Zariski's Main Theorem

1 Introduction

Zariski's Main Theorem, in the formulation that we present here is in essence a statement about the connectiveness of fibers, however the main essence of Zariski's Main Theorem is in fact about the so-called Theorem on Formal Functions which is a theorem which connects cohomology of small 'Hausdorff'-like neighborhoods (i.e. small neighborhoods that we see in analysis) with stalks of sheaves. Going through this route will allow us to see some beautiful constructions and learn about a central tool from cohomology theory.

2 Preliminary definitions

We will be following the treatment found in Hartshorne [3]. Hartshorne's treatment of Zariski's Main Theorem uses the language of schemes. Using this language will help simplify many of our arguments, so we will use it. We will try to use as little machinery of this subject as possible, and in fact we can get away with assuming that all the spaces we talk about are just plain old (projective) varieties (perhaps with some extra points added to the space). To begin, we will try to cover 'quickly' some of the advanced technology that we want to use. As a disclaimer, for the sake of length of this document, we will not provide full proofs for many of the background results and intermediate lemmas needed for the full proof of either Zariski's Main Theorem, nor the Theorem on Formal Functions.

2.1 Schemes

Let A be a commutative ring with 1. We define $\operatorname{Spec} A$ to be the set of all primes ideals of A. If $\mathfrak{a}$ is an ideal of A, then we define $V(\mathfrak{a}) = \{\mathfrak{p} \in \operatorname{Spec} A \mid \mathfrak{a} \subset \mathfrak{p}\}$. We can give $\operatorname{Spec} A$ the data of a topological space by taking closed subsets to be exactly $V(\mathfrak{a})$ for $\mathfrak{a}$ an ideal of A (this is exactly the same idea as the Zariski topology). These sets form a topology because:

1. $V(\mathfrak{a}) \cup V(\mathfrak{b}) = V(\mathfrak{a}\mathfrak{b})$, and
2. $\bigcap_i V(\mathfrak{a}_i) = V(\sum \mathfrak{a}_i)$.

We can also give $\operatorname{Spec} A$ a sheaf $\mathcal{O}$. For each open set U in $\operatorname{Spec} A$, we define $\mathcal{O}(U)$ to be the set of functions

$$s : U \to \coprod_{\mathfrak{p} \in U} A_{\mathfrak{p}}$$

with the conditions that $s(\mathfrak{p}) \in A_{\mathfrak{p}}$, and that for each $\mathfrak{p} \in U$ there is a neighborhood $V \subset U$ of $\mathfrak{p}$ and elements $a, f \in A_{\mathfrak{p}}$ such that for each $\mathfrak{q} \in V$, $f \notin \mathfrak{q}$ and $s(\mathfrak{q}) = a/f$ in $A_{\mathfrak{q}}$. The above definition is just a very long-winded way of basically describing the regular functions on an open set of a variety. The following theorem will help us with how to think of this sheaf of rings.

Proposition 1. Let A, $\operatorname{Spec} A$ and $\mathcal{O}$ be as above.

a) For any $\mathfrak{p} \in \operatorname{Spec} A$, the stalk $\mathcal{O}_{\mathfrak{p}}$ is isomorphic to $A_{\mathfrak{p}}$.

b) For any element $f \in A$, let $D(f) = \{\mathfrak{p} \mid \mathfrak{p} \notin V(f)\} = V(f)^c$. We have that $\mathcal{O}(D(f)) \simeq A_f$.

c) In particular $\Gamma(\operatorname{Spec} A, \mathcal{O}) \simeq A$.

Definition 1. A **ringed space** is a pair $(X, \mathcal{O}_X)$ consisting of a topological space X and a sheaf of rings $\mathcal{O}_X$. A morphism of ringed spaces $(X, \mathcal{O}_X) \to (Y, \mathcal{O}_Y)$ is a pair $(f, f^\#)$ of a continuous map $f : X \to Y$ and a map $f^\# : \mathcal{O}_Y \to f_*\mathcal{O}_X$ of sheaves of rings on Y.

$(X, \mathcal{O}_X)$ is a **locally ringed space** if the stalk $\mathcal{O}_{X,x}$ is a local ring for each $x \in X$. A morphism of locally ringed spaces is a morphism $(f, f^\#)$ of ringed spaces such that for each point $x \in X$, the induced map of local rings $f^\#_x : \mathcal{O}_{Y,f(x)} \to \mathcal{O}_{X,x}$ is a local homomorphism of local rings. For $x \in X$, the morphism of sheaves

$$f^\# : \mathcal{O}_Y \to f_*\mathcal{O}_X$$

induces homomorphisms of rings $\mathcal{O}_Y(U) \to \mathcal{O}_X(f^{-1}(U))$ for every open $U \subset Y$. If U ranges over every neighborhood of $f(x)$, then $f^{-1}(U)$ ranges over a subset of all neighborhoods of x. Taking direct limits, we have that

$$\mathcal{O}_{Y,f(x)} = \varinjlim_{U} \mathcal{O}_Y(U) \to \varinjlim_{U} \mathcal{O}(f^{-1}(U)) \to \mathcal{O}_{X,x}.$$

So by composing maps, we get a map $\mathcal{O}_{Y,f(x)} \to \mathcal{O}_{X,x}$ we need this to be a local homomorphism: if $\varphi : A \to B$ is a map of local rings, $\mathfrak{m}_A$ and $\mathfrak{m}_B$ are the local rings of A and B respectively, then $\varphi^{-1}(\mathfrak{m}_B) = \mathfrak{m}_A$.

