MOTIVATING (SHEAVES AND) SCHEMES

BRIAN OSSERMAN

This note is not intended to introduce schemes, but rather to accompany a presentation of their definition (such as [1, §II.1-2]) as an explanation of why they are defined the way they are.

Compared to a classical notion such as that of a projective variety, schemes differ in several key aspects:

- Schemes are constructed locally, using the machinery of sheaves.
- Rather than the usual notion of a projective variety, in which one considers only irreducible varieties, and treats two varieties to be the same if they have the same points (say, over an algebraically closed field), schemes consider objects which are reducible, and more substantively, non-reduced; i.e., rings of functions are allowed to have nilpotents.
- Points of schemes do not correspond to points on varieties in some projective space, and in particular include "generic points", corresponding to entire subvarieties.
- Schemes allow arbitrary "rings of functions", rather than finitely-generated k-algebras.

The last point requires little justification, as it allows one to treat at once both curves over fields and rings of integers. For instance, the number-theoretic fact that every ideal in a ring of integers can be uniquely factored as a product of prime ideals, and the geometric fact that every line bundle on a curve can be realized as the line bundle associated to a divisor (that is, a collection of points with multiplicities), are both manifestations of the general relationship in scheme theory between Weil divisors and Cartier divisors. Way cool.

That leaves the first three points, which we will discuss in turn, attempting to explain why it is natural and reasonable to take this approach.

1. Sheaves and local constructions

A projective variety is, by definition, the vanishing set of a collection of homogeneous polynomials in some projective space. This has the advantage of being a very concrete notion, but also has certain disadvantages. Specifically, one sometimes has an abstractly-defined object which looks locally like a projective algebraic variety, but isn't naturally imbedded in any particular projective space. Motivated by the definition of a manifold as a topological space which is locally homeomorphic to a subset of $\mathbb{R}^n$, it seems reasonable that an "abstract variety", i.e., something which is "locally an algebraic variety", should also be a good object to study.

For instance, if one considers a Grassmannian (the space parametrizing r-dimensional subspaces of a fixed d-dimensional vector space), although it is possible to give a presentation as a projective variety using Plücker coordinates, you might find it substantially easier (see Exercise 1 of Problem Set 1) to argue that every point has a neighborhood which looks like affine space. Thus, it is obtained by gluing together varieties, and we could consider it an "abstract variety" without knowing about Plücker coordinates.

In fact, Weil was motivated to introduce a notion of abstract variety for similar reasons: he used it to construct Picard varieties, which parametrize line bundles on smooth curves. As an early indication of the power of such abstractions, he then used his new theory of Picard varieties to prove the Weil conjectures for curves. See [1, Rem. II.4.10.2, §A.2].

Given that a local construction is a good idea, one might also ask why the machinery of sheaf theory should be used. In the definition of a manifold, this is not necessary: if we have a topological space, it is enough to require that every point have a neighborhood homeomorphic to an open subset of $\mathbb{R}^n$ (and that the total space is Hausdorff).

However, even in the definition of a differentiable manifold, we already see the benefits of sheaf theory, and we will use this case to motivate the use of sheaves in general. Classically, a differentiable manifold could be described as a topological space M, *together with* the data of an open covering $\{U_i\}$ of M, and maps $\varphi_i : U_i \to \mathbb{R}^n$ which are homeomorphisms onto open subsets V_i of $\mathbb{R}^n$. If we denote by $V_{i,j}$ the subset $V_i \cap \varphi_i^{-1}(U_j)$, the requirement is then the transition maps $\varphi_{i,j} : V_{i,j} \to U_i \cap U_j \to V_{j,i}$ are all differentiable (note that this makes sense because $V_{i,j}$ and $V_{j,i}$ are open subsets of $\mathbb{R}^n$).

This extra data is necessary, because we have to use the φ_i to decide which sorts of functions $M \to \mathbb{R}$ should be considered "differentiable". However, this definition has substantial disadvantages, mainly stemming from the fact that although the data of the U_i and φ_i determine a differentiable manifold, there is not a unique (or especially natural) choice of this data for any given manifold. That means that we have to explain when two different choices of $\{U_i\}$ and maps φ_i give the "same" (i.e., isomorphic) differentiable manifolds. More generally, it means that the notion of what constitutes a "good" map between differentiable manifolds is something of a mess to write down.

Sheaves address these drawbacks by providing a competely natural gadget to store all the extra data we need, in much the same way that remembering the complete collection of open sets provides a natural piece of data to determine a topological space. The basic idea is that the extra data we want should specify which functions we will consider "good" functions on our space. In the case of differentiable manifolds, these will be differentiable functions, while in the case of algebraic varieties, they will be algebraic functions. In principle, one could consider many kinds of functions, but it turns out that in general, it works well to consider functions to the line.

Once one has brought sheaves into the picture, it becomes necessary to introduce the sheaf (which we can denote by $\mathcal{O}_U$) associated to an open subset U of $\mathbb{R}^n$. However, this is easy enough: to each open subset V of U, we associate the ring of differentiable functions $V \to \mathbb{R}$. We can now define a differentiable manifold as a topological space M together with a sheaf $\mathcal{F}$ on M, such that every point of M has a neighborhood V on which $\mathcal{F}$ is isomorphic to $\mathcal{O}_V$ for some $V \subset \mathbb{R}^n$. That is, there should be a homeomorphism $U \overset{\sim}{\to} V$, inducing isomorphisms between the rings associated to all the open sets of U and V by the sheaves $\mathcal{F}$ and $\mathcal{O}_U$.

