Moroccan Journal of Algebra and Geometry with Applications
Supported by Sidi Mohamed Ben Abdellah University, Fez, Morocco

Volume 3, Issue 1 (2024), pp 1-22

ISSN: 2820-7114

Title :

An accessible calculation of the stalks of the structure sheaf of the affine scheme of an integral domain

Author(s):

David E. Dobbs

Moroccan Journal of Algebra and Geometry with Applications
Vol. 3(1) (2024), 1–22

Supported by Sidi Mohamed Ben Abdellah University, Fez, Morocco **ISSN: 2820-7114**

An accessible calculation of the stalks of the structure sheaf of the affine scheme of an integral domain

David E. Dobbs
Department of Mathematics, University of Tennessee, Knoxville, Tennessee 37996-1320
e-mail: *ddobbs1@utk.edu*

Communicated by Najib Mahdou
(Received 26 June 2023, Revised 16 September 2023, Accepted 19 Septembrer 2023)

Abstract. Assuming minimal background in algebra and topology, we give a proof that for a domain A, the stalk of the structure sheaf of the affine scheme Spec(A) at a point P is A_P. While being more accessible than the standard proof, the proof that is given here leaves few or no ambiguities or questions concerning the foundations of mathematics. Such ambiguities arise inevitably in the standard proof which considers, more generally, A to be an arbitrary commutative ring with 1. An appendix surveys some of the history involving such ambiguities in the mathematical and philosophical literature of the past 100 years.

Key Words: Integral domain, Zariski topology, localization, stalk, sheaf, Hilbert symbol, direct limit, commutative ring.
2010 MSC: Primary 13G05, 13A15, 13B30, 14A05; Secondary 18A30, 03A05, 14A15.

1 Introduction

All rings considered here are assumed to be associative and unital; except in Appendix II and in comments about Appendix II in this Introduction, all rings are also considered to be commutative. All inclusions of rings, ring extensions, subrings, algebras and ring/algebra homomorphisms will be assumed unital. Proper inclusions will be denoted by $\subset$. In connection with any commutative ring A, we will use the following standard notation: U(A) denotes the set of units of A; Spec(A) denotes the set of all prime ideals of A; and if $c \in A$, then A_c denotes the localization of A at the multiplicatively closed set generated by c (that is, at $\{c^n \mid n \geq 0\}$, where $c^0 := 1$). It will be convenient to refer to a (commutative) integral domain as a *domain*.

Let A be a ring and let $X = \text{Spec}(A)$ endowed with the Zariski topology. Recall that a basic open set in that topology is of the form X_a (more often nowadays denoted by $D(a)$), which for any element $a \in A$, is defined by

$$X_a := \{P \in X \mid a \notin P\}.$$

Now, let P be a point of the topological space X (that is, let P be a prime ideal of A). For more than 60 years, the fundamental fact that has allowed objects isomorphic to X (along with certain morphisms in some category) to constitute the affine foundations of modern algebraic geometry is that X can be given the structure of a local ringed space whose structure sheaf has its stalk at the point P given by the direct limit

$$\varinjlim_{P\in X_a} A_a = \varinjlim_{a\in A\setminus P} A_a$$

which is canonically isomorphic to A_P (as A-algebras).

The isomorphism that was just mentioned presents challenges in virtually every classroom where it is taught. (The same can be said of the implicit assumption in the preceding paragraph that student readers are familiar with terms such as “local ringed space", “structure sheaf", “stalk" and “direct limit".) The challenges to the students can be overwhelming. In Appendix I, I list 31 specific questions that can arise (and, in my experience, have often arisen) when an instructor presents a proof of the above isomorphism *verbatim* as it had been given in a well-respected, time-honored textbook. These questions are part of a blizzard of queries that many students (and their teachers) encounter when trying to understand the standard proof of the above isomorphism for the general context that was given above. The reality of the situation is that, except for the unusual class populated by students whose undergraduate studies included much of what most universities consider graduate-level material, the typical student in a beginning graduate-level course on modern algebraic geometry is simply not ready for a presentation emulating the austere sophistication of Grothendieck and Dieudonné (as in [15]). Put simply, in my experience, many students in such a course simply do not have the background to appreciate (that is, to understand) the above isomorphism in the generality that I have stated it. For instance, those students may not yet have heard of the notions of a direct limit or a sheaf (or the stalk of a sheaf or the "germ" of a function at a point). Most instructors should probably not simply assume that their students have already taken some relevant courses on subjects such as algebraic topology or differential geometry or high-dimensional real or complex analysis. Bearing in mind that in any lecture or conversation, a teacher should expect their audience to be able to carry away at most two or three of the most salient facts from that interaction, the following question naturally arises in the mind of someone planning to teach the above isomorphism. (I am now addressing some of the challenges that instructors must decide how to face.) How should a teacher (dare I say/insert, "best") first acquaint students with the just-mentioned isomorphism **if** those students have (essentially only) the following mathematical background: apart from fields (and possibly polynomial rings), the only rings that they have studied are domains; they are comfortable with fractions in the context of a fixed quotient field of a domain; they are familiar with prime ideals in the context of $\mathbb{Z}$ (perhaps also in the context of polynomial rings in one indeterminate over a field, perhaps more generally in the context of Euclidean domains, perhaps more generally in the context of principal ideal domains) and they *have* seen the definition of a prime ideal for some class of domains broader than the singleton set $\{\mathbb{Z}\}$? In short, while students in such a course have had some exposure to point-set topology (also known as general topology), it is often the case they have not studied algebraic topology or graduate-level analysis (so, to repeat, they typically have no knowledge of topics such as sheaves, direct limits, inverse limits, germs of functions, etc.).

Section 2 contains my suggested answer to the above question of how an instructor should/could/may best plan their first presentation of the isomorphism $\varinjlim_{P\in X_a} A_a \cong A_P$. That answer has worked well in classes populated with a majority of students having the kind of background described in the preceding paragraph. The detailed approach in Section 2 is occasionally presented in an informal, conversational style, somewhat as one may expect from time to time during a lecture, and readers should feel free to alter that specific content in accordance with their teaching style (and the composition and the perceived needs of their audience). As mentioned above, Appendix I mentions some of the ambiguities that can distract students who are trying to understand the standard proof for the general context. In my opinion, the proof in Section 2 avoids essentially all of those ambiguities. Of course, those ambiguities must be addressed at some time, but let us remember that "sufficient unto the moment is the complexity thereof". That maxim which I just "recalled" (honestly, I really just invented it) *is* part of the time-honored "cyclic method" approach to learning which we have all experienced and which most good teachers instinctively use in teaching most classes. Among teachers of calculus and analysis, there is general agreement that one should first learn about limits, continuity and ε-δ arguments for real-valued functions of one real variable, *then* cycle back to a deeper study (with teachers expecting deeper understanding from students) of these topics in subsequent courses (for instance, on advanced calculus) while studying real-valued functions of "several" (finitely many real) variables, and *only then* cycle back to yet deeper studies of these topics in a variety of courses (on subjects such as complex variables, metric spaces, differentiable manifolds, etc.).

Similarly, among teachers of topology, there is general agreement that students should have some of the just-mentioned experience before being placed into a course on point-set topology (or some deeper topic). There was a time, essentially when Birkhoff and MacLane wrote their 1941 text introducing the axiomatic "modern algebra" movement from Europe to an English-reading audience in North America, that algebraists were similarly devoted to the cyclic method of teaching. Indeed, even in the 1953 revised edition of their textbook, Birkhoff and Mac Lane scarcely speak of "rings", while emphasizing instead the study of $\mathbb{Z}$ and integral domains. When and why, I must earnestly ask, did teachers of algebra decide to emphasize almost-maximal generality in beginning courses? You may protest and say that some textbooks on abstract algebra nowadays still adopt a "domains first" approach – and you would be correct to assert that. But, now that we seem to have agreed on the usefulness and appropriateness of such an approach, why should we not also agree that there should be a time and place to implement it at the beginning of a course on modern algebraic geometry? How, a busy and harried teacher may well ask, can I do that –*where* should I look for advice on *how* to do that? I humbly suggest that Section 2 gives what is at least a start to the answer to such honorable questions.

Two other appendices should be mentioned here. In my experience while doing research on domains, I have encountered a significant number of workers in the field whose work avoids using any categorical or homological methods or references. In several cases, I have found these workers to be very intelligent and inspiringly creative, especially in constructing elaborate examples, but often without their being aware of some useful methods to generalize such constructions or their contexts. Sometimes, workers of this kind prefer ideal-theoretic, rather than module-theoretic, methods. Sometimes, they prefer their "domains" to be rngs (that is "domains which need not have a multiplicative identity"). Because I believe that workers such as these could offer more to the mathematical mainstream by adopting module-theoretic methods and the appropriateness of assuming that domains *should* have a multiplicative identity, I have written Appendix II. As I believe that "De gustibus non est disputandum," I cannot hope to *prove* that the just-mentioned colleagues have misplaced priorities or values. I can only hope that Appendix II will give food for thought to many. If any reader feels that my comments in this paragraph have insulted you or your mathematical heritage, please accept my sincere apology. My intent is honorable, even if you may conclude that my actual efforts have been clumsy or unseemly. The path to self-improvement can be strewn with reversals, misunderstanding and suspicion. I mean well and I wish you well.

Finally, let me say a few words about Appendix III. This has to do with a theme that underlies many of the above-mentioned 31 questions that often arise when students are shown the traditional proof that $\varinjlim_{P\in X_a} A_a \cong A_P$. For more than 100 years, serious scholars of (meta)mathematics have striven to find an appropriate universe of discourse and to understand how to arrange and access the objects of that universe. Many working algebraists are familiar with some of the history involving the Axiom of Choice and the Well-Ordering Principle, but I would expect that few readers of this article know much about Hilbert's attempt in 1923 to sidestep such topics by introducing what he called the operators ϵ and τ. I would also not expect that many readers would know that there is, to this day, ongoing research extending Hilbert's work and forming a school of "epsilontic calculus?. Appendix III gives a brief account of some current work of that school of thought, along with contributions due to Hilbert, Bourbaki and Grothendieck in regard to what I have described as "the above-mentioned ambiguities".

As usual, $|\mathcal{U}|$ denotes the cardinal number of a set $\mathcal{U}$. Any unexplained material is standard, as in [4], [12], [16].