Now we have that $(\operatorname{Spec} A, \mathcal{O})$ is a locally ringed space, and this construction is functorial in the following sense:

If $f : A \to B$ is a map of rings, then we get a map $\operatorname{Spec} B \to \operatorname{Spec} A$ given by $\mathfrak{q} \mapsto f^{-1}(\mathfrak{q})$, and furthermore the morphism $(\operatorname{Spec} B, \mathcal{O}_{\operatorname{Spec} B}) \to (\operatorname{Spec} A, \mathcal{O}_{\operatorname{Spec} A})$ is a morphism of locally ringed spaces (and in fact all such morphism of locally ringed spaces is induced by a homomorphism $A \to B$).

Now with all of these preliminary terminology out of the way, we can finally define schemes.

Definition 2. An **affine scheme** is a locally ringed space $(X, \mathcal{O}_X)$ which is isomorphic as locally ringed spaces to $(\operatorname{Spec} A, \mathcal{O})$ for some ring A. A **scheme** is a locally ringed space $(X, \mathcal{O}_X)$ in which every point has an open neighborhood U such that $(U, \mathcal{O}_X \mid_U)$ is isomorphic to $(\operatorname{Spec} A, \mathcal{O})$ for some ring A (i.e. X is covered by affine schemes). We call X the underlying topological space of the scheme, and $\mathcal{O}_X$ the structure sheaf.

2.2 Properties of schemes

We will now give some properties of schemes (which we will need later on).

Definition 3. A scheme $(X, \mathcal{O}_X)$ is **connected** if X is connected as a topological space, and similarly it is irreducible if X is irreducible.

Definition 4. A scheme X is **integral** if for all open $U \subset X$, $\mathcal{O}_X(U)$ is an integral domain.

Definition 5. A scheme X is **locally noetherian** if it can be covered by open affine subsets $\operatorname{Spec} A_i$ where each A_i is noetherian. X is **noetherian** if it is locally noetherian and compact. Equivalently, X is noetherian if it can be covered by finitely many $\operatorname{Spec} A_i$, each A_i noetherian.

Now we have to be a bit careful with this definition because it could be a priori that X can be covered by another affine cover $\{U_j\}$ such that each $U_j = \operatorname{Spec} B_j$ and that not every B_j is noetherian: we do not specify that every affine open subset of X has to be the spectrum of a noetherian ring. It is also not obvious that every affine noetherian scheme is the spectrum of a noetherian ring. It turns out that being noetherian is actually a local property, so these concerns are not too bad.

Proposition 2. A scheme X is locally noetherian if and only if for every open affine subset $U = \operatorname{Spec} A$, A is a noetherian ring. In particular, an affine scheme $A = \operatorname{Spec} A$ is noetherian if and only if A is noetherian.

Definition 6. A scheme X is **normal** if all of its local rings are integrally closed domains (in their field of fractions), i.e. these are normal domains.

Definition 7. A morphism $f : X \to Y$ of schemes is **locally of finite type** if there exists a covering of Y by open affine subsets $V_i = \operatorname{Spec} B_i$, such that for each i, $f^{-1}(V_i)$ can be covered by open affine subsets $U_{ij} - \operatorname{Spec} A_{ij}$, where each A_{ij} is a finitely generated B_i algebra. f is **of finite type** if in each $f^{-1}(V_i)$ can be covered by a finite number of the U_{ij}.

Definition 8. A **closed immersion** is a morphism $f : X \to Y$ of schemes such that f induces a homeomorphism of X with a closed subset of Y, and furthermore the induced morphism of sheaves $f^{\#} : \mathcal{O}_X \to f_* \mathcal{O}_X$ of sheaves on X is surjective.

Now in the world of schemes, we have this fibered product of schemes. That is, say we have maps $X \to S$ and $Y \to S$ of schemes. Then we have this fibered product of schemes $X \times_S Y$, which is defined in the usual way. This product is universal and unique up to unique morphism.

Definition 9. Let $f : X \to Y$ be a morphism of schemes and let $y \in Y$ be a point. Let $k(y)$ be the residue field of y and let $\operatorname{Spec} k(y) \to Y$ be the natural morphism. Then we define the **fiber** of the morphism f over the point y to be the subscheme

$$X_y = X \times_Y \operatorname{Spec} k(y).$$

The fiber X_y is homeomorphic as topological spaces to $f^{-1}(y)$ so we can identify these two objects together.