Thus, the general yoga of sheaves is that one first settles on what a space should look like locally (an open subset of $\mathbb{R}^n$, an open subset of $\mathbb{C}^n$, an affine variety...), then introduces an appropriate sheaf of functions on such spaces. This gives us a collection of sheaves on topological spaces which we can consider our "local models". Then, we will consider spaces which are topological spaces together with a sheaf, and such that every point has an open neighborhood isomorphic to one of our local models. The sheaf, defined appropriately, turns out to hold all the extra structure we need, whether we want to study real differentiable manifolds, C^∞ manifolds, complex manifolds, algebraic varieties, or schemes.

2. Non-reduced schemes

In the classical theory of algebraic varieties, one is only interested in the set of points where a collection of functions vanish, and so one only looks at the radical of any given ideal of functions. In scheme theory, we always consider the ideals themselves, so that different objects with the same vanishing set will not be considered equivalent.

For instance, the scheme $\operatorname{Spec} k[\epsilon]/\epsilon^2$, even though it has only a single point, is not the same as $\operatorname{Spec} k$.

Although this sometimes complicates the theory (for instance, it means that a map of schemes, even of finite type over an algebraically closed field, is not necessarily determined by the map on underlying topological spaces), it turns out to be an important and powerful generalization. We mention two examples to give a flavor of how non-reduced schemes arise naturally even in classical geometry.

First, if X is a variety over an algebraically closed field, and we choose a particular (closed) point $x \in X$, then the set of maps $\operatorname{Spec} k[\epsilon]/\epsilon^2 \to X$ with image x is precisely the tangent space of X at x (see Exercise 7 of Problem Set 1). Intuitively, one thinks of $\operatorname{Spec} k[\epsilon]/\epsilon^2$ as a "fat point" consisting of a point together with a tangent vector; saying that the map has image x specifies where the point goes, but the tangent vector is still free to map anywhere, as long as it based at x. We thus have a very simple way to describe something which otherwise would not even be obvious how to define.

Second, non-reduced schemes arise naturally in intersection theory. Recall the problem of intersection multiplicities: we want to say that curves C, D of degrees d, e in the projective plane (say, over an algebraically closed field) always intersect in exactly $d \cdot e$ points, but for this to be correct, we need to count points of tangency with the appropriate multiplicities. This part of the problem is local, so we will work in affine coordinates: say that C, D are distinct smooth curves, given by polynomials $f(x, y), g(x, y)$. The rings of functions for the curves are then $k[x, y]/(f(x, y))$ and $k[x, y]/(g(x, y))$. Intersection of C, D corresponds to taking the ideal generated by both of $f(x, y)$ and $g(x, y)$; let's denote this ideal by I. In classical geometry, we'd only be interested in the radical of I, which describes which points are in the intersection. But it turns out that simply by leaving I alone, we end up with a good description of intersection multiplicity. Our specific assertion is:

Proposition 2.1. *Let P be a point in $C \cap D$, corresponding to a maximal ideal of $\operatorname{Spec} k[x, y]/I$. Then the local ring $(k[x, y]/I)_P$ has finite dimension over k, and that dimension is the intersection multiplicity of C and D at P.*

3. Generic points

Finally, we discuss the systematic use of "generic points" (i.e., points which are not closed) in scheme theory.

One reason to consider generic points is that it makes the theory more natural. The way scheme theory is set up, we know that maps $\operatorname{Spec} B \to \operatorname{Spec} A$ correspond to ring homomorphisms $A \to B$. This is an elegant and powerful correspondence. However, let $k[t]$ be the polynomial ring over a field, and $k(t)$ its field of functions (i.e., the field of rational functions in one variable, with coefficients in k). Then $\operatorname{Spec} k(t)$ consists of a single point, and there should be a map $\operatorname{Spec} k(t) \to \operatorname{Spec} k[t]$ induced by the natural inclusion map of rings. There is of course such a map, but it sends the only point of $\operatorname{Spec} k(t)$ to the generic point of $\operatorname{Spec} k[t]$. Thus, if we didn't allow generic points, this morphism of schemes wouldn't even make sense as a map of the underlying set of points.

Generic points are also important because looking at them tells you what happens at a "generic point" of your scheme – hence the name. For instance, if one considers a pair of varieties X, Y as schemes (so that they have generic points), then a morphism $f : X \to Y$ is an isomorphism on a dense open subset of X if and only if it induces an isomorphism on the generic points. This sets up a convenient and powerful machinery for analyzing where on a given scheme a certain property holds.

Finally, we mention that we will describe a dimension theory for schemes that corresponds closely to dimension theory of rings. If we only looked at closed points, we could still get a good dimension theory for varieties, by looking at lengths of chains of closed subvarieties, but we see that this breaks down more generally. In particular, $X := \operatorname{Spec} k[t]_{(t)}$ should be one-dimensional, as it is the scheme associated to a local ring of dimension one. However, X has only a single closed point, so if we did not include the generic point, it would look no different from $\operatorname{Spec} k$. But the rings are quite different, and I hope that by the time we have studied, say, the valuative criterion for properness, I will have convinced you that it pays to consider X as being more similar to a curve than to a point!

References

1. Robin Hartshorne, *Algebraic geometry*, Springer-Verlag, 1977.