2 A proof for integral domains

Good day, students. Today, you will begin to understand what is perhaps the most important isomorphism at the heart of the "local" aspects of modern algebraic geometry. As you have often heard me say, this material, although it will be new to almost all of you, is being brought to you by the people who arranged the curricula for your earlier studies. So, in order to anticipate at least a part of what you should expect, let us begin with a special case, a very familiar context, where A is a (commutative unital integral) domain with quotient field K. As you know, $K = \{c/d \mid c \in A \text{ and } 0 \neq d \in A\}$. Recall that K is really a ring of fractions $A_{A\setminus 0}$. So, the elements of K are really equivalence classes. But, since domains do not have any interesting zero divisors, the underlying equivalence relation is especially simple, namely: if $c_1, d_1, c_2, d_2 \in A$ with both d_1 and d_2 being nonzero, then (c_1, d_1) is equivalent to (c_2, d_2) if and only if $c_1 d_2 = c_2 d_1$ in A (in which case, we have the same equivalence classes, $c_1/d_1 = c_2/d_2$). Now – and this is important if today's special case is going to be easily understood – I am going to ask you to forget about thinking of these fractions as equivalence classes. After all, you have been working with fractions (albeit, of integers) since elementary school. And I believe that you are very comfortable in working with them, without having to worry about where such fractions may "live". Inside that "home" where they live – which is the quotient field K that we fixed above – we will establish the kind of isomorphism that we want by building a special kind of union, called a directed union, of certain rings of fractions that are each subsets of that "home", K. In a later class, you will learn that when A is only a commutative ring, it is not so intuitively easy to understand where the various relevant rings of fractions live and the directed unions that we will see today will be generalized to "direct limit" processes by which these rings are somehow combined. Suffice it to say here that understanding direct limits will require you to do some additional foundational work. But, fortunately, none of that additional work will be necessary here today, where all of our rings of interest will be domains.

So, we're back to considering a domain A with quotient field K. Can you think of a way to build K as a union of some interesting rings that contain A? No? Well, let me suggest trying the rings of the form A_a. Recall that if $0 \neq a \in A$, then $A_a := \{c/a^n \in K \mid c \in A, n \geq 1\}$. Isn't it clear that $A \subseteq A_a \subseteq K$ for all such a, and also that $\cup_{0 \neq a \in A} A_a = K$? Yes? Yes! Good! What? Oh, you'd like to see an example. Sure! Let's consider $A := \mathbb{Z}$, so $K := \mathbb{Q}$, and let's take $a := 2$. Then in this example, $A_a = \mathbb{Z}_2 = \{c/2^n \in \mathbb{Q} \mid c \in \mathbb{Z}, n \geq 1\}$. And in this example, $3/4 \in A_a$ but $4/3 \notin A_a$. Is that all clear now? Good! Let's move on.

It would be nice if the "building blocks" A_a were all "comparable", in the sense that whenever a_1 and a_2 are nonzero elements of the domain A, then either $A_{a_1} \subseteq A_{a_2}$ or $A_{a_2} \subseteq A_{a_1}$. If that happens, then the set of the rings A_a is linearly ordered (some people call that sort of thing "totally ordered") and the building blocks would "line up" neatly. What a terrific way that would be to visualize K! Unfortunately, most familiar domains do not have those building blocks line up linearly. For instance, if $A = \mathbb{Z}$, then A_2 and A_3 are not comparable, since $3/2 \in A_2 \setminus A_3$ and $2/3 \in A_3 \setminus A_2$. But, for any domain A, the union of the building blocks is an example of what is called a "directed union", in the following sense: if a_1 and a_2 are any nonzero elements of A, there there exists some nonzero element $a \in A$ such that $A_{a_1} \subseteq A_a$ and $A_{a_2} \subseteq A_a$. Can anyone suggest how to find such an element a? What? Yes, taking $a := a_1 a_2$ *does* work. Thank you for that input. Do you all see why both A_{a_1} and A_{a_2} are contained in $A_{a_1 a_2}$? Some of you are shaking your heads. Well, please consider this: if $c \in A$ and $n \geq 1$, then $c/a_1^n = c a_2^n/(a_1 a_2)^n$. Right? Good – you're all nodding your heads. Isn't it great when we can use some old familiar algebra, even arithmetic, to validate a conjecture? Well, I'm glad that you're still with me.

Let's summarize what we've done so far. If A is a domain with quotient field K, then K is the directed union of the domains of the form A_a as a runs through the set $A \setminus \{0\}$. More formally, $K = \cup_{a \in A\setminus 0} A_a$. Let's spend some time explaining what it means for that index set to be "directed".

Most folks agree that a set I, equipped with a binary relation $\leq$ on I, is called a *directed set* if the following three conditions hold: $\leq$ is reflexive (you know that this means that $i \leq i$ for all $i \in I$); $\leq$ is transitive (you know that this means that if $i, j, k \in I$ satisfy $i \leq j$ and $j \leq k$, then $i \leq k$); and $\leq$ is directed (for most of you, this may be a new concept: this means that if $i, j \in I$, then there exists $k \in I$ such that $i \leq j$ and $j \leq k$). Isn't it clear that we have shown that K is a directed union of the domains A_a where $0 \neq a \in A$. What's that? Oh, you want to know how to define the relation $\leq$ in this case, right? Well, as in most cases involving sets with enriched structures (some people call these "concrete categories"), the relevant relation is either inclusion or reverse inclusion. These two kinds of relations are often directed because, if U and V are subsets of W, then $U \cap V$ is a subset of both U and V, while both U and V are subsets of $U \cup V$. Of course, the theory of an a enriched structure is often richer than set theory, since $U \cap V$ and/or $U \cup V$ may not share the same kind of enriched structure that U and V shared. In our example, if a_1 and a_a are nonzero elements of a domain A, then $A_{a_1} \cap A_{a_2}$ is a domain, but it may not be of the form A_a for some $a \in A$. Moreover, $A_{a_1} \cup A_{a_2}$ may not even be a domain. In fact, it may not be closed under addition – for homework, please construct an example showing this fact. Fortunately, our example does not need to use intersections or unions to establish the "directed" property. Do you recall that both A_{a_1} and A_{a_2} are subsets of $A_{a_1 a_2}$? Good! *That* is why we were able to view K as being a directed union of the rings A_a. What's that? Yes, I only verified the third axiom for a directed set. You see, the other two axioms are about reflexivity and transitivity, and those properties always hold because of basic set theory for any relation $\leq$ which has been induced by either inclusion or reverse inclusion. I apologize for not mentioning that earlier. Please keep it in mind for the future, because I probably won't remember to say it again!

You may be wondering if the above relation $\leq$ could have been described, perhaps using some equations, in terms of the "arithmetic" of the domain A. Yes, that can – and should – be done. We will do it below, in Proposition 2.1 (d).

Now, let's begin to generalize the above result to the context that really matters here: A is still a domain, but another piece of data is a prime ideal P of A. (Remember that can be summarized by writing $P \in \mathrm{Spec}(A)$.) You will come to see that what we did above really treated the case $P = 0$ (which *is* a prime ideal of A because A *is* a domain). The general fact that we are aiming for is the following:

$$\cup_{a \in A \setminus P} A_a = A_P$$

describes A_P as a directed union of the domains A_a as a ranges over the directed index set $A \setminus P$. You can easily modify the above reasoning to see that A_P is the just-displayed union. And that union is directed, once again because both A_{a_1} and A_{a_2} are contained in $A_{a_1 a_2}$. But this time, where P may not be 0, it may be less obvious why $a_1 a_2$ is admissible. Earlier (when $P = 0$), we just used the fact that A was assumed to be a domain to conclude that $a_1 a_2$, being the product of two nonzero elements of a domain, must be nonzero. Why, in the present situation, is $a_1 a_2$ admissible? In other words, if both a_1 and a_2 are elements of $A \setminus P$, why is $a_1 a_2$ also an element of $A \setminus P$? Thank you for that answer. It is absolutely right. The answer is: precisely because P is a prime ideal of A! And do you know what that suggests? That last fact did not use the "domain" property of A. Maybe some of this analysis could carry over more generally, to arbitrary commutative rings. Let's spend some time looking into that possibility. Don't worry – we will return to the context of domains long before any blizzard of ambiguities has been forecast by your local mathematical weatherperson.

Let's ease into the general case with a short paragraph involving some review and some topology, then get "radical" (sorry for the bad pun) in the following paragraph, and then get the result (Proposition 2.1) which holds the key to a better understanding of the index set for the above directed union(s).

Let A be a commutative (unital) ring. Consider the set $X := \mathrm{Spec}(A)$. For each $c \in A$, let $X_c := \{P \in X \mid c \notin P\}$. (So, for instance, $X_0 = \emptyset$ and $X_1 = X$.) Recall (cf. [4, Exercises 15 and 17, page 127]) that X can be given the structure of a topological space via the *Zariski topology*, by taking the sets of the form X_c (as c runs though the elements of A) as a basis for the open sets. Indeed, given the above information about X_0 and X_1, one gets this topological conclusion directly from the definition of a prime ideal of a commutative ring, as that easily gives that $X_a \cap X_b = X_{ab}$ for all $a, b \in A$.

It is well known (cf. [4, Proposition 1.14], [12, Corollary 2.10], [16, Theorem 26]) that if I is an ideal of a (commutative unital) ring A, then the *radical of I (in A)* is the following ideal of A:

$$\sqrt{I} := \{u \in A \mid \text{there exists an integer } n \geq 1 \text{ such that } u^n \in I\} =$$

$$\cap \{P \in \mathrm{Spec}(A) \mid I \subseteq P\}.$$

Part (b) of the next result shows that the above open basis of the Zariski topology can be described in terms of radicals of principal ideals. Part (d) of the next result shows that if the ambient commutative (unital) ring A is a domain, then the above open basis of the Zariski topology can also be described in terms of rings of fractions of the form A_a (with $a \in A$).

Proposition 2.1. (*a*) *Let A be a (commutative unital) ring. Let $a, b \in A$. Then $X_a \subseteq X_b$ if and only if $\sqrt{Aa} \subseteq \sqrt{Ab}$.*

(*b*) *Let A be a (commutative unital) ring. Let $a, b \in A$. Then $X_a = X_b$ if and only if $\sqrt{Ab} = \sqrt{Aa}$.*

(*c*) *Let A be a (commutative unital) domain, with quotient field K. Let $a, b \in A$ such that $a \neq 0$. Then $\sqrt{Aa} \subseteq \sqrt{Ab}$ (that is, $X_a \subseteq X_b$) if and only if $A_b \subseteq A_a$ (that is, if and only if A_b is a (unital) subring of A_a inside K).*

(*d*) *Let A be a (commutative unital) domain, with quotient field K. Let $a, b \in A$. Then $\sqrt{Aa} = \sqrt{Ab}$ (that is, $X_a = X_b$) if and only if $A_a = A_b$ (that is, if and only if A_a and A_b are (unital, but possibly zero) subrings of each other).*

Proof. (a) We have the following equivalences and implications: $X_a \subseteq X_b \Leftrightarrow X \setminus X_a \supseteq X \setminus X_b \Leftrightarrow \{P \in X \mid a \in P\} \supseteq \{P \in X \mid b \in P\} \Rightarrow \cap\{P \in X \mid a \in P\} \subseteq \cap\{P \in X \mid b \in P\} \Leftrightarrow \sqrt{Aa} \subseteq \sqrt{Ab} \Leftrightarrow a \in \sqrt{Ab} \Leftrightarrow$ there exists an integer $n \geq 1$ and an element $\alpha \in A$ such that $a^n = \alpha b$. This (more than) proves the "only if" assertion. To prove the converse, suppose that $\sqrt{Aa} \subseteq \sqrt{Ab}$. Our task is to prove that $X_a \subseteq X_b$; equivalently, that if P is a prime ideal of A such that $a \notin P$, then $b \notin P$. This, in turn, follows easily from P being a prime ideal of A, since the above reasoning gives an equation $a^n = \alpha b$ with $n \geq 1$ and $\alpha \in A$.

