Commutative Algebra

Preliminary version—comments welcome!

Ulrich Thiel

Email address: thiel@mathematik.uni-kl.de

Web address: https://ulthiel.com/math

University of Kaiserslautern, Department of Mathematics, 67653 Kaiserslautern, Germany

v0.2 (Oct 2021)

Contents

Introduction

These are my notes for a first course on commutative algebra. The material is fairly standard but I try to present a personal blend. One of my goals is to emphasize the geometric side as well. I assume you know basics about groups, rings, and vector spaces—but not more.

When I was a student, this course was my favorite—it really influenced me. I'll do my best to give you the same experience. You can help improving these notes by asking questions, pointing out mistakes, making suggestions, etc. Discussion is an integral part of mathematics, so please do it.

What is this about?

Well, surprise, this course is about commutative rings! Why is this interesting? *One* major motivation for studying commutative rings is that there's a dictionary between commutative rings and *algebraic geometry*, the latter being the study of solutions to systems of polynomial equations. This dictionary was developed in full generality in the 1950s by A. Grothendieck[1]. We will learn a tiny bit about this—you will see more in an algebraic geometry course—but the basic correspondence is like this:

Geometry	$\leftrightarrow$	**Algebra**
R^n	$\leftrightarrow$	$R[X_1, \ldots, X_n]$
common zero set of polynomials $f_1, \ldots, f_m$ in n variables over R	$\leftrightarrow$	$R[X_1, \ldots, X_n]/(f_1, \ldots, f_m)$

Here, R is a commutative ring, $n \in \mathbb{N}$, and R^n is the set of n-tuples $\boldsymbol{x} \coloneqq (x_1, \ldots, x_n)$ of elements of R. You probably know the polynomial ring $R[X]$ in one variable. More generally, we can consider polynomials over R in n variables, and these form a ring as well which is denoted by $R[X_1, \ldots, X_n]$. Given such a polynomial f and a point $\boldsymbol{x} \in R^n$ we can replace all the X_i in f by the value x_i for all i and get in this way an element $f(\boldsymbol{x}) \in R$, i.e. we simply evaluate f in $\boldsymbol{x}$. The common zero set of polynomials $f_1, \ldots, f_m$ is the subset of all points $\boldsymbol{x} \in R^n$ such that $f_i(\boldsymbol{x}) = 0$ for all $1 \leq i \leq m$. The dictionary tells us that to this geometric object there corresponds the quotient $R[X_1, \ldots, X_n]/(f_1, \ldots, f_n)$ of the polynomial ring by the ideal generated by the f_i.

It is a very important feature of the dictionary that the ring R can be arbitrary. For example, we can do "geometry" over $\mathbb{Z}$ or over the field $\mathbb{F}_2 = \mathbb{Z}/2\mathbb{Z}$ with two elements even though this is quite hard to imagine. But algebra doesn't care!

[1] If you've never heard this name before, go check it out on Wikipedia!

An easy fact that we will discover soon is that *any* commutative ring can be written as a quotient of a polynomial ring (in possibly infinitely many variables and the ideal we need to factor out may have infinitely many generators but this shouldn't bother us now). Hence, *any* commutative ring can be viewed geometrically! Many intuitive geometric properties of the zero set (e.g. its "dimension") translate to precise algebraic properties of the associated ring. E.g. there is the notion of *dimension* of a commutative ring. What do you think will be the dimension of the polynomial ring $\mathbb{R}[X_1, \ldots, X_n]$? If you said "$n$", congratulations! You should also be able to guess the dimension of $\mathbb{R}[X_1, X_2]/(X_1X_2)$ (Hint: draw the associated zero set). And now what about the dimension of $\mathbb{Z}$ or $\mathbb{Z}[X]$? You will learn!

I want to note that the real upshot of algebraic geometry is that you can *glue* zero sets of polynomials and what you obtain can in general not be described *globally* by a zero set—and thus by a commutative ring—anymore. So, commutative algebra is algebraic geometry *locally*. We will not consider such gluing things here because this is done in an algebraic geometry course.

It will be an important theme in this course that you can study (commutative) rings not just directly by studying the rings but also indirectly by studying their *modules*. A module is basically the same thing as a vector space over a field—only over a general ring and here the behavior becomes quite different, e.g. whereas any vector space has a basis, this is no longer true for modules.

This course is useful everywhere in algebra: commutative algebra (surprise!), algebraic geometry, algebraic number theory, representation theory, computer algebra, and life in general. So, let's go!

References

There is nothing original in these notes, all concepts and proofs are well-established—I just tried to create a personal blend. Standard textbooks are the books by Atiyah–Macdonald [2] (very concise, still contains almost everything but sometimes not so helpful for developing intuition) and Eisenbud [5] (very detailed with a lot of comments on geometric connections that help to build intuition, but also very long). I also recommend the lecture notes by Clark [4]. Once you know a fair deal of commutative algebra, useful references are the *Stacks Project* [12] and the book by Bourbaki [3]. Another book that is not so well-known but that I can recommend is the book [11] by Scheja and Storch (in German though but contains many exercises). I should note that it is amazing how much of commutative algebra is accessible via algorithms. In this course, I will not go into this direction and stay on the theoretical side (we will still see many explicit examples). But you will learn all the basics to go into this direction as well.

Conventions

The natural numbers $\mathbb{N}$ always include zero and we use $\mathbb{N}_{>0}$ to exclude it.

The protagonists

Whenever you dive into a new field of mathematics (or any other subject), it is very helpful to know the names of (some of) the protagonists, when they lived, and what they did. The following is a non-exhaustive list, sorted chronologically by year of birth.

Richard DEDEKIND (1831–1916): Dedekind domains (Definition 9.1.1).

David HILBERT (1862–1943): Hilbert's basis theorem (Theorem 7.3.5), Hilbert's Nullstellensatz (Corollary 6.1.9).

Emmy NOETHER (1882–1935): Noetherian rings and modules (Definition 7.3.1), Noether normalization (Theorem 8.3.9).

Emil ARTIN (1898–1962): Artinian rings and modules (Section 7.4).

Wolfgang KRULL (1899–1971): Krull dimension (Definition 8.2.1), Krull's principal ideal theorem (Theorem 8.5.1).

Oscar ZARISKI (1899–1986): Zariski topology (Section 2.5).

Nathan JACOBSON (1910–1999): Jacobson radical (Definition 2.6.7).

Alexander GROTHENDIECK (1928–2014): General dictionary between commutative rings and algebraic geometry (Section 2.4).

Acknowledgments

I would like to thank Gunter Malle and Johannes Schmitt for providing extensive comments and corrections on a preliminary version of these notes. Moreover, I would like to thank all of the following for providing comments and feedback: Marco Albert, Jeannine Berghänel, Cedric Brendel, Christoph Geis, Felix Wacker (from my 2020/2021 course).

CHAPTER 1

Review of basic ring theory

Before we really begin with commutative algebra it's best to review some basic ring theory. I will recall many concepts you already know (e.g., rings, morphisms, ideals, unique factorization) but also some you may not yet know (e.g., categories, universal properties, ideals under morphisms, algebras, polynomial rings in arbitrarily many variables). Nothing here is difficult but you should take your time to study everything carefully so that you're well prepared for the actual course.

1.1. Rings and ring morphisms

Recall the formal definition of a ring:

Definition 1.1.1. A **ring** is a set A with two operations:

$$+\colon A \times A \to A \quad (\textbf{addition}) \tag{1.1}$$

$$\cdot\colon A \times A \to A \quad (\textbf{multiplication}) \tag{1.2}$$

such that the following holds:

(1) $(A,+)$ is an abelian group,
(2) $(A,\cdot)$ is a **monoid**, i.e. **associative** with a **unit**,
(3) $\cdot$ is **distributive** with respect to addition, i.e.

$$a \cdot (a' + a'') = a \cdot a' + a \cdot a'' \tag{1.3}$$

$$(a + a') \cdot a'' = a \cdot a'' + a' \cdot a'' \tag{1.4}$$

for all $a, a', a'' \in A$.

One usually uses the shorthand notation aa' for $a \cdot a'$. The neutral element for addition is denoted by 0 and the neutral element for multiplication is denoted by 1. Note that both are uniquely determined by the property of a neutral element.

Remark 1.1.2. The reason you often see the letter A for a ring is that "ring" is "annulus" in Latin and "anneau" in French; also, it stands for an "algebra" which is a concept extending that of a ring, see Section 1.4.

Remark 1.1.3. Depending on the literature (especially before the 1960s) a ring may not be required to have a multiplicative unit—our rings are then called "unital rings" or "rings with identity". This is the case for example in the famous book on commutative algebra by Zariski and Samuel [14] from 1958. We will, however, follow the modern conventions and always assume that we have a multiplicative unit.

Example 1.1.4. The prime example of a ring is the ring of integers $\mathbb{Z}$.

Example 1.1.5. $A = \{0\}$ is a ring, called the **zero ring**. Convince yourself that this is the only ring in which $0 = 1$.

Example 1.1.6. Let X be a set and let R be a ring. Then the set $\mathrm{Maps}(X, R)$ of all maps $X \to R$ is a ring with respect to pointwise addition and multiplication, i.e.

$$(f + g)(x) \coloneqq f(x) + g(x) \tag{1.5}$$

$$(fg)(x) \coloneqq f(x)g(x) \tag{1.6}$$

for $f, g \in \mathrm{Maps}(X, R)$ and $x \in X$. The unit element in this ring is the constant function $1(x) = 1 \in R$.

Example 1.1.7. Let R be a ring and $n \in \mathbb{N}$. Then the set $\mathrm{Mat}_n(R)$ of $(n \times n)$-matrices with entries in R is a ring with the usual addition and matrix multiplication.

Definition 1.1.8. A ring A is **commutative** if $aa' = a'a$ for all $a, a' \in A$, i.e. $(A, \cdot)$ is a commutative monoid.

Example 1.1.9. The ring of integers $\mathbb{Z}$ and the zero ring are commutative. The ring $\mathrm{Maps}(X, R)$ from Example 1.1.6 is commutative if and only if R is commutative. The matrix ring $\mathrm{Mat}_n(R)$ is *never* commutative unless R is the zero ring or R is commutative and $n = 1$.

Assumption 1.1.10. Throughout this course, *all* rings are *commutative.* You should keep in mind though that *some* concepts we will discuss here can be generalized to non-commutative rings—in difficulty this ranges from straightforward to active research. I will make a few comments in this direction but I don't want to blow things up too much and will therefore assume commutativity everywhere. After all, it's a course on *commutative* algebra. Still, if you are interested in non-commutative algebra as well (which is used heavily for example in representation theory), then I recommend the textbook [9] for a start.

The difference between addition and multiplication in a ring is that multiplication does not necessarily have an inverse for each (non-zero) element.

Definition 1.1.11. A **unit** in a ring A is an invertible element of the monoid $(A, \cdot)$, i.e. there is $a^{-1} \in A$ such that $aa^{-1} = 1$.

The element a^{-1} is unique with this property. The set

$$A^\times \coloneqq \{a \in A \mid a \text{ is a unit}\} \tag{1.7}$$

is a sub*group* of the monoid $(A, \cdot)$, called the **unit group** of A.

Definition 1.1.12. A ring A is a **field** if $A^\times = A \setminus \{0\}$, i.e. A is non-zero and any non-zero element is a unit.

Example 1.1.13. The prime example of a field is the field of rational numbers $\mathbb{Q}$. Other examples are the field of real numbers $\mathbb{R}$ and of complex numbers $\mathbb{C}$.

Example 1.1.14. The units in $\mathbb{Z}$ are $\{\pm 1\}$, which form a cyclic group of order 2.

Remark 1.1.15. For non-commutative rings you want a unit to have both a left and a right inverse. It's easy to see that if this exists, left and right inverse are unique and coincide.

Ring morphisms are the structure preserving maps between rings. Here's the formal definition.

Definition 1.1.16. A **morphism** between rings A and B is a map $f\colon A \to B$ such that:

(1) $f\colon (A,+) \to (B,+)$ is a group morphism, i.e.
$$f(a + a') = f(a) + f(a') \tag{1.8}$$
for all $a, a' \in A$.

(2) $f\colon (A,\cdot) \to (B,\cdot)$ is a monoid morphism, i.e.
$$f(aa') = f(a)f(a') \quad \textbf{and} \quad f(1_A) = 1_B \tag{1.9}$$
for all $a, a' \in A$.

Note that (1.8) implies that
$$f(-a) = -f(a) \ , \quad \text{in particular } f(0) = 0 \ . \tag{1.10}$$
In contrast, since multiplication just forms a monoid (i.e. we don't necessarily have inverses), the equality $f(1) = 1$ is not automatic and we need to force it!

Example 1.1.17. For any ring A there is a unique ring morphism $\mathbb{Z} \to A$. Namely, a morphism must map $n \in \mathbb{N}$ to
$$n \cdot 1 \coloneqq \underbrace{1 + \ldots + 1}_{n \text{ times}} \in A \tag{1.11}$$
and then we have $(-n) \cdot 1 = -(n \cdot 1)$, which completely determines the morphism. It is clear that this assignment indeed defines a morphism, hence there is a unique ring morphism $\mathbb{Z} \to A$. Note that we can now define
$$n \cdot a \coloneqq (n \cdot 1) \cdot a \in A \tag{1.12}$$
for any $n \in \mathbb{Z}$ and $a \in A$, i.e. there is a canonical operation of $\mathbb{Z}$ on A. We'll come back to this in Section 1.4.

Clearly, if $f\colon A \to B$ and $g\colon B \to C$ are ring morphisms, so is their composition $g \circ f\colon A \to C$.

Definition 1.1.18. A ring morphism $f\colon A \to B$ is an **isomorphism** if there is a ring morphism $g\colon B \to A$ such that $f \circ g = \mathrm{id}_B$ and $g \circ f = \mathrm{id}_A$.

Lemma 1.1.19. *A ring morphism $f\colon A \to B$ is an isomorphism if and only if it is bijective.*

Proof. If f is an isomorphism, it has an inverse as a ring morphism, this is also an inverse as a map of sets, hence f is bijective. Conversely, if f is bijective, it has an inverse g as a map of sets. The map g is automatically a ring morphism and so f is a ring isomorphism:
$$f(g(b) + g(b')) = f(g(b)) + f(g(b')) = b + b' \ ,$$
and applying g yields
$$g(b) + g(b') = g(b + b')$$
for all $b, b' \in B$. Similarly you show that $g(bb') = g(b)g(b')$ and $g(1) = 1$. □

Example 1.1.20. Complex conjugation $\mathbb{C} \to \mathbb{C}$ is a ring isomorphism.

Many examples of rings arise as subrings of known rings. Here's the formal definition of this concept.

Definition 1.1.21. A **subring** of a ring B is a subset A of B such that:

(1) A is closed under $+$ and $-$, i.e. $(A,+)$ is a subgroup of $(B,+)$. Note that this implies $0 \in A$.
(2) A is closed under $\cdot$ and $1 \in A$, i.e. $(A,\cdot)$ is a submonoid of $(B,\cdot)$.

If A is a subring of B, then A itself is a ring with respect to the restriction of the operations and the inclusion $A \to B$ is an injective ring morphism.

Example 1.1.22. $\mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}$ is a chain of subrings.

Example 1.1.23. If you don't know what a topological space is (I'll recall this in Section 2.5) you can skip this example. Let X be a topological space (e.g. $\mathbb{R}^n$). Then the set $C(X,\mathbb{R})$ of continuous functions $X \to \mathbb{R}$ is a subring of $\mathrm{Maps}(X,\mathbb{R})$. Let $C_c(X,\mathbb{R}) \subseteq C(X,\mathbb{R})$ be the subset of continuous functions $f\colon X \to \mathbb{R}$ with *compact support*, i.e. the closure of the set $\{x \in X \mid f(x) \neq 0\} \subseteq X$ is compact. This is a subring of $C(X,\mathbb{R})$ if and only if X is compact because $1 \in C_c(X,\mathbb{R})$ if and only if X is compact.

Example 1.1.24. The zero ring 0 is *not* a subring of a ring A unless A itself is the zero ring.

Here is another general example of how subrings arise.

Lemma 1.1.25. *If $f\colon A \to B$ is a ring morphism, then the* ***image***

$$\operatorname{Im} f \coloneqq f(A) = \{f(a) \mid a \in A\} \tag{1.13}$$

is a subring of B.

Proof. If $b, b' \in \operatorname{Im} f$, then $b = f(a)$ and $b' = f(a')$ for some $a, a' \in A$. Hence,

$$\begin{aligned} b + b' &= f(a) + f(a') = f(a+a') \in \operatorname{Im} f\ , \\ -b &= -f(a) = f(-a) \in \operatorname{Im} f\ , \\ bb' &= f(a)f(a') = f(aa') \in \operatorname{Im} f\ , \\ 1 &= f(1) \in \operatorname{Im} f\ . \end{aligned}$$

□

Often, subrings are described by a (preferably small) set of generators which makes it easier to work with them. I will now define what "generators" means.

Lemma 1.1.26. *If $\mathcal{A}$ is a set of subrings of a ring B, then the intersection $\bigcap_{A \in \mathcal{A}} A$ is a subring of B as well.*

Proof. This is straightforward. □

Corollary 1.1.27. *For any subset $\boldsymbol{x}$ of a ring B there is a unique subring of B minimal among all subrings containing $\boldsymbol{x}$. We call this the subring of B* ***generated*** *by $\boldsymbol{x}$ and denote it by $\mathbb{Z}[\boldsymbol{x}]$. More explicitly, we have:*[1]

$$\mathbb{Z}[\boldsymbol{x}] = \left\{ \sum_{\mu \in \mathbb{N}^n} r_\mu x_1^{\mu_1} \cdots x_n^{\mu_n} \;\middle|\; \begin{array}{l} n \in \mathbb{N}, x_i \in \boldsymbol{x}, r_\mu \in \mathbb{Z} \\ \text{all but finitely many } r_\mu = 0 \end{array} \right\} . \tag{1.14}$$

Proof. Let $\mathcal{A}$ be the set of subrings of B containing $\boldsymbol{x}$. Then

$$\mathbb{Z}[\boldsymbol{x}] \coloneqq \bigcap_{A \in \mathcal{A}} A \tag{1.15}$$

[1] Here, we use the multiplication of ring elements by integers introduced in Example 1.1.17

is a subring of B containing $\boldsymbol{x}$. It is clear that $\mathbb{Z}[\boldsymbol{x}]$ is minimal among all subrings containing $\boldsymbol{x}$ and that it is unique with this property. The subring $\mathbb{Z}[\boldsymbol{x}]$ must surely contain the set on the right hand side of (1.14). On the other hand, this set is easily seen to be a subring containing $\boldsymbol{x}$. Hence, we have equality in (1.14). □

In the notation $\mathbb{Z}[\boldsymbol{x}]$ there is no indication of the big ring B anymore because it is mostly clear from the context. A subset $\boldsymbol{x}$ of a ring A such that $A = \mathbb{Z}[\boldsymbol{x}]$ is called a set of (ring) **generators** of A. Such a set always exists because we can take $A = \boldsymbol{x}$. A ring is said to be **finitely generated** if it admits a finite set of generators.

REMARK 1.1.28. The reason for the $\mathbb{Z}$ in the notation $\mathbb{Z}[\boldsymbol{x}]$ will become clearer later when we discuss algebras (Section 1.4). Basically, the idea is that one can have *scalars* in a ring and wants to have the minimal subring containing a set and which is closed under taking products with scalars. For a general ring, we only have $\mathbb{Z}$ as scalars and this is the $\mathbb{Z}$ in $\mathbb{Z}[\boldsymbol{x}]$.

EXAMPLE 1.1.29. Let $\mathrm{i} \in \mathbb{C}$ be the imaginary unit. Then

$$\mathbb{Z}[\mathrm{i}] = \{x + \mathrm{i}y \mid x, y \in \mathbb{Z}\} \subseteq \mathbb{C} \,. \tag{1.16}$$

This is called the ring of **Gaussian integers**.

Exercises.

EXERCISE 1.1.30. Which other rings do you know?

EXERCISE 1.1.31. If A is a ring and $a, b \in A$ with $ab \in A^\times$, show that $a, b \in A^\times$.

EXERCISE 1.1.32. If $f \colon A \to B$ is a ring morphism, show that $f(A^\times) \subseteq B^\times$ and that f induces a group morphism $A^\times \to B^\times$.

EXERCISE 1.1.33. Determine the units in $\mathbb{Z}[\mathrm{i}]$.

EXERCISE 1.1.34. Determine the units in the ring $C(\mathbb{R}, \mathbb{R})$ of continuous functions $\mathbb{R} \to \mathbb{R}$ from Example 1.1.23.

EXERCISE 1.1.35. Determine all ring morphisms $\mathbb{R} \to \mathbb{R}$.

EXERCISE 1.1.36. Show that two subrings of $\mathbb{Q}$ are isomorphic if and only if they are equal.

1.2. Categories and universal properties

You all know what a category is even if you've never heard of this before. Whenever you have an algebraic structure (like rings) you also consider morphisms between them (like ring morphisms). It would be great to have a concept encoding this efficiently. This is what a category is doing.

DEFINITION 1.2.1. A **category** $\mathcal{C}$ consists of:

1. A collection $\mathrm{Ob}_\mathcal{C}$ of **objects**. We simply write $X \in \mathcal{C}$ instead of $X \in \mathrm{Ob}_\mathcal{C}$.
2. A collection $\mathrm{Hom}_\mathcal{C}$ of **morphisms**, each morphism having a **source** object and a **target** object. We write $f \colon X \to Y$ if f is a morphism with source X and target Y, and write $\mathrm{Hom}_\mathcal{C}(X, Y)$ for the collection of such morphisms.
3. A **composition** $\circ \colon \mathrm{Hom}_\mathcal{C}(X, Y) \times \mathrm{Hom}_\mathcal{C}(Y, Z) \to \mathrm{Hom}_\mathcal{C}(X, Z)$, $(f, g) \mapsto g \circ f$, which is associative, i.e., $h \circ (g \circ f) = (h \circ g) \circ f$ whenever the composition is defined, and has an **identity** id_X for each object, i.e. $f \circ \mathrm{id}_X = f$ and $\mathrm{id}_X \circ g = g$ whenever the composition is defined.

Example 1.2.2. A standard example is the category Set of sets: the objects are sets, the morphisms are maps of sets, and the composition is the usual composition of maps. Similarly, with the usual composition, we have the category Grp of groups with group morphisms, the category Ring of rings with ring morphisms, and the category K-Vec of K-vector spaces over a field K with linear maps.

Remark 1.2.3. The reason that I used the term *collection* in Definition 1.2.1 is that often $\mathrm{Ob}_{\mathcal{C}}$ is not a *set*, e.g. by Russel's paradox we cannot form the set of all sets. One can avoid these problems by introducing a new hierarchy into set theory called *classes* which allows, e.g., to form the class of all sets. It won't be necessary for us to dive into these set-theoretic issues. I used the loose term "collection" to not write anything wrong here and avoid the technical term "class".

You're well familiar already with the following terms:

Definition 1.2.4. A morphism $f\colon X \to Y$ in a category $\mathcal{C}$ is called **isomorphism** if there is a morphism $g\colon Y \to X$ with

$$f \circ g = \mathrm{id}_Y \quad \text{and} \quad g \circ f = \mathrm{id}_X \ . \tag{1.17}$$

An **endomorphism** of an object X is a morphism $X \to X$; we write

$$\mathrm{End}_{\mathcal{C}}(X) \coloneqq \mathrm{Hom}_{\mathcal{C}}(X,X) \ . \tag{1.18}$$

An **automorphism** of X is an endomorphism which is also an isomorphism; we write $\mathrm{Aut}_{\mathcal{C}}(X)$ for the set of automorphisms on X and note that this is a group with respect to composition.

Categories are algebraic structures as well, so when you have two categories $\mathcal{C}$ and $\mathcal{D}$ you can consider structure preserving maps between them. Here's the formal definition.

Definition 1.2.5. A (covariant) **functor** $F\colon \mathcal{C} \to \mathcal{D}$ consists of:

(1) A map $\mathrm{Ob}_{\mathcal{C}} \to \mathrm{Ob}_{\mathcal{D}}$, written $X \mapsto F(X)$.
(2) A map $\mathrm{Hom}_{\mathcal{C}}(X,Y) \to \mathrm{Hom}_{\mathcal{D}}(F(X),F(Y))$, written $f \mapsto F(f)$, for any $X,Y \in \mathcal{C}$, and these maps have to be compatible with the composition and preserve the identity, i.e. $F(f \circ g) = F(f) \circ F(g)$ whenever the composition is defined, and $F(\mathrm{id}_X) = \mathrm{id}_{F(X)}$ for all X.

Example 1.2.6. Any ring has an underlying set, and any ring morphism is a map of the underlying sets. This yields a functor $\mathsf{Ring} \to \mathsf{Set}$, the so-called **forget functor** (which forgets about the ring structure).

Example 1.2.7. For any object X in a category $\mathcal{C}$ we have the (covariant) **Hom-functor**

$$\mathrm{Hom}_{\mathcal{C}}(X,-)\colon \mathcal{C} \to \mathsf{Set} \tag{1.19}$$

which maps an object $Y \in \mathcal{C}$ to the set[2] $\mathrm{Hom}_{\mathcal{C}}(X,Y)$, and which maps a morphism $f\colon Y \to Z$ in $\mathcal{C}$ to the set map

$$\begin{array}{rccc} \mathrm{Hom}_{\mathcal{C}}(X,f)\colon & \mathrm{Hom}_{\mathcal{C}}(X,Y) & \to & \mathrm{Hom}_{\mathcal{C}}(X,Z) \\ & g & \mapsto & f \circ g \end{array} \ . \tag{1.20}$$

[2] If you are very careful you may have noticed that this functor only maps to Set when $\mathrm{Hom}_{\mathcal{C}}(X,Y)$ is actually a set and not a too large collection, see Remark 1.2.3. Again, I don't want to go into set-theoretic issues here.

In practice, it often happens that you have mappings between objects and morphisms like for a functor but morphisms are “reversed”, i.e. you have mappings $\mathrm{Hom}_{\mathcal{C}}(X,Y) \to \mathrm{Hom}_{\mathcal{D}}(F(Y),F(X))$, which are compatible with composition and preserve the identity. Note that the compatibility with composition becomes here $F(f \circ g) = F(g) \circ F(f)$. In this case we speak of a **contravariant functor**.

Example 1.2.8. For any object Y of a category $\mathcal{C}$ we have the contravariant Hom-functor

$$\mathrm{Hom}_{\mathcal{C}}(-,Y) \colon \mathcal{C} \to \mathsf{Set} \tag{1.21}$$

defined like the covariant Hom-functor but with the argument in the first variable.

Categories and functors are not essential in this course but they really help to expose the key features of some constructions and phenomena we will encounter. I encourage you to check out the Wikipedia pages on categories for more examples.[3]

Let’s consider again the category of rings. We want to construct new rings from old ones. Taking direct products is one way to do this. I want to show you that there’s a general (categorical) idea underlying this construction and this is very helpful to know because it saves a lot of time later when we consider other categories. First, recall the direct product of rings.

Lemma 1.2.9. *Let $(A_\lambda)_{\lambda \in \Lambda}$ be a non-empty family of rings. Then their* ***direct product*** *$\prod_{\lambda \in \Lambda} A_\lambda$ is a ring as well with respect to component-wise operations:*

$$(a_\lambda)_{\lambda \in \Lambda} + (a'_\lambda)_{\lambda \in \Lambda} \coloneqq (a_\lambda + a'_\lambda)_{\lambda \in \Lambda} \ , \tag{1.22}$$

$$(a_\lambda)_{\lambda \in \Lambda} \cdot (a'_\lambda)_{\lambda \in \Lambda} \coloneqq (a_\lambda \cdot a'_\lambda)_{\lambda \in \Lambda} \ , \tag{1.23}$$

$$1 \coloneqq (1_{A_\lambda})_{\lambda \in \Lambda} \ , \tag{1.24}$$

$$0 \coloneqq (0_{A_\lambda})_{\lambda \in \Lambda} \ . \tag{1.25}$$

Proof. This is straightforward. □

The **projection**

$$\mathrm{p}_\mu \colon \prod_{\lambda \in \Lambda} A_\lambda \to A_\mu \tag{1.26}$$

onto A_μ is obviously a ring morphism for any $\mu \in \Lambda$. Note that we have in particular defined for any $n \in \mathbb{N}_{>0}$ the n**-fold product**

$$A^n \coloneqq \prod_{i=1}^{n} A \tag{1.27}$$

of a ring A.

I’m sure direct products of rings are nothing new for you. But I want you to think a bit more *categorical* about this. The direct product is not just an object (a ring) but it comes along with ring morphisms p_λ, and together they satisfy a **universal property**:

[3] See `https://en.wikipedia.org/wiki/Category_(mathematics)`.

Lemma 1.2.10. *If A is a ring and if for every $\mu \in \Lambda$ we have given a ring morphism $f_\mu \colon A \to A_\mu$, then there is a unique ring morphism*

$$f \colon A \to \prod_{\lambda\in\Lambda} A_\lambda \tag{1.28}$$

such that

$$\mathrm{p}_\mu \circ f = f_\mu \tag{1.29}$$

for all $\mu \in \Lambda$, i.e. the diagram

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[dr, "f_\mu"'] & \prod_{\lambda\in\Lambda} A_\lambda \arrow[d, "\mathrm{p}_\mu"] \\ & A_\mu \end{tikzcd} \tag{1.30}$$

commutes for all $\mu \in \Lambda$.

Proof. Define f via $f(a) \coloneqq (f_\lambda(a))_{\lambda\in\Lambda}$. This is a ring morphism satisfying $\mathrm{p}_\mu \circ f = f_\mu$. It is clear that this property also forces f to be defined like that. □

The nice thing about this categorical point of view is that it really completely characterizes the direct product! For clarity, let's set $X \coloneqq \prod_{\lambda\in\Lambda} A_\lambda$ in the following lemma.

Lemma 1.2.11. *The direct product $(X, (\mathrm{p}_\lambda)_{\lambda\in\Lambda})$ of a family $(A_\lambda)_{\lambda\in\Lambda}$ of rings is unique up to unique isomorphism, i.e. if $(X', (\mathrm{p}'_\lambda)_{\lambda\in\Lambda})$ is another pair of a ring X' and ring morphisms $\mathrm{p}'_\mu \colon X' \to A_\mu$ satisfying the property as in Lemma 1.2.10, then there is a unique ring isomorphism $X \to X'$ making the diagram*

$$\begin{tikzcd} X \arrow[rr] \arrow[dr, "\mathrm{p}_\mu"'] & & X' \arrow[dl, "\mathrm{p}'_\mu"] \\ & A_\mu & \end{tikzcd} \tag{1.31}$$

commutative for all $\mu \in \Lambda$.

Proof. We have morphisms $\mathrm{p}_\mu \colon X \to A_\mu$ for all $\mu \in \Lambda$. Hence, by the universal property of $(X', (\mathrm{p}'_\lambda)_{\lambda\in\Lambda})$, there is a unique morphism $f' \colon X \to X'$ making the diagram

$$\begin{tikzcd} X \arrow[rr, "f'"] \arrow[dr, "\mathrm{p}_\mu"'] & & X' \arrow[dl, "\mathrm{p}'_\mu"] \\ & A_\mu & \end{tikzcd} \tag{1.32}$$

commutative for all $\mu \in \Lambda$. Similarly, we have morphisms $\mathrm{p}'_\mu \colon X' \to A_\mu$ and by the universal property of $(X, (p_\lambda)_{\lambda\in\Lambda})$ there is a unique morphism $f \colon X' \to X$ making the diagram

$$\begin{tikzcd} X' \arrow[rr, "f"] \arrow[dr, "\mathrm{p}'_\mu"'] & & X \arrow[dl, "\mathrm{p}_\mu"] \\ & A_\lambda & \end{tikzcd} \tag{1.33}$$

commutative for all $\mu \in \Lambda$. Putting the latter two diagrams together, we get a commutative diagram

$$\begin{tikzcd} X \arrow[rr, "f \circ f'"] \arrow[dr, "\mathrm{p}_\mu"'] & & X \arrow[dl, "\mathrm{p}_\mu"] \\ & A_\mu & \end{tikzcd} \tag{1.34}$$

Finally, by the universal property of $(X, (\mathrm{p}_\lambda)_{\lambda\in\Lambda})$ applied to the $p_\mu \colon X \to A_\mu$, there is a unique morphism $i \colon X \to X$ making the diagram

$$\begin{tikzcd} X \arrow[rr, "i"] \arrow[dr, "\mathrm{p}_\mu"'] & & X \arrow[dl, "\mathrm{p}_\mu"] \\ & A_\mu & \end{tikzcd} \tag{1.35}$$

commutative. Of course, the identity $\mathrm{id}_X \colon X \to X$ would do, so $i = \mathrm{id}_X$ by uniqueness. But then also $f \circ f' = \mathrm{id}_X$ by uniqueness. Similarly, you deduce that $f' \circ f = \mathrm{id}_{X'}$. Hence, f is an isomorphism. The uniqueness is clear. □

Here's the upshot of all this categorical stuff:

(1) You can take the universal property Lemma 1.2.10 to *define* what a **direct product** should be in an arbitrary category $\mathcal{C}$. In this way you formalized direct products of sets, groups, rings, vector spaces, etc. all at once!
(2) You can consider many different kinds of universal properties (we will see some more later).
(3) A priori it's never clear that there is a solution to a universal property problem. For example in the category of finite groups there is no direct product for an infinite set Λ. Often, one proves existence by giving an explicit construction like we did for the direct product of rings in Lemma 1.2.10.
(4) But if there is a solution, it is already unique up to unique isomorphism. The proof always goes along the lines of the proof of Lemma 1.2.11.

Exercises.

Exercise 1.2.12. An **idempotent** in a ring A is an element $e \in A$ with $e^2 = e$. Show the following:

(1) $Ae = \{ae \mid a \in A\} \subseteq A$ is a ring with the addition and multiplication from A. But it is *not* a subring unless $e = 1$.
(2) $1 - e$ is an idempotent as well.
(3) As a ring, A is isomorphic to the product $Ae \times A(1-e)$.

1.3. Ideals

Another way to produce new rings from old ones is to take quotients. For this we need the notion of ideals, which is the ring-theoretic analogue of *normal* subgroups (but note that in contrast an ideal is *not* a subring satisfying an additional property).

Definition 1.3.1. An **ideal** in a ring A is a subset $I \subseteq A$ such that:

(1) I is closed under $+$ and $-$, i.e. $(I, +)$ is a subgroup of $(A, +)$. Note that this implies $0 \in I$.
(2) I is closed under multiplication with elements from A, i.e. $ax \in I$ for all $a \in A$ and $x \in I$.

In this case one also writes $I \trianglelefteq A$.

Example 1.3.2. In any ring A the sets $\{0\}$ and A are ideals. They are called the **trivial ideals**. An ideal $I \neq A$ is called a **proper** ideal.

Example 1.3.3. The only ideals in a field K are the trivial ideals $\{0\}$ and K.

Example 1.3.4. The ideals in $\mathbb{Z}$ are the subsets $n\mathbb{Z} = \{na \mid a \in \mathbb{Z}\}$ for $n \in \mathbb{Z}$. Here's how to prove this in case you don't remember. It's clear that $n\mathbb{Z}$ is an ideal for any $n \in \mathbb{Z}$, so we need to show that an arbitrary ideal I of $\mathbb{Z}$ is of the form $n\mathbb{Z}$ for some $n \in \mathbb{Z}$. If $I = 0$, we are done. Otherwise, there is a non-zero element $a \in I$. If a is negative, then $-a \in I$ as well, hence we can find a positive integer in I. Let n be the smallest positive integer contained in I. We claim that $I = n\mathbb{Z}$. It's clear that $n\mathbb{Z} \subseteq I$. Conversely, let $a \in I$. Doing division with remainder, we can write $a = qn + r$ for some $q \in \mathbb{Z}$ and $0 \leq r < n$. It follows that $r = a - qn \in I$. Since $r < n$ and n was the smallest positive integer contained in I, we must have $r = 0$, so $a = qn$, i.e. $a \in \mathbb{N}Z$.

Lemma 1.3.5. *If $f\colon A \to B$ is a morphism of rings, then the **kernel***

$$\operatorname{Ker}(f) \coloneqq \{a \in A \mid f(a) = 0\} \tag{1.36}$$

is an ideal of A. The map f is injective if and only if $\operatorname{Ker}(f) = 0$.

Proof. If $a, b \in \operatorname{Ker}(f)$, then $f(a+b) = f(a) + f(b) = 0$, so $a + b \in \operatorname{Ker}(f)$. If $a \in A$ and $b \in \operatorname{Ker}(f)$, then $f(ab) = f(a)f(b) = 0$, so $ab \in \operatorname{Ker}(f)$. It is clear that if f is injective, then $\operatorname{Ker}(f) = 0$. Conversely, if $f(a) = f(b)$, then $0 = f(a) - f(b) = f(a - b)$, so $a - b \in \operatorname{Ker}(f)$ and $a = b$ if $\operatorname{Ker}(f) = 0$. □

Since an ideal I in a ring A is a subgroup of the commutative group $(A, +)$, we can form the quotient A/I as additive groups. Recall that A/I is the set of equivalences classes under the relation

$$a \sim b \text{ if } a - b \in I \tag{1.37}$$

on A, so the classes are of the form

$$\overline{a} \coloneqq a + I\ . \tag{1.38}$$

The addition on A descends to an addition on A/I via

$$\overline{a} + \overline{b} \coloneqq \overline{a + b}\ , \tag{1.39}$$

and this makes A/I into a (commutative) group. Moreover, the multiplication on A descends to a multiplication on A/I via

$$\overline{a} \cdot \overline{b} \coloneqq \overline{ab} \tag{1.40}$$

and this makes A/I into a ring, called the **quotient** of A by I. The **quotient map** $q\colon A \to A/I$ is a surjective ring morphism with $\operatorname{Ker}(q) = I$. Combined with Lemma 1.3.5, this shows in particular that the ideals in A are precisely the kernels of ring morphisms out of A—analogous to normal subgroups.

Lemma 1.3.6. *The quotient A/I satisfies the following universal property: if $f\colon A \to B$ is a ring morphism with $I \subseteq \operatorname{Ker}(f)$, then there is a unique ring morphism*

$$\overline{f}\colon A/I \to B \tag{1.41}$$

making the diagram

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[d, "q"] & B \\ A/I \arrow[ur, "\overline{f}"'] & \end{tikzcd} \tag{1.42}$$

commutative, i.e.

$$f = \overline{f} \circ q \quad (\textit{one says “f factors through q”}) \ . \tag{1.43}$$

Proof. We must have $\overline{f}(\overline{a}) = f(a)$, this shows uniqueness of $\overline{f}$. On the other hand, $\overline{f}$ defined like that is a well-defined ring morphism. Namely, if $\overline{a} = \overline{b}$, then $a - b \in I \subseteq \mathrm{Ker}(f)$, so $0 = f(a-b) = f(a) - f(b)$, so $f(a) = f(b)$. We have

$$\overline{f}(\overline{a}\overline{b}) = \overline{f}(\overline{ab}) = f(ab) = f(a)f(b) = \overline{f}(\overline{a})\overline{f}(\overline{b}) \ .$$

Similarly, one shows the other properties of a ring morphism. □

Lemma 1.3.7 (**First isomorphism theorem**). *If $f\colon A \to B$ is a ring morphism, then f induces a ring isomorphism*

$$A/\,\mathrm{Ker}(f) \xrightarrow{\simeq} \mathrm{Im}(f) \ , \quad \overline{a} \mapsto f(a) \ . \tag{1.44}$$

Proof. This is straightforward. □

Example 1.3.8. We claim the units in the quotient ring $\mathbb{Z}/n\mathbb{Z}$ are all (the images of) $m \in \mathbb{Z}$ with $\gcd(m,n) = 1$. If $m \in \mathbb{Z}$ is a unit in $\mathbb{Z}/n\mathbb{Z}$, there is $a \in \mathbb{Z}$ such that $am = 1$ in $\mathbb{Z}/n\mathbb{Z}$, i.e. $am - 1 \in n\mathbb{Z}$, hence there is $b \in \mathbb{Z}$ with $am - 1 = bn$, so $1 = am - bn$ and this means that $\gcd(m,n) = 1$. This argument can also be read backwards. It follows in particular that $\mathbb{Z}/n\mathbb{Z}$ is a field if and only if $n = p$ is a prime number. We denote this **finite field** by $\mathbb{F}_p$.

Remark 1.3.9. The **characteristic** of a ring A is the minimal $n \in \mathbb{N}_{>0}$ such that

$$\underbrace{1 + \ldots + 1}_{n \text{ times}} = 0$$

if such a number n exists, and 0 otherwise. For example, the characteristic of $\mathbb{Z}/n\mathbb{Z}$ is equal to n. One can show that if K is a field, then the characteristic of K is either 0 or a prime number p. Hence, if K is a **finite field**, its characteristic is a prime number p. One can furthermore show that in this case the number of elements of K is equal to p^n for some n. Finally, one can show that for any power p^n of p there is an up to isomorphism unique field with p^n elements. This field is denoted by $\mathbb{F}_{p^n}$.

Remark 1.3.10. For non-commutative rings one distinguishes between left, right, and two-sided ideals, meaning from which side the subset is closed under multiplication with ring elements. Lemmas 1.3.5, 1.3.6, and 1.3.7 then hold for two-sided ideals. In commutative rings, these three notions of ideals are all the same.

Similar to subrings, ideals are often described by a (preferably small) set of generators.

Lemma 1.3.11. *If $\mathcal{I}$ is a set of ideals in A, then the intersection $\bigcap_{I \in \mathcal{I}} I$ is an ideal in A as well.*

Proof. This is straightforward. □

Corollary 1.3.12. *For any subset $\boldsymbol{x}$ of a ring A there is a unique ideal minimal among all ideals containing $\boldsymbol{x}$. We call this the ideal **generated** by $\boldsymbol{x}$ and denote it by $(\boldsymbol{x})$. More explicitly, we have*

$$(\boldsymbol{x}) = \left\{ \sum_{i=1}^{n} a_i x_i \mid n \in \mathbb{N}, a_i \in A, x_i \in \boldsymbol{x} \right\} . \tag{1.45}$$

Proof. This is analogous to the proof of Corollary 1.1.27. □

In the notation $(\boldsymbol{x})$ there is no indication of the ring A because it is mostly clear from the context. If there could be confusion, we will explicitly write $(\boldsymbol{x})_A$ or $A\boldsymbol{x}$. A set of **generators** of an ideal I is a subset $\boldsymbol{x} \subseteq I$ such that $I = (\boldsymbol{x})$. Such a set always exists since we can take $\boldsymbol{x} = I$. An ideal is said to be **finitely generated** if it admits a finite set of generators.

Remark 1.3.13. Some people also say “basis” instead of generators of an ideal. I don’t like this terminology because it suggests the representation of elements in the generators is unique. But in general it is not. Consider for example the ideal $I \coloneqq (2, 3)$ in $\mathbb{Z}$. Then $12 \in I$ and $12 = 6 \cdot 2 = 4 \cdot 3$ are two distinct representations of this element.

There are three elementary operations one can perform with ideals: intersections, sums, and products. The intersection and sum have a meaning in an order-theoretic sense. Recall that a **partial order** on a set X is a relation $\leq$ which is reflexive ($x \leq x$), antisymmetric (if $x \leq y$ and $y \leq x$ then $x = y$), and transitive (if $x \leq y$ and $y \leq z$ then $x \leq z$).

Example 1.3.14. The set $\mathrm{Ideals}(A)$ of ideals in a ring A is a partially ordered set with respect to the inclusion relation $\subseteq$.

A **lower bound** of a subset Y of a partially ordered set X is an element $x \in X$ such that $x \leq y$ for all $y \in Y$. An **infimum** of Y is a greatest lower bound of Y, i.e. a lower bound x of Y such that if x' is another lower bound of Y, then $x' \leq x$. Analogously, we define an **upper bound** and a **supremum** (least upper bound) of a subset. In a general partially ordered set a supremum or infimum does not have to exist. But in $\mathrm{Ideals}(A)$ any subset $\mathcal{I}$ has an infimum, namely the intersection $\bigcap_{I \in \mathcal{I}} I$. A first guess for an upper bound would be to take the union. However, in general the union of ideals is not necessarily an ideal, e.g. for $(2), (3) \trianglelefteq \mathbb{Z}$ we have $(2) \cup (3) = 2\mathbb{Z} \cup 3\mathbb{Z}$ but $3 - 2 = 1 \notin (2) \cup (3)$. But we can always consider the ideal *generated* by the union. This has the following more explicit description.

Lemma 1.3.15. *If $\mathcal{I}$ is a set of ideals in a ring A, then*

$$\left(\bigcup_{I \in \mathcal{I}} I \right) = \left\{ \sum_{I \in \mathcal{I}} a_I \mid a_I \in I, \text{ all but finitely many } a_I = 0 \right\} . \tag{1.46}$$

*We denote this ideal by $\sum_{I \in \mathcal{I}} I$ and call it the **sum** of $\mathcal{I}$.*

Proof. This is straightforward. □

The sum $\sum_{I \in \mathcal{I}} I$ is a supremum of $\mathcal{I}$ in the partially ordered set $\mathrm{Ideals}(A)$. A partially ordered set in which all subsets have a supremum and an infimum is called a **complete lattice**. So, $\mathrm{Ideals}(A)$ is a complete lattice.

We have one further operation on ideals.

Definition 1.3.16. The **product** IJ of two ideals I and J in a ring A is defined as the ideal generated by the set $\{xy \mid x \in I, y \in J\}$.

Note that we have in particular defined for $n \in \mathbb{N}_{>0}$ the **n-th power**

$$I^n \coloneqq \underbrace{I \cdot \ldots \cdot I}_{n \text{ times}} \tag{1.47}$$

of an ideal I. Obviously, we have

$$IJ \subseteq I \cap J \subseteq I + J \ . \tag{1.48}$$

The product of ideals makes $\mathrm{Ideals}(A)$ into a monoid with unit A.

One can say a few things about how ideals behave under morphisms and this is very useful. If $f \colon A \to B$ is a ring morphism and $I \trianglelefteq A$ is an ideal, then the image $f(I) \subseteq B$ is not necessarily an ideal. Consider for example the embedding $f \colon \mathbb{Z} \to \mathbb{Q}$. Then $\mathbb{Z} \trianglelefteq \mathbb{Z}$ but $f(\mathbb{Z}) = \mathbb{Z} \subseteq \mathbb{Q}$ is not an ideal. But of course we can always consider the ideal *generated* by the image and thus get an inclusion-preserving map

$$f_* \colon \mathrm{Ideals}(A) \to \mathrm{Ideals}(B) \ , \quad I \mapsto (f(I)) \ . \tag{1.49}$$

Inverse images behave a bit better.

Lemma 1.3.17. *If $f \colon A \to B$ is a ring morphism and $J \trianglelefteq B$ is an ideal, then the inverse image*

$$f^{-1}(J) \coloneqq \{a \in A \mid f(a) \in J\} \tag{1.50}$$

is an ideal.

Proof. If $a, a' \in f^{-1}(J)$, then $f(a), f(a') \in J$, hence $f(a+a') = f(a)+f(a') \in J$, so $a + a' \in f^{-1}(J)$. Also $f(-a) = -f(a) \in J$, so $-a \in f^{-1}(J)$. Finally, if $a \in A$ and $a' \in J$, then $f(aa') = f(a)f(a') \in J$, so $aa' \in f^{-1}(J)$. □

We thus have an inclusion-preserving map

$$f^* \colon \mathrm{Ideals}(B) \to \mathrm{Ideals}(A) \ , \quad J \mapsto f^{-1}(J) \ , \tag{1.51}$$

in the reverse direction. The two maps f_* and f^* obviously satisfy the following relation:

$$f_*(I) \subseteq J \Leftrightarrow I \subseteq f^*(J) \ . \tag{1.52}$$

Here's a general lemma about such maps.

Lemma 1.3.18. *Let $(X, \leq)$ and $(Y, \leq)$ be two partially ordered sets and let $F \colon X \to Y$ and $G \colon Y \to X$ be two order-preserving maps such that*

$$F(x) \leq y \Leftrightarrow x \leq G(y) \tag{1.53}$$

for all $x \in X$ and $y \in Y$. Then F and G restrict to pairwise inverse bijections between the subsets

$$GF(X) \subseteq X \quad \text{and} \quad FG(Y) \subseteq Y \ . \tag{1.54}$$

Proof. The claim follows at once from the two relations

$$FGF(x) = F(x) \quad \text{and} \quad GFG(y) = G(y) \ . \tag{1.55}$$

Let us show the first one. For any $x \in X$ we have $F(x) \leq F(x)$, hence $x \leq GF(x)$ by (1.53). Since F is order-preserving, we get $F(x) \leq FGF(x)$. On the other hand, we have $GF(x) \leq GF(x)$, hence $FGF(x) \leq F(x)$ by (1.53). Hence, $FGF(x) = F(x)$ as claimed. The second relation $GFG(y) = G(y)$ is proven similarly. □

A pair of maps (F, G) as in Lemma 1.3.18 is called a (monotone) **Galois connection** and the maps FG and GF are called the associated **closure operators**. So, (1.51) shows:

LEMMA 1.3.19. *For any ring morphism $f\colon A \to B$ the pair (f_*, f^*) is a Galois connection between the sets of ideals in A and those in B.* □

It is not so easy in general to describe the closure operators and thus the subsets of ideals on which (f_*, f^*) restrict to bijections. But there's one particular situation where we know this.

LEMMA 1.3.20. *If $f\colon A \to B$ is a* surjective *ring morphism, then:*

(1) *For every ideal I in A the image $f(I)$ is an ideal in B.*
(2) *(f_*, f^*) restrict to bijections*

$$\{I \trianglelefteq A \mid \mathrm{Ker}(f) \subseteq I\} \xleftrightarrow{\sim} \mathrm{Ideals}(B)\ . \tag{1.56}$$

(3) *For every ideal J in B, the morphism f induces an isomorphism*

$$A/f^{-1}(J) \simeq B/J\ . \tag{1.57}$$

PROOF.

(1): It is clear that $f(I)$ is an additive subgroup of B. The problem is only multiplication with elements from B. So, let $y = f(x) \in f(I)$ and let $b \in B$. Since f is surjective, there is $a \in A$ with $f(a) = b$. Then $by = f(a)f(x) = f(ax) \in f(I)$. Hence, $f(I)$ is an ideal.

(2): Since f is surjective, we have

$$f_*f^*(J) = f_*(f^{-1}(J)) = (ff^{-1}(J)) = (J) = J$$

for any ideal J of B. Conversely, we will show that $f^*f_*(I) = I + \mathrm{Ker}(f)$ for any ideal I of A, so the image of f^*f_* is the set of all ideals of A containing $\mathrm{Ker}(f)$, and this proves the claim using Lemma 1.3.18. Note that $f^*f_*(I) = f^{-1}(f(I))$ by (1). We have $f(I + \mathrm{Ker}(f)) = f(I)$, so $I + \mathrm{Ker}(f) \subseteq f^{-1}(f(I))$. Conversely, let $x \in f^{-1}(f(I))$. Then $f(x) \in f(I)$, so $f(x) = f(y)$ for some $y \in I$, implying $0 = f(x-y)$, so $x - y \in \mathrm{Ker}(f)$ and therefore $x \in y + \mathrm{Ker}(f) \subseteq I + \mathrm{Ker}(f)$.

(3): Composition of f with the quotient map yields a surjective ring morphism $A \to B/J$ with kernel $f^{-1}(J)$, hence by the first isomorphism theorem (Lemma 1.3.7) the morphism f induces an isomorphism $A/f^{-1}(J) \simeq B/J$. □

An application of Lemma 1.3.20 to a quotient map yields in particular:

COROLLARY 1.3.21. *If I is an ideal of A and $q\colon A \to A/I$ is the quotient map, then (q_*, q^*) yield bijections*

$$\{J \trianglelefteq A \mid I \subseteq J\} \to \mathrm{Ideals}(A/I)\ . \tag{1.58}$$

□

In the setting of Corollary 1.3.21 we will write

$$J/I \coloneqq q_*(J) = q(J) \trianglelefteq A/I \tag{1.59}$$

for an ideal J of A with $I \subseteq J$.

LEMMA 1.3.22 (**Third isomorphism theorem**). *For ideals I, J in a ring A with $I \subseteq J$ there is a canonical ring isomorphism*

$$(A/I)/(J/I) \simeq A/J\ . \tag{1.60}$$

Proof. Left for you as exercise Exercise 1.3.24. □

We note one further observation.

Lemma 1.3.23. *If A is a subring of B and J is an ideal in B, then $A \cap J$ is an ideal in A.*

Proof. If $f\colon A \to B$ denotes the inclusion, then $J \cap A = f^{-1}(J)$, so this is an ideal by Lemma 1.3.17. □

Exercises.

Exercise 1.3.24. Prove the third isomorphism theorem for rings (Lemma 1.3.22).

Exercise 1.3.25. Let A be a ring and let $I_1, \ldots, I_n$ be ideals in A. Prove the following:

(1) The quotient maps $q_i\colon A \to A/I_i$ taken together induce a ring morphism

$$\varphi\colon A \to \prod_{i=1}^{n} A/I_i \ . \tag{1.61}$$

(2) φ is injective if and only if $\bigcap_{i=1}^{n} I_i = 0$.

(3) φ is surjective if and only if the I_i are mutually **coprime**, i.e. $I_i + I_j = A$ for all $i \neq j$.

(4) If the I_i are mutually coprime, then φ induces a ring isomorphism

$$A/\bigcap_{i=1}^{n} I_i \simeq \prod_{i=1}^{n} A/I_i \ . \tag{1.62}$$

Moreover,

$$\bigcap_{i=1}^{n} I_i = \prod_{i=1}^{n} I_i \ . \tag{1.63}$$

The statement about the isomorphism is called the **Chinese remainder theorem**.

Exercise 1.3.26. Let $(A_\lambda)_{\lambda\in\Lambda}$ be a *finite* family of rings. Show that we have a natural bijection

$$\operatorname{Ideals}\left(\prod_{\lambda\in\Lambda} A_\lambda\right) \simeq \prod_{\lambda\in\Lambda} \operatorname{Ideals}(A_\lambda) \ . \tag{1.64}$$

Show that this bijection does not exist for an *infinite* family of rings.

Exercise 1.3.27. Let A be a ring whose characteristic is a prime number p. Show that the map $A \to A$, $a \mapsto a^p$ is a ring morphism. This is called the **Frobenius endomorphism** of A. Is it an automorphism?

Exercise 1.3.28. Let A be a ring. We consider operations on ideals of A.

(1) Show that $\cdot$ is distributive over $+$, i.e.

$$I(J+K) = IJ + IK \ . \tag{1.65}$$

(2) Show $\cap$ is *not* necessarily distributive over $+$.

(3) Show that $\cap$ and $+$ satisfy the **modular law**: if $J \subseteq I$ or $K \subseteq I$, then

$$I \cap (J+K) = (I \cap J) + (I \cap K) \ . \tag{1.66}$$

(4) Show that in general we do not have $IJ = I \cap J$ but that we do have equality if I and J are coprime, i.e. if $I + J = A$.

Exercise 1.3.29. Draw the partially ordered sets of ideals in $\mathbb{Z}/30\mathbb{Z}$ and in $(\mathbb{Z}/4\mathbb{Z}) \times (\mathbb{Z}/2\mathbb{Z})$.

1.4. Algebras and polynomials rings

Algebras are rings with extra structure and they extend the concept of rings.

Definition 1.4.1. Let R be a ring. An **R-algebra** (or **algebra** over R) is a ring A together with an operation $R \times A \to A$, $(r, a) \mapsto ra$, satisfying the following properties:

$$r(a + a') = ra + ra' \ , \tag{1.67}$$

$$(r + r')a = ra + r'a \ , \tag{1.68}$$

$$(rr')a = r(r'a) \ , \tag{1.69}$$

$$r(aa') = (ra)a' = a(ra') \tag{1.70}$$

for all $r, r' \in R$ and $a, a' \in A$.

The ring R is also called the **base ring** of the algebra A.

Example 1.4.2. The ring $\mathrm{Maps}(X, R)$ from Example 1.1.6 is naturally an R-algebra via

$$(rf)(x) \coloneqq rf(x) \tag{1.71}$$

for $r \in R$, $f \in \mathrm{Maps}(X, R)$, and $x \in X$.

Example 1.4.3. *Any* ring is uniquely a $\mathbb{Z}$-algebra via the operation from Example 1.1.17.

There's an alternative point of view of algebras which is very important to have in mind. Any R-algebra A has an associated ring morphism

$$\varphi\colon R \to A \ , \ r \mapsto r1 \ . \tag{1.72}$$

This morphism is called the **structure morphism** of the algebra. Conversely, *any* morphism $\varphi\colon R \to A$ of rings defines an R-algebra structure on A via

$$ra \coloneqq \varphi(r)a \ . \tag{1.73}$$

Both constructions are obviously inverse to each other, so both concepts—algebras via scalar operation and via structure morphism—are equivalent. The upshot of the structure morphism is that it highlights the *relative* nature of an algebra: it is a ring *over* R. Instead of writing a ring morphism $R \to A$ horizontally—as we just did and as everyone else does—you should write it *vertically*:

$$\begin{tikzcd} A \\ R \arrow[u, "\varphi"] \end{tikzcd} \tag{1.74}$$

This emphasizes that A is an algebra *over* R. This will make much more sense later when we can view A more geometrically as a "fiber bundle" over R.

Definition 1.4.4. A **morphism** between R-algebras $\varphi\colon R \to A$ and $\psi\colon R \to B$ is a ring morphism $f\colon A \to B$ such that the diagram

$$\begin{tikzcd}
A \arrow[rr, "f"] & & B \\
& R \arrow[ul, "\varphi"] \arrow[ur, "\psi"'] &
\end{tikzcd} \tag{1.75}$$

commutes, i.e.

$$f \circ \varphi = \psi \text{ or, equivalently, } f(ra) = rf(a) \tag{1.76}$$

for all $r \in R$ and $a \in A$, i.e. f is compatible with the scalar operation.

Again, the composition of R-algebra morphisms is an R-algebra morphism, so R-algebras together with R-algebra morphisms form a category denoted R-Alg. Note that

$$\mathbb{Z}\text{-}\mathsf{Alg} = \mathsf{Ring}\ , \tag{1.77}$$

i.e. $\mathbb{Z}$-algebras and their morphisms is exactly the same as rings and their morphisms. Many constructions and concepts for rings admit a straightforward extension to algebras:

(1) Direct products of rings (Lemma 1.2.9 and Lemma 1.2.10) work in the same way also for R-algebras. You additionally define a component-wise scalar action on the direct product of the rings via

$$r(a_\lambda)_{\lambda \in \Lambda} \coloneqq (ra_\lambda)_{\lambda \in \Lambda} \tag{1.78}$$

and then you can replace "ring morphism" by "R-algebra morphism" everywhere.

(2) Quotients of rings (Lemma 1.3.6) by ideals work in the same way also for R-algebras. Note that an ideal in an R-algebra A is automatically stable under the scalar operation, so we get an induced scalar operation on a quotient A/I and then you can replace "ring morphism" by "R-algebra morphism" everywhere.

(3) A **subalgebra** is a subring (Definition 1.1.21) that is also stable under the scalar operation. The image of an algebra morphism is a subalgebra. Given a subset $\boldsymbol{x}$ of an R-algebra B there is a unique subalgebra $R[\boldsymbol{x}]$ minimal among all subalgebras of B containing $\boldsymbol{x}$. Explicitly, we have

$$R[\boldsymbol{x}] = \left\{ \sum_{\mu \in \mathbb{N}^n} r_\mu x_1^{\mu_1} \cdots x_n^{\mu_n} \mid \begin{array}{l} n \in \mathbb{N}, x_i \in \boldsymbol{x}, r_\mu \in R, \\ \text{all but finitely many } r_\mu = 0 \end{array} \right\} . \tag{1.79}$$

If A is an R-algebra, then a subset $\boldsymbol{x} \subseteq A$ with $A = R[\boldsymbol{x}]$ is called a set of **generators** of A as an R-algebra. An R-algebra is said to be **finitely generated** if it admits a finite set of generators as an R-algebra.

Now, we come to the most important example of algebras in this course. You probably all know the polynomial ring $R[X]$ in one variable over a ring R (or at least over a field). It's straightforward to generalize this to several variables. Let $n \in \mathbb{N}$ and let $R[X_1, \ldots, X_n]$ be the set of **polynomials** in the variables $X_1, \ldots, X_n$ with coefficients in R, i.e. formal expressions like

$$f = \sum_{\mu \in \mathbb{N}^n} r_\mu X_1^{\mu_1} \cdots X_n^{\mu_n} \quad \text{with } r_\mu \in R, \text{ all but finitely many } r_\mu = 0\ . \tag{1.80}$$

"Formal" always means that you shouldn't think of this as functions or anything concrete, just as symbols. It is convenient to use the shorthand notation

$$\boldsymbol{X} \coloneqq \{X_1, \ldots, X_n\} \ , \quad R[\boldsymbol{X}] \coloneqq R[X_1, \ldots, X_n] \ , \tag{1.81}$$

and

$$\boldsymbol{X}^\mu \coloneqq X_1^{\mu_1} \cdots X_n^{\mu_n} \quad \text{for } \mu \in \mathbb{N}^n \ . \tag{1.82}$$

The elements $\boldsymbol{X}^\mu$ are called **monomials**, the tuple μ is called the **multidegree** of $\boldsymbol{X}^\mu$, and

$$\deg(\boldsymbol{X}^\mu) \coloneqq \sum_{i=1}^n \mu_i \in \mathbb{N} \tag{1.83}$$

is the **degree** of $\boldsymbol{X}^\mu$. The element $r_\mu \in R$ in (1.80) is called the **coefficient** of $\boldsymbol{X}^\mu$ in f and

$$\deg(f) \coloneqq \max\{\deg(\boldsymbol{X}^\mu) \mid r_\mu \neq 0\} \in \mathbb{N} \tag{1.84}$$

is called the **degree** of f.

We have an obvious addition and multiplication of polynomials:

$$\left(\sum_{\mu \in \mathbb{N}^n} r_\mu \boldsymbol{X}^\mu \right) + \left(\sum_{\mu \in \mathbb{N}^n} s_\mu \boldsymbol{X}^\mu \right) \coloneqq \sum_{\mu \in \mathbb{N}^n} (r_\mu + s_\mu) \boldsymbol{X}^\mu \tag{1.85}$$

$$\left(\sum_{\mu \in \mathbb{N}^n} r_\mu \boldsymbol{X}^\mu \right) \cdot \left(\sum_{\mu \in \mathbb{N}^n} s_\mu \boldsymbol{X}^\mu \right) \coloneqq \sum_{\mu \in \mathbb{N}^n} \sum_{\substack{\nu, \xi \in \mathbb{N}^n \\ \nu + \xi = \mu}} (r_\nu \cdot s_\xi) \boldsymbol{X}^\mu \ . \tag{1.86}$$

And we also have a scalar operation of R on $R[\boldsymbol{X}]$ via

$$r \left(\sum_{\mu \in \mathbb{N}^n} r_\mu \boldsymbol{X}^\mu \right) \coloneqq \sum_{\mu \in \mathbb{N}^n} r r_\mu \boldsymbol{X}^\mu \ . \tag{1.87}$$

In total, this makes $R[\boldsymbol{X}]$ into an R-algebra. Note that the structure morphism $R \to R[\boldsymbol{X}]$, $r \mapsto r \cdot 1$, is injective and allows us to identify R with a subring of $R[\boldsymbol{X}]$. The polynomials in this subring are called the the **constant** polynomials.

I defined the polynomial ring in n variables but actually there's no problem to define the polynomial ring $R[\boldsymbol{X}]$ for a family $\boldsymbol{X} \coloneqq (X_\lambda)_{\lambda \in \Lambda}$ of variables indexed by an *arbitrary* (possibly infinite) set Λ. Your monomials will be

$$\boldsymbol{X}^\mu \coloneqq \prod_{\lambda \in \Lambda} X_\lambda^{\mu_\lambda} \ , \tag{1.88}$$

associated to tuples $\mu \in \mathbb{N}^\Lambda$ which are zero in all but finitely many places (so that the product is finite). Then the rest works similarly—I will leave the details to you.

The polynomial ring satisfies a universal property:

LEMMA 1.4.5. *Let $\boldsymbol{X} \coloneqq (X_\lambda)_{\lambda \in \Lambda}$ be a family of variables. The polynomial ring $R[\boldsymbol{X}]$ satisfies the following universal property: for any map $\varphi \colon \Lambda \to A$ into an R-algebra A there is a unique R-algebra morphism*

$$\hat{\varphi} \colon R[\boldsymbol{X}] \to A \tag{1.89}$$

such that

$$\hat{\varphi}(X_\lambda) = \varphi(\lambda) \tag{1.90}$$

for all $\lambda \in \Lambda$.