Now we need to define some more properties of morphisms of schemes.

Definition 10. Let $f : X \to Y$ be a morphism of schemes. Consider the pullback $X \times_Y X$, which gives us the following commutative diagram.

$$\begin{array}{ccc} X \times_Y X & \xrightarrow{p_1} & X \\ \downarrow{\scriptstyle p_2} & & \downarrow{\scriptstyle f} \\ X & \xrightarrow{f} & Y \end{array}$$

There is a unique morphism $\Delta : X \to X \times_Y X$ such that $p_i \circ \Delta = id$. We call this the **diagonal morphism**. We say that the map f is separated if Δ gives a closed immersion, and in this case we also say that X is separated over Y. A scheme X is said to be **separated** if it is separated over $\operatorname{Spec} \mathbb{Z}$.

2.3 Proj and projective morphisms

So Spec allowed us to construct affine schemes, which we can view as some kind of augmentation of affine varieties. However, we also have projective varieties, and in fact we wish also to define projective morphisms. First we begin with Proj.

Let

$$S = \bigoplus_{d \geq 0} S_d$$

be a graded ring. The irrelevant ideal of S is the ideal of elements of positive degree

$$S_+ = \bigoplus_{d > 0} S_d.$$

Then we define

$$\operatorname{Proj} S = \{\mathfrak{p} \subset S \text{ homogenous prime ideal such that } S_+ \not\subset \mathfrak{p}\}.$$

We can endow $\operatorname{Proj} S$ with the Zariski topology in an analogous way to how we gave $\operatorname{Spec} A$ the Zariski topology for a ring A. And similarly, we can turn $\operatorname{Proj} S$ into a scheme by giving it a structure sheaf $\mathcal{O}$ in much the same way as we endowed $\operatorname{Spec} A$ with a sheaf.

Definition 11. We define the projective n-space over a ring A to be $\mathbb{P}^n_A = \operatorname{Proj} A[x_0, \ldots, x_n]$ (this is analogous to the construction of $\mathbb{P}^n_k$). If $A \to B$ is a morphism of rings, then we get a corresponding morphism of affine schemes $\operatorname{Spec} B \to \operatorname{Spec} A$, and then we have that

$$\mathbb{P}^n_B \simeq \mathbb{P}^n_A \times_{\operatorname{Spec} A} \operatorname{Spec} B,$$

(this construction can be thought of as a change of coordinates from A to B). In particular, since $\mathbb{Z}$ is the initial object in the category of rings, we always have a morphism $\mathbb{Z} \to A$ for any A and thus $\mathbb{P}^n_A \simeq \mathbb{P}^n_{\mathbb{Z}} \times_{\operatorname{Spec} \mathbb{Z}} \operatorname{Spec} A$. This construction generalizes to arbitrary schemes.

Definition 12. Let Y be any scheme, the **projective n-space over Y**, denoted $\mathbb{P}^n_Y$ is $\mathbb{P}^n_{\mathbb{Z}} \times_{\operatorname{Spec} \mathbb{Z}} Y$. A morphism of schemes $f : X \to Y$ is **projective** if it factors into a closed immersion $i : X \to \mathbb{P}^n_Y$ for some n followed by the projection $\mathbb{P}^n_Y \to Y$, i.e. the following diagram commutes.

$$\begin{array}{ccc} X & \xrightarrow{\ i\ } & \mathbb{P}^n_Y \\ & \searrow^{f} & \downarrow \\ & & Y \end{array}$$

2.4 Sheaf of modules

Another important object which we will need for Zariski's Main Theorem are sheaves of modules. In particular we will need the ideas of quasi-coherent and coherent sheaves.

Definition 13. Let A be a ring, and M an A-module. M can be turned into a sheaf $\tilde{M}$ on $\operatorname{Spec} A$ in a similar way to the structure sheaf on $\operatorname{Spec} A$.

Here is also an analogous theorem on this sheaf of modules to what we had for the structure sheaf.

Proposition 3. Let A be a ring, M an A-module, $\tilde{M}$ the sheaf associated to M on $X = \operatorname{Spec} A$. Then

a) $\tilde{M}$ is an $\mathcal{O}_X$ module;

b) for each $\mathfrak{p} \in X$, the stalk $(\tilde{M})_{\mathfrak{p}} \simeq M_{\mathfrak{p}}$;

c) for any $f \in A$, the A_f module $\tilde{M}(D(f)) \simeq M_f$ and in particular $\tilde{M}(X) = \Gamma(X, \tilde{M}) = M$.

This sheaf of modules will be the model for quasi-coherent sheaves. What you should have in your head when you think about quasi-coherent sheaves are vector bundles, and perhaps sections of vector bundles.