(b) It suffices to combine (a) with the assertion obtained by reversing the roles of a and b in (a).

(c) The first parenthetical comment follows from (a); the second parenthetical comment follows from the fact that the operations of addition and multiplication in both A_a and A_b are induced by the corresponding operations in K.

Let us first prove the "only if" assertion. Since $a \in \sqrt{Ab}$, there is an equation $a^n = \alpha b$ for some $n \geq 1$ and $\alpha \in A$. As $a \neq 0$ by hypothesis, then neither α nor b is 0 (since A is a domain). Therefore, as $A_{c^k} = A_c$ (as subsets of K) for all nonzero elements $c \in A$ and all integers $k \geq 1$, we have, in view of the assumption that $a \neq 0$, that $1/b = \alpha/(\alpha b) = \alpha/a^n$ in K, whence $1/b \in A_{a^n} = A_a$, and then it follows easily that $A_b \subseteq A_a$ (as subsets of K).

For the converse, suppose that $A_b \subseteq A_a$. Then, working in the quotient field K of A, we have $1/b = \alpha/a^n$ for some $\alpha \in A$ and some integer $n \geq 1$. Thus $a^n = \alpha b$, whence $a \in \sqrt{Ab}$, whence $\sqrt{Aa} \subseteq \sqrt{Ab}$, as desired.

(d) In view of (b), it suffices, if neither a nor b is 0, to apply (c).

It remains to consider the cases(s) where either $a = 0$ or $b = 0$ (or both). This situation requires separate treatment because of the existence of nilpotent elements. Indeed, notice that if A were only assumed to be a commutative (unital) ring, then $c \in A$ satisfies $X_c = \emptyset$ if and only if c is nilpotent; and, still assuming only that A is a commutative ring, notice that $c \in A$ satisfies $c \in \sqrt{A \cdot 0}$ if and only if c is nilpotent. As the present A is assumed to be a (commutative unital) domain, the assumption that $\sqrt{Aa} = \sqrt{Ab}$ (equivalently, $X_a = X_b$), when coupled with the assumption (of the prevailing case) that either $a = 0$ or $b = 0$, ensures that *both* a and b equal 0 in the domain A. Similarly, while working with the domain A, we have that the assumption that $A_a = A_b$, when coupled with the assumption that either $a = 0$ or $b = 0$, ensures that both A_a and A_b are zero rings, whence *both* a and b equal 0 in the domain A. Thus, under the assumption that either $a = 0$ or $b = 0$ (or both), we have:

$$\sqrt{Aa} = \sqrt{Ab} \Leftrightarrow a = 0 = b \Leftrightarrow A_a = A_b.$$

It is perhaps worth pointing out that when a and b are each equal to the same element $0 \in A$, the use of that element $0 \in A$ in the construction of both of the relevant rings of fractions, A_a and A_b, gives that $A_a = A_0 = A_b$, whence A_a and A_b *are* equal as rings, although *that* ring is a zero ring and not a (unital) subring of K. □

The hypothesis that A is a domain allows the conclusion, via Proposition 2.1 (d), that $\sqrt{Aa} = \sqrt{Ab} \subseteq A$ implies that A_a and A_b are equal rings, to be unambiguous. However, if A had been assumed only to be a commutative (unital) ring, we could hope to (at best) conclude that A_a and A_b are isomorphic rings. Consequently, if one attempts to apply a functor to an unspecified one of the pertinent rings that is isomorphic to A_a, it becomes unclear (that is, ambiguous; that is, known only up to isomorphism) as to what is meant by *the* alleged result of such an application. Yet, *that* is exactly the sort of thing that the literature does, many times over, in this general area when working with commutative (unital) rings A. I believe that during your *initial* exposure to the ring-theoretic foundations of modern algebraic geometry, there is no urgent reason for you to be bombarded with a blizzard of ambiguities. The term "blizzard" is not mere hyperbole here, as you will see if you read my critique in Appendix I of two well-respected expositions of the general case. Also, you will see, if you read Appendix III, that worries concerning the meaning and well-definedness of such applications of functions or functors to unspecified isomorphic copies of a known object have been the topic of ongoing studies for more than 100 years. To temporarily avoid (that is, to forestall) the ambiguities which arise in the general case, we will usually assume for the rest of this section that the ambient (commutative unital) ring A is a domain. Occasionally, we may pause to explain where/how that restriction to domains has simplified matters and avoided ambiguity, but typically we will leave it to you, the reader, to be alert to such instances. I believe that the following is a sound principle, both for students and for researchers: while reading each step of a proof, ask yourself if the step follows as indicated *and* also ask yourself if the conclusion of the step would have been possible under weaker assumptions.

Remark 2.2. Consider the form of the statement that $\varinjlim_{P \in X_a} A_a \cong A_P$. How could one come to understand this statement if it were expressed in its most efficient form? If the ring A is "far" from being a domain then, even if a and b are elements of A such that $X_a = X_b$, it is by no means clear that A_a and A_b are the same mathematical object, because there is no obvious universe containing both A_a and A_b within which one could compare A_a and A_b (in order to see if they are the same). As one can quickly see by tweaking the proof of Proposition 2.1, if $X_a = X_b$, then $A_a \cong A_b$. But that is palpably *not the same* as saying that $A_a = A_b$! Fortunately, we have seen in Proposition 2.1 (d) that if A is a domain, then any quotient field of A is the desirable kind of universe, as we showed that if $X_a = X_b$ for nonzero elements a and b of a domain A (with quotient field K), then we *do* have $A_a = A_b$ (as subsets of K). This suggests that a more efficient (or economical or elegant) description of $\varinjlim_{P \in X_a} A_a$ should be possible, especially if A is a domain, if one were to impose an appropriate equivalence relation of the index set. That is what we will do five paragraphs hence. This completes the remark.

For a fixed domain A (with given quotient field K) and a fixed prime ideal P of A, a reading of Proposition 2.1 (b) suggests (correctly) that it would be useful to define the following equivalence relation $\sim$ on $A \setminus P$. If $a, b \in A \setminus P$, we say that $a \sim b$ if and only if $\sqrt{Aa} = \sqrt{Ab}$; equivalently, if and only if $X_a = X_b$; equivalently (by Proposition 2.1 (c)), that $A_a = A_b$ (as A-subalgebras of K). Note also that by defining $\sim$ in this way, $c \in A \setminus P$ ensures that $c \neq 0$ (for the more general context where A is a commutative ring, $c \in A \setminus P$ would ensure that c is not nilpotent), so that the fussiness involving "such that $a \neq 0$" in the statement of Proposition 2.1 (c) will usually not be a concern as we work with (A and) P.

With A, K and P fixed as above, it may occasionally be necessary to denote the above equivalence relation $\sim$ by $\sim_{A\setminus P}$. If $a \in A \setminus P$, the $\sim$-equivalence class represented by a will be denoted by

$$[a] \text{ or } [a]_\sim \text{ or } [a]_{\sim_{A\setminus P}}$$

with the appropriate notation to be chosen in any given situation as simply as possible, solely in order to avoid ambiguity.

Let us examine more carefully our earlier description of A_P as the directed union $\cup_{a\in A\setminus P} A_a$. How, more precisely, can this directed union be understood to have been expressed in the form $\cup_{i\in I} R_i$ for some directed union of rings R_i indexed by some directed set I? Obviously, one should take $I := A \setminus P$, with the "dummy index" i being replaced by the dummy index a, and with the ring R_i, or rather R_a, then being taken to be A_a. But what is the precise order relation $\leq$ that is underlying this directed union? In other words, if a and b are elements of $A \setminus P$, what does/should it mean to say that $a \leq b$? The answer to this question comes from Proposition 2.1 (d). Indeed, if we say that for $a, b \in A \setminus P$, the definition of $a \leq b$ is that $A_a \subseteq A_b$ (in the quotient field K), then everything falls into place rigorously as desired, because *this* relation $\leq$ is, indeed, reflexive, transitive and directed (with the last of these properties holding since both A_a and A_b are subsets of A_{ab}). Notice also that if $a, b \in A \setminus P$ as above, then we have the following additional formulations of the above equivalence relation, thanks to Proposition 2.1: $a \leq b \Leftrightarrow X_b \subseteq X_a \Leftrightarrow \sqrt{Ab} \subseteq \sqrt{Aa}$.

The above understanding of A_P as the directed union $\cup_{a\in A\setminus P} A_a$ can be made "crisper" (some would say, "sharper" or "more economical" or "more elegant") by using the above equivalence relation $\sim = \sim_{A\setminus P}$. In a moment, I will explain how to do that. When that has been accomplished, I hope that you will agree that we will have a new description of A_P as a new directed union which merits the just-mentioned laudatory adjectives. But my main reason for getting to that new description has to do with some ambiguities in the literature. You see, the literature is not entirely uniform as to the definition of a directed index set. Of course, this fact affects the definition of a directed union (and it also affects, more generally, the definition of a direct limit). While the literature *does* agree that the binary relation $\leq$ on a directed set should be reflexive, transitive and directed (as in the definition that we have been working with here), a noticeable minority of the literature also requires $\leq$ to be antisymmetric (in the usual sense, namely, that if $i, j \in I$ satisfy $i \leq j$ and $j \leq i$, then $i = j$). Unfortunately, requiring the above relation $\leq$ on $A\setminus P$ to be antisymmetric would mean that whenever elements a and b of $A\setminus P$ satisfy $A_a = A_b$, then one would need to have $a = b$. *That* sad situation, for the prime ideal $P = 0$, would imply that $a^2 = a$ for each nonzero element of the domain A. And *that* would imply that $A \cong \mathbb{F}_2$, which is not all what we wanted in this attempt to say something interesting and useful about *all* domains A. So, to placate the above-mentioned minority, the promised "moment" has passed/come, and it is now time to introduce an equivalence relation $\preceq$ which will allow us to replace the index set $A \setminus P$ with the set of $\sim$-equivalence classes from $A \setminus P$. That will be done in the next paragraph.

