere else.

Example 9.3 (Constant presheaves and sheaves)**.** Let X be a space that contains two disjoint open subsets U_1 and U_2 (e.g. any nontrivial Hausdorff space). If S is a set, we can define the *constant presheaf* with value S by defining $\mathcal{F}(U) = S$ for all $U \subset X$, and the restriction maps to be the identity; this commutes with taking sections, and therefore is a presheaf.

However, this $\mathcal{F}$ is *not* a sheaf: we can pick distinct sections $s_1, s_2 \in S$, regarding $s_1 \in U_1$ and $s_2 \in U_2$. Since U_1 and U_2 are disjoint, then these have to come from a single section on $U_1 \cap U_2$ (they vacuously agree on the intersection), but the restriction maps are all the identity, so there's no way to do this.

However, we can tweak this into a sheaf. Now, for any $S \in \mathsf{Set}$, endow S with the discrete topology, and let $\underline{S}(U) = \mathrm{Hom}_{\mathsf{Top}}(U, S)$. These are continuous functions, and therefore form a sheaf $\underline{S}$, called the *constant sheaf*. Since S is totally disconnected, each map from a connected subset factors through a single point of S, and therefore the issue that the constant presheaf had doesn't arise.[31]

Example 9.4 (Sheaf of sections)**.** Let $\pi : Y \to X$ be a continuous map. Then, the *sheaf of sections* of π is defined by $\mathcal{F}(U)$ to be the sections of the map $\pi^{-1}(U) \to U$, i.e. continuous maps $s : U \to \pi^{-1}(U)$ such that $\pi \circ s = \mathrm{id}$. We can restrict sections, so this is a presheaf, but in fact sections are always a sheaf: if two sections agree on their overlap, they can be patched. That is, sections are local information. Notice that the codomain of this sheaf varies as U varies.

If $Y = X \times T$, and π is projecting onto the first factor, then sections of π are just maps $X \to T$ (regarded as its graph); in other words, the sheaf of sections generalizes the sheaf of maps. In fact, we'll see later that any sheaf can be regarded in this way: sections are actually sections.

Another good example is when $\pi : Y \to X$ is a covering space with fiber Γ (e.g. $\mathbb{R} \to S^1$ with fiber $\mathbb{Z}$); then, the sections of the covering map on a sufficiently small $U \subset X$ are the same thing as maps $U \to \Gamma$, because $\pi^{-1}(U) \cong U \times \Gamma$ for small enough U. Since Γ has the discrete topology, this means that for these small U, $\mathcal{F}(U) = \underline{\Gamma}(U)$: on small enough open sets, it looks like the constant sheaf. Geometrically, this means that a section of a covering map is a choice of one of the sheets along with the inverse of the projection. However, globally, we can't map $S^1 \to \mathbb{R}$ as a section of the covering map.

Every locally constant sheaf arises from a covering space in this way, though the definition of "covering space" may need to be expanded.

Definition. Let $U \subset X$ be open. Then, there's a functor $\mathsf{C}_X \to \mathsf{C}_U$, called *restriction* (of sheaves) that sends a sheaf $\mathcal{F}$ to the sheaf $\mathcal{F}|_U$ whose value at a $V \subset U$ is $\mathcal{F}(V)$. In exactly the same way, we can define restriction of presheaves.

This makes sense: all we do is forget about the opens not contained in U. And you can check this is functorial.

This allows us to formalize the covering example just above into an extremely useful class of sheaves.

Definition. A sheaf $\mathcal{F}$ is *locally constant* if there's an open cover $\mathfrak{U}$ of X such that for every $U \in \mathfrak{U}$, $\mathcal{F}|_U \cong \underline{S}_U$ is a constant sheaf.

Another perspective is that sheaves measure twisting: we know what local data looks like, and the sheaf tells us how these are twisted and glued together to obtain the total data. One interesting way

[31]Again, this is a little silly with the Zariski topology, as nonempty opens all intersect. Grothendieck resolved this by defining finer topologies on schemes; we'll just not deal with constant sheaves on the Zariski topology.

Definition. If $\mathcal{F}$ is a (pre)sheaf, its *global sections* $\Gamma(\mathcal{F}) = \Gamma(X, \mathcal{F})$ are just $\mathcal{F}(X) = \mathsf{C}$. Taking global sections defines a functor $\Gamma : \mathsf{C}_X \to \mathsf{C}$.

To be precise, the global sections of a sheaf are sections on an open subset that agree on overlaps, for any open cover (e.g. X itself is an open cover). If $\mathcal{F} \in \mathsf{Set}_X$, so F is a sheaf of sets, we can also write $\Gamma(\mathcal{F}) = \mathrm{Hom}_{\mathsf{Set}_X}(\underline{*}, X)$. Here, $\underline{*}$ is the constant sheaf valued in a point. The idea is that a map of sheaves is the data of a map $* \to \mathcal{F}(U)$ for each open $U \subset X$, i.e. a collection of $f_U \in \mathcal{F}(U)$ which agree on overlaps, which is exactly the data we needed. This may be confusing, but is sometimes useful: we know Hom commutes with limits, so $\Gamma : \mathsf{Set}_X \to \mathsf{Set}$ preserves limits! This is a common theme: if you can write a construction as an adjoint or an instance of Hom or a limit, you already know a bunch of properties of it.