Proof. A morphism satisfying (1.90) must satisfy

$$\hat{\varphi}\left(\sum_\mu r_\mu \boldsymbol{X}^\mu\right) = \hat{\varphi}\left(\sum_\mu r_\mu \prod_\lambda X_\lambda^{\mu_\lambda}\right) = \sum_\mu r_\mu \prod_\lambda \hat{\varphi}\left(X_\lambda\right)^{\mu_\lambda} = \sum_\mu r_\mu \prod_\lambda \varphi(\lambda)^{\mu_\lambda} \ .$$

Hence, $\hat{\varphi}$ is uniquely determined by φ. We just need to show that $\hat{\varphi}$ defined in this way is in fact an R-algebra morphism. This is straightforward. We have

$$\begin{aligned}\hat{\varphi}\left(\sum_\mu r_\mu \boldsymbol{X}^\mu + \sum_\mu s_\mu \boldsymbol{X}^\mu\right) = \hat{\varphi}\left(\sum_\mu (r_\mu + s_\mu)\boldsymbol{X}^\mu\right) &= \sum_\mu (r_\mu + s_\mu)\left(\prod_\lambda \varphi(\lambda)^{\mu_\lambda}\right) \\ &= \sum_\mu r_\mu \prod_\lambda \varphi(\lambda)^{\mu_\lambda} + \sum_\mu s_\mu \prod_\lambda \varphi(\lambda)^{\mu_\lambda} \\ &= \hat{\varphi}\left(\sum_\mu r_\mu \boldsymbol{X}^\mu\right) + \hat{\varphi}\left(\sum_\mu s_\mu \boldsymbol{X}^\mu\right)\end{aligned}$$

and similarly

$$\begin{aligned}\hat{\varphi}\left(\sum_\mu r_\mu \boldsymbol{X}^\mu \cdot \sum_\mu s_\mu \boldsymbol{X}^\mu\right) &= \hat{\varphi}\left(\sum_\mu \sum_{\nu+\xi=\mu} r_\nu s_\xi \boldsymbol{X}^\mu\right) \\ &= \sum_\mu \left(\sum_{\nu+\xi=\mu} r_\nu s_\xi\right) \prod_\lambda \varphi(\lambda)^{\mu_\lambda} \\ &= \sum_\mu r_\mu \prod_\lambda \varphi(\lambda)^{\mu_\lambda} \cdot \sum_\mu s_\mu \prod_\lambda \varphi(\lambda)^{\mu_\lambda} \\ &= \hat{\varphi}\left(\sum_\mu r_\mu \boldsymbol{X}^\mu\right) \cdot \hat{\varphi}\left(\sum_\mu s_\mu \boldsymbol{X}^\mu\right) \ .\end{aligned}$$

Obviously, $\hat{\varphi}$ maps 1 to 1 and is compatible with the scalar operation, so $\hat{\varphi}$ is an R-algebra morphism. □

Example 1.4.6. Let R be a ring, let $\boldsymbol{X} \coloneqq (X_\lambda)_{\lambda\in\Lambda}$, and let $\Lambda' \subseteq \Lambda$ be a subset. Define

$$\boldsymbol{X}' \coloneqq (X_\lambda)_{\lambda\in\Lambda'} \quad \text{and} \quad \boldsymbol{X} \setminus \boldsymbol{X}' \coloneqq (X_\lambda)_{\lambda\in\Lambda\setminus\Lambda'} \ . \tag{1.91}$$

By the universal property of the polynomial ring we get an R-algebra morphism

$$\begin{array}{rcl} R[\boldsymbol{X}'] & \to & R[\boldsymbol{X}] \\ X_\lambda & \mapsto & X_\lambda \ . \end{array} \tag{1.92}$$

This morphism is injective, hence we can identify $R[\boldsymbol{X}']$ with a subring of $R[\boldsymbol{X}]$. This subring consists of all polynomials in which none of the variables X_λ with $\lambda \in \Lambda \setminus \Lambda'$ occurs. To get all of $R[\boldsymbol{X}]$ from this subring, we simply need to add all the variables X_λ with $\lambda \in \Lambda \setminus \Lambda'$. In other words, there is a canonical R-algebra isomorphism

$$R[\boldsymbol{X}] \simeq \left(R[\boldsymbol{X}']\right)[\boldsymbol{X} \setminus \boldsymbol{X}'] \ . \tag{1.93}$$

So, e.g. we have

$$R[X_1,\ldots,X_n] \simeq R[X_1,\ldots,X_{n-1}][X_n]\ . \tag{1.94}$$

This isomorphism is often used to deduce properties about polynomial rings in n variables from properties about polynomial rings in one variable—keep this in mind!

Example 1.4.7. Let R be a ring, let $\boldsymbol{X}$ be a set of variables, and let I be an ideal of R. Then there is a canonical R-algebra morphism

$$\begin{array}{rcl} R[\boldsymbol{X}] & \to & (R/I)[\boldsymbol{X}] \\ \sum_\mu r_\mu \boldsymbol{X}^\mu & \mapsto & \sum_\mu \overline{r}_\mu \boldsymbol{X}^\mu\ , \end{array} \tag{1.95}$$

where $\overline{r}_\mu$ denotes the image of R in R/I. This morphism is called **reduction** of coefficients modulo I. Its kernel consists precisely of the polynomials whose coefficients are contained in I, and this is equal to the ideal in $R[\boldsymbol{X}]$ generated by I. We thus have an R-algebra isomorphism

$$R[\boldsymbol{X}]/(R[\boldsymbol{X}]\cdot I) \simeq (R/I)[\boldsymbol{X}]\ . \tag{1.96}$$

Another (equivalent) point of view of the universal property of a polynomial ring is that for any family $\boldsymbol{x} \coloneqq (x_\lambda)_{\lambda\in\Lambda}$ of elements of an R-algebra A we get an R-algebra morphism

$$\mathrm{ev}_{\boldsymbol{x}} \colon R[\boldsymbol{X}] \to A \tag{1.97}$$

associated to the map $\lambda \mapsto x_\lambda$. We call this the **evaluation map** in $\boldsymbol{x}$ because for a polynomial

$$f = \sum_\mu r_\mu \boldsymbol{X}^\mu \in R[\boldsymbol{X}] \tag{1.98}$$

we have

$$f(\boldsymbol{x}) \coloneqq \mathrm{ev}_{\boldsymbol{x}}(f) = \sum_\mu r_\mu \boldsymbol{x}^\mu \in A\ , \tag{1.99}$$

where

$$\boldsymbol{x}^\mu \coloneqq \prod_{\lambda\in\Lambda} x_\lambda^{\mu_\lambda}\ , \tag{1.100}$$

i.e. for each λ we simply plug in x_λ for the variable X_λ. This gets particularly exciting when $\boldsymbol{x}$ is a family of generators of the R-algebra A. Then the image of $\mathrm{ev}_{\boldsymbol{x}}$ is precisely the subalgebra $R[\boldsymbol{x}]$ of A generated by the x_λ, so $\mathrm{ev}_{\boldsymbol{x}}$ is surjective. We thus conclude:

Lemma 1.4.8. *Any R-algebra is isomorphic to a quotient of a polynomial ring over R.* □

Clearly, A is a quotient of a polynomial ring in finitely many variables if and only if A is finitely generated as an R-algebra. The kernel of the morphism $\mathrm{ev}_{\boldsymbol{x}}$ consists of all polynomials $f \in R[\boldsymbol{X}]$ such that $f(\boldsymbol{x}) = 0$, i.e. the kernel describes precisely all the (polynomial) **relations** between the generators x_λ. When there are no relations, then the generating set is said to be **free** and A is said to be a **free** R-algebra—such a generating set exists if and only if A is isomorphic to a polynomial ring.

Compare the universal property of the polynomial ring to what you know from vector spaces: if you take a basis $(b_\lambda)_{\lambda\in\Lambda}$ of a vector space V, then for any vector space W and any map $\varphi\colon \Lambda \to W$ you get a well-defined linear map $\hat{\varphi}\colon V \to W$ mapping b_λ to $\varphi(\lambda)$, i.e. you can choose where the basis elements should map to and you get unique a linear map doing exactly this. The reason this works is because you don't have any (linear) relations between the basis elements. The polynomial ring $R[\boldsymbol{X}]$ satisfies the analogous property in the category of R-algebras: there are no (polynomial) relations between the X_λ, so morphisms out of $R[\boldsymbol{X}]$ are completely determined by where the X_λ map to. In categorical terms one says that $R[\boldsymbol{X}]$ is the **free object** over the set $\boldsymbol{X}$ in the category R-Alg.

Exercises.

Exercise 1.4.9. This exercise is a preparation for Exercise 1.4.10 below on the units in the polynomial ring. Let A be a ring. An element $x \in A$ is called **nilpotent** if there is $n \in \mathbb{N}$ with $x^n = 0$. Show the following:

(1) If $u \in A$ is a unit and $x \in A$ is nilpotent, then $u + x$ is a unit.
(2) If $x, y \in A$ are nilpotent, so is $x + y$.

Exercise 1.4.10. Let R be a ring. Use Exercise 1.4.9 to prove the following claim: a polynomial $f = r_0 + r_1X + \ldots + r_nX^n \in R[X]$ is a unit if and only if r_0 is a unit and all $r_1, \ldots, r_n$ are nilpotent. Now, determine the units in $\mathbb{Z}[X]$.

Exercise 1.4.11. What are the free objects in the category of groups?

1.5. Divisibility and factorization

The absence of multiplicative inverses leads to the problem of divisibility—this is what makes rings so interesting.

Definition 1.5.1. Let A be a ring and let $a, b \in A$. We say that a **divides** b (or that a is a **divisor** of b), written $a \mid b$, if there is $x \in A$ such that $ax = b$.

The zero $0 \in A$ plays a special role in divisibility. First, note that any element $a \in A$ is a divisor of 0 since $ax = 0$ with $x = 0$. Things get more interesting when this happens for $x \neq 0$.

Definition 1.5.2. An element $a \in A$ is called a **zero-divisor** if there is $0 \neq x \in A$ with $ax = 0$.

Note that being a zero-divisor means precisely that the **multiplication map**

$$\mu_a \colon A \to A\ , \quad x \mapsto ax\ , \tag{1.101}$$

is *not* injective. Conversely, a **non-zero-divisor** is an element $a \in A$ such that $ax = 0$ implies $x = 0$, i.e. the multiplication map μ_a is injective. Note that $0 \in A$ is a zero-divisor if and only if $A \neq 0$. Because "non-zero-divisor" is such an unwieldy word, one also says an element is **regular** if it is a non-zero-divisor.

Definition 1.5.3. An **integral domain** is a non-zero ring which has no zero-divisors except 0.

Example 1.5.4. $\mathbb{Z}$ is an integral domain but $\mathbb{Z} \times \mathbb{Z}$ is not: $(1, 0) \cdot (0, 1) = (0, 0)$.

Example 1.5.5. Any field is an integral domain.

Example 1.5.6. $\mathbb{Z}/n\mathbb{Z}$ is an integral domain if and only if n is a prime number.

Example 1.5.7. Let R be a ring. Then the polynomial ring $R[X]$ is an integral domain if and only if R is an integral domain. The proof is left as Exercise 1.5.27. Using Example 1.4.6 it follows that if R is an integral domain, then $R[X_1, \ldots, X_n]$ is an integral domain as well.

To get a better understanding of a zero-divisor a, it is often helpful to look at all the elements $x \in A$ "annihilating" it.

Definition 1.5.8. Let A be a ring. The **annihilator** of a subset $S \subseteq A$ is

$$\operatorname{Ann}_A(S) \coloneqq \{x \in A \mid xs = 0 \text{ for all } s \in S\} . \tag{1.102}$$

If the ring A is clear from the context, we just write $\operatorname{Ann}(S)$. It is clear that the annihilator $\operatorname{Ann}(S)$ is an ideal in A and that $\operatorname{Ann}(S) = \operatorname{Ann}((S))$, where the latter is the annihilator of the ideal generated by S. For an element $a \in A$ we write

$$\operatorname{Ann}(a) \coloneqq \operatorname{Ann}(\{a\}) = \{x \in A \mid xa = 0\} . \tag{1.103}$$

So, $a \in A$ is a zero-divisor if and only if $\operatorname{Ann}(a) \neq 0$.

Divisibility of elements can be expressed as an inclusion between the ideals they generate: we have

$$a \mid b \Longleftrightarrow (b) \subseteq (a) . \tag{1.104}$$

Definition 1.5.9. An ideal of the form (a) is called a **principal** ideal. A **principal ideal ring** (PIR) is a ring in which every ideal is principal; a **principal ideal domain** (PID) is an integral domain which is a principal ideal ring.

Example 1.5.10. $\mathbb{Z}$ is a principal ideal domain by Example 1.3.4. Any quotient of $\mathbb{Z}$ by an ideal is a principal ideal ring.

Example 1.5.11. Let R be a ring. Then the polynomial ring $R[X]$ is a principal ideal domain if and only if R is a field. The proof is left as Exercise 1.5.28.

The divisibility relation $\mid$ defines an equivalence relation $\sim$ on A via

$$a \sim b \Longleftrightarrow a \mid b \text{ and } b \mid a \Longleftrightarrow (a) = (b) . \tag{1.105}$$

In this case, we say that a and b are **associates**.

Lemma 1.5.12. *If $b = au$ with a unit u, then a and b are associates. The converse holds if a and b are regular.*

Proof. If $b = au$, then clearly $a \mid b$, but also $b \mid a$ since $a = bu^{-1}$. Conversely, suppose that $a, b \in A$ are associates, so $ax = b$ and $by = a$ for some x, y. Then $axy = by = a$, so $a(xy - 1) = 0$. If a is regular, this implies $xy = 1$, so x and y are units. □

You all know what a **prime number** is: it's a number $p > 1$ whose only (positive) divisors are 1 and p, i.e. you cannot further factorize $p = ab$. Why do you care? Because prime numbers are the "atoms" of numbers: any integer can be factorized into a product of prime numbers and this factorization is unique up to permutation of the factors.

It would be excellent to have something like this in a general ring, so let's take a look at this. Things get a bit awkward and difficult when you want to consider (unique) factorizations of zero-divisors as well. One can make all this work in general rings—to some extent and there are competing theories—but I don't want to delve into this here, see e.g. [1] if you want to know more about this.

Assumption 1.5.13. For the rest of this section we assume that A is an integral domain.

We begin by generalizing the characterizing property of a prime number.

DEFINITION 1.5.14. An element $p \in A$ is called **irreducible** if it is a non-zero non-unit and if $p = ab$ for some $a, b \in A$ implies that a or b is a unit.

So, p being irreducible means that there is no non-trivial factorization of p.

EXAMPLE 1.5.15. The irreducible elements in $\mathbb{Z}$ are precisely $\pm p$ for prime numbers p.

LEMMA 1.5.16. *A non-zero non-unit $p \in A$ is irreducible if and only if $p = ab$ for some $a, b \in A$ implies $p \sim a$ or $p \sim b$.*

PROOF. If p is irreducible and $p = ab$, then by definition a or b is a unit, so $p \sim a$ or $p \sim b$ by Lemma 1.5.12. Conversely, suppose $p = ab$ implies $p \sim a$ or $p \sim b$. Without loss of generality, assume $p \sim a$. By Lemma 1.5.12 this means $a = pu$ with a unit u (we are in an integral domain). Hence, $p = ab = pub$, so $p(1 - ub) = 0$, implying that $ub = 1$ (we are in an integral domain), i.e. b is a unit. This shows that p is irreducible. □

Now, what we would like to have in our ring A is that:

(1) Every non-zero non-unit $a \in A$ admits a factorization into irreducible elements, i.e. $a = p_1 \cdots p_r$ with p_i irreducible. If this holds, we call A **atomic**.
(2) All factorizations of a non-zero non-unit a into irreducible elements are **isomorphic** in the following sense: if

$$a = \prod_{i=1}^{r} p_i = \prod_{i=1}^{s} q_i \tag{1.106}$$

are two such factorizations, then $r = s$ and there is a permutation $\sigma \in S_r$ such that $p_i \sim q_{\sigma(i)}$ for all i. If this holds, we call A a **unique factorization domain**.

The properties won't hold in general. Let's start with the first.

LEMMA 1.5.17. *A principal ideal domain is atomic.*

PROOF. We need to show that any non-zero non-unit $a \in A$ admits a factorization into irreducible elements. Assume $a \in A$ is a non-zero non-unit not admitting a factorization into irreducible elements. We inductively construct a sequence $(a_n)_{n \in \mathbb{N}}$ of non-zero non-units in A not admitting a factorization into irreducible elements and this will eventually lead to a contradiction. We set $a_0 \coloneqq a$. Now suppose, we have constructed a_n for some $n \in \mathbb{N}$. The element a_n cannot be irreducible as otherwise we would have a factorization into irreducibles contrary to the assumption. Hence, we can write $a_n = a_{n+1} b_{n+1}$ with non-zero non-units a_{n+1}, b_{n+1}. If both these elements would admit a factorization into irreducibles, so would a_n, which is a contradiction. Hence, without loss of generality, a_{n+1} does not admit a factorization into irreducibles. By construction, we have $a_{n+1} \mid a_n$, so $(a_n) \subseteq (a_{n+1})$. We cannot have $a_n \mid a_{n+1}$ since then $a_n \sim a_{n+1}$, hence $a_n = a_{n+1} u$ for a unit $u \in A$ by Lemma 1.5.12, therefore $a_n = a_{n+1} u = a_{n+1} b$ which implies $a_{n+1}(u - b) = 0$, so $u - b = 0$ since a_{n+1} is non-zero and therefore b is a unit, contrary to our assumption. We thus have an *infinite* strictly ascending ideal chain

$$(a_0) \subsetneq (a_1) \subsetneq (a_2) \subsetneq \cdots . \tag{1.107}$$

We claim that such an infinite chain cannot exist in a principal ideal domain. Let $I_n := (a_n)$. Then $I := \bigcup_{n\in\mathbb{N}} I_n$ is an ideal (note: we really take the union of sets; it's an ideal because the I_n form a chain). Hence, $I = (x)$ for some $x \in A$. There is $n \in \mathbb{N}$ such that $x \in I_n$. But then $I \subseteq I_n$, so $I_m \subseteq I_n$ for all $m \in \mathbb{N}$, i.e. the chain becomes stationary, i.e. is not infinite—a contradiction. □

Note that a key ingredient in the proof was to show that there is no infinite ascending chain of (principal) ideals. Such **chain conditions** play a central role in commutative algebra. We will see more of this later (Chapter 7) and see that the class of rings satisfying this property is huge—basically any ring you can think of will satisfy it (if you don't have bad thoughts). So, existence of a factorization is not really a big problem. The problem is uniqueness.

Example 1.5.18. Consider the subring $\mathbb{Z}[\sqrt{-5}] = \{x + \sqrt{-5}y \mid x, y \in \mathbb{Z}\} \subseteq \mathbb{C}$. We have two distinct factorizations

$$6 = 2 \cdot 3 = (1 + \sqrt{-5}) \cdot (1 - \sqrt{-5}) \tag{1.108}$$

We claim that all the elements above are irreducible. A great tool is the norm map $N\colon \mathbb{Z}[\sqrt{-5}] \to \mathbb{N}$ defined by

$$N(x + \sqrt{-5}y) := (x + \sqrt{-5}y)(x - \sqrt{-5}y) = x^2 + 5y^2 \ . \tag{1.109}$$

You can easily convince yourself that the norm is multiplicative, i.e.

$$N(ab) = N(a)N(b) \tag{1.110}$$

for any $a, b \in \mathbb{Z}[\sqrt{-5}]$. In particular, if $a \in \mathbb{Z}[\sqrt{-5}]$ is a unit, then

$$1 = N(1) = N(aa^{-1}) = N(a)N(a^{-1}) \ .$$

But since $N(x + \sqrt{-5}y) = x^2 + 5y^2$, the only way to get this equal to 1 is $y = 0$ and $x = \pm 1$, so

$$N(a) = 1 \Longleftrightarrow a \in \mathbb{Z}[\sqrt{-5}]^\times \ , \quad \text{hence} \quad \mathbb{Z}[\sqrt{-5}]^\times = \{\pm 1\} \ . \tag{1.111}$$

Now, suppose that $2 \in \mathbb{Z}[\sqrt{-5}]$ were reducible, so we could write $2 = ab$ with non-units a, b. Then

$$4 = N(2) = N(ab) = N(a)N(b) \ .$$

Since a and b are non-units, this forces $N(a) = \pm 2$ and $N(b) = \pm 2$. But there is no way to get $x^2 + 5y^2$ equal to ± 2 for $x, y \in \mathbb{Z}$, so this is a contradiction and 2 must be irreducible. Similarly, you prove irreducibility of the other elements in (1.108). Since the units in $\mathbb{Z}[\sqrt{-5}]$ are just ± 1, it follows that in (1.108) we really have two distinct factorizations of the element 6.

How can we ensure uniqueness of the factorization? The answer is to strengthen the notion of irreducibility.

Definition 1.5.19. An element $p \in A$ is called **prime** if p is a non-zero non-unit and whenever $p \mid ab$ for some $a, b \in A$, then $p \mid a$ or $p \mid b$.

Lemma 1.5.20. *Every prime element is irreducible.*

Proof. Let p be prime. Suppose, we have a factorization $p = ab$. Then $a \mid p$ and $b \mid p$. On the other hand, $p \mid ab$ and since p is prime, we have $p \mid a$ or $p \mid b$. Hence, $p \sim a$ or $p \sim b$, so p is irreducible by Lemma 1.5.16. □

Lemma 1.5.21. *If a non-zero non-unit $a \in A$ admits a factorization into prime elements, then all factorizations of a into irreducible elements are isomorphic and all factors of such factorizations are already prime.*

Proof. Let $a = p_1 \cdots p_r$ be a factorization into prime elements. Let $a = q_1 \cdots q_s$ be another factorization into irreducible elements. Since p_1 is prime and $p_1 \mid a = q_1 \cdots q_s$, it follows that $p_1 \mid q_j$ for some j, so $p_1 x = q_j$ for some $x \in A$. Since q_j is irreducible and p_1 is a non-unit (since it is prime), the element x must be a unit, so $p_1 \sim q_j$ and q_j is also prime. Inductively, we conclude that $r = s$, that there is a permutation $\sigma \in S_r$ such that $p_i \sim q_{\sigma(i)}$ for all i, and all q_j are prime as well. □

Hence, prime elements are a better generalization of prime numbers than irreducible elements. In unique factorization domains both notions become the same.

Lemma 1.5.22. *In a unique factorization domain, irreducible elements are prime.*

Proof. Let p be an irreducible element. Suppose $p \mid ab$ for some $a, b \in A$, so $px = ab$ for some $x \in A$. Let $a = p_1 \cdots p_r$, $b = p_{r+1} \cdots p_{r+s}$, and $x = y_1 \ldots y_t$ be factorizations into irreducible elements. Then

$$p_1 \cdots p_{r+s} = p y_1 \cdots y_t$$

are two factorizations of an element into irreducible elements. By uniqueness, p is an associate of one of the elements p_i, so $p \mid p_i$ for some i, hence $p \mid a$ or $p \mid b$. This shows that p is prime. □

We have thus obtained a strategy to prove that an integral domain is a unique factorization domain:

Corollary 1.5.23. *Let A be an integral domain. If A is atomic and every irreducible element of A is prime, then A is a unique factorization domain.* □

This is what we'll do to prove the next lemma.

Lemma 1.5.24. *Every principal ideal domain is a unique factorization domain.*

Proof. Let A be a principal ideal domain. From Lemma 1.5.17 we already know that A is atomic. So, we just need to show that every irreducible element $p \in A$ is prime. Suppose that $p \mid ab$ for some $a, b \in A$ and suppose that $p \nmid a$. We need to show that $p \mid b$. Since A is a principal ideal domain, there is $g \in A$ such that $(p, a) = (g)$. Then $g \mid p$ and since p is irreducible, this forces g to be a unit. Hence, $(p, a) = (1)$ and so there is $x, y \in A$ with $1 = ax + py$. Multiplying by b yields $b = bax + bpy$. Since $p \mid ab$, it follows that p divides the whole right-hand-side of the equation, hence $p \mid b$. □

Example 1.5.25. $\mathbb{Z}$ is a unique factorization domain.

Example 1.5.26. Let K be a field. Then the polynomial ring $K[X]$ in one variable is a principal ideal domain by Example 1.5.11, hence it is a unique factorization domain by Lemma 1.5.24. More generally, we will show in Lemma 8.1.2 that the polynomial ring $R[X]$ over a unique factorization domain R is again a unique factorization domain (the proof is elementary, we could have proven this here). This implies in particular, that $K[X_1, \ldots, X_n]$ is a unique factorization domain using Example 1.4.6.

Exercises.

Exercise 1.5.27. Prove the claim in Example 1.5.7.

Exercise 1.5.28. Prove the claim in Example 1.5.11.

CHAPTER 2

Prime spectrum

Recall the ring $\mathbb{Z}[\sqrt{-5}]$ and the two distinct factorizations

$$6 = 2 \cdot 3 = (1+\sqrt{-5}) \cdot (1-\sqrt{-5}) \tag{2.1}$$

of 6 into irreducible elements in this ring. So, $\mathbb{Z}[\sqrt{-5}]$ is not a unique factorization domain (and thus not a principal ideal domain). But you know what's funny? If instead of elements you consider their principal ideals, the situation looks much better! Consider the ideals

$$P \coloneqq (2, 1+\sqrt{-5}) \;, \quad Q_1 \coloneqq (3, 1+\sqrt{-5}) \;, \quad Q_2 \coloneqq (3, 1-\sqrt{-5}) \tag{2.2}$$

in $\mathbb{Z}[\sqrt{-5}]$. Then you can check that

$$(2) = P \cdot P \;, \quad (3) = Q_1 \cdot Q_2 \;, \quad (1+\sqrt{-5}) = P \cdot Q_1 \;, \quad (1-\sqrt{-5}) = P \cdot Q_2 \;. \tag{2.3}$$

And now look what happens to (2.1) when we pass to ideals:

$$P \cdot P \cdot Q_1 \cdot Q_2 = (2) \cdot (3) = (6) = (1+\sqrt{-5}) \cdot (1-\sqrt{-5}) = P \cdot Q_1 \cdot P \cdot Q_2 \;. \tag{2.4}$$

Both factorizations become identical! Where exactly the P and Q_i come from shouldn't matter right now. What matters are the following two observations:

(1) Prime elements seem to be the correct generalization of prime numbers to general rings.

(2) Ideals seem to behave better than numbers. In fact, this is the story behind the concept and terminology of ideals: Dedekind considered ideals as "ideal numbers" which behave better than actual numbers.

So, taken together, we should look for a notion of "prime ideal"!

2.1. Prime ideals

Recall that we can rephrase the divisibility $a \mid b$ in terms of ideal containment:

$$a \mid b \iff (b) \subseteq (a) \tag{2.5}$$

Hence, we can define a prime element more ideal-theoretic as a non-zero non-unit p such that $(a)(b) \subseteq (p)$ implies $(a) \subseteq (p)$ or $(b) \subseteq (p)$. Let's be naive and transport this definition to a general ring and to general ideals. Whereas in the last section we restricted to integral domains because element factorization of zero-divisors is awkward, we will make no assumption on A being an integral domain here anymore.

Definition 2.1.1. A **prime ideal** in a ring A is an ideal $P \neq A$ such that $IJ \subseteq P$ implies $I \subseteq P$ or $J \subseteq P$ for any ideals I and J of A.

Note that—in contrast to the definition of prime elements—we did not assume $P \neq 0$, and it's correct not to do so. It's basically just a historical convention that prime numbers are non-zero, see Example 2.3.4 for more on this. Let's play with the definition. To prove that a given ideal is prime, one often makes use of the following equivalent characterizations.

Lemma 2.1.2. *For an ideal $P \trianglelefteq A$ the following are equivalent:*

(1) *P is a prime ideal.*
(2) *$P \neq A$ and if $ab \in P$, then $a \in P$ or $b \in P$.*
(3) *$P \neq A$ and if $a \notin P$ and $b \notin P$, then $ab \notin P$.*
(4) *A/P is an integral domain.*

Proof.

(1) ⇒ (2): if $ab \in P$, then $(a)(b) = (ab) \subseteq P$, so $(a) \subseteq P$ or $(b) \subseteq P$ since P is prime, hence $a \in P$ or $b \in P$.

(2) ⇒ (1): Let I, J be ideals with $IJ \subseteq P$. Suppose $I \not\subseteq P$ and $J \not\subseteq P$. Then there is $a \in I \setminus P$ and $b \in J \setminus P$. We have $ab \in IJ \subseteq P$. Hence, $a \in P$ or $b \in P$ by assumption, which is a contradiction. So, we must have $I \subseteq P$ or $J \subseteq P$.

(2) ⇔ (3): The second statement is just the negation of the third.

(3) ⇔ (4): Clear. □

Corollary 2.1.3. *The zero ideal in a ring A is a prime ideal if and only if A is an integral domain.*

Corollary 2.1.4. *If A is an integral domain, an element $p \in A$ is a prime element if and only if (p) is a non-zero prime ideal.*

Proof. Let p be a prime element. By definition, p is a non-zero non-unit, so $0 \neq (p) \neq A$. If $ab \in (p)$, then $p \mid ab$, hence $p \mid a$ or $p \mid b$, so $a \in (p)$ or $b \in (p)$. This proves by Lemma 2.1.2 that (p) is a non-zero prime ideal. Conversely, suppose that (p) is a non-zero prime ideal. Then p is non-zero. A prime ideal is by definition not the whole ring, so p is a non-unit. Suppose that $p \mid ab$. Then $(a)(b) = (ab) \subseteq (p)$, hence $(a) \subseteq (p)$ or $(b) \subseteq (p)$, i.e. $p \mid a$ or $p \mid b$. This proves that p is a prime element. □

Corollary 2.1.5. *If A is a principal ideal domain, the prime ideals in A are precisely the zero ideal and the ideals (p) for p a prime element.*

Corollary 2.1.6. *The prime ideals in $\mathbb{Z}$ are precisely the zero ideal and the ideals (p) for p a prime number.*

So, prime ideals generalize prime elements, which generalize prime numbers—great!

Example 2.1.7. The ideals $P, Q_1, Q_2 \trianglelefteq \mathbb{Z}[\sqrt{-5}]$ from (2.2) are prime ideals. Let's show this for $P = (2, 1+\sqrt{-5})$. We get $\mathbb{Z}[\sqrt{-5}]$ from $\mathbb{Z}$ by adding $\sqrt{-5}$. How's this element characterized? It's a root of the polynomial $X^2 + 5 \in \mathbb{Z}[X]$. So,

$$\mathbb{Z}[\sqrt{-5}] \simeq \mathbb{Z}[X]/(X^2+5) \,. \tag{2.6}$$

You can check that an isomorphism is given by mapping X to $\sqrt{-5}$. So,

$$\begin{aligned}\mathbb{Z}[\sqrt{-5}]/(2, 1+\sqrt{-5}) &\simeq \left(\mathbb{Z}[X]/(X^2+5)\right)/(2, 1+X) \\ &\overset{(a)}{\simeq} \mathbb{Z}[X]/(X^2+5, 2, 1+X) \\ &\overset{(a)}{\simeq} \left(\mathbb{Z}[X]/(2)\right)/(X^2+5, 1+X) \\ &\overset{(b)}{\simeq} \mathbb{F}_2[X]/(X^2+5, X+1)\end{aligned}$$

$$\begin{aligned}&\simeq \mathbb{F}_2[X]/(X^2+1, X+1)\\&\simeq \mathbb{F}_2\,.\end{aligned}$$

In the isomorphisms marked with (a) we have used the third isomorphism theorem (Lemma 1.3.22). Recall from Example 1.3.8 that $\mathbb{F}_2 = \mathbb{Z}/2\mathbb{Z}$ is a field. The isomorphism (b) is an application of Example 1.4.7. Since $\mathbb{F}_2$ is a field and thus an integral domain, it follows that P is a prime ideal by Lemma 2.1.2. Similarly you can prove this for Q_1 and Q_2.

Hence, (2.4) gives a factorization of the ideal $(6) \trianglelefteq \mathbb{Z}[\sqrt{-5}]$ into prime ideals, and this factorization repaired the non-uniqueness on the element level. But now it is important to believe me that we are not so much interested in factorizations of ideals into prime ideals. We will come to this only much later (Chapter 9) as it holds only in very special rings like $\mathbb{Z}[\sqrt{-5}]$ etc. The concept of a prime ideal itself—ignoring any factorization questions—is a fundamental one for *any* commutative ring. So, let's forget about the factorization questions and study prime ideals more closely.

Example 2.1.8. Let K be a field. Then for any $1 \leq r \leq n$ the ideal $(X_1, \ldots, X_r)$ is a prime ideal in $K[X_1, \ldots, X_n]$ because

$$K[X_1, \ldots, X_n]/(X_1, \ldots, X_r) \simeq K[X_{r+1}, \ldots, X_n] \tag{2.7}$$

is an integral domain. We thus have a chain

$$(0) \subsetneq (X_1) \subsetneq (X_1, X_2) \subsetneq \cdots \subsetneq (X_1, \ldots, X_n) \tag{2.8}$$

of prime ideals in $K[X_1, \ldots, X_n]$. Notice that in, e.g. $\mathbb{Z}[X_1, \ldots, X_n]$, we can find an even longer chain, e.g.

$$(0) \subsetneq (p) \subsetneq (X_1) \subsetneq (X_1, X_2) \subsetneq \cdots \subsetneq (X_1, \ldots, X_n) \tag{2.9}$$

for any prime number p.

Remark 2.1.9. We have defined prime elements in integral domains only. Believing in the concept of prime ideals, we could use Corollary 2.1.4 to define the notion of a prime element in arbitrary rings: an element of a ring is prime if it is non-zero and generates a prime ideal. This actually boils down to exactly the same defining divisibility property as before in Definition 1.5.19. But still, element factorization of zero-divisors remains an intricate problem not readily solved by prime elements and we're also not interested in this here, see Example 2.2.3.

Exercises.

Exercise 2.1.10. Let K be a field. Let $p \in K[X]$ and let $\phi\colon K[X_1, X_2] \to K[X]$ be the morphism defined by $X_1 \mapsto X$ and $X_2 \mapsto p$, i.e. $\phi(f) = f(X, p)$. Show that $\operatorname{Ker}\phi = (X_2 - p(X_1))$ and conclude that this is a prime ideal in $K[X_1, X_2]$. Use this to show that $(X_1 + X_2 - 1) \subseteq K[X_1, X_2]$ is a prime ideal.

2.2. Functoriality

Let A be a ring. The set $\operatorname{Spec}(A)$ of all prime ideals in A is called the **prime spectrum** of A. A first indication that $\operatorname{Spec}(A)$ is an important footprint of A is the following lemma on "functoriality".

Lemma 2.2.1. *If $f\colon A \to B$ is a ring morphism, then the map*

$$\begin{array}{rcl} f^*\colon \mathrm{Ideals}(B) & \to & \mathrm{Ideals}(A)\ , \\ J & \mapsto & f^{-1}(J)\ , \end{array} \tag{2.10}$$

from (1.51) *restricts to a map*

$$f^*\colon\ \mathrm{Spec}(B) \to \mathrm{Spec}(A)\ . \tag{2.11}$$

Proof. We just need to show that if Q is a prime ideal in B, then $f^{-1}(Q)$ is a prime ideal in A. The composition g of $f\colon A \to B$ and the quotient map $B \to B/Q$ is a morphism $A \to B/Q$ with kernel equal to $f^{-1}(Q)$. Hence, we have an isomorphism $A/f^{-1}(Q) \simeq \mathrm{Im}(g) \subseteq B/Q$. Since Q is a prime ideal in B, the quotient B/Q is an integral domain, hence the subring $\mathrm{Im}(g)$ is an integral domain, hence $A/f^{-1}(Q)$ is an integral domain, hence $f^{-1}(Q)$ is a prime ideal in A. □

Recall that you should view this map vertically:

$$\begin{tikzcd} \mathrm{Spec}(B) \arrow[d, "f^*"] \\ \mathrm{Spec}(A) \end{tikzcd} \tag{2.12}$$

If $f\colon A \to B$ and $g\colon B \to C$ are two ring morphisms, then you can easily check that

$$(g \circ f)^* = f^* \circ g^*\colon\ \mathrm{Spec}(C) \to \mathrm{Spec}(A)\ . \tag{2.13}$$

Hence, mapping a ring A to its prime spectrum $\mathrm{Spec}(A)$ and mapping a ring morphism $f\colon A \to B$ to $f^*\colon\ \mathrm{Spec}(B) \to \mathrm{Spec}(A)$ combines to a contravariant functor

$$\mathrm{Spec}\colon \mathsf{Ring} \to \mathsf{Set}\ . \tag{2.14}$$

Because of the functoriality, we can completely understand the spectrum of a quotient:

Corollary 2.2.2. *If $f\colon A \to B$ is a* surjective *morphism of rings, then (f_*, f^*) restrict to bijections*

$$\mathrm{Spec}(B) \simeq \{P \in \mathrm{Spec}(A) \mid \mathrm{Ker}(f) \subseteq P\}\ . \tag{2.15}$$

In particular, for an ideal $I \trianglelefteq A$ we have

$$\mathrm{Spec}(A/I) \simeq \{P \in \mathrm{Spec}(A) \mid I \subseteq P\}\ . \tag{2.16}$$

Proof. From Lemma 1.3.20 we know that (f_*, f^*) restrict to bijections between the set of ideals in B and the set of ideals in A containing $\mathrm{Ker}(f)$. From Lemma 2.2.1 we know that if Q is a prime ideal in B, then $f^*(Q)$ is a prime ideal in A. On the other hand, if P is a prime ideal in A, then $f^*f_*(P) = P$ and by Lemma 1.3.20 we have an isomorphism

$$A/P = A/f^*f_*(P) \simeq B/f_*(P)\ .$$

Since P is prime, A/P is an integral domain, hence $B/f_*(P)$ is an integral domain and therefore $f_*(P)$ is a prime ideal. This proves the claim. □

Example 2.2.3. By Corollary 2.2.2 the prime ideals in $\mathbb{Z}/6\mathbb{Z}$ are in bijection with the prime ideals of $\mathbb{Z}$ containing (6). We know from Corollary 2.1.6 that the non-zero prime ideals of $\mathbb{Z}$ are of the form (p) for p a prime number. Since $(6) \subseteq (p)$ means $p \mid 6$, this only leaves (2) and (3). Hence,

$$\operatorname{Spec}(\mathbb{Z}/6\mathbb{Z}) \simeq \{(2), (3)\} . \tag{2.17}$$

I find this absolutely makes sense: you mod out 6, and the prime ideals remaining are (2) and (3). Taking up Remark 2.1.9, you can consider 2 and 3 as prime elements in $\mathbb{Z}/6\mathbb{Z}$. But we have

$$4 = 2^2 = 2^4 = \dots , \tag{2.18}$$

so a factorization of the zero-divisor 4 into prime elements is not unique. Because of such issues we're not really interested in element factorizations—prime ideals themselves are interesting.

Example 2.2.4. Let K be a field and consider $K[X]/(X^2)$. Prime ideals in this ring correspond to prime ideals in $K[X]$ containing (X^2). Let P be such a prime ideal. Since $(X^2) \subseteq P$, we have $X^2 = X \cdot X \in P$, hence $X \in P$ because P is prime. So, $(X) \subseteq P$. But $K[X]/(X) \simeq K$, and since a field just has the trivial ideals, we have $P = (X)$, i.e.

$$\operatorname{Spec}(K[X]/(X^2)) = \{(X)\} . \tag{2.19}$$

Functoriality also gives us an important source of prime ideals—more about this in a bit.

Corollary 2.2.5. *If $f \colon A \to B$ is a ring morphism into an integral domain B, then* $\operatorname{Ker}(f) \in \operatorname{Spec}(A)$.

Proof. If B is an integral domain, then $(0) \in \operatorname{Spec}(B)$ by Corollary 2.1.3, hence $\operatorname{Ker}(f) = f^{-1}(0) \in \operatorname{Spec}(A)$ by Lemma 2.2.1. □

2.3. Maximal ideals

If prime ideals are really so fundamental, there should be a good supply of them in any ring. This is what we'll show now. It's one of the first big theorems in commutative algebra.

Definition 2.3.1. An ideal M in a ring A is called **maximal** if $M \neq A$ and there is no ideal I in A with $M \subsetneq I \subsetneq A$.

Lemma 2.3.2. *An ideal M in A is maximal if and only if A/M is a field.*

Proof. Note that by definition and Corollary 1.3.21, an ideal M is maximal if and only if $\operatorname{Ideals}(A/M) = \{(0), A/M\}$. The lemma thus follows from the following claim: a non-zero ring A is a field if and only if it has precisely two ideals (necessarily the trivial ideals). Namely, if A is a field, this is obviously true. Conversely, suppose A has only two ideals. If $a \in A$ is not a unit, then $(a) \neq A$, so $(a) = (0)$ necessarily, i.e. $a = 0$; so, any non-zero element is a unit, i.e. A is a field. □

Corollary 2.3.3. *Maximal ideals are prime ideals.*

Example 2.3.4. In $\mathbb{Z}$ the ideals (p) for p a prime number are maximal: $(p) \subseteq (q)$ implies $q \mid p$, and since p is irreducible, we then have $q \sim p$, and therefore $(p) = (q)$, or $(q) = \mathbb{Z}$. Since maximal ideals are prime ideals and the only other prime ideal is the zero ideal, these are all the maximal ideals in $\mathbb{Z}$. More generally, in any principal ideal domain the maximal ideals are precisely the (p) for p a prime element and the only other non-maximal prime ideal is the zero ideal.

This brings us back to the question why the zero ideal is included as a prime ideal but a prime number is non-zero by definition. Maybe one should have said "maximal number" instead of "prime number" and should have defined a "prime number" to be a "maximal number" or 0—but back in the old times when prime numbers were introduced one couldn't see that far. In the end, it's just a convention.

The only problem is somehow: why should there be a maximal ideal at all? If the set of ideals is infinite, weird things could happen. But they cannot, and this follows from basic set theory. Let's consider a general partially ordered set $(X, \leq)$, e.g. Ideals(A). A **chain** in X is a subset Y of X which is totally ordered with respect to $\leq$, i.e. $y \leq y'$ or $y' \leq y$ for all $y, y' \in Y$. A **maximal element** in X is an element $x \in X$ such that there is no $x' \in X$ with $x < x'$.

Lemma 2.3.5 (**Zorn's lemma**). *If X is a non-empty partially ordered set in which every chain has an upper bound in X, then X has a maximal element.*

Before I come to the proof (which I will only sketch), let's take this lemma into action. Many proofs in commutative algebra go along similar lines, so really make sure you understand it.

Theorem 2.3.6 (Krull, 1929). *Every non-zero ring A has a* ***maximal ideal****.*

Proof. Let Σ be the set of proper ideals of A. This is a partially ordered set with respect to inclusion and it is non-empty since A is non-zero. We want to show that every chain in Σ has an upper bound. By Zorn's lemma, this implies that Σ has a maximal element, and this is a maximal ideal in A by definition. So, let $(I_\lambda)_{\lambda \in \Lambda}$ be a chain in Σ. Then $I \coloneqq \bigcup_{\lambda \in \Lambda} I_\lambda$ is an ideal (again, as in the proof of Lemma 1.5.17, this works because we take a union of a chain of ideals). Since $1 \notin I_\lambda$ for all $\lambda \in \Lambda$, we also have $1 \notin I$, i.e. $I \in \Sigma$. Obviously, I is an upper bound of the chain.[1] □

Corollary 2.3.7. *Any proper ideal is contained in a maximal ideal.*

Proof. This follows immediately from Theorem 2.3.6 in combination with Corollary 1.3.21. □

Corollary 2.3.8. *Every non-unit is contained in a maximal ideal.*

Great, all that remains to prove is Zorn's lemma. But this is a bit tricky since it relies on the **axiom of choice** from set theory which states that the product $\prod_{\lambda \in \Lambda} X_\lambda$ of a family of non-empty sets X_λ is non-empty, so you can choose an element $x_\lambda \in X_\lambda$ for any $\lambda \in \Lambda$. This statement sounds pretty obvious but when Λ is infinite, things can get weird and the term "axiom" already suggests that this does not follow from basic set theoretic axioms. Some mathematicians (in general) actually don't want to assume the axiom of choice or always point out explicitly when they assume it. But, you not only need the axiom of choice for proving Zorn's lemma—Zorn's lemma is even equivalent to the axiom of choice! To make things weirder, the existence of a maximal ideal in any non-zero ring is equivalent to Zorn's lemma, and thus to the axiom of choice! Without the axiom of choice, we wouldn't get far in commutative algebra. Hence:

[1] Note: when we work with ideals we can always take the union of ideals in a chain to get an ideal which bounds the chain. But the assumption in Zorn's lemma is that this bound lies *inside* the set of ideals in which we want to find a maximal element. In this case it's the set of all proper ideals, so we need to show that the union is a proper ideal. Make sure you understand this delicate point.

Assumption 2.3.9. We will always assume the axiom of choice.

Sketch of proof of Zorn's lemma. Let X be a partially ordered set satisfying the assumption, i.e. every chain has an upper bound. Assume, X does not have a maximal element. We build up a contradiction as follows. Since X is non-empty, we can pick an arbitrary $x_0 \in X$. Since X does not have a maximal element, we can pick some $x_1 \in X$ with $x_0 < x_1$. We continue like this and construct a sequence $x_0 < x_1 < x_2 \cdots$ in X indexed by all $i \in \mathbb{N}$. This is a chain in X, so by assumption it has an upper bound x_ω. Here, ω denotes the first infinite ordinal coming after all natural numbers. Now, we can start the process all over: we can choose $x_\omega < x_{\omega+1} < x_{\omega+2} < \cdots$ etc and get an element for the next limit ordinal 2ω by choosing an upper bound for the chain. In this way, by transfinite induction, we get a chain (x_α) indexed by *all* ordinals α. To make this work, i.e. to choose the elements, we really needed the axiom of choice. Now, the thing is that all the elements x_α are distinct, so we get an injection from the class Ord of all ordinals into X. But the class Ord is a proper class (it is not a set, it is too big—like the set of all sets) and X is a set. So, this is not possible. □

So, now we know that $\operatorname{Spec}(A) \neq \emptyset$ for any non-zero ring A. By $\operatorname{Max}(A) \subseteq \operatorname{Spec}(A)$ we denote the set of maximal ideals in A and call it the **maximal spectrum** of A. In contrast to prime ideals, maximal ideals are not functorial in general, i.e. if $f\colon A \to B$ is a ring morphism and $N \in \operatorname{Max}(B)$, then we don't necessarily have $f^{-1}(N) \in \operatorname{Max}(A)$. Consider for example the inclusion $f\colon \mathbb{Z} \to \mathbb{Q}$. We have $0 \in \operatorname{Max}\mathbb{Q}$ but $f^{-1}(0) = 0$ is not a maximal ideal of $\mathbb{Z}$. The situation is better for surjective morphisms:

Lemma 2.3.10. *If $f\colon A \to B$ is* surjective *and $N \in \operatorname{Max}(B)$, then $f^{-1}(N) \in \operatorname{Max}(A)$.*

Proof. This follows immediately from Lemma 1.3.20. □

Exercises.

Exercise 2.3.11. Show that a ring A has a unique maximal ideal if and only if $A \setminus A^\times$ is an ideal in A. In this case, the unique maximal ideal is equal to $A \setminus A^\times$. Such rings are called **local**—more about this later. Find examples of local rings!

2.4. A glimpse of algebraic geometry

Prime ideals have a geometric meaning and this lies at the heart of algebraic geometry. It will take you some time to develop a good intuition for this because it goes beyond things you can easily imagine or visualize—it's like general relativity. But once you get used to it, you have an extremely powerful tool! Algebraic geometry is a very old subject but the modern idea of employing prime ideals in full generality—and thereby bending the common intuition about reality—goes back to A. Grothendieck's *Éléments de géométrie algébrique* (EGA) developed in the 1950s. I highly recommend reading the introduction of [7]. Please get well seated during this section and don't expect to get all the ideas the first time you read it—that's okay!

Algebraic geometry is the study of solutions ("zeros") of systems of polynomials in several variables. Let R be a ring and consider the polynomial ring $R[\boldsymbol{X}]$ in variables

$\boldsymbol{X} \coloneqq (X_\lambda)_{\lambda \in \Lambda}$ indexed by some set Λ. Let $\Bbbk$ be an R-algebra (not necessarily—but often—a field). Recall from (1.97) that by the universal property of the polynomial ring $R[\boldsymbol{X}]$ we have for any point $\boldsymbol{x} \coloneqq (x_\lambda)_{\lambda \in \Lambda} \in \Bbbk^\Lambda$ an evaluation map

$$\mathrm{ev}_{\boldsymbol{x}} \colon R[\boldsymbol{X}] \to \Bbbk \; . \tag{2.20}$$

This is an R-algebra morphism and for a polynomial

$$f = \sum_\mu r_\mu \boldsymbol{X}^\mu \in R[\boldsymbol{X}] \tag{2.21}$$

we have

$$f(\boldsymbol{x}) = \mathrm{ev}_{\boldsymbol{x}}(f) = \sum_\mu r_\mu \boldsymbol{x}^\mu \in \Bbbk \; . \tag{2.22}$$

So, another way to state the universal property is that we have a natural bijection

$$\begin{array}{rcl} \Bbbk^\Lambda & \leftrightarrow & \mathrm{Hom}_{R\text{-}\mathsf{Alg}}(R[\boldsymbol{X}], \Bbbk) \\ \boldsymbol{x} & \mapsto & \mathrm{ev}_{\boldsymbol{x}} \\ (\varphi(X_\lambda)_{\lambda \in \Lambda}) & \mapsfrom & \varphi \; . \end{array} \tag{2.23}$$

Now, take a subset $S \subseteq R[\boldsymbol{X}]$. This is going to be our system of polynomials we want to study. We are interested in the common zeros of S over $\Bbbk$, i.e. points $\boldsymbol{x} \in \Bbbk^\Lambda$ such that $f(\boldsymbol{x}) = 0$ for all $f \in S$. We call

$$Z_S(\Bbbk) \coloneqq \{ \boldsymbol{x} \in \Bbbk^\Lambda \mid f(\boldsymbol{x}) = 0 \text{ for all } f \in S \} \subseteq \Bbbk^\Lambda \tag{2.24}$$

the set of **$\Bbbk$-points** of S. Subsets of $\Bbbk^\Lambda$ of this form are called **algebraic sets**. So, when varying $\Bbbk$, a system S of polynomials yields a map

$$Z_S \colon R\text{-}\mathsf{Alg} \to \mathsf{Set} \; . \tag{2.25}$$

As my notation already suggests, this is in fact a functor: if $\psi \colon \Bbbk \to \Bbbk'$ is an R-algebra morphism, then we get a map

$$\begin{array}{rcl} Z_S(\psi) \colon Z_S(\Bbbk) & \to & Z_S(\Bbbk') \\ \boldsymbol{x} & \mapsto & \psi(\boldsymbol{x}) \coloneqq (\psi(x_\lambda))_{\lambda \in \Lambda} \end{array} \tag{2.26}$$

because if $f(\boldsymbol{x}) = 0$, then also $f(\psi(\boldsymbol{x})) = \psi(f(\boldsymbol{x})) = 0$ since ψ is a morphism of R-algebras. We call this the **point functor** associated to the system S. Since $Z_{\{0\}}(\Bbbk) = \Bbbk^\Lambda$ for any R-algebra $\Bbbk$, this functor is called the (Λ-dimensional) **affine space** over R and is denoted by $\mathbb{A}^\Lambda_\Bbbk$. Any functor Z_S is a **subfunctor** (think about the definition) of $\mathbb{A}^\Lambda_\Bbbk$.

You are used to studying zeros of polynomials but maybe not in the generality that I have introduced. The big question you may have asked is: why on earth do we look at zeros in any possible R-algebra, i.e. why do we study a *functor*, not just the zeros over a *fixed* field? The answer is: only in this way you can really see the whole structure of S!

Example 2.4.1. Let $S \coloneqq \{X^2 - 2\} \subseteq \mathbb{Q}[X]$. Then $Z_S(\mathbb{Q}) = \emptyset$ but $Z_S(\mathbb{R}) = \{\pm\sqrt{2}\}$. If you would just consider the zero set over $\mathbb{Q}$, this would just look like the zeros of a trivial system $\{1\} \subseteq \mathbb{Q}[X]$. Only when you consider extensions, more structure appears!

This should convince that you the functor Z_S is the right thing to look at. Recall from (2.23) that we have a canonical bijection

$$\begin{array}{ccc} \Bbbk^\Lambda = Z_{\{0\}}(\Bbbk) & \overset{\simeq}{\to} & \mathrm{Hom}_{R\text{-}\mathsf{Alg}}(R[\boldsymbol{X}], \Bbbk) \\ \boldsymbol{x} & \mapsto & \mathrm{ev}_{\boldsymbol{x}} \; . \end{array} \tag{2.27}$$

This bijection is **functorial** in the following sense. Recall the Hom-functor

$$\mathrm{Hom}_{R\text{-}\mathsf{Alg}}(R[\boldsymbol{X}], -) \colon R\text{-}\mathsf{Alg} \to \mathsf{Set} \tag{2.28}$$

from Example 1.2.7. If $\psi \colon \Bbbk \to \Bbbk'$ is an R-algebra morphism, then we have a commutative diagram

$$\begin{tikzcd} Z_{\{0\}}(\Bbbk') \arrow[r, "\simeq"] & \mathrm{Hom}_{R\text{-}\mathsf{Alg}}(R[\boldsymbol{X}], \Bbbk') \\ Z_{\{0\}}(\Bbbk) \arrow[r, "\simeq"] \arrow[u, "Z_{\{0\}}(\psi)"] & \mathrm{Hom}_{R\text{-}\mathsf{Alg}}(R[\boldsymbol{X}], \Bbbk) \arrow[u, "\mathrm{Hom}_{R\text{-}\mathsf{Alg}}(R[\boldsymbol{X}],\psi)"'] \end{tikzcd} \tag{2.29}$$

Because of this, one also says that there is an **isomorphism** of functors

$$Z_{\{0\}}(-) \simeq \mathrm{Hom}_{R\text{-}\mathsf{Alg}}(R[\boldsymbol{X}], -) \; . \tag{2.30}$$

Now, it is a trivial but important observation that $Z_S = Z_{(S)}$, where (S) denotes the ideal in $R[\boldsymbol{X}]$ generated by S. Hence, without loss of generality we can assume that our system of polynomials forms an ideal $I \trianglelefteq R[\boldsymbol{X}]$. We then have the following commutative diagram:

$$\begin{tikzcd} \Bbbk^\Lambda = Z_{\{0\}}(\Bbbk) \arrow[r, "\simeq"] & \mathrm{Hom}_{R\text{-}\mathsf{Alg}}(R[\boldsymbol{X}], \Bbbk) \\ Z_I(\Bbbk) \arrow[r, "\simeq"] \arrow[u, hook] & \{\varphi \in \mathrm{Hom}_{R\text{-}\mathsf{Alg}}(R[\boldsymbol{X}], \Bbbk) \mid I \subseteq \mathrm{Ker}(\varphi)\} \arrow[u, hook] \\ & \mathrm{Hom}_{R\text{-}\mathsf{Alg}}(R[\boldsymbol{X}]/I, \Bbbk) \arrow[u, "\simeq"] \end{tikzcd} \tag{2.31}$$

We thus have an isomorphism of functors

$$Z_I(-) \simeq \mathrm{Hom}_{R\text{-}\mathsf{Alg}}(R[\boldsymbol{X}]/I, -) \; . \tag{2.32}$$

Using the Hom-functor is simply a more fancy and algebraic way to look at the point functor. But now recall from Lemma 1.4.8 that any R-algebra A is (non-canonically) isomorphic to a quotient $R[\boldsymbol{X}]/I$. It thus makes sense to associate to any R-algebra A a point functor

$$Z_A(-) \coloneqq \mathrm{Hom}_{R\text{-}\mathsf{Alg}}(A, -) \; . \tag{2.33}$$

This is very similar to the notion of an abstract manifold and choosing an embedding into some affine space. You can think of the functor Z_A as encoding the abstract "geometry" of A. You can make it concrete by choosing an isomorphism $A \simeq R[\boldsymbol{X}]/I$. We call a map $\varphi \in Z_A(\Bbbk) = \mathrm{Hom}_{R\text{-}\mathsf{Alg}}(A, \Bbbk)$ a **$\Bbbk$-point** of A.

Recall from Example 1.2.7 the Hom-functor $\mathrm{Hom}_{\mathcal{C}}(X, -)$ for an object X in a category $\mathcal{C}$. It is a standard fact in category theory—which follows from the so-called **Yoneda lemma**—that this functor completely determines the object X up to isomorphism. Let's just take this as a fact here (the proof is not difficult) and conclude that the R-algebra A is completely determined by the functor Z_A, i.e.

$$A \simeq B \text{ as } R\text{-algebras} \iff Z_A \simeq Z_B \text{ as functors} \; . \tag{2.34}$$

So, the "geometry" of A encoded by Z_A "sees" everything of A.

Example 2.4.2. The ideals $I \coloneqq (X)$ and $J \coloneqq (X^2)$ in $\mathbb{Q}[X]$ are distinct. Hence the associated functors Z_I and Z_J should be distinct as well. We have

$$Z_I(\mathbb{Q}) = \{0\} = Z_J(\mathbb{Q}) \ .$$

But now consider solutions over $\Bbbk \coloneqq \mathbb{Q}[X]/(X^2)$. Then indeed

$$Z_I(\Bbbk) = \{0\} \quad \text{but} \quad Z_J(\Bbbk) = \{\alpha X \mid \alpha \in \mathbb{Q}\} \ .$$

If in Example 2.4.2 we only consider points in fields $\Bbbk$, we cannot distinguish Z_I and Z_J—we need zero-divisors for this. Nevertheless, looking just at points in fields is still the most natural thing to do in the beginning. So, fix an R-algebra A. We consider the set of all field-valued points of A and call this the set of **places**[23] of A, i.e.

$$\mathrm{Pl}(A) \coloneqq \coprod_{\substack{\Bbbk \text{ an } R\text{-algebra} \\ \text{which is a field}}} Z_A(\Bbbk) \ . \tag{2.35}$$

If you have a $\Bbbk$-point and consider it in a bigger field $\Bbbk'$, then it's still the "same" point. To take care of this, we define an equivalence relation on $\mathrm{Pl}(A)$ as follows. Let $\varphi \in Z_A(\Bbbk) = \mathrm{Hom}_{R\text{-}\mathsf{Alg}}(A, \Bbbk)$ and $\varphi' \in Z_A(\Bbbk') = \mathrm{Hom}_{R\text{-}\mathsf{Alg}}(A, \Bbbk')$ be two points of A with values in some R-algebras $\Bbbk$ and $\Bbbk'$ which are fields. Then we write $\varphi \sim \varphi'$ if there is a commutative diagram

$$\begin{tikzcd}
& \Bbbk'' & \\
\Bbbk \arrow[ur] & & \Bbbk' \arrow[ul] \\
& A \arrow[ul, "\varphi"] \arrow[ur, "\varphi'"'] &
\end{tikzcd} \tag{2.36}$$

of morphisms of R-algebras, where $\Bbbk''$ is some R-algebra which is a field.

Now, for any $\varphi \in Z_A(\Bbbk) = \mathrm{Hom}_{R\text{-}\mathsf{Alg}}(A, \Bbbk)$ the kernel

$$P_\varphi \coloneqq \mathrm{Ker}(\varphi) \trianglelefteq A \tag{2.37}$$

is a *prime* ideal in A by Corollary 2.2.5. If $\varphi \sim \varphi'$, then obviously $P_\varphi = P_{\varphi'}$. Hence, we get a well-defined map

$$\mathrm{Pl}(A)/\!\sim\, \to \mathrm{Spec}(A) \ . \tag{2.38}$$

This gives us a connection between field-valued points and prime ideals! This is in fact a bijection. For every prime ideal $P \in \mathrm{Spec}(A)$ the quotient A/P is an integral domain and you'll learn in Section 4.1 that you can artificially add fractions to an integral domain to get a field $k_A(P) \supseteq A/P$. This is like passing from $\mathbb{Z}$ to $\mathbb{Q}$. So, to P you can associate the $k_A(P)$-valued point $\varphi_P \colon A \to A/P \to k_A(P)$, and this is an inverse to (2.38). We thus have a bijection between the prime ideal spectrum and the field valued points of A (or of the polynomial system defined by a presentation of A) up to equivalence. Nice! All the field-valued points of A have been compressed

[2] This is not official terminology. In [7] these things are called *geometric points* but this is more commonly used for points with values in an algebraically closed field.

[3] Note that there is a little set-theoretic issue in this definition since the collection of all R-algebras does not form a set. But we can still work with this as a so-called *class* and can basically work with this like a set.

into the spectrum. The prime spectrum *is* geometry!

Notice the following: if $\varphi\colon A \to \Bbbk$ is a $\Bbbk$-valued point of A and if $\sigma\colon \Bbbk \to \Bbbk$ is an R-algebra automorphism (a **Galois automorphism** of $\Bbbk$ over R), then we have a commutative diagram

$$\begin{tikzcd}
 & \Bbbk & \\
\Bbbk \arrow[ur, "\sigma"] & & \Bbbk \arrow[ul, "\mathrm{id}"'] \\
 & A \arrow[ul, "\varphi"] \arrow[ur, "\sigma\circ\varphi"'] &
\end{tikzcd} \tag{2.39}$$

This means, the points φ and $\sigma\circ\varphi$ are equivalent. In other words, the prime spectrum identifies points in the same **Galois orbit**.

Here's something strange: in $\mathrm{Spec}(A)$ we can have an inclusion $P \subseteq Q$ of prime ideals. Using the bijection (2.38) we have just constructed, this means we have an "inclusion" of field-valued points of A. This doesn't really make sense intuitively if you think about these points like usual points in $\mathbb{R}^n$. But it does! Let's make things concrete and look at ideals $J \subseteq I$ in $R[\boldsymbol{X}]$. We then have $J \subseteq I$ and this means we have an inclusion

$$Z_I(\Bbbk) \subseteq Z_J(\Bbbk) \tag{2.40}$$

because J, viewed of as a polynomial system, is smaller than I, so there are less constraints, so the zero set gets bigger. We also say that J is more **generic** than I and that I is more **special** than J. We used this already in (2.31) in the special case $\{0\} \subseteq I$. Now, if P is a prime ideal of $R[\boldsymbol{X}]$ containing J, this means it corresponds to a field-valued point of the system J. If P contains I, this means it is also a point of the more special system I. Hence, an inclusion of prime ideals $P \subseteq Q$ means that Q also corresponds to a field-valued point of a more special system than P does. We thus say that Q is more **special** than P and that P is more **generic** than Q. A maximal ideal is as special as it can get and a minimal prime ideal is as generic as it can get—in case the ring is an integral domain, the zero ideal is the unique minimal prime ideal.

So, to wrap up:

(1) Studying the field-valued points of a system $I \subseteq R[\boldsymbol{X}]$ of polynomials amounts to studying the prime ideals of $A := R[\boldsymbol{X}]/I$, or, equivalently, the prime ideals of $R[\boldsymbol{X}]$ containing I.
(2) An abstract algebra A gets such a geometric interpretation by choosing an isomorphism $A \simeq R[\boldsymbol{X}]/I$.

All this takes a bit of time and many examples to get used to.

Example 2.4.3. A typical situation one often encounters is the case where R is a field and $\Bbbk = R$. In this case any $\Bbbk$-valued point $\varphi \in Z_A(\Bbbk) = \mathrm{Hom}_{\Bbbk\text{-}\mathsf{Alg}}(A, \Bbbk)$ of A is already surjective and it follows from Lemma 2.3.10 that the associated prime ideal $P_\varphi \in \mathrm{Spec}(A)$ is maximal. We can describe this ideal more explicitly. To this end, we choose an isomorphism $A \simeq \Bbbk[\boldsymbol{X}]/I$ and view ideals in A as ideals in $\Bbbk[\boldsymbol{X}]$ containing I as usual. Then φ is the morphism induced by $\mathrm{ev}_{\boldsymbol{x}}$ for a point $\boldsymbol{x} \in Z_I(\Bbbk)$

and we are interested in

$$P_{\boldsymbol{x}} := P_{\mathrm{ev}_{\boldsymbol{x}}} = \mathrm{Ker}(\mathrm{ev}_{\boldsymbol{x}}) \trianglelefteq \Bbbk[\boldsymbol{X}] \ . \tag{2.41}$$

I claim that

$$P_{\boldsymbol{x}} = (\{X_\lambda - x_\lambda \mid \lambda \in \Lambda\}) \trianglelefteq \Bbbk[\boldsymbol{X}] \ . \tag{2.42}$$

Let J be the ideal on the right hand side of (2.42). It is clear that the evaluation of the polynomial $X_\lambda - x_\lambda$ in $\boldsymbol{x}$ is zero, so J is contained in $P_{\boldsymbol{x}}$. But you can easily see that also $\Bbbk[\boldsymbol{X}]/J \simeq \Bbbk$, so J is maximal as well, forcing $J = P_{\boldsymbol{x}}$.

Note that since we assume $\Bbbk = R$, there are no non-trivial Galois automorphisms of $\Bbbk$ over R. Mapping $\boldsymbol{x}$ to $P_{\boldsymbol{x}}$ thus yields an injection

$$Z_A(\Bbbk) \hookrightarrow \mathrm{Max}(A) \ . \tag{2.43}$$

As we will see in Example 2.4.4, this is in general not a bijection, i.e. there are maximal ideals in A not of the form (2.42).

Example 2.4.4. Suppose you want to study the (field-valued) points of the polynomial $X^2 - 2 \in \mathbb{Q}[X]$. We now know this means finding prime ideals in $A := \mathbb{Q}[X]/(X^2-2)$, or, equivalently, prime ideals in $\mathbb{Q}[X]$ containing $X^2 - 2$. We have an isomorphism

$$\mathbb{Q}[X]/(X^2-2) \simeq \mathbb{Q}(\sqrt{2}) \tag{2.44}$$

by sending X to $\sqrt{2}$. Because the right hand side is a field, this means that (X^2-2) is a maximal ideal in $\mathbb{Q}[X]$ and that this is the only prime ideal containing $X^2 - 2$. Hence, there is just one field-valued point, described by the maximal ideal (X^2-2), which is the kernel of the point $\mathbb{Q}[X] \to \mathbb{Q}(\sqrt{2})$ mapping X to $\sqrt{2}$.