Definition 14. Let $(X, \mathcal{O}_X)$ be a scheme. A sheaf of $\mathcal{O}_X$ modules $\mathscr{F}$ is **quasi-coherent** if X can be covered by affine open sets $U_i = \operatorname{Spec} A_i$ such that for each i, there is an A_i module M_i with $\mathscr{F}|_{U_i} \simeq \tilde{M}_i$. $\mathscr{F}$ is **coherent** if in addition, each M_i is a finitely generated A_i module.

An important class of coherent sheaves are the sheaves $\mathcal{O}_X(q)$ for $q \in \mathbb{Z}$. These are defined as $\mathcal{O}_X(q) = \bigotimes^q \mathcal{O}_X(1)$, so now we need to explain what $\mathcal{O}_X(1)$ is. So first we need to consider a graded ring

$$S = \bigoplus_{d \geq 0} S_d.$$

We define the module $S(1)$ as the module obtained by giving S a new grading. More specifically, the degree d elements of $S(1)$ are given by

$$S(1)_d = S_{d+1}.$$

Then we define $\mathcal{O}_X(1) = S(1)^\sim$, i.e. we sheafify this module. In general, we have that $\mathcal{O}_X(n) \simeq S(n)^\sim$.

2.5 Some blurb about derived functors

We know that given a coherent sheaf $\mathscr{F}$, we can find an injective resolution $\mathscr{I}^\bullet$

$$0 \to \mathcal{F} \to \mathcal{I}^0 \to \mathcal{I}^1 \to \dots$$

Forgetting about $\mathscr{F}$, we get a co-chain complex

$$0 \to \mathcal{I}^0 \to \mathcal{I}^1 \to \mathcal{I}^3 \to \dots.$$

Then we define the ***i*-th right derived functor** $R^i\mathcal{F}$ to be the i-th cohomology of this chain complex. That is, we let $d^i : \mathcal{I}^i \to \mathcal{I}^{i+1}$ and then define $R^i\mathcal{F} = \ker d^i / \operatorname{im} d^{i-1}$ (after sheafifying the latter).

2.6 Flatness

One more important notion we need from scheme theory is the notion of flatness. Recall that a module M is flat over some ring R if $- \otimes_R M$ is an exact functor. A module map $f : A \to B$ is a flat morphism of rings if f makes B into a flat A-module.

Definition 15. Let $f : X \to Y$ be a morphism of schemes and $\mathscr{F}$ be a quasi-coherent sheaf of $\mathcal{O}_X$ modules.

1) f is **flat a point** $x \in X$ if $\mathcal{O}_{X,x}$ is flat as a $\mathcal{O}_{Y,f(x)}$-module (by the natural induced map of f).

2) f is a **flat morphism** if it is flat at each point $x \in X$.

3) $\mathscr{F}$ is **flat over Y at a point** $x \in X$ if $\mathscr{F}_x$ (the stalk) is a flat $\mathcal{O}_{Y,f(x)}$ module via the natural map $f^\# \mathcal{O}_{Y,f(x)} \to \mathcal{O}_{X,x}$.

4) $\mathscr{F}$ is **flat** over Y if it is flat at every point $x \in X$.

Flatness is important for our purposes mainly because in some sense it commutes with cohomology. This will be an important lemma for the proof of the Theorem on Formal Functions, and we will state it in the proof of the Theorem itself.

2.7 Completions

One final piece of background construction that we need are completions from commutative algebra. Completions are one of the central ideas of the Theorem on Formal Functions. For a more thorough treatment of completions, see either Aityah-MacDonald [1] and Eisenbud [2].

Let R be an abelian group, and let $R = \mathfrak{m}_0 \supset \mathfrak{m}_1 \supset \mathfrak{m}_2 \supset \ldots$ be a descending filtration. The completion of R with respect to $\mathfrak{m}_i$ is denoted as $\hat{R}$ and is defined by

$$\begin{aligned}\hat{R} &= \varprojlim R/\mathfrak{m}_i \\ &= \{(g_1, g_2, g_3, \ldots) \in \prod_i R/\mathfrak{m}_i \mid g_j \equiv g_i \mod \mathfrak{m}_i \text{ for all } j > i\}.\end{aligned}$$

This construction is very analogous to the completion of a metric space by taking equivalence class of Cauchy sequences, and we can think of each element of $\hat{R}$ in a similar fashion as an equivalence class of Cauchy sequences.

Now we need some conditions for completions to play nice with homological algebra, since as is, the inverse limit is only left exact, so completion is only left exact. So quickly lets define inverse systems. An inverse system of (say abelian groups) is a collection (A_n) with homomorphisms $\varphi_{n',n} : A_{n'} \to A_n$ for $n' \geq n$ such that for each $n'' \geq n' \geq n$ we have that

$$\varphi_{n',n} \circ \varphi_{n'',n'} = \varphi_{n'',n}.$$

So then the inverse limit $\varprojlim A_n$ consists of sequences $(a_n) \in \prod_n A_n$ such that $\varphi_{n',n}(a_{n'}) = a_n$ for all $n' \geq n$. This system satisfies the Mittag-Leffler condition if for each n, there is $n_0 \geq n$ such that for all $n', n'' \geq n_0$,

$$\varphi_{n',n}(A_{n'}) = \varphi_{n'',n}(A_{n''})$$

as subgroups of A_n (so if we continue the analogy with sequences, this kind of sounds like that eventually the tail of the sequence is going to be contained in some neighborhood of the limit). So the important fact about satisfying the Mittag-Leffler condition is that it makes $\varprojlim$ into an exact functor.