Global sections are an example of something more general: sheaves propagate from one space to another.

Definition. Let $\pi : Y \to X$ be continuous, and $\mathcal{F}$ be a sheaf (resp. presheaf) on Y. Then, we can define its *pushforward* $\pi_*\mathcal{F}$, which is a sheaf (resp. presheaf) on X, by $\pi_*\mathcal{F}(U) = \mathcal{F}(\pi^{-1}(U))$. Since π^{-1} commutes with restriction, this is a presheaf, and if $\mathcal{F}$ is a sheaf, then a cover of U pulls back under π^{-1} to a cover of its preimage, so we can glue on $\pi^{-1}(U)$ by elements of π^{-1} of its covers.

This very useful operation on sheaves defines a functor $\pi_* : \mathsf{C}_Y \to \mathsf{C}_X$, which does need to be checked: if one has a map of sheaves $\mathcal{F} \to \mathcal{G}$ and $V \subset U$, does the following diagram commute?

$$\begin{tikzcd}
\mathcal{F}(\pi^{-1}(U)) \arrow[r] \arrow[d] & \mathcal{G}(\pi^{-1}(U)) \arrow[d] \\
\mathcal{F}(\pi^{-1}(V)) \arrow[r] & \mathcal{G}(\pi^{-1}(V))
\end{tikzcd}$$

This generalizes several things we've already seen.

- Global sections are a special case of pushforward: for any topological space X, there's a unique map $\pi : X \to *$, and $\pi_* : \mathsf{C}_X \to \mathsf{C}_* = \mathsf{C}$ just takes sections that agree on all of X, i.e. $\Gamma(\mathcal{F})$.
- Skyscraper sheaves are also pushforwards: consider the map $i_x : * \hookrightarrow X$ sending $* \mapsto x$. Then, $i_{x,*} : \mathsf{C} \to \mathsf{C}_X$ is the pushforward of the constant sheaf S on $*$, as $i_{x,*}S(U) = S(i_x^{-1}(U))$, which agrees with its definition in Example 9.2.
- The sheaf of sections can also be realized in this way: if $\pi : Y \to X$ is continuous, let $\underline{*}$ denote the constant sheaf on Y; then, the sheaf of sections is just $\pi_*(\underline{*})$: for any open $V \subset Y$, $\underline{*}(V) = *$, so for any $U \subset X$, $\pi_*(\underline{*})(U) = \pi^{-1}(U)$.

We want to write any sheaf $\mathcal{F}$ as a sheaf of sections of a map $\pi : Y \to X$, and we'll do this by building Y out of the stalks of $\mathcal{F}$.

Definition. Let $\mathcal{F}$ be a (pre)sheaf and $x \in X$. Then, the *stalk* of $\mathcal{F}$ at x, $\mathcal{F}_x \in \mathsf{C}$, is the object of sections of $\mathcal{F}$ on some open subset containing X: any two neighborhoods of x intersect in a smaller neighborhood, and we would like to identify sections that agree on the intersection. If we had a minimal neighborhood of x, that would be where the stalk takes its sections, but instead we do the next best thing.

To be precise, the stalk is $\mathcal{F}_x = \varinjlim_{x \in U} \mathcal{F}(U)$. What does this mean? We have a poset of opens containing X, and if $x \in V \subset U$, then restrictions $\mathcal{F}(U) \to \mathcal{F}(V)$ define a filtered system, so we're just taking the filtered colimit, which tries to be the miniaml element.

Since we worked out filtered colimits, we can write this as the quotient

$$\mathcal{F}_x = \varinjlim_{x \in U} \mathcal{F}(U) = \left\{ f_U \in \mathcal{F}(U) \text{ where } x \in U / f_U \sim f_V \text{ if } \mathrm{res}_U^{U\cap V} f_U = \mathrm{res}_V^{U \cap V} f_V \right\}.$$

If $\mathcal{F}$ is the shead of C^∞ (or similar) functions, then its stalks $\mathcal{F}_x$ are the *germs* of functions at x: smooth functions on neighborhoods of x, where we identify functions that agree on a neighborhood of x. Interestingly, if $\mathcal{F}$ is the sheaf of holomorphic functions on C, then by analytic continuation, $\mathcal{F}_0$ is the ring of Taylor series with nonzero radii of convergence.

Not all filtered colimits exist, but in the categories we'll care about (sets, abelian groups, rings, and such), all filtered colimits exist and are fairly well-behaved.

Episode X.