Now you should complain: there are two zeros of X^2-2, namely $+\sqrt{2}$ and $-\sqrt{2}$, why do we see just one? Recall that we noticed above that the prime spectrum identifies points on the same Galois orbit, and $\sqrt{2}$ and $-\sqrt{2}$ are Galois conjugate! So, everything makes perfect sense. Also note that (X^2-2) is a maximal ideal which is not of the form $P_{\boldsymbol{x}}$ as in (2.42).

Example 2.4.5. There are typical settings where the map $Z_A(\Bbbk) \hookrightarrow \mathrm{Max}(A)$ from (2.43) actually *is* a bijection. Let's consider $A = \mathbb{C}[X]$. Since A is a principal ideal domain, we know from Corollary 2.1.5 that the prime ideals of A are either the zero ideal or the ideals (f) where f is an irreducible polynomial. From Example 2.3.4 we know that the latter are already maximal. Now, since $\mathbb{C}$ is algebraically closed, the irreducible polynomials are all of the form $X - x$ for $x \in \mathbb{C}$. Hence, we indeed have

$$\mathbb{C} = Z_A(\mathbb{C}) \simeq \mathrm{Max}(A) \ . \tag{2.45}$$

We will (much) later—in Corollary 6.1.8—prove that $Z_A(\Bbbk) \hookrightarrow \mathrm{Max}(A)$ is a bijection for any finitely generated algebra A over an algebraically closed field $\Bbbk$. This is the famous "Nullstellensatz" and this fact is the reason why classically in algebraic geometry only maximal ideals were considered.

Example 2.4.6. Recall Example 2.4.2. We looked at the polynomials X and X^2 in $\mathbb{Q}[X]$. We noticed that we cannot distinguish the associated functors $Z_{\{X^2\}}$ and $Z_{\{X\}}$ by just looking at field-valued points. Let's confirm this from the prime spectrum point of view. We have $\mathbb{Q}[X]/(X) \simeq \mathbb{Q}$, so

$$\mathrm{Spec}(\mathbb{Q}[X]/(X)) = \{(0)\} \tag{2.46}$$

consists of just one point (which makes sense). In Example 2.2.4 we already determined that

$$\operatorname{Spec}(\mathbb{Q}[X]/(X^2)) = \{(X)\}\ . \tag{2.47}$$

Hence, as sets, the prime spectra of $\mathbb{Q}[X]/(X)$ and $\mathbb{Q}[X]/(X^2)$ look the same!

Remark 2.4.7. If you really want a perfect dictionary between algebra and geometry, i.e. if you want to recover A from the associated geometry like the functor Z_A does, then the conclusion from Example 2.4.6 is that the prime spectrum is just not enough–you need additional information. What you do in algebraic geometry is to additionally put the ring of "polynomial functions" defined by A on top of the spectrum $\operatorname{Spec}(A)$, which basically means you just record A itself and just view its elements more geometrically. There is no other way if you want to recover A.

Example 2.4.8. Suppose you want to study the (field-valued) points of the polynomial $X_1X_2 \in \mathbb{Q}[X_1, X_2]$. This means finding the prime ideals in the ring $A \coloneqq \mathbb{Q}[X_1, X_2]/(X_1X_2)$, or, equivalently, the prime ideals in $\mathbb{Q}[X_1, X_2]$ containing (X_1X_2). Clearly, (X_1) and (X_2) are two such prime ideals since these ideals obviously contain (X_1X_2), and the quotients

$$\mathbb{Q}[X_1, X_2]/(X_1) \simeq \mathbb{Q}[X_2] \quad \text{and} \quad \mathbb{Q}[X_1, X_2]/(X_2) \simeq \mathbb{Q}[X_1] \tag{2.48}$$

are integral domains, so (X_1) and (X_2) are prime ideals. Now, let P be any prime ideal containing (X_1X_2). Then $X_1X_2 \in P$. Since P is prime, we must have $X_1 \in P$ or $X_2 \in P$. Hence, $(X_1) \subseteq P$ or $(X_2) \subseteq P$. This means (X_1) and (X_2) are the unique minimal prime ideals in A. We can thus split the hunt for all the prime ideals in A into two cases: those containing (X_1) and those containing (X_2). In the first case, the prime ideals correspond to prime ideals in

$$\mathbb{Q}[X_1, X_2]/(X_1) \simeq \mathbb{Q}[X_2]\ . \tag{2.49}$$

This is a polynomial ring in one variable over a field, hence a principal ideal domain by Example 1.5.11, and we know from Corollary 2.1.5 that the prime ideals are of the form (f) where f is zero or an irreducible polynomial in $\mathbb{Q}[X_2]$. So, the prime ideals in $\mathbb{Q}[X_1, X_2]$ containing (X_1) are of the form (X_1, f) for $f \in \mathbb{Q}[X_2]$ either zero or irreducible. The prime ideals containing X_2 have a similar description. So for example, we have prime ideals

$$(X_1, X_2 - x_2) \quad \text{and} \quad (X_1 - x_1, X_2) \tag{2.50}$$

for all $x_1, x_2 \in \mathbb{Q}$. But we also have prime ideals like

$$(X_1, X_2^2 - 2)\ . \tag{2.51}$$

What does all this mean? You should always try to draw a real picture of the situation. What does $Z_{\{X_1X_2\}}(\mathbb{R}) \subseteq \mathbb{R}^2$ look like? It's the zero set of the polynomial X_1X_2, so it consists of all points $(x_1, x_2) \in \mathbb{R}^2$ such that $x_1 = 0$ or $x_2 = 0$. Hence, $Z_{\{X_1X_2\}}(\mathbb{R})$ is the union of the two coordinate axes in $\mathbb{R}^2$. We thus expect to have points of the form $(0, x_2)$ and $(x_1, 0)$. This is precisely what the prime ideals (2.50) describe! But note that these describe $\mathbb{Q}$-valued points only. The prime spectrum looks at general field-valued points, so you'd also expect points like $(0, \sqrt{2})$. And this is exactly what you get from (2.51). It's all there!

What about the prime ideals (X_1) and (X_2)? They are more generic than those in (2.50). They describe the lines $\{X_1 = 0\}$ and $\{X_2 = 0\}$ in their *entirety*! What are the corresponding points? Associated to (X_1) is the quotient map $\mathbb{Q}[X_1, X_2] \to$

$\mathbb{Q}[X_1, X_2]/(X_1) \simeq \mathbb{Q}[X_2]$. Now, recall that we can artificially add fractions (more on this later) to produce a field

$$\mathbb{Q}(X_2) \coloneqq \left\{ \frac{f}{g} \mid f, g \in \mathbb{Q}[X_2], g \neq 0 \right\} \tag{2.52}$$

containing $\mathbb{Q}[X_2]$. So, to (X_1) corresponds the field-valued point

$$\mathbb{Q}[X_1, X_2] \to \mathbb{Q}(X_2) \ . \tag{2.53}$$

And indeed this point describes the zero $(0, X_2)$ of the polynomial $X_1 X_2$ in the huge field $\mathbb{Q}(X_2)$! It's a zero you would normally not look at—but prime ideals see them! The prime spectrum describes what the zero set of $X_1 X_2$ is really made of—from the largest to the smallest scales!

Example 2.4.9. The prime spectrum of $\mathbb{Z}$ consists of the maximal ideals (p) for p a prime number and the zero ideal. The corresponding field-valued points are the maps $\mathbb{Z} \twoheadrightarrow \mathbb{Z}/p\mathbb{Z} = \mathbb{F}_p$ and $\mathbb{Z} \hookrightarrow \mathbb{Q}$. It's a bit strange viewing this geometrically. Note that in contrast to the examples above, we have here points taking values in fields of **mixed characteristic**. This happens when you have $\mathbb{Z}$ as base ring.

Example 2.4.10. [4] Suppose you want to study the (field-valued) points of $X^2 - 4 \in \mathbb{Z}[X]$. This means finding the prime ideals in $A \coloneqq \mathbb{Z}[X]/(X^2 - 4)$. Let P be such a prime ideal. Let $\varphi \colon \mathbb{Z} \to \mathbb{Z}[X]/(X^2 - 4)$ be the natural map. Then $\varphi^{-1}(P) \trianglelefteq \mathbb{Z}$ is a prime ideal, hence of the form (p) for a prime number p or for $p = 0$. We have $p \in P$, so P corresponds to a prime ideal in the quotient $A/(p) \simeq \mathbb{F}_p[X]/(X^2 - 4)$. We distinguish three cases:

(1) $p = 2$. Then $A/(2) \simeq \mathbb{F}_2[X]/(X^2)$ and from Example 2.4.6 (in which $\mathbb{Q}$ was the base field but the argument works for any base field) we know that this quotient has just one prime ideal, namely (X). Hence, we must have $P = (2, X)$.

(2) $p > 2$. The prime ideals in $A/(p) \simeq \mathbb{F}_p[X]/(X^2 - 4)$ are the prime ideals in $\mathbb{F}_p[X]$ containing $X^2 - 4 = (X-2)(X+2)$. Hence, such prime ideals must contain $X - 2$ or $X + 2$. Since $\mathbb{F}_p[X]$ is a principal ideal domain, all non-zero prime ideals are already maximal, hence there are just two such prime ideals, namely $(X-2)$ and $(X+2)$. In conclusion, P is either $(p, X-2)$ or $(p, X+2)$.

(3) $p = 0$. Then P corresponds to a prime ideal in $\mathbb{Z}[X]$ containing $X^2 - 4 = (X+2)(X-2)$. So, P contains $X + 2$ or $X - 2$. Hence, P corresponds to a prime ideal in $\mathbb{Z}[X]/(X-2)$ or in $\mathbb{Z}[X]/(X+2)$ whose intersection with $\mathbb{Z}$ is (0). Since $\mathbb{Z}[X]/(X-2) \simeq \mathbb{Z}$ and similarly $\mathbb{Z}[X]/(X+2) \simeq \mathbb{Z}$, there is just one such prime ideal, namely the zero ideal. Hence, $P = (X+2)$ or $P = (X-2)$.

Notice how all these prime ideals describe field-valued solutions of $X^2 - 4$: we have just one solution (with multiplicity 2) over a field of characteristic $p = 2$, and we have two solutions over a field of characteristic $p > 2$, and two solutions over a field of characteristic $p = 0$.

[4] This and some more examples are discussed in `https://stacks.math.columbia.edu/tag/00EX`.

Figure 2.1 shows you how to visualize the spectrum. Why did I draw this one very fat point?[5]

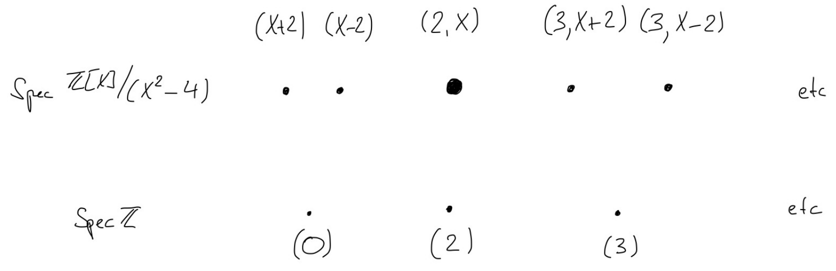

FIGURE 2.1.

Exercises.

EXERCISE 2.4.11. Let K be a field. Describe $\operatorname{Spec}(K[X_1,X_2]/(X_1^2-X_2^2))$.

2.5. Zariski topology

We'll do something crazy now. I'm not sure whether you're familiar with topology, so here's the deal. Topology is an abstract setting for talking about continuous maps. In $\mathbb{R}^n$ we can consider around any point x and for any $\varepsilon \in \mathbb{R}_{>0}$ the set

$$U_\varepsilon(x) \coloneqq \{ y \in \mathbb{R}^n \mid \|x-y\| < \varepsilon \} \ , \tag{2.54}$$

where $\|\cdot\|$ denotes the euclidean distance. Such sets are called **basic open** subsets of $\mathbb{R}^n$. An arbitrary union of basic open subsets is called **open**.

LEMMA 2.5.1. *A subset $U \subseteq \mathbb{R}^n$ is open if and only if for each $x \in U$ there is $\varepsilon > 0$ such that $U_\varepsilon(x) \subseteq U$.*

PROOF. Suppose that U is open. By definition, there is $y \in U$ and $\delta > 0$ such that $x \in U_\delta(y) \subseteq U$. Set $\varepsilon \coloneqq \delta - \|x-y\| > 0$. If $z \in U_\varepsilon(x)$, then

$$\|z-y\| = \|z-x+x-y\| \leq \|z-x\| + \|x-y\| < \varepsilon + \|x-y\| = \delta \ , \tag{2.55}$$

hence $U_\varepsilon(x) \subseteq U_\delta(y) \subseteq U$. Conversely, if for each $x \in U$ there is $\varepsilon_x > 0$ such that $U_{\varepsilon_x}(x) \subseteq U$, then $U = \bigcup_{x \in U} U_{\varepsilon_x}(x)$ is open by definition. □

Now, we make the following observation.

LEMMA 2.5.2.

(1) *$\emptyset$ and $\mathbb{R}^n$ are open.*
(2) *Arbitrary unions of open subsets are open.*
(3) Finite *intersections of open subsets are open.*

PROOF. $\mathbb{R}^n$ is open because $\mathbb{R}^n = \bigcup_{\varepsilon > 0} U_\varepsilon(0)$, and $\emptyset \subseteq \mathbb{R}^n$ is open because it is the empty union. Unions of open subsets are open by definition. For the last claim, it is enough to show this for two open subsets U_1 and U_2. Let $x \in U_1 \cap U_2$. Then there is $\varepsilon_1, \varepsilon_2 > 0$ such that $U_{\varepsilon_1}(x) \subseteq U_1$ and $U_{\varepsilon_2}(x) \subseteq U_2$. Set $\varepsilon \coloneqq \min\{\varepsilon_1, \varepsilon_2\}$. Then $U_\varepsilon(x) \subseteq U_1 \cap U_2$, hence $U_1 \cap U_2$ is open. □

[5] If you like pictures, you should already check out `https://ulthiel.com/math/wp-content/uploads/other/Spec-Zx.pdf`.

Recall that a map $f\colon X \to \mathbb{R}^m$ from an open subset $X \subseteq \mathbb{R}^n$ is continuous in $x \in X$ if for each $\varepsilon > 0$ there is a $\delta > 0$ such that for all $y \in X$ with $\|y - x\| < \delta$ we have $\|f(x) - f(y)\| < \varepsilon$. We can rephrase this in terms of open subsets.

Lemma 2.5.3. *The map $f\colon X \to \mathbb{R}^m$ is continuous in $x \in X$ if and only if for any open subset $V \subseteq \mathbb{R}^m$ containing $f(x)$ there is an open subset U of X containing x such that $f(U) \subseteq V$.*

Proof. This is actually quite obvious but let's prove it anyways. Suppose that f is continuous in x. Let $V \subseteq \mathbb{R}^m$ be open and containing $f(x)$. Since V is open, there is $\varepsilon > 0$ such that $U_\varepsilon(f(x)) \subseteq V$. Since f is continuous, there is $\delta > 0$ such that $\|f(x) - f(y)\| < \varepsilon$ if $\|x - y\| < \delta$. Hence, $f(U_\delta(x)) \subseteq U_\varepsilon(f(x)) \subseteq V$.

Conversely, suppose the other condition holds. Let $\varepsilon > 0$. Then $V \coloneqq U_\varepsilon(f(x))$ is an open subset containing $f(x)$. Hence, there is an open subset U of X containing x such that $f(U) \subseteq V$. Since U is open, there is $\delta > 0$ with $U_\delta(x) \subseteq U$. Then $f(U_\delta(x)) \subseteq V = U_\varepsilon(f(x))$, hence $\|f(x) - f(y)\| < \varepsilon$ if $\|x - y\| < \delta$. □

So far, nothing really new. But let's take Lemma 2.5.2 and Lemma 2.5.3 and make them abstract.

Definition 2.5.4. A **topology** on a set X is a collection of subsets of X, called **open** subsets, satisfying the following properties:

(1) $\emptyset$ and X are open.
(2) The union of open subsets is open.
(3) *Finite* intersections of open subsets are open.

A set equipped with a topology is called a **topological space**.

Definition 2.5.5. A map $f\colon X \to Y$ of topological spaces is **continuous** in $x \in X$ if for any open subset $V \subseteq Y$ containing $f(x)$ there is an open subset $U \subseteq X$ containing x such that $f(U) \subseteq V$. The map is **continuous** if it is continuous in every point.

Notice that the composition of continuous maps is continuous, hence we get a category Top of topological spaces.

Lemma 2.5.6. *A map $f\colon X \to Y$ between topological spaces is continuous if and only if for any open subset $V \subseteq Y$ the preimage $f^{-1}(V) \subseteq X$ is open.*

Proof. Left for you as Exercise 2.5.18. □

Example 2.5.7. If X is a topological space, we can equip any subset $A \subseteq X$ with a topology by defining the open subsets of A to be the sets $A \cap U$ for $U \subseteq X$ open. This is called the **subspace topology** on A. If nothing else is mentioned, subsets are always considered with the subspace topology. Note that if $U \subseteq X$ is open, then the open subsets of U in the subspace topology are precisely the open subsets of X contained in U.

Example 2.5.8. The set $\mathbb{R}^n$ with the notion of open subsets defined above (unions of the basic open subsets) is a topological space. A map $f\colon X \to \mathbb{R}^m$ from an open subset $X \subseteq \mathbb{R}^n$ is continuous in the usual ε-δ-sense if and only if it is continuous as a map of topological spaces as in Definition 2.5.5 when equipping X with the subspace topology. Similarly, you can equip any metric space with a topology, called **metric topology**, and the ε-δ-concept of continuity is the same as the topological one.

The thing with topological spaces is that you can do crazy things now and talk about continuity without a metric or anything similar.

Example 2.5.9. *Any* set X can be equipped with a topology by taking as open subsets *all* subsets of X. This is the so-called **discrete topology**. In this case, *any* map $f\colon X \to Y$ to a topological space is continuous.

Example 2.5.10. *Any* set X can be equipped with the **trivial topology** just consisting of the open sets $\emptyset$ and X. Note that this is crazy because we cannot separate two points by open subsets anymore, i.e. given x, y we cannot find disjoint open subsets U and V with $x \in U$ and $y \in V$. Spaces satisfying this separation property are called **Hausdorff** and this is the least you would expect of a topological space when resembling anything of what you intuitively think of as "space".

Example 2.5.11. Consider a 2-element set $X \coloneqq \{0, 1\}$. This can be equipped with a topology by defining the sets $\emptyset, \{1\}$, and X to be open. This is the smallest example of a topology which is neither trivial nor discrete.

We were talking about the prime spectrum, so what's the point of all this here? Maybe you can guess that we're going to define a topology on the spectrum! There's no metric or anything, so where does the topology come from? Let's first talk about an equivalent definition of a topology. A subset Z of a topological space X is called **closed** if its complement $X \setminus Z$ is open. Then using de Morgan's laws, the properties of open sets in Definition 2.5.4 become the following properties of closed sets:

(1) $\emptyset$ and X are closed.
(2) The intersection of closed sets is closed.
(3) *Finite* unions of closed sets are closed.

You can thus equivalently describe a topology in terms of its closed sets satisfying the above properties—and sometimes this approach is more convenient. Using Lemma 2.5.6, also the notion of continuity can be rephrased in terms of closed sets: a map $f\colon X \to Y$ between topological spaces is continuous if and only if the preimage of closed sets is closed.

Remark 2.5.12. "Sets are not doors"[6]: it is not necessarily true that a subset of a topological space is either open or closed. For example, $A \coloneqq [0, 1) \subseteq \mathbb{R}$ is not open since $0 \in A$ does not have an open neighborhood contained in A. But A is also not closed: if it would be closed, then $B \coloneqq \mathbb{R} \setminus A = (-\infty, 0) \cup [1, \infty)$ would be open but again, $1 \in B$ does not have a neighborhood contained in B.

Remark 2.5.13. Let X be a topological space and let $A \subseteq X$. Recall from Example 2.5.7 that the open subsets in the subspace topology on A are of the form $A \cap U$ with $U \subseteq X$ open. We have

$$A \setminus (A \cap U) = A \setminus U = (X \setminus U) \cap A\,.$$

Since $X \setminus U$ is closed in X, it thus follows that the closed subsets of A are of the form $A \cap Z$ for $Z \subseteq X$ closed.

We still didn't talk about the prime spectrum, so let's come to this now. Consider a polynomial $f \in \mathbb{R}[X_1, \ldots, X_n]$. Evaluation of f in points of $\mathbb{R}^n$ defines a continuous map $f\colon \mathbb{R}^n \to \mathbb{R}$. It is not hard to see that the complement of a point $\{x\}$ in $\mathbb{R}^n$ is an open subset, so a point $\{x\}$ is closed. In particular, $\{0\}$ is closed, hence the zero

[6] I've read that this idiom is due to James Munkres.

set $Z_{\{f\}}(\mathbb{R}^n) = f^{-1}(0)$ of f is a closed subset of $\mathbb{R}^n$. More generally, if you have a system S of polynomials, then

$$Z_S(\mathbb{R}^n) = \bigcap_{f \in S} Z_{\{f\}}(\mathbb{R}^n)$$

is closed.

Now, recall that the spectrum $\mathrm{Spec}(A)$ of a ring A can be considered as a zero set of a system of polynomials after choosing a presentation of A. Given a subset S of A, the prime ideals of A containing S can be considered as field-valued solutions of the polynomial system defined by S. This bring us to the (crazy?) idea of considering the sets

$$\mathrm{V}(S) \coloneqq \{P \in \mathrm{Spec}(A) \mid S \subseteq P\} \tag{2.56}$$

for subsets $S \subseteq A$ as the closed sets of a topology on $\mathrm{Spec}(A)$. The letter V here stands for **variety**, which is a more fancy word for "zero set". We just need to check whether they satisfy all the necessary properties.

We have $\mathrm{V}(\{0\}) = \mathrm{Spec}(A)$ and $\mathrm{V}(A) = \emptyset$. We obviously have $\mathrm{V}(S) = \mathrm{V}((S))$, the latter being the variety of the ideal (S) generated by S. It is thus sufficient to consider sets $\mathrm{V}(I)$ where I is an ideal. For a family $(I_\lambda)_{\lambda \in \Lambda}$ of ideals we claim that

$$\mathrm{V}\left(\bigcup_{\lambda \in \Lambda} I_\lambda\right) = \bigcap_{\lambda \in \Lambda} \mathrm{V}(I_\lambda)\ , \tag{2.57}$$

which implies that intersections of $\mathrm{V}(I)$ are closed. Indeed,

$$P \in \mathrm{V}\left(\bigcup_\lambda I_\lambda\right) \Leftrightarrow \bigcup_\lambda I_\lambda \subseteq P \Leftrightarrow I_\lambda \subseteq P\ \forall \lambda \Leftrightarrow P \in \mathrm{V}(I_\lambda)\ \forall \lambda \Leftrightarrow P \in \bigcap_\lambda \mathrm{V}(I_\lambda)\ .$$

Moreover, we claim that

$$\mathrm{V}(I \cap J) = \mathrm{V}(IJ) = \mathrm{V}(I) \cup \mathrm{V}(J)\ , \tag{2.58}$$

which implies in particular that finite unions of the $\mathrm{V}(I)$ are closed. The equality $\mathrm{V}(I \cap J) = \mathrm{V}(I) \cup \mathrm{V}(J)$ is clear and the other equality follows from

$$P \in \mathrm{V}(IJ) \Leftrightarrow IJ \subseteq P \Leftrightarrow I \subseteq P \text{ or } J \subseteq P \text{ (since } P \text{ is prime)} \Leftrightarrow P \in \mathrm{V}(I) \cup \mathrm{V}(J)\ .$$

We conclude:

Corollary 2.5.14. *The sets* $\mathrm{V}(S)$ *for* $S \subseteq A$ *define the closed sets of a topology on* $\mathrm{Spec}(A)$*, called the* ***Zariski topology***.

We will abbreviate

$$\mathrm{V}(f_1, \ldots, f_r) \coloneqq \mathrm{V}((f_1, \ldots, f_r)) \tag{2.59}$$

for elements $f_i \in A$. The prime spectrum is functorial in the topological category:

Lemma 2.5.15. *If* $f \colon A \to B$ *is a ring morphism, then* $f^* \colon \mathrm{Spec}(B) \to \mathrm{Spec}(A)$ *is continuous. Hence,* Spec *defines a contravariant functor* $\mathsf{Ring} \to \mathsf{Top}$.

Proof. We claim that

$$(f^*)^{-1}(\mathrm{V}(I)) = \mathrm{V}(f(I))\ , \tag{2.60}$$

which shows that preimages of closed sets are closed, hence f^* is continuous. Indeed, we have

$$\begin{aligned}(f^*)^{-1}(\mathrm{V}(I)) &= \{J \in \operatorname{Spec}(B) \mid f^*(J) \in \mathrm{V}(I)\}\\ &= \{J \in \operatorname{Spec}(B) \mid f^{-1}(J) \in \mathrm{V}(I)\}\\ &= \{J \in \operatorname{Spec}(B) \mid I \subseteq f^{-1}(J)\}\\ &= \{J \in \operatorname{Spec}(B) \mid f(I) \subseteq J\}\\ &= \mathrm{V}(f(I))\ . \end{aligned} \qquad \square$$

The Zariski topology is extremely weird—it's nothing like what you're used to.

EXAMPLE 2.5.16. Let's consider Spec($\mathbb{Z}$). By definition, the closed subsets are of the form $\mathrm{V}(I)$ for $I \trianglelefteq \mathbb{Z}$ an ideal. Any ideal is of the form (n) for a natural number n. We have

$$\mathrm{V}(0) = \operatorname{Spec}(\mathbb{Z}) \quad \text{and} \quad \mathrm{V}(1) = \emptyset\ . \tag{2.61}$$

If $n \neq 0$ and $n = p_1 \cdots p_r$ is the prime factorization of n, then by (2.58) we have

$$\mathrm{V}(n) = \mathrm{V}(p_1 \cdots p_r) = \bigcup_{i=1}^{r} \mathrm{V}(p_i) = \{(p_1), \ldots, (p_r)\}\ , \tag{2.62}$$

where $\mathrm{V}(p_i) = \{(p_i)\}$ since (p_i) is a maximal ideal. We thus know all the closed subsets explicitly. But here's something strange: the point $\{(0)\}$ is not among these sets—it's *not* closed! In fact, its closure—i.e., the smallest closed subset containing it—is equal to the whole space! How crazy is this?

EXAMPLE 2.5.17. With the same arguments as in Example 2.5.16 we can describe the spectrum of any principal ideal domain, especially of $K[X]$ for a field K. Let's take a closer look at $A := \mathbb{C}[X]$. Recall from Example 2.4.5 that $\mathbb{C} \simeq \operatorname{Max} A$. Hence, we can transfer the Zariski topology on the subspace Max(A) of Spec(A) to a topology on $\mathbb{C}$. How does this look like? A maximal ideal $(X - x)$ in A corresponds to a point $x \in \mathbb{C}$, so the closed subsets in the Zariski topology on $\mathbb{C}$ are of the form $\{x_1, \ldots, x_r\}$ for $x_i \in \mathbb{C}$. Hence, the open subsets are complements of a finite set of points. In contrast to the metric topology on $\mathbb{C}$, the Zariski-open subsets are huge and there are just very few of them—the Zariski topology is much *coarser* than the metric topology!

Nonetheless, the Zariski topology is an extremely valuable tool to study prime ideals and getting some order into the prime spectrum!

Exercises.

EXERCISE 2.5.18. Proof Lemma 2.5.6.

EXERCISE 2.5.19. Let A be a ring and $I \trianglelefteq A$ an ideal. Show that the topological spaces Spec(A/I) and V(I) are **homeomorphic**, i.e. isomorphic as topological spaces.

EXERCISE 2.5.20. Let A be a ring. Show that the sets

$$\mathrm{D}(f) := \operatorname{Spec}(A) \setminus \mathrm{V}(f) \tag{2.63}$$

define a **basis** for the Zariski topology on Spec(A), i.e. every open subset is a union of such sets.

2.6. The Galois connection between closed subsets and radical ideals

We have a way to associate to a subset $S \subseteq A$ a (closed) subset $\mathrm{V}(S) \subseteq \operatorname{Spec}(A)$. Is there also a construction the other way around? Basically, you would like to associate to a (closed) subset $Y \subseteq \operatorname{Spec}(A)$ the "system" defining the "zero set" Y. Thinking a bit, this should be the maximal generic "system" having all the P as points, i.e.

$$\mathrm{I}(Y) := \bigcap_{P \in Y} P \, . \tag{2.64}$$

We call this the **ideal** associated to Y. Now we have two maps

$$\begin{tikzcd} \text{Ideals in } A \arrow[r, shift left, "\mathrm{V}"] & \text{Subsets of } \operatorname{Spec}(A) \arrow[l, shift left, "\mathrm{I}"] \end{tikzcd} \; . \qquad (2.65)$$

Both maps are inclusion-*reversing*, i.e.

$$I \subseteq I' \Rightarrow \mathrm{V}(I) \supseteq \mathrm{V}(I') \tag{2.66}$$

$$Y \subseteq Y' \Rightarrow \mathrm{I}(Y) \supseteq \mathrm{I}(Y') \, . \tag{2.67}$$

This makes sense: a larger system defines more constraints and thus has a smaller zero set. Note that in general the map V, defined on the whole set of ideals in A, is not injective: e.g., in the polynomial ring $\mathbb{Q}[X]$ we have $\mathrm{V}(X) = \mathrm{V}(X^2)$; this is exactly what we discussed in Example 2.4.6. But recall the notion of a Galois connection from Lemma 1.3.18. We only discussed the monotone (inclusion-preserving) version but there's an obvious antitone (inclusion-reversing) analogue.

Lemma 2.6.1. *The maps* (I, V) *form a (antitone) Galois connection, i.e.*

$$I \subseteq \mathrm{I}(Y) \Leftrightarrow Y \subseteq \mathrm{V}(I) \, . \tag{2.68}$$

and consequently the maps induce bijections between the images of $\mathrm{V}\,\mathrm{I}$ *and* $\mathrm{I}\,\mathrm{V}$.

Proof. We have

$$I \subseteq \mathrm{I}(Y) \Leftrightarrow I \subseteq \bigcap_{P \in Y} P \Leftrightarrow I \subseteq P \; \forall P \in Y \Leftrightarrow Y \subseteq \mathrm{V}(I) \, .$$

The claim about induced bijections follows as in the proof of Lemma 1.3.18. □

We want to give a more explicit description of the associated closure operators $\mathrm{V}\,\mathrm{I}$ and $\mathrm{I}\,\mathrm{V}$. For $\mathrm{V}\,\mathrm{I}$ it's really the topological closure: the **closure** of a subset A of a topological space X is the smallest closed subset of X containing A, i.e.

$$\overline{A} := \bigcap_{\substack{Z \subseteq X \text{ closed} \\ A \subseteq Z}} Z \, . \tag{2.69}$$

Lemma 2.6.2. *For any subset* $Y \subseteq \operatorname{Spec}(A)$ *we have*

$$\mathrm{V}\,\mathrm{I}(Y) = \overline{Y} \, . \tag{2.70}$$

Proof. If $P \in Y$, then $\mathrm{I}(Y) \subseteq P$, so $P \in \mathrm{V}\,\mathrm{I}(Y)$, hence $Y \subseteq \mathrm{V}\,\mathrm{I}(Y)$. So, $\mathrm{V}\,\mathrm{I}(Y)$ is a closed subset of $\operatorname{Spec}(A)$ containing Y. Now, let Z be any closed subset of $\operatorname{Spec}(A)$ containing Y. Then $Z = \mathrm{V}(I)$ for some ideal $I \trianglelefteq A$. Since $Y \subseteq \mathrm{V}(I)$, we have $I \subseteq P$ for all $P \in Y$, hence $I \subseteq \bigcap_{P \in Y} P = \mathrm{I}(Y)$. This implies $Z = \mathrm{V}(I) \supseteq \mathrm{V}\,\mathrm{I}(Y)$, i.e. $\mathrm{V}\,\mathrm{I}(Y)$ is the smallest closed subset containing Y. □

Let's come to the other closure operator $\mathrm{I}\,\mathrm{V}$. From the example $\mathrm{V}(X) = \mathrm{V}(X^2)$ in $\mathbb{Q}[X]$ it's clear that V, and thus $\mathrm{I}\,\mathrm{V}$, doesn't care about powers of elements in an ideal I. So, $\mathrm{I}\,\mathrm{V}(I)$ will probably add to I all elements $x \in A$ such that some power x^n is contained in I, i.e.

$$\mathrm{I}\,\mathrm{V}(I) \overset{!?}{=} \sqrt{I} \coloneqq \{x \in A \mid x^n \in I\} \,. \tag{2.71}$$

The set $\sqrt{I}$ is called the **radical** of I. To prove this claim, recall from Exercise 1.4.9 that an element $x \in A$ is called **nilpotent** if $x^n = 0$ for some $n \in \mathbb{N}$. The reason we now talk about nilpotent elements is that we clearly have

$$\sqrt{I} = \{x \in A \mid x \text{ nilpotent in } A/I\} \,. \tag{2.72}$$

So, let's define the **nilradical** of a ring A as

$$\mathrm{Nil}(A) \coloneqq \{x \in A \mid x \text{ is nilpotent}\} \,. \tag{2.73}$$

LEMMA 2.6.3. *The nilradical* $\mathrm{Nil}(A)$ *is an ideal in* A *and*

$$\mathrm{Nil}(A/\mathrm{Nil}(A)) = \{0\} \,. \tag{2.74}$$

PROOF. In Exercise 1.4.9 you have proven that $\mathrm{Nil}(A)$ is closed under addition. It's clear that $\mathrm{Nil}(A)$ is also closed under taking negatives and multiplication with elements from A. Hence, $\mathrm{Nil}(A)$ is an ideal. If $\overline{x} \in A/\mathrm{Nil}(A)$ is nilpotent, then $x^n \in \mathrm{Nil}(A)$ for some $n \in \mathbb{N}$, hence $(x^n)^m = 0$ for some $m \in \mathbb{N}$. But this means x is nilpotent, so $\overline{x} = 0$. □

Now, we make the connection to the Zariski topology.

THEOREM 2.6.4. *We have*

$$\mathrm{Nil}(A) = \bigcap_{P \in \mathrm{Spec}(A)} P \,. \tag{2.75}$$

PROOF. Let $N \coloneqq \bigcap_{P \in \mathrm{Spec}(A)} P$. We know that N is an ideal. Let $x \in \mathrm{Nil}(A)$. Then $x^n = 0$ for some $n \in \mathbb{N}$. Hence, if $P \in \mathrm{Spec}(A)$, then $0 = x^n \in P$. Since P is prime, it follows that $x \in P$ by induction. This implies that $x \in N$, so $\mathrm{Nil}(A) \subseteq N$.

Conversely, let $x \in A \setminus \mathrm{Nil}(A)$. We want to show that $x \notin N$. Let Σ be the set of ideals I in A with the property that $x^n \notin I$ for all $n > 0$. We have $\{0\} \in \Sigma$, so $\Sigma \neq \emptyset$. Let $(I_\lambda)_{\lambda \in \Lambda}$ be a chain in Σ with respect to inclusion. Then $I \coloneqq \bigcup_{\lambda \in \Lambda} I_\lambda$ is an ideal with $x^n \notin I$ for all $n > 0$. Hence, I is an upper bound of the chain $(I_\lambda)_{\lambda \in \Lambda}$ in Σ. Now, Zorn's lemma Lemma 2.3.5 implies that Σ has a maximal element P. We claim that P is prime. Before proving this, note that $x \notin P$ since $P \in \Sigma$, hence we have found a prime ideal not containing x, hence $x \notin N$ which is what we wanted to show.

So, let's prove that P is prime. First note that $P \neq A$ since $x \notin P$. Let $a, b \in A$ with $a \notin P$ and $b \notin P$. We need to show that $ab \notin P$. The ideals $P + (a)$ and $P + (b)$ are strictly larger than P, so because of the maximality of P in Σ, they are not contained in Σ. Hence, there are $m, n \in \mathbb{N}$ with $x^n \in P + (a)$ and $x^m \in P + (b)$. We thus have $x^n = f + ac$ and $x^m = f' + bc'$ for some $f, f' \in P$ and $c, c' \in A$. Then

$$x^{n+m} = \underbrace{ff' + fbc + f'ac}_{\in P} + \underbrace{cc'ab}_{\in (ab)} \,,$$

so $x^{n+m} \in P + (ab)$ and therefore $P + (ab) \notin \Sigma$. But then also $ab \notin P$ since otherwise $P + (ab) = P \in \Sigma$. This shows that P is prime. □

We can now give an explicit description of the closure operator $\mathrm{I}\,\mathrm{V}$:

COROLLARY 2.6.5. *The radical $\sqrt{I}$ of an ideal $I \trianglelefteq A$ is an ideal as well and*

$$\sqrt{I} = \bigcap_{\substack{P \in \mathrm{Spec}(A) \\ P \supseteq I}} P = \mathrm{I}\,\mathrm{V}(I)\ . \tag{2.76}$$

PROOF. Recall from (2.72) that $\sqrt{I} = \{x \in A \mid \overline{x} \in \mathrm{Nil}(A/I)\}$. Applying Theorem 2.6.4 to A/I yields $\mathrm{Nil}(A/I) = \bigcap_{P \in \mathrm{Spec}(A/I)} P$. The claim follows from the correspondence between prime ideals in A/I and prime ideals in A containing I. □

In particular, the image of the closure operator $\mathrm{I}\,\mathrm{V}$ consists of ideals of the form $\sqrt{I}$. Such ideals are called **radical ideals**. Note that $\sqrt{\sqrt{I}} = \sqrt{I}$ by definition of the radical, so I being a radical ideal means $\sqrt{I} = I$. From Corollary 2.6.5 it is clear that prime ideals are radical ideals. We finally deduce:

COROLLARY 2.6.6. *The maps* I *and* V *restrict to bijections*

$$\begin{tikzcd} \textit{Radical ideals in } A \arrow[r, shift left, "\mathrm{V}"] & \textit{Closed subsets of } \mathrm{Spec}(A)\ . \arrow[l, shift left, "\mathrm{I}"] \end{tikzcd} \tag{2.77}$$

PROOF. This is clear now. □

Recall that the nilradical of a ring A is equal to the intersection of all the prime ideals in A. There's the following important enlargement:

DEFINITION 2.6.7. The **Jacobson radical** of A is

$$\mathrm{Jac}(A) \coloneqq \bigcap_{M \in \mathrm{Max}(A)} M\ . \tag{2.78}$$

Whereas the nilradical consists of the nilpotent elements, the Jacobson radical is about units.

LEMMA 2.6.8. *We have $x \in \mathrm{Jac}(A)$ if and only if $1 - xy$ is a unit for all $y \in A$.*

PROOF. Let $x \in \mathrm{Jac}(A)$. Suppose that $1 - xy$ would not be a unit. Then we know from Corollary 2.3.8 that there is a maximal ideal M with $1 - xy \in M$. But $x \in \mathrm{Jac}(A) \subseteq M$, so $xy \in M$ and therefore $1 \in M$, which is a contradiction. Hence, $1 - xy$ must be a unit.

Conversely, assume that $1 - xy$ is a unit for all $y \in A$. Suppose that $x \notin \mathrm{Jac}(A)$. Then there is a maximal ideal M with $x \notin M$. Since M is maximal, we must have $M + (x) = A$. Hence, there is $m \in M$ and $y \in A$ such that $1 = m + xy$, so $1 - xy \in M$ and therefore $1 - xy$ cannot be a unit (since otherwise $M = A$), which is a contradiction. Hence, $x \in M$ for all maximal ideals M and therefore $x \in \mathrm{Jac}(A)$. □

Exercises.

EXERCISE 2.6.9. Let A be a ring.

(1) Show that $\sqrt{I^n} = \sqrt{I}$ for any $I \trianglelefteq A$ and $n \in \mathbb{N}_{>0}$.
(2) Show that $\sqrt{I \cap J} = \sqrt{I} \cap \sqrt{J}$.
(3) $\sqrt{I + J} = \sqrt{\sqrt{I} + \sqrt{J}}$ for any $I, J \trianglelefteq A$.
(4) Find an example where $\sqrt{I + J} \neq \sqrt{I} + \sqrt{J}$.

EXERCISE 2.6.10. Let K be a field. Determine the following radicals:

(1) $\sqrt{(X^3 - X^2 - X + 1)}$ in $K[X]$.
(2) $\sqrt{(X_1^2 - X_2X_3, X_1(1 - X_3)}$ in $K[X_1, X_2, X_3]$.

Exercise 2.6.11. Let A be a ring. Show that the topological space $\mathrm{Spec}(A)$ is **quasi-compact**, i.e. every open cover of $\mathrm{Spec}(A)$ has a *finite* subcover.

Exercise 2.6.12. Let R be a ring. Show that $\mathrm{Nil}(R[X]) = \mathrm{Jac}(R[X])$. What is the Jacobson radical of $\mathbb{Z}[X]$?

2.7. Irreducible components

Recall the spectrum of $A \coloneqq \mathbb{Q}[X_1, X_2]/(X_1X_2)$ from Example 2.4.8. Geometrically, this ring describes the union of the two coordinate axes $\{X_1 = 0\}$ and $\{X_2 = 0\}$. The discussion in Example 2.4.8 shows that (X_1) and (X_2) are the unique minimal prime ideals in A and we can write

$$\mathrm{Spec}(A) = \mathrm{V}(X_1) \cup \mathrm{V}(X_2) .$$

So, we can decompose the spectrum into two closed subsets—not necessarily disjoint but that's alright—and this really reflects the geometric picture.

Can we do something like this for *any* ring? The answer is...yes! It's helpful to discuss this for a general topological space.

Definition 2.7.1. A topological space X is called **irreducible** if $X \neq \emptyset$ and X cannot be written as the union of two proper closed subsets, i.e. $X = Z_1 \cup Z_2$ with closed subsets Z_i implies $Z_1 = X$ or $Z_2 = X$.

If X is not irreducible, it's called **reducible** (double negation). Note that the definition of irreducible is very similar to the definition of a prime ideal—indeed we will soon prove a connection.

Example 2.7.2. The topological notion of irreducibility is really made for the Zariski topology—it's basically useless for "usual", i.e. metric, topologies. For example, the complex line $\mathbb{C}$ equipped with the metric topology is reducible since

$$\mathbb{C} = \{x \mid |x| \leq 1\} \cup \{x \mid |x| \geq 1\} \tag{2.79}$$

is a union of two non-trivial closed subsets—which doesn't fit the picture we want. The metric topology is way too fine for irreducibility to be a useful notion—but this is very different for the (much coarser) Zariski topology.

Lemma 2.7.3. *For a topological space $X \neq \emptyset$ the following are equivalent:*

(1) *X is irreducible.*
(2) *Any two non-empty open subsets of X have non-empty intersection.*

Proof. Suppose that X is irreducible and let U_1, U_2 be two non-empty open subsets. Then $Z_i \coloneqq X \setminus U_i$ is a proper closed subset. Since X is irreducible, we must have $X \neq Z_1 \cup Z_2$, hence

$$\emptyset \neq X \setminus (Z_1 \cup Z_2) = (X \setminus Z_1) \cap (X \setminus Z_2) = U_1 \cap U_2 .$$

This argument can be read backwards, proving the converse. □

Definition 2.7.4. A subset A of a topological space is **irreducible** if it is an irreducible topological space with respect to the subspace topology.

Lemma 2.7.5. *A subset A of a topological space X is irreducible if and only if its closure $\overline{A}$ is irreducible.*

Proof. We can assume $A \neq \emptyset$. Assume that A is irreducible. Suppose we can decompose $\overline{A} = Z_1 \cup Z_2$ with two closed subsets $Z_i \subseteq \overline{A}$. By Remark 2.5.13, we have $Z_i = \overline{A} \cap Z_i'$ with $Z_i' \subseteq X$ closed. Hence,

$$A = A \cap \overline{A} = A \cap (Z_1 \cup Z_2) = (A \cap Z_1) \cup (A \cap Z_2) = (A \cap Z_1') \cup (A \cap Z_2')\ .$$

Since $A \cap Z_i'$ is closed in A by Remark 2.5.13 and A is irreducible, we have $A = A \cap Z_i'$ for some i. Then $A \subseteq Z_i'$ and since Z_i' is closed, it follows that $\overline{A} \subseteq Z_i'$. Hence, $Z_i = \overline{A} \cap Z_i' = \overline{A}$. This shows that $\overline{A}$ is irreducible.

Conversely, assume that A is reducible, so $A = Z_1 \cup Z_2$ with $Z_i \subsetneq A$ closed. We claim that $\overline{A} = \overline{Z_1} \cup \overline{Z_2}$. In fact, $Z_i \subseteq Z_1 \cup Z_2$, so $\overline{Z_i} \subseteq \overline{Z_1 \cup Z_2} = \overline{A}$, hence $\overline{Z_1} \cup \overline{Z_2} \subseteq \overline{A}$. Conversely, $\overline{Z_1} \cup \overline{Z_2}$ is a closed set containing Z_1 and Z_2, hence $\overline{A} = \overline{Z_1 \cup Z_2} \subseteq \overline{Z_1} \cup \overline{Z_2}$. It remains to show that $\overline{Z_i} \neq \overline{A}$. We have

$$A \cap \overline{Z}_i = A \cap \bigcap_{\substack{Z \supseteq Z_i \\ Z \text{ closed}}} Z = \bigcap_{\substack{Z \supseteq Z_i \\ Z \text{ closed}}} (Z \cap A) = Z_i\ ,$$

since Z_i is closed in A. It follows that $\overline{Z}_i \neq \overline{A}$ since $\overline{A} = \overline{Z}_i$ implies $A = \overline{A} \cap A = \overline{Z}_i \cap A = Z_i$, which is a contradiction. Hence, $\overline{A}$ is reducible. □

Definition 2.7.6. An **irreducible component** of a topological space X is a maximal irreducible subset.

Proposition 2.7.7. *Let X be a topological space. Then:*

(1) *Every irreducible subset is contained in an irreducible component of X.*
(2) *X is the union of its irreducible components.*
(3) *The irreducible components are closed subsets.*

Proof. (1): Let $A \subseteq X$ be an irreducible subset. We want to show that A is contained in an irreducible component of X, i.e. in a maximal irreducible subset. Let Σ be the set of irreducible subsets of X containing A. We have $\Sigma \neq \emptyset$ since $A \in \Sigma$. Let $(Z_\lambda)_{\lambda \in \Lambda}$ be a chain in Σ. Set $Z := \bigcup_{\lambda \in \Lambda} Z_\lambda$. We claim that Z is irreducible. Suppose that $Z = Z_1 \cup Z_2$ with $Z_i \subseteq Z$ closed. For every $\lambda \in \Lambda$ we have

$$Z_\lambda = Z_\lambda \cap (Z_1 \cup Z_2) = (Z_\lambda \cap Z_1) \cup (Z_\lambda \cap Z_2)\ .$$

Hence, since Z_λ is irreducible, we must have

$$Z_\lambda \cap Z_1 = Z_\lambda\ (\Rightarrow Z_\lambda \subseteq Z_1) \quad \text{or} \quad Z_\lambda \cap Z_2 = Z_\lambda\ (\Rightarrow Z_\lambda \subseteq Z_2)\ . \tag{2.80}$$

If $Z_\lambda \subseteq Z_1$ for *all* $\lambda \in \Lambda$, then $Z \subseteq Z_1$, implying $Z = Z_1$. On the other hand, if there is $\mu \in \Lambda$ with $Z_\mu \not\subseteq Z_1$, then we must have $Z_\mu \subseteq Z_2$ by (2.80). But then $Z_\lambda \subseteq Z_2$ for *all* $\lambda \in \Lambda$ for the following reason: because the Z_λ form a chain, we have $Z_\lambda \subseteq Z_\mu$ or $Z_\mu \subseteq Z_\lambda$; if $Z_\lambda \subseteq Z_\mu$, then clearly $Z_\lambda \subseteq Z_2$; if $Z_\mu \subseteq Z_\lambda$ but $Z_\lambda \not\subseteq Z_2$, then $Z_\lambda \subseteq Z_1$ by (2.80), so $Z_\mu \subseteq Z_1$, which is a contradiction. In conclusion, Z is irreducible, i.e. $Z \in \Sigma$. Now, Zorn's Lemma (Lemma 2.3.5) implies that Σ has a maximal element. This is a maximal irreducible subset of X, hence an irreducible component of X, and it contains A by construction.

(2): For every $x \in X$ the set $\{x\}$ is irreducible, hence it is contained in an irreducible component by (1). Consequently, X is the union of its irreducible components.

(3): If Z is an irreducible component, then also its closure $\overline{Z}$ is irreducible by Lemma 2.7.5. Because of the maximality of Z, we must have $Z = \overline{Z}$, i.e. Z is closed. □

In particular, the prime spectrum $\operatorname{Spec}(A)$ of a ring A is the union of its irreducible components. But what do irreducible subsets and the irreducible components look like in this case?

LEMMA 2.7.8. *Let A be a ring. A subset $Y \subseteq \operatorname{Spec}(A)$ is irreducible if and only if $\mathrm{I}(Y)$ is a prime ideal. In this case, $\overline{Y}$ is equal to the closure of the point $\mathrm{I}(Y) \in Y$.*

PROOF. Suppose Y is irreducible. Then $\overline{Y}$ is irreducible by Lemma 2.7.5 and by Lemma 2.6.2 we have $\mathrm{V}\,\mathrm{I}(Y) = \overline{Y}$. Let $a, b \in A$ with $ab \in \mathrm{I}(Y)$. Then

$$\mathrm{V}\,\mathrm{I}(Y) \subseteq \mathrm{V}(ab) = \mathrm{V}(a) \cup \mathrm{V}(b)\ .$$

Since $\overline{Y} = \mathrm{V}\,\mathrm{I}(Y)$ is irreducible, we must have $\mathrm{V}\,\mathrm{I}(Y) \subseteq \mathrm{V}(a)$ or $\mathrm{V}\,\mathrm{I}(Y) \subseteq \mathrm{V}(b)$. Using the Galois connection property, this implies

$$\mathrm{I}\,\mathrm{V}\,\mathrm{I}(Y) \supseteq (a) \quad \text{or} \quad \mathrm{I}\,\mathrm{V}\,\mathrm{I}(Y) \supseteq (b)\ .$$

By the relations (1.55) of a Galois connection we have $\mathrm{I}\,\mathrm{V}\,\mathrm{I}(Y) = \mathrm{I}(Y)$, hence $\mathrm{I}(Y) \supseteq (a)$ or $\mathrm{I}(Y) \supseteq (b)$. This proves that $\mathrm{I}(Y)$ is a prime ideal.

Conversely, assume that P is a prime ideal. Then by Lemma 2.6.2 we have

$$\overline{\{P\}} = \mathrm{V}\,\mathrm{I}(P) = \mathrm{V}(P)\ ,$$

using that $\mathrm{I}(P) = \bigcap_{Q \supseteq P} Q = P$. □

Hence, using the Galois connection from Corollary 2.6.6, we conclude:

COROLLARY 2.7.9. *The map $P \mapsto V(P) = \overline{\{P\}}$ is a bijection between $\operatorname{Spec}(A)$ and the irreducible closed subsets of $\operatorname{Spec}(A)$.*

This tells us that prime ideals really are the "atoms" of a ring—in a geometric sense. You see, this is not about factorizations. Since the irreducible components are the maximal irreducible subsets, it follows that the irreducible components of $\operatorname{Spec}(A)$ correspond to the **minimal prime ideals** in A. In particular, minimal prime ideals exist. Moreover, it follows that every irreducible closed subset $Z \subseteq \operatorname{Spec}(A)$ contains a unique point $z \in Z$ such that $Z = \overline{\{z\}}$. This point is called the **generic point** of Z.

COROLLARY 2.7.10. *For any ideal $I \neq A$ of a ring A there is a minimal prime ideal over I, i.e. a prime ideal P minimal with the property that $P \supseteq I$.*

PROOF. This follows at once since the minimal prime ideals over I correspond precisely to the irreducible components of the space $\mathrm{V}(I)$. □

EXAMPLE 2.7.11. Recall the example $A \coloneqq \mathbb{Q}[X_1, X_2]/(X_1X_2)$. We know from Example 2.4.8 that the minimal prime ideals are (X_1) and (X_2). Hence, the irreducible components of $\operatorname{Spec}(A)$ are $\mathrm{V}(X_1)$ and $\mathrm{V}(X_2)$. This is exactly what we wanted.

EXAMPLE 2.7.12. If A is an integral domain, then

$$\overline{\{(0)\}} = \mathrm{V}(0) = \operatorname{Spec}(A) \tag{2.81}$$

because (0) is a prime ideal. Hence, $\operatorname{Spec}(A)$ is irreducible with generic point (0).

Exercises.

EXERCISE 2.7.13. Let K be a field and consider the ideal

$$I \coloneqq (X_1X_3 - X_2^3, X_1^3 - X_2X_3) \trianglelefteq K[X_1, X_2, X_3]\ .$$

Decompose $\mathrm{V}(I)$ into irreducible components.

CHAPTER 3

Modules

3.1. The category of modules

The definition of a module over a ring is exactly the same as that of a vector space—you only replace the base field by a general ring.

Definition 3.1.1. Let A be a ring. An **A-module** is an abelian group $(V,+)$ together with an operation

$$A \times V \to V \ , \quad (a,v) \mapsto av \ , \tag{3.1}$$

of A on V which is:

(1) distributive, i.e.

$$a(v+v') = av + av' \quad \text{and} \quad (a+a')v = av + a'v \tag{3.2}$$

for all $a, a' \in A$ and $v, v' \in V$;

(2) compatible with the multiplication of A, i.e.

$$(aa')v = a(a'v) \quad \text{and} \quad 1v = v \tag{3.3}$$

for all $a, a' \in A$ and $v \in V$.

The operation of A on V is also called a **scalar operation**. As for vector spaces, you can derive simple relations like

$$0v = (0+0)v = 0v + 0v, \text{ hence } 0 = 0v \tag{3.4}$$

for any $v \in V$.

Definition 3.1.2. A **morphism** between A-modules V and W is a map $f\colon V \to W$ such that:

(1) f is a group morphism $(V,+) \to (W,+)$, i.e.

$$f(v+v') = f(v) + f(v') \tag{3.5}$$

for all $v, v' \in V$;

(2) f is **A-linear**, i.e.

$$f(av) = af(v) \tag{3.6}$$

for all $a \in A$ and $v \in V$.

It is clear that the composition of A-module morphisms is again an A-module morphism, hence we get a category A-Mod of A-modules. Because we'll use it so many times, we'll abbreviate

$$\operatorname{Hom}_A(V,W) \coloneqq \operatorname{Hom}_{A\text{-}\mathsf{Mod}}(V,W) \ , \tag{3.7}$$

$$\operatorname{End}_A(V) \coloneqq \operatorname{Hom}_A(V,V) \ . \tag{3.8}$$

As for vector spaces, you can show that an A-module morphism is an **isomorphism** if and only if it is bijective.

Example 3.1.3. If K is a field, then K-modules are precisely the K-vector spaces and K-module morphisms are precisely the K-linear maps of vector spaces. Hence,

$$K\text{-}\mathsf{Mod} = K\text{-}\mathsf{Vec}\ , \tag{3.9}$$

the latter being the category of K-vector spaces.

Example 3.1.4. Any ring A is naturally an A-module with respect to the multiplication as operation, i.e. $A \times A \to A$, $(a, a') \mapsto aa'$. But caution: A-module morphisms $A \to A$ and ring morphisms $A \to A$ are different things: you want $f(aa') = f(a)f(a')$ for a ring morphism but $f(aa') = af(a')$ for a module morphism.

Example 3.1.5. The scalar operation of a ring R on an R-algebra A makes A into an R-module. Now you see: we have (at least) two module structures on A: we can view it as an A-module and as an R-module. Sometimes one needs to clearify which module structure one is considering.

Example 3.1.6. Any abelian group G is naturally a $\mathbb{Z}$-module via

$$ng \coloneqq \underbrace{g + \ldots + g}_{n \text{ times}} \tag{3.10}$$

and $(-n)g = -(ng)$ for $n \in \mathbb{N}$ and $g \in G$. Conversely, any $\mathbb{Z}$-module is of course an abelian group. Moreover, morphisms of abelian groups are the same as $\mathbb{Z}$-module morphisms for the $\mathbb{Z}$-module structure defined above. We conclude:

$$\mathbb{Z}\text{-}\mathsf{Mod} = \mathsf{Ab}\ , \tag{3.11}$$

the latter denoting the category of abelian groups.

Example 3.1.7. Let K be a field and consider the polynomial ring $K[X]$ in one variable. Let V be a $K[X]$-module. Since $K \subseteq K[X]$, we have a natural action of K on V, so V is naturally a K-vector space. The action of X on V defines a vector space endomorphism $f\colon V \to V$ via $v \mapsto Xv$. So, we can associate to a $K[X]$-module V a pair (V, f) of a K-vector space and a vector space endomorphism. Conversely, any such pair can be "upgraded" to a $K[X]$-module by defining the action of X via $Xv \coloneqq f(v)$ and then extend to $K[X]$ via

$$pv \coloneqq p(f)v \coloneqq a_0 v + a_1 f(v) + a_2 f^2(v) \ldots + a_n f^n(v) \tag{3.12}$$

for $p = \sum_{i=0}^n a_i X^i$. Things like the Jordan normal form actually follow from structural results about modules over principal ideal domains.

Example 3.1.8. If V and W are A-modules, then $\mathrm{Hom}_A(V, W)$ is naturally an A-module with respect to pointwise operations, i.e.

$$(f + g)(v) \coloneqq f(v) + g(v)\ , \tag{3.13}$$

$$(af)(v) \coloneqq af(v)\ , \tag{3.14}$$

for all $f, g \in \mathrm{Hom}_A(V, W)$, $v \in V$ and $a \in A$.

Example 3.1.9. If $\varphi\colon A \to B$ is a ring morphism and W is a B-module, then W naturally becomes an A-module via

$$aw \coloneqq \varphi(a)w\ . \tag{3.15}$$

One also writes W_A when considering W in this way as an A-module (dropping φ in the notation). A B-module morphism $f\colon W \to W'$ is then naturally an A-module morphism $W_A \to W'_A$. We thus get a functor

$$(-)_A\colon B\text{-}\mathsf{Mod} \to A\text{-}\mathsf{Mod}\ . \tag{3.16}$$

This process is called **scalar restriction** because one often uses this in case φ is the embedding of a subring—but in general we do not need to assume that φ is injective.

Remark 3.1.10. One can use the same definition of modules over a non-commutative ring. One distinguishes here between left and right modules, depending on whether A acts from the left, $A \times V \to V$, or from the right, $V \times A \to V$. If A is commutative, any left module can naturally be viewed as a right module and vice versa.

3.2. Basic constructions that work like for vector spaces

Many of the basic constructions for modules are completely analogous to what you already know for vector spaces.

We can form the **direct product**

$$\prod_{\lambda\in\Lambda} W_\lambda \coloneqq \{(w_\lambda)_{\lambda\in\Lambda} \mid w_\lambda \in W_\lambda\} \tag{3.17}$$

of a family $(W_\lambda)_{\lambda\in\Lambda}$ of A-modules (the operation on the product is componentwise), and this satisfies the universal property of a direct product (see Lemma 1.2.10) in the category of A-modules, i.e. given A-module morphisms $f_\mu\colon V \to W_\mu$ for all $\mu \in \Lambda$, there is a unique A-module morphism $f\colon V \to \prod_{\lambda\in\Lambda} W_\lambda$ such that the diagram

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[dr, "f_\mu"'] & \prod_{\lambda\in\Lambda} W_\lambda \arrow[d, "\mathrm{p}_\mu"] \\ & W_\mu \end{tikzcd} \tag{3.18}$$

commutes for all $\mu \in \Lambda$. Here,

$$\mathrm{p}_\mu\colon \prod_{\lambda\in\Lambda} W_\lambda \to W_\mu \tag{3.19}$$

is the **projection** onto W_μ.

We can also form the **direct sum**

$$\bigoplus_{\lambda\in\Lambda} V_\lambda = \left\{ (v_\lambda)_{\lambda\in\Lambda} \in \prod_{\lambda\in\Lambda} V_\lambda \mid v_\lambda = 0 \text{ for all but finitely many } \lambda \right\}\ . \tag{3.20}$$

This satisfies the following universal property: given A-module morphisms $f_\mu\colon V_\mu \to W$ for all $\mu \in \Lambda$ there is a unique A-module morphism $f\colon \bigoplus_{\lambda\in\Lambda} V_\lambda \to W$ such that the diagram

$$\begin{tikzcd} \bigoplus_{\lambda\in\Lambda} V_\lambda \arrow[r, "f"] & W \\ V_\mu \arrow[u, "\mathrm{i}_\mu"] \arrow[ur, "f_\mu"'] & \end{tikzcd} \tag{3.21}$$

commutes for all $\mu \in \Lambda$. Here,

$$\mathrm{i}_\mu \colon V_\mu \to \bigoplus_{\lambda\in\Lambda} V_\lambda \tag{3.22}$$

is the **injection** of V_μ into the direct sum. Note that the universal property of the direct sum is just like that of the direct product but with arrows reversed. One therefore also says that the direct sum is the **coproduct** in the category of A-modules. For a finite family Λ the direct product and the direct sum is the same thing—one says it's a **biproduct**.

Recall that we can view a ring A as an A-module. We have now in particular defined the products

$$A^\Lambda \coloneqq \prod_{\lambda\in\Lambda} A \ , \tag{3.23}$$

$$A^{(\Lambda)} \coloneqq \bigoplus_{\lambda\in\Lambda} A \ . \tag{3.24}$$

A **submodule** of an A-module V is a subgroup U of $(V, +)$ such that $AU \subseteq U$. The operation of A on V then restricts to an operation of A on U and this turns U into an A-module. Any A-module V has the **trivial submodules** 0 and V. A **proper** submodule of V is a submodule $U \neq V$.

Example 3.2.1. Submodules of a vector space are precisely the subspaces.

Example 3.2.2. When considering a ring A with the natural A-module structure, see Example 3.1.4, then submodules of A are precisely the ideals in A.

If V is an A-module and $U \subseteq V$ is a submodule, then A induces an operation on the **quotient** V/U of additive groups via

$$a\,\overline{v} \coloneqq \overline{av} \ , \tag{3.25}$$

and this makes V/U into an A-module. We have a **quotient map** $V \to V/U$ satisfying the usual universal property in the category of A-modules, see Lemma 1.3.6.

If $f\colon V \to W$ is a morphism of A-modules, then we have a (monotone) Galois connection

$$\begin{array}{rcl} \mathrm{Sub}(V) & \leftrightarrow & \mathrm{Sub}(W) \\ V' & \mapsto & f(V') \\ f^{-1}(W') & \leftarrow\!\shortmid & W' \end{array} \tag{3.26}$$

between the partially ordered sets of submodules. Especially, the **kernel**

$$\mathrm{Ker}(f) \coloneqq f^{-1}(0) \tag{3.27}$$

is a submodule of V and the **image**

$$\mathrm{Im}(f) \coloneqq f(V) \tag{3.28}$$

is a submodule of W. If f is *surjective*, the maps above restrict to bijections

$$\{V' \in \mathrm{Sub}(V) \mid \mathrm{Ker}(f) \subseteq V'\} \simeq \mathrm{Sub}(W) \ . \tag{3.29}$$

In particular, for a submodule $U \subseteq V$ we have bijections

$$\{V' \in \mathrm{Sub}(V) \mid U \subseteq V'\} \simeq \mathrm{Sub}(V/U) \ . \tag{3.30}$$

The usual **isomorphism theorems**isomorphism theorem!for modules you know for vector spaces hold similarly for A-modules.

As for vector spaces, the **intersection** $\bigcap_{\lambda \in \Lambda} U_\lambda$ of a family of submodules U_λ of V is a submodule. Hence, given a subset $\boldsymbol{v} := \{v_\lambda\}_{\lambda \in \Lambda}$ of an A-module V, there is a unique submodule of V minimal among all submodules containing $\boldsymbol{v}$, namely

$$A\boldsymbol{v} := \bigcap_{\substack{U \in \mathrm{Sub}(V) \\ \boldsymbol{v} \subseteq U}} U = \left\{ \sum_{\lambda \in \Lambda} a_\lambda v_\lambda \mid a_\lambda \in A, \text{ all but finitely many} = 0 \right\} , \tag{3.31}$$

i.e. $A\boldsymbol{v}$ consists of the finite A-linear combinations of elements of $\boldsymbol{v}$. We call this the submodule of V **generated** by $\boldsymbol{v}$. If $V = A\boldsymbol{v}$, we call $\boldsymbol{v}$ a set of A-module **generators** of V. Such a set always exists since we can take $\boldsymbol{v} = V$. We say that V is **finitely generated** if it admits a finite set of generators. If $\boldsymbol{v}$ is a set of generators of V, we get a surjective A-module morphism

$$A^{(\Lambda)} \to V \ , \quad e_\lambda \mapsto v_\lambda \ , \tag{3.32}$$

where e_λ is the tuple having the entry 1 in position λ and being 0 elsewhere. In particular, any A-module is a quotient of $A^{(\Lambda)}$ for appropriate Λ. Clearly, V is finitely generated if and only if a finite Λ can be chosen. The kernel of the morphism (3.32) is called the **syzygy module** of $\boldsymbol{v}$ and is denoted by $\mathrm{Syz}_A(\boldsymbol{v})$. This module describes the **relations** between the generators in $\boldsymbol{v}$, i.e. tuples $\boldsymbol{a} := (a_\lambda)_{\lambda \in \Lambda} \in A^{(\Lambda)}$ such that

$$0 = \boldsymbol{a} \cdot \boldsymbol{v}^t = \sum_{\lambda \in \Lambda} a_\lambda v_\lambda \ . \tag{3.33}$$

As for ideals (and vector spaces), we define the **sum** $\sum_{\lambda \in \Lambda} U_\lambda$ of a family $(U_\lambda)_{\lambda \in \Lambda}$ of submodules of V to be the submodule generated by the union $\bigcup_{\lambda \in \Lambda} U_\lambda$. Explicitly, we have

$$\sum_{\lambda \in \Lambda} U_\lambda = \left\{ \sum_{\lambda \in \Lambda} a_\lambda u_\lambda \mid u_\lambda \in U_\lambda, a_\lambda \in A, \text{ all but finitely many} = 0 \right\} . \tag{3.34}$$

The set of submodules of V is a complete lattice with respect to intersection and sum.