Proposition 4. Let

$$0 \longrightarrow A_n \xrightarrow{f_n} B_n \xrightarrow{g_n} C_n \longrightarrow 0$$

be a short exact sequence of inverse systems (this means that f_n and g_n respects the inverse system morphisms). Then we have that

1. if B_n satisfies the Mittag-Leffler condition, then so does C_n, and
2. if A_n satisfies the Mittag-Leffler condition, then the inverse limit is exact, i.e. the sequence

   $$0 \to \varprojlim A_n \to \varprojlim B_n \to \varprojlim C_n \to 0$$

   is exact.

3 Theorem on Formal Functions

Now with the preliminary definitions out of the way. We are now ready to tackle the Theorem on Formal Functions. Let $f : X \to Y$ be a projective morphism of noetherian schemes. Let $\mathscr{F}$ be a coherent sheaf on X and $y \in Y$ is a point. Now we want to look at some 'Hausdorff' like neighborhood near each such y, and moreover we would like to perhaps also look at the preimage of such 'Hausdorff' neighborhoods. Since Zariski open sets are so large, we can only approximate such small neighborhoods. So we define

$$X_n = X \times_Y \operatorname{Spec} \mathcal{O}_y/\mathfrak{m}_y^n.$$

Now recall that $\mathcal{O}_y$ is a local ring and $\mathfrak{m}_y$ is its max ideal. This gives us a filtration $\mathcal{O}_y \supset \mathfrak{m}_y \supset \mathfrak{m}_y^2 \supset \dots$ which we can use to take a completion $\widehat{\mathcal{O}}_y$. Now this completion we can think of the set of convergent power series around the point y, and thus we can think of $\operatorname{Spec} \widehat{\mathcal{O}}_y$ as a small Hausdorff-like neighborhood around the point y. Correspondingly, X_n can then be thought of as some sort of "thickened fiber" over the point y. For $n = 1$, we can identify X_1 with the fiber X_y (or $f^{-1}(y)$). Now we have natural maps that make the Cartesian square commute.

$$\begin{array}{ccc} X_n & \xrightarrow{p_n} & X \\ \downarrow{\scriptstyle f'} & & \downarrow{\scriptstyle f} \\ \operatorname{Spec} \mathcal{O}_y/\mathfrak{m}_y^n & \longrightarrow & Y \end{array}$$

Define $\mathscr{F}_n = p_n^* \mathscr{F}$, then there are natural maps for each n

$$R^i f_*(\mathscr{F}) \otimes \mathcal{O}_y/\mathfrak{m}_y^n \to R^i f'_*(\mathscr{F}_n).$$

Now there is a handy theorem about the right hand side that we can use.

Proposition 5. Let X be a noetherian scheme, and let $f : X \to \operatorname{Spec} A$ be a morphism of schemes. Then for any quasi-coherent sheaf $\mathscr{F}$ on X, we have

$$R^i f_*(\mathscr{F}) \simeq H^i(X, \mathscr{F})^{\sim}.$$

So the above turns into a map

$$R^i f_*(\mathscr{F}) \otimes \mathcal{O}_y/\mathfrak{m}_y^n \to H^i(X_n, \mathscr{F}_n)$$

since f'_* is a map from noetherian X to affine scheme $\operatorname{Spec} \mathcal{O}_y/\mathfrak{m}_y^n$. But now both sides are inverse systems (in n), so we can take inverse limits and get a natural map

$$\widehat{R^i f_*(\mathscr{F})}_y \to \varprojlim H^i(X_n, \mathscr{F}_n).$$

Now a priori, this is only an approximation of $\widehat{R^i f_*(\mathscr{F})}$, but it turns out that this approximation is actually an isomorphism—this is the magic of the Theorem of Formal Functions.

Theorem 1 (Theorem on Formal Functions)**.** Let $f : X \to Y$ be a projective morphism of noetherian schemes, let $\mathscr{F}$ be a coherent sheaf on X, and let $y \in Y$. Then the natural map we had above

$$\widehat{R^i f_*(\mathscr{F})}_y \to \varprojlim H^i(X_n, \mathscr{F}_n)$$

is an isomorphism for all $i \geq 0$.

The proof of this theorem relies on a type of induction argument on sheaves over noetherian schemes.

Proof. Since f is a projective morphism, let us first embed X into $\mathbb{P}^n_Y$ for some N and consider $\mathscr{F}$ as a coherent sheaf on $\mathbb{P}^N_Y$. So this is now just the same case as considering $X = \mathbb{P}^N_Y$.