Revenge of the Sheaf, II: 2/18/16

"Ravi says some people swear by [the espace étalé]. I haven't met them."

Today, we're going to talk more about sheaves. Recall that these generalize the notion of functions on a space. If X is a topological space and C is a category, a C-valued sheaf is an association of an object of C called $F(U)$ to every open $U \subset X$, and with restriction maps $F(U) \to F(V)$ when $V \subset U$, compatible with gluing across intersecting opens. (For concreteness, you can think of everything using $\mathsf{C} = \mathsf{Set}$.)

The most common example is the sheaf of functions, $U \mapsto \mathrm{Maps}(U, \mathbb{R})$. We can also talk about "twisted functions" such as the sheaf of sections or a covering space: the local structure is $U \mapsto \mathrm{Maps}(U, T)$ for some target T, but the global structure is different. This is a common way to think about sheaves.

The most important measurement we extract from a sheaf is its stalks, which allow us to understand how the sheaf behaves on non-open subsets. For example, if $x \in X$, we'd like to understand it through the not-quite-open set $\bigcap_{x \in U} U$, and therefore we get the stalk $\mathcal{F}_x = \varinjlim_{x \in U} \mathcal{F}(U)$, where the open sets are indexed by restriction. This is a filtered colimit, and therefore can be described explicitly as equivalence classes of functions in neighborhoods, where $f \sim g$ if they're equivalent on some common neighborhood of x. Elements of a stalk are called *germs of sections*.

The notion of a stalk still makes sense for presheaves, and today we'll talk about how to determine whether a presheaf is a sheaf using its set of stalks.

Recall that $\mathbb{R}[\![x]\!]$ is the ring of power series with coefficients in $\mathbb{R}$. $C^\infty(\mathbb{R})$ is a sheaf on $\mathbb{R}$, and its stalk at the origin is C_0^∞, the germs of functions at the origin. Since every function has a Taylor series, there's a surjective ring homomorphism $\mathbb{C}_0^\infty \twoheadrightarrow \mathbb{R}[\![x]\!]$. However, if we use the sheaf of analytic functions $C^\omega(\mathbb{R})$, the stalk C_0^ω is the ring of Taylor series with nonzero radius of convergence, and therefore maps *in*jectively into $\mathbb{R}[\![x]\!]$.

Let S be an object of C; we'd like to understand $\mathrm{Hom}_{\mathsf{C}}(\mathcal{F}_x, S)$. We know this is $\mathrm{Hom}_{\mathsf{C}}(\varinjlim_{x \in U} \mathcal{F}(U), S) = \varprojlim \mathrm{Hom}_{\mathsf{C}}(\mathcal{F}(U), S)$, and in fact this is $\mathrm{Hom}_{\mathsf{C}_X}(\mathcal{F}, i_{x,*}S)$. This is because the sections of the skyscraper sheaf $i_{x,*}(U)$ for U containing x are just S, so these are just maps between sections of these sheaves, compatible with restriction.

The point is, *stalks are left adjoint to skyscrapers*. That is, there's an adjoint pair $-_x : \mathsf{C}_X \to \mathsf{C} : i_{x,*}$. In particular, stalks will preserve colimits (and therefore stuff like coproducts and cokernels), and skyscrapers preserve limits.

Now, given a map $\pi : Y \to X$, we'd like to understand how sheaves pass back and forth from X and Y. We already have the pushforward $\pi_* : \mathsf{C}_Y \to \mathsf{C}_X$, and it would be pretty cool if it has a left adjoint. It'll be called $\pi^{-1} : \mathsf{C}_X \to \mathsf{C}_Y$, but we can't define it in the same way: the image of an open subset may not be open, so there's no canonical open to associate with a $U \subset Y$. If π is an open embedding, then we already have $\mathcal{F} \mapsto \mathcal{F}|_Y$; we'll have to generalize this, in a way reminiscent of stalks. Given a $U \subset Y$, $\pi(U)$ may not be open, but the open subsets of X containing it is an inverse system: if $V, W \supset \pi(U)$, then $V \cap W$ does too. Since we can't literally take intersections, let's take a colimit again, and define

$$\pi^{-1}\mathcal{F}(U) = \varinjlim_{\substack{\pi(U) \subset W \subset X \\ \text{open}}} \mathcal{F}(W).$$

For example, if $Y \hookrightarrow X$ is a closed embedding, this is a notion of "germs along Y;" that is, functions that extend to some open neighborhood of Y, with the same notion of equivalence. This is very like a stalk, but along any subset.

For example, there's a unique map $\pi : Y \to \mathrm{pt}$, and any set S defines a sheaf over $*$. Then, $\pi^{-1}S(U) = S$, so $\pi^{-1}S$ is the constant presheaf. The point is, this pullback operation is only defined for presheaves. We'll have to do something else, called sheafification, to make sheaves. That said, (π^{-1}, π_*) are are still an adjoint pair on presheaves. We will be able to bump this into an adjoint pair of presheaves, and therefore conclude that π_* commutes with limits (e.g. global sections are a special case of pushforward, so as a corollary, global sections will be right exact!).