Exercises.

Exercise 3.2.3. Let A be a ring. Show that for a family $(V_\lambda)_{\lambda \in \Lambda}$ of A-modules and an A-module W there is a canonical isomorphism

$$\mathrm{Hom}_A(\bigoplus_{\lambda \in \Lambda} V_\lambda, W) \simeq \prod_{\lambda \in \Lambda} \mathrm{Hom}_A(V_\lambda, W) \tag{3.35}$$

of A-modules.

Exercise 3.2.4. Let A be a ring. Show that for an A-module V and a family $(W_\lambda)_{\lambda \in \Lambda}$ of A-modules there is a canonical isomorphism

$$\mathrm{Hom}_A(V, \prod_{\lambda \in \Lambda} W_\lambda) \simeq \prod_{\lambda \in \Lambda} \mathrm{Hom}_A(V, W_\lambda) \tag{3.36}$$

of A-modules.

EXERCISE 3.2.5. Let $(V_\lambda)_{\lambda\in\Lambda}$ be a family of submodules and for each λ let $U_\lambda \subseteq V_\lambda$ be a submodule. Then $\bigoplus_{\lambda\in\Lambda} U_\lambda$ is naturally a submodule of $\bigoplus_{\lambda\in\Lambda} V_\lambda$ and there is a canonical A-module isomorphism

$$\left(\bigoplus_{\lambda\in\Lambda} V_\lambda\right) / \left(\bigoplus_{\lambda\in\Lambda} U_\lambda\right) \simeq \bigoplus_{\lambda\in\Lambda} V_\lambda/U_\lambda \ . \tag{3.37}$$

3.3. A whirlwind of emotions

So far, so nice. (Un)fortunately, there are several twists and traps in the module story and you need to be very careful when using your intuition for vector spaces also for general modules. Let's start with the most important thing you know about vector spaces: they have a basis.

DEFINITION 3.3.1. Let V be an A-module. A subset $\boldsymbol{v} := \{v_\lambda\}_{\lambda\in\Lambda}$ of V is said to be **linearly independent** if $\mathrm{Syz}_A(\boldsymbol{v}) = 0$, i.e. a relation

$$\sum_{\lambda\in\Lambda'} a_\lambda v_\lambda = 0 \tag{3.38}$$

for a finite subset $\Lambda' \subseteq \Lambda$ implies that $a_\lambda = 0$ for all $\lambda \in \Lambda'$. A linearly independent generating set is called a **basis**. A module admitting a basis is called **free**.

It's clear that a set $\boldsymbol{v}$ is a basis of V if and only if the morphism

$$A^{(\Lambda)} \to V \ , \quad e_\lambda \mapsto v_\lambda \ , \tag{3.39}$$

from (3.32) is an isomorphism. Especially, up to isomorphism the free A-modules are precisely those of the form $A^{(\Lambda)}$. It's clear from the definition that if V is free with basis $\boldsymbol{v}$, then every $v \in V$ has a *unique* expression as $v = \sum_{\lambda\in\Lambda} a_\lambda v_\lambda$, i.e. if also $v = \sum_{\lambda\in\Lambda} a'_\lambda v_\lambda$, then $a_\lambda = a'_\lambda$ for all $\lambda \in \Lambda$. Moreover, a free module V with basis $\{v_\lambda\}_{\lambda\in\Lambda}$ satisfies the following universal property: for any A-module W and any map $\varphi\colon \Lambda \to W$ there is a unique A-module morphism $f\colon V \to W$ mapping v_λ to $\varphi(\lambda)$. Namely, you define

$$f\left(\sum_{\lambda\in\Lambda} a_\lambda v_\lambda\right) := \sum_{\lambda\in\Lambda} a_\lambda \varphi(\lambda) \ . \tag{3.40}$$

In particular, morphisms between free modules can, after choosing bases, be described by (possibly infinite) matrices over R—just like for vector spaces.

As for vector spaces, we can show that all bases of a free module have the same cardinality so that we get a well-defined notion of rank of a free module. To prove this, we'll cook up a vector space from an A-module V as follows (we'll use this construction many times). Let I be an ideal in A. Then the submodule of V generated by elements of the form av for $a \in I$ and $v \in V$ is given by

$$IV := A\{av \mid a \in I, v \in V\} = \left\{\sum_{i=1}^n a_i v_i \mid n \in \mathbb{N}, a_i \in I, v_i \in V\right\} \ . \tag{3.41}$$

The action of A on the quotient module V/IV induces an action of A/I via

$$\overline{a}\ \overline{v} := \overline{av} \ . \tag{3.42}$$

This makes V/IV naturally into an (A/I)-module. Especially, if M is a maximal ideal in A, then V/MV is naturally a K-vector space, where $K := A/M$. So, from a general module you can obtain a whole **vector bundle** $(V/MV)_{M\in\mathrm{Max}(A)}$.

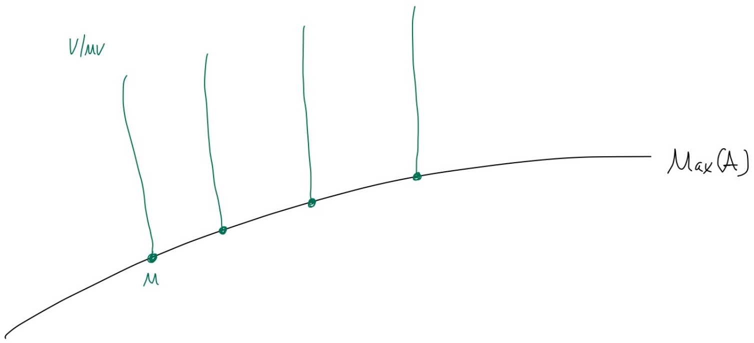

FIGURE 3.1. One way to think about a module: as a vector bundle.

THEOREM 3.3.2. *Let $A \neq 0$ and let V be a free A-module. Then all bases of V have the same cardinality, which is called the* ***rank*** *of V and is denoted by* $\operatorname{rk}_A(V)$.

PROOF. Since $A \neq 0$, there is a maximal ideal in A by Theorem 2.3.6. So, let M be any maximal ideal. Let $K := A/M$, which is a field. Recall from above that the quotient $\overline{V} := V/MV$ is naturally a K-vector space. Let $\boldsymbol{v} := \{v_\lambda\}_{\lambda\in\Lambda}$ be a generating set of V. Consider the image $\overline{\boldsymbol{v}} := \{\overline{v}_\lambda\}_{\lambda\in\Lambda}$ in $\overline{V}$. Then clearly, $\overline{\boldsymbol{v}}$ is a generating set of $\overline{V}$ as a K-vector space. Hence, $\#\Lambda \geq \dim_K(\overline{V})$.

Now, assume that $\boldsymbol{v}$ is a basis. We claim that $\overline{\boldsymbol{v}}$ is a basis of $\overline{V}$ as a K-vector space. We already know that $\overline{\boldsymbol{v}}$ is a generating set. Suppose that we have a relation $\sum_{\lambda\in\Lambda} \overline{a}_\lambda \overline{v}_\lambda = 0$. This means $\sum_{\lambda\in\Lambda} a_\lambda v_\lambda \in MV$. Since $\boldsymbol{v}$ is a generating set of V and M is an ideal in A, we have

$$MV = \left\{ \sum_\lambda m_\lambda v_\lambda \mid m_\lambda \in M, \text{ all but finitely many} = 0 \right\} .$$

Hence, $\sum_{\lambda\in\Lambda} a_\lambda v_\lambda = \sum_{\lambda\in\Lambda} m_\lambda v_\lambda$ for certain $m_\lambda \in M$. Since $\boldsymbol{v}$ is a basis, it follows that $a_\lambda = m_\lambda$ for all $\lambda \in \Lambda$, hence $\overline{a}_\lambda = 0$ for all $\lambda \in \Lambda$. This shows that $\overline{\boldsymbol{v}}$ is linearly independent and thus a basis. We conclude that $\#\Lambda = \dim_K(\overline{V})$ for any basis $\boldsymbol{v}$. □

COROLLARY 3.3.3. *A finitely generated free module has a finite basis.*

PROOF. If V is finitely generated, then V/MV is a finitely generated K-vector space, hence has finite dimension and therefore $\operatorname{rk}_A(V)$ is finite as well by the proof of Theorem 3.3.2. □

REMARK 3.3.4. The zero ring is an exception: if $A = 0$ then both the empty set $\emptyset$ and $\{0\}$ are bases of A as an A-module, and $A^{(\Lambda)} \simeq A = 0$ for any Λ, which is weird. That's why we excluded the zero ring in Theorem 3.3.2 (who cares about the zero ring anyways?).

REMARK 3.3.5. The conclusion of Theorem 3.3.2 does not necessarily hold over non-commutative rings! There are non-zero rings A such that $A^n \simeq A$ as A-modules for any $n > 0$ and so there is no well-defined rank of a free module. Rings for which free modules have a well-defined rank are said to satisfy the **invariant basis number** property.

Nothing new so far. But now here are a few observations that begin to tell us that not all things are like for vector spaces.

Example 3.3.6. It is not true that any generating set has a subset that is a basis: consider the free $\mathbb{Z}$-module $\mathbb{Z}$ and the generating set $\{2, 3\}$.

Example 3.3.7. It is not true that any linearly independent subset can be extended to a basis: consider the free $\mathbb{Z}$-module $\mathbb{Z}$ and the linearly independent subset $\{2\}$.

And here's the final blow: a module does not necessarily have a basis (otherwise we wouldn't need the notion of a "free" module of course). One particular problem that prevents a module from being free—and that cannot arise over a field—is torsion.

Definition 3.3.8. A **torsion element** of an A-module V is an element $v \in V$ such that there is a regular[1] element $a \in A$ with $av = 0$.

Example 3.3.9. Consider $\mathbb{Z}/n\mathbb{Z}$ as a $\mathbb{Z}$-module for $n > 1$. Then any $v \in \mathbb{Z}/n\mathbb{Z}$ is a torsion element since $nv = 0$.

Lemma 3.3.10. *The set* $\mathrm{T}(V)$ *of all torsion elements in* V *is a submodule of* V.

Proof. We clearly have $0 \in \mathrm{T}(V)$. Let $v, v' \in \mathrm{T}(V)$. Then there are regular elements $a, a' \in A$ with $av = 0$ and $a'v' = 0$. The product aa' is regular as well and

$$aa'(v + v') = aa'v + aa'v' = a'av + aa'v' = 0 \;,$$

so $v + v' \in \mathrm{T}(V)$. Also $a(-v) = -(av) = 0$, so $-v \in \mathrm{T}(V)$. Finally, for any $a'' \in A$ we have $a(a''v) = a''(av) = 0$, so $a''v \in \mathrm{T}(V)$. □

Definition 3.3.11. A module V with $\mathrm{T}(V) = \{0\}$ is called **torsion-free**.

Lemma 3.3.12. *A free module is torsion-free.*

Proof. Let V be a free A-module. Let $\boldsymbol{v} \coloneqq \{v_\lambda\}_{\lambda\in\Lambda}$ be a basis of V. Let $v \in \mathrm{T}(V)$ and let $a \in A$ be regular such that $av = 0$. We can write $v = \sum_{\lambda\in\Lambda} a_\lambda v_\lambda$ for certain $a_\lambda \in A$, so $0 = av = \sum_{\lambda\in\Lambda} aa_\lambda v_\lambda$. Since $\boldsymbol{v}$ is a basis, we must have $aa_\lambda = 0$ for all λ. Since a is regular, this forces $a_\lambda = 0$ for all λ, hence $v = 0$. □

Example 3.3.13. For $n > 1$ the $\mathbb{Z}$-module $\mathbb{Z}/n\mathbb{Z}$ has torsion, hence is not free.

Alright, when modules are not free in general, the next most natural general thing to consider would be minimal generating sets.

Definition 3.3.14. A generating set of an A-module V is **minimal** if it is minimal with respect to inclusion, i.e. one cannot remove a generator and still have a generating set. The **minimal generating number** of V is

$$\mu_A(V) \coloneqq \min\{\#\boldsymbol{v} \mid \boldsymbol{v} \text{ is a generating set of } V\} \;. \tag{3.43}$$

The minimal generating number is well-defined since cardinal numbers are well-ordered, so any set of cardinal numbers has a unique minimum.

Lemma 3.3.15. *Let* $A \neq 0$ *and let* V *be a free* A*-module. Then a basis of* V *is a minimal generating set and* $\mathrm{rk}_A(V) = \mu_A(V)$.

Proof. Let $\boldsymbol{v} \coloneqq \{v_\lambda\}_{\lambda\in\Lambda}$ be a basis. Suppose there is $\Lambda' \subsetneq \Lambda$ such that $\{v_\lambda\}_{\lambda\in\Lambda'}$ is still a generating set. Let $\mu \in \Lambda \setminus \Lambda'$. Then $v_\mu = \sum_{\lambda\in\Lambda'} a_\lambda v_\lambda$ for certain $a_\lambda \in A$. But then $v_\mu - \sum_{\lambda\in\Lambda'} a_\lambda v_\lambda = 0$, which contradicts the linear independence of $\boldsymbol{v}$. The claim that $\mathrm{rk}_A(V) = \mu_A(V)$ follows from the proof of Theorem 3.3.2. □

[1] Recall: this means non-zero-divisor.

Great, this sounds like a minimal generating set is a good generalization of a basis. (Un)fortunately, minimal generating sets do not behave like bases at all.

EXAMPLE 3.3.16. Consider the $\mathbb{Z}$-module $\mathbb{Z}$. Then $\{2,3\}$ is a minimal generating set because we cannot remove a generator and still have a generating set. But we have $\mu_{\mathbb{Z}}(\mathbb{Z}) = 1$ since $\{1\}$ is a generating set.

EXAMPLE 3.3.17. The $\mathbb{Z}$-module $\mathbb{Q}$ does not have a minimal generating set.[2] Suppose there is such a set $\boldsymbol{v}$. Recall that $\mathbb{Z}$-modules are the same thing as abelian groups. Take $v \in \boldsymbol{v}$ and let H be the submodule (subgroup) of $\mathbb{Q}$ generated by $\boldsymbol{v} \setminus \{v\}$. Because $\boldsymbol{v}$ is minimal, H is a proper subgroup of $\mathbb{Q}$. Let $G \coloneqq \mathbb{Q}/H$. Then G is non-trivial and cyclic since it is generated by the class of v in G. Hence, $G \simeq \mathbb{Z}/n\mathbb{Z}$ for some $n \in \mathbb{N}$. Note that for any non-zero $a \in \mathbb{Z}$ the multiplication map

$$\mu_a \colon \mathbb{Q} \to \mathbb{Q}\ , \quad x \mapsto ax\ , \tag{3.44}$$

is *surjective* (in general, abelian groups with this property are called **divisible**; you can generalize this concept to modules). This is obviously still true when we pass to a quotient, i.e. $\mu_a \colon G \to G$ should be surjective for any $a \in \mathbb{Z}$. But since $G = \mathbb{Z}/n\mathbb{Z}$, this is obviously not true—a contradiction!

Is torsion the only obstruction to being free? No!

EXAMPLE 3.3.18. The $\mathbb{Z}$-module $\mathbb{Q}$ is torsion-free. But it does not have a minimal generating set by Example 3.3.17, so from Lemma 3.3.15 we conclude that $\mathbb{Q}$ is not a free $\mathbb{Z}$-module.

We have constructed a hierarchy

$$\text{free} \subsetneq \text{torsion-free} \subsetneq \text{all modules} \tag{3.45}$$

and we'll refine this a bit more in the next sections. Note that we have strict inclusions only in general—there are rings where the hierarchy collapses, e.g. for fields.

I'll finish this section with more examples of odd behavior.

EXAMPLE 3.3.19. A submodule of a free module is not necessarily free: consider $A \coloneqq \mathbb{Z}/6\mathbb{Z}$ as a module over itself and take the submodule $U \coloneqq 2A$. This submodule cannot be free since $U \simeq A^{(\Lambda)}$ cannot work because of cardinality reasons.

EXAMPLE 3.3.20. A free proper submodule of a free module of finite rank does not necessarily have lower rank: consider $\mathbb{Z}$ as a $\mathbb{Z}$-module and take the submodule $2\mathbb{Z}$.

EXAMPLE 3.3.21. A submodule of a finitely generated module is not necessarily finitely generated: consider the polynomial ring $\mathbb{Z}[X_i \mid i \in \mathbb{N}]$ in infinitely many variables as a module over itself and take the submodule generated by $\{X_i \mid i \in \mathbb{N}\}$.

Exercises.

EXERCISE 3.3.22. Let A be a non-zero ring such that every ideal is a free A-module. Show that A is a principal ideal domain.

[2] I found this proof (which is much shorter than the one I had once in mind) on `https://math.stackexchange.com/questions/487820/additive-group-of-rationals-has-no-minimal-generating-set`.

3.4. Tensor products

We'll come to an important construction: tensor products. I'm not sure whether you've seen this in linear algebra already, so I'll review this here. It's basically just about bilinear maps and how to view them as linear maps—but it's a powerful tool!

Definition 3.4.1. Let A be a ring and let V_1, V_2, W be A-modules. A map $f\colon V_1 \times V_2 \to W$ is **A-bilinear** if it is A-linear in both components, i.e. the maps

$$f(v_1, -)\colon V_2 \to W \tag{3.46}$$

$$f(-, v_2)\colon V_1 \to W \tag{3.47}$$

are linear for all $v_1 \in V_1$ and $v_2 \in V_2$.

Let $\mathrm{Bil}_A(V_1, V_2; W)$ be the set of all A-bilinear maps $V_1 \times V_2 \to W$. This is an A-module with respect to pointwise operation. Here's a question: is there an A-module T such that

$$\mathrm{Bil}_A(V_1, V_2; W) \simeq \mathrm{Hom}_A(T, W) \quad ? \tag{3.48}$$

This would allow us to derive properties about bilinear maps from those of linear maps without creating any new theory. The answer is...yes!

Proposition 3.4.2. *Given A-modules V_1 and V_2, there is a pair (T, τ) consisting of an A-module T and an A-bilinear map $\tau\colon V_1 \times V_2 \to T$ which satisfies the following universal property: if W is an A-module and $f \in \mathrm{Bil}_A(V_1, V_2; W)$, then there is a unique morphism $\widetilde{f} \in \mathrm{Hom}_A(T, W)$ such that the diagram*

$$\begin{tikzcd} T \arrow[r, "\widetilde{f}"] & W \\ V_1 \times V_2 \arrow[u, "\tau"] \arrow[ur, "f"'] & \end{tikzcd} \tag{3.49}$$

commutes (one also says "f factors through τ"). The maps

$$\begin{array}{rcl} \mathrm{Bil}_A(V_1, V_2; W) & \simeq & \mathrm{Hom}_A(T, W) \\ f & \mapsto & \widetilde{f} \\ g \circ \tau & \mapsfrom & g\ , \end{array} \tag{3.50}$$

are mutually inverse isomorphisms.

Proof. Let C be the free A-module with basis $V_1 \times V_2$, i.e.

$$C := A^{(V_1 \times V_2)} = \bigoplus_{(v_1, v_2) \in V_1 \times V_2} A\ .$$

This is a huge module but that's okay. Let D be the submodule of C generated by elements of the following forms:

$$(v_1 + v_1', v_2) - (v_1, v_2) - (v_1', v_2)\ , \tag{3.51}$$

$$(v_1, v_2 + v_2') - (v_1, v_2) - (v_1, v_2')\ , \tag{3.52}$$

$$(av_1, v_2) - a(v_1, v_2)\ , \tag{3.53}$$

$$(v_1, av_2) - a(v_1, v_2)\ . \tag{3.54}$$

where $a \in A$, $v_1, v_1' \in V_1$, and $v_2, v_2' \in V_2$. Define $T := C/D$. We will write $v_1 \otimes v_2$ for the image of the basis element $(v_1, v_2) \in C$ in T. Then T is generated by the $v_1 \otimes v_2$ for $v_1 \in V_1$ and $v_2 \in V_2$ and we have:

$$(v_1 + v_1') \otimes v_2 = v_1 \otimes v_2 + v_1' \otimes v_2 \ , \tag{3.55}$$

$$v_1 \otimes (v_2 + v_2') = v_1 \otimes v_2 + v_1 \otimes v_2' \ , \tag{3.56}$$

$$(av_1) \otimes v_2 = a(v_1 \otimes v_2) = v_1 \otimes (av_2) \ . \tag{3.57}$$

It is then clear that the map

$$\tau \colon V_1 \times V_2 \to T \ , \quad (v_1, v_2) \mapsto v_1 \otimes v_2 \ , \tag{3.58}$$

is A-bilinear. Now, for any $f \in \mathrm{Bil}_A(V_1, V_2; W)$ we get a unique A-module morphism $\hat{f} \colon C \to W$ mapping the basis element (v_1, v_2) to $f(v_1, v_2)$. Since f is bilinear, we have $D \subseteq \mathrm{Ker}(\hat{f})$, so $\hat{f}$ induces a morphism

$$\widetilde{f} \colon T \to W \ , \quad v_1 \otimes v_2 \mapsto f(v_1, v_2) \ , \tag{3.59}$$

and it is clear that the diagram (3.49) commutes. The map $\widetilde{f}$ is also uniquely determined by this diagram. The last claim is clear by construction. □

As you have learned in Section 1.2, if there is a solution to a universal property problem, then it is already unique up to unique isomorphism.

Corollary 3.4.3. *The pair (T, τ) is unique up to unique isomorphism, i.e. if (T', τ') is another pair as in Proposition 3.4.2, then there is a unique isomorphism $j \colon T \to T'$ making the diagram*

$$\begin{tikzcd}
T \arrow[rr, "j"] & & T' \\
& V_1 \times V_2 \arrow[ul, "\tau"] \arrow[ur, "\tau'"'] &
\end{tikzcd}$$

commutative.

Definition 3.4.4. The (unique) pair (T, τ) associated to V_1 and V_2 in Proposition 3.4.2 is called the **tensor product** of V_1 and V_2, and one writes $V_1 \otimes_A V_2$ for the A-module T.

Keep in mind that in the proof of Proposition 3.4.2 we have given an explicit construction for the tensor product $V_1 \otimes_A V_2$, namely as the A-module generated by symbols $v_1 \otimes v_2$ for $v_1 \in V_1$ and $v_2 \in V_2$ satisfying the bilinearity relations (3.55) to (3.57), together with the map

$$\tau \colon V_1 \times V_2 \to V_1 \otimes_A V_2 \ , \quad (v_1, v_2) \mapsto v_1 \otimes v_2 \ . \tag{3.60}$$

The elements $v_1 \otimes v_2$ are called **elementary tensors**. It is *very* important to keep in mind that these are just generators of $V_1 \otimes_A V_2$; an arbitrary element of $V_1 \otimes_A V_2$ is not an elementary tensor but a linear combination of elementary tensors. Also, when you want to specify a morphism $V_1 \otimes_A V_2 \to W$ it's not sufficient to just say where the generators $v_1 \otimes v_2$ map to because you also have to check that your map satisfies the bilinearity relations (3.55) to (3.57), i.e. the map needs to be well-defined. It's best to start with a map $V_1 \times V_2 \to W$ and check that it is bilinear—then you get an induced map $V_1 \otimes_A V_2 \to W$ doing what you want. You'll see an example of this procedure in the proof of Lemma 3.4.6. When you work with tensor products you'll both use the explicit construction and the universal property—whatever works best.

Remark 3.4.5. Analogously, you can also consider **multilinear** maps $V_1 \times \ldots \times V_n \to W$ and construct a tensor product $V_1 \otimes_A \ldots \otimes_A V_n$ such that multilinear maps out of $V_1 \times \ldots \times V_n$ correspond to linear maps out of the tensor product.

The tensor product of modules really behaves like a product in a ring: it's commutative, associative, distributive, and has a unit—but all this is only up to canonical isomorphism of course. In fancy terms: A-Mod is a **tensor category**.

Lemma 3.4.6. *Let V_1, V_2, V_3 be A-modules. Then there are the canonical isomorphisms:*

1. $V_1 \otimes_A V_2 \to V_2 \otimes_A V_1$, $v_1 \otimes v_2 \mapsto v_2 \otimes v_1$;
2. $(V_1 \otimes_A V_2) \otimes_A V_3 \to V_1 \otimes_A (V_2 \otimes_A V_3) \to V_1 \otimes_A V_2 \otimes_A V_3$, $(v_1 \otimes v_2) \otimes v_3 \mapsto v_1 \otimes (v_2 \otimes v_3) \mapsto v_1 \otimes v_2 \otimes v_3$;
3. $(V_1 \oplus V_2) \otimes_A V_3 \to (V_1 \otimes_A V_3) \oplus (V_2 \otimes_A V_3)$, $(v_1, v_2) \otimes v_3 \mapsto (v_1 \otimes v_3, v_2 \otimes v_3)$;
4. $A \otimes_A V_1 \to V_1$, $a \otimes v_1 \mapsto a v_1$.

Proof. We'll only do this for $V_1 \otimes V_2 \to V_2 \otimes V_1$ here, the rest is proven analogously. Clearly, the map $f \colon V_1 \times V_2 \to V_2 \otimes_A V_1$ with $(v_1, v_2) \mapsto v_2 \otimes v_1$ is bilinear. We thus get an induced morphism $\widetilde{f} \colon V_1 \otimes_A V_2 \to V_2 \otimes_A V_1$ with $\widetilde{f}(v_1 \otimes v_2) = f(v_1, v_2) = v_2 \otimes v_1$. Analogously, we get a morphism $\widetilde{g} \colon V_2 \otimes V_1 \to V_1 \otimes V_2$ with $\widetilde{g}(v_2 \otimes v_1) = v_1 \otimes v_2$. By construction, we have $\widetilde{f}\widetilde{g} = \mathrm{id}$ and $\widetilde{g}\widetilde{f} = \mathrm{id}$. □

Lemma 3.4.7. *If V and W are free A-modules with bases $\boldsymbol{v} \coloneqq \{v_\lambda\}_{\lambda \in \Lambda}$ and $\boldsymbol{w} \coloneqq \{w_\sigma\}_{\sigma \in \Sigma}$, then $V \otimes_A W$ is a free A-module with basis*

$$\boldsymbol{v} \otimes \boldsymbol{w} \coloneqq \{v_\lambda \otimes w_\sigma\}_{\lambda \in \Lambda, \sigma \in \Sigma} \, . \tag{3.61}$$

In particular,

$$\mathrm{rk}_A(V \otimes_A W) = \mathrm{rk}_A(V) \cdot \mathrm{rk}_A(W) \, . \tag{3.62}$$

Proof. The bases give us isomorphisms $V \simeq A^{(\Lambda)}$ and $W \simeq A^{(\Sigma)}$. Generalizing Lemma 3.4.6, you can show that

$$\begin{aligned} V \otimes_A W \simeq A^{(\Lambda)} \otimes_A A^{(\Sigma)} = \left(\bigoplus_{\lambda \in \Lambda} A \right) \otimes_A \left(\bigoplus_{\sigma \in \Sigma} A \right) \simeq \bigoplus_{\substack{\lambda \in \Lambda \\ \sigma \in \Sigma}} A \otimes_A A \\ \simeq \bigoplus_{\substack{\lambda \in \Lambda \\ \sigma \in \Sigma}} A \simeq A^{(\Lambda \times \Sigma)} \, . \end{aligned}$$

If you go through the isomorphisms, you get the claimed basis of $V \otimes_A W$. □

So far, so nice. But there are some strange things that can happen with tensor products and you need to be careful. Here's a typical example.

Example 3.4.8. The $\mathbb{Z}$-module $(\mathbb{Z}/3\mathbb{Z}) \otimes_{\mathbb{Z}} (\mathbb{Z}/5\mathbb{Z})$ is 0. Namely, let $f \colon \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/5\mathbb{Z} \to W$ be a $\mathbb{Z}$-bilinear map to a $\mathbb{Z}$-module W. Then we need to have

$$3f(x,y) = f(3x,y) = f(0,y) = 0 \quad \text{and} \quad 5f(x,y) = f(x,5y) = f(x,0) = 0$$

for all $(x,y) \in \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/5\mathbb{Z}$. Hence,

$$f(x,y) = 1f(x,y) = (2 \cdot 3 + (-1) \cdot 5) f(x,y) = 2 \cdot 3f(x,y) + (-1)5f(x,y) = 0 \, ,$$

so every bilinear map out of $\mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/5\mathbb{Z}$ is 0, i.e. the tensor product is 0.

What can we do with the tensor product? Recall from Example 3.1.9 the scalar restriction functor

$$(-)_A \colon B\text{-}\mathsf{Mod} \to A\text{-}\mathsf{Mod} \tag{3.63}$$

associated to a ring morphism $\varphi\colon A \to B$. You may have immediately asked yourself whether there's an "inverse" to this construction. The answer is "yes" and now we can give one. Let V be an A-module. We can view B as an A-module via φ and then the tensor product

$$V^B \coloneqq B \otimes_A V \tag{3.64}$$

is naturally a B-module via

$$b(b' \otimes v) \coloneqq (bb') \otimes v \tag{3.65}$$

for $b, b' \in B$ and $v \in V$. Moreover, if $f\colon V \to W$ is a morphism of A-modules, we get an induced morphism

$$f^B \colon V^B \to W^B \;, \quad b \otimes v \mapsto b \otimes f(v) \;. \tag{3.66}$$

In total, we have defined a functor

$$(-)^B \colon A\text{-}\mathsf{Mod} \to B\text{-}\mathsf{Mod} \tag{3.67}$$

which is called **scalar extension**. Again, note that we have dropped φ from the notation and that we do not necessarily really have an "extension" of scalars, e.g. $\mathbb{Z}/3\mathbb{Z} \otimes_{\mathbb{Z}} \mathbb{Z}/5\mathbb{Z} = 0$ as we have seen in Example 3.4.8.

Example 3.4.9. If V is an $\mathbb{R}$-vector space, one often considers its **complexification** $\mathbb{C} \otimes_{\mathbb{R}} V$. If $\boldsymbol{v}$ is an $\mathbb{R}$-basis of V, then $1 \otimes \boldsymbol{v}$ is a $\mathbb{C}$-basis of $\mathbb{C} \otimes_R V$.

Now, what about scalar extension being an "inverse" to scalar restriction? Notice that for an A-module V and a B-module W we have canonical isomorphisms

$$\begin{array}{rcl} \operatorname{Hom}_B(V^B, W) & \simeq & \operatorname{Hom}_A(V, W_A) \\ (f\colon V^B \to W) & \mapsto & (v \mapsto f(1 \otimes v)) \\ (b \otimes v \mapsto bf(v)) & \leftarrow\!\shortmid & (f\colon V \to W_A) \end{array} \tag{3.68}$$

of A-modules. This is an example of an extremely important categorical concept: if we have functors $F\colon \mathcal{C} \to \mathcal{D}$ and $G\colon \mathcal{D} \to \mathcal{C}$ between two categories $\mathcal{C}$ and $\mathcal{D}$ such that there is an isomorphism

$$\operatorname{Hom}_{\mathcal{D}}(F(-), -) \simeq \operatorname{Hom}_{\mathcal{C}}(-, G(-)) \tag{3.69}$$

of (bi-)functors (i.e. there are isomorphisms for any objects you plug in for the blanks and these isomorphisms are compatible with morphisms), then (F, G) is said to be a pair of **adjoint functors**. It is not true in general that G is some sort of inverse of F, but at least for morphisms there is a nice relationship between F and G, and there are countless of examples of adjoint functors. We have just shown that $(-)^B$ and $(-)_A$ is a pair of adjoint functors.

Exercises.

Exercise 3.4.10. Let $\varphi\colon A \to B$ be a ring morphism and let V be an A-module. Show that if V is free with basis $\boldsymbol{v}$, then V^B is a free B-module with basis $1 \otimes \boldsymbol{v}$.

Exercise 3.4.11. Let I be an ideal in a ring A and let V be an A-module. Show that there is a canonical isomorphism

$$V/IV \simeq (A/I) \otimes_A V \tag{3.70}$$

of (A/I)-modules, where the scalar extension is taken with respect to the quotient map $A \to A/I$.

Exercise 3.4.12. Show that $(\mathbb{Z}/m\mathbb{Z}) \otimes_{\mathbb{Z}} (\mathbb{Z}/n\mathbb{Z}) \simeq \mathbb{Z}/d\mathbb{Z}$ for $m, n \in \mathbb{N}$ and $d \coloneqq \gcd(m, n)$.

Exercise 3.4.13. Let V be a $\mathbb{Z}$-module with $\mathrm{T}(V) = V$, e.g. a finite abelian group or $\mathbb{Q}/\mathbb{Z}$. Show that $\mathbb{Q} \otimes_{\mathbb{Z}} V = 0$, i.e. scalar extension to $\mathbb{Q}$ "kills torsion".

Exercise 3.4.14. Let R be a ring. Show that there is a canonical R-algebra isomorphism $R[X_1, X_2] \simeq R[X_1] \otimes_R R[X_2]$. (Question: How do you equip the tensor product of two R-algebras with an R-algebra structure?)

Exercise 3.4.15. Let $\varphi\colon A \to B$ be a ring morphism and let V and W be two A-modules. Show that there is a canonical isomorphism

$$(V \otimes_A W)^B \simeq V^B \otimes_B W^B \tag{3.71}$$

of B-modules.

3.5. Exact sequences and exact functors

We want to refine our hierarchy of modules and to this end, we'll use a general machinery. A core theme of ring and module theory is **homological algebra**: the study of exact sequences and how functors act on them. Consider a **sequence**

$$\cdots \xrightarrow{f_{i-2}} V_{i-1} \xrightarrow{f_{i-1}} V_i \xrightarrow{f_i} V_{i+1} \xrightarrow{f_{i+1}} \cdots \tag{3.72}$$

of A-module morphisms indexed by integers i in some interval I. If $i \in I$ is such that also $i+1 \in I$, then the sequence is called **exact at position** i if

$$\mathrm{Im}(f_i) = \mathrm{Ker}(f_{i+1}) \ , \tag{3.73}$$

and it is called **exact** if it is exact at all positions $i \in I$ with $i + 1 \in I$.

Example 3.5.1. The sequence $0 \to V \xrightarrow{f} W$ is exact if and only if f is injective.

Example 3.5.2. The sequence $V \xrightarrow{f} W \to 0$ is exact if and only if f is surjective.

Example 3.5.3. A **short exact sequence** is an exact sequence of the form

$$0 \longrightarrow U \xrightarrow{j} V \xrightarrow{q} W \longrightarrow 0 \ . \tag{3.74}$$

This means that j is injective, q is surjective, and $W \simeq V/\mathrm{Im}(j)$. Conversely, for any submodule U of a module V we get a short exact sequence $0 \to U \to V \to V/U \to 0$.

We want to introduce some terminology to study how functors $F\colon A\text{-}\mathsf{Mod} \to \mathcal{A}$ into some category $\mathcal{A}$ act on short exact sequences. To make this work, we need to be able to talk about exact sequences in $\mathcal{A}$ as well, so we assume $\mathcal{A}$ is a category of modules over some ring as well, usually $\mathcal{A} = \mathsf{Ab}$ is the category of abelian groups (there's an abstract notion of **abelian categories** where you can take kernels, images, etc., and speak about exact sequences). Moreover, we want F to be **additive**, which means that F preserves finite direct sums, i.e.

$$F\Big(\bigoplus_{i=1}^{n} V_i\Big) \simeq \bigoplus_{i=1}^{n} F(V_i) \ . \tag{3.75}$$

This isomorphism needs to be canonical and compatible with the projections and inclusions (I leave it up to you to formalize this). In particular,

$$F(0) \simeq 0\,. \tag{3.76}$$

Example 3.5.4. For an A-module V we have a functor

$$\begin{array}{rcl} V \otimes_A -\colon A\text{-}\mathsf{Mod} & \to & A\text{-}\mathsf{Mod} \\ W & \mapsto & V \otimes_A W \\ (f\colon W \to W') & \mapsto & (v \otimes w \mapsto v \otimes f(w))\,. \end{array} \tag{3.77}$$

It follows from Lemma 3.4.6 that this functor is additive.

Example 3.5.5. Recall from Example 1.2.7 that for any A-module V we have a functor $\mathrm{Hom}_A(V,-)\colon A\text{-}\mathsf{Mod} \to \mathsf{Set}$. Recall from Example 3.1.8 that $\mathrm{Hom}_A(V,W)$ is naturally an A-module itself. It's straightforward to check that for an A-module morphism f the induced map $\mathrm{Hom}_A(V,f) = f \circ -$ is an A-module morphism as well, so $\mathrm{Hom}_A(V,-)$ is actually a functor $A\text{-}\mathsf{Mod} \to A\text{-}\mathsf{Mod}$. Moreover, this functor is easily seen to be additive (check also Exercise 3.2.4).

Now, if $F\colon A\text{-}\mathsf{Mod} \to \mathcal{A}$ is an additive functor, then for every short exact sequence as in (3.74) we get an induced sequence

$$0 \longrightarrow F(U) \xrightarrow{F(j)} F(V) \xrightarrow{F(q)} F(W) \longrightarrow 0\,. \tag{3.78}$$

If this sequence were exact, then we could immediately deduce some nice properties of F, for example that F commutes with taking quotients:

$$F(V/U) \simeq F(V)/F(U)\,. \tag{3.79}$$

The problem is that in general the induced sequence is not necessarily exact as we will see in several examples. To describe what F is doing to short exact sequences, we introduce the following terminology:

Definition 3.5.6. The functor F is called:

(1) **exact** if $0 \to F(U) \to F(V) \to F(W) \to 0$ is exact for every ses (3.74);
(2) **left-exact** if $0 \to F(U) \to F(V) \to F(W)$ is exact for every ses (3.74);
(3) **right-exact** if $F(U) \to F(V) \to F(W) \to 0$ is exact for every ses (3.74).

Lemma 3.5.7. *For every A-module V the functor $V \otimes_A -\colon A\text{-}\mathsf{Mod} \to A\text{-}\mathsf{Mod}$ is right-exact.*

Proof. We need to show that for every short exact sequence

$$0 \longrightarrow W' \xrightarrow{f} W \xrightarrow{g} W'' \longrightarrow 0 \tag{3.80}$$

of A-modules the induced sequence

$$V \otimes_A W' \xrightarrow{V \otimes_A f} V \otimes_A W \xrightarrow{V \otimes_A g} V \otimes_A W'' \longrightarrow 0 \tag{3.81}$$

is exact. To simplify notation, we set $\widetilde{f} \coloneqq V \otimes_A f$ and $\widetilde{g} \coloneqq V \otimes_A g$. Recall that $\widetilde{f}(v \otimes w') = v \otimes f(w')$ and $\widetilde{g}(v \otimes w) = v \otimes g(w)$. So, we need to show that $\widetilde{g}$ is surjective and that $\mathrm{Im}(\widetilde{f}) = \mathrm{Ker}(\widetilde{g})$.

Let's first show that $\widetilde{g}$ is surjective. An element of $V \otimes_A W''$ is of the form $\sum_{i=1}^n a_i(v_i \otimes w_i'')$. Since (3.80) is exact, g is surjective and so there is $w_i \in W$ with $g(w_i) = w_i''$. We then have

$$\widetilde{g}\left(\sum_{i=1}^n a_i(v_i \otimes w_i)\right) = \sum_{i=1}^n a_i(v_i \otimes g(w_i)) = \sum_{i=1}^n a_i(v_i \otimes w_i'') \ ,$$

i.e. $\widetilde{g}$ is surjective.

Next, we show that $\operatorname{Im}(\widetilde{f}) \subseteq \operatorname{Ker}(\widetilde{g})$. Since $\operatorname{Im}(f) \subseteq \operatorname{Ker}(g)$ we have

$$\widetilde{g} \circ \widetilde{f}(v \otimes w') = \widetilde{g}(v \otimes f(w')) = v \otimes g(f(w')) = 0 \ .$$

Hence, $\widetilde{g} \circ \widetilde{f} = 0$, so $\operatorname{Im}(\widetilde{f}) \subseteq \operatorname{Ker}(\widetilde{g})$.

Finally, we show that $\operatorname{Im}(\widetilde{f}) \supseteq \operatorname{Ker}(\widetilde{g})$. This is the most difficult part. Note that it is not sufficient to show that $\widetilde{g}(v \otimes w) = 0$ implies $v \otimes w \in \operatorname{Im}(\widetilde{f})$. We prove the claim as follows. Since we already know that $\operatorname{Im}(\widetilde{f}) \subseteq \operatorname{Ker}(\widetilde{g})$, the morphism $\widetilde{g}$ induces a morphism

$$\overline{\widetilde{g}}\colon (V \otimes_A W)/\operatorname{Im}(\widetilde{f}) \to V \otimes_A W'' \ . \tag{3.82}$$

This map is surjective since we have already proven that $\widetilde{g}$ is surjective. If we can show that $\overline{\widetilde{g}}$ is an isomorphism, then we can conclude that

$$0 = \operatorname{Ker}(\overline{\widetilde{g}}) = \operatorname{Ker}(\widetilde{g})/\operatorname{Im}(\widetilde{f}) \ ,$$

so $\operatorname{Im}(\widetilde{f}) = \operatorname{Ker}(\widetilde{g})$. To prove that (3.82) is an isomorphism, we'll construct an inverse as follows. For $(v, w'') \in V \times W''$ we can choose $w \in W$ with $g(w) = w''$ since g is surjective. This yields a map

$$V \times W'' \to (V \otimes_A W)/\operatorname{Im}(\widetilde{f}), \quad (v, w'') \mapsto \overline{v \otimes w} \ . \tag{3.83}$$

This map is well-defined since if $w_1, w_2 \in W$ with $g(w_1) = g(w_2)$, then $w_1 - w_2 \in \operatorname{Ker}(g) = \operatorname{Im}(f)$, so $v \otimes (w_1 - w_2) \in \operatorname{Im}(\widetilde{f})$ and therefore $\overline{v \otimes w_1} = \overline{v \otimes w_2}$. The map is moreover bilinear, it thus induces a morphism $h\colon V \otimes_A W'' \to (V \otimes_A W)/\operatorname{Im}(\widetilde{f})$. We have

$$\overline{\widetilde{g}} \circ h(v \otimes w'') = \overline{\widetilde{g}}(\overline{v \otimes w}) = v \otimes g(w) = v \otimes w'' \ ,$$

hence $\overline{\widetilde{g}} \circ h = \mathrm{id}$. Moreover,

$$h \circ \overline{\widetilde{g}}(\overline{v \otimes w}) = h(v \otimes g(w)) = \overline{v \otimes w} \ ,$$

since $w \in g^{-1}(g(w))$. Hence, $h \circ \overline{\widetilde{g}} = \mathrm{id}$. □

Example 3.5.8. The functor $V \otimes_A -$ is not necessarily exact. Consider the sequence $0 \to \mathbb{Z} \xrightarrow{\mu} \mathbb{Z} \to \mathbb{Z}/n\mathbb{Z} \to 0$ of $\mathbb{Z}$-modules, where μ is the multiplication map by an integer $n > 1$. Applying $(\mathbb{Z}/n\mathbb{Z}) \otimes_{\mathbb{Z}} -$ to this sequence yields the exact sequence

$$(\mathbb{Z}/n\mathbb{Z}) \otimes_{\mathbb{Z}} \mathbb{Z} \xrightarrow{(\mathbb{Z}/n\mathbb{Z}) \otimes_{\mathbb{Z}} \mu} (\mathbb{Z}/n\mathbb{Z}) \otimes_{\mathbb{Z}} \mathbb{Z} \longrightarrow (\mathbb{Z}/n\mathbb{Z}) \otimes_{\mathbb{Z}} (\mathbb{Z}/n\mathbb{Z}) \longrightarrow 0$$

But the map $(\mathbb{Z}/n\mathbb{Z}) \otimes_{\mathbb{Z}} \mu$ is not injective since it sends $\overline{1} \otimes 1$ to $\overline{1} \otimes n = n(\overline{1} \otimes 1) = \overline{n} \otimes 1 = 0$.

Exercises.

Exercise 3.5.9. Show that taking the torsion submodule defines a left-exact functor $A\text{-}\mathsf{Mod} \to A\text{-}\mathsf{Mod}$.

Exercise 3.5.10. Show that for any two submodules U, U' of an A-module V there is a canonical short exact sequence

$$0 \longrightarrow U \cap U' \longrightarrow U \oplus U' \longrightarrow U + U' \longrightarrow 0 \, . \tag{3.84}$$

Exercise 3.5.11. The **cokernel** of an A-module morphism $f\colon V \to W$ is defined as

$$\operatorname{Coker}(f) := W/\operatorname{Im}(f) \, . \tag{3.85}$$

Show that there is a canonical exact sequence

$$0 \longrightarrow \operatorname{Ker}(f) \longrightarrow V \xrightarrow{f} W \longrightarrow \operatorname{Coker}(f) \longrightarrow 0 \, . \tag{3.86}$$

3.6. Flat modules

Now that you have learned that $V \otimes_A -\colon A\text{-}\mathsf{Mod} \to A\text{-}\mathsf{Mod}$ is not exact in general, you can ask: well, for which V is it exact? This sorts out some nice modules!

Definition 3.6.1. An A-module V is called **flat** if $V \otimes_A -$ is an exact functor.

In Example 3.5.8 we have seen that the $\mathbb{Z}$-module $\mathbb{Z}/n\mathbb{Z}$ for $n > 1$ is not flat. The problem is again torsion.

Lemma 3.6.2. *Flat modules are torsion-free.*

Proof. Let V be a flat A-module. Let $x \in A$ be regular. Then the multiplication map $\mu_x\colon A \to A$ mapping a to ax is injective. Since V is flat, also

$$\begin{tikzcd}
V \otimes_A A \arrow[r, "V\otimes_A \mu_x"] \arrow[d, "\simeq"] & V \otimes_A A \arrow[d, "\simeq"] \\
V \arrow[r, "v \mapsto xv"] & V
\end{tikzcd}$$

is injective. Hence, V is torsion-free. □

Lemma 3.6.3. *Free modules are flat.*

Proof. Let V be a free A-module and choose an isomorphism $V \to A^{(\Lambda)}$. If $0 \to W' \xrightarrow{f} W \xrightarrow{g} W'' \to 0$ is a short exact sequence, then we obtain a commutative diagram

$$\begin{tikzcd}
0 \arrow[r] & V \otimes_A W' \arrow[r] \arrow[d, "\simeq"] & V \otimes_A W \arrow[r] \arrow[d, "\simeq"] & V \otimes_A W'' \arrow[r] \arrow[d, "\simeq"] & 0 \\
0 \arrow[r] & A^{(\Lambda)} \otimes_A W' \arrow[r] \arrow[d, "\simeq"] & A^{(\Lambda)} \otimes_A W \arrow[r] \arrow[d, "\simeq"] & A^{(\Lambda)} \otimes_A W'' \arrow[r] \arrow[d, "\simeq"] & 0 \\
0 \arrow[r] & \bigoplus_{\lambda\in\Lambda} A \otimes_A W' \arrow[r] \arrow[d, "\simeq"] & \bigoplus_{\lambda\in\Lambda} A \otimes_A W \arrow[r] \arrow[d, "\simeq"] & \bigoplus_{\lambda\in\Lambda} A \otimes_A W'' \arrow[r] \arrow[d, "\simeq"] & 0 \\
0 \arrow[r] & \bigoplus_{\lambda\in\Lambda} W' \arrow[r, "\bigoplus_\lambda f"] & \bigoplus_{\lambda\in\Lambda} W \arrow[r, "\bigoplus_\lambda g"] & \bigoplus_{\lambda\in\Lambda} W'' \arrow[r] & 0
\end{tikzcd}$$

The bottom sequence is exact since it is just a direct sum of copies of the exact sequence we started with. Hence, also the top sequence is exact and therefore V is flat. □

Example 3.6.4. There are flat modules which are not free. We know from Example 3.3.18 that $\mathbb{Q}$ is not a free $\mathbb{Z}$-module. But it is flat. Let $f\colon W' \to W$ be an injective morphism of $\mathbb{Z}$-modules. An arbitrary element of $\mathbb{Q} \otimes_{\mathbb{Z}} W'$ is of the form $\sum_{i=1}^n \frac{r_i}{s_i} \otimes w_i'$ with $r_i \in \mathbb{Z}$, $0 \neq s_i \in \mathbb{Z}$, and $w_i' \in W'$. Suppose that

$$0 = \mathbb{Q} \otimes_{\mathbb{Z}} f \left(\sum_{i=1}^n \frac{r_i}{s_i} \otimes w_i' \right) = \sum_{i=1}^n \frac{r_i}{s_i} \otimes f(w_i') \ .$$

Let $s \coloneqq \prod_{i=1}^n s_i$ and $s_i' \coloneqq \prod_{j \neq i} s_j$. Then $s \neq 0$ and multiplying the above equation by s yields

$$0 = \sum_{i=1}^n r_i s_i' f(w_i') = f(\sum_{i=1}^n r_i s_i' w_i') \ .$$

Since f is injective, it follows that $\sum_{i=1}^n r_i s_i' w_i' = 0$. Multiplying by $0 \neq s^{-1}$ yields $\sum_{i=1}^n \frac{r_i}{s_i} \otimes w_i' = 0$. Hence, $\mathbb{Q} \otimes_{\mathbb{Z}} f$ is injective and therefore $\mathbb{Q}$ is a flat $\mathbb{Z}$-module.

Example 3.6.5. There are torsion-free modules which are not flat. Let K be a field, let $A \coloneqq K[X_1, X_2]$, and let $V \coloneqq (X_1, X_2)$. Clearly, V is torsion-free. Let $B \coloneqq A/(X_1) \simeq K[X_2]$ and let $q\colon A \to B$ be the quotient map. Suppose that V would be a flat A-module. Then V^B would be a flat B-module by Exercise 3.6.7. But we have

$$\begin{aligned} V^B &= B \otimes_A (X_1, X_2) = A/(X_1) \otimes_A (X_1, X_2) \\ &\overset{*}{\simeq} (X_1, X_2)/(X_1)(X_1, X_2) \simeq (X_1, X_2)/(X_1^2, X_1 X_2) \ , \end{aligned}$$

where the isomorphism * comes from Exercise 3.4.11. We now see that X_1 is a torsion-element in V^B, so this module has torsion and therefore it cannot be flat by Lemma 3.6.2.

Hence, we have refined our module hierarchy into

$$\text{free} \subsetneq \text{flat} \subsetneq \text{torsion-free} \subsetneq \text{all modules} \ . \tag{3.87}$$

Remark 3.6.6. Why are flat modules called "flat"? It's a bit difficult going into the details but this comes from geometry. Let $\varphi\colon A \to B$ be a ring morphism. We get an induced map

$$\begin{tikzcd} \mathrm{Spec}(B) \arrow[d, "\varphi^*"] \\ \mathrm{Spec}(A) \end{tikzcd}$$

You know that you can view $\mathrm{Spec}(A)$ and $\mathrm{Spec}(B)$ as describing zero sets (in some general sense). Using the above morphism we can decompose $\mathrm{Spec}(B)$ into the fibers (preimages of a point) of φ^*. Geometrically, this means that over each point of $\mathrm{Spec}(A)$ there's a selection of points of $\mathrm{Spec}(B)$. So, you can view $\mathrm{Spec}(B)$ as a family of zero sets varying over the points of $\mathrm{Spec}(A)$. Now, if B is a flat A-module, the fibers in the family "vary smoothly" with the points of $\mathrm{Spec}(A)$. This is basically the intuition but there are examples where this intuition is not really correct. Flatness is one of the few notions in algebraic geometry that is motivated by algebra and not by geometry!

Exercises.

Exercise 3.6.7. Let $\varphi\colon A \to B$ be a ring morphism. Show that if V is a flat A-module, then V^B is a flat B-module.

3.7. Projective modules

We'll chuck in another class of modules into the hierarchy between free and flat. I only give one definition and leave the work to you this time (it's a nice project to work on)!

Definition 3.7.1. An A-module P is called **projective** if $\operatorname{Hom}_A(P, -)$ is exact.

From the exercises below you will conclude that we have a hierarchy

$$\text{free} \subsetneq \text{projective} \subsetneq \text{flat} \subsetneq \text{torsion-free} \subsetneq \text{all modules} . \tag{3.88}$$

Exercises.

Exercise 3.7.2. Show that the functor $\operatorname{Hom}_A(V, -)\colon A\text{-}\mathsf{Mod} \to A\text{-}\mathsf{Mod}$ is left-exact.

Exercise 3.7.3. Show that the functor $\operatorname{Hom}_A(V, -)\colon A\text{-}\mathsf{Mod} \to A\text{-}\mathsf{Mod}$ is in general not exact. Hint: consider $\operatorname{Hom}_{\mathbb{Z}}(\mathbb{Z}/2\mathbb{Z}, -)$ and the quotient map $q\colon \mathbb{Z} \to \mathbb{Z}/4\mathbb{Z}$.

Exercise 3.7.4. Prove that the following are equivalent for an A-module P:

(1) P is projective;

(2) P is a direct summand of a free A-module, i.e. there is an A-module Q such that $P \oplus Q \simeq A^{(\Lambda)}$ for some Λ;

(3) Every short exact sequence of the form

$$0 \longrightarrow V' \longrightarrow V \xrightarrow{g} P \longrightarrow 0 \tag{3.89}$$

splits, i.e. there is an A-module morphism $s\colon P \to V$ such that $gs = \mathrm{id}_P$ (such an s is called a **section** of g);

(4) For every morphism $h\colon P \to W$ and every *surjective* morphism $f\colon V \to W$ there is a morphism $\widetilde{h}\colon P \to V$ such that the diagram

$$\begin{tikzcd} & V \arrow[d, two heads, "f"] \\ P \arrow[ur, dashed, "\widetilde{h}"] \arrow[r, "h"] & W \end{tikzcd} \tag{3.90}$$

commutes.

Exercise 3.7.5. Show that every free module is projective.

Exercise 3.7.6. Show that projective modules are flat.

Exercise 3.7.7. Let $\varphi\colon A \to B$ be a ring morphism. Show that if V is a projective A-module, then V^B is a projective B-module.

Exercise 3.7.8. Show that there are projective modules which are not free. Hint: let $A := \mathbb{Z}/6\mathbb{Z}$ and consider the A-module $\mathbb{Z}/2\mathbb{Z}$.

Exercise 3.7.9. Show that there are flat modules which are not projective. Hint: consider the $\mathbb{Z}$-module $\mathbb{Q}$.

3.8. Specialties about finitely generated modules

Most of the time we will work with finitely generated modules. In this section, we will discuss some specialties about such modules which in general don't hold for arbitrary modules.

Recall Krull's theorem (Theorem 2.3.6) about the existence of a maximal ideal in a ring A containing a given proper ideal. Note that A as an A-module is finitely generated and that submodules of A are exactly the ideals. The following is a generalization of Krull's theorem to finitely generated modules.

Lemma 3.8.1. *If V is a finitely generated A-module, then every proper submodule of V is contained in a maximal submodule.*

Proof. Let $\{v_1, \ldots, v_n\}$ be a generating set of V and let $U \subsetneq V$ be a proper submodule. Let Σ be the set of all proper submodules of V containing U. Then $U \in \Sigma$, so $\Sigma \neq \emptyset$. Let $(U_\lambda)_{\lambda \in \Lambda}$ be a chain in Σ. Then $U' \coloneqq \bigcup_{\lambda \in \Lambda} U_\lambda$ is a submodule of V (as for ideals, this works because we take the union of a chain). We claim that U' is a proper submodule. Suppose that $U' = V$. For every $i = 1, \ldots, n$ there is λ_i such that $v_i \in U_{\lambda_i}$. Since the U_λ form a chain, we can find a largest U_λ among the U_{λ_i}. Then $v_1, \ldots, v_n \in U_\lambda$, so $V = U_\lambda$, which is a contradiction to $U_\lambda \in \Sigma$ and thus being a proper submodule. Hence, $U' \in \Sigma$. We can now apply Zorn's lemma to get a maximal element in Σ, and this is a maximal submodule in V containing U. □

Remark 3.8.2. One can show that the $\mathbb{Z}$-module $\mathbb{Q}$ does not have a maximal submodule, so Lemma 3.8.1 is not necessarily true when we drop finitely generated.

Maybe you remember the **Cayley–Hamilton theorem** from linear algebra: when you have a matrix $\boldsymbol{M}$ and plug it into its characteristic polynomial p, you get zero, i.e. $p(\boldsymbol{M}) = 0$. The following theorem is a generalization of this in module-theoretic terms.

Theorem 3.8.3 (Cayley–Hamilton). *Let V be a finitely generated A-module, generated by n elements. Let $I \trianglelefteq A$ be an ideal and let $f \in \mathrm{End}_A(V) \coloneqq \mathrm{Hom}_A(V, V)$ be an endomorphism with $f(V) \subseteq IV$. Then there is a* monic *polynomial*

$$p = X^n + a_1 X^{n-1} + \ldots + a_n \in A[X] \tag{3.91}$$

with

$$0 = p(f) = f^n + a_1 f^{n-1} + \ldots + a_n \; . \tag{3.92}$$

Moreover, $a_i \in I^i$.

For the proof, we will need an elementary theorem about matrices. First note that we can define and do the basic arithmetic with matrices over any ring in the same way as you are used to over a field. Let $\boldsymbol{M}$ be an $(n \times n)$-matrix over a ring A. Then we can define its determinant $\det(\boldsymbol{M}) \in A$ in the usual way as a sum over the permutations in the symmetric group S_n (the Laplace formula), and this satisfies the usual rules and you can compute it via expansion along a row or column. Now, something you may not know is the so-called **adjugate** matrix $\mathrm{adj}(\boldsymbol{M})$. The (i,j) entry of this matrix is $(-1)^{i+j}$ multiplied by the determinant of the $(n-1) \times (n-1)$ submatrix of $\boldsymbol{M}$ obtained by deleting row i and column j. Using expansion for the determinant you can then prove[3] that

$$\operatorname{adj}(\boldsymbol{M}) \cdot \boldsymbol{M} = \det(\boldsymbol{M})\mathbf{1}_n \; . \tag{3.93}$$

Proof of Theorem 3.8.3. Let $\{v_1, \ldots, v_n\}$ be a generating set of V. We can write

$$f(v_i) = \sum_{j=1}^{n} a_{ij} v_j \tag{3.94}$$

for certain $a_{ij} \in I$. Define the $(n \times n)$-matrix $\boldsymbol{M} \coloneqq (a_{ij})$ and let $\boldsymbol{v} \coloneqq (v_1, \ldots, v_n) \in V^n$. As in Example 3.1.7, we can view V as an $A[X]$-module with X acting by f, i.e. $Xv = f(v)$ for $v \in V$. We can then rewrite (3.94) as

$$\underbrace{(X\mathbf{1}_n - \boldsymbol{M})}_{\text{matrix over } A[X]} \boldsymbol{v} = 0 \; , \tag{3.95}$$

where $\mathbf{1}_n$ is the identity matrix of size n. Now, we multiply this equation with the adjugate matrix of $X\mathbf{1}_n - \boldsymbol{M} \in \operatorname{Mat}_n(A[X])$ and obtain

$$\underbrace{\det(X\mathbf{1}_n - \boldsymbol{M})}_{\coloneqq p \in A[X]} \boldsymbol{v} = 0 \; . \tag{3.96}$$

This means $pv_i = 0$ for all i, hence $p(f) = 0$ since X acts via f. By the Laplace formula for the determinant, the polynomial p is monic and the coefficient of X^{n-i} is contained in I^i. □

We will now derive a whole series of corollaries that are used frequently.

Corollary 3.8.4. *Let V be a finitely generated A-module and let $f \colon V \to V$ be an A-module morphism. If f is surjective, then f is already bijective.*

Proof. As in the proof of Theorem 3.8.3 we consider V as an $A[X]$-module with X acting via f. Let $I \coloneqq (X) \trianglelefteq A[X]$. Since f is surjective, we have $IV = V$. The Cayley–Hamilton theorem applied to the $A[X]$-module morphism $\mathrm{id}_V \colon V \to V$ gives us a polynomial

$$p = Y^n + a_1 Y^{n-1} + \ldots + a_n \in (A[X])[Y] \tag{3.97}$$

with $a_i \in I$ for all i and $p(\mathrm{id}_V) = 0$, i.e.

$$0 = \mathrm{id}_V^n + a_1 \mathrm{id}_V^{n-1} + \ldots + a_n \mathrm{id}_V^0 = \mathrm{id}_V + a_1 \mathrm{id}_V + \ldots + a_n \mathrm{id}_V \; . \tag{3.98}$$

Since $a_i \in I = (X)$, there are $a_1', \ldots, a_n' \in A[X]$ with $a_i' X = a_i$. Let

$$q \coloneqq -(a_1' + \ldots + a_n') \in A[X] \; . \tag{3.99}$$

Then $0 = (1 - qX)v$, i.e. $v = (qX)v$, for all $v \in V$. Since X acts via f, this means

$$0 = 1 - q(f)f \in \operatorname{End}_A(V) \; , \tag{3.100}$$

so $q(f)f = 1$. But this means f has an inverse, namely $q(f)$. □

Corollary 3.8.5. *Let V be a finitely generated free A-module and let $n \coloneqq \mathrm{rk}_A(V)$. Then every generating system of V consisting of n elements is already a basis.*

[3] You can find a short proof at `https://proofwiki.org/wiki/Matrix_Product_with_Adjugate_Matrix`.

Proof. Let $\{v_1, \ldots, v_n\}$ be a generating set. This defines a surjective A-module morphism $f\colon A^n \to V$. Since V is free of rank n, there is an isomorphism $g\colon V \to A^n$. We thus get a surjective morphism $gf\colon A^n \to A^n$. By Corollary 3.8.4, this is already an isomorphism. But then also $f = g^{-1}(gf)$ is an isomorphism and $\{v_1, \ldots, v_n\}$ is a basis. □

Corollary 3.8.6. *Let V be a finitely generated A-module and let $I \trianglelefteq A$ be an ideal with $IV = V$. Then there is $a \in A$ with $a \equiv 1 \bmod I$ and $aV = 0$.*

Proof. Let $f \coloneqq \mathrm{id}_V \in \mathrm{End}_A(V)$. Then by Theorem 3.8.3 there is $p = X^n + a_1X^{n-1} + \ldots + a_n \in A[X]$ with $p(f) = 0$ and $a_i \in I$ for all i. Then $a \coloneqq 1 + a_1 + \ldots + a_n$ satisfies the claimed properties. □

The next corollary is called **Nakayama's lemma**—it is a simple but fundamental result in commutative algebra. You need to recall the Jacobson radical $\mathrm{Jac}(A)$ of a ring A from Definition 2.6.7.

Corollary 3.8.7 (Nakayama). *Let V be a finitely generated A-module and $I \trianglelefteq A$ an ideal with $I \subseteq \mathrm{Jac}(A)$. If $IV = V$, then $V = 0$.*

Proof. By Corollary 3.8.6 there is $a \in A$ with $a \equiv 1 \bmod I$ and $aV = 0$. Since $1 - a \in I \subseteq \mathrm{Jac}(A)$, it follows from Lemma 2.6.8 that a is a unit. Since $aV = 0$, we must have $V = 0$. □

Corollary 3.8.8. *Let V be a finitely generated A-module, let $U \subseteq V$ be a submodule and $I \trianglelefteq A$ be an ideal with $I \subseteq \mathrm{Jac}(A)$. If $V = IV + U$, then already $U = V$.*

Proof. The assumption $V = IV + U$ implies $V/U = I(V/U)$. Hence, $V/U = 0$ by Corollary 3.8.7, so $V = U$. □

Recall from Exercise 2.3.11 that a ring A is called **local** if it has a unique maximal ideal M. In this case we have $\mathrm{Jac}(A) = M$, so Nakayama's lemma says something about the action of M on an A-module V.

Corollary 3.8.9. *Suppose that A is a local ring with maximal ideal M. Let $K \coloneqq A/M$. If V is a finitely generated A-module and $v_1, \ldots, v_n$ are such that their images $\overline{v}_1, \ldots, \overline{v}_n$ in $\overline{V} \coloneqq V/MV$ form a K-basis, then $\{v_1, \ldots, v_n\}$ is a minimal generating set of V.*

Proof. Let $U \coloneqq A\{v_1, \ldots, v_n\} \subseteq V$. Then, by assumption, $(U + MV)/MV = V/MV$, so $U + MV = V$. Hence, $U = V$ by Corollary 3.8.8, so $\{v_1, \ldots, v_n\}$ is a generating set of V. The minimality follows from the fact that a vector space basis is minimal. □

Exercises.

Exercise 3.8.10. Let A be a local ring. Let V and W be finitely generated A-modules. Show that if $V \otimes_A W = 0$, then already $V = 0$ or $W = 0$.

CHAPTER 4

Localization

In this chapter, we will discuss how we can create from a ring A and a subset $S \subseteq A$ another ring $S^{-1}A$ together with a morphism $j\colon A \to S^{-1}A$ such that the elements in $j(S) \subseteq A$ become invertible. The process is similar to how you get from the integers $\mathbb{Z}$ to the rational numbers $\mathbb{Q}$—you only have to be careful when there are zero-divisors in S. This process is called localization. But why is it called like this? A special case of localization is to take the complement $S = A \setminus P$ of a prime ideal P in A. This produces a ring $A_P \coloneqq S^{-1}A$ whose ideals are under j in correspondence with ideals in A contained in P. In particular, A_P is local with maximal ideal $j(P)$. If you think geometrically, this really means you throw away all points not having anything to do with P—you localize in P! Localization is a fundamental tool in commutative algebra.