Given a domain A and a prime ideal P of A, we can define a binary relation on the equivalence classes of the equivalence relation $\sim = \sim_{A\setminus P}$ as follows. If $[a]$ and $[b]$ are such equivalence classes, let us say that $[a] \preceq [b]$ if and only if $a \leq b$. (To avoid ambiguity, you may occasionally prefer to use the notation "$\preceq_{A\setminus P}$" instead of "$\preceq$".) Notice that the binary relation $\preceq$ has been well defined (for if $[a_1] = [a_2]$ and $[b_1] = [b_2]$ with $a_1 \sim b_1$, then we have $A_{a_1} = A_{a_2}$ and $A_{b_1} = A_{b_2}$, along with $A_{a_1} \subseteq A_{b_1}$, whence $A_{a_2} \subseteq A_{b_2}$.) Moreover, it is easy to see (please check this, but do not hand it in as homework, as it really is very easy) that $\preceq$ inherits each of the properties of reflexivity, transitivity and directedness from $\preceq$, and so $\preceq$ endows the set of $\sim_{A\setminus P}$-equivalence classes with the structure of a directed set. Furthermore, this structure has the additional property that is cherished by the "noticeable minority", namely, that $\preceq$ is antisymmetric. Indeed, if $[a] := [a]_{\sim_{A\setminus P}}$ and $[b] := [b]_{\sim_{A\setminus P}}$ satisfy $[a] \preceq [b]$ and $[b] \preceq [a]$, then $a \leq b$ and $b \leq a$, whence $A_a \subseteq A_b$ and $A_b \subseteq A_a$, whence $A_a = A_b$, whence $a \sim_{A\setminus P} b$, whence $[a] = [b]$, as desired. Thus, we now have what *everyone* can agree is a description of A_P as a "directed union" (and, by taking $P := 0$, one would get a description of the quotient field of A as "a directed union"), namely,

$$A_P = \bigcup_{[a] \text{ is a } \sim_{A\setminus P}\text{-equivalence class}} A_a.$$

Ignoring the mini-controversy concerning the definition of a directed set, let us consider a claim to the effect that the last display is more "elegant" than our earlier result that (if P is a prime ideal of a domain A, then) $A_P = \cup_{a \in A\setminus P} A_a$. By sending each a to its equivalence class $[a]$, one obtains a surjection from the second index set to the first index set. (One could, instead, have noted that the Axiom of Choice gives an injection from the first index set to the second index set.) However, it would be wrong to conclude that, in general, the first index set is "smaller than" the second index set. While the cardinal number of the first index set *is* less than or equal to the cardinal number of the second index set, those cardinal numbers could be, depending on A and P, equal infinite cardinal numbers. Consider, for instance, $A := \mathbb{Z}$ and $P := 2A$. Since the set of odd integers is denumerable, this example satisfies $|A \setminus P| = \aleph_0$. In other words, the second index set in this example has cardinal number $\aleph_0$. So, in view of the above-mentioned injection, the first index set in this example is either finite or denumerable. In fact, that first index set is denumerable, since it has a fairly prominent denumerable subset. Let's pause a moment. Did you find or guess what that denumerable subset is? No? Well, thanks for trying. The subset that I noticed is the set of $\preceq_{A\setminus P}$-equivalence classes represented by odd prime numbers. The underlying fact is a gem from elementary number theory: if q and r are distinct odd prime numbers, then $\mathbb{Z}_q \neq \mathbb{Z}_r$. (You can check that this follows from the Fundamental Theorem of Arithmetic.) Since any subset of a finite set is finite, we have proved that the first index set in this example is denumerable; that is, it has cardinal number $\aleph_0$. I will leave this example by asking you to ponder the following question: should you call the first index set (in this example) "more elegant" than the second index set (in this example) *even though* these sets have the same cardinal number?

In looking at various books for the main result that we proved today, you may have come across statements such as

$$\varinjlim_{P \in X_a} \mathcal{F}(X_a) \cong A_P \text{ or } \varinjlim_{a \in A\setminus P} \mathcal{F}(X_a) \cong A_P.$$

So, you know that "$\varinjlim$" is a standard notation for direct limit, and *that* is a generalization of directed union. You may have realized that $\mathcal{F}$ is what is usually called the structure sheaf of the affine scheme $X := \text{Spec}(A)$. (Most algebraic geometers denote $\mathcal{F}$ by $\underline{O}_X$.) Given that we have focused on the result that $\cup_{a \in A\setminus P} A_a = A_P$ (when P is a prime ideal of a domain A), you have probably also surmised that $\mathcal{F}(X_a) = A_a$ (although it may not yet be clear to you whether that equation is a definition or a proven fact). It would be natural for you to wonder what sort of binary relation is being imposed on the index set $A \setminus P$ in the just-displayed statements from the literature. (Let's skim over the technical but important difference between a direct limit and a directed set, and agree that there is something like an underlying ordering on $A \setminus P$ going on in those statements from the literature.) Remember (cf. Proposition 2.1) that when P is a prime ideal of a domain A, $a \leq b$ used to mean that $A_a \subseteq A_b$, and for that context, that this condition was equivalent to $X_b \subseteq X_a$. If b is "later than" a in the relevant directed union or direct limit process (that is, if a is "less than or equal to" b in some sense), it is traditional to have $X_b \subseteq X_a$, so that "later" indexes give "smaller" neighborhoods, and the "functorial" behavior of the sheaf $\mathcal{F}$ gives a "restriction map" from $\mathcal{F}(X_a)$ to $\mathcal{F}(X_b)$, that is, from A_a to A_b. For our familiar domain-theoretic context, this is just the inclusion map $A_a \hookrightarrow A_b$. For more general ring-theoretic contexts, you will eventually develop enough intuition, based on repeated exposure and familiarity with examples, to appreciate whether/when/why restriction maps can or should be regarded as inclusion maps. It will take a while, but I am confident that you can do it. In time, the general setting will seem as natural to you as it was today when you worked with fractions inside a fixed quotient field.

I would like to point out that it is possible to extend the above reasoning in today's main result by generalizing from the multiplicatively closed set $A \setminus P$ (where P is a prime ideal of a domain A) to the case where S is an arbitrary (nonempty) multiplicatively closed subset of a domain A such that $0 \notin S$. As homework, please do the following. Show, under these conditions, that the ring of fractions A_S is a directed union of the domains A_a as the elements a run through the set S. It will be part of the assignment for you to decide what the ordering is on the relevant directed set I (remember to identify I and its ordering and to prove that I *is* directed!). You will also need to explain how, if a and b are suitably related by that ordering, one has that $A_a \subseteq A_b$ (inside the given quotient field K of A). If you wish, for extra credit, you may also try to find conditions under which you can reduce the size of the index set by setting up a suitable equivalence relation on S and letting the "new" indices be the corresponding equivalence classes (instead of the "old" indices which were elements of S).

Next, as an additional step in preparing you for some of the "wrinkles" that may arise when A is a commutative (unital) ring, but not necessarily a domain, let us notice some of what may be gained by slightly tweaking the proof of Proposition 2.1. Suppose that $a, b \in A$. Then: $X_b \subseteq X_a \Rightarrow$ there exists an integer $n \geq 1$ and an element $\alpha \in A$ such that $b^n = \alpha a$. Now suppose that, in fact, $X_b \subseteq X_a$. It will be desirable for the resulting A-algebra homomorphism $f : A_a \to A_b$ (which sends c/a^m to $c\alpha^m/b^{nm}$, for all $c \in A$ and integers $m \geq 1$) to be injective. (Notice that such f exists, thanks to the universal mapping property of rings of fractions, since $a/1$ is a unit in A_b; indeed, it has multiplicative inverse α/b^n there.) To accomplish this injectivity, some fine-tuning will be necessary in regard to the ring elements a and b that will come under consideration in the general case (when A is not necessarily a domain). Rather than delving into that fine-tuning here, let us simply notice why there was no such difficulty in case A is a domain (with quotient field K). In *that* context, with a and b each nonzero elements of a domain A such that $b^n = \alpha a$ for some integer $n \geq 1$ and some (necessarily nonzero) element $\alpha \in A$, we were actually working with a (directed) union in the above argument. Indeed, one gets that $A_a \subseteq A_b$ there since, if $c \in A$ and m is a positive integer, then $c/a^m = c\alpha^m/b^{nm}$ in K and, hence, in A_b.

There will be more fine-tuning as you continue to study the affine scheme Spec(A) (for an arbitrary commutative ring A) and its role in the "local" part of modern algebraic geometry. In addition to what we have just seen here, you will learn new machinery, involving things called direct limits, sheaves and stalks. You will also learn why $\mathcal{F}(U)$ is called the "sections of the sheaf $\mathcal{F}$ over the open set U". That should help you to understand better or more easily some material that you have seen or will see in some courses on topology or analysis (especially, in regard to the "germs" of functions at a point). You will certainly learn that if $\mathcal{F}$ is the structure sheaf of the affine scheme $X = \text{Spec}(A)$, then "the ring of global sections" $\mathcal{F}(X)$ is isomorphic to A. (Here's one final exercise: give a quick proof of this fact, using only information from today's class.) That may lead you to study other historically important representation theorems where a given ring is realized (up to isomorphism) as the ring of global sections of some sheaf on some topological space.

Congratulations! You are about to dive into the really geometric part of algebraic geometry. But that's enough for today.

3 Appendix I: Comparison with the traditional proof

For many years during the 1960s, graduate students in an algebraic geometry course learned the modern approach to the basics of that subject by reading some notes [19] (with red front and back covers) that were affectionately known as "the red Mumford". (The only alternative at that time was to read the variety of French language material that was being produced by Grothendieck and his followers.) In retrospect, Mumford did a good job at getting to the basics, while providing background and examples that were sufficient for his intended audience. I would like to begin this appendix by reviewing Mumford's five-line proof in [19, page 40] that (to use the notation of our Section 2 but now for an arbitrary commutative ring A), the stalk of the structure sheaf (I will denote that sheaf by $\mathcal{F}$) at a point P of $X := \mathrm{Spec}(A)$ (that is, at a prime ideal P of A) is A_P. Apart from some notational changes, the following is only a slight rewording of Mumford's proof:
"Since the sets of the form X_a give a basis of the Zariski topology of Spec(A), we have that the stalk of the sheaf $\mathcal{F}$ [of course, Mumford denotes $\mathcal{F}$ by $\underline{O}_X$] at the point P [of course, Mumford denotes P by x] is

$$\varinjlim_{P\in U} \mathcal{F}(U) = \varinjlim_{P\in X_a} \mathcal{F}(X_a) = \varinjlim_{a(P)\neq 0} A_a.$$

Since all restriction maps in our sheaf are injective, this is just $\cup_{a(P)\neq 0} A_a$, which is clearly A_P."