Lemma 1. Let $f : X \to Y$ be a separated morphism of finite type of noetherian schemes. Let $\mathscr{F}$ be a quasi-coherent sheaf on X and let $u : Y' \to Y$ be a flat morphism of noetherian schemes. So we get the following cartesian square

$$\begin{array}{ccc} X' & \xrightarrow{p} & X \\ \downarrow{\scriptstyle g} & & \downarrow{\scriptstyle f} \\ Y' & \xrightarrow{u} & Y. \end{array}$$

Then for all $i \geq 0$, there are natural isomorphisms

$$u^* R^i f_*(\mathscr{F}) \simeq R^i g_*(p^* \mathscr{F}).$$

Now let $A = \mathcal{O}_y$, and make the flat base extension $\operatorname{Spec} A \to Y$ (i.e. a flat morphism $u : \operatorname{Spec} A \to Y$ as in the above). So this really reduces our problem to the cause $Y = \operatorname{Spec} A$, and y is the generic/closed point of Y. So we have that from the above lemma, $u = id$ by our reduction, $g = f'$ and $p = p_n$, so we have then that

$$id^* R^i f_*(\mathscr{F}) = R^i f_*(\mathscr{F}) \simeq R^i f'_*(p^* \mathscr{F}) = R^i f'_*(\mathscr{F}_n).$$

Now using the fact that we have $f : X \to Y$ is now a map to affine scheme $Y = \operatorname{Spec} A$, and $\mathscr{F}$ is quasi-coherent, we get that

$$R^i f_*(\mathscr{F}) \simeq H^i(X, \mathscr{F})^{\sim},$$

and

$$R^i f'_*(\mathscr{F}_n) \simeq H^i(X_n, \mathscr{F}_n)^{\sim}.$$

So we can just restate our result to be an isomorphism of A-modules

$$\widehat{H^i(X, \mathscr{F})} \xrightarrow{\sim} \varprojlim H^i(X_n, \mathscr{F}_n).$$

Now first we establish the result for $\mathscr{F} = \mathcal{O}_X(q)$ for some $q \in \mathbb{Z}$ on $X = \mathbb{P}^N_A$ ($\mathbb{P}^N_Y$ is now $\mathbb{P}^N_A$). So now $\mathscr{F}_n$ becomes $\mathcal{O}_{X_n}(q)$ for $X_n = \mathbb{P}^N_{A/\mathfrak{m}^n}$. Now it is a result that we have to take on faith that

$$H^i(X_n, \mathscr{F}_n) \simeq H^i(X, \mathscr{F}) \otimes_A A/\mathfrak{m}^n$$

for each n. So by definition of completion, we get that

$$H^i\widehat{(X_n, \mathscr{F}_n)} = \varprojlim H^i(X_n, \mathscr{F}_n) \simeq \varprojlim H^i(X, \mathscr{F}) \otimes A/\mathfrak{m}^n = H^i\widehat{(X, \mathscr{F})}.$$

This proves the theorem for $\mathcal{O}(q)$. This also holds for finite direct sums of $\mathcal{O}(q_i)$ since everything distributes over finite direct sums.

Now that we established the theorem of $\mathcal{O}(q)$, we can prove it now for arbitrary coherent sheaves on X. We will use a descending induction argument on i. Now for $i > N$, both sides are 0, so assume the theorem is proven for $i + 1$.

Now since we are dealing with noetherian schemes, something magical happens. We refer to the following lemma.

Lemma 2. Let X be projective over a noetherian ring A. Then any coherent sheaf $\mathscr{F}$ on X can be written as a quotient of a sheaf $\mathscr{E}$, where $\mathscr{E} = \bigoplus \mathcal{O}(n_i)$ for $n_i \in \mathbb{Z}$.

So let $\mathscr{F}$ now be a coherent sheaf. Let $\mathscr{E}$ be as in the above lemma, and let $\mathscr{R}$ be the kernel so that we get the short exact sequence of sheaves

$$0 \to \mathscr{R} \to \mathscr{E} \to \mathscr{F} \to 0.$$

Now we can tensor the above with $\mathcal{O}_{X_n}$ to define $\mathscr{R}_n = \mathscr{R} \otimes \mathcal{O}_{X_n}$ and $\mathscr{E}_n = \mathscr{E} \otimes \mathcal{O}_{X_n}$. Since tensoring is only right exact, we now have a right short exact sequence

$$\mathscr{R}_n \to \mathscr{E}_n \to \mathscr{F}_n \to 0$$

of sheaves on X_n for each n. Now define $\mathscr{T}_n$ to be the image and $\mathscr{S}_n$ to be the kernel of the map $\mathscr{R}_n \to \mathscr{E}_n$. This gives us short exact sequences