4.1. Field of fractions

We'll start with the nicest case of localization, namely we'll find for an integral domain A a (minimal) field into which A can be embedded. The general idea is similar to how we get from $\mathbb{Z}$ to $\mathbb{Q}$ by adding a formal inverse $\frac{1}{n}$ for every non-zero $n \in \mathbb{Z}$ with the obvious arithmetic rules. You simply define a relation $\sim$ on $A \times (A \setminus \{0\})$ via

$$(a,s) \sim (b,t) :\Leftrightarrow at = bs \quad (\text{i.e. “}\frac{a}{s} = \frac{b}{t}\text{”}) \ . \tag{4.1}$$

This is an equivalence relation: it's obviously reflexive and symmetric; it's also transitive since if $(a,s) \sim (b,t)$ and $(b,t) \sim (c,u)$, then $at = bs$ and $bu = ct$, hence $atu = bsu = cts$ and since A is an integral domain, we conclude that $au = cs$, i.e. $(a,s) \sim (c,u)$. We will write $\frac{a}{s}$ for the equivalence class of (a,s). The set

$$\mathrm{Frac}(A) \coloneqq (A \times A \setminus \{0\})/\sim \tag{4.2}$$

of equivalence classes becomes a ring via

$$\frac{a}{s} + \frac{b}{t} \coloneqq \frac{at+bs}{st} \ , \tag{4.3}$$

$$\frac{a}{s} \cdot \frac{b}{t} \coloneqq \frac{ab}{st} \ . \tag{4.4}$$

The unit element is $1 = \frac{1}{1}$. It's not just a ring but actually a field since

$$\frac{a}{s} \cdot \frac{s}{a} = 1 \tag{4.5}$$

for $a \neq 0$. Moreover, the map

$$j\colon A \to \mathrm{Frac}(A) \ , \quad a \mapsto \frac{a}{1} \ , \tag{4.6}$$

is an injective ring morphism.

Definition 4.1.1. The field $\mathrm{Frac}(A)$ is called the **field of fractions** (or **fraction field**) of A.

Remark 4.1.2. In the literature, you'll also see $\mathrm{Quot}(A)$ instead of $\mathrm{Frac}(A)$ and people also call this the **field of quotients** or **or quotient field** of A. This term is a bit misleading since "quotient" may also mean the quotient by an ideal.

Example 4.1.3. $\mathrm{Frac}(\mathbb{Z}) = \mathbb{Q}$ (or $\mathrm{Frac}(\mathbb{Z}) \simeq \mathbb{Q}$ wherever you're coming from).

Example 4.1.4. If K is already a field, then $\mathrm{Frac}(K) \simeq K$ canonically.

Example 4.1.5. If K is a field, then the polynomial ring $K[X]$ is an integral domain and we can form

$$K(X) \coloneqq \mathrm{Frac}(K[X]) = \left\{ \frac{f}{g} \mid f,g \in K[X], g \neq 0 \right\} \; . \tag{4.7}$$

This is also called the **rational function field** over K. You can do this analogously in arbitrarily many variables.

Example 4.1.6. Let A be any ring and let P be a prime ideal of A. Then A/P is an integral domain and we can thus form

$$k_A(P) \coloneqq \mathrm{Frac}(A/P) \; . \tag{4.8}$$

This is also called the **residue class field** of A in P. If M is a maximal ideal, then $k_A(M) = A/M$. For example, we have

$$k_{\mathbb{Z}}((0)) = \mathbb{Q} \; , \quad k_{\mathbb{Z}}((p)) = \mathbb{F}_p \; , \quad k_{\mathbb{Z}[X]}((X)) = \mathbb{Q} \; , \quad k_{\mathbb{Z}[X]}((p)) = \mathbb{F}_p(X) \; . \tag{4.9}$$

The residue class field is precisely the construction that we needed in (2.38) to get a bijection between the prime spectrum of a ring and the places up to equivalence.

The fraction field of an integral domain A is the minimal field containing $\mathrm{Frac}(A)$. Let's make this precise:

Lemma 4.1.7. *If $h\colon A \to K$ is an injective ring morphism into a field, then there is a unique ring morphism $\widetilde{h}\colon \mathrm{Frac}(A) \to K$ making the diagram*

$$\begin{tikzcd}
\mathrm{Frac}(A) \arrow[r, hook, "\widetilde{h}"] & K \\
A \arrow[u, hook, "j"] \arrow[ur, hook, "h"'] &
\end{tikzcd} \tag{4.10}$$

commutative. The map $\widetilde{h}$ is injective as well.

Proof. Since h is injective, it follows that for any $0 \neq s \in A$ the element $h(s) \in K$ is non-zero, hence invertible. Assuming $\widetilde{h}$ exists, we can now write

$$\widetilde{h}\left(\frac{a}{s}\right) = \widetilde{h}\left(\frac{a}{1} \cdot \left(\frac{s}{1}\right)^{-1}\right) = \widetilde{h}\left(\frac{a}{1}\right)\widetilde{h}\left(\frac{s}{1}\right)^{-1} = \widetilde{h}j(a) \cdot (\widetilde{h}j(s))^{-1} = h(a) \cdot h(s)^{-1} \; .$$

Hence, $\widetilde{h}$ is uniquely determined. To prove existence of $\widetilde{h}$, we simply define it by the above equation. This is well-defined since if $\frac{a}{s} = \frac{b}{t}$, then $at = bs$, hence $h(a)h(t) = h(b)h(s)$ and therefore $h(a)h(s)^{-1} = h(b)h(t)^{-1}$. The map $\widetilde{h}$ is a ring morphism making the diagram commutative. □

4.2. Localization of rings

If A has zero-divisors, we cannot embed A into a field, so the construction from the previous section won't work. But there is a natural generalization. Suppose we want to make all elements of a subset $S \subseteq A$ invertible, and this should happen in a minimal way. This leads us to the following concept.

Definition 4.2.1. The **localization** of a ring A in a subset $S \subseteq A$ is a ring $S^{-1}A$ together with a ring morphism $j \colon A \to S^{-1}A$ such that:

(1) $j(s)$ is a unit in $S^{-1}A$ for any $s \in S$;
(2) if $h \colon A \to B$ is a ring morphism mapping elements from S to units in B, then there is a unique ring morphism $\widetilde{h} \colon S^{-1}A \to B$ making the diagram

$$\begin{tikzcd} S^{-1}A \arrow[r, "\widetilde{h}"] & B \\ A \arrow[u, "j"] \arrow[ur, "h"'] & \end{tikzcd} \tag{4.11}$$

commutative.

Lemma 4.2.2. *The localization $j \colon A \to S^{-1}A$ exists and is unique up to unique isomorphism.*

Proof. The uniqueness claim follows in the same way as for any other universal property. We'll prove existence by giving an explicit construction of $S^{-1}A$. First, let's record the following observation. Suppose that $S^{-1}A$ exists. If $s, t \in S$, then $j(s), j(t)$ are units in $S^{-1}A$, hence also $j(st) = j(s)j(t)$ is a unit in $S^{-1}A$. Moreover, $1 = j(1)$ is a unit. We thus have

$$S^{-1}A = (\overline{S})^{-1}A \ , \tag{4.12}$$

where $\overline{S}$ is the **multiplicative closure** of S, i.e. the smallest subset of A containing S and which is closed under taking products (which also implies that $1 \in \overline{S}$ since this is the empty product). Now, we define a relation $\sim$ on $A \times \overline{S}$ via

$$(a,s) \sim (b,t) :\Leftrightarrow atu = bsu \text{ for some } u \in \overline{S} \ . \tag{4.13}$$

This looks a bit different than the relation (4.1) that we used in the construction of the fraction field. If you go back you will notice that when we proved that (4.1) is transitive, we used the fact that our ring is an integral domain. Here, we do not want to assume that A is an integral domain. The relation (4.13) is the right thing to use in this generality because it is always transitive. We define

$$S^{-1}A \coloneqq (A \times \overline{S}) / \sim \tag{4.14}$$

and we write $\frac{a}{s}$ for the equivalence class of (a,s) in $S^{-1}A$. Then $S^{-1}A$ is a ring via

$$\frac{a}{s} + \frac{b}{t} \coloneqq \frac{at+bs}{st} \ , \tag{4.15}$$

$$\frac{a}{s} \cdot \frac{b}{t} \coloneqq \frac{ab}{st} \ . \tag{4.16}$$

The map

$$j \colon A \to S^{-1}A \ , \quad a \mapsto \frac{a}{1} \ , \tag{4.17}$$

is a ring morphism mapping elements of S to units in $S^{-1}A$. If $h\colon A \to B$ is a ring morphism mapping the elements of S to units in B, then a ring morphism $\widetilde{h}\colon S^{-1}A \to B$ making the diagram in Definition 4.2.1 commutative must satisfy

$$\widetilde{h}\left(\frac{a}{s}\right) = h(a) \cdot h(s)^{-1} \ ,$$

similarly to what we had for the fraction field. To prove existence of $\widetilde{h}$, we need to check if $\widetilde{h}$ defined as in the above equation is a well-defined ring morphism. This is straightforward. □

Example 4.2.3. If A is an integral domain, then $S = A \setminus \{0\}$ is multiplicatively closed and $S^{-1}A = \mathrm{Frac}(A)$ is the fraction field.

Example 4.2.4. We'll now come to the most fundamental example of localization. Let A be any ring and let P be a prime ideal of A. Then the complement $A \setminus P$ is multiplicatively closed and we call

$$A_P \coloneqq (A \setminus P)^{-1}A \tag{4.18}$$

the **localization** of A in P. So, explicitly

$$A_P = \left\{\frac{a}{s} \mid a \in A, s \notin P\right\} \ . \tag{4.19}$$

For example, for a prime number $p \in \mathbb{Z}$ we have

$$\mathbb{Z}_{(p)} = \left\{\frac{a}{s} \mid s \text{ coprime to } p\right\} \subseteq \mathbb{Q} \ . \tag{4.20}$$

Example 4.2.5. For an element f of a ring A we define

$$A_f \coloneqq \{f\}^{-1}A = (\overline{\{f\}})^{-1}A = \left\{\frac{a}{f^n} \mid a \in A, n \in \mathbb{N}\right\} \ . \tag{4.21}$$

For example, for a prime number p we obtain

$$\{p\}^{-1}\mathbb{Z} = \left\{\frac{a}{p^n} \mid a \in \mathbb{Z}, n \in \mathbb{N}\right\} \ . \tag{4.22}$$

Unfortunately, in this particular case the shorthand notation $\mathbb{Z}_p$ for $\{p\}^{-1}\mathbb{Z}$ that we just introduced is problematic because there are *three* objects that people like to denote by $\mathbb{Z}_p$. First, there is $\mathbb{Z}/p\mathbb{Z}$. I think it's generally a bad idea to denote this by $\mathbb{Z}_p$ since there is the (non-confusing and thus better) notation $\mathbb{F}_p$. Second, there is the localization $\{p\}^{-1}\mathbb{Z}$ that we just saw. Third, there is the ring of so-called p-adic integers (which are basically formal power series in p). This ring is really everywhere denoted by $\mathbb{Z}_p$ and since it is very important and used much more often than the localization of $\mathbb{Z}$ in $\{p\}$, we will reserve the notation $\mathbb{Z}_p$ for the p-adic integers and write $\{p\}^{-1}\mathbb{Z}$ if we really have to consider this special case (we won't).

Example 4.2.6. We have $S^{-1}A = \{0\}$ if and only if $0 \in \overline{S}$.

This last example shows you in particular that the canonical map $j\colon A \to S^{-1}A$ may *not* be injective, we'll take a more detailed look at this in Lemma 4.2.9. But I want to tell you one categorical property that the map j still satisfies. It's a bit weird because it's actually a generalization of *surjective* maps but the category of rings is simply a bit weird in this respect.

Definition 4.2.7. A morphism $f\colon X \to Y$ in a category $\mathcal{C}$ is called an **epimorphism** if every commutative diagram

$$\begin{tikzcd} X \arrow[r, "f"] & Y \arrow[r, shift left, "g_1"] \arrow[r, shift right, "g_2"'] & Z \end{tikzcd} \tag{4.23}$$

implies $g_1 = g_2$, i.e. f is **right-cancellative**. Dually, one defines the notion of a **monomorphism**.

In the category Set of sets you can easily convince yourself that monomorphisms are precisely the injective maps and the epimorphisms are precisely the surjective maps. This is still true in the category Grp of groups, but it's non-trivial to see that epimorphisms are surjective. You can check out my category theory lecture notes [13] for a proof. In the category Ring of (commutative) rings, things get really strange: monomorphisms are the same as injective ring morphisms, and surjective ring morphisms are epimorphisms; but there are epimorphisms which are not necessarily surjective—namely:

Lemma 4.2.8. *The localization map $j\colon A \to S^{-1}A$ is an epimorphism in* Ring.

Proof. Suppose we have a diagram

$$\begin{tikzcd} A \arrow[r, "j"] & S^{-1}A \arrow[r, shift left, "g_1"] \arrow[r, shift right, "g_2"'] & B\,. \end{tikzcd}$$

Then $g_1(j(a)) = g_2(j(a))$ for all $a \in A$. Since $j(s)$ is a unit for $s \in \overline{S}$, it follows that $g_i(j(s))$ is a unit in B with inverse $g_i(j(s)^{-1})$, hence $g_1(j(s)^{-1}) = g_2(j(s)^{-1})$. It follows that

$$g_1\left(\frac{a}{s}\right) = g_1\left(\frac{a}{1}\cdot\frac{1}{s}\right) = g_1(j(a))\cdot g_1(j(s)^{-1}) = g_2(j(a))\cdot g_2(j(s)^{-1}) = g_2\left(\frac{a}{s}\right)\ ,$$

i.e. $g_1 = g_2$. □

We can explicitly describe the kernel of the localization map.

Lemma 4.2.9. *The kernel of $j\colon A \to S^{-1}A$ is equal to*

$$\operatorname{Ker}(j) = \bigcup_{s\in\overline{S}} \operatorname{Ann}_A(s) = \{a \in A \mid sa = 0 \textit{ for some } s \in \overline{S}\}\ . \tag{4.24}$$

Proof. Suppose that $sa = 0$. Then $0 = j(sa) = \frac{s}{1}\cdot\frac{a}{1}$, hence $0 = \frac{1}{s}\cdot\frac{s}{1}\cdot\frac{a}{1} = \frac{a}{1} = j(a)$, so $a \in \operatorname{Ker}(j)$. Suppose conversely that $j(a) = 0$. Then $\frac{a}{1} = \frac{0}{1}$ in $S^{-1}A$. Hence, there is $u \in \overline{S}$ such that $a\cdot 1\cdot u = 0\cdot 1\cdot u = 0$, i.e. $au = 0$. □

Localization is transitive in the following sense: if S and T are subsets of A with $S \subseteq T$, then we have a canonical map

$$S^{-1}A \to T^{-1}A\ ,\quad \frac{a}{s} \mapsto \frac{a}{s}\ , \tag{4.25}$$

and the diagram

$$\begin{tikzcd} T^{-1}A \\ S^{-1}A \arrow[u] \\ A \arrow[u, "j_S"] \arrow[uu, bend right=60, "j_T"'] \end{tikzcd} \tag{4.26}$$

commutes. Moreover, the map $S^{-1}A \to T^{-1}A$ above is also the localization map of $S^{-1}A$ in $j_S(T)$ and

$$T^{-1}A \simeq U^{-1}(S^{-1}A) \ , \quad U := j_S(T) \ . \tag{4.27}$$

Keep in mind that none of these maps have to be injective, so localizing more and more doesn't necessarily create bigger and bigger rings.

But there is one nice situation where localization maps are injective and everything behaves nicely. Namely, we can take any ring A but consider subsets S which consist of non-zero-divisors. It is then clear from Lemma 4.2.9 that $j_S \colon A \to S^{-1}A$ is injective and

$$\frac{a}{s} = \frac{b}{t} \Leftrightarrow at = bs \ , \tag{4.28}$$

so we're back at the simpler relation (4.13) that we used for the fraction field. We can in particular consider the set of *all* non-zero-divisors. This set is already multiplicatively closed. The corresponding localization of A is called the **total ring of fractions** of A and we denote it by $\mathrm{Frac}(A)$ since this gives precisely the fraction field in case A is an integral domain. All localizations of A in subsets consisting of non-zero-divisors take place inside the total ring of fractions. Namely, let S consist of non-zero divisors and consider in $\mathrm{Frac}(A)$ the set of all elements of the form $\frac{a}{s}$ for $s \in S$. Then you can check that this subring together with the inclusion from A satisfies the universal property of the localization of A in S, hence is naturally isomorphic to $S^{-1}A$. We used this already when describing $\mathbb{Z}_{(p)}$ and $\{p\}^{-1}\mathbb{Z}$ as subrings of $\mathbb{Q}$ in Example 4.2.4 and Example 4.2.5.

Recall from Lemma 1.3.19 that a ring morphism $f \colon A \to B$ yields a Galois connection (f_*, f^*) between the set of ideals of A and those of B. In case of the localization map we have a very good understanding of the induced bijections.

Proposition 4.2.10. *Let A be a ring, let $S \subseteq A$ be any subset, and let $j \colon A \to S^{-1}A$ be the localization map. Let $\mathrm{Ideals}_S(A)$ be the set of all ideals I in A such that every $s \in \overline{S}$ is a non-zero-divisor in A/I. Then the Galois connection (j_*, j^*) restricts to bijections*

$$\begin{array}{rcl} \mathrm{Ideals}_S(A) & \leftrightarrow & \mathrm{Ideals}(S^{-1}A) \\ I & \mapsto & IS^{-1}A := j_*(I) = j(I)S^{-1}A \\ j^*(J) = j^{-1}(J) & \leftarrow\!\shortmid & J \end{array} \tag{4.29}$$

Moreover, it restricts to isomorphisms

$$\mathrm{Spec}(S^{-1}A) \simeq \mathrm{Spec}_S(A) := \{P \in \mathrm{Spec}(A) \mid P \cap \overline{S} = \emptyset\} \tag{4.30}$$

of topological spaces.

Proof. We know from Lemma 1.3.17 that $J \mapsto j^{-1}(J)$ is a map $\mathrm{Ideals}(S^{-1}A) \to \mathrm{Ideals}(A)$. We claim that $j^{-1}(J)$ is contained in $\mathrm{Ideals}_S(A)$. Let $a \in A$ and let $s \in \overline{S}$ such that $as \equiv 0 \bmod j^{-1}(J)$, i.e. $as \in j^{-1}(J)$. Then $\frac{as}{1} = j(as) \in J$, hence $\frac{a}{1} = \frac{1}{s} \cdot \frac{as}{1} \in J$, so $a \in j^{-1}(J)$ and therefore $a \equiv 0 \bmod j^{-1}(J)$. Hence, $j^{-1}(J) \in \mathrm{Ideals}_S(A)$.

Next, we show that $j_* j^*(J) = J$ for any ideal J of $S^{-1}A$. The inclusion $j(j^{-1}(J))S^{-1}A \subseteq J$ is clear. Conversely, let $\frac{a}{s} \in J$. Then $\frac{a}{1} = \frac{s}{1} \cdot \frac{a}{s} \in J$, hence $a \in j^{-1}(J)$, so $\frac{a}{1} \in j(j^{-1}(J))$, and therefore $\frac{a}{s} = \frac{a}{1} \cdot \frac{1}{s} \in j(j^{-1}(J))S^{-1}A$.

Now, we show that $j^*j_*(I) = I$ for any ideal $I \in \mathrm{Ideals}_S(A)$. The inclusion $j^{-1}(j(I)S^{-1}A) \supseteq I$ is clear. Conversely, let $a \in j^{-1}(j(I)S^{-1}A)$. Then $\frac{a}{1} \in j(I)S^{-1}A$ and therefore $\frac{a}{1} = \sum_{i=1}^n \frac{c_i}{1} \cdot \frac{a_i}{s_i}$ for certain $c_i \in I$, $a_i \in A$, and $s_i \in \overline{S}$. We can rewrite this $\frac{a}{1} = \sum_{i=1}^n \frac{c_i s_i' a_i}{s}$ with $s_i' := s_1 \cdots s_{i-1} s_{i+1} \cdots s_n$ and $s := s_1 \cdots s_n$. Hence, $\frac{a}{1} = \frac{c}{s}$ for some $c \in I$ and $s \in \overline{S}$. It follows that there is $u \in \overline{S}$ with $asu = cu \in I$. But since $su \in \overline{S}$ and $I \in \mathrm{Ideals}_S(A)$, it follows that $a \in I$.

We have now proven that (j_*, j^*) restricts to bijections between $\mathrm{Ideals}_S(A)$ and $\mathrm{Ideals}(S^{-1}A)$.

If Q is a prime ideal in $S^{-1}A$, we know from Lemma 2.2.1 that $j^*(Q)$ is a prime ideal in A. By what we proved above, $j^*(Q)$ is contained in

$$\mathrm{Ideals}_S(A) \cap \mathrm{Spec}(A) = \{P \in \mathrm{Spec}(A) \mid P \cap \overline{S} = \emptyset\} \ .$$

Conversely, let $P \in \mathrm{Spec}_S(A)$. We need to show that $j_*(P)$ is a prime ideal in $S^{-1}A$. Suppose that $\frac{a}{s} \cdot \frac{b}{t} \in j_*(P)$. Then $\frac{ab}{1} = st \cdot \frac{ab}{st} \in j_*(P)$, hence $ab \in j^*j_*(P) = P$. Since P is prime, it follows that $a \in P$ or $b \in P$. Without loss of generality, we can assume $a \in P$. Then $\frac{a}{1} \in j_*(P)$, hence $\frac{1}{s} \cdot \frac{a}{1} = \frac{a}{s} \in j_*(P)$. This proves that $j_*(P)$ is prime. We have now proven that (j_*, j^*) restricts to bijections between $\mathrm{Spec}_S(A)$ and $\mathrm{Spec}(S^{-1}A)$.

It remains to prove that this is topological. We already know from Lemma 2.5.15 that j^* is a continuous map, so we just need to prove that j_* is continuous as well. Let $Z \subseteq \mathrm{Spec}(S^{-1}A)$ be a closed subset. By definition, this means

$$Z = \mathrm{V}(J) = \{Q \in \mathrm{Spec}(S^{-1}A) \mid Q \supseteq J\}$$

for some ideal J in $S^{-1}A$. Using the Galois connection (j_*, j^*), we obtain

$$\begin{aligned}(j_*)^{-1}(Z) &= \{j^*(Q) \mid Q \in \mathrm{Spec}(S^{-1}A), Q \supseteq J\} \\ &= \{P \in \mathrm{Spec}_S(A) \mid P \supseteq j^*(J)\} \\ &= \mathrm{V}(j^*(J)) \cap \mathrm{Spec}_S(A) \ ,\end{aligned}$$

and this is a closed subset of $\mathrm{Spec}_S(A)$ by definition. □

So, the ideal theory of $S^{-1}A$ is just a portion—and thus a simplification—of the ideal theory of A.

COROLLARY 4.2.11. *If $P \in \mathrm{Spec}(A)$, then A_P is a* local *ring with maximal ideal PA_P and*

$$\mathrm{Spec}(A_P) \simeq \{Q \in \mathrm{Spec}(A) \mid Q \subseteq P\} \ . \tag{4.31}$$

PROOF. Since $A \setminus P$ is multiplicatively closed, Proposition 4.2.10 tells us that

$$\mathrm{Spec}(A_P) \simeq \{Q \in \mathrm{Spec}(A) \mid Q \cap (A \setminus P) = \emptyset\} = \{Q \in \mathrm{Spec}(A) \mid Q \subseteq P\} \ . \quad □$$

Note that it is only in the special case of localization in a prime ideal that we actually get a *local* ring. This is where the terminology "localization" comes from but it is also used beyond this case.

Exercises.

EXERCISE 4.2.12. Let A be a ring and let $f \in A$. Show that $\mathrm{Spec}(A_f)$ is homeomorphic to $\mathrm{D}(f) = \mathrm{Spec}(A) \setminus \mathrm{V}(f)$, i.e. there is an isomorphism of topological spaces.

Exercise 4.2.13. Let A be a ring and let $f \in A$. Show that A_f is isomorphic to $A[X]/(fX-1)$ as A-algebras.

Exercise 4.2.14. Let K be a field and consider the polynomial ring $K[X]$ in one variable. The localization

$$K[X, X^{-1}] \coloneqq K[X]_X = \{X\}^{-1}K[X] \tag{4.32}$$

is called the **Laurent polynomial** ring over K. Describe this ring explicitly and show that it is a principal ideal domain.

4.3. Localization of modules

Let A be a ring and let $S \subseteq A$ be a subset. The construction of $S^{-1}A$ can also be performed analogously for an A-module V. We define $S^{-1}V$ as the set $V \times \overline{S}$ modulo the equivalence relation

$$(v,s) \sim (w,t) :\Leftrightarrow \exists u \in \overline{S} \text{ with } vtu = wsu \; . \tag{4.33}$$

We write $\frac{v}{s}$ for the equivalence class of (v,s). We then get a well-defined A-module structure on $S^{-1}V$ via

$$a \cdot \frac{v}{s} \coloneqq \frac{av}{s} \; , \quad \frac{v}{s} + \frac{w}{t} \coloneqq \frac{vt+ws}{st} \; . \tag{4.34}$$

The canonical map

$$j \colon V \to S^{-1}V \; , \quad v \mapsto \frac{v}{1} \; , \tag{4.35}$$

is an A-module morphism. We call $S^{-1}V$ the **localization** of V in S. Note that $S^{-1}V$ is not just an A-module but naturally an $S^{-1}A$-module via

$$\frac{a}{s} \cdot \frac{v}{t} \coloneqq \frac{av}{st} \; . \tag{4.36}$$

We have two particularly important cases of localization:

$$V_P \coloneqq (A \setminus P)^{-1}V \quad \text{and} \quad V_f \coloneqq \{f\}^{-1}V \tag{4.37}$$

for a prime ideal P in A and an element $f \in A$.

Note that when viewing A as an A-module we have two constructions of localizations: the localization $S^{-1}A$ of the *ring* A and the localization $S^{-1}A$ of the A-*module* A. Both are clearly the same. Similarly as in Lemma 4.2.9 you prove that

$$\mathrm{Ker}(j \colon V \to S^{-1}V) = \{ v \in V \mid sv = 0 \text{ for some } s \in \overline{S} \} \; . \tag{4.38}$$

The ideal correspondence $\mathrm{Ideals}_S(A) \simeq \mathrm{Ideals}(S^{-1}A)$ from Proposition 4.2.10 generalizes to a correspondence

$$\mathrm{Sub}_S(V) \simeq \mathrm{Sub}(S^{-1}V) \; , \tag{4.39}$$

where here $\mathrm{Sub}(S^{-1}V)$ denotes the set of $S^{-1}A$-submodules of $S^{-1}V$ and $\mathrm{Sub}_S(V)$ is the set of all submodules U of V such that V/U is $\overline{S}$-torsion-free, i.e. if $s \in \overline{S}$ and $v \in V$ such that $sv \in U$, then $v \in U$.

From (4.38) you deduce:

Lemma 4.3.1. *If V is torsion-free, then $j \colon V \to S^{-1}V$ is injective.*

Remark 4.3.2. In contrast to the situation of localization of rings, the localization morphism $j \colon V \to S^{-1}V$ is in general *not* an epimorphism in the category of modules (as these are precisely the surjective morphisms).

If $f\colon V \to W$ is an A-module morphism, we get an $S^{-1}A$-module morphism

$$S^{-1}f\colon S^{-1}V \to S^{-1}W \ , \quad \frac{v}{s} \mapsto \frac{f(v)}{s} \ . \tag{4.40}$$

You can now convince yourself that localization in S defines a functor

$$S^{-1}-\colon A\text{-}\mathsf{Mod} \to S^{-1}A\text{-}\mathsf{Mod} \ . \tag{4.41}$$

LEMMA 4.3.3. *The functors $S^{-1}A \otimes_A -$ and $S^{-1}-$ are isomorphic, i.e. for every A-module V there is an $S^{-1}A$-module isomorphism*

$$\varphi_V\colon S^{-1}A \otimes_A V \to S^{-1}V \tag{4.42}$$

such that for every A-module morphism $f\colon V \to W$ the diagram

$$\begin{tikzcd}
S^{-1}A \otimes_A V \arrow[r, "\varphi_V"] \arrow[d, "S^{-1}A \otimes_A f"'] & S^{-1}V \arrow[d, "S^{-1}f"] \\
S^{-1}A \otimes_A W \arrow[r, "\varphi_W"'] & S^{-1}W
\end{tikzcd} \tag{4.43}$$

commutes.

PROOF. The map

$$S^{-1}A \times V \to S^{-1}V \ , \quad \left(\frac{a}{s}, v\right) \mapsto \frac{av}{s} \tag{4.44}$$

is A-linear, hence induces an A-module morphism

$$\varphi_V\colon S^{-1}A \otimes_A V \to S^{-1}V \ , \quad \frac{a}{s} \otimes v \mapsto \frac{av}{s} \ . \tag{4.45}$$

This map is obviously $S^{-1}A$-linear. The map

$$\psi_V\colon S^{-1}V \to S^{-1}A \otimes_A V \ , \frac{v}{s} \mapsto \frac{1}{s} \otimes v \tag{4.46}$$

is obviously an inverse to φ_V. Moreover, the diagram

$$\begin{tikzcd}
\frac{a}{s} \otimes v \arrow[rrr, mapsto] \arrow[ddd, mapsto] & & & \frac{av}{s} \arrow[ddd, mapsto] \\
 & S^{-1}A \otimes_A V \arrow[r, "\varphi_V"] \arrow[d, "S^{-1}A \otimes_A f"'] & S^{-1}V \arrow[d, "S^{-1}f"] & \\
 & S^{-1}A \otimes_A W \arrow[r, "\varphi_W"'] & S^{-1}W & \\
\frac{a}{s} \otimes f(v) \arrow[rrr, mapsto] & & & \frac{af(v)}{s}
\end{tikzcd}$$

commutes. □

Now, we come to a fundamental property of localization.

LEMMA 4.3.4. *The localization functor $S^{-1}\colon A\text{-}\mathsf{Mod} \to S^{-1}A\text{-}\mathsf{Mod}$ is exact, i.e. $S^{-1}A$ is a flat A-module.*

Proof. Let $V' \xrightarrow{f} V \xrightarrow{g} V''$ be an exact sequence of A-modules. Then

$$S^{-1}g \circ S^{-1}f\left(\frac{v'}{s}\right) = S^{-1}g\left(\frac{f(v')}{s}\right) = \frac{gf(v')}{s} = 0 \;, \tag{4.47}$$

hence $\operatorname{Im} S^{-1}f \subseteq \operatorname{Ker} S^{-1}g$. On the other hand, let $\frac{v}{s} \in \operatorname{Ker} S^{-1}g$. This means $\frac{g(v)}{s} = 0 = \frac{0}{1}$, so there is $u \in \overline{S}$ with $0 = g(v)u = g(vu)$, hence $uv \in \operatorname{Ker} g$. Because our initial sequence was exact, there is $v' \in V'$ with $uv = f(v')$. Hence,

$$\frac{v}{s} = \frac{uv}{su} = \frac{f(v')}{su} = S^{-1}f\left(\frac{v'}{su}\right) \in \operatorname{Im} S^{-1}f \;. \qquad \Box$$

Corollary 4.3.5. *If $U \subseteq V$ is a submodule, then $S^{-1}U \to S^{-1}V$ is injective, i.e. we can identify $S^{-1}U$ with a submodule of $S^{-1}V$. Moreover, there is a canonical isomorphism*

$$S^{-1}(V/U) \simeq S^{-1}V/S^{-1}U \tag{4.48}$$

of $S^{-1}A$-modules.

Localization also commutes with the tensor product:

Lemma 4.3.6. *If V and W are two A-modules, then there is a canonical isomorphism*

$$S^{-1}V \otimes_{S^{-1}A} S^{-1}W \simeq S^{-1}(V \otimes_A W) \tag{4.49}$$

of $S^{-1}A$-modules.

Proof. The map

$$S^{-1}V \times S^{-1}W \to S^{-1}(V \otimes_A W) \;, \quad \left(\frac{v}{s}, \frac{w}{t}\right) \mapsto \frac{v \otimes w}{st} \tag{4.50}$$

is $S^{-1}A$-bilinear, hence induces a morphism

$$\varphi\colon S^{-1}V \otimes_{S^{-1}A} S^{-1}W \to S^{-1}(V \otimes_A W) \;. \tag{4.51}$$

We claim that this is an isomorphism. First, we prove surjectivity. An element of $S^{-1}(V \otimes_A W)$ is of the form $\frac{1}{s}\sum_{i=1}^n v_i \otimes w_i$. We have

$$\varphi\Big(\sum_{i=1}^n \frac{v_i}{s} \otimes \frac{w_i}{1}\Big) = \frac{1}{s}\sum_{i=1}^n v_i \otimes w_i \;,$$

hence φ is surjective. To prove injectivity, suppose that $\varphi(\sum_{i=1}^n \frac{v_i}{s_i} \otimes \frac{w_i}{t_i}) = 0$, i.e.

$$\sum_{i=1}^n \frac{v_i \otimes w_i}{s_i t_i} = 0 \in S^{-1}(V \otimes_A W) \;. \tag{4.52}$$

Let $u_i := s_1 t_1 \cdots s_{i-1}t_{i-1}s_{i+1}t_{i+1} \cdots s_n t_n$ and $u := s_1 t_1 \cdots s_n t_n = u_i s_i t_i$. Multiplication of (4.52) by u yields

$$0 = \sum_{i=1}^n \frac{u_i v_i \otimes w_i}{1} \in S^{-1}(V \otimes_A W) \;.$$

Hence,

$$\sum_{i=1}^n u_i v_i \otimes w_i \in \operatorname{Ker}(j\colon V \otimes_A W \to S^{-1}(V \otimes_A W)) \;.$$

So, by Equation 4.38 there is $t \in \overline{S}$ with

$$0 = t\sum_{i=1}^{n} u_i v_i \otimes w_i = \sum_{i=1}^{n} t u_i v_i \otimes w_i \in V \otimes_A W \ .$$

The map

$$V \times W \to S^{-1}V \otimes_{S^{-1}A} S^{-1}W \ , \quad (v,w) \mapsto \frac{v}{1} \otimes \frac{w}{1}$$

is A-bilinear, hence induces a morphism

$$g \colon V \otimes_A W \to S^{-1}V \otimes_{S^{-1}A} S^{-1}W \ .$$

It follows that

$$0 = g\left(\sum_{i=1}^{n} t u_i v_i \otimes w_i\right) = \sum_{i=1}^{n} \frac{t u_i v_i}{1} \otimes \frac{w_i}{1} \in S^{-1}V \otimes_{S^{-1}A} S^{-1}W \ ,$$

hence

$$0 = \frac{1}{t} \cdot \frac{1}{u} \cdot \sum_{i=1}^{n} \frac{t u_i v_i}{u} \otimes \frac{w_i}{1} = \sum_{i=1}^{n} \frac{u_i v_i}{u} \otimes \frac{w_i}{1} = \sum_{i=1}^{n} \frac{v_i}{s_i t_i} \otimes \frac{w_i}{1} = \sum_{i=1}^{n} \frac{v_i}{s_i} \otimes \frac{w_i}{t_i} \ .$$

□

We want to show that localization also commutes with the Hom-functor—but this only works for modules with a finiteness condition that we'll introduce now. Recall that an A-module V is finitely generated if and only if there is an exact sequence $A^n \to V \to 0$ of A-modules for some $n \in \mathbb{N}$. This is fine, but if you really want to describe the module V constructively (e.g. in a computer) you not just want a finite generating set but also the relations between the generators—the syzygy module—should be finitely generated. This brings us to the following definition.

Definition 4.3.7. An A-module V is **finitely presented** if there is an exact sequence

$$A^m \to A^n \to V \to 0 \tag{4.53}$$

for some $m, n \in \mathbb{N}$, i.e. the kernel of $A^n \to V$ is finitely generated as well.

Example 4.3.8. Every free module of finite rank is finitely presented.

Lemma 4.3.9. *Every finitely generated projective module is finitely presented.*

Proof. Let V be finitely generated and projective. Then we have a surjective morphism $f \colon A^n \to V$. Since V is projective, there is by Exercise 3.7.4 a section $s \colon V \to A^n$, i.e. $f \circ s = \mathrm{id}_V$. We then have $A^n = \mathrm{Im}(s) \oplus \mathrm{Ker}(f) \simeq V \oplus \mathrm{Ker}(f)$. The projection $A^n \to \mathrm{Ker}(f)$ is surjective, hence $\mathrm{Ker}(f)$ is finitely generated, i.e. V is finitely presented. □

Now, here is what we want to prove:

Theorem 4.3.10. *Let B be an A-algebra and let V and W be two A-modules. Then there is a B-module morphism*

$$\begin{array}{rcl} \alpha_{V,W} \colon B \otimes_A \mathrm{Hom}_A(V,W) & \to & \mathrm{Hom}_B(B \otimes_A V, B \otimes_A W) \ , \\ 1 \otimes f & \mapsto & (1 \otimes v \mapsto 1 \otimes f(v)) \ . \end{array} \tag{4.54}$$

If B is a flat A-module and V is finitely presented, then $\alpha_{V,W}$ is an isomorphism.

Before we come to the proof, let's write down an important application.

Corollary 4.3.11. *If $S \subseteq A$ and V is a finitely presented A-module, then for any A-module W there is an isomorphism*

$$S^{-1}\operatorname{Hom}_A(V,W) \simeq \operatorname{Hom}_{S^{-1}A}(S^{-1}V, S^{-1}W) \tag{4.55}$$

of $S^{-1}A$-modules.

Proof. This follows immediately from Theorem 4.3.10 because $S^{-1}A$ is a flat A-module by Lemma 4.3.4. □

Now, we come to the proof of Theorem 4.3.10. The proof is really beautiful because it uses many of the things we discussed about modules. You should spend some time trying to understand all the arguments. Afterwards you'll be an expert on modules!

Proof of Theorem 4.3.10. It's easy to see that $\alpha \coloneqq \alpha_{V,W}$ is a B-module morphism. We'll proceed in three steps.

First, assume that $V = A$. Then $\operatorname{Hom}_A(V,W) \simeq W$ via $f \mapsto f(1)$ and

$$\operatorname{Hom}_B(B \otimes_A V, B \otimes_A W) = \operatorname{Hom}_B(B \otimes_A A, B \otimes_A W) \simeq B \otimes_A W\ .$$

The morphism α is in this case simply the identity $B \otimes_A W \to B \otimes_A W$.

Next, consider $V = A^n$. Since Hom and $\otimes_A$ commute with finite direct sums by Exercise 3.2.3 and Lemma 3.4.6, we can deduce this case from the previous case.

Finally, let V be general. Choose a presentation

$$A^m \xrightarrow{\ f\ } A^n \xrightarrow{\ g\ } V \longrightarrow 0\ . \tag{4.56}$$

Since $B \otimes_A -$ is right-exact by Lemma 3.5.7, we obtain a presentation

$$\begin{tikzcd}
B \otimes_A A^m \arrow[r] \arrow[d, "\simeq"] & B \otimes_A A^n \arrow[r] \arrow[d, "\simeq"] & B \otimes_A V \arrow[r] \arrow[d, equal] & 0 \arrow[d, equal] \\
B^m \arrow[r, ":=g'"] & B^n \arrow[r, ":=f'"] & B \otimes_A V \arrow[r] & 0
\end{tikzcd} \tag{4.57}$$

Hence, $V' \coloneqq B \otimes_A V$ is a finitely presented B-module. Let $W' \coloneqq B \otimes_A W$. In Exercise 3.7.2 you have shown that the Hom-functor with the argument in the second variable is left-exact. Similarly, you prove that the *contra*variant functor $\operatorname{Hom}_A(-,W)$ with the argument in the first variable is left-exact, which means that applying this functor to the presentation (4.56) yields an exact sequence

$$0 \longrightarrow \operatorname{Hom}_A(V,W) \xrightarrow{:=g^\vee} \operatorname{Hom}_A(A^n,W) \xrightarrow{:=f^\vee} \operatorname{Hom}_A(A^m,W)\ . \tag{4.58}$$

Analogously, applying $\operatorname{Hom}_B(-,W')$ to the presentation (4.57) of V' yields an exact sequence

$$0 \longrightarrow \operatorname{Hom}_B(V',W') \xrightarrow{:=g'^\vee} \operatorname{Hom}_B(B^n,W') \xrightarrow{:=f'^\vee} \operatorname{Hom}_B(B^m,W')\ . \tag{4.59}$$

Since B is a flat A-module by assumption, the functor $B \otimes_A -$ is exact, hence applying it to (4.58) yields an exact sequence

$$0 \to B \otimes_A \operatorname{Hom}_A(V,W) \overset{:=g^{\vee\prime}}{\longrightarrow} B \otimes_A \operatorname{Hom}_A(A^n,W) \overset{:=f^{\vee\prime}}{\longrightarrow} B \otimes_A \operatorname{Hom}_A(A^m,W)\ . \tag{4.60}$$

Combining (4.59) and (4.60), we get the commutative diagram

$$\begin{tikzcd}
0 \arrow[r] \arrow[d, equal] & B \otimes_A \operatorname{Hom}_A(V, W) \arrow[r, "g^{\vee\prime}"] \arrow[d, "\alpha_{V,W}"] & B \otimes_A \operatorname{Hom}_A(A^n, W) \arrow[r, "f^{\vee\prime}"] \arrow[d, "\alpha_{A^n,W}", "\simeq"'] & B \otimes_A \operatorname{Hom}_A(A^m, W) \arrow[d, "\alpha_{A^m,W}", "\simeq"'] \\
0 \arrow[r] & \operatorname{Hom}_B(V', W') \arrow[r, "g'^{\vee}"] & \operatorname{Hom}_B(B^n, W') \arrow[r, "f'^{\vee}"] & \operatorname{Hom}_B(B^m, W')
\end{tikzcd} \tag{4.61}$$

The rows are exact and the two vertical morphisms on the right are isomorphisms by the second case. It now follows from the five lemma which you'll prove below in Exercise 4.3.12 that $\alpha_{V,W}$ is an isomorphism as well (you add a zero column on the left of this diagram to get the required five modules). □

Exercises.

Exercise 4.3.12. Consider a commutative diagram

$$\begin{tikzcd}
V \arrow[r, "f"] \arrow[d, two heads, "l"] & W \arrow[r, "g"] \arrow[d, "m"] & X \arrow[r, "h"] \arrow[d, "n"] & Y \arrow[r, "j"] \arrow[d, "p"] & Z \arrow[d, hook, "q"] \\
V' \arrow[r] & W' \arrow[r] & X' \arrow[r] & Y' \arrow[r] & Z'
\end{tikzcd} \tag{4.62}$$

of module morphisms, where:

(1) the two rows are exact;
(2) m and p are isomorphisms;
(3) l is surjective;
(4) q is injective.

Show that n is an isomorphism. This is the so-called **five lemma**.

4.4. Local properties

We'll come to an extremely important concept in commutative algebra. Recall that the localization A_P of a ring A in a prime ideal P really has the geometric meaning of throwing away everything not having anything to do with the point P. Now, suppose you're standing in a huge dark room full of stuff and you're trying to get a full picture of this room. Luckily, you have a flash light which helps you to understand some local spots of the huge room. If you have gathered local information about the room from every possible spot, you may hope to piece together a full picture of the room. This is what we're trying to do here.

More specifically, let X be either a ring A, or a module V over a ring A, or a morphism $f\colon V \to W$ of modules over a ring A. Then for a prime ideal P in A we have defined the localization X_P of X in P, namely

$$X_P = \begin{cases} A_P = (A \setminus P)^{-1} A \\ V_P = (A \setminus P)^{-1} V \\ f_P \coloneqq (A \setminus P)^{-1} f\colon V_P \to W_P \end{cases} \tag{4.63}$$

The idea is now to study X via the X_P, which are (hopefully) simpler to study, and then deduce information about X itself. This only works really well for "local properties":

Definition 4.4.1. A property $\mathcal{P}$ of rings/modules/module morphisms is called **local** if the following holds: X has property $\mathcal{P}$ if and only if X_P has property $\mathcal{P}$ for all $P \in \operatorname{Spec}(A)$.

I'm not going to give a formal definition of what I mean by "property" because this will become clear from the examples and it's not helpful formalizing this. First, let's prove the following "meta-lemma":

Lemma 4.4.2. *If $\mathcal{P}$ is a local property, then it is sufficient to check it only in the maximal ideals, i.e. if X_M has property $\mathcal{P}$ for all maximal ideals M in A, then X has property $\mathcal{P}$.*

Proof. We need to show that X_P holds for all $P \in \operatorname{Spec}(A)$. Choose a maximal ideal M in A with $P \subseteq M$. By assumption, X_M has property $\mathcal{P}$. But then, since $\mathcal{P}$ is local, also $(X_M)_{PA_M} \simeq X_P$ has property $\mathcal{P}$. Here, we have used the transitivity of localization. □

This is all very abstract, so let's finally look at some specific examples.

Lemma 4.4.3. *The property of modules to be the zero module is a local property, i.e. for an A-module V the following are equivalent:*

(1) $V = 0$*;*
(2) $V_P = 0$ *for all* $P \in \operatorname{Spec}(A)$*;*
(3) $V_M = 0$ *for all* $M \in \operatorname{Max}(A)$*.*

Proof. The claims are obviously true for $A = 0$, so we assume that $A \neq 0$. The implications (1) ⇒ (2) ⇒ (3) are clear. Suppose that (3) holds but that $V \neq 0$. Take $0 \neq v \in V$. Then $I \coloneqq \operatorname{Ann}_A(v)$ is a proper ideal of A, hence it is contained in a maximal ideal M. Since $V_M = 0$ by assumption, it follows that $\frac{v}{1} = \frac{0}{1} \in V_M$, so there is $u \in A \setminus M$ with $vu = 0$. But this means $u \in I \subseteq M$—a contradiction. Hence, we must have $V = 0$. This proves (3) ⇒ (1). □

This brings us to the following:

Definition 4.4.4. For an A-module V the set

$$\operatorname{Supp}(V) \coloneqq \{P \in \operatorname{Spec}(A) \mid V_P \neq 0\} \tag{4.64}$$

is called the **support** of V.

Corollary 4.4.5. *Equality of modules is a local property, i.e. if V is an A-module and U is a submodule, then the following are equivalent:*

(1) $U = V$*;*
(2) $U_P = V_P$ *for all* $P \in \operatorname{Spec}(A)$*;*
(3) $U_M = V_M$ *for all* $M \in \operatorname{Max}(A)$*.*

Proof. Since localization is exact, we have $(V/U)_P \simeq V_P/U_P$. The claim now follows immediately from Lemma 4.4.3 applied to V/U. □

Corollary 4.4.6. *Exactness of sequences of module morphisms is a local property, i.e. for a sequence $V' \to V \to V''$ of A-module morphisms the following are equivalent:*

(1) *the sequence* $V' \to V \to V''$ *is exact;*
(2) *the sequence* $V'_P \to V_P \to V''_P$ *is exact for all* $P \in \operatorname{Spec}(A)$*;*
(3) *the sequence* $V'_M \to V_M \to V''_M$ *is exact for all* $M \in \operatorname{Max}(A)$*.*

Proof. The implication (1) ⇒ (2) follows from the exactness of localization. The implication (2) ⇒ (3) is clear. Let's consider (3) ⇒ (1). If f and g denote the morphisms in the sequence $V' \to V \to V''$, then we need to show that $\operatorname{Im}(f) =$

$\operatorname{Ker}(g)$. By assumption, we have $\operatorname{Im}(f)_M = \operatorname{Ker}(g)_M$ for all $M \in \operatorname{Max}(A)$. But $\operatorname{Im}(f)_M = \operatorname{Im}(f_M)$ and $\operatorname{Ker}(g)_M = \operatorname{Ker}(g_M)$ by exactness of localization. Using Corollary 4.4.5, this implies exactness of $V' \to V \to V''$. □

Corollary 4.4.7. *Injectivity, surjectivity, and bijectivity of module morphisms are local properties.*

Remark 4.4.8. You need to use this corollary correctly. It just says that if you start with a morphism $f\colon V \to W$ you can check whether it's, say, an isomorphism by checking this for the local morphisms f_M. It does not mean that if V_M and W_M are isomorphic by *some* isomorphism for all M, then V and W are isomorphic—you need to start with a "global" morphism $V \to W$.

Lemma 4.4.9. *Flatness is a local property, i.e. for an A-module V the following are equivalent:*

(1) *V is a flat A-module;*
(2) *V_P is a flat A_P-module for all $P \in \operatorname{Spec}(A)$;*
(3) *V_M is a flat A_M-module for all $M \in \operatorname{Max}(A)$.*

Proof. (1) ⇒ (2): Since A_P is a flat A-module, the scalar extension $V_P \simeq A_P \otimes_A V$ is a flat A_P-module by Exercise 3.6.7. The implication (2) ⇒ (3) is clear.

(3) ⇒ (1): Let $W' \to W \to W''$ be an exact sequence of A-modules. We need to prove that $V \otimes_A W' \to V \otimes_A W \to V \otimes_A W''$ is exact. Since V_M is a flat A_M-module, the top row in the commutative diagram

$$\begin{tikzcd}
V_M \otimes_{A_M} W'_M \arrow[r] \arrow[d, "\simeq"] & V_M \otimes_{A_M} W_M \arrow[r] \arrow[d, "\simeq"] & V_M \otimes_{A_M} W''_M \arrow[d, "\simeq"] \\
(V \otimes_A W')_M \arrow[r] & (V \otimes_A W)_M \arrow[r] & (V \otimes_A W'')_M
\end{tikzcd} \tag{4.65}$$

is exact. The vertical morphisms are the isomorphisms from Lemma 4.3.6. Hence, also the lower row is exact. Now, exactness is a local property by Corollary 4.4.6, hence $V \otimes_A W' \to V \otimes_A W \to V \otimes_A W''$ is exact. □

Freeness and projectivity on the other hand are *not* local properties. This follows from the following important theorem.

Theorem 4.4.10. *For* finitely generated *modules over a* local *ring flatness, projectivity, and freeness are all equivalent.*

Proof. Let A be a local ring and let V be a finitely generated A-module. We already know that freeness implies projectivity and that projectivity implies flatness. So, assume that V is flat. We need to show that V is already free. Let M be the maximal ideal of A. We will prove the following: if $\overline{v}_1, \dots, \overline{v}_n$ are linearly independent elements of the A/M-vector space V/MV, then representatives $v_1, \dots, v_n$ in V are linearly independent over A as well. If then $\overline{v}_1, \dots, \overline{v}_n$ is a basis, we know from Corollary 3.8.9 (which follows from Nakayama's lemma) that $v_1, \dots, v_n$ is a basis of V, hence V is free.

So, suppose that $\sum_{i=1}^n a_i v_i = 0 \in V$. Let $I \coloneqq (a_1, \dots, a_n) \trianglelefteq A$. Since V is flat, the map

$$I \otimes_A V \to A \otimes_A V \simeq V\ , \quad a \otimes v \mapsto av\ , \tag{4.66}$$

is injective. Hence, we have $\sum_{i=1}^{n} a_i \otimes v_i = 0 \in I \otimes_A V$. Let $g\colon A^n \to I$ be the map sending the i-th standard basis element $e_i \in A^n$ to a_i. Let $K \coloneqq \mathrm{Ker}(g)$ and consider the exact sequence

$$0 \longrightarrow K \xrightarrow{f} A^n \xrightarrow{g} I \longrightarrow 0 \,, \tag{4.67}$$

where $f\colon K \to A^n$ is the inclusion. Since V is flat, the induced sequence

$$0 \longrightarrow K \otimes_A V \xrightarrow{f'} A^n \otimes_A V \xrightarrow{g'} I \otimes_A V \longrightarrow 0 \tag{4.68}$$

is exact. We have

$$g'(\sum_{i=1}^{n} e_i \otimes v_i) = \sum_{i=1}^{n} a_i \otimes v_i = 0 \,,$$

hence, there is $\sum_{j=1}^{m} a'_j \otimes v'_j \in K \otimes_A V$ with

$$\sum_{j=1}^{m} a'_j \otimes v'_j = \sum_{i=1}^{n} e_i \otimes v_i \in A^n \otimes_A V \,.$$

We can write $a'_j = \sum_{i=1}^{n} a_{ij} e_i$ for certain $a_{ij} \in A$. Then

$$\sum_{i=1}^{n} e_i \otimes v_i = \sum_{j=1}^{m} a'_j \otimes v'_j = \sum_{j=1}^{m} \left(\sum_{i=1}^{n} a_{ij} e_i \right) \otimes v'_j = \sum_{i=1}^{n} e_i \otimes \left(\sum_{j=1}^{m} a_{ij} v'_j \right) \,,$$

hence,

$$v_i = \sum_{j=1}^{m} a_{ij} v'_j \tag{4.69}$$

since

$$A^n \otimes_A V \simeq (\bigoplus_{i=1}^{n} A) \otimes_A V \simeq \bigoplus_{i=1}^{n} (A \otimes_A V) \simeq V^n \,.$$

Since $a'_j = \sum_{i=1}^{n} a_{ij} e_i \in \mathrm{Ker}(g)$, we have

$$0 = g(a'_j) = \sum_{i=1}^{n} a_{ij} g(e_i) = \sum_{i=1}^{n} a_{ij} a_i \tag{4.70}$$

for all j. Our goal was to prove that the a_i are zero and we'll now come to this by induction on n.

First, consider $n = 1$. In this case $a_1 v_1 = 0$. By (4.69) we have $v_1 = \sum_{j=1}^{m} a_{1j} v'_j$ and by (4.70) we have $a_{1j} a_1 = 0$ for all j. Since $\overline{v}_1 \neq 0$, we have $v_1 \notin MV$. Hence, there must be an index k with $a_{1k} \notin M$, i.e. $a_{1k} \in A \setminus M = A^\times$ by Exercise 2.3.11. Hence, $a_{1k} a_1 = 0$ implies $a_1 = 0$.

Now, consider $n > 1$. By (4.69) we have $v_n = \sum_{j=1}^{m} a_{nj} v'_j$. As before, since $\overline{v}_n \neq 0$, we have $v_n \notin MV$, hence, there must be an index k with $a_{nk} \notin M$, so $a_{nk} \in A^\times$. By (4.70) we have $0 = \sum_{i=1}^{n} a_{ik} a_i$, hence $a_{nk} a_n = -\sum_{i=1}^{n-1} a_{ik} a_i$, from which we get

$$a_n = -\sum_{i=1}^{n-1} \frac{a_{ik}}{a_{nk}} a_i = \sum_{i=1}^{n-1} c_i a_i$$

with $c_i := -\frac{a_{ik}}{a_{nk}}$. Now, we get

$$\begin{aligned}0 &= \sum_{i=1}^{n} a_i v_i = \sum_{i=1}^{n-1} a_i v_i + a_n v_n = \sum_{i=1}^{n-1} a_i v_i + \left(\sum_{i=1}^{n-1} c_i a_i \right) v_n \\ &= a_1(v_1 + c_1 v_n) + \ldots + a_{n-1}(v_{n-1} + c_{n-1} v_n) \ .\end{aligned}$$

The elements $\overline{v_i + c_i v_n} \in V/MV$ for $i = 1, \ldots, n-1$ are linearly independent. Hence, by induction we conclude that $a_1 = \ldots = a_{n-1} = 0$. But then also $a_n = 0$. □

COROLLARY 4.4.11. *If V is a finitely generated flat module over a (not necessarily local) ring A, then V_P is already a free A_P-module for all $P \in \operatorname{Spec}(A)$.*

PROOF. By Exercise 3.6.7, flatness is preserved under scalar extension, hence V_P is a flat (and finitely generated) A_P-module, thus free by Theorem 4.4.10. □

REMARK 4.4.12. There are finitely generated flat modules which are not projective, see Exercise 4.4.18. This implies that freeness and projectivity are *not* local properties because then every finitely generated flat module would already be free (and thus projective) by Corollary 4.4.11.

REMARK 4.4.13. In the proof of Theorem 4.4.10 we have proven the following: if V is a flat module over a (not necessarily local) ring A and $\sum_{i=1}^{n} a_i v_i = 0$ is a relation in V, this relation is already trivial in the sense that there are $v'_1, \ldots, v'_m \in V$ and $a_{ij} \in A$ such that

$$v_i = \sum_{j=1}^{m} a_{ij} v'_j \quad \text{and} \quad \sum_{i=1}^{n} a_{ij} a_i = 0 \tag{4.71}$$

for all j. One can show that this property (for all relations) is actually equivalent to flatness of V!

REMARK 4.4.14. Let A be a local ring. Theorem 4.4.10 tells us that if V is finitely generated and projective, then V is already free. This particular implication actually holds without the assumption that V is finitely generated! This is a theorem by Kaplansky. For flatness, however, we cannot remove this assumption.

Exercises.

EXERCISE 4.4.15.

(1) Let A be a ring and let V be an A-module. Show that for a subset $S \subseteq A$ consisting of non-zero-divisors we have

$$\mathrm{T}(S^{-1}V) = S^{-1}\,\mathrm{T}(V) \ , \tag{4.72}$$

where $\mathrm{T}(-)$ denotes the torsion submodule.

(2) Deduce that torsion-freeness of modules over an integral domain is a local property.

EXERCISE 4.4.16. Show that projectivity of *finitely presented* modules is a local property.

EXERCISE 4.4.17. Show that for a module V over an arbitrary ring the following are equivalent:

(1) V is finitely presented and flat.
(2) V is finitely generated and projective.

Exercise 4.4.18. In this exercise you will construct a finitely generated flat module which is not projective. This shows that we need the assumption "finitely presented" in Exercise 4.4.17 and that freeness and projectivity are not local properties (see Remark 4.4.12). This example is due to Vasconcelos and the following piecemeal approach is due to Lam.

(1) First, you need to prove **Schanuel's Lemma**: if

$$0 \longrightarrow K \longrightarrow P \xrightarrow{f} M \longrightarrow 0$$

and

$$0 \longrightarrow K' \longrightarrow P' \xrightarrow{f'} M \longrightarrow 0$$

are short exact sequences of module morphisms over a ring A with P and P' projective, then there is an isomorphism

$$K' \oplus P \simeq K \oplus P' \ .$$

Hint: consider $X := \{(p, p') \in P \oplus P' \mid f(p) = f'(p')\}$ and show that $X \simeq K' \oplus P$ and $X \simeq K \oplus P'$.

(2) Show that if A is a ring and V is a finitely presented A-module, then the kernel $\operatorname{Ker}(f)$ of a surjective morphism $f \colon W \twoheadrightarrow V$ from a finitely generated A-module W is finitely generated as well.
Hint: since W is finitely generated, there is a surjective morphism $g \colon A^k \twoheadrightarrow W$ for some $k \in \mathbb{N}$. From this you get an exact sequence $0 \to K' \to A^k \overset{fg}{\to} V \to 0$. Moreover, we have an exact sequence $0 \to K \to A^n \to V \to 0$ with K finitely generated since V is finitely presented. Now, use Schanuel's Lemma and note that $g(\operatorname{Ker}(fg)) = \operatorname{Ker}(f)$.

(3) Consider the $\mathbb{Z}$-module $A_0 := \bigoplus_{n \in \mathbb{N}} (\mathbb{Z}/2\mathbb{Z})$. With respect to component-wise addition and multiplication this is a ring *without* unit. But $A := \mathbb{Z} \oplus A_0$ becomes a ring *with* unit $(1, 0)$ with respect to component-wise addition and the multiplication defined by $(n, a_0) \cdot (n', a_0') := (nn', na_0' + n'a_0 + a_0 a_0')$.

(4) Let $a := (2, 0) \in A$ and $V := (a) \trianglelefteq A$. This is a finitely generated A-module. We have a short exact sequence $0 \to \operatorname{Ann}_A(a) \to A \overset{\varphi}{\to} V \to 0$, where φ is multiplication by a. Show that the ideal $\operatorname{Ann}_A(V) \trianglelefteq A$ is not finitely generated and conclude that V is not finitely presented.

(5) Show that V is not projective.
Hint: Can a finitely generated but not finitely presented module be projective?

(6) Show that V is flat.
Hint: Show that V is locally flat, i.e. V_P is flat for all $P \in \operatorname{Spec}(A)$. Treat the cases $A_0 \not\subseteq P$ and $A_0 \subseteq P$ separately.

CHAPTER 5

Integrality

In this chapter, we'll be concerned with particularly nice ring extensions $A \subseteq B$, so-called *integral* ring extensions. General examples are *finite* extensions, i.e. where B is a finitely generated A-module, and particular examples of those are extensions like $\mathbb{Z} \subseteq \mathbb{Z}[i]$ that you consider in number theory. The nice feature about integral ring extensions is that the morphism $\mathrm{Spec}(B) \to \mathrm{Spec}(A)$, $Q \mapsto Q \cap A$, associated to the inclusion $A \hookrightarrow B$ is *surjective* and has *finite* fibers. This allows you to bring some order into the spectrum of B and you can draw nice pictures like in Figure 5.1. The number-theoretic examples show some beautiful special features that we're not going to address here—we'll try to keep things rather general for now.

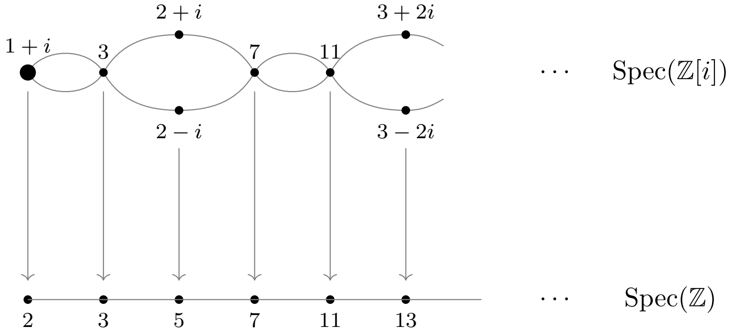

FIGURE 5.1. Visualization of $\mathrm{Spec}(\mathbb{Z}[i]) \to \mathrm{Spec}(\mathbb{Z})$.

5.1. Integral elements

Integrality is an important finiteness condition on elements and ring extensions which generalizes the module-theoretic finiteness (i.e. B is a finitely generated A-module) and includes some infinite extensions as well. Throughout, let $A \subseteq B$ be a ring extension.

DEFINITION 5.1.1. An element $b \in B$ is called **integral** over A if there is a *monic* polynomial

$$p = X^n + a_1 X^{n-1} + \ldots + a_{n-1} X + a_n \in A[X] \tag{5.1}$$

with the property that $p(b) = 0$.

Note two crucial things:

(1) Integrality is a *relative* notion: $b \in B$ is integral *over* A.

(2) We want the polynomial p to be *monic*, i.e. having leading coefficient equal to 1.

By $\mathrm{Int}_A(B) \subseteq B$ we denote the set of all elements of B which are integral over A.

EXAMPLE 5.1.2. Every $a \in A \subseteq B$ is integral over A since it satisfies the monic polynomial $X - a$. Hence,

$$A \subseteq \mathrm{Int}_A(B) \subseteq B \ . \tag{5.2}$$

Because of this observation we make the following definitions.

DEFINITION 5.1.3. One calls $\mathrm{Int}_A(B)$ the **integral closure** of A in B.

(1) If $\mathrm{Int}_A(B) = B$, then B is said to be **integral** over A; alternatively one says that the extension $A \subseteq B$ is **integral**.
(2) If $\mathrm{Int}_A(B) = A$, then A is said to be **integrally closed** in B.

EXAMPLE 5.1.4. The integral closure of $\mathbb{Z}$ in $\mathbb{Q}$ is just $\mathbb{Z}$, and this is one of the reasons why integral elements are called *integral.* Namely, let $\frac{r}{s} \in \mathbb{Q}$. Without loss of generality we can assume that r and s are coprime. Let $p = X^n + a_1X^{n-1} + \ldots + a_n \in \mathbb{Z}[X]$ with $p(\frac{r}{s}) = 0$. This means:

$$0 = \left(\frac{r}{s}\right)^n + a_1\left(\frac{r}{s}\right)^{n-1} + \ldots + a_n \ ,$$

hence

$$0 = r^n + a_1r^{n-1}s + \ldots + a_{n-1}rs^{n-1} + a_ns^n \ ,$$

from which we get

$$r^n = s(-a_1r^{n-1} - \ldots - a_{n-1}rs^{n-2} - a_ns^{n-1}) \ .$$

We thus conclude that s divides r^n. Since r and s were assumed to be coprime, it follows that s is a unit in $\mathbb{Z}$ and therefore $\frac{r}{s} \in \mathbb{Z}$.

EXAMPLE 5.1.5. With the same argumentation as in Example 5.1.4 you prove more generally that any unique factorization domain is integrally closed in its field of fractions.

We want to give a module-theoretic characterization of integrality. This will allow us to prove that $\mathrm{Int}_A(B)$ is actually a sub*ring* of B. To this end, we'll need the following concept.

DEFINITION 5.1.6. An A-module V is called **faithful** if $\mathrm{Ann}_A(V) = 0$, i.e. if $0 \neq a \in A$, then $aV \neq 0$.

In the following, for an element $b \in B$ we denote as usual by $A[b]$ the A-subalgebra of B generated by b.