One interesting property about sheaves is that since gluing satisfies an existence and uniqueness, then $\mathcal{F}$ is determined by $\mathcal{F}(U_\alpha)$, where $\{U_\alpha\}$ is a basis for the topology on X. The sheaf property is that it's determined by open covers. As a corollary, we can think of the stalk $\mathcal{F}_x$ as the values of $\mathcal{F}$ on a "basis" of tiny open sets around x. Of course, there's no smallest such open set, but we can think of X as having this "basis" of infinitesimal open sets. This doesn't really exist, but it's motivation for the following properties of stalks. In particular, this whole idea is bunk for presheaves.

If $\mathcal{F}$ is a sheaf, then there's a map $\mathcal{F}(U) \hookrightarrow \prod_{x \in U} \mathcal{F}_x$: a section defines a germ at every point on U, and in particular this is unique: every germ is defined on some open subset, giving us a cover on which the germs agree on intersections, so it pulls back to a unique section. And if $\varphi : \mathcal{F} \to \mathcal{G}$ is a morphism of sheaves, functoriality gives us a map $\varphi_x : \mathcal{F}_x \to \mathcal{G}_x$, so in particular $\mathrm{Hom}_{\mathsf{C}_X}(\mathcal{F}, \mathcal{G}) \hookrightarrow \prod_{x \in X} \mathrm{Hom}_{\mathsf{C}}(\mathcal{F}_x, \mathcal{G}_x)$. We can use this to understand some properties pointwise.

Lemma 10.1. *A map $\varphi : \mathcal{F} \to \mathcal{G}$ of sheaves on X is an isomorphism iff $\varphi_x : \mathcal{F}_x \to \mathcal{G}_x$ is an isomorphism for all $x \in X$.*

This will allow us to prove something we stated last lecture.

Claim. If $\mathcal{F}$ is a sheaf of sets on X, then it's isomorphic to the sheaf of sections of a map $Y_{\mathcal{F}} \to X$, called the *espace étalé* of $\mathcal{F}$.

This espace étalé tends to be more useful as a conceptual object than for doing stuff; it's completely insane unless your sheaf already looks like a cover. It will feel like a covering map, but won't be one technically.

Proof. The idea is that the fiber will be the stalks, and the fact that these are defined on a neighborhood of a point give us the topology.

The points of $Y_{\mathcal{F}}$ will be $\coprod_{x \in X} \mathcal{F}_x$, and there's a map of sets $Y_{\mathcal{F}} \to X$ sending $\mathcal{F}_x \to x$. Thus, for every open set $U \subset X$ and $f \in \mathcal{F}(U)$, we'd like the map $f : U \to \pi^{-1}(U) \subset Y_{\mathcal{F}}$ sending $x \mapsto f_x$ to be continuous. Thus, give $Y_{\mathcal{F}}$ the weakest topology making this so. Intuitively, we're parallel-transporting a germ to the points on nearby fibers that are represented by the same $f \in \mathcal{F}(U)$. Thus, it's really a covering-like topology: associated to each stalk is a copy of U, and there's no more topology. And $\mathcal{F}(U)$ is exactly the continuous maps $U \to Y_{\mathcal{F}}$ commuting with π. ⊠

In the case of a covering space, we recover the covering space again; for stuff like vector bundles, though, we end up with a similar cover, but where the fibers are discrete. This is bizarre, yes, but the point is that you can reconstruct a sheaf from its stalks. Anyways, now that we've done this, we can put it back in a box and never talk about it again.

Sheafification. Since sheaves on X are presheaves, we have a forgetful functor $\mathrm{For} : \mathsf{C}_X^{\mathrm{pre}} \to \mathsf{C}_X$. We'd like to have a 'free" functor $-_{\mathrm{sh}} : \mathsf{C}_X^{\mathrm{pre}} \to \mathsf{C}_X$ which is left adjoint to For, which will be called sheafification.

That is, this will satisfy the universal property that if $\mathcal{F}$ is a presheaf, $\mathcal{G}$ is a sheaf, and $\varphi : \mathcal{F} \to \mathcal{G}$ is a map of presheaves, there's a unique map of sheaves $\widetilde{\varphi}$ making the following diagram commute.

$$\begin{tikzcd} \mathcal{F} \arrow[r] \arrow[dr, "\varphi"'] & \mathcal{F}_{\mathrm{sh}} \arrow[d, dashed, "\exists!\,\widetilde{\varphi}"] \\ & \mathcal{G} \end{tikzcd}$$

Stated in terms of adjoints, we'd like a natural identification $\mathrm{Hom}_{\mathsf{C}_X^{\mathrm{pre}}}(\mathcal{F}, \mathrm{For}(\mathcal{G})) = \mathrm{Hom}_{\mathsf{C}_X}(\mathcal{F}_{\mathrm{sh}}, \mathcal{G})$.