An objective reading of the above argument raises, in order, 31 questions. (The Introduction promised a "blizzard of queries"!) These questions are collected, together with some comments, as the following seven items:
(i) What does the first equal sign in the display really mean? If two objects (such as the partners in that asserted equality) are each only defined up to isomorphism, what sense does it make to say that those objects are equal? Would it not make more sense to say, instead, that those objects are isomorphic? Or, given that those objects are each only defined up to isomorphism, would an assertion of isomorphism here mean the same thing as your assertion of equality here? With respect, I must ask: is your assertion of equality even meaningful? Was it intentional or was it a typo?
(ii) I suppose that the answer to last part of the above set of questions is that you intended to use that equal sign and there was no typo, as I have now seen that you have continued to use equal signs two more times. Let me ask next about a quantity on the right-hand side of the first equal sign in the display. What does "$\mathcal{F}(X_a)$" really mean? I understand that for each relevant element a, the set X_a is well defined and so is $\mathcal{F}(X_a)$. But in reading about direct limits over directed sets, it was not clear to me if a directed set must be asymmetric. (I know that it must be reflexive, transitive and directed.) Looking online, I see that many people are confused, like I am, about how direct limits are defined, asking whether the index set that we are discussing is what they call an "order" or what they call a "preorder". Since P is given and a is somehow varying, should the subscript of "$\varinjlim$" on the right-hand side of the second equal sign in the display be, instead, a specific statement about the behavior of a, or perhaps, about the behavior of an equivalence class (for some equivalence relation that you have not mentioned) represented by a? If the answer to the last question is "Yes", what is that equivalence relation and what sense does it then make to speak of "$\mathcal{F}(X_a)$"? After all, if elements $a_1, a_2 \in A$ are such that $X_{a_1} = X_{a_2}$, I can probably believe that $\mathcal{F}(X_{a_1}) \cong \mathcal{F}(X_{a_2})$ – would you please give or assign a proof of this fact? – but I would need to be convinced that $\mathcal{F}(X_{a_1}) = \mathcal{F}(X_{a_2})$. Was that a hidden part of the message that you were trying to convey here? Is this question somehow linked to the questions listed under (i)? Will some or all of these concerns in (ii) dissipate if we just decide to not worry whether the index set is asymmetric? I wish that I had asked this part of my question earlier when you covered direct limits, but it only occurred to me now when I saw how you were using them. Maybe no one has ever asked you this before, so perhaps your lesson plans did not anticipate such a question from the audience. If so, please excuse me and I'll wait a bit longer for you to think about this before you answer.

(iii) I have a comment about the index set on the right-hand side of the second equal sign in the display. This is the kind of "specific statement about the behavior of a" that I mentioned in (ii). Would it have been better – or possible –to go directly from the first direct limit to the third direct limit in the display? If so, that would eliminate some (maybe all) of my concerns in (ii) above. By the way, thank you for explaining the notation "$a(P)$" earlier (on the preceding page of [19]), when you said/wrote that "The elements of $\mathcal{F}(U)$ can be viewed as functions on U." I was certainly ready to understand that part of the display!
(iv) I am going to have several questions about your phrase "This is just ...." First, what did you intend the word "This" to refer to?
(v) Next, still about your phrase "This is just", I would like to ask: did you intend the word "is" to refer to "is equal to" or "is isomorphic to"?
(vi) Here are my final questions about your phrase "This is just". What did you mean by the word "just" there? Did you mean "equal to" or "isomorphic to"? I remember that when we covered direct limits, you mentioned that every directed union is isomorphic to a direct limit over a directed set. But is the converse true? In particular, are there some instances of "$\varinjlim_{P\in X_a} \mathcal{F}(X_a)$" that should not be viewed as being isomorphic to directed unions? Would your answer to this last question depend on whether we had grouped the various elements a into equivalence classes as suggested above? If so, what is the relevant equivalence relation?
(vii) Why is your union "$\cup_{a(P)\neq 0} A_a$" well defined? I was always taught that a meaningful union of the form $\cup_{j\in J} W_j$ requires that the objects W_j be well understood sets and that there exists a universe which contains each of these sets W_j as a subset. Is that really the case here? I do not know whether you intended the elements a to range over a subset of A or over certain equivalences classes (again: if so, what is the relevant equivalence relation?), but regardless of your answer to that question, I do not see how there could exist some universe containing each of the relevant sets A_a because each of these "sets" is only defined up to isomorphism. What sense doers it make to talk about a union of things that are only defined up to isomorphism? And what sense would it make for such a union, if it *were* well defined, to be equal to something like A_P, which is itself only defined up to isomorphism? Are these kinds of questions related to what I was asking about in (i) (and occasionally later)? Has some group of mathematical leaders somehow agreed that mathematicians are working in an ideal Platonic world where all isomorphic objects are equal? If so, I did not get that memo and, unlike Leibnitz, I do not necessarily believe in a "pre-established harmony"! Who or what has somehow ordered everything to work so well together? Is some kind of well ordering being supposed and used? Excuse me, I am only trying to learn and understand, but I must add, with respect: your saying "which is clearly" did not seem at all clear to me.

As I recall, the courses in modern algebraic geometry that I took (which were taught by Professors George Rinehart and Stephen Lichtenbaum) presented the above proof (and prepared the class for it) almost exactly as in [19].

Let us next review how, a few years later, Atiyah-Macdonald approached the above result. One should note that although Mumford's notes may have been written under time pressure and were compiled into a "Preliminary version" of the first three chapters of a foreseen book, Atiyah and Macdonald had the advantage of the passage of time and their book was not explicitly a "preliminary version". Also, because of the intentionally small size of [4], much of the substance of that book is to be found in its exercises. Prior to the actual exercise stating, in effect, that $\varinjlim_{P\in U} \mathcal{F}(U) \cong A_P$, Atiyah-Macdonald did a good job of covering the Zariski topology (having the sets of the form X_a as an open basis), direct limits and rings of fractions. We next essentially reproduce the five parts of [4, Exercise 23, page 47]. As before, in order to assure uniformity of notation for comparison purposes, the following summary is the result of only a light editing of what Atiyah and Macdonald wrote in that exercise. As before, we are considering a commutative ring A, $X := \mathrm{Spec}(A)$ has been equipped with the Zariski topology, and the structure sheaf of this affine scheme is being denoted by $\mathcal{F}$. Here,

then, are the five parts of the pertinent exercise from [4]:
(i) If $U = X_a$ for some $a \in A$, show that $\mathcal{F}(U) := A_a$ depends only on U and not on a.
(ii) Let $U^{'} = X_b$ be another basic open set in X such that $U^{'} \subseteq U$ $(= X_a)$. Show that there is an equation of the form $b^n = ua$ for some integer $n > 0$ and some $u \in A$, and use this to define a homomorphism $\rho : \mathcal{F}(U) \to \mathcal{F}(U^{'})$ (that is, $A_a \to A_b$) by $c/a^m \mapsto cu^m/b^{mn}$. Show that ρ depends only on U and $U^{'}$. This homomorphism is called the *restriction* homomorphism.
(iii) If $U = U^{'}$, then ρ is the identity map.
(iv) If $U \supseteq U^{'} \supseteq U^{''}$ are basic open sets in X, show that the composite of the restriction homomorphisms $\mathcal{F}(U) \to \mathcal{F}(U^{'})$ and $\mathcal{F}(U^{'}) \to \mathcal{F}(U^{''})$ is the restriction homomorphism $\mathcal{F}(U) \to \mathcal{F}(U^{''})$.
(v) If P is a prime ideal of A, show that $\varinjlim_{P \in U} \mathcal{F}(U) \cong A_P$.

In commentary after (v), Atiyah-Macdonald went on to state (again, I will lightly edit their notation) the following: "The assignment sending each basic open set U to the ring $\mathcal{F}(U)$, together with the restriction homomorphisms ρ satisfying the conditions in (iii) and (iv), constitutes a *presheaf of rings* on the basis of open sets $\{X_a \mid a \in A\}$" and that "(v) says that the stalk of this presheaf at P is the (quasi-)local ring A_P."

For the most part, I would prefer to let the reader decide the following three things: which, if any, of the earlier 31 questions about the presentation in [19] applies to the presentation in [4]; whether any new questions arise as a result of that presentation in [4]; and whether the presentations in [19] and/or [4] would be preferable to the presentation that I gave (for domains A) in Section 2, when the reader is considering how to present the "stalk" result to his/her/their class. Before leaving instructor/readers with such weighty matters (after all, you surely know your students, their background and their needs better than I do!), I would like to close this appendix by making three sets of points.

First, in teaching graduate courses on commutative ring theory several times at two state universities, I have often given a fuller treatment, than in either [19] or [4], of the identification (of A_P as) the stalk of the structure sheaf of Spec(A) at a prime ideal P of the commutative ring A. Occasionally, because of time pressure in such a course, I have covered only the special case for domains A, as in Section 2 above. But in *all* those courses, I took/found the time to explain what a sheaf is and why the construction at hand actually produces a sheaf. I did so, in part, because I have found the categorical concepts of an equalizer and a coequalizer to be helpful and illuminating, both for research and in teaching, on several occasions. Also, graduate students specializing in analysis, topology and differential geometry have told me that my comments along those lines had been helpful to them in their research. Speaking of teaching, I am uncomfortable in speaking of "the stalk of a presheaf at a point" (as Atiyah-Macdonald did), but perhaps this sort of worry is a personal one that the reader need not be concerned about.

Second, while exercises do not have the same purpose as lecture material, each should be clearly stated, and so, if only for the sake of completeness and fairness, I would like to raise a few questions and/or comments in regard to the presentation in [4, Exercise 23, page 47]. (Yes, that process did begin in the preceding paragraph – thank you for noticing that.) It seems natural to me to ask what Atiyah-Macdonald intended to mean by the phrase "depends on" in (i)? One could ask a similar question about Atiyah-Macdonald's (ii). In regard to their (ii), one could also ask if the name "*restriction* homomorphism" should be appended by something like "from $\mathcal{F}(U)$ to $\mathcal{F}(U^{'})$". Given the previous sentence, I must admit that I found it heartening to see the plural in "restriction homomorphisms" in Atiyah-Macdonald's commentary after (v).

Third, the following advice/principle will, I hope, meet with universal acceptance. Any graduate course on modern algebraic geometry should cover in detail the "stalk" result for the general context of an arbitrary commutative ring A. Whether or not that coverage should be preceded (either in class or as homework) with the special case where A is a domain (as given in Section 2) is a decision that should be up to the instructor (or instructors) who is (are) responsible for such a course. While someone in my position can offer advice, please let me repeat: you surely know your students, their background and their needs better than I do!

4 Appendix II: Should a domain have a multiplicative identity element?

Let R be an rng. (Some authors write "a rng" instead of "an rng". Which option is appropriate grammatically depends on how one pronounces "rng" – should it be like "urng" or like "rung"? – and there is no universal agreement about that pronunciation.) Then with respect to addition, R is an abelian group, that is, a $\mathbb{Z}$-module. Consider the mathematical object $\mathcal{R} := \mathbb{Z} \oplus R$, the external direct sum of $\mathbb{Z}$ and R as an abelian group under addition, but also equipped with a multiplication, given by

$$(n_1, r_1)(n_2, r_2) = (n_1 n_2, n_1 r_2 + n_2 r_1 + r_1 r_2) \text{ for all } n_1, n_2 \in \mathbb{Z} \text{ and } r_1, r_2 \in R.$$

It is straightforward to check that $\mathcal{R}$ is a (unital) ring under the above operations, with multiplicative identity element $(1,0)$. Also, there is an injective rng homomorphism $\theta : R \to \mathcal{R}$, given by $r \mapsto (0,r)$ for all $r \in R$. It is customary to use θ to view R as a subrng of the (unital) ring $\mathcal{R}$. Of course, there would be no practical reason to use the above construction if the given rng R is known to be a ring (that is, if it is known to have a multiplicative identity element, say 1_R). Indeed, in that case, the above rng homomorphism θ would not be unital, the point being that $\theta(1_R) = (0, 1_R) \in \mathcal{R}$ is distinct from the multiplicative identity element $(1,0)$ of $\mathcal{R}$.