$$0 \to \mathscr{S}_n \to \mathscr{R}_n \to \mathscr{T}_n \to 0$$

and

$$0 \to \mathscr{T}_n \to \mathscr{E}_n \to \mathscr{F}_n \to 0.$$

Now consider the following diagram:

$$\begin{array}{ccccccccc}
H^i\widehat{(X, \mathscr{R})} & \longrightarrow & H^i\widehat{(X, \mathscr{E})} & \longrightarrow & H^i\widehat{(X, \mathscr{F})} & \longrightarrow & H^{i+1}\widehat{(X, \mathscr{R})} & \longrightarrow & H^{i+1}\widehat{(X, \mathscr{E})} \\
\downarrow{\scriptstyle \alpha_1} & & & & & & \downarrow{\scriptstyle \alpha_4} & & \\
\varprojlim H^i(X_n, \mathscr{R}_n) & & \downarrow{\scriptstyle \alpha_2} & & \downarrow{\scriptstyle \alpha_3} & & \varprojlim H^{i+1}(X_n, \mathscr{R}_n) & & \downarrow{\scriptstyle \alpha_5} \\
\downarrow{\scriptstyle \beta_1} & & & & & & \downarrow{\scriptstyle \beta_2} & & \\
\varprojlim H^i(X_n, \mathscr{T}_n) & \longrightarrow & \varprojlim H^i(X_n, \mathscr{E}_n) & \longrightarrow & \varprojlim H^i(X_n, \mathscr{F}_n) & \longrightarrow & \varprojlim H^{i+1}(X_n, \mathscr{T}_n) & \longrightarrow & \varprojlim H^{i+1}(X_n, \mathscr{E}_n)
\end{array}$$

The top row of this diagram comes from completing the associated long exact sequence of homology that we got from the sequence $0 \to \mathscr{R} \to \mathscr{E} \to \mathscr{F} \to 0$. Since these are all coherent sheaves, each $H^i(X, \_)$ is a finitely generated A-module in that line, so completion is exact and thus the first row of the diagram is exact. The bottom row comes from long exact sequence that you get from the sequence $0 \to \mathscr{T}_n \to \mathscr{E}_n \to \mathscr{F}_n \to 0$ and then inverse limits. These modules are by construction finitely generated $A/\mathfrak{m}^n$ modules so we have that they satisfy the descending chain condition. Such modules satisfy the so-called Mittag-Leffler condition which basically means that inverse limits is an exact functor on such modules, thus the last row is exact.

Now the maps $\alpha_1, \ldots, \alpha_5$ are the maps from the theorem. We have that α_2 and α_5 are isomorphisms because $\mathscr{E} = \bigoplus \mathcal{O}_X(q_i)$, and we have established the theorem already for such sheaves. α_4 is an isomorphism by the induction hypothesis (in fact this also implies that α_5 is an isomorphism). Now β_1 and β_2 are the maps that are induced by the sequence $0 \to \mathscr{S}_n \to \mathscr{R}_n \to \mathscr{T}_n \to 0$. We want to see that these two induced maps are indeed isomorphisms.

To see that β_i's are isomorphisms, look at the sequence $0 \to \mathscr{S}_n \to \mathscr{E}_n \to \mathscr{T}_n \to 0$. Taking cohomology of this sequence, and passing to the inverse limit, we get

$$\ldots \to \varprojlim H^i(X_n, \mathscr{S}_n) \to \varprojlim H^i(X_n, \mathscr{E}_n) \to \varprojlim H^i(X_n, \mathscr{T}_n) \to \ldots,$$

and this sequence is exact because all the modules are finitely generated so satisfy the Mittag-Leffler condition. So then if each $\varprojlim H^i(X_n, \mathscr{S}_n) = 0$, then $\varprojlim H^i(X_n, \mathscr{E}_n) \to \varprojlim H^i(X_n, \mathscr{T}_n)$ would be an isomorphism. To see this, we want to see that for any n, there is $n' > n$ such that the map of sheaves $\mathscr{S}_{n'} \to \mathscr{S}_n$ is the zero map. Now X is a noetherian scheme, and noetherian schemes are (quasi-)compact, so the question is really local on X (since X is a finite union of noetherian affines we can just look at each affine piece separately). So say that $X = \operatorname{Spec} B$. Denote by R, E, S_n the B-modules corresponding to the sheaves $\mathscr{R}, \mathscr{E}, \mathscr{S}_n$ (i.e. global sections) and let $\mathfrak{a} = \mathfrak{m}B$ the ideal $\mathfrak{m}$ in B. Now $\mathscr{R}$ is the kernel of the map $\mathscr{E} \to \mathscr{F}$, so R is a submodule of E and that since $\mathscr{S}_n$ is the kernel of the map $\mathscr{R}_n \to \mathscr{T}_n$ or really the kernel of the map $\mathscr{R}_n \to \mathscr{E}_n$, so

$$S_n = \ker(R/\mathfrak{a}^n R \to E/\mathfrak{a}^n E).$$

So S_n is the preimage of $\mathfrak{a}^n E$, i.e.

$$S_n = (R \cap \mathfrak{a}^n E)/\mathfrak{a}^n R.$$

Now it is a result from commutative algebra that there is $n' > n$ such that

$$R \cap \mathfrak{a}^{n'} E \subset \mathfrak{a}^n R,$$

which would imply that the map $S_{n'} \to S_n$ is 0, which implies that the map $\mathscr{S}_{n'} \to \mathscr{S}_n$ is 0 also.