From this universal property, we know that if $\mathcal{F}$ is already a sheaf, then $\mathcal{F}_{\mathrm{sh}} = \mathcal{F}$, and therefore $-_{\mathrm{sh}} \circ -_{\mathrm{sh}} = -_{\mathrm{sh}}$. In this sense, it's idempotent, so it's sort of a projection. This will be extremely useful, because we're going to do all sorts of operations on sheaves, and if they end up constructing presheaves, that's all right, and we can sheafify it right back: sheafification allows us to ignore the difference between sheaves and presheaves.

The construction won't change $\mathcal{F}$ much locally, since the issue is with gluing, which is more global. That is, we'll keep the same local data, and re-glue it to satisfy the sheaf axioms. To be precise, we'll construct a new sheaf from the stalks of $\mathcal{F}$. In fact, the espace étalé for $\mathcal{F}$ gives you such a construction: take its sheaf of sections, and you're done.

More concretely, let $\mathcal{F}_{\mathrm{sh}}(U)$ be the set of *compatible* sections $(f_x \in \mathcal{F}_x)_{x\in U} \in \prod_{x\in X}\mathcal{F}_x$; that is, for all $x \in U$, there's a $V \subset U$ containing x and an $f_V \in \mathcal{F}(V)$ such that $f_V|_y = f_y$ for all $y \in V$.

This is a way of saying that a compatible section is a collection of germs of the same function: over a small neighborhood of any point, they all come from the same section. This will allow us to glue, though one has to prove this. Another way to write this is that $\mathcal{F}_{\mathrm{sh}}(U)$ is the equalizer of $\mathcal{F}_x$ for all $x \in U$.

This is not as grotesque as it sounds. Recall that for any set S, the constant presheaf $\underline{S}_{\mathrm{pre}}$; its sheafification is the constant sheaf $\underline{S}$: compatible sections are elements of S on connected subsets of X, which is the only way to glue stalks of the constant presheaf.

Yet another way to word this: we can do everything for sheaves with stalks. Since presheaves also have stalks, this says that if you think of a presheaf through its stalks, you're really thinking of its sheafification.

Now, we have an adjunction $(-_{\mathrm{sh}}, \mathrm{For})$, so limits of sheaves are the same as the limits of the underlying presheaves. In particular, kernels are always the same. However, to calculate colimits of sheaves, you have to sheafify: sheafification preserves colimits, so you can calculate colimits in presheaves, but then you have to fix them. This is more important than it looks.

We'll describe some examples of kernels and cokernels of sheaves next time, but before that, a little more abstract nonsense. There's an analogy between the adjunction $(-_{\mathrm{sh}}, \mathrm{For})$ of presheaves and sheaves with (S^{-1}, For) between R-modules and $S^{-1}R$-modules. This is because both of these realize of the latter as a full subcategory of the former, and so the left adjoint is idempotent (if you localize twice, nothing happens). This notion is called a *categorical localization*, which can be thought of in many ways, including as an idempotent left adjoint.

To wit, let's describe localizations more categorically. Localization of modules can be thought of as a subset not just of R, but as a collection $\mathcal{S}$ of arrows $s : M \to M$ for $s \in S$ (given by multiplication by s). Then, we can *localize* the category Mod_R by making all the arrows in $\mathcal{S}$ invertible, by formally adding their inverses. The resulting category, denoted $\mathcal{S}^{-1}(\mathsf{Mod}_R)$, is equivalent to $\mathsf{Mod}_{S^{-1}R}$. This analogy exploits the same one we used for Yoneda's lemma: that a category is not unlike a noncommutative ring.

Now, we can consider the *multiplicative* (meaning closed under composition) set $\mathcal{S}$ of morphisms of presheaves that induce isomorphisms on all stalks. Then, it turns out that localizing $\mathsf{C}_X^{\mathrm{pre}}$ by $\mathcal{S}$ gives one the category of sheaves!

The same idea is used in homotopy theory, where one localizes Top at the set of maps that are *weak equivalences*, i.e. inducing isomorphisms on homotopy groups, and therefore obtains the *homotopy category*.

Episode XI.

Locally Ringed Spaces: 2/23/16

"I didn't give it a good notation because I didn't like it."

Recall that last time, we defined sheafification, which can be thought of projecting presheaves onto sheaves in a particularly nice way. This allows us to forget the difference between sheaves and presheaves, so to speak; we'll use this to understand colimits of sheaves.

Example 11.1. First, a quick digression, since we got confused last time, on the espace étalé of a skyscraper sheaf. Directly from the sheaf axioms, one can show that if $\mathcal{F}$ is a C-valued sheaf, then $\mathcal{F}(\varnothing)$ is the terminal object (a point for Set, 0 for Ab, and so on). This follows from abstract nonsense: the empty product $\prod_\varnothing S$ is necessarily the terminal object (there's more to think through here). This is what motivates the definition of the skyscraper sheaf $i_*S = i_{x,*}S$ in Example 9.2. For simplicity, assume $x \in X$ is a closed point.