THEOREM 5.1.7. *For an element $b \in B$ the following are equivalent:*

(1) *b is integral over A;*
(2) *$A[b]$ is a finitely generated A-module;*
(3) *$A[b]$ is contained in a subring A' of B such that A' is a finitely generated A-module;*
(4) *There is a faithful $A[b]$-module which is finitely generated as an A-module.*

PROOF. (1) $\Rightarrow$ (2): There is a monic polynomial $p = X^n + a_1X^{n-1} + \ldots + a_n \in A[X]$ with $p(b) = 0$, i.e. $b^n + a_1b^{n-1} + \ldots + a_n = 0$. Hence,

$$b^{n+r} = -(a_1b^{n+r-1} + \ldots + a_nb^r) \tag{5.3}$$

for all $r \geq 0$. From this you can inductively conclude that for every $m \in \mathbb{N}$ the element b^m is contained in the A-submodule of B generated by $1, b, \ldots, b^{n-1}$, i.e.

$$A[b] = A\{1, b, \ldots, b^{n-1}\} . \tag{5.4}$$

In particular, $A[b]$ is a finitely generated A-module.

(2) $\Rightarrow$ (3): Simply take $A' := A[b]$.

(3) $\Rightarrow$ (4): Take $V := A'$. This is an A-module and it is faithful since $1 \in A'$.

(4) $\Rightarrow$ (1): We will use the Cayley–Hamilton Theorem 3.8.3. Let V be a faithful $A[b]$-module which is finitely generated as an A-module. Consider the A-module morphism $f \colon V \to V$, $v \mapsto bv$. Let $I := A$. We have $f(V) = bV \subseteq V$ since V is an $A[b]$-module. Hence, we are in the setting of the Cayley–Hamilton theorem and conclude that there is a monic polynomial $p = X^n + a_1X^{n-1} + \ldots + a_n \in A[X]$ with $p(f) = 0$. Since f is multiplication by b, this means that $(b^n + a_1b^{n-1} + \ldots + a_n)v = 0$ for all $v \in V$. Since V is faithful, it follows that $b^n + a_1b^{n-1} + \ldots + a_n = 0$, i.e. b is integral over A. □

We quickly need to record a simple general lemma before coming to corollaries of the theorem.

LEMMA 5.1.8. *If V is a finitely generated B-module and B is a finitely generated A-module, then V is also finitely generated as an A-module.*

PROOF. Let $\{v_1, \ldots, v_n\}$ be a generating system of V as a B-module and let $\{b_1, \ldots, b_m\}$ be a generating system of B as an A-module. If $v \in V$, then $v = \sum_{i=1}^n c_iv_i$ for certain $c_i \in B$. Moreover, $c_i = \sum_{j=1}^m a_{ij}b_j$ for certain $a_{ij} \in A$. Hence,

$$v = \sum_{i=1}^n \left(\sum_{j=1}^m a_{ij}b_j \right) v_i = \sum_{i,j} a_{ij}b_jv_i ,$$

i.e. $V = A\{b_jv_i\}_{i,j}$ is finitely generated. □

COROLLARY 5.1.9. *If $b_1, \ldots, b_n \in B$ are integral over A, then the A-algebra $A[b_1, \ldots, b_n] \subseteq B$ is a finitely generated A-module.*

PROOF. We prove this by induction on n. The case $n = 1$ is Theorem 5.1.7. Now, let $n > 1$. Let $A_r := A[b_1, \ldots, b_r]$. By induction, A_{n-1} is a finitely generated A-module. Moreover, by the $n = 1$ case, $A_n = A_{n-1}[b_n]$ is a finitely generated A_{n-1}-module. Hence, A_n is a finitely generated A-module by Lemma 5.1.8. □

COROLLARY 5.1.10. *The integral closure $\mathrm{Int}_A(B)$ of A in B is a subring of B.*

PROOF. It's clear that $1 \in \mathrm{Int}_A(B)$. Let $b, b' \in \mathrm{Int}_A(B)$. Then $A[b, b']$ is a finitely generated A-module by Corollary 5.1.9. In particular, the elements $b+b', b-b', b\cdot b' \in A[b, b']$ are contained in a subring of B which is finitely generated as an A-module, hence they are all integral over A by Theorem 5.1.7. □

As a special case of integral extensions we obtain:

LEMMA 5.1.11. *If B is a finitely generated A-module, then B is integral over A.*

Proof. This follows immediately from Theorem 5.1.7 by taking $A' = B$. □

A ring extension $A \subseteq B$ is said to be **finite** if B is a finitely generated A-module.

Example 5.1.12. The element $i = \sqrt{-1} \in \mathbb{C}$ is integral over $\mathbb{Z}$ because it is a zero of the monic integral polynomial $X^2 + 1 \in \mathbb{Z}[X]$. Moreover, $\mathbb{Z} \subseteq \mathbb{Z}[i]$ is a finite extension by Theorem 5.1.7. In particular, it is an integral extension.

Lemma 5.1.13. *Integrality is transitive: if $A \subseteq B$ and $B \subseteq C$ are two integral ring extensions, then $A \subseteq C$ is integral as well.*

Proof. Let $c \in C$. Since $B \subseteq C$ is integral, there is a monic polynomial $p = X^n + b_1X^{n-1} + \ldots + b_n \in B[X]$ with $p(c) = 0$. Since $b_i \in B = \operatorname{Int}_A(B)$, the subalgebra $B' \coloneqq A[b_1, \ldots, b_n]$ is a finitely generated A-module by Corollary 5.1.9. Clearly, c is integral over B', hence $B'[c]$ is a finitely generated B'-module. But then $B'[c]$ is also finitely generated as an A-module by Lemma 5.1.8, hence c is integral over A by Theorem 5.1.7. □

And now we can justify why $\operatorname{Int}_A(B)$ is called the integral *closure* of A in B:

Corollary 5.1.14. *The integral closure of A in B is integrally closed in B.*

Proof. The claim is that $\operatorname{Int}_{\operatorname{Int}_A(B)}(B) = \operatorname{Int}_A(B)$. We have

$$A \xhookrightarrow{\text{integral}} \operatorname{Int}_A(B) \xhookrightarrow{\text{integral}} \operatorname{Int}_{\operatorname{Int}_A(B)}(B)\ . \tag{5.5}$$

Hence, by Lemma 5.1.13 also $A \subseteq \operatorname{Int}_{\operatorname{Int}_A(B)}(B)$ is integral. This means that every $b \in \operatorname{Int}_{\operatorname{Int}_A(B)}(B)$ is integral over A, i.e. $b \in \operatorname{Int}_A(B)$. □

We finish this section by showing that taking the integral closure commutes with localization.

Lemma 5.1.15. *Let $S \subseteq A$. Then*

$$S^{-1}\operatorname{Int}_A(B) = \operatorname{Int}_{S^{-1}A}(S^{-1}B)\ . \tag{5.6}$$

Proof. Recall from Lemma 4.3.4 that S^{-1} is an exact functor. Hence, the canonical map $S^{-1}A \to S^{-1}B$, $\frac{a}{s} \mapsto \frac{a}{s}$, induced by the inclusion $A \hookrightarrow B$ is injective. We can thus view $S^{-1}A$ as a subring of $S^{-1}B$. If $b \in \operatorname{Int}_A(B)$, then surely $\frac{b}{1} \in \operatorname{Int}_{S^{-1}A}(S^{-1}B)$ since $\frac{b}{1}$ satisfies any polynomial that b does. Since $\operatorname{Int}_{S^{-1}A}(S^{-1}B)$ is a subring containing $S^{-1}A$, also $\frac{1}{s} \cdot \frac{b}{1} = \frac{b}{s} \in \operatorname{Int}_{S^{-1}A}(S^{-1}B)$ for any $s \in \overline{S}$. This shows that $\operatorname{Int}_A(B) \subseteq \operatorname{Int}_{S^{-1}A}(S^{-1}B)$.

Conversely, let $\frac{b}{s} \in \operatorname{Int}_{S^{-1}A}(S^{-1}B)$ with $b \in B$ and $s \in \overline{S}$. Then there is a polynomial

$$p = X^n + \frac{a_1}{s_1}X^{n-1} + \ldots + \frac{a_{n-1}}{s_{n-1}}X + \frac{a_n}{s_n} \in S^{-1}A[X]$$

with $a_i \in A$ and $s_i \in \overline{S}$ such that

$$0 = p\left(\frac{b}{s}\right) = \left(\frac{b}{s}\right)^n + \frac{a_1}{s_1}\left(\frac{b}{s}\right)^{n-1} + \ldots + \frac{a_n}{s_n} \in S^{-1}B\ .$$

Multiplication with $(ss_1 \cdots s_n)^n$ yields

$$0 = \frac{(bs_1 \cdots s_n)^n}{1} + \frac{a_1(ss_2 \cdots s_n)}{1}\,\frac{(bs_1 \cdots s_n)^{n-1}}{1} + \ldots + \frac{a_n(ss_1 \cdots s_{n-1})^n s_n^{n-1}}{1} \in S^{-1}B.$$

Hence, there is $u \in \overline{S}$ such that

$$0 = u\left((bs_1 \cdots s_n)^n + a_1(ss_2 \cdots s_n)(bs_1 \cdots s_n)^{n-1} + \ldots + a_n(ss_1 \cdots s_{n-1})^n s_n^{n-1}\right) \in B$$

and multiplication with u^{n-1} yields

$$0 = (bs_1 \cdots s_n u)^n + a_1(ss_2 \cdots s_n u)(bs_1 \cdots s_n u)^{n-1} + \ldots + a_n(ss_1 \cdots s_{n-1})^n s_n^{n-1} u^n \ ,$$

which shows that $bs_1 \cdots s_n u \in \operatorname{Int}_A(B)$. Since $ss_1 \cdots s_n u \in \overline{S}$, it follows that $\frac{b}{s} \in S^{-1} \operatorname{Int}_A(B)$. □

5.2. Normal domains

We consider a special situation of integrality.

Definition 5.2.1. The integral closure of a ring A in its total ring of fractions $\operatorname{Frac}(A)$ is called the **normalization** of A. A ring is called **normal** (or **integrally closed**) if it is integrally closed in its total ring of fractions.

Example 5.2.2. We have seen in Example 5.1.5 that any unique factorization domain is normal. So, the class of normal domains is an extension of the class of unique factorization domains—which is a nice but rather special class of rings.

Example 5.2.3. Let K be a field and let $B := K[t]$ be the polynomial ring in one variable t. Consider the subalgebra $A := K[t^2, t^3] \subseteq B$. Then $\operatorname{Int}_A(B) = B$. This is because t is a zero of the polynomial $X^2 - t^2 \in A[X]$, hence $t \in \operatorname{Int}_A(B)$ and therefore $\operatorname{Int}_A(B) = B$ since $\operatorname{Int}_A(B)$ is a ring. The fraction field of B is the rational function field of $K(t)$. But also $\operatorname{Frac}(A) = K(t)$ since $t = \frac{t^3}{t^2}$. It follows that A is *not* normal. Moreover,

$$B = \operatorname{Int}_A(B) \subseteq \operatorname{Int}_A(\operatorname{Frac}(B)) \subseteq \operatorname{Int}_B(\operatorname{Frac}(B)) = B \ , \tag{5.7}$$

the latter equality follows from the fact that B is a unique factorization domain and thus normal by Example 5.1.5. Hence, the normalization of A is B. This has a geometric interpretation. Remember that the geometry of a ring is best read off from a presentation of that ring, i.e. you write it as a quotient of a polynomial ring. In this case, we have

$$\begin{array}{rcl} A = K[t^2, t^3] & \xrightarrow{\sim} & K[X_1, X_2]/(X_2^2 - X_1^3) \\ t^2 & \mapsto & X_1 \\ t^3 & \mapsto & X_2 \end{array} \tag{5.8}$$

The morphism $\operatorname{Spec}(B) \to \operatorname{Spec}(A)$ induced by the inclusion $A \subseteq B$ can now be depicted as in Figure 5.2. You see that A has a “singularity” but its normalization is “smooth”—it’s just the affine line. You can thus view normalization as a kind of “resolution of singularities”. All this can be made precise but we’ll leave that to an algebraic geometry class. I just want to clarify that normalization only gives an actual resolution of singularities for curves; in higher dimensions you can still have singularities after normalizing. The high-brow statement is that normal domains are “smooth in codimension 1”.

Rings like the Gaussian integers $\mathbb{Z}[\mathrm{i}]$ studied in algebraic number theory are constructed as follows. You start with a finite extension field L of $\mathbb{Q}$. Such a field is called an **algebraic number field**. Then the **ring of integers** in L is the integral

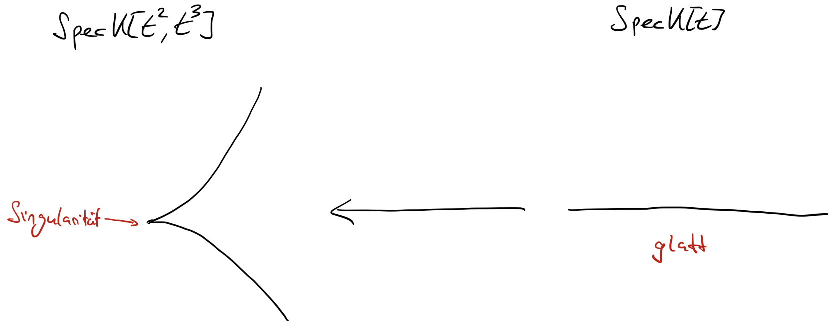

FIGURE 5.2. The normalization of $K[t^2, t^3]$.

closure $\mathcal{O}_L$ of $\mathbb{Z}$ in L. For example, $\mathbb{Z}[\mathrm{i}]$ is the integral closure of $\mathbb{Z}$ in $\mathbb{Q}(\mathrm{i})$. In Exercise 5.2.5 you will explicitly describe the ring of integers in a **quadratic field** $\mathbb{Q}(\sqrt{d})$.

Rings of integers are always normal. This is implied by the following general result.

LEMMA 5.2.4. *Let A be an integral domain with fraction field K and let L be an extension field of K which is algebraic*[1] *over L. Then the integral closure of A in L is a normal domain with fraction field L.*

PROOF. Let $\beta \in L$. Since $K \subseteq L$ is algebraic, there is a polynomial $p = X^n + \alpha_1 X^{n-1} + \ldots + \alpha_n \in K[X]$ with $p(\beta) = 0$. Since $K = \mathrm{Frac}(A)$, we can write $\alpha_i = \frac{a_i}{s_i}$ with $a_i, s_i \in A$, $s_i \neq 0$. Let $s \coloneqq s_1 \cdots s_n$ and $s_i' \coloneqq s_1 \cdots s_{i-1} s_{i+1} \cdots s_n$. Then

$$sp = sX^n + a_1 s_1' X^{n-1} + a_2 s_2' X^{n-2} + \ldots + a_n s_n' \in A[X] \;,$$

hence

$$s^n p = s^n X^n + a_1 s_1' s^{n-1} X^{n-1} + a_2 s_2' s^{n-1} X^{n-2} + \ldots + a_n s_n' s^{n-1} \in A[X] \;.$$

We have

$$0 = s^n p(\beta) = s^n \beta^n + a_1 s_1' s^{n-1} \beta^{n-1} + a_2 s_2' s^{n-1} \beta^{n-2} + \ldots + a_n s_n' s^{n-1} \;.$$

Let

$$\tilde{p} \coloneqq X^n + a_1 s_1' X^{n-1} + a_2 s_2' s X^{n-2} + \ldots + a_{n-1} s_{n-1}' s^{n-2} X + a_n s_n' s^{n-1} \in A[X] \;.$$

Then $\tilde{p}(s\beta) = 0$, i.e. $s\beta \in \mathrm{Int}_A(L)$. Hence,

$$\beta = \frac{1}{s} \cdot s\beta \in \mathrm{Frac}(\mathrm{Int}_A(L)) \;.$$

This shows that $L = \mathrm{Frac}(\mathrm{Int}_A(L))$. We know that $\mathrm{Int}_A(L)$ is integrally closed in L, hence $\mathrm{Int}_A(L)$ is normal. □

You can now ask many more questions about the ring of integers $\mathcal{O}_L$ in an algebraic number field L:

[1] This means that $K \subseteq L$ is integral (the monic condition is irrelevant for a field).

(1) Is $\mathcal{O}_L$ factorial? In general the answer is “no”, e.g. $\mathbb{Z}[\sqrt{-5}] = \mathcal{O}_{\mathbb{Q}(\sqrt{-5})}$ is not factorial (Example 1.5.18). But $\mathcal{O}_L$ is always a so-called Dedekind domain, meaning that every ideal has a unique factorization into prime ideals.
(2) Is $\mathbb{Z} \subseteq \mathcal{O}_L$ a finite extension? One can show that the answer is “yes” and this implies that $\mathcal{O}_L$ is a free $\mathbb{Z}$-module.
(3) How do the prime ideals of $\mathcal{O}_L$ look like? This question is difficult to answer explicitly in general.

We’ll not go into details about rings of integers here because this is the domain of algebraic number theory. We’ll be more concerned with general properties of integral ring extensions, but whatever we’ll prove can be applied to rings of integers as well of course.

Exercises.

Exercise 5.2.5. Let $d \in \mathbb{Z}$ be square-free, i.e. no prime number occurs with a power > 1 in the prime factorization of d. Let $L := \mathbb{Q}(\sqrt{d})$ and let $\mathcal{O}_L$ be the integral closure of $\mathbb{Z}$ in L. Show that

$$\mathcal{O}_L = \begin{cases} \mathbb{Z}[\sqrt{d}] & \text{if } d \equiv 2, 3 \bmod 4 \\ \mathbb{Z}[\frac{1+\sqrt{d}}{2}] & \text{if } d \equiv 1 \bmod 4 \ . \end{cases} \tag{5.9}$$

Hint: the proof is elementary but it involves a bit of fiddling. You can get some help from [5, p. 138].

Exercise 5.2.6. Let K be a field. In this exercise, you will compute the normalization of $A := K[X_1, X_2]/(X_2^2 - X_1^2 - X_1^3)$.
(1) Draw a picture of A.
(2) Show that A is an integral domain.
Hint: show that $X_2^2 - X_1^2 - X_1^3$ is irreducible. Why is this sufficient?
(3) Let x_i be the image of $X_i \in K[X_1, X_2]$ in A. Show that $\frac{x_2}{x_1} \in \mathrm{Frac}(A)$ is integral over A.
(4) Show that $A \subseteq K[\frac{x_2}{x_1}] \subseteq \mathrm{Frac}(A)$.
Hint: can you express x_1 and x_2 in terms of $\frac{x_2}{x_1}$?
(5) Show that $K[\frac{x_2}{x_1}]$ is the normalization of A.
Hint: we have a surjective map $\varphi\colon K[t] \to K[\frac{x_2}{x_1}]$ from the polynomial ring $k[t]$. Hence, $K[\frac{x_2}{x_1}]$ is a principal ideal domain, hence?
(6) Show that $K[\frac{x_2}{x_1}]$ is (isomorphic to) a polynomial ring in one variable and conclude that the normalization of A is the polynomial ring in one variable.
Hint: consider again the surjection $\varphi\colon K[t] \to K[\frac{x_2}{x_1}]$. Then $K[t]/\operatorname{Ker}(\varphi) \simeq K[\frac{x_2}{x_1}]$ and we want to show that $\operatorname{Ker}(\varphi) = 0$. If $\operatorname{Ker}(\varphi)$ were non-zero, then $\operatorname{Ker}(\varphi)$ would be a maximal ideal, hence...

Exercise 5.2.7. Show that being normal is a local property for integral domains.

5.3. Fibers

Let $\varphi\colon A \to B$ be a ring morphism. We want to give a more explicit description of the **fibers** of the morphism $\varphi^*\colon \mathrm{Spec}(B) \to \mathrm{Spec}(A)$ between spectra, i.e. for $P \in \mathrm{Spec}(A)$ we want to describe

$$(\varphi^*)^{-1}(P) = \{Q \in \mathrm{Spec}(B) \mid \varphi^{-1}(Q) = P\} \ . \tag{5.10}$$

Recall that for a prime ideal P of A the fraction field $k(P) = \mathrm{Frac}(A/P)$ of A/P is called the residue field of A in P. We have a morphism

$$A \twoheadrightarrow A/P \hookrightarrow k(P) \tag{5.11}$$

making $k(P)$ into an A-algebra. Since localization commutes with taking quotients we have

$$A_P/PA_P \simeq (A/P)_{P/P} = \mathrm{Frac}(A/P) = k(P)\ , \tag{5.12}$$

where you need to note that P/P is the zero ideal in A/P. If we have a ring morphism $\varphi\colon A \to B$, then we can consider B as an A-module via φ and thus localize B and φ in P:

$$B_P := (A \setminus P)^{-1}B \simeq A_P \otimes_A B\ , \tag{5.13}$$

$$\varphi_P := (A \setminus P)^{-1}\varphi\colon A_P \to B_P\ . \tag{5.14}$$

The module B_P is in fact a ring and φ_P is a ring morphism since $B_P = (\varphi(A \setminus P)^{-1})B$ is also a localization of B as a ring. Using some canonical isomorphisms you deduce that

$$k(P) \otimes_A B \simeq (A_P/PA_P) \otimes_A B = ((A_P/PA_P) \otimes_{A_P} A_P) \otimes_A B \tag{5.15}$$

$$\simeq (A_P/PA_P) \otimes_{A_P} (A_P \otimes_A B) \simeq (A_P/PA_P) \otimes_{A_P} B_P \tag{5.16}$$

$$\simeq B_P/PB_P\ . \tag{5.17}$$

We now get the following commutative diagram:

$$\begin{tikzcd}
B_Q/QB_Q = k(Q) & B_Q \arrow[l] & B \arrow[l] \arrow[r] & B/Q \arrow[r] & k(Q) \\
B_P/PB_P = k(P)\otimes_A B \arrow[u] & B_P \arrow[l] \arrow[u] & B \arrow[l] \arrow[r] \arrow[u, equal] & B/PB \arrow[r] \arrow[u] & B_P/PB_P \arrow[u] \\
A_P/PA_P = k(P) \arrow[u] & A_P \arrow[l] \arrow[u, "\varphi_P"] & A \arrow[l] \arrow[r] \arrow[u, "\varphi"'] & A/P \arrow[r] \arrow[u] & k(P) \arrow[u]
\end{tikzcd}$$

Here, we get the upper row for any prime ideal Q of B with $\varphi^{-1}(Q) = P$. When applying the Spec-functor to this diagram, we get

$$\begin{tikzcd}
\overset{(0)}{\mathrm{Spec}(k(Q))} \arrow[r] \arrow[d] & \overset{Q_Q}{\mathrm{Spec}(B_Q)} \arrow[r] \arrow[d] & \overset{Q}{\mathrm{Spec}\,B} \arrow[r] \arrow[d, equal] & \overset{(0)}{\mathrm{Spec}(k(Q))} \arrow[d] \\
\overset{X}{\mathrm{Spec}(k(P)\otimes_A B)} \arrow[r] \arrow[d] & \overset{Q_P}{\mathrm{Spec}(B_P)} \arrow[r] \arrow[d, "\varphi_P^*"'] & \overset{Q}{\mathrm{Spec}\,B} \arrow[d, "\varphi^*"] & \overset{X}{\mathrm{Spec}(k(P)\otimes_A B)} \arrow[l] \arrow[d] \\
\overset{(0)}{\mathrm{Spec}(k(P))} \arrow[r] & \overset{PA_P}{\mathrm{Spec}(A_P)} \arrow[r] & \overset{P}{\mathrm{Spec}\,A} & \overset{(0)}{\mathrm{Spec}(k(P)} \arrow[l]
\end{tikzcd}$$

Here, the horizontal morphisms always induce homeomorphisms into their image since $\mathrm{Spec}(A/P) \simeq \mathrm{V}(P)$ and $\mathrm{Spec}(A_P) \simeq \mathrm{Spec}_{A \setminus P}(A)$ by Exercise 2.5.19 and Corollary 4.2.11. The outer vertical morphisms are identical, i.e. we can glue the left and right sides of the diagram. Form this we deduce:

Lemma 5.3.1. *There is a canonical bijection*

$$\mathrm{Spec}(k(P) \otimes_A B) \simeq (\varphi^*)^{-1}(P)\ . \tag{5.18}$$

In particular,

$$(\varphi^*)^{-1}(P) \neq \emptyset \Leftrightarrow k(P) \otimes_A B \neq 0 \Leftrightarrow PB_P \neq B_P \ . \tag{5.19}$$

The $k(P)$-algebra $k(P) \otimes_A B$ is called the **scheme-theoretic fiber** of φ^* in P. Note that this has more structure than simply being a set like the usual set-theoretic fiber.

5.4. Prime ideals in integral ring extensions

It's in general very hard to say much about the fibers of $\varphi^* \colon \operatorname{Spec}(B) \to \operatorname{Spec}(A)$. But in case of an integral extension $A \subseteq B$ and the inclusion $\varphi \colon A \to B$ we can prove several fundamental results which are used all the time in commutative algebra. They are called: **lying over**, **going-up**, and **incomparability**. Throughout this section, we assume we are in this integral setting.

Theorem 5.4.1. *The morphism $\varphi^* \colon \operatorname{Spec}(B) \to \operatorname{Spec}(A)$ is* surjective*, i.e. for every $P \in \operatorname{Spec}(A)$ there is $Q \in \operatorname{Spec}(B)$* ***lying over*** *$P$, i.e.*

$$P = \varphi^*(Q) = Q \cap A \ . \tag{5.20}$$

Proof. By Lemma 5.3.1 it is enough to show that $PB_P \neq B_P$. Suppose that $PB_P = B_P$. Then we can write $1 = \sum_{i=1}^n f_i b_i$ for some $f_i \in P$ and $b_i \in B_P$. Let B' be the A_P-subalgebra of B_P generated by $b_1, \ldots, b_n$. Since $1 \in PB'$, we have $PB' = B'$. Since $A \subseteq B$ is integral, also $A_P \subseteq B_P$ is integral by Lemma 5.1.15. Hence, B' is a finitely generated A_P-module by Theorem 5.1.7. Since $PB' = B'$, we have $(PA_P)B' = B'$ and since A_P is local with maximal ideal PA_P, Nakayama's lemma (Corollary 3.8.7) implies that $B' = 0$, which is a contradiction. □

Theorem 5.4.2. *Let $P_1 \subseteq \ldots \subseteq P_n$ be a chain in $\operatorname{Spec}(A)$ and let $Q_1 \subseteq \ldots \subseteq Q_m$ with $m < n$ be a chain in $\operatorname{Spec}(B)$ with $Q_i \cap A = P_i$ for all $i = 1, \ldots, m$. Then $Q_1 \subseteq \ldots \subseteq Q_m$ can be extended to a chain $Q_1 \subseteq \ldots \subseteq Q_n$ with $Q_i \cap A = P_i$ for all $i = 1, \ldots, n$. This process is called* ***going-up****.*

Proof. By induction and using Theorem 5.4.1, it is enough to consider the case $n = 2$ and $m = 1$, i.e. we need to show that we can complete the diagram

$$\begin{tikzcd} Q_1 \arrow[r, hook] \arrow[d, dash] & Q_2? \arrow[d, dash] \\ P_1 \arrow[r, hook] & P_2 \end{tikzcd} \tag{5.21}$$

Let $\overline{A} \coloneqq A/P_1$ and $\overline{B} \coloneqq B/Q_1$. Since $Q_1 \cap A = P_1$, we have $\overline{A} \subseteq \overline{B}$. The extension $\overline{A} \subseteq \overline{B}$ is integral since $A \subseteq B$ is integral. As $P_1 \subseteq P_2$, the image $\overline{P}_2$ of P_2 in $\overline{A}$ is a prime ideal. Hence, by Theorem 5.4.1 there is $\overline{Q}_2 \in \operatorname{Spec}(\overline{B})$ with $\overline{Q}_2 \cap \overline{A} = \overline{P}_2$. We can write $\overline{Q}_2 = Q_2/Q_1$ for some $Q_2 \in \operatorname{Spec}(B)$ with $Q_1 \subseteq Q_2$, and we have $Q_2 \cap A = P_2$. □

Theorem 5.4.3. *The prime ideals in a fiber of φ^* are* ***incomparable****, i.e. if $Q_1, Q_2 \in (\varphi^*)^{-1}(P)$ are distinct, then $Q_1 \nsubseteq Q_2$ and $Q_2 \nsubseteq Q_1$.*

For the proof, we'll need an elementary lemma.

Lemma 5.4.4. *Let $A \subseteq B$ be an integral extension of integral domains. If $0 \neq b \in B$, then there is $b' \in B$ such that $0 \neq bb' \in A$.*

Proof. Let $p = X^n + a_1X^{n-1} + \ldots + a_n \in A[X]$ with $p(b) = 0$. Since $b \neq 0$ and B is an integral domain, we must have $a_i \neq 0$ for some i. Let k be the largest index with $a_k \neq 0$, i.e. $a_i = 0$ for all $i > k$. Then we can write

$$p = X^n + a_1X^{n-1} + \ldots + a_kX^{n-k} = X^{n-k}(X^k + a_1X^{k-1} + \ldots + a_k) ,$$

hence

$$0 = b^{n-k}(b^k + a_1b^{k-1} + \ldots + a_{k-1}b + a_k) .$$

Since B is an integral domain, this implies

$$0 = b^k + a_1b^{k-1} + \ldots + a_{k-1}b + a_k .$$

We can thus write

$$a_k = b^k + a_1b^{k-1} + \ldots + a_{k-1}b = b(b^{k-1} + a_1b^{k-2} + \ldots + a_{k-1}) \in A \setminus \{0\} . \quad \square$$

Proof of Theorem 5.4.3. First, suppose that B is an integral domain and that $P = 0$. The claim is that there is no non-zero prime ideal in the fiber $(\varphi^*)^{-1}(0)$. Let $Q \in (\varphi^*)^{-1}(0)$, i.e. $Q \cap A = 0$. Suppose that $Q \neq 0$. Then we can find $0 \neq b \in Q$. By Lemma 5.4.4 there is $b' \in B$ such that $0 \neq bb' \in A$. But also $bb' \in Q$, i.e. $0 \neq bb' \in Q \cap A = 0$, which is a contradiction.

Now, consider the general case. Suppose that $Q_1, Q_2 \in (\varphi^*)^{-1}(P)$ are comparable. Without loss of generality we can assume that $Q_1 \subseteq Q_2$. Since $A \cap Q_1 = P$, we have a ring extension $A/P \subseteq B/Q_1$. This extension is integral because $A \subseteq B$ is integral. But in $A/P \subseteq B/Q_1$ both Q_1/Q_1 and Q_2/Q_1 are lying over P/P. By the first case that we discussed this implies $Q_2/Q_1 = Q_1/Q_1$ and thus $Q_2 = Q_1$. □

Corollary 5.4.5. *If $A \subseteq B$ is an integral extension of integral domains, then A is a field if and only if B is a field.*

Proof. Because of lying over and incomparability we have $\mathrm{Spec}(A) = \{(0)\}$ if and only if $\mathrm{Spec}(B) = \{(0)\}$. The claim follows from the fact that an integral domain is a field if and only if (0) is the only prime ideal. □

There's an opposite of going-up called **going-down**. This does not hold in arbitrary integral extensions, however. Here's one rather general setting where it holds.

Theorem 5.4.6. *Suppose that B is an* integral domain *and that A is* normal. *Let $P_1 \supseteq \ldots \supseteq P_n$ be a descending chain in* Spec(A) *and let $Q_1 \supseteq \ldots \supseteq Q_m$ with $m < n$ be a chain in* Spec(B) *with $Q_i \cap A = P_i$ for all $i = 1, \ldots, m$. Then $Q_1 \supseteq \ldots \supseteq Q_m$ can be extended to a chain $Q_1 \supseteq \ldots \supseteq Q_n$ with $Q_i \cap A = P_i$ for all $i = 1, \ldots, n$. This process is called* ***going-down****.*

Proof. I won't give a proof here but it's possible with the tools we have and it's not too hard. You can find a proof in [2, Theorem 5.1.6]. □

Exercises.

Exercise 5.4.7. Let $A \subseteq B$ be an integral extension. Show that if $Q \in \mathrm{Spec}(B)$ lies over $P \in \mathrm{Spec}(A)$, then Q is maximal if and only if P is maximal.

Exercise 5.4.8. Let $\varphi\colon A \to B$ be an **integral** ring morphism, i.e. the extension $\varphi(A) \subseteq B$ is integral (this generalizes integral extensions). Show that the morphism $\varphi^*\colon \mathrm{Spec}(B) \to \mathrm{Spec}(A)$ is **closed**, i.e. φ^* maps closed subsets to closed subsets.

Exercise 5.4.9. In this exercise, you will show that finite morphisms have finite fibers.

(1) Show that if A is a finite-dimensional algebra over a field and A is also an integral domain, then A is already a field.
Hint: multiplication with an element from A defines a vector space endomorphism $A \to A$.

(2) Show that if A is a finite-dimensional algebra over a field, then $\operatorname{Spec}(A)$ is finite and consists only of maximal ideals.
Hint: maximality of all prime ideals follows from the previous part. To prove finiteness note that if we have r maximal ideals, then $\dim_K A \geq r$ by the Chinese remainder theorem.

(3) Let $\varphi\colon A \to B$ be a **finite** morphism (i.e. B is a finitely generated A-module via φ; this generalizes finite extensions). Show that the fibers of $\varphi^*\colon \operatorname{Spec}(B) \to \operatorname{Spec}(A)$ are finite.

CHAPTER 6

Nullstellensatz

In this chapter, we will prove the famous **Nullstellensatz** by Hilbert which gives a complete description of the maximal ideals in a polynomial ring in finitely many variables over an algebraically closed field. You may want to look at Example 2.4.3 and Example 2.4.5 to recall some context. The Nullstellensatz is a key theorem in (classical) algebraic geometry. There are several ways to prove it. I will present here a more general approach via Jacobson algebras that I found in [5] and that I like because it applies to some arithmetic situations as well.

6.1. The Nullstellensatz via Jacobson algebras

Recall from Corollary 2.6.5 that for the radical $\sqrt{I}$ of an ideal I in a ring A we have the relation

$$\sqrt{I} = \bigcap_{\substack{P \in \mathrm{Spec}(A) \\ P \supseteq I}} P \; . \tag{6.1}$$

This basically has the interpretation that the zero set described by I is "composed" of all the field valued solutions of I—which somehow makes sense. To get this a bit more into a classical setting it would be nice if you could see everything already with *maximal* ideals (the most special points). This brings us to the following definition.

Definition 6.1.1. A ring A is called **Jacobson** if

$$\sqrt{I} = \bigcap_{\substack{M \in \mathrm{Max}(A) \\ M \supseteq I}} M$$

for every ideal I of A.

Using (6.1), the Jacobson condition is equivalent to

$$P = \bigcap_{\substack{M \in \mathrm{Max}(A) \\ M \supseteq P}} M \tag{6.2}$$

for all prime ideals P of A.

Example 6.1.2. Every field is a Jacobson ring.

Example 6.1.3. Let A be a principal ideal domain. We know from Example 2.3.4 that all non-zero prime ideals in A are maximal, so (6.2) clearly holds for all non-zero prime ideals. Hence, A is Jacobson if and only if

$$(0) = \bigcap_{M \in \mathrm{Max}(A)} M \; . \tag{6.3}$$

But the expression on the right hand side is precisely the Jacobson radical $\mathrm{Jac}(A)$ of A from Definition 2.6.7. It follows that a principal ideal domain A is a Jacobson ring if and only if $\mathrm{Jac}(A) = 0$. This shows for example that $\mathbb{Z}$ is Jacobson. Moreover, from Exercise 2.6.12 we know that if K is field, then $\mathrm{Jac}(K[X]) = \mathrm{Nil}(K[X]) = 0$, so $K[X]$ is Jacobson as well.

Example 6.1.4. A quotient of a Jacobson ring is Jacobson.

Remark 6.1.5. We will not need what I say in this remark but maybe it helps to build intuition and explains what I mean by saying that maximal ideals "see" everything of the geometry if the ring is Jacobson. A morphism $f\colon X \to Y$ of topological spaces is called a **quasi-homeomorphism** if the map $V \mapsto f^{-1}(V)$ is a bijection between the open subsets of Y and the open subsets of X. A subset X of a topological space Y is called **very dense** if the inclusion $X \to Y$ is a quasi-homeomorphism. In Exercise 6.1.12 you prove that a ring A is Jacobson if and only if the subset $\mathrm{Max}(A)$ is very dense in $\mathrm{Spec}(A)$.

We want to prove the following theorem which will imply everything else.

Theorem 6.1.6. *If R is a Jacobson ring, then any finitely generated*[1] *R-algebra A is a Jacobson ring as well. Moreover, if $\varphi\colon R \to A$ denotes the structure morphism of A and $M \in \mathrm{Max}(A)$, then $\varphi^{-1}(M) \in \mathrm{Max}(R)$ and $R/\varphi^{-1}(M) \subseteq A/M$ is a* finite *field extension.*

It will require some work proving this so it's good to have some motivation first in the form of important corollaries. In all of the following, we let K be a field and consider the polynomial ring

$$K[\boldsymbol{X}] \coloneqq K[X_1, \dots, X_n]$$

in finitely many variables.

Corollary 6.1.7. *If K is a field, then $K[\boldsymbol{X}]$ is Jacobson. Moreover, if $M \in \mathrm{Max}(K[\boldsymbol{X}])$, then $K \subseteq K[\boldsymbol{X}]/M$ is a finite field extension. The same is true for any quotient of $K[\boldsymbol{X}]$.*

Proof. Clear from Theorem 6.1.6, only note that $\varphi^{-1}(M) = (0) \in \mathrm{Max}(K)$. □

We can now deduce a complete description of the maximal ideals in a polynomial ring in finitely many variables over an algebraically closed field.

Corollary 6.1.8. *Let K be an* algebraically closed *field. If $M \in \mathrm{Max}(K[\boldsymbol{X}])$, then $K[\boldsymbol{X}]/M \simeq K$. Moreover, the map*

$$\begin{array}{ccc} K^n & \to & \mathrm{Max}(K[\boldsymbol{X}]) \\ (\alpha_1, \dots, \alpha_n) & \mapsto & (X_1 - \alpha_1, \dots, X_n - \alpha_n) \end{array} \tag{6.4}$$

is a bijection.

Proof. We already know from Example 2.4.3 (but it's basically clear) that $(X_1 - \alpha_1, \dots, X_n - \alpha_n)$ is a maximal ideal. Conversely, let M be a maximal ideal in $K[\boldsymbol{X}]$. Then $K \subseteq K[\boldsymbol{X}]/M$ is finite by Corollary 6.1.7 and since K is algebraically closed, we actually have equality $K = K[\boldsymbol{X}]/M$. Let $\psi\colon K[\boldsymbol{X}] \to K[\boldsymbol{X}]/M = K$ be the quotient map and let $\alpha_i \coloneqq \psi(X_i) \in K$. We obviously have

$$(X_1 - \alpha_1, \dots, X_n - \alpha_n) \subseteq \mathrm{Ker}(\psi) = M\ .$$

[1] This means finitely generated as an *R-algebra*, i.e. there is a surjective R-algebra morphism $R[X_1, \dots, X_n] \to A$.

Since the left hand side is a maximal ideal, we must have equality. □

And now we can finally solve the mystery why this whole topic is called the "Nullstellensatz".

Corollary 6.1.9. *If K is an* algebraically closed *field and $S \subseteq K[\boldsymbol{X}]$ is a set of polynomials with $(S) \neq K[\boldsymbol{X}]$, then there is $(\alpha_1, \dots, \alpha_n) \in K^n$ with*

$$f(\alpha_1, \dots, \alpha_n) = 0 \textit{ for all } f \in S\ . \tag{6.5}$$

Proof. Since $(S) \neq K[\boldsymbol{X}]$, there is $M \in \operatorname{Max}(K[\boldsymbol{X}])$ with $S \subseteq M$ by Corollary 2.3.7. From Corollary 6.1.8 we know that $M = (X_1 - \alpha_1, \dots, X_n - \alpha_n)$ for some $\alpha_i \in K$. Hence, if $f \in S$, then $f \in M$ and $f(\alpha_1, \dots, \alpha_n) = 0$. □

So, Corollary 6.1.9 says that any (consistent) system of polynomials in finitely many variables over an algebraically closed field has a common zero. The German word for a zero of a polynomial is "Nullstelle"—and now you know why Corollary 6.1.9 is called the **Nullstellensatz**. This is a vast generalization of the fundamental theorem of algebra. Nowadays, one also calls Nullstellensatz any of the more general results above that led us to this conclusion—up to Theorem 6.1.6.

The "only" thing that remains to be done is to prove Theorem 6.1.6, so let's get to this. One of the key conclusions of Theorem 6.1.6 is a finiteness property, namely for the residue field extension $R/\varphi^{-1}(M) \subseteq A/M$. Being Jacobson is thus likely something about finiteness. This is basically what the characterization in the following lemma is saying—even though this is not directly visible.

Lemma 6.1.10. *A ring A is Jacobson if and only if the following property holds: if $P \in \operatorname{Spec}(A)$ and $B := A/P$ contains an element $b \neq 0$ such that $B_b = \{b\}^{-1}B$ is a field, then B is already a field.*

Proof. Let's first assume that A is Jacobson. Then also $B = A/P$ is Jacobson by Example 6.1.4. Since B is an integral domain, the zero ideal $(0) \subseteq B$ is prime and therefore $(0) = \bigcap_{M \in \operatorname{Max}(B)} M$ by the Jacobson property. By Proposition 4.2.10 we have

$$\operatorname{Spec}(B_b) \simeq \operatorname{Spec}_{\{b\}} B = \{Q \in \operatorname{Spec}(B) \mid Q \cap \overline{\{b\}} = \emptyset\} = \{Q \in \operatorname{Spec}(B) \mid b \notin Q\}\ . \tag{6.6}$$

Since B_b is a field by assumption, we have $\operatorname{Spec}_{\{b\}}(B) = \{(0)\}$. It thus follows from (6.6) that $b \in Q$ for all $0 \neq Q \in \operatorname{Spec}(B)$. Suppose that $(0) \in \operatorname{Spec}(B)$ would not be maximal. Then

$$(0) = \bigcap_{M \in \operatorname{Max}(B)} M = \bigcap_{\substack{M \in \operatorname{Max}(B) \\ M \neq 0}} M \ni b\ ,$$

i.e. $b = 0$, so $B_b = 0$ which is a contradiction to B_b being a field. Hence, (0) is maximal in B and it follows that B is a field.

Conversely, assume that the stated property holds and let's show that A is Jacobson. Let $Q \in \operatorname{Spec}(A)$ and set

$$I := \bigcap_{\substack{M \in \operatorname{Max}(A) \\ M \supseteq Q}} M\ . \tag{6.7}$$

We need to show that $I = Q$. Suppose that $I \neq Q$. Then $I \supsetneq Q$. Let $b \in I \setminus Q$. Let Σ be the set of all prime ideals in A containing Q but not b. Clearly, $\Sigma \neq \emptyset$ since $Q \in \Sigma$.

The set Σ is partially ordered with respect to inclusion and it has the property that any chain $(P_i)_{i\in\mathbb{N}}$ in Σ has an upper bound. To see this, let $J := \bigcup_{i\in\mathbb{N}} P_i$. We claim that J is a prime ideal. We already know that J is an ideal because we take a union of a chain. But this is also what makes J a prime ideal since if $aa' \in J$ for some $a, a' \in A$, then $aa' \in P_i$ for some i and since P_i is prime, it follows that $a \in P_i$ or $a' \in P_i$, hence $a \in J$ or $a' \in J$. It's clear that J does not contain b, so $J \in \Sigma$. We can thus apply Zorn's lemma to get a maximal element $P \in \Sigma$. The prime ideal P is not a maximal ideal since if it were, then P would be among the intersection in (6.7), which then would imply $b \notin I$. Hence, $B := A/P$ is *not* a field. The image $\overline{b}$ of b in B is non-zero and by Proposition 4.2.10 we have

$$\operatorname{Spec}(A_b) \simeq \{P' \in \operatorname{Spec}(A) \mid b \notin P'\} \supseteq \Sigma \ni P\ .$$

Since P is maximal in Σ, it is also maximal in $\operatorname{Spec}(A_b)$ since if $P' \in \operatorname{Spec}(A_b)$ with $P' \supseteq P$, then $P' \supseteq Q$, so $P' \in \Sigma$. Now, P being maximal in $\operatorname{Spec}(A_b)$ means that

$$A_b/PA_b \simeq (A/P)_b = B_b$$

is a field. But since B is not a field, this contradicts the property of A we're assuming. □

The proof of Theorem 6.1.6 relies on a closer study of a special case as given in the following lemma.

Lemma 6.1.11. *Suppose we are in the following situation:*

1. *$R \subseteq A$ and $\varphi\colon R \to A$ is the inclusion;*
2. *R and A are integral domains;*
3. *A is generated as an R-algebra by a single element;*
4. *there is $0 \neq b \in A$ such that A_b is a field.*

Then R and A are already fields and $R \subseteq A$ is finite.

Before we prove this lemma, let's first show how to use it to prove the theorem.

Proof of Theorem 6.1.6. We will prove the theorem by induction on the number of R-algebra generators $a_1, \ldots, a_r$ of A.

Let's start with $r = 1$. Let $P \in \operatorname{Spec}(A)$ and suppose there is $0 \neq b \in B := A/P$ such that B_b is a field. Note that this is obviously the case if P is maximal. Let $N := \varphi^{-1}(P)$ and let $\overline{R} := R/N$. Both $\overline{R}$ and B are integral domains and $\overline{R} \subseteq B$. Moreover, B is an $\overline{R}$-algebra generated by a single element. Finally, $\overline{R}$ is Jacobson as a quotient of a Jacobson ring. We can thus apply Lemma 6.1.11 to $\overline{R} \subseteq B$ and conclude that both $\overline{R}$ and B are fields and that $\overline{R} \subseteq B$ is finite. This proves that A is Jacobson by Lemma 6.1.10. Moreover, in case P is maximal, also $N = \varphi^{-1}(P)$ is maximal (since $\overline{R}$ is a field) and $R/N = \overline{R} \subseteq B = A/P$ is finite. Hence, we have proven the theorem in case A is generated by a single element.

Now, assume that $r > 1$. Let A' be the R-subalgebra of A generated by $a_1, \ldots, a_{r-1}$. By induction, A' is Jacobson and also $A = A'[a_r]$ is Jacobson. Let $M \in \operatorname{Max}(A)$. Then by induction $M' := A' \cap M \in \operatorname{Max}(A')$ and $A'/M' \subseteq A/M$ is finite. Let $\varphi'\colon R \to A'$. By induction, $N := (\varphi')^{-1}(M') \in \operatorname{Max}(R)$ and $R/N \subseteq A'/M'$ is finite. We have $N = \varphi^{-1}(M)$ and then also $R/N \subseteq A/M$ is finite. □

So, this wasn't too difficult—except for we didn't prove the special case that this proof is based on yet. I advise you get a coffee first because this last piece of the puzzle really requires some work.

Proof of Lemma 6.1.11. Since A is generated by a single element and A is an integral domain, we can assume that $A = R[X]/Q$ for a prime ideal Q in $R[X]$. Since $R \subseteq A$ by assumption, we have $Q \cap R = 0$. Let $S \coloneqq R \setminus \{0\}$. Then $S^{-1}R = \operatorname{Frac}(R) =: K$ and $S^{-1}(R[X]) = K[X]$. Let $j\colon R[X] \to K[X]$ be the localization map. Since $Q \cap S = \emptyset$, we have $QK[X] \in \operatorname{Spec}(K[X])$ and $j^{-1}(QK[X]) = Q$ by Proposition 4.2.10. Hence, j induces an injective morphism

$$A = R[X]/Q \hookrightarrow K[X]/QK[X] \, . \tag{6.8}$$

We claim that $Q \neq 0$. Suppose that $Q = 0$. Then $A = R[X]$ and by assumption $A_b = R[X]_b$ would be a field. But since $R \subseteq R[X]_b$ and the latter is a field, we must have $K \subseteq R[X]_b$ and therefore $R[X]_b = K[X]_b$ is a field. But since $K[X]$ is Jacobson by Example 6.1.3, this would mean that $K[X]$ is a field as well by Lemma 6.1.10 (applied to $P = 0$), and this is clearly a contradiction. Hence, indeed $Q \neq 0$. But then $QK[X] \in \operatorname{Spec}(K[X])$ is already maximal, so $K[X]/QK[X]$ is a field and therefore

$$A_b = (R[X]/Q)_b = K[X]/QK[X] \, , \tag{6.9}$$

where the equality sign is due to the fact that $R \subseteq A \subseteq A_b$: since A_b is a field, we have $K \subseteq A_b$, hence $(K[X]/QK[X])_b \subseteq A_b$, which implies the equality. Since $K[X]$ is a principal ideal domain, we have $QK[X] = (p)$ for an irreducible polynomial $p = \alpha_n X^n + \alpha_{n-1}X^{n-1} + \ldots + \alpha_1 X + \alpha_0 \in K[X]$. When we denote the image of X in $A = R[X]/Q$ by t, then $p(t) = 0$. Note that

$$\dim_K(K[X]/QK[X]) \leq \deg(p) < \infty \, . \tag{6.10}$$

We can multiply p by the product of the denominators of its coefficients and obtain a polynomial

$$\tilde{p} \coloneqq r_n X^n + r_{n-1}X^{n-1} + \ldots + r_0 \in R[X]$$

with $\tilde{p}(t) = 0$. Hence, inverting r_n in A, we get that $t \in A_{r_n} = \{r_n\}^{-1}A$ is integral over $R_{r_n} = \{r_n\}^{-1}R$. This implies that the whole extension $R_{r_n} \subseteq A_{r_n}$ is integral since A_{r_n} is generated as an R_{r_n}-algebra by t. Since $b \in A_b = K[X]/QK[X]$ and $\dim K[X]/QK[X] < \infty$, there is $q \coloneqq q_m X^m + \ldots + q_0 \in K[X]$ with $q(b) = 0$. Again we can multiply with the product of the denominators of the coefficients of q to get a polynomial

$$\tilde{q} \coloneqq s_m X^m + s_{m-1}X^{m-1} + \ldots + s_1 X + s_0 \in R[X]$$

with $\tilde{q}(b) = 0$. Since A is an integral domain, we can assume that $s_0 \neq 0$. Let $\beta \coloneqq b^{-1} \in K[X]/QK[X]$. Multiplication of $\tilde{q}(b) = 0$ with $(s_0 b^m)^{-1}$ yields

$$0 = \frac{s_m}{s_0} + \frac{s_{m-1}}{s_0}\beta + \ldots + \frac{s_1}{s_0}\beta^{m-1} + \beta^m \, ,$$

i.e. $\beta = b^{-1}$ is integral over R_{s_0}. Hence, $A_b = A[b^{-1}]$ is integral over $A_{s_0} \supseteq R_{s_0}$, in particular it is integral over the (even larger) ring $\{r_n, s_0\}^{-1}A = A_{r_n s_0}$. Since $R_{r_n} \subseteq A_{r_n}$ is integral, also $R_{r_n s_0} \subseteq A_{r_n s_0}$ is integral. We thus have a chain $R_{r_n s_0} \subseteq A_{r_n s_0} \subseteq A_b$ of integral extensions. From the transitivity of integrality (Lemma 5.1.13) it follows that $R_{r_n s_0} \subseteq A_b$ is integral. But then, since A_b is a field, $R_{r_n s_0}$ is a field as well by Corollary 5.4.5. By assumption, R is Jacobson and we now have $r_n s_0 \in R = R/(0)$ such that $R_{r_n s_0}$ is a field. It thus follows from Lemma 6.1.10 that R is a field as well. We thus have $R = K$ and since $\dim_K(A_b) = \dim_K(K[X]/QK[X]) < \infty$, we also have $\dim_K(A) < \infty$ because $A \subseteq A_b$. Hence, A is finite over K. But then A is a field as well by Corollary 5.4.5. □

Exercises.

Exercise 6.1.12. Show that a ring A is Jacobson if and only if the subset $\mathrm{Max}(A)$ is very dense in $\mathrm{Spec}(A)$.

CHAPTER 7

Chain conditions

Recall from Example 3.3.21 that a submodule of a finitely generated module is not necessarily finitely generated. This chapter is concerned with finiteness conditions on rings and modules which ensure that such strange things do not happen.

7.1. Chain conditions for partially ordered sets

Important finiteness conditions come from conditions on the partially ordered set $\mathrm{Sub}(V)$ of submodules of a module V. From this point of view it makes sense to discuss such properties for partially ordered sets in general. Recall the notion of a *chain* that we discussed in the context of Zorn's lemma (Lemma 2.3.5).

Definition 7.1.1. A partially ordered set $(X, \leq)$ is called **noetherian** if it satisfies the **ascending chain condition**: every ascending chain

$$x_1 \leq x_2 \leq \dots \tag{7.1}$$

in X becomes **stationary**, i.e. there is $n \in \mathbb{N}$ such that $x_n = x_m$ for all $m \geq n$.

So, the ascending chain condition is some sort of finiteness condition. Here's an equivalent formulation of this property.

Lemma 7.1.2. *A partially ordered set is noetherian if and only if every non-empty subset has a maximal element.*

Proof. Let $(X, \leq)$ be a partially ordered set. Assume that X is noetherian. Suppose that there is a non-empty subset $T \subseteq X$ which does not have a maximal element. Choose an element $x_1 \in T$. Since x_1 is not maximal in T, there is an element $x_2 \in T$ with $x_1 < x_2$. Continuing like this yields an infinite chain in X, contradicting the assumption that X is noetherian.

Conversely, assume that every non-empty subset of X has a maximal element. Then in particular for any chain $x_1 \leq x_2 \leq \dots$ the set $T \coloneqq \{x_i \mid i \in \mathbb{N}\}$ has a maximal element. But this means that the chain becomes stationary. □

Analogously, we say that $(X, \leq)$ is **artinian** if it satisfies the **descending chain condition**: every descending chain $x_1 \geq x_2 \geq \dots$ becomes stationary. Similarly as in Lemma 7.1.2 you can prove that a partially ordered set is artinian if and only if any non-empty subset has a *minimal* element.

7.2. Noetherian modules

Let's apply these concepts to modules—starting with the noetherian property. Throughout, A denotes a ring.

Definition 7.2.1. An A-module V is **noetherian** if the partially ordered set $\mathrm{Sub}(V)$ of submodules of V is noetherian, i.e. any ascending chain of submodules of V becomes stationary.

Example 7.2.2. If V has just finitely many elements, then clearly V is noetherian. In particular, finite abelian groups are noetherian (as $\mathbb{Z}$-modules).

Example 7.2.3. The $\mathbb{Z}$-module $\mathbb{Z}$ is noetherian since if $I_1 \subseteq I_2 \subseteq \ldots$ is a chain of ideals (submodules) of $\mathbb{Z}$, then we can write $I_i = (a_i)$ for some $a_i \in \mathbb{Z}$ and the relation $I_1 \subseteq I_i$ means that $a_i \mid a_1$ for all i. Since a_1 has only finitely many divisors, the chain becomes stationary. More generally, it follows from the proof of Lemma 1.5.17 that any principal ideal domain A is noetherian as an A-module.

Example 7.2.4. If R is a non-zero ring, then the polynomial ring $R[X_i \mid i \in \mathbb{N}]$ in infinitely many variables is *not* noetherian as a module over itself since

$$(X_1) \subsetneq (X_1, X_2) \subsetneq \ldots \tag{7.2}$$

is an infinite ascending chain of ideals.

The previous example shows that the noetherian property rules out our Example 3.3.21 of a finitely generated module having a non-finitely generated submodule. Indeed, being noetherian is precisely about this property:

Theorem 7.2.5. *An A-module V is noetherian if and only if every submodule of V is finitely generated.*

Proof. Assume that V is noetherian and let $U \subseteq V$ be a submodule. We need to show that U is finitely generated. Let Σ be the set of all finitely generated submodules of U. Then $\Sigma \neq \emptyset$ since $0 \in \Sigma$. Since V is noetherian, every chain in Σ has an upper bound (because it becomes stationary). Hence, we can apply Zorn's lemma and get a maximal element U_0 in Σ. Suppose that $U_0 \neq U$. Let $x \in U \setminus U_0$. Then $U_0 + Ax \in \Sigma$, and this is strictly larger than U_0, contradicting the maximality of U_0 in Σ. Hence, $U_0 = U$. In particular, U is finitely generated.

Conversely, suppose that any submodule of V is finitely generated. Let $U_1 \subseteq U_2 \subseteq \ldots$ be a chain of submodules in V. Because this is a chain, also $U \coloneqq \bigcup_{i \in \mathbb{N}} U_i$ is a submodule of V. By assumption, U is finitely generated. Let $u_1, \ldots, u_r$ be generators. Then for each i there is $n_i \in \mathbb{N}$ such that $u_i \in U_{n_i}$. Hence, setting $n \coloneqq \max_i \{n_i\}$, we have $u_i \in U_n$ for all i, implying that $U_n = U$ and therefore the chain becomes stationary. □

How do we prove that a given module is noetherian? Here are some tools.

Lemma 7.2.6. *If $0 \to V' \to V \to V'' \to 0$ is an exact sequence of A-modules, then V is noetherian if and only if V' and V'' are noetherian.*

Proof. If V is noetherian, then an ascending chain of submodules in V' or in V'' yields an ascending chain of submodules in V, hence becomes stationary, i.e. V' and V'' are noetherian. Conversely, suppose that V' and V'' are noetherian. Let $V_1 \subseteq V_2 \subseteq \ldots$ be a chain of submodules in V. Denote by $f \colon V' \to V$ and $g \colon V \to V''$ the morphisms in the exact sequence. Let $V_i' \coloneqq f^{-1}(V_i)$ and $V_i'' \coloneqq g(V_i)$. Then $V_1' \subseteq V_2' \subseteq \ldots$ is a chain in V', hence becomes stationary. Analogously, $V_1'' \subseteq V_2'' \subseteq \ldots$ is a chain in V'', hence becomes stationary. We can thus find $n \in \mathbb{N}$

with $V'_n = V'_m$ and $V''_n = V''_m$ for all $m \geq n$. We claim that also $V_n = V_m$ for all $m \geq n$. To this end, it is sufficient to show that $V_{n+1} \subseteq V_n$. Let $v_{n+1} \in V_{n+1}$. Then

$$g(v_{n+1}) \in g(V_{n+1}) = V''_{n+1} = V''_n = g(V_n)\ .$$

Hence, there is $v_n \in V_n$ with $g(v_n) = g(v_{n+1})$. But then $v_n - v_{n+1} \in \mathrm{Ker}(g) = \mathrm{Im}(f)$, so there is $v' \in V'$ with $v_n - v_{n+1} = f(v')$. Since $v_n - v_{n+1} \in V_{n+1}$, we have $v' \in f^{-1}(V_{n+1}) = V'_{n+1} = V'_n$. It follows that

$$v_n - v_{n+1} = f(v') \in f(V'_n) = f(f^{-1}(V_n)) \subseteq V_n\ ,$$

i.e. $v_{n+1} = v_n - f(v') \in V_n$. This shows that $V_{n+1} \subseteq V_n$. □

Exercises.

Exercise 7.2.7. Show that a finite direct sum of noetherian modules is noetherian.

Exercise 7.2.8. Show that if V is a noetherian A-module and $S \subseteq A$, then $S^{-1}V$ is a noetherian $S^{-1}A$-module.

Exercise 7.2.9. Show that a surjective endomorphism of a noetherian module is already an isomorphism. (Hint: consider the kernel of powers of the endomorphism.)

7.3. Noetherian rings

It would be great if we could ensure that any (necessarily finitely generated) module over a ring A is noetherian. This is a property on the ring A.

Definition 7.3.1. A ring A is **noetherian** if it is noetherian as an A-module.

Example 7.3.2. Obviously, any field is noetherian. We have seen in Example 7.2.3 that principal ideal domains are noetherian. Quotients of noetherian rings are noetherian by Lemma 7.2.6 and localizations of noetherian rings are noetherian by Exercise 7.2.8. In Example 7.2.4 we have seen that the polynomial ring in infinitely many variables is not noetherian.

Lemma 7.3.3. *A ring A is noetherian if and only if any finitely generated A-module is noetherian.*

Proof. If any finitely generated A-module is noetherian, then also A as an A-module is noetherian, i.e. A is a noetherian ring. Conversely, assume that A is noetherian and let V be a finitely generated A-module. Then there is a surjective A-module morphism $f\colon A^n \to V$ from which we get a short exact sequence $0 \to \mathrm{Ker}(f) \to A^n \to V \to 0$. Since A is a noetherian A-module, it follows from Exercise 7.2.7 that also A^n is noetherian. Hence, V is noetherian by Lemma 7.2.6. □

Corollary 7.3.4. *If A is noetherian, then any finitely generated A-module is already finitely* presented*.*

Proof. In the proof of Lemma 7.3.3 it follows from Lemma 7.2.6 that also $\mathrm{Ker}(f)$ is noetherian and thus in particular finitely generated. □

So far, basically the only examples of noetherian rings we know are fields and principal ideal domains. Maybe being noetherian is quite a special property? No: noetherian rings are abundant! This is the upshot of the following famous theorem. You should compare it (very roughly) with the abstract Nullstellensatz (Theorem 6.1.6).

Theorem 7.3.5 (**Hilbert basis theorem**). *If R is a noetherian ring, then any finitely generated R-algebra is noetherian as well.*

Proof. It is sufficient to prove the claim for $A \coloneqq R[X]$. The general claim then follows by induction and the fact that a quotient of a noetherian ring is noetherian as well. By Theorem 7.2.5 we need to show that any ideal I in $A = R[X]$ is finitely generated. The leading coefficients of polynomials $f \in I$ form an ideal J in R. Since R is noetherian, this ideal is finitely generated, i.e. $J = (r_1, \dots, r_n)$ for some $r_i \in R$. For every i choose some $f_i \in I$ with leading coefficient r_i. Let $I' \coloneqq (f_1, \dots, f_n) \trianglelefteq A$. For each i let $d_i \coloneqq \deg(f_i)$ and let $d \coloneqq \max_i\{d_i\}$. Consider the R-submodule V of A generated by $\{1, X, \dots, X^{d-1}\}$. We claim that

$$I = (I \cap V) + I' \ . \tag{7.3}$$

Before we prove this equation, we note that it implies that I is finitely generated. Namely, since V is a finitely generated R-module and R is noetherian by assumption, also V is noetherian by Lemma 7.3.3. Hence, any R-submodule of V is finitely generated. In particular, $I \cap V$ is finitely generated as an R-module. But then $I \cap V$ is obviously also finitely generated as an A-module and since I' is a finitely generated A-module by construction, it follows that I is a finitely generated A-module.

All that remains to be done is to prove (7.3). We obviously have $I \supseteq (I \cap V) + I'$. Conversely, let $f \in I$. If $\deg(f) < d$, then $f \in I \cap V \subseteq (I \cap V) + I'$. So, suppose that $\deg(f) \geq d$. Let r be the leading coefficient of f. Then $r \in J$, hence $r = \sum_{i=1}^n s_i r_i$ for some $s_i \in R$. Let

$$f' \coloneqq \sum_{i=1}^n s_i f_i X^{\deg(f) - d_i} \ .$$

Then $f' \in I'$, the degree of f' is equal to $\deg(f)$, and the leading coefficient of f' is equal to $\sum_{i=1}^n s_i r_i = r$. Hence, $\deg(f - f') < \deg(f)$ and $f - f' \in I$. Continuing like this one eventually obtains $f'' \in I'$ with $\deg(f - f'') < d$. We then have $f - f'' \in I \cap V$, i.e. $f \in (I \cap V) + I'$. □

In particular, the polynomial ring in finitely many variables over a field—and any quotient thereof—is noetherian. Fantastic!

Exercises.

Exercise 7.3.6. Find an example showing that a subring of a noetherian ring is *not* necessarily noetherian.

Exercise 7.3.7. Find an example showing that the property of rings being noetherian is *not* a local property. (Hint: consider the ring $A \coloneqq \prod_{i \in \mathbb{N}} (\mathbb{Z}/2\mathbb{Z})$ and show that this ring is locally a field, i.e. A_P is a field for any $P \in \operatorname{Spec}(A)$. To this end, use the fact that $x^2 = x$ for all $x \in A$.)

Exercise 7.3.8. Show that for a finitely generated module V over a noetherian ring A the following are equivalent:

(1) V is flat;
(2) V is projective;
(3) V is **locally free**, i.e. V_P is a free A_P-module for all $P \in \operatorname{Spec}(A)$.

Exercise 7.3.9. A topological space X is called **noetherian** if the partially ordered set of open subsets of X is noetherian. This is obviously equivalent to the partially ordered set of *closed* subsets being *artinian*. Prove the following:

(1) A subspace of a noetherian space is noetherian as well.
(2) A noetherian space has only finitely many irreducible components. (Hint: we already know that irreducible components are closed subsets. Consider the set Σ of all closed subsets not having finitely many irreducible components. The assumption $\Sigma \neq \emptyset$ leads to a contradiction after choosing a minimal element of Σ (why does this exist?).)
(3) If A is a noetherian ring, then $\operatorname{Spec}(A)$ is a noetherian space.
(4) A noetherian ring has only finitely many minimal prime ideals.

7.4. Artinian modules

So far, we considered the *ascending* chain condition. Let's turn things around. A module is called **artinian** if its partially ordered set of submodules is artinian, i.e. every *descending* chain of submodules becomes stationary.

Example 7.4.1. The $\mathbb{Z}$-module $\mathbb{Z}/n\mathbb{Z}$ is artinian for $n \geq 1$. But $\mathbb{Z}$ itself is *not* an artinian $\mathbb{Z}$-module since for example the descending chain

$$(2) \supsetneq (4) \supsetneq (8) \supsetneq \dots \tag{7.4}$$

does not become stationary. Similarly, $K[X]$ is not an artinian $K[X]$-module since the descending chain

$$(X) \supsetneq (X^2) \supsetneq (X^3) \supsetneq \dots \tag{7.5}$$

does not become stationary.

Example 7.4.2. If K is a field and V is a finite-dimensional K-vector space, then V is an artinian K-module since if $V = V_0 \supseteq V_1 \supseteq \dots$ is a descending chain, then $\dim V_i \geq \dim V_{i+1}$ and therefore the chain must become stationary.