At one time, many writers of textbooks on "ring theory" appreciated that one should introduce students early to constructions like the one given in the preceding paragraph. It seemed to take until the late 1960s, or perhaps even the mid-1970s, until a sizable majority of the community coalesced around the idea that a "ring" should have a multiplicative identity element, whereas an "rng" possesses all the properties of a ring except possibly that of having a multiplicative identity element. So, when one is reading textbooks on ring theory that were written long ago, one must take care to understand which definition of "ring" the author is using. Yet, there is often much to be gained from reading old textbooks. One such book that comes to mind is [18]. After I finished the master's degree and just before I moved to the United States to study for the doctorate, a former professor suggested that I should read [18] to learn something about rings. (He was aware that although I had extensive knowledge of group theory and matrix theory, I had never heard the word "ring" uttered in a classroom.) In reading [18], I developed a quick respect for ring theory. Even the second section of the first chapter of [18] had a couple of challenging homework problems. That same section (to be precise, page 8 of that book) contained McCoy's version of the construction in the preceding paragraph. (To be accurate, I should point out the following ultimately insignificant difference: where the above construction used the external direct sum $\mathbb{Z} \oplus R$, McCoy had used the external direct sum $R \oplus \mathbb{Z}$, with the necessary concomitant changes in the definition of multiplication.) McCoy's version of "embedding a ring [I would say "an rng"] in a ring with unity" [I would omit "with unity"] was actually better tailored to the ring R at hand. Indeed, it used essentially what we did in the above paragraph if R has characteristic 0, but replaced $\mathbb{Z}$ with the ring of integers modulo n if the characteristic of R is some positive integer n. (By definition, the characteristic of an rng is the smallest positive integer n such that the sum of n copies of each element of R is 0, if such an n exists; and the characteristic of an rng R is taken to be 0 if no such n exists. Notice that this definition of the characteristic of an rng is, in the case of positive finite characteristic, really talking about the exponent of the additive group of R. Notice also that the characteristic of a/the zero rng is 1, a situation that has led some workers to argue that the notations $\mathbb{Z}/1\mathbb{Z}$ or $\mathbb{Z}_1$ should be used for "the" zero ring, since one could then say that the characteristic of $\mathbb{Z}/n\mathbb{Z}$ is n for all positive integers n. It is worth recording that no one has been foolish enough to suggest extending this practice by letting the notations $\mathbb{Z}/0\mathbb{Z}$ or $\mathbb{Z}_0$ stand for the ring of integers, although $\mathbb{Z}$ *does* have characteristic 0.) It is interesting to note that the definition of "characteristic" had not been finalized as recently as 1953, as that is when the revised edition of the classic textbook, "A survey of modern algebra," by Birkhoff and Mac Lane defined the characteristic of $\mathbb{Z}$ to be ∞ (while acknowledging in a footnote that "most" authors had decided to say that the usage "characteristic ∞" should be replaced by "characteristic 0").

When I was a pre-doctoral student trying to learn about rings and linear transformations on my own, an appealing aspect of the approach to modern algebra in the above-mentioned textbook by Birkhoff and Mac Lane is that (as opposed to the "groups first" approach in the popular textbook, "Topics in algebra," by Herstein) they defined "characteristic" only for domains. That "domains first" approach suited me well, as I was quite comfortable with fields and I was becoming comfortable with some domains (notably, $\mathbb{Z}$ and polynomials rings in one indeterminate over a field). I wonder if the researcher in commutative algebra that I became (spending many years studying many classes of domains) would agree with his younger self that the embedding of an rng in a ring is worth teaching to today's students.

Following [12], a "domain" is defined to be a nonzero commutative rng with no nonzero zero-divisors. With that usage, a "domain" need not be unital, that is, it need not be a ring. For me (and, I suspect, for most readers of this article), a "domain" is unital, that is, it is a ring. That has also been my belief ever since I learned about domains. As my doctoral research had been on what is nowadays called arithmetic algebraic geometry, it was natural for me to see a "domain" as an example of a (necessarily unital) commutative ring. I had not heard of multiplicative ideal theory (or [12] or Robert Gilmer) until almost the end of my postdoctoral year at UCLA (and it was a noncommutative ring theorist, Julius Zelmanowitz, who informed me of the area and who suggested that I familiarize myself with [12]). The fact that a significant number of commutative ring theorists and other mathematicians do not believe that a "domain" needs to be unital was brought to my attention in an anecdote that I relate in the next paragraph. (That anecdote does not reflect me in the highest moral light, but I do find it amusing and instructive – I hope that you will, too.)

One weekday around noon several decades ago, I left my office and went to the mathematics department's mail room to see if the daily mail had been delivered. Discovering that the current issue of the MAA's Monthly magazine had just arrived, I took my copy of that magazine back to my office and quickly turned to the problem section. I found a problem that seemed to be in commutative algebra (that was a rare occurrence in that section of the Monthly during that period of history) and I set to work on it. The problem was about domains, and within moments, I had solved the problem by using a standard tool, the ring-theoretic generalization of the classical result on extending valuations (as in [12, Theorem 19.6] or [16, Theorem 56]). I quickly printed (by hand) my solution, handed it to a secretary to be typed appropriately (professors did not have typewriters or computers at that time, but we did have secretaries to type for us), received the typed copy for proofreading just a few minutes later, found the typing to be perfect (that is, accurate), and managed to get my submitted solution mailed to the MAA before the departmental mail was picked up that day. Surely, I thought, with the MAA office just one or two days away by normal mail, my solution had a good chance of being the first to be received. I eagerly awaited the eventual issue of the Monthly magazine, expecting to see my name next to the published solution. (I promised you that this anecdote would make me look all too human.) To my dismay, my name only appeared in the alphabetic list under the heading "also solved by". The published solution did not look familiar to me. But when I saw that the solution was due to Robert Gilmer, my dismay disappeared. At that point in time, Gilmer was indisputably the world leader in multiplicative ideal theory (and perhaps, more generally, on the topic of domains). My feelings were further assuaged when I read Gilmer's solution and realized that it differed significantly from my solution. In fact, Gilmer's solution seemed slightly longer than mine. (Yes, more human frailty is on exhibit here, but the story is nearly over.) More importantly, Gilmer's solution did not use the assumption that the ambient domain was unital. (Remember that the definition of a "domain" in [12] does not require the unital property.) So, quite likely, I had achieved a Pyrrhic victory, in that my solution was probably the first to arrive at the office of the MAA Monthly, but Gilmer's solution was eventually deemed to be better by the powers that be, presumably because his solution assumed less (and, therefore by the standards that most mathematics would use for such matters, was the "better", more elegant solution). This experience taught me the following valuable lesson: although some things are easier to do when an rng happens to be unital, one should always be alert to the possibility that a result that has just been proved could be susceptible to a different line of reasoning, perhaps coming from a different mathematical genre, leading to a more elegant/economical proof.

The last few paragraphs may have caused some readers who are *aficionados* of domains to wonder if "domains" really should be unital. At this point, I cannot claim that the above material has convincingly presented the case for an affirmative answer. I do think, however, that the next paragraph will help to make that case (especially in the minds of any of the just-mentioned *aficionados* who are not yet convinced about this matter). I also think that the development of algebra during the past 60 years will also help to make that case. In that regard, following the next paragraph, please see the subsequent seven paragraphs. There, you will find what I consider to be the most convincing reasons why domains should be unital. Those seven paragraphs give what was, in my experience as a student and a young professional, the beginning of a series of critical observations about modules. The material in the initial five of those seven paragraphs comes from an exercise (that I recall working nearly 60 years ago) from van der Waerden's classic textbook "Modern Algebra."

First, recall that the comments at the beginning of this section embedded any rng R as a subrng of some (unital) ring $\mathcal{R}$. However, if R happened to be a domain, then that construction could not be guaranteed to produce a ring $\mathcal{R}$ which is a domain. Indeed, if the characteristic of R is some prime number p, then that ring $\mathcal{R}$ is definitely *not* a domain, the point being that if r is any nonzero element of R, then $(0,r)\cdot(p,0) = (0,pr) = (0,0) = 0$ in $\mathcal{R}$. However, we show next that a more suitable embedding is available. For clarity, let us change notation and begin with a domain D (in the sense of [12]) which is definitely not unital. To avoid trivialities, one supposes that $D \neq \{0\}$, since the/a zero ring cannot be a unital subring of any unital domain. According to [12], D has a quotient field, say K. (More generally, I learned from a seminar talk by Kaplansky at UCLA in the spring quarter of 1970 that special cases of what we now call rings of fractions R_S were anticipated (long ago, before I was born) by workers such as Grell, with the role of 1 in R_S being played by the fraction s/s for any element $s \in S$. For more about this, see [12] and [18, pages 138-139].) I will next show that D can be embedded as a subrng of some domain $\mathcal{D}$ such that $\mathcal{D}$ is a unital domain and $\mathcal{D}$ also has K as its quotient field. (As [12] emphasizes the importance of such an "overring" extension in multiplicative ideal theory, I find this result, whose proof will follow next, to be especially persuasive.) Observe that K is a ring, with multiplicative identity element $1 = s/s$ for any nonzero element s of D. Take $\mathcal{D}$ to be the subring of K that is generated by D and 1. (In other words, take $\mathcal{D}$ to be the intersection of all the subrings of K which contain D and, necessarily, 1.) Then $\mathcal{D}$ is a (unital) subring of K (since $\mathcal{D}$ is a subrng of K such that $1 \in \mathcal{D}$), so $\mathcal{D}$ is a "domain with 1". Of course, we also have that D is a subrng of $\mathcal{D}$ and that K is a quotient field of $\mathcal{D}$.

The benefits of changing from predominantly ideal-theoretic reasoning to module-theoretic reasoning in commutative ring theory were widely recognized and took hold during the late 1950s and 1960s, producing many useful generalizations and new methods. Prior to that, in part because of the embedding result discussed in the first paragraph of this section, there was natural interest in deciding whether a "module" over a (unital) ring should, by definition, be required to be unital. Many mathematicians were convinced that this question should be answered in the affirmative (and I concur with them) because of the following result from van der Waerden's textbook. Let R be a not necessarily commutative (but unital) ring and let M be a left module over R. Then M can be uniquely expressed as an internal direct sum of (not necessarily unital left) R-modules, $M = M_1 \oplus M_2$, where M_1 *is* a unital (left) R-module and the action of R on M_2 is like the action of a zero ring on M_2 (in the sense that $r \cdot m = 0$ for all $r \in R$ and all $m \in M_2$).