Now, let's construct its espace étalé $\pi : Y_{i_*S} \to X$; for any $y \in X$, $\pi^{-1}(y)$ is the stalk of (i_*S) at y, which is S if $y = x$ or the terminal object $*$ otherwise. Thus, Y_{i_*S} is as a set a copy of X, but with S over the basepoint x instead of a single point; then, we glue each of these points of S to the rest of Y_{i_*S} as if they were all x. The result is X with multiple basepoints, so to speak, and is not at all Hausdorff. However, as topological spaces, we have a pullback diagram

$$\begin{tikzcd}
(U \setminus \{x\}) \times S \arrow[r] \arrow[d] & U \times S \arrow[d] \\
U \setminus \{x\} \arrow[r] & Y_{i_* S}.
\end{tikzcd}$$

We can also use the espace étalé to define sheafification: the sheafification $\mathcal{F}_{\text{sh}}$ is just the sheaf of sections of $Y_{\mathcal{F}}$.

Kernels and Cokernels. Before discussing limits and colimits more generally, let's focus on kernels and cokernels. Let $\varphi : \mathcal{F} \to \mathcal{G}$ be a morphism of sheaves of abelian groups on a space X, and for every open $U \subset X$, define $(\ker \varphi)(U) = \ker(\varphi|_U)$. It's easy to check that this is a presheaf, and a little more work to check that it's a sheaf, too. And this is actually the kernel in Ab_X, in that it satisfies the universal property: it fits into the diagram

$$\begin{tikzcd}
\ker \varphi \arrow[r] \arrow[d] & \mathcal{F} \arrow[d, "\varphi"] \\
0 \arrow[r] & \mathcal{G},
\end{tikzcd}$$

and any other sheaf $\mathcal{H}$ that fits into the same place in the above diagram has a unique map to $\ker \varphi$.

Likewise, a morphism in Ab_X is injective (meaning a monomorphism) exactly when $\varphi|_U : \mathcal{F}(U) \to \mathcal{G}(U)$ is injective for all open $U \subset X$.

Cokernels are a little more interesting. The sheaf assigning $U \mapsto \operatorname{coker}(\varphi|_U)$ is a presheaf, and is the cokernel in the category of presheaves, but it is *not* the cokernel in the category of sheaves; it fails to satisfy the universal property. This is where some of the interesting nuances of sheaf theory pop up.

Example 11.2. We'll let $X = \mathbb{C}$, and let $\mathcal{O}$ be the sheaf of holomorphic functions and $\mathcal{O}^*$ be the sheaf of "invertible," i.e. nonvanishing, holomorphic functions (an abelian group under multiplication). The exponential map $f(z) \mapsto e^{f(z)}$ sends holomorphic functions to nonvanishing holomorphic functions, and commutes with restriction, so it's a morphism $\exp : \mathcal{O} \to \mathcal{O}^*$ in $\mathsf{Ab}_{\mathbb{C}}$.

If a function maps to 1 in $\mathcal{O}^*$, then it must be an integer multiple of $2\pi i$, so it must be locally constant, Thus, it's constant on each connected component of the given open set. Thus, $\ker(\exp) = 2\pi i \underline{\mathbb{Z}}$: the constant sheaf, not the constant presheaf. This agrees with what we just learned about kernels.

Then, $\operatorname{Im}(\exp)(U)$ is the $f^* \in \mathcal{O}^*(U)$ such that $f = e^{2\pi i g}$ for some $g \in \mathcal{O}(U)$. That is, $\log f$ must have a well-defined branch on U. In particular, if $U = \mathbb{C}^*$ and $f = z$, then $f \notin \operatorname{Im}(\exp(U)$. This is a problem: $\mathbb{C}^*$ can be covered by simply connected open sets on which the logarithm exists, but the gluing axiom fails.

However, since $\exp : \mathcal{O} \to \mathcal{O}^*$ is surjective on simply connected open sets, then it's surjective on the level of stalks, even though it's not surjective as a map of sheaves. In other words, we want the sequence

$$\begin{tikzcd}
0 \arrow[r] & \underline{\mathbb{Z}} \arrow[r, "j"] & \mathcal{O} \arrow[r, "\exp"] & \mathcal{O}^* \arrow[r, dashed] & 0
\end{tikzcd}$$

to be a short exact sequence of sheaves, but if we naïvely define the cokernel like the kernel, it isn't. This means that *to define the sheaf cokernel, we sheafify the presheaf cokernel*. In this case, the sheafification of the presheaf cokernel $\operatorname{coker}(j)$ stitches together the stalks, but on stalks exp is surjective, so since a sheaf is completely determined by stalks, this is just $\mathcal{O}^*$ again, which jives with the idea of surjectivity. In the same way, we get that $\operatorname{coker}(\exp) = 0$, as one would expect.