Now that we have that β_1, and β_2 are isomorphisms, lets summarize what we know now about the diagram. We have that $\alpha_2, \alpha_4, \alpha_5$ are isomorphisms, and β_1 and β_2 are isomorphisms, which implies that $\beta_2\alpha_4$ form an isomorphism as well. From here we are going to appeal to the 5-lemma. α_2 and $\beta_2\alpha_4$ are surjetive, and α_5 is injective, thus we have that α_3 is surjective. But $\mathscr{F}$ is an arbitrary coherent sheaf on X, so this holds for all coherent sheaves on X. $\mathscr{R}$ is a coherent sheaf on X, so this must hold for $\mathscr{R}$ also, hence α_1 is surjective as well. So this implies that $\beta_1\alpha_1$ is surjective. So $\beta_1\alpha_1$ is surjective, while α_2 and $\beta_2\alpha_4$ are injective, which implies that α_3 must be injective as well. This proves the theorem! □

Corollary 1. Let $f : X \to Y$ be a projective morphism of noetherian schemes, and assume that $f_*\mathcal{O}_X = \mathcal{O}_Y$. Then $f^{-1}(y)$ is connected for every $y \in Y$.

Proof. Suppose for contradiction that $f^{-1}(y) = X' \cup X''$ as a disjoint union of two closed sets X' and X''. So then for each n we have that

$$H^0(X_n, \mathcal{O}_{X_n}) = H^0(X'_n, \mathcal{O}_{X_n}) \oplus H^0(X''_n, \mathcal{O}_{X_n}).$$

So then by the theorem we have that

$$\begin{aligned} \widehat{\mathcal{O}}_y = \widehat{(f_*\mathcal{O}_X)}_y &= \varprojlim H^0(X_n, \mathcal{O}_{X_n}) \\ &= \varprojlim (H^0(X'_n, \mathcal{O}_{X_n}) \oplus H^0(X''_n, \mathcal{O}_{X_n})). \end{aligned}$$

Since H^0 are just rings and inverse limits distribute over (finite) direct sums of rings, we have then that

$$\widehat{\mathcal{O}}_y \simeq \varprojlim H^0(X'_n, \mathcal{O}_{X_n}) \oplus \varprojlim H^0(X''_n, \mathcal{O}_{X_n}).$$

But $\widehat{\mathcal{O}}_y$ is a local ring, so it cannot be the direct sum of two rings. Indeed, suppose that $(R, \mathfrak{m})$ is a local ring such that $R \simeq R_1 \oplus R_2$, then we have that $(1,0) + (0,1) = 1 \in R$, but on the other hand $(0,1) \cdot (1,0) = 0 \in R$, so they are nonunits and are thus contained in $\mathfrak{m}$, but $1 \notin \mathfrak{m}$, so impossible. □

Theorem 2 (Zariski's Main Theorem)**.** Let $f : X \to Y$ be a birational projective morphism of noetherian integral schemes, and assume that Y is normal. Then for every $y \in Y$, $f^{-1}(y)$ is connected.

Proof. By the previous corollary, we need only show that $f_*\mathcal{O}_X = \mathcal{O}_Y$. Now $f^{-1}(y)$ being connected is a local property, so we can assume that $Y = \operatorname{Spec} A$ (i.e. y is in a single affine chart). Then $f_*\mathcal{O}_X$ is a coherent sheaf of $\mathcal{O}_Y$ algebras (since we have a ring map $f^\# : \mathcal{O}_Y \to f_*\mathcal{O}_X$), so $B = \Gamma(Y, f_*\mathcal{O}_X) = H^0(Y, f_*\mathcal{O}_X)$ is a finitely generated A module. But A and B are integral domains with the same quotient field and since Y is normal, A is integrally closed (normal domain), thus $B = A$, but since Y is affine, this implies then that $f_*\mathcal{O}_X = \mathcal{O}_Y$. □

What does this theorem say? Intuitively this means that every such birational morphism only has one branch (think in terms of complex analysis, e.g. $x \mapsto x^2$). This has applications in for instance the resolution of singularities.

References

[1] Michael Atiyah and Ian MacDonald. *Introduction to Commutative Algebra*. CRC Press, 2018.

[2] David Eisenbud. *Commutative Algebra With a View Toward Algebraic Geometry*, volume 150. Springer Science & Business Media, 2013.

[3] Robin Hartshorne. *Algebraic Geometry*, volume 52. Springer Science & Business Media, 2013.