Proof of uniqueness: Suppose that $M = M_1 \oplus M_2 = N_1 \oplus N_2$, where M_1 and N_1 are each unital (left) R-modules and R acts as a zero ring on both M_2 and N_2. We will prove that $M_1 = N_1$ and $M_2 = N_2$. Suppose first that $u \in M_1$. By hypothesis, $u = v + w$ for some uniquely determined $v \in N_1$ and $w \in N_2$. Then $u - v = w$ satisfies

$$u - v = 1 \cdot u - 1 \cdot v = 1 \cdot (u - v) = 1 \cdot w = 0 \cdot w = 0,$$

whence $u = v$. Hence $M_1 \subseteq N_1$. Similarly, $N_1 \subseteq M_1$. Thus $M_1 = N_1$.

Suppose next that $x \in M_2$. Then $x = y + z$ for some uniquely determined $y \in N_1$ and $z \in N_2$. Then $x - z = y$ satisfies

$$x - z = y = 1 \cdot y = 1 \cdot (x - z) = 1 \cdot x - 1 \cdot z = 0 \cdot x - 0 \cdot z = 0 - 0 = 0,$$

whence $x = z$. Hence $M_2 \subseteq N_2$. Similarly, $N_2 \subseteq M_2$. Thus $M_2 = N_2$. This completes the proof of the uniqueness assertion.

Proof of existence: Let $M_1 := \{x \in M \mid 1 \cdot x = x\}$. It is straightforward to check that M_1 contains 0 and is closed under scalar multiplication from R, sums and differences, and so M_1 is a not necessarily unital R-submodule of M. But M_1 is then also clearly a unital R-module. Next, let $M_2 := \{y \in M \mid 1 \cdot y = 0\}$. It is straightforward to check that M_2 contains 0 and is closed under scalar multiplication from R, sums and differences, and so M_2 is a not necessarily unital R-submodule of M. In fact, R acts as a zero ring on M_2 since, if $r \in R$ and $y \in M_2$, then $r \cdot y = (r \cdot 1) \cdot y = r \cdot (1 \cdot y) = r \cdot 0 = 0$. It remains only to prove that M is the internal direct sum of M_1 and M_2, that is, that $M_1 + M_2 = M$ and $M_1 \cap M_2 = 0$.

Let $u \in M$. Put $v := 1 \cdot u$ and $w := u - v$. Observe that $1 \cdot v = 1 \cdot (1 \cdot u) = (1 \cdot 1) \cdot u = 1 \cdot u = v$, whence $v \in M_1$; and, since we have just noted that $1 \cdot u = v = 1 \cdot v$, we have $1 \cdot w = 1 \cdot u - 1 \cdot v = v - v = 0$, whence $1 \cdot w = 0$, whence $w \in M_2$. Hence $u = v + w \in M_1 + M_2$, and so $M \subseteq M_1 + M_2$. The reverse inclusion is obvious, and so $M_1 + M_2 = M$. Finally, we need only show that if $z \in M_1 \cap M_2$, then $z = 0$. This, in turn, holds since $0 = 1 \cdot z$ (as $z \in M_2$) and $1 \cdot z = z$ (as $z \in M_1$). This completes the proof of the existence assertion. This completes the proof.

I would suggest that the main point to be gleaned from the result in the past five paragraphs is this. Because of the nature of the direct summands in the direct sum decomposition $M = M_1 \oplus M_2$, that result has reduced the study of non necessarily unital modules to the following two studies: the study of unital modules and the study of abelian groups (because a not necessarily unital module on which the ambient ring acts as a zero ring is nothing more than an abelian group). Hence, from the point of view of a ring-theorist, "modules" should be unital, as *other* considerations involving "not necessarily unital modules" have been reduced to (abelian) group theory. If a reader believes that my conclusion is outlandish, I can assure you that it is torn from the pages of history. Specifically (yes, here comes another anecdote): each academic year during the late 1960s and early 1970s, UCLA's mathematics department hosted promising postdocs, some folks on sabbatical, some mid-career specialists and senior leaders in a particular field of mathematics (the field varied annually). The field in 1969-70 was "Algebra", and I was lucky enough to be invited to participate as a Visiting Professor for the entire year. Many of the visitors were present for only three weeks, during which such visitors were obliged to give three lectures per week. One of the year-long visitors, S. A. Amitsur, gave three lectures a week for the entire academic year. More than half of those lectures were devoted to a theorem that he had only recently proved. The statement of the theorem could be given in many formulations, some of which involved noncommutative ring theory (and were thus of interest to many of those present for the "Algebra year") and one of which involved classical geometries (and hence was of interest to me, largely because of my masters studies in 1964-65 in Canada). At the end of his last lecture, Amitsur declared that, from the point of view of a ring theorist, he had just completed the solution of the overall problem that his lectures had been devoted to. There was a stunned silence in the crowd, as none of us in attendance could "connect the dots". Amitsur sensed our confusion (perhaps he had anticipated it) and then, with a twinkle in his eye, he added a fuller explanation. His analysis had reduced the overall problem at hand to a problem in group theory and so, he concluded, our interest in it, as ring-theorists, was now at an end. One by one, the audience members grinned as the wisdom of Amitsur's comment sank into their understanding, and we rose in applause of Amitsur's great accomplishment. While the preceding five paragraphs concern much, much lower-level mathematics, I suggest that they have made a similar point, hopefully as convincingly as Amitsur did in 1970.

The late 1950s and 1960s witnessed what has been called an "invasion" (I would prefer the term "infusion") of homological algebra into many areas of algebra. This use of homological and categorical predilections has continued and, in my opinion, has enriched much of algebra and its applications. A principal effect has been that there is now widespread agreement that ring homomorphisms and algebra homomorphisms should be unital. Of course, one would argue, algebras should be unital since rings should be unital and, after all, an extension involving commutative rings is an example of an algebra, is it not? More generally, given a commutative ring R and an R-algebra S (for a commutative ring R, this means that there is a ring homomorphsim f from R to the center of S), it has long been traditional to view S as a (left) R-module via $r \cdot s := f(r)s$ for all $r \in R$ and all $s \in S$. By taking $s := 1$, we see that the *only* way for this module to be unital (and we have been arguing that modules *should* be unital) is for f to be unital. Once one agrees that algebra homomorphisms should be unital, one must agree that ring homomorphisms should be unital (the point being that every ring is a $\mathbb{Z}$-algebra).

I hope that this section has given the reader some food for thought. When it comes to a discussion of values, one cannot hope to *prove* that one's views are "correct" and that others' views are "wrong." I can only hope that this section will be of help to anyone who is hesitating as to whether their rings (or their modules or their homomorphisms) should be unital. I will have accomplished my goal for this section if such readers understand better what they may expect to gain or lose as a result of any particular decision they may make about such matters.

5 Appendix III: Some professional preoccupations with beginners' angst

Some of the questions that were raised in Appendix I indicate that many beginning students of category theory and/or algebraic geometry express concerns about the use of the definite article "the" instead of the indefinite articles "a" or "an" in describing a mathematical object that is only well defined up to isomorphism. (Such angst is often manifested in regard to constructions such as A_S or $\varinjlim_{i \in \mathbf{I}} A_i$, and it is only compounded by the use of notation such as $\varinjlim_{P \in X_a} A_a$, which contains at least two such stimuli for concern.) As a beginning graduate student and later in doing my doctoral research, such worries arose naturally in the course of my reading and my research. For instance, the n^{th} piece of Amitsur's cochain complex (cf. [3]) is obtained by applying the units functor U (also known as G_m) to the tensor product, over a given field K, of $n+1$ copies of a field extension L of K. It is natural to ask what it means to apply a functor to something that is only defined up to isomorphism, and so I had some concern about the well-definedness of Amitsur's cochains. That concern compounded when I needed to address the (co)homology groups inferred from Amitsur's cochain complex, since *the* n^{th} such group was defined as *the* quotient group of *the* group of n^{th} cocycles modulo *the* group of n^{th} coboundaries. It is natural to ask what it means to be *the* factor group G/N when a group G and its normal subgroup N are each only defined up to isomorphism. Such concerns intensified during the first week of my doctoral research, as part of my assignment for that week was to read [7] where, *inter alia*, Amitsur's field extension $K \subset L$ was generalized to any (perhaps one should add "faithful") commutative algebra (over a commutative ring) and the units functor U was generalized to any abelian group-valued functor on a suitable category of algebras. Of course, the relevant cohomology groups were generalized. For an R-algebra S and a functor F, *the* associated n^{th} cohomology group was denoted by $\mathrm{H}^n(S/R, F)$. That was quite a first week of work, as my assignment also included reading a book about profinite groups. (That actually was not as difficult, even though it mixed algebra with topology, because the relevant inverse limit defining a profinte completion was truly *the* inverse limit of some unambiguous things indexed by an unambiguous directed set.) My unease was triply compounded, even that first week, because I knew that my area of doctoral research was not going to be Amitsur cohomology – it was going to be Cech cohomology and "*the*" n^{th} Cech cohomology group of a given object R and a given functor/presheaf F (in something like a Grothendieck topology $\mathcal{T}$ – yes, I also had to quickly absorb M. Artin's 1962 Harvard notes on Grothendieck topologies) is *the* direct limit of *the* corresponding Amitsur cohomology groups $\mathrm{H}^n(S/R, F)$ as S ranges over some appropriate directed set of objects drawn from $\mathcal{T}$. I quickly realized that my advisor should not be bothered with my triply-provoked concerns, but I resolved to identify the secret by which mathematicians had decided that some super version of the Axiom of Choice could be used to turn all of those occurrences of what should perhaps have been "*a*" or *an*" into occurrences of "*the*".

The semester before beginning my doctoral research, I took a very stimulating course on homological algebra. It was taught by Professor Len Silver and its official textbook was the classic work by Cartan and Eilenberg. As that work was already 10 years old by then, I realized that it would be advisable for me to try to understand many of the ideas in Professor Silver's class in a more general categorical setting. Fortunately, one of the sources that I chose to read in order to learn more about category theory during my "spare time" (what graduate student ever has any spare time?) was Grothendieck's classic paper [14], which was then widely known as "Tohoku". Fortunately, in reading (and re-reading) [14], I came across a passage that stuck in my memory. It is on page 133 of [14] and it is quoted in the next paragraph. By remembering that passage, I was able (a few months later, when I began my doctoral research) to unlock the "secret" that I had resolved to identify. It turns out that the "super Axiom of Choice" that I supposed must lie at the crux of the secret has to do with a well ordered set-theoretic universe. The availability of that universe is due (depending on one's point of view) to one or both of the following: Hilbert's desire to have the benefits of a rather strong Axiom of Choice, without explicitly committing himself to such an axiom, but instead introducing (c. 1923) certain operators, dubbed τ and ϵ, which had certain desirable properties; and Gödel's construction (barely 10 years later) of the model V for ZFC set theory which featured a well -ordered universe. In the next two paragraphs, I will say a little more about the first of these matters, having to do with Grothendieck's use of the Hilbert symbol τ. The final three paragraphs will discuss, *inter alia*, well-ordered universes.