In other words, a surjective map of sheaves (categorically, an epimorphism), is surjectivity on stalks, but *not* surjectivity on all open subsets. Injectivity is equivalent to injectivity on stalks and on open subsets, though.

Since sheafification preserves colimits, this can be generalized: the colimit of a diagram of sheaves is the sheafification of the presheaf colimit (which is just the colimit on every open set).

Example 11.3. This next example is in some sense the same example. Let X be a smooth manifold, $\mathcal{F}$ be the sheaf of smooth maps to S^1, C^∞ be the smooth maps to $\mathbb{R}$ (so just the smooth functions), and $\underline{\mathbb{Z}}$ is the constant sheaf (which is also smooth maps to $\mathbb{Z}$, since $\mathbb{Z}$ is discrete); each of these is a sheaf of abelian groups.

We'd like to understand that $S^1 = \mathbb{R}/\mathbb{Z}$. This comes from the sequence

$$0 \longrightarrow \underline{\mathbb{Z}} \longrightarrow C^\infty \longrightarrow \mathcal{F} \longrightarrow 0,$$

which is short exact. The injectivity of $\underline{\mathbb{Z}} \hookrightarrow C^\infty$ comes from the fact that every map to $\mathbb{Z}$ can be lifted to a smooth map to $\mathbb{R}$, and surjectivity comes from the fact that germs of functions can be lifted on a small neighborhood, so it's surjective on stalks. However, there are open subsets where functions can't be lifted: if $X = S^1$, then the identity map $S^1 \to S^1$ can't be lifted to a map to $\mathbb{R}$. Thus, this is surjective, even though it's not so on the level of open sets.

Example 11.4. Our next example will be the de Rham complex. Let X be a smooth manifold. Let $\underline{\mathbb{R}}$ denote the constant sheaf on $\mathbb{R}$ (locally constant functions) and Ω^1 denote the sheaf of one-forms on X. The exterior derivative gives us an exact sequence

$$0 \longrightarrow \underline{\mathbb{R}} \longrightarrow C^\infty \xrightarrow{\mathrm{d}} \Omega^1 \xrightarrow{\mathrm{d}} \Omega^2 \xrightarrow{\mathrm{d}} \dots$$

However, this is not in general short exact; if Ω^1_{cl} denotes the space of closed one-forms, then the Poincaré lemma just states that the following sequence is short exact.

$$0 \longrightarrow \underline{\mathbb{R}} \longrightarrow C^\infty \xrightarrow{\mathrm{d}} \Omega^1_{\mathrm{cl}} \longrightarrow 0$$

In other words, even considering something very simple about short exact sequences of sheaves gives us cohomology. This can be used to define sheaf cohomology, though we won't return to that anytime soon.

In fact, Example 11.2 is a special case of this, since $\mathrm{d}z/z \in \Omega^1(\mathbb{C}^*)$ is a closed form that's not exact.

Ringed Space. Anyways, we were going to talk about schemes, right? These are not just topological spaces, but ringed spaces: topological spaces with a notion of a ring of functions.

Definition. A *ringed space* is the data $(X, \mathcal{O}_X)$, where X is a topological space and $\mathcal{O}_X$ is a sheaf of rings on X.

The motivating examples are a topological space with continuous functions to $\mathbb{R}$ (since these form a ring), or a smooth manifold with the sheaf C^∞, or an analytic manifold with C^ω (analytic functions). Thus, there are definitely different notions of "function" on a manifold, but the ringed space structure means knowing what kinds of functions (geometric structure) is.

We'd also like to know how to evaluate functions on a ringed space. For an arbitrary $x \in U$ and $f \in \mathcal{O}_X(U)$, it's not clear how to define $f(x)$; we have stalks, but then what? In each of our examples (continuous functions, smooth functions, analytic functions, holomorphic functions, etc.), the stalks $\mathcal{O}_{X,x}$ aren't just rings, but local rings,[32] with the maximal ideal $\mathfrak{m}_x$ of functions which vanish at x. $\mathfrak{m}_x$ is unique, because if $f \in \mathcal{O}_{X,x} \setminus \mathfrak{m}_x$, then $f(x) \neq 0$, so it's nonzero on a neighborhood of x, and therefore invertible in that subset! Thus, $f \in \mathcal{O}_{X,x}^\times$, so $\mathfrak{m}_x$ must be unique.

The point is, evaluating at x is exactly quotienting by $\mathfrak{m}_x$, producing an element of $\mathbb{R}$. The sheaves we care about have local rings for stalks, which is what makes this evaluation work. We'll turn this into a definition of something much more useful than a ringed space.

Definition. A *locally ringed space* is a ringed space $(X, \mathcal{O}_X)$ such that for every $x \in X$, the stalk $\mathcal{O}_{X,x}$ is a local ring.