Whereas a noetherian module is always finitely generated (see Theorem 7.2.5), this is not necessarily true for artinian modules. In Exercise 7.4.4 you can work out an example by yourself.

Analogously to Lemma 7.2.6, Exercise 7.2.7, Exercise 7.2.8, and Exercise 7.2.9 you can prove:

Lemma 7.4.3.

(1) *If* $0 \to V' \to V \to V'' \to 0$ *is an exact sequence of* A*-modules, then* V *is artinian if and only if* V' *and* V'' *are artinian.*
(2) *Finite direct sums of artinian modules are artinian.*
(3) *If* V *is an artinian* A*-module and* $S \subseteq A$*, then* $S^{-1}V$ *is an artinian* $S^{-1}A$*-module.*
(4) *An injective endomorphism of an artinian module is already an isomorphism.*

Exercises.

Exercise 7.4.4. Let p be a prime number and let $V_p := \{p\}^{-1}\mathbb{Z}/\mathbb{Z}$. Then V_p is an artinian but not noetherian (in particular not finitely generated) $\mathbb{Z}$-module. (Hint: Show that a proper submodule of V_p is generated by $1/p^n$ for some $n \in \mathbb{N}$.)

7.5. Modules of finite length

In this section, we'll take both the noetherian and artinian property together. We'll see that this gives a very nice class of modules which are "composed" of finitely many so-called simple modules. Throughout, let A be a ring.

Definition 7.5.1. Let V be an A-module. If $V_0 \supsetneq V_1 \supsetneq \ldots \supsetneq V_n$ is a chain of submodules of V with strict inclusions, then we call n the **length** of the chain. The **length** of V, denoted by $\ell_A(V)$, is the supremum of lengths of chains of submodules of V with strict inclusions. If $\ell_A(V) < \infty$, then V is said to be of **finite length**.

If the base ring A is clear from the context, we'll simply write $\ell(V)$ for $\ell_A(V)$. Note that if V is of finite length, then every ascending and every descending chain becomes stationary, i.e. V is both noetherian and artinian. Our goal is to prove that the converse holds as well.

Example 7.5.2. If V is a vector space over a field K, then $\ell_K(V) = \dim_K(V)$.

Example 7.5.3. If V is of finite length, then V is noetherian and artinian, thus in particular finitely generated. But recall from Example 7.4.1 that there are finitely generated modules which are not artinian, hence not of finite length.

There's just one module of length 0, namely the zero module. Modules of length 1 have the following equivalent characterization.

Definition 7.5.4. An A-module V is called **simple** if $V \neq 0$ and 0 and V are the only submodules of V.

Example 7.5.5. If K is a field, then the one-dimensional K-vector space K is (up to isomorphism) the only simple K-module.

Remark 7.5.6. Note that if V is simple and $0 \neq v \in V$, then $V = Av$ since otherwise we would get a non-trivial submodule. Hence, simple modules are **cyclic**, i.e. generated by a single element.

Example 7.5.7. If M is a maximal ideal in a ring A, then A/M is a simple A-module. In fact, up to isomorphism all simple A-modules are of this form. Namely, let S be a simple A-module. Then for any $0 \neq x \in S$ we have $S = Ax$ and therefore the morphism $\varphi_x \colon A \to S$, $a \mapsto ax$, is surjective. The kernel of φ_x is equal to the annihilator $\operatorname{Ann}_A(S)$ and therefore $A/\operatorname{Ann}_A(S) \simeq S$. Since S is simple, the ideal $\operatorname{Ann}_A(S)$ is maximal. If we denote by $\operatorname{Simp}(A)$ the collection of isomorphism classes of simple A-modules, then we get pairwise inverse bijections

$$\begin{array}{rcl} \operatorname{Simp}(A) & \to & \operatorname{Max}(A) \\ S & \mapsto & \operatorname{Ann}_A(S) \\ A/M & \leftarrow\!\shortmid & M\,. \end{array} \tag{7.6}$$

The thing about finite length modules is that they are "composed" of simple modules. This is made precise by the notion of composition series.

Definition 7.5.8. A **composition series** of an A-module V is a *maximal* finite chain of submodules of V with strict inclusions, i.e. a chain of the form

$$V = V_0 \supsetneq V_1 \supsetneq \ldots \supsetneq V_n = 0 \tag{7.7}$$

which cannot be further refined, i.e. the quotients V_i/V_{i+1} are simple. If S is a simple A-module, then

$$\#\{0 \leq i < n \mid S \simeq V_i/V_{i+1}\} \tag{7.8}$$

is called the **multiplicity** of S in the chain.

THEOREM 7.5.9. *For an A-module V the following are equivalent:*

(1) *V is of finite length.*
(2) *V is noetherian and artinian.*
(3) *V has a composition series.*

PROOF. We already noticed that if V is of finite length, then V is noetherian and artinian. Suppose that V is noetherian and artinian. We'll construct a composition series of V is follows. We set $V_0 \coloneqq V$. Since V is noetherian, we can find a maximal submodule V_1 of V_0. Note that V_1 being a maximal submodule of V_0 means precisely that the quotient V_0/V_1 is simple. We continue like this and obtain a chain $V = V_0 \supsetneq V_1 \supsetneq \ldots$. Since V is artinian, this chain becomes stationary, i.e. there is $n \in \mathbb{N}$ with $V_n = 0$. Our chain is thus a composition series.

Now, suppose that V has a composition series. We want to show that V is of finite length. Denote by $\ell'(V)$ the minimum of the lengths of composition series in V. Obviously, $\ell'(V) \leq \ell(V)$. We will show that $\ell'(V) = \ell(V)$, which implies that $\ell(V) < \infty$. First, we claim that if $V' \subsetneq V$, then $\ell'(V') < \ell'(V)$. To this end, let $V = V_0 \supsetneq V_1 \supsetneq \ldots \supsetneq V_n = 0$ be a composition series of minimal length, i.e. $n = \ell'(V)$. Let $V_i' \coloneqq V' \cap V_i$. We then have a chain $V' = V_0' \supseteq V_1' \supseteq \ldots \supseteq V_n' = 0$. The kernel of the canonical map $V_i' \to V_i \to V_i/V_{i+1}$ is $V_i' \cap V_{i+1} = V_{i+1}'$. We thus get an injective map $V_i'/V_{i+1}' \to V_i/V_{i+1}$, i.e. we can identify V_i'/V_{i+1}' with a submodule of V_i/V_{i+1}. Since the latter is simple by assumption, we must have $V_i'/V_{i+1}' = 0$ or $V_i'/V_{i+1}' = V_i/V_{i+1}$, i.e. $V_i' = V_{i+1}'$ or V_i'/V_{i+1}' is simple. Hence, if we remove from the chain $V' = V_0' \supseteq V_1' \supseteq \ldots V_n' = 0$ all terms ocurring multiple times, we get a compositon series of V'. This shows that $\ell'(V') \leq \ell'(V)$. But suppose that $\ell'(V') = \ell'(V)$. Then none of the V_i' occurs multiple times, so $V_i'/V_{i+1}' = V_i/V_{i+1}$ for all i. This implies $V_{n-1}' = V_{n-1}'/V_n' = V_{n-1}/V_n = V_{n-1}$ and inductively $V' = V_0' = V_0 = V$, which is a contradiction. Hence, $\ell'(V') < \ell'(V)$. Now, if $V = V_0 \supsetneq V_1 \supsetneq \ldots \supsetneq V_n = 0$ is any chain in V, then by what we have just proved we have

$$\ell'(V) > \ell'(V_1) > \ldots > \ell'(V_n) = 0 \ ,$$

i.e. $\ell'(V) \geq n$ and therefore $\ell'(V) \geq \ell(V)$. Hence, $\ell'(V) = \ell(V)$. □

COROLLARY 7.5.10. *If V is of finite length, then:*

(1) *All composition series of V have the same length, namely $\ell(V)$.*
(2) *Every chain of submodules in V can be refined to a composition series.*

PROOF. Consider a composition series of V and denote by n its length. Since $\ell(V)$ is by definition the supremum of lengths of chains of submodules, we must have $n \leq \ell(V)$. But from the proof of Theorem 7.5.9 we know that $\ell(V)$ is also equal to the minimum of the lengths of composition series of V, hence $\ell(V) \leq n$ and therefore we have equality $n = \ell(V)$.

Now, consider any chain of submodules in V and let n be its length. If $n = \ell(V)$, this chain is of maximal length and therefore it is a composition series. If $n < \ell(V)$, the chain is not maximal, i.e. we can insert further terms until $n = \ell(V)$ and then the chain is a composition series. □

Theorem 7.5.11 (**Jordan–Hölder**, strong version). *Let V be an A-module of finite length. Then there is a canonical isomorphism*

$$V \simeq \bigoplus_{M \in \operatorname{Max}(A)} V_M \tag{7.9}$$

of A-modules. If $V = V_0 \supsetneq V_1 \supsetneq \ldots \supsetneq V_n = 0$ is a composition series of V and $M \in \operatorname{Max}(A)$, then

$$\#\{0 \leq i < n \mid V_i/V_{i+1} \simeq A/M\} = \ell_{A_M}(V_M)\ . \tag{7.10}$$

Before proving the theorem, let's note that it implies in particular that the multiplicity of a simple A-module S in a composition series of V is *independent* of the choice of composition series. We denote this multiplicity by $[V\colon S]$ and call the simple modules S with $[V\colon S] > 0$ the **composition factors** of V. By the theorem the composition factors are in bijection with $\operatorname{Supp}(V) \cap \operatorname{Max}(A)$, where $\operatorname{Supp}(V) = \{P \in \operatorname{Spec}(A) \mid V_P \neq 0\}$ is the support of V, see Definition 4.4.4.

Proof of Theorem 7.5.11. The proof works by induction on the length of V. First, suppose that $\ell(V) = 1$. This means that V is simple, so $V \simeq A/N$ for some maximal ideal N of A by Example 7.5.7. Since A/N is a field, the complement $A \setminus N$ acts by units on A/N. Therefore, V is naturally an A_N-module and $V_N \simeq (A/N)_N \simeq A/N \simeq V$ as A_N-modules. Let M be any other maximal ideal of A with $M \neq N$. To prove the claim we need to show that $V_M = 0$. We have $N \not\subseteq M$ and $M + N = A$. In particular, there is $n \in N$ with $n \notin M$. We then have $\frac{1}{n} \in N_M$. If $m \in M$, then $m = mn \cdot \frac{1}{n} \in N_M$, i.e. $M \subseteq N_M$ and therefore $M_M \subseteq N_M$. Hence,

$$A_M = (M+N)_M \subseteq M_M + N_M \subseteq N_M\ ,$$

i.e. $A_M = N_M$. We thus get $V_M \simeq (A/N)_M \simeq A_M/N_M = 0$.

Now, suppose that $\ell(V) > 1$. Let $V = V_0 \supsetneq V_1 \supsetneq \ldots \supsetneq V_n = 0$ be a composition series. Let $M \in \operatorname{Max}(A)$. Then by localization we get a chain

$$V_M = (V_0)_M \supseteq (V_1)_M \supseteq \ldots \supseteq (V_n)_M = 0 \tag{7.11}$$

in the A_M-module V_M. Since V_i/V_{i+1} is simple, we have $\ell(V_i/V_{i+1}) = 1$. By induction we know that

$$(V_i)_M/(V_{i+1})_M \simeq (V_i/V_{i+1})_M = \begin{cases} V_i/V_{i+1} & \text{if } M = \operatorname{Ann}_A(V_i/V_{i+1}) \\ 0 & \text{else} \end{cases} \tag{7.12}$$

as A_M-modules. We thus obtain from (7.11) a composition series of V_M as A_M-module by just keeping the terms with $M = \operatorname{Ann}_A(V_i/V_{i+1})$, i.e. $V_i/V_{i+1} \simeq A/M$. This implies in particular that

$$\#\{i \mid V_i/V_{i+1} \simeq A/M\} = \ell_{A_M}(V_M)\ . \tag{7.13}$$

It remains to show that we have a canonical isomorphism $V \simeq \bigoplus_{M \in \operatorname{Max}(A)} V_M$. By the above we know that $V_M = 0$ for all but finitely many M, hence the direct sum is finite. The localization maps $V \to V_M$ thus induce a morphism

$$\alpha\colon V \to \bigoplus_{M \in \operatorname{Max}(A)} V_M\ . \tag{7.14}$$

We have $(V_M)_M = V_M$. On the other hand, we know from the above that the only composition factor of V_M is A/M and therefore $(V_M)_N = 0$ for any maximal ideal

N with $N \neq M$. Hence, the localization

$$\alpha_N \colon V_N \to \left(\bigoplus_{M \in \operatorname{Max}(A)} V_M \right)_N \simeq \bigoplus_{M \in \operatorname{Max}(A)} (V_M)_N = V_N$$

is the identity. This shows that α is locally an isomorphism and now we use Corollary 4.4.7 to conclude that α is an isomorphism. □

Exercises.

EXERCISE 7.5.12. Show that the length ℓ of modules is an **additive function**, i.e. if $0 \to V' \to V \to V'' \to 0$ is a short exact sequence, then

$$\ell(V) = \ell(V') + \ell(V'') \ . \tag{7.15}$$

7.6. Artinian rings

Similarly as for the noetherian property, we consider rings with the artinian property.

DEFINITION 7.6.1. A ring A is **artinian** if it is artinian as an A-module.

EXAMPLE 7.6.2. Any finite-dimensional algebra over a field is artinian by Example 7.4.2. Quotients and localizations of artinian rings are again artinian by Lemma 7.4.3. The ring $\mathbb{Z}$ is not artinian by Example 7.4.1.

Similar to Lemma 7.3.3 you can prove:

LEMMA 7.6.3. *A ring A is artinian if and only if any finitely generated A-module is artinian.*

Recall from Exercise 7.4.4 that there are artinian modules which are not noetherian. For *rings* the situation is a bit nicer:

THEOREM 7.6.4. *A ring A is artinian if and only if it is noetherian and all prime ideals are maximal. In this case,* $\operatorname{Spec}(A)$ *is finite.*

PROOF. Assume that A is artinian. We first show that the zero ideal is a product of maximal ideals. Let Σ be the set of all ideals which are a product of maximal ideals. Obviously, $\Sigma \neq \emptyset$ and since A is artinian, there is a minimal element $J \in \Sigma$. We show that $J = 0$. If $M \in \operatorname{Max}(A)$, then due to the minimality of J we must have $MJ = J$. In particular, $J \subseteq M$, so

$$J \subseteq \bigcap_{M \in \operatorname{Max}(A)} M = \operatorname{Jac}(A) \ . \tag{7.16}$$

Since $J^2 \in \Sigma$, we have $J^2 = J$ because of the minimality of J. Suppose that $J \neq 0$. Let Ω be the set of all ideals I with $IJ \neq 0$. Then $\Omega \neq \emptyset$ because $J \in \Omega$ due to $J^2 = J \neq 0$. Since A is artinian, there is a minimal element I in Ω. We have $(IJ)J = IJ^2 = IJ \neq 0$, so $IJ \in \Omega$ and since $IJ \subseteq I$, we conclude that $IJ = I$ due to the minimality of I. Since $IJ \neq 0$, there is $f \in I$ with $fJ \neq 0$. Hence, $(f) \in \Omega$ and therefore $I = (f)$ because of the minimality of I. Since $fJ = IJ = I = (f)$, there is $g \in J$ with $gf = f$, i.e. $(1-g)f = 0$. We have $g \in J \subseteq \operatorname{Jac}(A)$, so $1-g$ is a unit by Lemma 2.6.8. But then we must have $f = 0$, which is a contradiction.

We have now proven that $0 = M_1 \cdots M_t$ for certain $M_i \in \operatorname{Max}(A)$. For any index s we have $M_1 \cdots M_{s+1} \subseteq M_{s+1}$ and therefore

$$V_s \coloneqq (M_1 \cdots M_s)/(M_1 \cdots M_{s+1}) \tag{7.17}$$

is an A/M_{s+1}-module. Note that $K_s \coloneqq A/M_{s+1}$ is a field, so V_s is a vector space over K_s. A subspace in V_s corresponds to an ideal in A containing $M_1 \cdots M_{s+1}$, and so a chain of subspaces in V_s corresponds to a chain of ideals in A containing $M_1 \cdots M_{s+1}$. Since A is artinian, it follows that V_s is an artinian A/M_{s+1}-module, i.e. V_s is a finite-dimensional vector space over K_s. Hence,

$$\ell_{K_s}(V_s) = \dim_{K_s}(V_s) < \infty$$

and therefore V_s has a composition series as a K_s-module. Such a composition series corresponds to a maximal chain of ideals in A between $M_1 \cdots M_{s+1}$ and $M_1 \cdots M_s$. Considering the chain

$$0 = M_1 \cdots M_t \subseteq M_1 \cdots M_{t-1} \subseteq M_1 \cdots M_{t-2} \subseteq \cdots \subseteq M_1 \subseteq A$$

it follows that the composition of the former chains for all the indices $1 \leq s < t$ yields a maximal chain of ideals in A between 0 and A, i.e. a composition series of A as an A-module. This shows that A is noetherian by Theorem 7.5.9. Finally, let $P \in \operatorname{Spec}(A)$. Then $P \supseteq 0 = M_1 \cdots M_t$ and since P is prime, it follows that $P \supseteq M_i$ for some i. Hence, $P = M_i$ is maximal and it follows that there are only finitely many prime ideals, namely the $M_1, \ldots, M_t$, which are all maximal.

Conversely, assume that A is noetherian and all prime ideals are maximal. Suppose that A is not artinian. Then the A-module A is not of finite length by Theorem 7.5.9. Let Σ be the set of all ideals I in A such that the A-module A/I is not of finite length. Then $\Sigma \neq \emptyset$ since $0 \in \Sigma$. Since A is noetherian, there is a maximal element I in Σ. We claim that I is prime. Let $a, b \in A$ with $ab \in I$ but $a \notin I$. We need to show that $b \in I$. Let

$$(I \colon a) \coloneqq \{x \in A \mid xa \in I\} \trianglelefteq A\ . \tag{7.18}$$

We then get an exact sequence

$$0 \longrightarrow A/(I \colon a) \xrightarrow{\cdot a} A/I \longrightarrow A/(I + (a)) \longrightarrow 0\ . \tag{7.19}$$

Since $I + (a) \supsetneq I$ and $I \in \Sigma$ is maximal, we have $I + (a) \notin \Sigma$, i.e. $A/(I + (a))$ is of finite length. Now, suppose that $b \notin I$. Since $ab \in I$, we have $b \in (I \colon a)$, i.e. $(I \colon a) \supsetneq I$ and therefore also $A/(I \colon a)$ is of finite length. The exact sequence (7.19) now implies that A/I is also of finite length, which is a contradiction to $I \in \Sigma$. We therefore must have $b \in I$, showing that I is prime.

Since all prime ideals of A are maximal by assumption, we conclude that I is maximal. Then A/I is a field and is therefore of finite length, which is a contradiction to $I \in \Sigma$. Hence, A must be of finite length and therefore A is artinian by Theorem 7.5.9. □

Exercises.

EXERCISE 7.6.5. Show that an artinian ring A is a finite direct product of local artinian rings—more precisely, if $P_1, \ldots, P_n$ are the prime ideals of A, then

$$A \simeq \prod_{i=1}^{n} A_{P_i}\ . \tag{7.20}$$

CHAPTER 8

Dimension theory

We have discussed how to view any ring geometrically. In geometry, we have the fundamental concept of *dimension* of a space. If you think about the affine space $\mathbb{R}^n$, $\mathbb{C}^n$, more generally K^n for a field K, I'm sure you would say the dimension is equal to n. Hence, the associated ring $K[X_1, \ldots, X_n]$ should have dimension equal to n. If you consider an (irreducible) polynomial $f \in K[X_1, X_2]$ and draw the zero set, you'll see a curve in the plane K^2. Hence, the associated ring $K[X_1, X_2]/(f)$ should have dimension equal to 1. If you take an (irreducible) polynomial $f \in K[X_1, X_2, X_3]$ and draw the zero set, you'll see a surface in 3-space. Hence the associated ring $K[X_1, X_2, X_3]/(f)$ should have dimension equal to 2. And so on.

But how can we make these ideas algebraic and precise for arbitrary commutative rings? Let's first think about why we actually consider K^n to be n-dimensional. We probably have in mind the n linearly independent directions we can move in at every point of K^n. For a general "curved" zero set we can still consider the possible directions we can move in at every point: these are the so-called "tangent directions". Intuitively, tangent directions at a point are the directions you'll fly in when you're driving with a car on your zero set, you're too fast, lose control precisely at the point and become airborne. It thus seems like the number of linearly independent tangent directions at a point is a good candidate for the notion of dimension—assuming we can formalize the notion of tangent directions (we can). But there's a little problem with this idea. Consider the example in Figure 8.1. Clearly, this is a curve, so something 1-dimensional. At any point except the intersection point marked with a red circle there's precisely 1 tangent direction—that's what we want. But in this one special point—a singularity—there are *two* linearly independent tangent directions! This change in dimension of the number of tangent direction is actually the characterizing property of a singularity. What now?

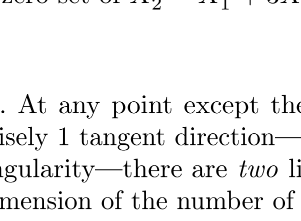

FIGURE 8.1. The zero set of $X_2^2 - X_1^3 + 3X_1 - 2 \in K[X_1, X_2]$

8.1. Prelude: the prime spectrum of $K[X_1, X_2]$

The correct notion of dimension lies in the “complexity” of the prime spectrum. Before we dive into general dimension theory and make this precise, I want to discuss a non-trivial example that will help to guide us. We have a complete understanding of the prime spectrum of a field and of a polynomial ring $K[X]$ in one variable, corresponding to the zero-dimensional and to the one-dimensional affine space over K, respectively. What about one dimension higher: the polynomial ring $K[X_1, X_2]$? It is a beautiful coincidence that for this ring too one can give a complete description of the prime spectrum! The reason this can be done is that $K[X_1, X_2]$ is isomorphic to the polynomial ring $K[X_1][X_2]$, which is a polynomial ring in *one* variable over a principal ideal domain. We can in fact give a complete description of the prime spectrum of $R[X]$ for any principal ideal domain R. This covers not just $K[X_1, X_2]$ but $\mathbb{Z}[X]$ as well!

Theorem 8.1.1. *Let R be a principal ideal domain. Then the prime ideals P in $R[X]$ are precisely the following (without overlap):*

(0) *$P = 0$,*
(1) *$P = (f)$ with $f \in R[X]$ irreducible.*
(2) *$P = (p, f)$ with $p \in R$ irreducible and $f \in R[X]$ irreducible in $(R/(p))[X]$.*

Inclusions between these only occur in the form (0) $\subsetneq$ (1) $\subsetneq$ (2). The ideals in (2) are maximal.

For the proof we’ll need some elementary facts about polynomials in one variable over a principal ideal domain (they hold more generally over a unique factorization domain) that we could have proven already in Section 1.5 and that you’ll probably know already from basic algebra. First, note that in a unique factorization domain R any two elements $a, b \in R$ have a **greatest common divisor**, i.e. an element $d \in R$ which divides both a and b, and such that every other common divisor of a and b divides d. Namely, let $p_1, \ldots, p_r$ be the collection of prime factors of a and b, so that we can write

$$a = u \prod_{i=1}^{r} p_i^{\alpha_i} , \quad b = v \prod_{i=1}^{r} p_i^{\beta_i} \tag{8.1}$$

for suitable exponents $\alpha_i, \beta_i \in \mathbb{N}$ and units u, v. Then a greatest common divisor is given by

$$\gcd(a, b) := \prod_{i=1}^{r} p_i^{\min\{\alpha_i, \beta_i\}} . \tag{8.2}$$

The collection of greatest common divisors of two elements forms a single equivalence class under the associates relation (1.105) and we usually speak of *the* greatest common divisor even though it’s really only unique up to associates. Inductively you can define the greatest common divisor of a finite set of elements. Now, if $R[X]$ is the polynomial in one variable over a unique factorization domain R and $f \in R[X]$ is a polynomial, we define the **content** $\operatorname{cont}(f)$ of f to be the greatest common divisor of the coefficients of f. One says that f is **primitive** if the content is a unit. Equipped with these concepts, we can now prove:

Lemma 8.1.2. *Let R be a unique factorization domain with fraction field K and let $R[X]$ be the polynomial ring in one variable over R.*

(1) *If $f, g \in R[X]$ are primitive, then so is fg. This is called **Gauss’s lemma**.*

(2) *If $f, g \in R[X]$, then the relation*

$$\operatorname{cont}(f) \cdot \operatorname{cont}(g) = \operatorname{cont}(fg) \tag{8.3}$$

holds (up to associates).

(3) *If $f \in R[X]$ is irreducible and non-constant, then f is also irreducible in $K[X]$.*

(4) *$R[X]$ is a unique factorization domain as well.*

Proof.

(1): Let $f = \sum_{i\in\mathbb{N}} r_i X^i$ and $g = \sum_{i\in\mathbb{N}} s_i X^i$. Then

$$fg = \sum_i \left(\sum_{k+l=i} r_k s_l \right) X^i \ .$$

Let $p \in R$ be an arbitrary prime element. Since f is primitive, there is k with $p \nmid r_k$. Similarly, there is l with $p \nmid s_l$. Let

$$k_0 \coloneqq \min\{k \mid p \nmid r_k\} \quad \text{and} \quad l_0 \coloneqq \min\{l \mid p \nmid s_l\} \ .$$

Let's have a look at the coefficient of $X^{k_0+l_0}$ in fg. This coefficient is equal to $\sum_{k+l=k_0+l_0} r_k s_l$. There are three cases to distinguish for the summands. If $k < k_0$, then $p \mid r_k$ by definition of k_0. If $k > k_0$, then $l < l_0$ and therefore $p \mid s_l$ by definition of l_0. Hence, $p \mid r_k s_l$ if $k \neq k_0$. If however $k = k_0$, then $l = l_0$, hence $p \nmid r_k$ and $p \nmid s_l$, i.e. $p \nmid r_k s_l$. We thus conclude that p does not divide the coefficient of $X^{k_0+l_0}$ in fg. Since p was an arbitrary prime element, it follows that the content of fg must be a unit, i.e. fg is primitive.

(2): We can write $f = \operatorname{cont}(f) \cdot \tilde{f}$ and $g = \operatorname{cont}(g) \cdot \tilde{g}$ for primitive polynomials $\tilde{f}, \tilde{g}$. We then have $fg = \operatorname{cont}(f) \cdot \operatorname{cont}(g) \cdot \tilde{f}\tilde{g}$. Since $\tilde{f}\tilde{g}$ is primitive by (1), the claim follows.

(3): Let $f \in R[X]$ be irreducible and non-constant. Then f is primitive because otherwise $f = \operatorname{cont}(f) \cdot \tilde{f}$ is a factorization of f into non-zero non-units. Now, suppose we can factorize $f = g_1 h_1$ in $K[X]$ into non-zero non-units $g_1, h_1 \in K[X]$. Then g_1 and h_1 are of positive degree by 1.4.10. By clearing the denominators of the coefficients, we can find $r \in R$ such that $rf = g_2 h_2$ for some $g_2, h_2 \in R[X]$. By (2) we have

$$r = r \cdot 1 = r \cdot \operatorname{cont}(f) = \operatorname{cont}(rf) = \operatorname{cont}(g_2) \cdot \operatorname{cont}(h_2) \ ,$$

hence

$$f = \frac{1}{r} \cdot g_2 h_2 = \tilde{g}_2 \cdot \tilde{h}_2 \ ,$$

where

$$\tilde{g}_2 \coloneqq \frac{1}{\operatorname{cont}(g_2)} g_2 \quad \text{and} \quad \tilde{h}_2 \coloneqq \frac{1}{\operatorname{cont}(h_2)} h_2$$

are polynomials in $R[X]$ of positive degree, thus non-units by 1.4.10 and this contradicts the irreducibility of f in $R[X]$.

(4): By 1.5.23 we need to show that $R[X]$ is atomic (every non-zero non-unit admits a factorization into irreducible elements) and that every irreducible element is prime. Let's start with the atomic property. Let $f \in R[X]$ be a non-zero non-unit. We can write $f = \operatorname{cont}(f) \cdot \tilde{f}$. Since R is a unique factorization domain, the content $\operatorname{cont}(f) \in R$ admits a factorization into irreducible elements in R. Clearly, irreducible elements in R are irreducible in $R[X]$ when considered as constant polynomials. We can thus assume that f is primitive and proceed by induction on the degree d

of f. If $d = 0$, then f is a unit and thus admits a factorization (the empty one). So, suppose that $d > 0$. If f is irreducible, there is nothing to prove. Otherwise, write $f = gh$ with non-zero non-units g, h. If one of the two, say g, has degree zero, then $1 = \operatorname{cont}(f) = g \cdot \operatorname{cont}(h)$ which contradicts the assumption that g is a non-unit. Hence, both g and h are of positive degree, therefore of degree $< d$ since R is an integral domain, and so by induction they admit a factorization into non-zero non-units. Putting these together, yields a factorization of f.

It remains to show that irreducible elements are prime. Let $f \in R[X]$ be irreducible. Then f is irreducible in $K[X]$ by (3). Hence, $f \in K[X]$ is a prime element since $K[X]$ is a principal ideal domain by 1.5.11, thus a unique factorization domain by 1.5.24 and in a unique factorization domain irreducible elements are prime by Lemma 1.5.22. We still need to show that f is prime as an element in $R[X]$. Suppose $f \mid gh$ in $R[X]$. Then $f \mid gh$ in $K[X]$ as well and since f is prime in $K[X]$ it follows that $f \mid g$ or $f \mid h$ in $K[X]$. Without loss of generality, we can assume that $f \mid g$. Then $g = qf$ for some $q \in K[X]$. After clearing denominators, we get $rg = hf$ in $R[X]$ for some $r \in R$ and $h \in R[X]$. Taking contents, we deduce $r \cdot \operatorname{cont}(g) = \operatorname{cont}(h)$, where we used that f is primitive (because it is irreducible). This means that r divides all the coefficients of h, so $\tilde{h} := \frac{1}{r} \cdot h \in R[X]$. Now, $g = \tilde{h} f \in R[X]$, i.e. $f \mid g$ in $R[X]$. □

An inductive application of Lemma 8.1.2 yields:

COROLLARY 8.1.3. *If K is a field, then $K[X_1, \ldots, X_n]$ is a unique factorization domain.*

But now let's come back to what we initially wanted to prove in this section.

PROOF OF THEOREM 8.1.1. If R is a field, then $R[X]$ is a principal ideal domain by Example 1.5.11 and we know from Corollary 2.1.5 that the prime ideals are precisely as claimed (case (2) doesn't exist). So, assume that R is a principal ideal domain which is not a field. All the listed ideals are indeed prime ideals:

(0) $P = 0$ is prime since $R[X]$ is an integral domain by Example 1.5.7.

(1) Since R is a principal ideal domain, it is a unique factorization domain by Lemma 1.5.24, so $R[X]$ is a unique factorization domain by Lemma 8.1.2. Hence, an irreducible element $f \in R[X]$ is prime by Lemma 1.5.22, hence (f) is a prime ideal in $R[X]$ by Corollary 2.1.4.

(2) We have
$$R[X]/(p, f) \simeq (R/(p))[X]/(f)\ . \tag{8.4}$$
Since R is a principal ideal domain, we know that (p) is a maximal ideal in R by Example 2.3.4. Hence, $R/(p)$ is a field, so $(R/(p))[X]$ is a principal ideal domain and since $f \in (R/(p))[X]$ is irreducible by assumption, it follows that (f) is a maximal ideal in $(R/(p))[X]$. Consequently, the quotient (8.4) is a field and therefore (p, f) is a maximal ideal in $R[X]$.

Next, we'll show that any prime ideal P in $R[X]$ belongs to one of the cases (0), (1) or (2). We can assume that $P \neq 0$. Let $\varphi\colon R \to R[X]$ be the inclusion. We know that $P \cap R = \varphi^{-1}(P) \in \operatorname{Spec}(R)$ and since R is a principal ideal domain, it follows that
$$P \cap R = (p) \quad \text{or} \quad P \cap R = (0) \tag{8.5}$$
for an irreducible element $p \in R$. Let's look at these two cases:

- $P \cap R = (p)$. Consider the quotient map $\psi\colon R[X] \to (R/(p))[X]$. Since $P \supseteq (p)$ and ψ is surjective, it follows that $\psi(P)$ is a prime ideal in $(R/(p))[X]$ and $\psi^{-1}(\psi(P)) = P$. Hence, $\psi(P) = (0)$ or $\psi(P) = (\overline{f})$ for some $f \in R[X]$ which is irreducible in $(R/(p))[X]$. This implies:
$$P = \psi^{-1}(\psi(P)) = (p) \quad \text{or} \quad P = \psi^{-1}(\psi(p)) = (p, f)\ , \tag{8.6}$$
which corresponds to the cases (1) and (2), respectively.
- $P \cap R = (0)$. Since $P \neq 0$ and $R[X]$ is noetherian by Theorem 7.3.5, it follows that P contains an irreducible element f. Namely, choose $0 \neq f' \in P$. If f' is irreducible, we can take $f = f'$. If f' is not irreducible, then $f' = ab$, hence $a \in P$ or $b \in P$. Without loss of generality, we can assume $a \in P$. Now, we continue like this with a and because $R[X]$ is noetherian, this process eventually stops and yields an irreducible element f contained in P. Suppose that $P \neq (f)$. There is an irreducible element $g \in P$ with $g \notin (f)$. Namely, choose $g' \in P$ with $g' \notin (f)$. If g' is irreducible, we take $g = g'$. If g' is not irreducible, then $g' = ab$, hence $a \in P$ or $b \in P$. Without loss of generality we assume $a \in P$. Then we cannot have $a \in (f)$ since then $g' \in (f)$. Now, we continue with a and eventually get an irreducible element $g \in P$ with $g \notin (f)$. If K denotes the fraction field of R, then both f and g are irreducible in $K[X]$ by Lemma 8.1.2. Since $(f) \neq (g)$ and $K[X]$ is a principal ideal domain, there are $a, b \in K[X]$ with $1 = af + bg$. Let $r \in R$ be the product of the denominators of the coefficients of a and b. We then get an equation $r = raf + rbg \in P$ in R. Hence, $r \in P \cap R = (0)$, i.e. $r = 0$, which is absurd. Hence, our assumption $P \neq (f)$ was wrong and we have $P = (f)$, belonging to case (1).

Finally, we argue that all the ideals are distinct and inclusions only occur in the form $(0) \subsetneq (1) \subsetneq (2)$:

- In case (1) an inclusion $(f) \subseteq (g)$ means there is h with $f = gh$ but since f and g are irreducible, this implies that h is a unit and therefore $(f) = (g)$.
- In case (2) an inclusion $(p, f) \subseteq (q, g)$ implies $p = aq + bg$ for some $a, b \in R[X]$. Since g is irreducible in $(R/(q))[X]$, it must be of positive degree. But since $\deg(p) = 0$, it follows that $b = 0$, i.e. $p = aq$. Now, since p is irreducible and q is a non-unit, it follows that a is a unit. Hence, $(p, f) = (q, f)$. Next, $(p, f) = (q, f) \subseteq (q, g)$ implies that $f = cq + dg$ for some $c, d \in R[X]$. Reducing mod q yields $\overline{f} = \overline{d}\overline{g}$ in $(R/(q))[X]$. Since $\overline{f}$ and $\overline{g}$ are irreducible by assumption, it follows that $\overline{d}$ is a unit, i.e. there is $e \in R[X]$ such that $\overline{e}\overline{f} = \overline{g}$. This means $ef = g + qh$ for some $h \in R[X]$ and therefore $g \in (q, f) = (p, f)$, i.e. $(p, f) = (q, g)$.
- Finally, suppose there is an overlap between case (1) and (2), i.e. $(f) = (q, g)$. Then $q = af$ for some $a \in R[X]$. Since q and f are irreducible, it follows that a is a unit. But then $a \in R$ and since $\deg(f) > 0$ because f is irreducible, it follows that $\deg(q) > 0$, which is a contradiction. □

Exercises.

EXERCISE 8.1.4. Theorem 8.1.1 not just applies to $K[X_1, X_2]$ but also applies to $\mathbb{Z}[X]$. David Mumford created a famous visualization of the spectrum of $\mathbb{Z}[X]$ that you can find in Figure 8.2. Explain everything you can see in the picture!

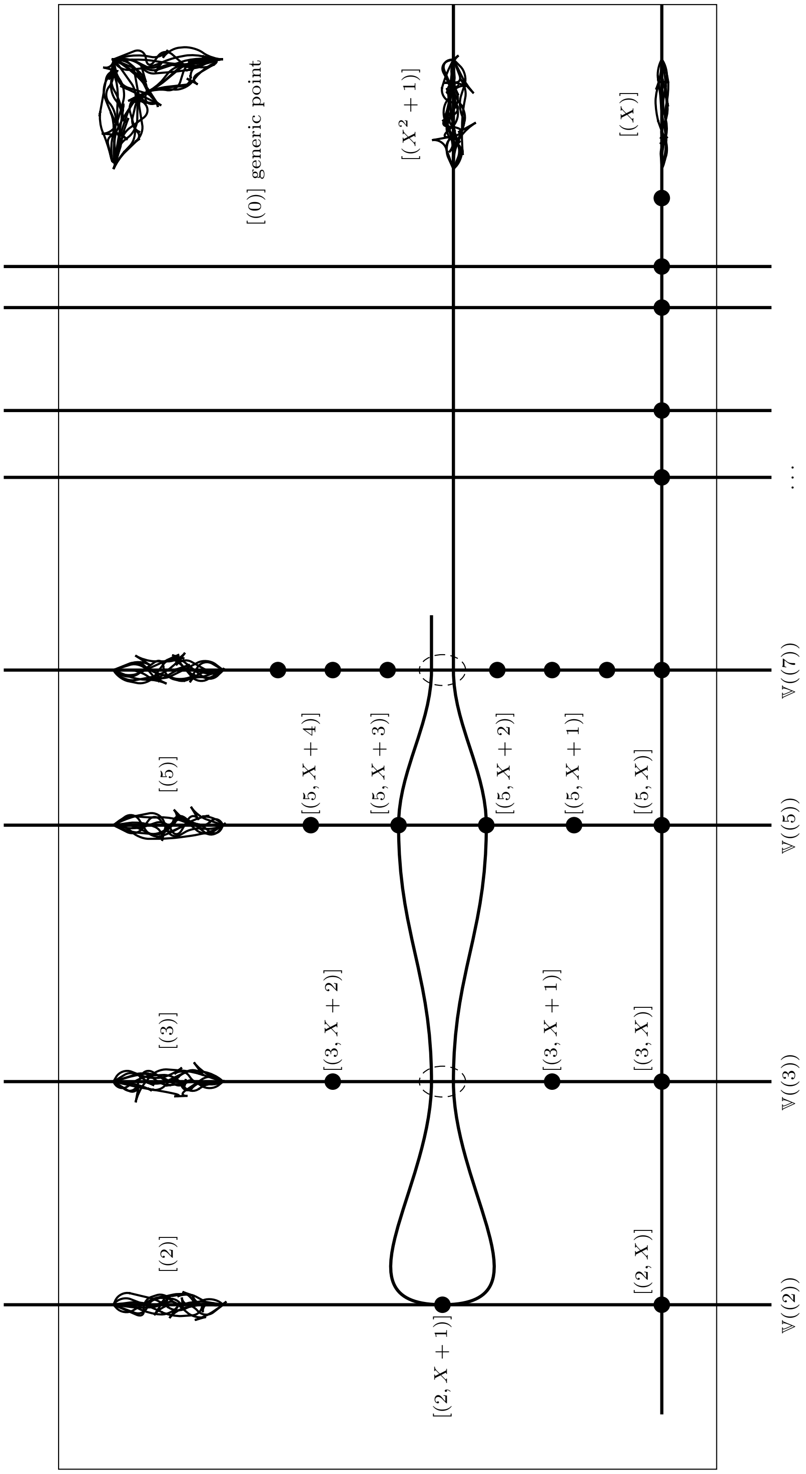

FIGURE 8.2. Mumford's visualization of $\mathrm{Spec}(\mathbb{Z}[X])$. This beautiful LaTeXversion is due to Pieter Belmans. Rotated by 90 degrees.

8.2. Krull dimension of a ring

Have a look again at Theorem 8.1.1 in the case of the polynomial ring $K[X_1, X_2]$ in two variables over a field K. We conclude that there are chains of prime ideals

$$(0) \subsetneq (f) \subsetneq (p, f) \tag{8.7}$$

in $K[X_1, X_2]$ but there is no longer chain. Hence, the supremum of lengths of chains of prime ideals in $K[X_1, X_2]$ gives us the magic number 2 which is the dimension of the corresponding affine space K^2. Interesting. Moreover, if $f \in K[X_1, X_2]$ is an irreducible polynomial, then in $K[X_1, X_2]/(f)$ we have chains

$$(f) \subsetneq (p, f) \tag{8.8}$$

and there is no longer chain. Hence, the supremum of lengths of chains of prime ideals in $K[X_1, X_2]/(f)$ gives us the magic number 1 which is the "correct" dimension of a curve as in Figure 8.1, no matter if there are singularities or not! Very interesting! This brings us—more precisely, Krull—to the following idea.

Definition 8.2.1. The (Krull) **dimension** $\dim(A)$ of a ring A is the supremum of the lengths of chains of prime ideals in A.

Example 8.2.2. The dimension of a field K, corresponding to the 0-dimensional affine space over K, is equal to 0 because there's just one prime ideal in K. More generally, by Theorem 7.6.4 the dimension of an artinian ring (e.g. a finite-dimensional algebra over a field) is equal to 0.

Example 8.2.3. We already know that in the polynomial ring $K[X]$ over a field K, corresponding to the 1-dimensional affine space over K, all non-zero prime ideals are already maximal. Hence, the dimension of $K[X]$ is equal to 1. More generally, the dimension of a principal ideal domain is equal to 1.

The beauty of the notion of dimension is that it works for *any* ring. For example, what is the dimension of $\mathbb{Z}$? It is equal to 1. Hence, $\mathbb{Z}$ is a *curve* from a dimension-theoretic point of view! This is why I (and everyone else) have drawn the prime spectrum of $\mathbb{Z}$ in Figure 5.1 in the form of a line. The following lemma gives a useful tool to determine the dimension of some rings.

Lemma 8.2.4. *If $A \subseteq B$ is an integral ring extension, then* $\dim(A) = \dim(B)$.

Proof. Left for you as Exercise 8.2.14. □

This implies that a ring of integers $\mathcal{O}_L$ in an algebraic number field L is 1-dimensional as well! You see this again in Figure 5.1 for the Gaussian integers. This is why many concepts that were developed for curves in (classical) algebraic geometry exist similarly for rings of integers in algebraic number theory—these two worlds are very similar!

Example 8.2.5. By Theorem 8.1.1, the dimension of $K[X_1, X_2]$ is equal to 2. More generally, if R is a principal ideal domain, then the dimension of $R[X]$ is equal to $2 = \dim(R) + 1$. This increase by 1 in the dimension when passing to the polynomial ring in one variable is in fact a general behavior that we'll show later.

Example 8.2.6. In the polynomial ring $K[X_1, \ldots, X_n]$ we have a chain of prime ideals

$$(0) \subsetneq (X_1) \subsetneq (X_1, X_2) \subsetneq \ldots \subsetneq (X_1, \ldots, X_n) \,, \tag{8.9}$$

hence the dimension of $K[X_1, \ldots, X_n]$ is *at least* n. Of course, this was a special choice of chain—there may be a longer chain and then Krull's notion of dimension would be a bad one. But check out the next theorem!

THEOREM 8.2.7. *If K is a field, then*

$$\dim(K[X_1, \ldots, X_n]) = n \ . \tag{8.10}$$

For the proof, we'll need one general fact.

LEMMA 8.2.8. *If A is a unique factorization domain, then all minimal non-zero primes in A are principal.*

PROOF. Let P be a minimal non-zero prime in A. Choose $0 \neq x \in P$ and let $x = up_1 \cdots p_n$ be the factorization into a unit u and prime elements p_i. Since P is prime and $x \in P$, we have $p_i \in P$ for some i. Since p_i is a prime element, the ideal (p_i) is a prime ideal. This prime ideal is non-zero and contained in P. Hence, $(p_i) = P$ by minimality of P. □

PROOF OF THEOREM 8.2.7. The proof will proceed by induction on n. The case $n = 0$ is clear, so assume $n > 0$. We have seen in Example 8.2.6 that the dimension of $K[X_1, \ldots, X_n]$ is at least equal to n. Now, let $P_0 \subsetneq P_1 \subsetneq \ldots \subsetneq P_r$ be an arbitrary chain of prime ideals. We will show that $r \leq n$, which proves the claim. If $P_0 \neq 0$, we can extend the chain to include the zero ideal, so without loss of generality we can assume that $P_0 = 0$. If P_1 is not a minimal non-zero prime ideal, we choose a minimal prime ideal P in A_{P_1} (which exists by Corollary 2.7.10) and insert this into the chain before P_1, so we can without loss of generality assume that P_1 is a minimal non-zero prime. Since $K[X_1, \ldots, X_n]$ is a unique factorization domain by Corollary 8.1.3, it follows from Lemma 8.2.8 that $P_1 = (f)$ for some $0 \neq f \in P_1$.

We will now perform a kind of coordinate transformation and discuss its properties. Let $e \in \mathbb{N}_{>1}$ and set

$$Y_i \coloneqq X_i - X_n^{e^i} \tag{8.11}$$

for $1 \leq i < n$. If $\boldsymbol{X}^\mu = X_1^{\mu_1} \cdots X_n^{\mu_n}$ is a monomial, we can rewrite it in the variables $Y_1, \ldots, Y_{n-1}, X_n$ as

$$\begin{aligned} \boldsymbol{X}^\mu &= X_1^{\mu_1} \cdots X_n^{\mu_n} = (Y_1 + X_n^e)^{\mu_1} \cdot (Y_2 + X_n^{e^2})^{\mu_2} \cdots (Y_{n-1} + X_n^{e^{n-1}})^{\mu_{n-1}} X_n^{\mu_n} \\ &= \underbrace{X_n^{\mu_1 e + \mu_2 e^2 + \ldots + \mu_{n-1} e^{n-1} + \mu_n}}_{\text{(I)}} + \underbrace{\ldots + Y_1^{\mu_1} \cdots Y_{n-1}^{\mu_{n-1}} X_n^{\mu_n}}_{\text{(II)}} \ . \end{aligned}$$

The degree of summand (I) is equal to

$$d_{\mu,e} \coloneqq \mu_1 e + \mu_2 e^2 + \ldots + \mu_{n-1} e^{n-1} + \mu_n \ . \tag{8.12}$$

The degree of summand (II) is smaller than the degree of summand (I), hence $\boldsymbol{X}^\mu$ is a monic polynomial in the variables $Y_1, \ldots, Y_{n-1}, X_n$ with leading term $X_n^{d_{\mu,e}}$. If we choose e such that $e > \mu_i$ for all i, then the μ_i are the digits in the base-e representation of the number $d_{\mu,e}$. Hence, if $\boldsymbol{X}^\mu$ and $\boldsymbol{X}^\nu$ are two distinct monomials and we choose $e > \mu_i, \nu_i$ for all i, then also $d_{\mu,e} \neq d_{\nu,e}$.

Recall that $P_1 = (f)$. We will now apply the above coordinate transformation to f. We choose e such that $e > \mu_i$ for all i and for all μ such that $\boldsymbol{X}^\mu$ is a monomial in f. Then f, written in the variables $Y_1, \dots, Y_{n-1}, X_n$, has leading term αX_n^d for some $0 \neq \alpha \in K$ and

$$d \coloneqq \max_{\mu} d_{\mu,e} \ . \tag{8.13}$$

If $\alpha \neq 1$, we can replace f by $\alpha^{-1} f$ and still have $P_1 = (f)$ and now f is a monic polynomial in the variables $Y_1, \dots, Y_{n-1}, X_n$. Let $g \in (K[Y_1, \dots, Y_{n-1}])[X]$ be the polynomial we get from f after replacing X_n by a new variable X. This is a monic polynomial and $g(X_n) = f$, i.e. $g(X_n) - f = 0$. But this means that $X_n \in K[X_1, \dots, X_n]$ is integral over the subring

$$A \coloneqq K[Y_1, \dots, Y_{n-1}, f] \subseteq K[X_1, \dots, X_n] \ . \tag{8.14}$$

Hence, the whole ring $A[X_n]$ is integral over A by Theorem 5.1.7. But since $Y_i = X_i - X_n^{e_i}$, we have

$$A[X_n] = K[Y_1, \dots, Y_{n-1}, f, X_n] = K[X_1, \dots, X_n] \ , \tag{8.15}$$

i.e. $K[X_1, \dots, X_n]$ is integral over $K[Y_1, \dots, Y_{n-1}, f] = A$. Remember that we had a chain

$$0 = P_0 \subsetneq P_1 = (f) \subsetneq P_2 \subsetneq \dots \subsetneq P_r \tag{8.16}$$

of prime ideals in $K[X_1, \dots, X_n]$. From this we obtain a chain

$$0 = P'_0 \subseteq P'_1 \subseteq \dots \subseteq P'_r \tag{8.17}$$

of prime ideals in A with $P'_i \coloneqq P_i \cap A$. Since $A \subseteq K[X_1, \dots, X_n]$ is integral, it follows from the incomparability Theorem 5.4.3 that the inclusions in (8.17) are still strict. Since $f \neq 0$, taking the quotient by (f) yields a chain of prime ideals of length $r - 1$ in

$$A/(f) = K[Y_1, \dots, Y_{n-1}, f]/(f) \simeq K[Y_1, \dots, Y_{n-1}] \ . \tag{8.18}$$

This algebra is generated by $n - 1$ elements, hence it is a quotient of the polynomial ring $K[X_1, \dots, X_{n-1}]$. The Krull dimension of a quotient of a ring is obviously at most equal to the dimension of the ring. By induction we know that $\dim(K[X_1, \dots, X_{n-1}]) = n - 1$, hence $\dim(A/(f)) \leq n - 1$ and consequently $r - 1 \leq n - 1$, i.e. $r \leq n$. □

Corollary 8.2.9. *If A is a finitely generated algebra over a field K, then $\dim(A) < \infty$.*

Remark 8.2.10. In the proof of Theorem 8.2.7 we have shown the following statement: if $f \in K[X_1, \dots, X_n]$ is a non-constant polynomial, then there are elements $Y_1, \dots, Y_{n-1} \in K[X_1, \dots, X_n]$ such that $K[X_1, \dots, X_n]$ is a finitely generated module over $K[Y_1, \dots, Y_{n-1}, f]$. One can choose $Y_i = X_i - X_n^{e^i}$ for e large enough.

We can formulate the Krull dimension of a ring A also in terms of its associated topological space Spec(A): the Krull dimension of A is the supremum of lengths of chains of irreducible closed subsets of Spec(A). We can generalize the concept of Krull dimension to arbitrary topological spaces:

Definition 8.2.11. The (Krull) **dimension** $\dim(X)$ of a topological space X is the supremum of lengths of chains of irreducible closed subsets of X.

We then recover the Krull dimension of a ring A via

$$\dim(A) = \dim(\operatorname{Spec}(A)) \ . \tag{8.19}$$

Even though we will just work with rings and their spectrum it is sometimes easier and more conceptual to work with general topological spaces when studying general aspects of dimension. From this point of view, the following is also very natural. To an ideal I of A there corresponds the closed subset $\mathrm{V}(I) = \{P \in \operatorname{Spec}(A) \mid I \subseteq P\}$ of $\operatorname{Spec}(A)$ and we define

$$\dim(I) \coloneqq \dim(\mathrm{V}(I)) \ . \tag{8.20}$$

By definition, $\dim(I)$ is the supremum of lengths of chains of irreducible closed subsets of $\mathrm{V}(I)$; which is the same as the supremum of lengths of chains

$$I \subseteq P_0 \subsetneq \ldots \subsetneq P_n \tag{8.21}$$

of prime ideals in A containing I. We obviously have

$$\dim(I) \leq \dim(A) \ . \tag{8.22}$$

Recall that $\mathrm{V}(I) \simeq \operatorname{Spec}(A/I)$ as topological spaces, so

$$\dim(I) = \dim(A/I) \ . \tag{8.23}$$

The Krull dimension is a local concept in the following sense:

Lemma 8.2.12. *For any ring A the relation*

$$\dim(A) = \sup_{P \in \operatorname{Spec}(A)} \dim(A_P) \tag{8.24}$$

holds.

Proof. We know that $\operatorname{Spec}(A_P) \simeq \{Q \in \operatorname{Spec}(A) \mid Q \subseteq P\}$. A chain of length n in $\operatorname{Spec}(A_P)$ thus corresponds to a chain of length n in $\operatorname{Spec}(A)$, hence $\dim(A_P) \leq \dim(A)$. Conversely, a chain $P_0 \subsetneq P_1 \subsetneq \ldots \subsetneq P_n \coloneqq P$ of length n in $\operatorname{Spec}(A)$ yields a chain of length n in $\operatorname{Spec}(A_P)$. □

Remark 8.2.13. You need to be a bit careful with (non-)finiteness of the dimension: there are noetherian rings A with $\dim(A) = \infty$ and there are non-noetherian rings A with $\dim(A) < \infty$.

Exercises.

Exercise 8.2.14. Prove Lemma 8.2.4.

Exercise 8.2.15. Determine the dimension of the following rings:

(1) $K[X_1, X_2, X_3]/(X_1X_2 - X_3^2)$
(2) $\mathbb{Z}_{(2)}[X]/(2X - 1)$

Exercise 8.2.16. Let X be a topological space. Prove the following:

(1) If $Y \subseteq X$, then $\dim(Y) \leq \dim(X)$.
(2) If X is irreducible with $\dim(X) < \infty$ and $Y \subsetneq X$ is a proper closed subset, then $\dim(Y) < \dim(X)$.
(3) $\dim(X) = \sup_\lambda \dim(X_\lambda)$, where the X_λ are the irreducible components of X.

8.3. Another view on Krull dimension: transcendence degree

For rings which are both an integral domain and a finitely generated algebra over a field (this is what you usually work with in algebraic geometry), there's another characterization of dimension which is not based on prime ideals but on a generalization of the dimension of a field extension: the transcendence degree. This is what people used before thinking about prime ideals and it allows one to prove many important results about the dimension because it's often easier to handle.

Throughout, let K be a field. Consider the fraction field $K(X_1, \dots, X_n)$ of the polynomial ring $K[X_1, \dots, X_n]$. This is an extension field of K but for $n > 0$ it is not of finite dimension over K. Still, we would somehow like to say that it is of "dimension" n over K—in some infinite sense. We can make this precise.

Definition 8.3.1. A family $\boldsymbol{a} \coloneqq \{a_\lambda\}_{\lambda \in \Lambda}$ of elements in a K-algebra A is called **algebraically dependent** over K if there is a non-zero polynomial $f \in K[X_\lambda \mid \lambda \in \Lambda]$ such that $f(\boldsymbol{a}) = 0$, i.e. there is a polynomial relation between the a_λ. If $\boldsymbol{a}$ is not algebraically dependent, then $\boldsymbol{a}$ is said to be **algebraically independent** over K.

Note that $\boldsymbol{a}$ is algebraically independent if and only if the K-algebra morphism $K[X_\lambda \mid \lambda \in \Lambda] \to A$ determined by $X_\lambda \mapsto a_\lambda$ is injective.

Definition 8.3.2. Let L be an extension field of K. A **transcendence basis** of L over K is a maximal element in the set of algebraically independent subsets of L over K ordered by inclusion.

Example 8.3.3. Let $\boldsymbol{X} \coloneqq \{X_1, \dots, X_n\}$. Then $\boldsymbol{X}$ is surely an algebraically independent subset of $K(\boldsymbol{X})$. We claim that $\boldsymbol{X}$ is in fact a transcendence basis of $K(\boldsymbol{X})$ over K. To this end, we need to show that $\boldsymbol{X}$ is a maximal algebraically independent subset of $K(\boldsymbol{X})$, i.e. when we add another element to $\boldsymbol{X}$, the resulting set becomes algebraically dependent. An arbitrary element of $K(\boldsymbol{X})$ is of the form $\frac{f}{g}$ with $f, g \in K[\boldsymbol{X}]$ and $g \neq 0$. Now, the map

$$\begin{array}{rcl} K[Y_1, \dots, Y_{n+1}] & \mapsto & K(\boldsymbol{X}) \\ Y_i & \mapsto & X_i \text{ for } 1 \leq i \leq n \\ Y_{n+1} & \mapsto & \frac{f}{g} \end{array} \tag{8.25}$$

is not injective since

$$g(Y_1, \dots, Y_n) \cdot Y_{n+1} - f(Y_1, \dots, Y_n) \mapsto g \cdot \frac{f}{g} - f = 0 \ . \tag{8.26}$$

Hence, the set $\boldsymbol{X} \cup \{\frac{f}{g}\}$ is algebraically dependent over K.

You need to be careful with the "basis" in "transcendence basis". For example, the following properties are obviously equivalent:

1. $K \subseteq L$ is algebraic;
2. the set $\{x\}$ is algebraically dependent over K for any $x \in L$;
3. all subsets of L are algebraically dependent over K;
4. the empty set is a transcendence basis of L over K.

The notion of transcendence basis is really made for non-algebraic (and thus infinite) extensions. In fact:

Lemma 8.3.4. *Let L be an extension field of K. An algebraically independent subset $\boldsymbol{a} \subseteq L$ over K is a transcendence basis of L over K if and only if the extension $K(\boldsymbol{a}) \subseteq L$ is algebraic, where $K(\boldsymbol{a})$ denotes the smallest extension field of K in L containing $\boldsymbol{a}$.*

Proof. This is straightforward. □

Nonetheless, there are several similarities between transcendence bases and vector space bases. Here's the big theorem about transcendence bases.

Theorem 8.3.5. *Let L be an extension field of K.*

1. *L has a transcendence basis over K.*
2. *Any two transcendence bases of L over K have the same cardinality.*
3. *Any algebraically independent subset of L over K can be extended to a transcendence basis of L over K.*

Proof.

(1): Let $\mathcal{B}$ be the set of all algebraically independent subsets of L over K. By definition, a maximal element in $\mathcal{B}$ is a transcendence basis of L over K. To deduce existence of a maximal element, we want to apply Zorn's lemma and we thus have to check the necessary assumptions on $\mathcal{B}$. Since $\emptyset \in \mathcal{B}$, the set $\mathcal{B}$ is non-empty. If $(B_\lambda)_\lambda$ is a chain in $\mathcal{B}$, then clearly $\bigcup_{\lambda \in \Lambda} B_\lambda$ is contained in $\mathcal{B}$ as well and this is a supremum of the chain in $\mathcal{B}$. We can thus apply Zorn's lemma to $\mathcal{B}$ and we are done.

(2): Let B and B' be two transcendence bases of L over K. Without loss of generality we can assume that $|B'| \leq |B|$. To prove equality, we distinguish two cases: $|B| < \infty$ and $|B| = \infty$.

First, assume that $|B| < \infty$. Let $B = \{\alpha_1, \ldots, \alpha_n\}$ and $B' = \{\beta_1, \ldots, \beta_m\}$. If $m = 0$, then $K \subseteq L$ is algebraic, hence also $n = 0$. We now assume that $m > 0$. If $B = B'$, the claim is obvious. We thus assume $B \neq B'$. Then there is β_i with $\beta_i \notin B$ as otherwise we would have $B' \subsetneq B$, contradicting the maximality of B'. Without loss of generality we can assume that $i = 1$. Then the set $B \cup \{\beta_1\}$ is algebraically dependent due to the maximality of B. Hence, there is $f \in K[X_1, \ldots, X_n, Y]$ with $f(\alpha_1, \ldots, \alpha_n, \beta_1) = 0$. Since B is algebraically independent, the variable Y must occur in f. Moreover, since β_1 is not algebraic over K, one of the variables X_j must occur in f. We assume without loss of generality that $j = 1$. Let $B^* := \{\alpha_2, \ldots, \alpha_n, \beta_1\}$. We now have a tower

$$K(B^*) \subseteq K(B^* \cup \{\alpha_1\}) \subseteq L \tag{8.27}$$

of extensions of K. The first extension $K(B^*) \subseteq K(B^* \cup \{\alpha_1\})$ is algebraic because of the relation we have involving f. The second extension $K(B^* \cup \{\alpha_1\}) \subseteq L$ is algebraic as well because $B^* \cup \{\alpha_1\}$ contains the transcendence basis B. We claim that B^* is algebraically independent over K. Assume it is not. Then there is $g \in K[X_2, \ldots, X_n, Y]$ with $g(\alpha_2, \ldots, \alpha_n, \beta_1) = 0$. Since $\{\alpha_2, \ldots, \alpha_n\}$ is algebraically independent, the variable Y must occur in g. But then β_1 is algebraic over $K(\alpha_2, \ldots, \alpha_n)$, so $K(\alpha_2, \ldots, \alpha_n) \subseteq K(B^*)$ is algebraic, hence $K(\alpha_2, \ldots, \alpha_n) \subseteq L$ is algebraic by the algebraic tower above, and this implies α_1 is algebraic over $K(\alpha_2, \ldots, \alpha_n)$ which is a contradiction. So, B^* is algebraically independent over K. Hence, $\{\alpha_2, \ldots, \alpha_n\}$ is algebraically independent over $K(\beta_1)$, and so is $\{\beta_2, \ldots, \beta_m\}$. Since the extensions

$$K(\beta_1)(\alpha_2, \ldots, \alpha_n) = K(B^*) \subseteq L \tag{8.28}$$

and

$$K(\beta_1)(\beta_2,\dots,\beta_m) = K(B') \subseteq L \tag{8.29}$$

are algebraic, the sets $\{\alpha_2,\dots,\alpha_n\}$ and $\{\beta_2,\dots,\beta_m\}$ are both transcendence bases of L over $K(\beta_1)$. The induction assumption implies that $m-1=n-1$. Hence, $m=n$, i.e. $|B'|=|B|$.

Now, we take care of the case $|B|=\infty$. Since B is a transcendence basis, the extension $K(B)\subseteq L$ is algebraic. In particular, any $\beta\in B'$ is algebraic over $K(B)$. Since an algebraic dependence relation is given by a polynomial and a polynomial involves only finitely many terms, we can find for any $\beta\in B'$ a finite subset $B_\beta\subseteq B$ such that β is algebraic over $K(B_\beta)$. Let $B^* := \bigcup_{\beta\in B'} B_\beta$. We have $B^*\subseteq B$ and claim that we actually have equality. Namely, suppose that $B^*\neq B$. Then there is $\alpha\in B\setminus B^*$. Since B' is a transcendence basis, we know that α is algebraic over $K(B')$. Moreover, by construction, $K(B')$ is algebraic over $K(B^*)$. Hence, α is algebraic over $K(B^*)$. But this means there is an algebraic dependence relation between the elements of B, contradicting the assumption that B is algebraically independent. We conclude

$$B = B^* = \bigcup_{\beta\in B'} B_\beta\ . \tag{8.30}$$

This shows that if B' were finite, then B would be finite as well, which is a contradiction. Hence, $|B'|=\infty$.

(3): Let B be an algebraically independent subset of L over K. If B is maximal in $\mathcal{B}$ (as defined in part (1)), then by definition B is a transcendence basis. If B is not maximal, then (by the proof of) part (1), B is contained in a maximal element of $\mathcal{B}$, i.e. B can be extended to a transcendence basis. □

We can now make the following definition.

Definition 8.3.6. Let A be a K-algebra which is also an integral domain. The **transcendence degree** of A over K, written $\mathrm{trdeg}_K(A)$, is the cardinality of a (any) transcendence basis of the fraction field $\mathrm{Frac}(A)$ of A over K.

Example 8.3.7. By Example 8.3.3, the transcendence degree of $K[X_1,\dots,X_n]$ over K is equal to n.

Hence, for a polynomial ring in finitely many variables over K the transcendence degree over K is the same as the Krull dimension. This is not a coincidence and holds more generally:

Theorem 8.3.8. *Let A be a finitely generated K-algebra which is also an integral domain. Then*

$$\dim(A) = \mathrm{trdeg}_K(A)\ . \tag{8.31}$$

For the proof, we'll need an extremely important tool:

Theorem 8.3.9 (**Noether normalization**, 1926). *Let A be a finitely generated K-algebra.*

(1) *There exists an algebraically independent subset $\boldsymbol{Y} := \{Y_1,\dots,Y_d\}\subseteq A$ over K such that the extension $K[\boldsymbol{Y}]\subseteq A$ is finite. Any such ring $K[\boldsymbol{Y}]$ is called a* ***Noether normalization*** *of A.*

(2) *If $I_1 \subsetneq \ldots \subsetneq I_m$ is a chain of ideals in A with dimensions $d_j := \dim(I_j)$ such that $d_1 > d_2 > \ldots > d_m > 0$, then the Y_i above can be chosen in such a way that*

$$I_j \cap K[Y_1, \ldots, Y_d] = (Y_{d_j+1}, \ldots, Y_d) \tag{8.32}$$

for all $1 \leq j \leq m$.

Note that by definition a Noether normalization $K[\boldsymbol{Y}]$ is a polynomial ring, and since the extension $K[\boldsymbol{Y}] \subseteq A$ is finite, we know that the corresponding morphism

$$\operatorname{Spec}(A) \to \operatorname{Spec}(K[\boldsymbol{Y}]) \tag{8.33}$$

is a closed and surjective map with finite fibers and we have $\dim(A) = \dim(K[\boldsymbol{Y}])$ (this follows from Theorem 5.4.1, Exercise 5.4.9, Exercise 5.4.8, and Lemma 8.2.4). In particular, the number d in the theorem is equal to the dimension of A and thus uniquely determined. Geometrically, one can think of a finite morphism like (8.33) as a "branched covering": in each fiber there are finitely many points and they "trace out some branches in Spec(A) when moving through Spec($K[\boldsymbol{Y}]$)"; this is similar to how we depicted the arithmetic example of the Gaussian integers over the integers in Figure 5.1. So, Noether normalization says that any finitely generated K-algebra is a "branched covering" of an affine space over K. The way to get there is to make an appropriate change of coordinates—this is the idea of the proof. You will notice that there are traces of Noether normalization and its proof already in the proof of Theorem 8.2.7.