In [14, page 133], Grothendieck addressed and dismissed some concerns similar to the ones that were mentioned in the first sentence of this appendix. He focused on the well-definedness of direct limits in the following passage (the rather literal translation is mine, but the usage of italics is from the original): "In particular, two direct limits of the same directed system are canonically isomorphic (in an evident sense), also it is natural to choose, for each directed system that admits a direct limit, one such direct limit (for example by means of Hilbert's symbol τ), which we will then denote by $\varinjlim \mathbf{A}$ or $\varinjlim_{i\in\mathbf{I}} A_i$ and which we will call *the* direct limit of the given directed system. If $\mathbf{I}$ and $\mathbf{C}$ are such that $\varinjlim \mathbf{A}$ exists for *every* directed system $\mathbf{A}$ indexed by $\mathbf{I}$ with values in $\mathbf{C}$, it follows from the above that $\varinjlim \mathbf{A}$ is a *covariant functor* defined on the category of directed systems indexed by $\mathbf{I}$ in $\mathbf{C}$, with values in $\mathbf{C}$." I can only suppose that in referring to "Hilbert's symbol τ", Grothendieck was assuming familiarity with an earlier (French language) edition of the appropriate chapter of [6].

My online searches in April 2023 indicated that this year (2023) marks the centennial of Hilbert's introduction of the operator τ. In this regard, I would like to mention some recent work of M. Abrusci and his collaborators having to do with some philosophical/mathematical questions concerning quantification and proof. First, one should acknowledge that there seems to be a widespread impression online to the effect that "Hilbert's symbol τ" had really originally been Hilbert's symbol "ϵ" and that various workers had decided to change the notation "ϵ" to "τ" some time before the original French edition of the relevant chapter of [6], presumably in order to avoid confusion between "ϵ" and the set-theoretic symbol "$\in$". However, this widespread belief seems to have been refuted by Abrusci in [1], as can be seen from the following beginning of the author's (that is, Abrusci's) summary of that work: "In section 1, I expose in an informal way the rules – and the logical rules – on the proofs of the universal statements and existential statements, and the rules – and the logical rules – on the deductions from these statements. In section 2, I show how Hilbert's operators τ and ϵ allow a representation of the universal statements and existential statements which is strictly related to the logical rules on the proofs of these statements and to the logical rules on the deductions from these statements, so that we may say that Hilbert in the introduction of the operators τ and ϵ aimed to propose a kind of proof-theoretical representation of the universal statements and existential statements." In joint work [2] with Pasquali and Retoré two years earlier, Abrusci makes clear that the set-theoretic foundational concerns of the late 19^{th} century which mathematicians typically associate with people such as Cantor and Frege (concerns which were only heightened by Hilbert's formalist pronouncement in Königsberg in 1930 that "Wir müssen wiesen. Wir werden wissen" – a belief that was shattered by Gödel's incompleteness results shortly afterward) are shared and are still being examined further to this day in some serious research (however remote such research may seem to be from our daily activities as mathematicians). A sense of the flavor and scope of [2] can be gotten from its Math. Review by B. H. Mayoh: "Quantifiers are ubiquitous in natural language. This paper presents many approaches to capturing the complexity of natural language quantification and suggests a new proof-theoretic approach. First, the authors discuss the classical universal and existential quantifiers and why G. F. L. Frege rejected the appealing idea of domain restriction. Next they present individual concepts, second-order logic and various Hilbert operators. Finally, they present a section on generalized quantifiers. Many problems remain." If there are any readers who wish to learn more about some serious, current, professional studies related to the τ and ϵ operators, I would encourage them to look into the extensive literature on what is nowadays called the "epsilontic calculus?.

In my experience, a working algebraist can occasionally benefit by attention to foundational matters. Consider, for example, the following result in category theory: a functor is a categorical equivalence if (and only if) it is fully faithful and essentially surjective. As a doctoral student, I first came across this result when I read its use by Bass in [5, Chapter II, 1.2] for some work on algebraic K-theory. Although Bass did not mention any foundational issues that may arise when using that categorical result, the only proof that I know of that result requires that some well ordering be applicable to the domain category of the given functor (certainly a well ordering of the class of objects of that category, perhaps also something like – or more than – the well ordering of each set of morphisms with a given domain and a given codomain in that category). Thanks to a famous result of Gödel [13], there *is* a model satisfying the ZFC (Zermelo-Frankel and the Axiom of Choice) foundations whose universe is well ordered.

Some mathematicians have occasionally used the above characterization of a categorical equivalence to conclude that every category is equivalent to a skeletal category, that is, to a category in which any two isomorphic objects are equal. While this would be acceptable (assuming ZFC) for a *small* category (that is, a category whose class of objects is a *set*), the famous paradoxes of intuitive set theory have led several mathematicians to conclude that many important categories are not small. In reading authors such as Grothendieck or Mac Lane (see, especially, [17, pages 23-24 and 30]), I have often had the impression that they preferred the meaning of "set" to be placed on a "sliding scale", that is, to be adjusted in accordance with the data for the problem at hand. It has been said that although most mathematicians profess to be formalists in their official pronouncements on foundational matters, we tend to think, create and act like Platonists, as though the objects of our professional attention are "ideal" things, in the spirit of "The Republic of Plato." Is there a better way to guarantee access to such ideal things than to have a well-ordered universe?

The fact that having a well ordered universe is consistent with ZFC allowed me to access and use what I called "chosen fields" to construct a functor in [8, Definition 3.8, page 24] which had several cohomologically useful applications (cf. [8, Chapter I, Theorems 3.10, 3.13 and 5.9]). The "chosen fields" were also instrumental in my proof of a very useful result [9, Theorem 2.2] stating that for any field k, in the étale toplogy for $\mathrm{Spec}(k)$, there is a left adjoint functor sending presheaves to what may be called "additive presheaves" in a way that is analogous to the "sheafification" functor that sends presheaves to sheaves (in a more general context, of course). My research has perhaps had only two other noteworthy interactions with mathematical logic: in [11, Proposition 2.5 (a)], A. Hetzel and I worked with countable models to prove the "lifting" result that a ring homomorphism is a chain morphism if (and only if) it is an n-chain morphism for every positive integer n; and in [10], R. C. Heitmann and I showed that the answer to a certain question depends on which model of ZFC is being used. That question asked to determine those infinite cardinal numbers $\aleph_\alpha$ for which there exists a field extension $K \subset L$ such that $\aleph_\alpha$ is the supremum of the set of cardinalities that arise as lengths of chains of intermediate fields contained between K and L. Regardless of whether the reader has found my anecdotes to be interesting or merely self-indulgent, I should close by pointing out that there have been several (I would add "other") interesting questions in algebra whose answers depend on the model of ZFC that is being used. To be brief, let me mention just two of them (in chronological order). In [20], B. L. Osofsky proved that the global dimension of a countable direct product of fields is $k+1$ if and only if $2^{\aleph_0} = \aleph_k$. In [21], S. Shelah proved that the Whitehead Problem is undecidable; that is, he proved that there are two axioms, each of which is consistent with ZFC, that give different answers to the question which asks whether an abelian group A such that $\mathrm{Ext}^1_{\mathbb{Z}}(A, \mathbb{Z}) = 0$ must be a free abelian group.

References

[1] V. M. Abrusci, Hilbert's τ and ϵ in proof theory: a proof-theoretical representation of universal and existential statements. *From arithmetic to metaphysics*, 1?21, Philos. Anal., 73, De Gruyter, Berlin, 2018.

[2] M. Abrusci, F. Pasquali and C. Retoré, Quantification in ordinary language and proof theory, Philos. Sci. (Paris) 20 (1) (2016), 185–205.

[3] S. A. Amitsur, Simple algebras and cohomology groups of arbitrary fields, Trans. Amer. Math. Soc. 90 (1959), 73–112.

[4] M. F. Atiyah and I. G. Macdonald, Introduction to Commutative Algebra, Addison-Wesley, Reading, MA, 1969.

[5] H. Bass, Lectures on Topics in Algebraic K-theory, Tata Institute of Fundamental Research, Bombay, 1967.

[6] N. Bourbaki, Elements of mathematics – Theory of sets, English translation of Théorie des ensembles (Hermann, Publishers in Arts and Science, Paris), Addison-Wesley, Reading (MA)-London-Don Mills (Ont.), 1968.

[7] S. U. Chase, D. K. Harrison and A. Rosenberg, Galois theory and Galois cohomology of commutative rings, Mem. Amer. Math. Soc., Volume 52 (1965).

[8] D. E. Dobbs, Cech cohomological dimensions for commutative rings, Lecture Notes in Math., Volume 147, Springer-Verlag, Berlin-Heidelberg-New York, 1970.

[9] D. E. Dobbs, Amitsur cohomology in additive functors, Can. Math. Bull. 16 (1973), 417–426.

[10] D. E. Dobbs and R. C. Heitmann, Realizing infinite cardinal numbers via maximal chains of intermediate fields, Rocky Mountain J. Math. 44 (5) (2014), 1471–1503.

[11] D. E. Dobbs and A. J. Hetzel, On chain morphisms of commutative rings, Rend. Circ. Mat. Palermo (2), 53 (1) (2004), 71?-84.

[12] R. Gilmer, Multiplicative Ideal Theory, Dekker, New York, 1972.

[13] K. Gödel, The consistency of the axiom of choice and of the generalized-continuum hypothesis with the axioms of set theory, Princeton University Press, Princeton, 1940.

[14] A. Grothendieck, Sur quelques points d'algèbre homologique, Tohoku Math. J. 9 (2) (1957), 119–221.

[15] A. Grothendieck and J. A. Dieudonné, Éléments de Géométrie Algébrique, I, Springer-Verlag, 1971.

[16] I. Kaplansky, Commutative Rings, rev. ed., Univ. of Chicago Press, Chicago, 1974.

[17] S. Mac Lane, Categories for the Working Mathematician, second edition, Graduate Texts in Math., Volume 5, Springer-Verlag, New York, 1998.

[18] N. H. McCoy, The Theory of Rings, Macmillan, New York, 1964.

[19] D. Mumford, Introduction to Algebraic Geometry (preliminary version of first 3 chapters), undated notes from Harvard University [my copy was obtained in January 1967].

[20] B. L. Osofsky, Homological dimension and cardinality, Trans. Amer. Math. Soc. 151 (1970), 641–649.

[21] S. Shelah, Infinite abelian groups – Whitehead problem and some constructions, Israel J. Math. 18 (1974), 243–256.