Thus, all of our basic examples are locally ringed spaces, and in general, given an $f \in \mathcal{O}_X(U)$, we can define $V(f) = \{x \in U : f(x) = 0\}$, and this will end up being a closed set.

Schemes are particular examples of locally ringed spaces. We'll have to define how to produce a sheaf of functions, which we'll probably do next time, but we're almost there. One major takeaway is that schemes behave somewhat like these examples we already have.

We also need to define morphisms. An isomorphism is evident: $(X, \mathcal{O}_X) \cong (Y, \mathcal{O}_Y)$ is the data of a homeomorphism $f : X \to Y$ that identifies the sheaves, i.e. for all open $U \subset Y$, there's an isomorphism $f_* : \mathcal{O}_Y(U) \to \mathcal{O}_X(f^{-1}(U))$.

[32] Recall that a *local ring* is a ring with a unique maximal ideal.

It's less obvious how to define morphisms in general; clearly, we need a continuous $f : X \to Y$, and we want to compare $\mathcal{O}_X$ and $\mathcal{O}_Y$. Functions pull back (because the preimage of an open set is open); in the examples we had before, we checked that the pullbacks of continuous (smooth, etc.) functions were continuous (smooth, etc.). More generally, given an open $U \subset Y$, we have the two rings $\mathcal{O}_Y(U)$ and $\mathcal{O}_X(f^{-1}(U))$, and we want the pullback of functions $f_* : \mathcal{O}_Y(U) \to \mathcal{O}_X(f^{-1}(U))$ to be a ring homomorphism. This is exactly how we defined the pushforward of a sheaf.

Definition. A *morphism of ringed spaces* is a pair $(f, f^\sharp) : (X, \mathcal{O}_X) \to (Y, \mathcal{O}_Y)$ in which
- $f : X \to Y$ is continuous, and
- $f^\sharp : \mathcal{O}_X \to f_*\mathcal{O}_X$ is a morphism in Ring_Y.

That is: for every open subset, we can pull functions back into that open subset. But we can say that more concisely with the sheaf theory we have developed.

It's worth remembering that nilpotents on affine schemes give us functions that aren't determined by their values (well, we do have to set up the structure of a locally ringed space first, but we'll get there), so a function isn't quite a bunch of values at points; it's something that we care to pull back.

This is cool, but we care about ringed spaces. What about these maximal ideals? They tell us what it means for a function to vanish. Back in the world of smooth functions, if $\varphi(y) = 0$ and $x \in f^{-1}(y)$, then $(f^*\varphi)(x) = \varphi(f(x))$ had better be 0 too. This is not preserved by morphisms of ringed spaces (since evaluation isn't defined for germs of functions on ringed spaces), so we need an additional axiom.

If $(f, f^\sharp)$ is a morphism of ringed spaces, passing to colimits induces a map $\mathcal{O}_{Y,y} \to \mathcal{O}_{X,x}$, whenever $f(x) = y$ (this is generally true for a map of sheaves, thanks to the property of colimits). Then, we want this map to send $\mathfrak{m}_y \to \mathfrak{m}_x$.

Definition. A morphism $(f, f^\sharp) : (X, \mathcal{O}_X) \to (Y, \mathcal{O}_Y)$ is a morphism of locally ringed spaces if for every $x \in X$, $y \in Y$ such that $f(x) = y$, the induced map $f_x^\sharp : \mathcal{O}_{Y,y} \to \mathcal{O}_{X,x}$ maps $\mathfrak{m}_y$ into $\mathfrak{m}_x$.

This is actually all the data that we'll need to define schemes: schemes are a full subcategory of locally ringed spaces; specifically, they are the ones that are locally isomorphic to $(\operatorname{Spec} R, \mathcal{O}_{\operatorname{Spec} R})$ (as soon as we define the locally ringed space structure on $\operatorname{Spec} R$), i.e. there are actual isomorphisms on an open cover.

Does this look weird? It's actually not unfamiliar: a smooth manifold is a locally ringed space that's locally isomorphic to $(\mathbb{R}^n, C^\infty)$. This encodes a lot of information; in particular, a continuous map of manifolds is smooth iff it pulls smooth functions back to smooth functions. In the same way, a topological manifold is a locally ringed space locally isomorphic to $(\mathbb{R}^n, C)$ (the sheaf of continuous functions). All the structure of an atlas is encapsulated in this notion of locally ringed spaces.

This notion is extremely general. For example, we can define a complex analytic manifold to be a locally ringed space locally isomorphic to $(U \subset \mathbb{C}^n, \mathrm{Hol})$ (since small discs in $\mathbb{C}^n$ aren't necessarily biholomorphic to all of $\mathbb{C}^n$). In all of the cases we've seen, though, $\mathcal{O}_X(U)$ is always a subset of set maps $U \to \mathbb{R}$ (or $\mathbb{C}$), and in particular functions are determined by their values. This is something that will not be true for schemes.

Next time, we will define $\operatorname{Spec} R$, as a scheme.