One immediate and less philosophical consequence of Noether normalization is:

Proof of Theorem 8.3.8. Let $K[\boldsymbol{Y}]$ be a Noether normalization of A. Since $K[\boldsymbol{Y}] \subseteq A$ is finite, we have

$$\begin{aligned} \operatorname{trdeg}_K(A) &\overset{8.3.4}{=} \operatorname{trdeg}_K(K[\boldsymbol{Y}]) \overset{8.3.7}{=} \#\boldsymbol{Y} \\ &\overset{8.2.7}{=} \dim(K[\boldsymbol{Y}]) \overset{8.2.4}{=} \dim(A) . \end{aligned}$$ □

We should thus prove that Noether normalization indeed works.

Proof of Theorem 8.3.9. We will prove both statements at the same time. First, we assume that $A = K[X_1, \ldots, X_d]$ is a polynomial ring. Starting with $e = d$, we will inductively construct elements $Y_1^{(e)}, \ldots, Y_e^{(e)}$ and Y_{e+1} for each $0 \leq e \leq d$ satisfying the following properties:

(1) A is a finitely generated module over the subalgebra

$$B_e := K[Y_1^{(e)}, \ldots, Y_e^{(e)}, Y_{e+1}, \ldots, Y_d] ; \tag{8.34}$$

(2) $I_j \cap B_e \supseteq (Y_{h_j^{(e)}}, \ldots, Y_d)$ for all $1 \leq j \leq m$, where

$$h_j^{(e)} := \max\{d_j + 1, e + 1\} . \tag{8.35}$$

Let the games begin (better get a coffee). For $e = d$ we set $Y_i^{(e)} := X_i$ for $1 \leq i \leq d$ and $Y_{e+1} := 1$ (we actually don't need that latter element). Now, we do the induction step $e \to e - 1$. We distinguish two cases: $e > d_m$ and $e \leq d_m$.

Let's first consider the case $e > d_m$. Let k be the smallest index with $e > d_k$. Then $h_k^{(e)} = e + 1$ and $I_k \cap B_e \supseteq (Y_{e+1}, \ldots, Y_d)$. We claim that

$$I_k \cap K[Y_1^{(e)}, \ldots, Y_e^{(e)}] \neq 0 . \tag{8.36}$$

Suppose this is not true. Then $I_k \cap B_e = (Y_{e+1}, \ldots, Y_d)$ as ideals in B_e. Hence,

$$\begin{aligned}
&\dim(I_k \cap B_e)\\
&= \dim(Y_{e+1}, \ldots, Y_d)\\
&= \dim\left(K[Y_1^{(e)}, \ldots, Y_e^{(e)}, Y_{e+1}, \ldots, Y_d]/(Y_{e+1}, \ldots, Y_d)\right)\\
&= \dim(K[Y_1^{(e)}, \ldots, Y_e^{(e)}])\\
&\overset{*}{=} e\ .
\end{aligned}$$

The equality $*$ holds because $\{Y_1^{(e)}, \ldots, Y_e^{(e)}\}$ is algebraically independent over K (otherwise we would have $\operatorname{trdeg}_K(B_e) < d = \operatorname{trdeg}_K(A)$, contradicting that $B_e \subseteq A$ is finite by assumption). Now, the fact that $B_e \subseteq A$ is finite and thus integral implies that $\dim(I_k) = \dim(I_k \cap B_e)$. Hence,

$$d_k = \dim(I_k) = \dim(I_k \cap B_e) = e\ ,$$

contradicting that $e > d_k$. Hence, (8.36) holds. We can thus find a non-constant polynomial $Y_e \in I_k \cap K[Y_1^{(e)}, \ldots, Y_e^{(e)}]$. Now, recall Remark 8.2.10 (concerning the proof of Theorem 8.2.7): there are elements $Y_1^{(e-1)}, \ldots, Y_{e-1}^{(e-1)}$ such that

$$K[Y_1^{(e-1)}, \ldots, Y_{e-1}^{(e-1)}, Y_e] \subseteq K[Y_1^{(e)}, \ldots, Y_e^{(e)}] \tag{8.37}$$

is finite. With these elements and B_{e-1} defined in terms of them we get that $B_{e-1} \subseteq B_e$ is finite. Hence, $B_{e-1} \subseteq A$ is finite and therefore property (1) holds. Moreover, property (2) holds since B_{e-1} contains the elements $Y_e, \ldots, Y_d$ and $h_j^{(e)} - h_j^{(e-1)} \leq 1$ so that in the intersections we get in step $e \to e-1$ at most the additional element Y_e. This completes the induction step $e \to e-1$ in case $e > d_m$.

The other case we have to consider is $e \leq d_m$. Here, we set $Y_i^{(e-1)} := Y_i^{(e)}$ for all $1 \leq i \leq e-1$ and $Y_e := Y_e^{(e)}$. These elements obviously satisfy all the properties.

After completing this induction, we have found a subset $\boldsymbol{Y} := \{Y_1, \ldots, Y_d\}$ of A such that $K[\boldsymbol{Y}] \subseteq A$ is finite and $I_j \cap K[\boldsymbol{Y}] \supseteq (Y_{d_j+1}, \ldots, Y_d)$ for all j. That's almost exactly the claim—the only thing that remains to be proven is that we actually have equality $I_j \cap K[\boldsymbol{Y}] = (Y_{d_j+1}, \ldots, Y_d)$. As above, we can prove that

$$\dim(I_j \cap K[\boldsymbol{Y}]) = \dim(Y_{d_j+1}, \ldots, Y_d)\ . \tag{8.38}$$

Suppose that $I_j \cap K[\boldsymbol{Y}] \supsetneq (Y_{d_j+1}, \ldots, Y_d)$. Since $(Y_{d_j+1}, \ldots, Y_d)$ is a prime ideal in $K[\boldsymbol{Y}]$, we can extend every chain of prime ideals in $K[\boldsymbol{Y}]$ beginning at $I_j \cap K[\boldsymbol{Y}]$ by this additional prime ideal—but this is a contradiction to the equality in (8.38). We have now finally proven the theorem in case that A is a polynomial ring.

We still need to take care of the general case, i.e. A is any finitely generated K-algebra. Up to isomorphism we have $A = K[X_1, \ldots, X_n]/I$ for some ideal I. Let I'_j be the preimage of I_j under the quotient morphism $K[X_1, \ldots, X_d] \to A$. Let $I'_0 := I$. We then have a chain of ideals $I'_0 \subsetneq I'_1 \subsetneq \ldots \subsetneq I'_m$ in $K[X_1, \ldots, X_d]$. By the polynomial case we have proven above, there is a set $\boldsymbol{Y}' := \{Y_1, \ldots, Y_d\}$ such that $K[\boldsymbol{Y}'] \subseteq K[X_1, \ldots, X_d]$ is finite and $I'_j \cap K[\boldsymbol{Y}'] = (Y_{d_j+1}, \ldots, Y_d)$ for all j, where $d_j := \dim(I'_j) = \dim(I_j)$. In particular, setting $\boldsymbol{Y} := \{Y_1, \ldots, Y_{d_0}\}$, we have

$$I'_j \cap K[\boldsymbol{Y}] = (Y_{d_j+1}, \ldots, Y_{d_0})\ , \tag{8.39}$$

hence

$$I_0' \cap K[\boldsymbol{Y}] = 0 \ . \tag{8.40}$$

Since $K[\boldsymbol{Y}'] \subseteq K[X_1, \ldots, X_d]$ is finite, the extension

$$K[\boldsymbol{Y}] = K[Y_1, \ldots, Y_d]/(I \cap K[Y_1, \ldots, Y_d]) \subseteq K[X_1, \ldots, X_d]/I = A \tag{8.41}$$

is finite as well. The claim of the theorem holds for the subalgebra $K[\boldsymbol{Y}]$ and we are finally done. □

Remark 8.3.10. Using Noether normalization, one can also give a more down-to-earth proof of the Nullstellensatz—but of course "only" for finitely generated algebras over a field. As I said, I like our more general Nullstellensatz since it also applies to arithmetic settings.

8.4. Beware of the codimension!

If Z is a closed subset of a topological space X, then instead of considering the "absolute" dimension $\dim(Z)$ of Z it also makes sense to consider its "relative" dimension inside X. This brings us to the following dual notion of dimension.

Definition 8.4.1. Let X be a topological space and let Z be a closed subset of X. The **codimension** $\operatorname{codim}(Z, X)$ of Z in X is defined as follows:

(1) If Z is irreducible, then $\operatorname{codim}(Z, X)$ is the supremum of lengths of chains $Z = Z_0 \subsetneq Z_1 \subsetneq \ldots \subsetneq Z_n$ of irreducible closed subsets of X containing Z.
(2) If Z is arbitrary, then $\operatorname{codim}(Z, X)$ is the infimum of the codimension of the irreducible components of Z.

If the space X is clear from the context, one often simply writes $\operatorname{codim}(Z)$ instead of $\operatorname{codim}(Z, X)$. But it's important to keep in mind that—in contrast to the dimension—the codimension is really a *relative* notion and we will shortly see examples where this makes a difference.

Let's apply the codimension concept to $\operatorname{Spec}(A)$. We define the codimension of a prime ideal P in A as

$$\operatorname{codim}(P, A) \coloneqq \operatorname{codim}(\mathrm{V}(P), \operatorname{Spec}(A)) \ . \tag{8.42}$$

By definition, this is the supremum of lengths of chains

$$P_0 \subsetneq \ldots \subsetneq P_n = P \tag{8.43}$$

of prime ideals in A *ending* in P. We obviously have the relation

$$\operatorname{codim}(P, A) = \dim(A_P) \ . \tag{8.44}$$

We define the codimension of an arbitrary ideal I in A as

$$\operatorname{codim}(I, A) \coloneqq \operatorname{codim}(\mathrm{V}(I), \operatorname{Spec}(A)) = \inf_{\substack{P \in \operatorname{Spec}(A) \\ P \supseteq I \\ \text{minimal}}} \operatorname{codim}(P, A) \ . \tag{8.45}$$

Another word for codimension of an ideal is **height**.

What is the relation between the dimension and the codimension? There's the following—obvious but fundamental—inequality.

Lemma 8.4.2. *For any non-empty closed subset Z of a topological space X the relation*

$$\dim(Z) + \operatorname{codim}(Z, X) \leq \dim(X) \tag{8.46}$$

holds.

Proof. First, assume that Z is irreducible. If $Z_0 \subsetneq Z_1 \subsetneq \ldots \subsetneq Z_n = Z$ is a chain of irreducible closed subsets of Z and $Z = Z_n \subsetneq Z_{n+1} \subsetneq \ldots \subsetneq Z_{n+m}$ is a chain of irreducible closed subsets of X, then putting both chains together we get a chain of irreducible closed subsets of X of length $n+m$. Hence, $n + m \leq \dim(X)$. Since this is true for arbitrary chains, it follows that $\dim(Z) + \operatorname{codim}(Z, X) \leq \dim(X)$. Now, let Z be not necessarily irreducible. If Z_λ is an irreducible component of Z, then

$$\dim(Z_\lambda) + \operatorname{codim}(Z_\lambda, X) \leq \dim(X) \ .$$

Hence, the supremum over λ is bounded by $\dim(X)$ and therefore in particular also

$$\dim(Z) + \operatorname{codim}(Z, X) \leq \dim(X) \ . \qquad \square$$

Example 8.4.3. In linear algebra there's also a notion of **codimension** of a subspace U of a (finite-dimensional) vector space V over a field K, namely

$$\operatorname{codim}_K(U, V) \coloneqq \dim_K(V) - \dim_K(U) = \dim_K(V/U) \ . \tag{8.47}$$

I claim that the Krull versions of dimension and codimension give in the linear setting precisely the usual notions from linear algebra. After choosing a basis $\{x_1, \ldots, x_n\}$ of V we have an isomorphism $V \simeq K^n$. The ring corresponding to this affine space is the polynomial ring $K[\boldsymbol{X}] \coloneqq K[X_1, \ldots, X_n]$ and we already know that $\dim(K[\boldsymbol{X}]) = n = \dim_K(V)$. Now, a subspace U of V of codimension d can be defined by d linear equations in the coordinates of the basis, i.e. U is the common zero set of d linear polynomials. In particular, to U corresponds the prime ideal $P_U \in \operatorname{Spec}(K[\boldsymbol{X}])$ generated by these linear polynomials. After a change of basis we can assume without loss of generality that U is the subspace spanned by $\{x_{d+1}, \ldots, x_n\}$, and then $P_U = (X_1, \ldots, X_d)$. Now, it's clear that

$$\begin{aligned} \dim(P_U) &= \dim(K[X_1, \ldots, X_n]/(X_1, \ldots, X_d)) &\quad (8.48)\\ &= \dim(K[X_{d+1}, \ldots, X_n]) = n - d = \dim_K(U) \ . &\quad (8.49) \end{aligned}$$

Moreover, we have a chain

$$(0) \subsetneq (X_1) \subsetneq (X_1, X_2) \subsetneq \ldots \subsetneq (X_1, \ldots, X_d) = P_U \tag{8.50}$$

of prime ideals ending at P_U. This chain is of length d, hence $\operatorname{codim}(P_U, K[\boldsymbol{X}]) \geq d$. But

$$n = \dim(K[\boldsymbol{X}]) \geq \dim(P_U) + \operatorname{codim}(P_U, K[\boldsymbol{X}]) \geq n - d + d = n \ .$$

Therefore, we have equality

$$\dim(P_U) + \operatorname{codim}(P_U, K[\boldsymbol{X}]) = \dim(K[\boldsymbol{X}]) \tag{8.51}$$

and

$$\operatorname{codim}(P_U, K[\boldsymbol{X}]) = \operatorname{codim}_K(U, V) \ . \tag{8.52}$$

It would be great if we always had equality in (8.46) so that we can easily switch between dimension and codimension like in linear algebra. But unfortunately, things can go wrong sometimes.

Example 8.4.4. Let K be a field and let $A := K[X_1, X_2, X_3]$. Consider the prime ideals

$$P_1 := (X_1) \ , \quad P_2 := (X_2, X_3) \ , \quad P_3 := (X_1 - 1, X_2, X_3) \ . \tag{8.53}$$

Note that $P_2 \subsetneq P_3$. Let

$$B := A/(P_1 \cdot P_2) \simeq K[X_1, X_2]/(X_1X_2, X_1X_3) \ . \tag{8.54}$$

It follows that

$$\mathrm{Spec}(B) \simeq \mathrm{V}(P_1 \cdot P_2) = \mathrm{V}(P_1) \cup \mathrm{V}(P_2) \ , \tag{8.55}$$

i.e. $\mathrm{Spec}(B)$ has the irreducible components $\mathrm{V}(P_1)$ and $\mathrm{V}(P_2)$. Let's determine the dimension of the ideals P_i. We have

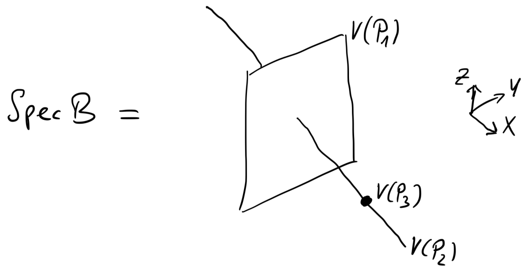

Figure 8.3. Picture of Spec(B).

$$\dim(P_1) = \dim(K[X_1, X_2, X_3]/(X_1)) = \dim(K[X_2, X_3]) = 2 \ , \tag{8.56}$$
$$\dim(P_2) = \dim(K[X_1, X_2, X_3]/(X_2, X_3)) = \dim(K[X_1]) = 1 \ , \tag{8.57}$$
$$\dim(P_3) = \dim(K[X_1, X_2, X_3]/(X_1 - 1, X_2, X_3)) \tag{8.58}$$
$$= \dim(K[X_1]/(X_1 - 1)) = \dim(K) = 0 \ . \tag{8.59}$$

Hence, using Exercise 8.2.16, we deduce that

$$\dim(B) = \sup_{i=1,2} \dim(\mathrm{V}(P_i)) = 2 \ . \tag{8.60}$$

Now, we look at codimensions. Since there is a chain

$$(0) \subsetneq (X_2) \subsetneq (X_2, X_3) \subsetneq (X_1 - 1, X_2, X_3) = P_3 \ , \tag{8.61}$$

we must have $\mathrm{codim}(P_3, A) \geq 3$. But $\dim(A) = 3$, hence

$$\mathrm{codim}(P_3, A) = 3 \tag{8.62}$$

and we have equality

$$\dim(P_3) + \mathrm{codim}(P_3, A) = \dim(A) \ . \tag{8.63}$$

Everything looking good. Since $P_2 \subseteq P_3$, we have $P_1P_2 \subseteq P_3$, so $P_3 \in \mathrm{Spec}(B)$ as well and we can also look at the situation inside $\mathrm{Spec}(B)$. Here, things become ugly. By definition, $\mathrm{codim}(P_3, B)$ is equal to the supremum of lengths of chains of prime ideals in A between P_1P_2 and P_3. Since $P_1 \not\subseteq P_3$, this is equal to the supremum of chains of prime ideals between P_2 and P_3. But

$$(A/P_2)/(P_3/P_2) \simeq K[X_1]/(X_1 - 1) \simeq K \ , \tag{8.64}$$

i.e. P_3/P_2 is maximal in A/P_2. This means $P_2 \subsetneq P_3$ is the only chain of prime ideals between P_2 and P_3, and therefore

$$\operatorname{codim}(P_3, B) = 1\ . \tag{8.65}$$

We conclude:

$$\dim(P_3) + \operatorname{codim}(P_3, B) = 0 + 1 < 2 = \dim(B)\ . \tag{8.66}$$

Because of this observation we make the following definition.

Definition 8.4.5. Let X be a topological space. We say that a closed subset Z of X satisfies the **dimension formula** in X if there is equality

$$\dim(Z) + \operatorname{codim}(Z, X) = \dim(X)\ . \tag{8.67}$$

We say that X itself satisfies the **dimension formula** if $\dim(X) < \infty$ and any non-empty[1] closed subset of X satisfies the dimension formula in X.

Lemma 8.4.6. *If $\dim(X) < \infty$ and every irreducible closed subset of X satisfies the dimension formula in X, then X satisfies the dimension formula.*

Proof. Let Z be a non-empty closed subset and let Z_λ denote the irreducible components. By assumption, we have

$$\dim(Z_\lambda) = \dim(X) - \operatorname{codim}(Z_\lambda, X)\ . \tag{8.68}$$

It follows that if λ is such that $\dim(Z_\lambda)$ is maximal (this exists since $\dim(X) < \infty$ and thus $\dim(Z) < \infty$), then $\operatorname{codim}(Z_\lambda, X)$ is smallest. For such λ we then have $\dim(Z_\lambda) = \dim(Z)$ and $\operatorname{codim}(Z_\lambda, X) = \operatorname{codim}(Z, X)$. This shows that Z satisfies the dimension formula in X. □

What causes the dimension formula to fail? If we look again at Example 8.4.4 we see that we have a 2-dimensional ring B but a maximal chain $P_2 \subsetneq P_3$ of prime ideals in B of length 1. Whenever there are maximal chains of prime ideals of distinct length, then obviously there is a prime ideal for which the dimension formula fails. We should thus eliminate such behavior! We will now introduce several conditions on a topological space that will allow us to get a hold on the dimension formula.

Definition 8.4.7. A topological space X is said to be:

(1) **equidimensional** (or **pure dimensional**) if all irreducible components of X (maximal irreducible closed subsets) have the same dimension;
(2) **equicodimensional** (or **pure codimensional**) if all minimal irreducible closed subsets of X have the same codimension in X;
(3) **catenary** if for all irreducible closed subsets $Y \subseteq Z$ of X all maximal chains of irreducible closed subsets between Y and Z have the same length;
(4) **bi-equidimensional** if $\dim(X) < \infty$ and all maximal chains of irreducible closed subsets of X have the same length.

Moreover, we introduce the following relative version of equicodimensional:

Definition 8.4.8. A closed subset Z of X is said to be **equicodimensional** (or **pure codimensional**) in X, if all irreducible components of Z are of the same codimension in X.

[1] The dimension formula doesn't really make much sense if $Z = \emptyset$: if $X \neq \emptyset$, then $\dim(Z) = \sup\emptyset = -\infty$ and $\operatorname{codim}(Z, X) = \inf\emptyset = +\infty$.

We use all these terms for a ring A, respectively an ideal I, as well by applying them to $\mathrm{Spec}(A)$, respectively to $\mathrm{V}(I)$. Be aware that the mapping V from prime ideals in A to irreducible closed subsets of $\mathrm{Spec}(A)$ reverses inclusions, so A being equidimensional means that all *minimal* primes in A are of the same dimension and equicodimensional means that all *maximal* ideals in A are of the same codimension in A. Note that the latter means in particular that

$$\dim(A) = \dim(A_M) \quad \text{for all } M \in \mathrm{Max}(A)\ , \tag{8.69}$$

i.e. we can read off the dimension of A from the dimension of the localization in an arbitrary maximal ideal—this is quite nice!

Your head must be spinning from this load of terminology. The point of all this is that a bi-equidimensional space has everything we can ask for:

Lemma 8.4.9. *Let X be a bi-equidimensional topological space. Then:*

1. *X is equidimensional, equicodimensional, catenary, and satisfies the dimension formula.*
2. *A closed subspace Z of X which is equicodimensional in X (e.g. if Z is irreducible) is bi-equidimensional as well.*

Proof. In this proof, "chain" will always mean "chain of irreducible closed subsets".

If X_λ is an irreducible component of X, then a maximal chain in X_λ is also a maximal chain in X, and since all these are of the same length $\dim(X)$, it follows that $\dim(X_\lambda) = \dim(X)$, i.e. X is equidimensional.

Similarly, if X_0 is a minimal irreducible closed subset in X, then a maximal chain between X_0 and X is a maximal chain in X, hence $\mathrm{codim}(X_0, X) = \dim(X)$, i.e. X is equicodimensional.

Next, let $Y \subseteq Z$ be irreducible closed subsets of X. Since $\dim(X) < \infty$, we can find maximal chains $Y_0 \subsetneq \ldots \subsetneq Y_m = Y$ and $Z = Z_0 \subsetneq \ldots \subsetneq Z_n \subseteq X$. Then a maximal chain between Y and Z of length l can be complemented by these two chains to a maximal chain in X of length $m + n + l$. Since all maximal chains in X have the same length $\dim(X)$, it follows that $l = \dim(X) - m - n$ is independent of the chain, i.e. X is catenary.

To prove the dimension formula, it is by Lemma 8.4.6 enough to show that any irreducible closed subset Z of X satisfies the dimension formula in X. We apply what we just discussed in case $Y = Z$ by taking a maximal chain in Z of length $m = \dim(Z)$ and a maximal chain between Z and X of length $l = \mathrm{codim}(Z, X)$. Then $\mathrm{codim}(Z, X) = \dim(X) - \dim(Z)$.

Finally, let Z be equicodimensional in X. Consider a maximal chain $Z_0 \subsetneq Z_1 \subsetneq \ldots \subsetneq Z_n$ in Z. Then Z_n is an irreducible component in Z. By assumption, we can find a chain $Z = Z_n \subsetneq Z_{n+1} \subsetneq \ldots \subsetneq Z_{n+m}$ in X with $m = \mathrm{codim}(Z, X)$. The chain $Z_0 \subsetneq \ldots \subsetneq Z_{n+m}$ is a maximal chain in X, hence $n + \mathrm{codim}(Z, X) = \dim(X)$ and it follows that n is independent of the chain, i.e. Z is bi-equidimensional. □

Example 8.4.10. We have already noticed above that the ring B in Example 8.4.4 has maximal chains of prime ideals of distinct length, hence B is not bi-equidimensional. In fact B is not even equidimensional: it has an irreducible component of dimension 1 and an irreducible component of dimension 2.

Now we come to the big question: which rings are bi-equidimensional? Here's a very satisfying class of examples.

Theorem 8.4.11. *Let K be a field and let A be a finitely generated K-algebra which is also an integral domain. Then A is bi-equidimensional. In particular, for any proper ideal I in A the relation*

$$\dim(I) + \operatorname{codim}(I, A) = \dim(A) \tag{8.70}$$

holds, and for any maximal ideal M of A the relation

$$\dim(A) = \dim(A_M)\ . \tag{8.71}$$

holds.

Proof. Let's first prove the claim for a polynomial ring $A = K[\boldsymbol{X}]$, where $\boldsymbol{X} = \{X_1, \ldots, X_n\}$. We need to show that all maximal chains of prime ideals in $K[\boldsymbol{X}]$ have the same length, i.e. length n. Let $Q_0 \subsetneq \ldots \subsetneq Q_m$ be a chain of prime ideals in $K[\boldsymbol{X}]$. Then $m \leq n$. We will show that if $m < n$, then we can extend the chain. This will obviously prove the claim. By Noether normalization (Theorem 8.3.9) we can find an algebraically independent subset $\boldsymbol{Y} \coloneqq \{Y_1, \ldots, Y_n\}$ of $K[\boldsymbol{X}]$ over K such that

$$P_i \coloneqq Q_i \cap K[\boldsymbol{Y}] = (Y_{d_i+1}, \ldots, Y_n)\ , \quad d_i \coloneqq \dim(Q_i)\ . \tag{8.72}$$

Since $m < n$, it follows that there is an i such that a prime ideal P can be inserted between P_{i-1} and P_i. Let

$$B \coloneqq K[\boldsymbol{Y}]/P_{i-1}, \quad B' \coloneqq K[\boldsymbol{X}]/Q_{i-1}\ . \tag{8.73}$$

By Noether normalization, the extension $B \subseteq B'$ is finite and thus integral. Moreover, B is a polynomial ring, hence a unique factorization domain by Lemma 8.1.2 and therefore normal by Example 5.1.5. We can thus apply the going-down theorem (Theorem 5.4.6) to the extension $B \subseteq B'$ to deduce the existence of a prime ideal $Q \in K[\boldsymbol{X}]$ fitting into the following picture:

$$\begin{tikzcd}
\mathrm{Spec}(B') \arrow[d, mapsto] & Q_i/Q_{i-1} \arrow[d, mapsto] & \exists Q/Q_{i-1} \arrow[l, hook'] \arrow[d, mapsto] \\
\mathrm{Spec}(B) & P_i/P_{i-1} & P/P_{i-1} \arrow[l, hook']
\end{tikzcd} \tag{8.74}$$

The prime ideal Q satisfies $Q_{i-1} \subsetneq Q \subsetneq Q_i$, i.e. we were able to extend the chain.

It remains to take care of the general case where A is a finitely generated K-algebra which is also an integral domain. Up to isomorphism, A is then a quotient $K[\boldsymbol{X}]$ by a prime ideal P of $K[\boldsymbol{X}]$ and $\mathrm{Spec}(A) \simeq \mathrm{V}(P)$. It thus follows from Lemma 8.4.9 that $\mathrm{Spec}(A)$ is bi-equidimensional as well. □

Remark 8.4.12. This section was a bit more technical (in terms of terminology and intuition at least) and maybe more exotic than usual commutative algebra literature. But I felt it is important to show that: a) the concept of codimension can be quite tricky; b) the dimension formula fails in general; c) the dimension formula holds for the algebras you usually work with in (classical) algebraic geometry. I want to make a few more comments.

Definition 8.4.5 on the dimension formula is not official terminology but I think it makes sense for a conceptual discussion. The only general context I know which implies the dimension formula (globally) is when all maximal chains of irreducible closed subsets are of the same length—this is what we called bi-equidimensional.

Nagata [10, §34] called this property (in the context of rings) the **first chain condition** on prime ideals. The terminology "bi-equidimensional" is due to Grothendieck [6, §14.3]. But why did Grothendieck call it like this? In [6, §14.3] it was shown that a space which is equidimensional, equicodimensional, and catenary has the property that all maximal chains of irreducible closed subsets have the same length—and I guess this is why Grothendieck called such a space "bi-equidimensional". However, it was *recently* pointed out by Heinrich [8] that this implication is not correct! Since we want the stronger property that all maximal chains of irreducible closed subsets are of the same length we call this property "bi-equidimensional" instead. All this confusion is a bit unfortunate. Even worse: not many people actually use this terminology, and some even mean by bi-equidimensional only the combination of equidimensional and equicodimensional—and even these concepts may be defined differently (e.g. involving associated primes). You thus need to be careful!

Lemma 8.4.6 is basically [6, Corollaire 14.3.5]. The whole discussion in loc. cit. is done for a topological space which is noetherian and Kolmogorov. But for what we discussed here (basically only Lemma 8.4.6 and Lemma 8.4.9) I don't see that we need these assumptions and I'm fine with $\dim(X) < \infty$.

I do not like so much how the absolute notion of equicodimensional in Definition 8.4.7 plays along with the relative one I introduced in Definition 8.4.8 because it's not compatible. But it's what you find in the literature. I think if you emphasize that something is equicodimensional *in* something, things should be clear. Even better is saying that Z is "pure of codimension n in X"; then there should be no confusion.

Finally, it seems to be a (common?) misconception that catenary is enough to imply the dimension formula. But this is not true, see Exercise 8.4.13!

Exercises.

Exercise 8.4.13. In this exercise you will see that even for rather nice rings the dimension formula may fail.[2] Let p be a prime number and let A be the localization of $\mathbb{Z}$ in the prime ideal (p). We will work in the ring $A[X]$. Show the following:

(1) $M_1 \coloneqq (pX-1)$ and $M_2 \coloneqq (p, X)$ are maximal ideals in $A[X]$.
(2) $\operatorname{codim}(M_1, A[X]) = 1$ and $\operatorname{codim}(M_2, A[X]) = 2$.
(3) $\dim(M_1) + \operatorname{codim}(M_1, A[X]) < \dim(A[X])$.
(4) $A[X]$ is catenary.

8.5. Krull's principal ideal theorem

Recall that, if K is a field, the ideal (f) generated by an irreducible polynomial $f \in K[X_1, X_2]$ is of dimension 1, i.e. of dimension one less than the surrounding ring. This is in fact just an example of a very general phenomenon.

Theorem 8.5.1 (**Krull's principal ideal theorem**). *Let A be a noetherian ring and let $x \in A$. Then*

$$\operatorname{codim}(P, A) \leq 1 \tag{8.75}$$

[2] I learned about this from an answer by Georges Elencwajg to a post in `https://math.stackexchange.com/questions/49136/is-operatornameheight-mathfrakp-dim-a-mathfrakp-dim-a-true`.

for any prime P minimal above (x). In particular, if x is not a unit[3]*, then*

$$\operatorname{codim}((x), A) \leq 1 \ . \tag{8.76}$$

If x is neither a unit nor a zero-divisor, then (x) is pure of codimension 1 in A.

Note that the theorem only makes a statement about the *co*dimension. But if A satisfies the dimension formula (e.g. if A is bi-equidimensional), we can convert this into a statement about the dimension:

(1) $\dim(x) \geq \dim(A) - 1$, i.e. "the zero set of a single element reduces the dimension by at most 1";
(2) if x is neither a unit nor a zero-divisor, then (x) is pure of dimension $\dim(A) - 1$.

By Theorem 8.4.11 the second conclusion holds in particular for any non-constant polynomial f in a polynomial ring $K[X_1, \ldots, X_n]$ over a field K. Moreover, we will see in Proposition 8.5.12 that also for *local* rings we can deduce the above statement about dimensions regardless of the dimension formula (this will require some additional work of course).

We still need to prove the theorem though. We'll use a new construction for prime ideals that will also play an important role later on.

Definition 8.5.2. Let A be a ring. The **n-th symbolic power** of a prime ideal P in A is

$$P^{(n)} := \{a \in A \mid as \in P^n \text{ for some } s \in A \setminus P\} \ . \tag{8.77}$$

Lemma 8.5.3. *$P^{(n)}$ is the preimage of the power $(PA_P)^n \trianglelefteq A_P$ under the localization map $A \to A_P$, and therefore*

$$P^{(n)} A_P = (PA_P)^n \ . \tag{8.78}$$

Proof. Let $a \in P^{(n)}$. Then by definition, there is $s \in A \setminus P$ with $as \in P^n$. Hence,

$$(PA_P)^n \ni as \cdot \frac{1}{s} = \frac{a}{1} \ ,$$

i.e. a maps into $(PA_P)^n$ under the localization map. Conversely, let $a \in A$ such that $\frac{a}{1} \in (PA_P)^n$. Then $\frac{a}{1} = \frac{x}{s}$ for some $x \in P^n$ and $s \in A \setminus P$. Hence, there is $t \in A \setminus P$ with

$$a \cdot \underbrace{ts}_{\in A \setminus P} = \underbrace{xt}_{\in P^n}$$

and therefore $a \in P^{(n)}$. The second claim follows from the ideal correspondence Proposition 4.2.10. □

We need one more little lemma on minimal primes.

Lemma 8.5.4. *Let A be a noetherian local ring with maximal ideal P. For an ideal I in A the following are equivalent:*

(1) *P is minimal over I.*
(2) *$P^n \subseteq I$ for some n, i.e. P is nilpotent modulo I.*

In particular, for such an ideal I the quotient A/I is artinian.

[3] If x is a unit, there is no minimal prime above (x), hence $\operatorname{codim}((x), A) = +\infty$ by convention. No one cares about this case though.

Proof. Let P be minimal over I. Since P is maximal, it follows that A/I is 0-dimensional, hence A/I is artinian by Theorem 7.6.4 (here, we use that A is noetherian). Consequently, the descending chain

$$P/I \supseteq P^2/I \supseteq \dots \tag{8.79}$$

of ideals in A/I becomes stationary, i.e. there is $n \in \mathbb{N}$ such that $P^n/I = P^{n+1}/I$, hence

$$P/I \cdot P^n/I = P^n/I \ . \tag{8.80}$$

Since A/I is noetherian, the ideal P^n/I in A/I is finitely generated and we can apply Nakayama's lemma (Corollary 3.8.7) to deduce that the equation above implies $P^n/I = 0$, i.e. $P^n \subseteq I$.

Conversely, assume that $P^n \subseteq I$ for some n. Let $Q \in \operatorname{Spec}(A)$ with $I \subseteq Q \subseteq P$. Since $P^n \subseteq I$, also $P^n \subseteq Q$ and therefore $P \subseteq Q$ because P is a prime ideal. But P is maximal, so $P = Q$. This shows that P is minimal over I. □

Now, we're prepared.

Proof of Theorem 8.5.1. Let P be a minimal prime over (x). We need to show that $\operatorname{codim}(P, A) \leq 1$, i.e. $\dim(A_P) \leq 1$. This is equivalent to $\dim(A_Q) = 0$ for any $Q \in \operatorname{Spec}(A)$ with $Q \subsetneq P$. By replacing A by A_P (which is still noetherian), we can thus assume without loss of generality that A is local with maximal ideal P. Then $A/(x)$ is artinian by Lemma 8.5.4. Consequently, the descending chain

$$Q^{(1)} + (x) \supseteq Q^{(2)} + (x) \supseteq \dots \tag{8.81}$$

of ideals becomes stationary, i.e. there is $n \in \mathbb{N}$ such that

$$Q^{(n)} + (x) = Q^{(n+1)} + (x) \ . \tag{8.82}$$

In particular, $Q^{(n)} \subseteq Q^{(n+1)} + (x)$. For any $f \in Q^{(n)}$ we can thus find $a \in A$ and $g \in Q^{(n+1)} \subseteq Q^{(n)}$ with $f = ax + g$, i.e.

$$ax = f - g \in Q^{(n)} \ . \tag{8.83}$$

Now, by definition of $Q^{(n)}$, there is $s \in A \setminus Q$ with $sax \in Q^n$. Since P is minimal over (x) and $Q \subsetneq P$, we must have $x \in A \setminus Q$, so $(sx)a \in Q^n$ means that $a \in Q^{(n)}$. This shows that

$$Q^{(n)} = (x)Q^{(n)} + Q^{(n+1)} \ . \tag{8.84}$$

The ideals in the above equation are all finitely generated because A is noetherian. Moreover, $x \in P = \operatorname{Jac}(A)$ because A is local. We can thus apply Nakayama's lemma (Corollary 3.8.7) to deduce that

$$Q^{(n)} = Q^{(n+1)} \ . \tag{8.85}$$

Recall that $Q^{(m)}A_Q = (QA_Q)^m$ for any m by Lemma 8.5.3. Hence, from the equation above we get

$$(QA_Q)^n = (QA_Q)^{n+1} = QA_Q \cdot (QA_Q)^n = \operatorname{Jac}(A_Q) \cdot (QA_Q)^n \ . \tag{8.86}$$

Nakayama's lemma thus implies

$$(QA_Q)^n = 0 \ . \tag{8.87}$$

This means the maximal ideal QA_Q in A_Q is nilpotent and now Lemma 8.5.4 shows that QA_Q is minimal over the zero ideal in A_Q and therefore $\dim(A_Q) = 0$. We have now proven that $\operatorname{codim}(P, A) \leq 1$ for any minimal prime P above (x). If x is not a unit, this clearly implies $\operatorname{codim}((x), P) \leq 1$.

Next, let x be neither a unit nor a zero-divisor. Let P be a prime minimal above (x). The claim is that $\operatorname{codim}(P, A) = 1$. Suppose that $\operatorname{codim}(P, A) = 0$. Since $\operatorname{codim}(P, A) = \dim(A_P)$, it follows that $\dim(A_P) = 0$, i.e. A_P is artinian by Theorem 7.6.4. Consequently, the chain

$$PA_P \supsetneq (PA_P)^2 \supsetneq \dots \tag{8.88}$$

of ideals in A_P becomes stationary, i.e. there is $n \in \mathbb{N}$ such that

$$(PA_P)^n = (PA_P)^{n+1} . \tag{8.89}$$

As above we deduce using Nakayama's lemma that PA_P is nilpotent. Since $x \in P$, it thus follows that $\frac{x}{1} \in A_P$ is nilpotent, i.e. there is $m \in \mathbb{N}$ with $\frac{x^m}{1} = 0 \in A_P$. This means there is $s \in A \setminus P$ with

$$0 = sx^m = (sx^{m-1})x \in A . \tag{8.90}$$

Since x is not a zero-divisor, this implies $sx^{m-1} = 0$ and then inductively $sx = 0$ which is a contradiction to x not being a zero-divisor. Hence, we must have $\operatorname{codim}(P, A) = 1$. □

Remark 8.5.5. In Exercise 8.5.13 you can see that it may happen that $\operatorname{codim}((x), A) = 1$ also for a zero-divisor $x \in A$.

By induction, we can extend Krull's principal ideal theorem to arbitrary ideals.

Theorem 8.5.6 (**Krull's ideal theorem**). *Let A be a noetherian ring and let $I = (x_1, \dots, x_c)$ be an ideal in A. Then*

$$\operatorname{codim}(P, A) \leq c \tag{8.91}$$

for any prime P minimal above I. In particular, if $I \neq A$, then

$$\operatorname{codim}(I, A) \leq c . \tag{8.92}$$

Proof. As in the proof of Theorem 8.5.1, we can assume without loss of generality that A is local with maximal ideal P. We prove the claim by induction over c. The case $c = 1$ is Theorem 8.5.1. Now, let $c > 1$. Let

$$P_0 \subsetneq P_1 \subsetneq \dots \subsetneq P_n = P \tag{8.93}$$

be a chain of prime ideals in A. We need to show that $n \leq c$. Since A is noetherian, we can assume without loss of generality that the chain above is maximal (if the chain is not maximal, we can always extend, and this eventually has to stop because of the noetherian property). Since $P \supsetneq P_{n-1}$ and P is minimal above $(x_1, \dots, x_c)$, not all x_i can be contained in P_{n-1}. Without loss of generality we can assume that $x_c \notin P_{n-1}$. Then P is minimal over (P_{n-1}, x_c). Hence, by Lemma 8.5.4 there is $r \in \mathbb{N}$ with

$$P^r \subseteq (P_{n-1}, x_c) . \tag{8.94}$$

For each i the element x_i^r is contained in P^r and so there are elements $a_i \in A$ and $y_i \in P_{n-1}$ with

$$x_i^r = a_i x_c + y_i . \tag{8.95}$$

By assumption, P is minimal over $(x_1, \dots, x_c)$, hence again by Lemma 8.5.4 there is $s \in \mathbb{N}$ with

$$P^s \subseteq (x_1, \dots, x_c) . \tag{8.96}$$

Let $N \coloneqq c \cdot r$. Then

$$P^{s+N} \subseteq (x_1^r, \dots, x_c^r) \subseteq (y_1, \dots, y_{c-1}, x_c) , \tag{8.97}$$

and therefore P is minimal over $(y_1,\dots,y_{c-1},x_c)$ by Lemma 8.5.4. Setting $J :=(y_1,\dots,y_{c-1})$, this means that P/J is minimal over the image of x_c in A/J. We can now apply Theorem 8.5.1 to deduce that

$$\operatorname{codim}(P/J, A/J) \leq 1\ . \tag{8.98}$$

Since $J \subseteq P_{n-1} \subsetneq P$, we also have $P_{n-1}/J \subsetneq P/J$ and therefore

$$\operatorname{codim}(P_{n-1}/J, A/J) = 0\ . \tag{8.99}$$

This means that P_{n-1} is minimal over $J = (y_1,\dots,y_{c-1})$. We can now apply the induction assumption to deduce that

$$\operatorname{codim}(P_{n-1}, A) \leq c-1\ . \tag{8.100}$$

This means $n-1 \leq c-1$, hence $n \leq c$. □

In a noetherian ring any ideal is finitely generated, so Theorem 8.5.6 implies:

Corollary 8.5.7. *For any proper ideal I in a noetherian ring A we have*

$$\operatorname{codim}(I, A) \leq \textit{number of generators of } I < \infty\ . \tag{8.101}$$

What about the dimension? Recall from Remark 8.2.13 that there are noetherian rings of infinite Krull dimension, so unfortunately we cannot deduce that the dimension is finite. But for *local* rings the situation is better:

Corollary 8.5.8. *If A is a noetherian* local *ring with maximal ideal M, then*

$$\dim(A) \leq \textit{number of generators of } M < \infty. \tag{8.102}$$

Proof. Since A has a *unique* maximal ideal M, we have

$$\dim(A) = \operatorname{codim}(M, A) \tag{8.103}$$

and now the claim follows immediately from Corollary 8.5.7. □

There's a kind of converse of Krull's ideal theorem:

Proposition 8.5.9. *Let A be a noetherian ring and let $P \in \operatorname{Spec}(A)$. If $\operatorname{codim}(P, A) = c$, then P is minimal over an ideal in A generated by c elements.*

For the proof we'll need a nice lemma that has many more applications.

Lemma 8.5.10 (**Prime avoidance**). *Let $I_1,\dots,I_n$ be ideals in a ring A and let J be another ideal not contained in any of the I_i. If at least $n-2$ of the I_i are prime, then*

$$J \not\subseteq \bigcup_{i=1}^{n} I_i\ . \tag{8.104}$$

Proof. Without loss of generality we can assume that the ideals I_i are numbered in such a way that I_i is prime for $i > 2$. We prove the claim by induction on n. The case $n = 1$ is clear, so let $n > 1$. Then by induction hypothesis we know that for each $1 \leq j \leq n$ we have

$$J \not\subseteq \bigcup_{\substack{i=1\\ i\neq j}}^{n} I_i\ . \tag{8.105}$$

Hence, for each j there is $z_j \in J \setminus \bigcup_{i\neq j} I_i$. If $z_j \notin I_j$ for some j, then we are done. Otherwise, we proceed as follows. Our assumption is that $z_j \in I_j$ for all j. Let

$$z := z_1 \cdots z_{n-1} + z_n \in J\ . \tag{8.106}$$

We claim that $z \notin I_i$ for all i. First, suppose that $z \in I_i$ for some $1 \le i < n$. Then

$$z_n = \underbrace{z}_{\in I_i} - \underbrace{z_1 \cdots z_{n-1}}_{\in I_i \text{ (since } z_i \in I_i)} \in I_i \ , \tag{8.107}$$

contradicting the choice of z_n. Next, suppose that $z \in I_n$. Then

$$z_1 \cdots z_{n-1} = z - z_n \in I_n \ . \tag{8.108}$$

If $n = 2$, then $z = z_1 + z_2 \in I_2$ and therefore $z_1 \in I_2$, which contradicts the choice of z_1. If $n > 2$, then $z_i \in I_n$ for some $1 \le i < n-1$ since I_n is prime, and this again contradicts the choice of z_i. □

Proof of Proposition 8.5.9. If $\operatorname{codim}(P, A) = 0$, then P is a minimal prime in A and therefore P is minimal over the zero ideal. The zero ideal is generated by the empty set, hence P is minimal over an ideal generated by $c = 0$ elements, proving the claim in this case.

We now assume that P is not minimal. We will inductively construct a sequence of elements $x_1, \ldots, x_r$ in P with $\operatorname{codim}((x_1, \ldots, x_r), A) = r$. For $r = c$ we then get

$$\operatorname{codim}((x_1, \ldots, x_c), A) = c = \operatorname{codim}(P, A) \ . \tag{8.109}$$

This implies in particular that P is minimal over $(x_1, \ldots, x_c)$.

Let $r = 1$. We have seen in Exercise 7.3.9 that a noetherian ring has only finitely many minimal primes. Let $P_1, \ldots, P_n$ be the minimal primes of A. Since P is not minimal, we have $P \not\subseteq P_i$ for all i. Hence, $P \not\subseteq \bigcup_{i=1}^n P_i$ by prime avoidance (Lemma 8.5.10). We can thus find an element $x_1 \in P$ with $x_1 \notin P_i$ for all i. Then certainly $\operatorname{codim}((x_1), A) \ge 1$. But $\operatorname{codim}((x_1), A) \le 1$ by Theorem 8.5.1, hence $\operatorname{codim}((x_1), A) = 1$.

Assume that $r > 1$ and that the elements $x_1, \ldots, x_{r-1}$ have been constructed. The construction of x_r is similar to the case $r = 1$ above. Let $\mathcal{M}$ be the set of prime ideals in A minimal over $(x_1, \ldots, x_{r-1})$. Since $A/(x_1, \ldots, x_{r-1})$ is noetherian, the set $\mathcal{M}$ is finite. Since $\operatorname{codim}((x_1, \ldots, x_{r-1}), A) = r - 1$ by construction, we have $\operatorname{codim}(Q, A) \le r-1$ for any $Q \in \mathcal{M}$. But also $\operatorname{codim}(Q, A) \ge r-1$ by Theorem 8.5.6, hence

$$\operatorname{codim}(Q, A) = r - 1 \quad \text{for all } Q \in \mathcal{M} \ . \tag{8.110}$$

Since $\operatorname{codim}(P, A) = c > r - 1$, we must have $P \not\subseteq Q$ for all $Q \in \mathcal{M}$. Hence, $P \not\subseteq \bigcup_{Q \in \mathcal{M}} Q$ by prime avoidance. We can thus find an element $x_r \in P$ with $x_r \notin Q$ for all $Q \in \mathcal{M}$. If Q' is a prime above $(x_1, \ldots, x_r)$, then Q' is of course also a prime above $(x_1, \ldots, x_{r-1})$. Hence, there is $Q \in \mathcal{M}$ with $Q \subseteq Q'$. Since $x_r \notin Q$, we actually have $Q \subsetneq Q'$. This implies

$$\operatorname{codim}(Q', A) > \operatorname{codim}(Q, A) = r - 1 \ , \tag{8.111}$$

hence $\operatorname{codim}((x_1, \ldots, x_r), A) \ge r$. On the other hand, $\operatorname{codim}((x_1, \ldots, x_r), A) \le r$ by Theorem 8.5.6, and we thus conclude that $\operatorname{codim}((x_1, \ldots, x_r), A) = r$. □

With the converse of Krull's ideal theorem we can now prove a special property of the maximal ideal in a noetherian local ring.

Lemma 8.5.11. *Let A be a noetherian local ring with maximal ideal M. For a subset $\{x_1, \ldots, x_n\}$ of M the following are equivalent:*

(1) $\dim((x_1, \ldots, x_n)) = 0$*;*
(2) *M is minimal over $(x_1, \ldots, x_n)$;*

(3) *$M^N \subseteq (x_1, \dots, x_n)$ for some $N > 0$;*
(4) *$\sqrt{(x_1, \dots, x_n)} = M$.*

If the above properties hold, then $\dim(A) \le n$. *Moreover, there exists such a subset with* $\dim(A) = n$. *Any such set is called a* ***system of parameters*** *for A.*

PROOF. The equivalences follow at once from Lemma 8.5.4. Note that since M is the *unique* maximal ideal, we have $\dim(A) = \operatorname{codim}(M, A)$. If M is minimal over $(x_1, \dots, x_n)$, then by Theorem 8.5.6 we have $\operatorname{codim}(M, A) \le n$, hence $\dim(A) \le n$. On the other hand, we know from Proposition 8.5.9 that M is minimal over an ideal generated by $\operatorname{codim}(M, A)$ many elements. □

With the existence of a system of parameters in a local ring we can now prove a version of Krull's principal ideal theorem in terms of dimension (without assuming the dimension formula):

PROPOSITION 8.5.12. *Let A be a noetherian* local *ring and let $x \in A$. If x is not a unit, then*

$$\dim(x) \ge \dim(A) - 1 \, . \tag{8.112}$$

If x is neither a unit nor a zero-divisor, then (x) is pure of dimension $\dim(A) - 1$.

PROOF. Let M be the maximal ideal in A. Note that $x \in M$ since x is not a unit by assumption. The quotient $\overline{A} := A/(x)$ is local as well with maximal ideal $\overline{M} := M/(x)$. Let $x_1, \dots, x_n \in A$ be chosen such that their images $\overline{x}_1, \dots, \overline{x}_n$ in $\overline{A}$ form a system of parameters for $\overline{A}$ (which exists by Lemma 8.5.11). Then

$$\dim(\overline{A}/(\overline{x}_1, \dots, \overline{x}_n)) = 0 \tag{8.113}$$

by Lemma 8.5.11. Since

$$\overline{A}/(\overline{x}_1, \dots, \overline{x}_n) = (A/(x)) \,/\, ((x, x_1, \dots, x_n)/(x)) \simeq A/(x, x_1, \dots, x_n) \, , \tag{8.114}$$

it follows that

$$\dim(A/(x, x_1, \dots, x_n)) = 0 \, , \tag{8.115}$$

hence $\dim(A) \le n + 1$ by Lemma 8.5.11. We conclude that

$$\dim(A) \le n + 1 = \dim(\overline{A}) + 1 = \dim(A/(x)) + 1 = \dim(x) + 1 \, . \tag{8.116}$$

Next, assume that x is neither a unit nor a zero-divisor. Then $\operatorname{codim}(P, A) = 1$ by Theorem 8.5.1 for any prime P minimal above (x). Since $\dim(P) + \operatorname{codim}(P, A) \le \dim(A)$, we have $\dim(x) \le \dim(P) \le \dim(A) - 1$. Since $\dim(x) \ge \dim(A) - 1$ by above, we thus have equality

$$\dim(P) = \dim(A) - 1 \, , \tag{8.117}$$

i.e. (x) is pure of dimension $\dim(A) - 1$. □

Exercises.

EXERCISE 8.5.13. Let K be a field and let $A := K[X, Y]/(X^2, XY)$. We denote by $\overline{X}$ and $\overline{Y}$ the image of X and Y, respectively, in A. Show the following:

(1) $\dim(A) = 1$.
(2) $(\overline{X})$ is the unique minimal prime in A.
(3) $\overline{Y}$ is a zero-divisor in A.
(4) $\operatorname{codim}(\overline{Y}, A) = 1$.

8.6. Regular sequences

In Theorem 8.5.1 we could conclude that if x is neither a unit nor a zero-divisor in a noetherian ring A, then (x) is *pure* of codimension 1 in A. The idea of regular sequences is to inductively apply this result to conclude that a (sufficiently nice) ideal in A is of *pure* codimension in A.

Definition 8.6.1. A sequence $x_1, \ldots, x_c$ of elements in a ring A is called **regular** if:

1. $(x_1, \ldots, x_c) \neq A$;
2. x_i is not a zero-divisor in $A/(x_1, \ldots, x_{i-1})$ for all $1 \leq i \leq c$.

Proposition 8.6.2. *If $x_1, \ldots, x_c$ is a regular sequence in a noetherian ring A, then the ideal $(x_1, \ldots, x_c)$ is pure of codimension c in A.*

Proof. We prove this by induction on c. The case $c = 1$ is covered by Theorem 8.5.1. So, let $c > 1$. Let $\overline{x}_c$ be the image of x_c in

$$\overline{A} \coloneqq A/(x_1, \ldots, x_{c-1}) \ . \tag{8.118}$$

By assumption, $\overline{x}_c$ is not a zero-divisor. Moreover, $\overline{x}_c$ is not a unit: otherwise there would be an element $y \in A$ with $x_c y \in 1 + (x_1, \ldots, x_{c-1})$, hence $1 \in (x_1, \ldots, x_c)$ which contradicts the assumption. Let P be a prime in A minimal over $(x_1, \ldots, x_c)$. Then

$$\overline{P} \coloneqq P/(x_1, \ldots, x_{c-1}) \tag{8.119}$$

is a prime in $\overline{A}$ minimal over

$$(x_1, \ldots, x_c)/(x_1, \ldots, x_{c-1}) = (\overline{x}_c) \tag{8.120}$$

and therefore Theorem 8.5.1 implies that

$$\operatorname{codim}(\overline{P}, \overline{A}) = 1 \ . \tag{8.121}$$

In particular, there is $\overline{Q} \in \operatorname{Spec}(\overline{A})$ with $\overline{Q} \subsetneq \overline{P}$ and $\overline{Q}$ minimal. We can write $\overline{Q} = Q/(x_1, \ldots, x_{c-1})$ for some $Q \in \operatorname{Spec}(A)$. Then $Q \subsetneq P$ is a maximal chain between Q and P, and Q is minimal over $(x_1, \ldots, x_{c-1})$. The induction assumption now implies that $\operatorname{codim}(Q, A) = c-1$, hence $\operatorname{codim}(P, A) \geq c$. On the other hand, we have $\operatorname{codim}(P, A) \leq c$ by Corollary 8.5.7 and thus conclude that $\operatorname{codim}(P, A) = c$. □

Definition 8.6.3. An ideal which is generated by a regular sequence is called a **complete intersection**.

What is the intuition behind this terminology? Recall from (2.57) that

$$\mathrm{V}(x_1, \ldots, x_c) = \bigcap_{i=1}^{c} \mathrm{V}(x_i) \ , \tag{8.122}$$

i.e. the zero set of $(x_1, \ldots, x_c)$ is the intersection of the zero sets of the various x_i. Now, if $(x_1, \ldots, x_c)$ is a complete intersection, then by Proposition 8.6.2 the codimension drops by precisely 1 when intersecting $\mathrm{V}(x_1, \ldots, x_{i-1})$ with the next $\mathrm{V}(x_i)$. Hence, a *complete* intersection is a particularly nice kind of intersection.

Example 8.6.4. In $K[X_1, \ldots, X_n]$ the sequence $X_1, \ldots, X_n$ is regular.

Example 8.6.5. In $K[X_1, X_2, X_3]$ the sequence $X_1, X_2(1 - X_1), X_3(1 - X_1)$ is regular. In particular,

$$\dim(\mathrm{V}(X_1, X_2(1 - X_1), X_3(1 - X_1))) = 3 - 3 = 0 \ ,$$

using Proposition 8.6.2 and the dimension formula. Obviously, the zero set is just the origin, so this makes sense.

Remark 8.6.6. A regular sequence really needs to be considered as a *sequence* and not just as a set, i.e. the order of the elements matters. For example, the sequence $X_2(1 - X_1), X_3(1 - X_1), X_1$ is a permutation of the sequence in Example 8.6.5 but it is not regular anymore. Nonetheless, one can show that in a *local* noetherian ring any permutation of a regular sequence is still regular.

8.7. Regular rings

Let A be a ring. You know that A has the geometric interpretation as zeros (over varying fields) of a system of polynomials. A point $P \in \mathrm{Spec}(A)$ corresponds to a (generalized) zero of this system. The localization A_P corresponds to a "neighborhood" around P: you throw away everything not contained in P. The maximal ideal $M_P := PA_P$ in A_P can be thought of as "residues" of polynomial functions locally around P that vanish in P. So far nothing new.

Now, let's look at the quotient M_P/M_P^2. This is a vector space over the residue field A_P/M_P of A in P. In the quotient we kill all higher powers of polynomial functions. Hence, what's left in the quotient can be thought of as *linear* functions locally around P that vanish in P. Such functions can be thought of as tangent directions to A in P as described briefly in the introduction to Chapter 8. For naturality reasons it's actually better to consider the *dual* vector space as tangent directions, and view M_P/M_P^2 as *co*tangent directions—but this should not bother us too much. The space M_P/M_P^2 is called the (Zariski) **cotangent space** of A in P.

This space—especially its dimension—provides some interesting local information in the point P. Let's assume from now on that A is noetherian. Then Corollary 8.5.8 implies that $\dim(A_P)$ is bounded from above by the (minimal) number of generators of the ideal M_P. What is this number? The answer is in Corollary 3.8.9, which was a corollary of Nakayama's lemma (Corollary 3.8.7): the minimal number of generators of M_P is equal to the dimension of the vector space M_P/M_P^2 over the field A_P/M_P. This is the dimension of the Zariski cotangent space! Let's write this down once more.

Corollary 8.7.1. *If A is noetherian, then*

$$\dim(A_P) \leq \textit{minimal number of generators of } M_P = \dim_{A_P/M_P}(M_P/M_P^2) \ . \tag{8.123}$$

Our discussion around Figure 8.1 motivates that a "singularity"—a "non-regular" point—is characterized by the property that there are more linearly independent tangent directions in the point than the dimension of the local ring in this point. This leads to the following definition.

Definition 8.7.2. A **regular local ring** is a local noetherian ring A such that

$$\dim(A) = \dim_{A/M}(M/M^2) \ , \tag{8.124}$$

where M is the maximal ideal of A. A (not necessarily local) noetherian ring A is called **regular** in $P \in \operatorname{Spec}(A)$ if A_P is a regular local ring; and A itself is called **regular** if A is regular in any $P \in \operatorname{Spec}(A)$.

Remark 8.7.3. One can show that the localization of a regular local ring in any prime ideal is again a regular local ring. This is actually non-trivial and the proof uses a homological characterization of regularity. Anyways, this implies that a noetherian local ring is a "regular local ring" if and only if it is a "regular ring" so that the terminology is compatible (this is quite subtle, think about it).

Intuitively, a regular local ring should be something very nice with nothing weird going on. In particular, a regular local ring should not contain zero-divisors. This is indeed true but it's not obvious:

Theorem 8.7.4. *A regular local ring is an integral domain.*

Proof. Let A be a regular local ring. We know that $\dim(A) < \infty$ and we will prove the claim by induction on $\dim(A)$.

First, suppose that $\dim(A) = 0$. Then $0 = \dim(A) = \dim_{A/M}(M/M^2)$, hence $M = M^2$. Nakayama's lemma (Corollary 3.8.7) thus implies that $M = 0$. We know from Exercise 2.3.11, that in a local ring the set of units is the complement of the maximal ideal. Hence, all non-zero elements of A are units, i.e. A is a field. In particular, A is an integral domain.

Now, let $n \coloneqq \dim(A) > 0$. Then $M \neq M^2$ as otherwise $0 = \dim_{A/M}(M/M^2) = \dim(A)$. Because A is noetherian, the set $\mathcal{M}$ of minimal prime ideals in A is finite by Exercise 7.3.9. Since $\dim(A) > 0$, the maximal ideal M is not minimal, hence $M \not\subseteq Q$ for all $Q \in \mathcal{M}$. Moreover, $M \not\subseteq M^2$. Hence, by prime avoidance (Lemma 8.5.10) we have

$$M \not\subseteq M^2 \cup \bigcup_{Q \in \mathcal{M}} Q \, . \tag{8.125}$$

We can thus find an element

$$x \in M \setminus \left(M^2 \cup \bigcup_{Q \in \mathcal{M}} Q \right) \, . \tag{8.126}$$

The quotient $\overline{A} \coloneqq A/(x)$ is local with maximal ideal $\overline{M} \coloneqq M/(x)$. Since $x \notin Q$ for all $Q \in \mathcal{M}$, we must have $\dim(\overline{A}) < \dim(A)$. Hence,

$$\dim(\overline{A}) = \dim(A) - 1 = n - 1 \tag{8.127}$$

by Proposition 8.5.12. Let $K \coloneqq A/M$ and $V \coloneqq M/M^2$. Since A is regular, we have $n = \dim(A) = \dim_K(V)$. Let $\overline{x}$ be the image of x in V. Since $x \notin M^2$, the image $\overline{x}$ is non-zero. Hence,

$$\dim_K(V/(\overline{x})) = \dim_K(V) - 1 = \dim(\overline{A}) \, . \tag{8.128}$$

Note that

$$\overline{A}/\overline{M} = (A/(x)) \, / \, (M/(x)) \simeq A/M = K \tag{8.129}$$

and

$$\overline{M}/\overline{M}^2 = (M/(x)) \, / \, \big((M^2 + (x))/(x)\big) \simeq V/(x) \, . \tag{8.130}$$

Hence,

$$\dim_K(\overline{M}/\overline{M}^2) = \dim_K(V/(x)) = \dim(\overline{A}) \, , \tag{8.131}$$

which shows that $\overline{A}$ is regular. Since $\dim(\overline{A}) < n$, we can thus apply the induction assumption to conclude that $\overline{A}$ is an integral domain. This in turn implies that (x) is a prime ideal in A. The ideal (x) cannot be a minimal prime ideal because $x \notin Q$ for all $Q \in \mathcal{M}$ by choice of x. Hence, there is $Q \in \mathcal{M}$ with $Q \subsetneq (x)$. For any $y \in Q$ we can thus find $a \in A$ with $y = ax$. Since $x \notin Q$ and Q is prime, it follows that $a \in Q$ and therefore $y \in xQ$. This shows that $Q = xQ$. Now, $x \in M = \mathrm{Jac}(A)$, so Nakayama's lemma (Corollary 3.8.7) implies $Q = 0$. In particular, the zero ideal in A is prime, which means that A is an integral domain. □

That was quite a bit of work and an interesting route to proving that a ring is an integral domain, right? But how do you prove that a local ring is actually regular? This is a big topic and I don't really want to go into this here. I only want to mention one alternative characterization of regularity that I find quite nice and that sometimes helps you to prove regularity. Recall the concept of regular sequences from Section 8.6. A regular sequence in a local ring must necessarily live inside the maximal ideal because the elements in a regular sequence are by assumption non-units. The question is: when is the maximal ideal generated by a regular sequence? Can you guess?

Theorem 8.7.5. *Let A be a noetherian local ring with maximal ideal M. Then A is regular if and only if M is generated by a regular sequence. In this case, any preimage of a basis of M/M^2 is a regular sequence generating M.*

Proof. Let $K \coloneqq A/M$. Suppose that M is generated by a regular sequence $x_1, \ldots, x_c$. Then we know from Proposition 8.6.2 that $c = \mathrm{codim}(M, A)$. Since $\mathrm{codim}(M, A) = \dim(A)$, this implies $\dim(A) = c$. From Corollary 3.8.9 we know that $\dim_K(M/M^2)$ is equal to the minimal number of generators of M, hence

$$\dim_K(M/M^2) \leq c = \dim(A) \ . \tag{8.132}$$

But also $\dim_K(M/M^2) \geq \dim(A)$ by Corollary 8.5.8 and therefore we have equality, i.e. A is regular.

Conversely, assume that A is regular. Let $x_1, \ldots, x_n \in A$ be such that their images in M/M^2 form a K-basis. It follows from Corollary 3.8.9 that $x_1, \ldots, x_n$ is a minimal generating set of M. Moreover, since A is regular, we have $n = \dim(A)$. Let

$$A_i \coloneqq A/(x_1, \ldots, x_{i-1}) \ . \tag{8.133}$$

This is a local ring with maximal ideal

$$M_i \coloneqq M/(x_1, \ldots, x_{i-1}) = (x_1, \ldots, x_n)/(x_1, \ldots, x_{i-1}) \ . \tag{8.134}$$

The ideal M_i is generated by $x_i, \ldots, x_n$ and the images of these elements in M_i/M_i^2 form a K-basis. Hence, $x_i, \ldots, x_n$ is a minimal generating set of M_i by Corollary 3.8.9. In particular,

$$\dim(A_i) \leq n - i + 1 = \dim_K(M_i/M_i^2) \ . \tag{8.135}$$

On the other hand,

$$\dim(A_i) \geq \dim(A) - i + 1 = n - i + 1 \tag{8.136}$$

by (an inductive application of) Proposition 8.5.12. We conclude:

$$\dim(A_i) = \dim_K(M_i/M_i^2) \ , \tag{8.137}$$

i.e. A_i is regular. Hence, A_i is an integral domain by Theorem 8.7.4. This shows that $x_1, \ldots, x_n$ is a regular sequence. □

Exercises.

EXERCISE 8.7.6. Let K be a field. Show that $K[X_1, \ldots, X_n]$ is regular in the origin.

EXERCISE 8.7.7. Let K be a field. Show that $K[X_1, X_2, X_3]/(X_1X_2 - X_3^2)$ is not regular in the origin. To this end, prove the following general statement: if A is a regular local ring with maximal ideal M and $0 \neq x \in M^2$, then $A/(x)$ is not regular.

EXERCISE 8.7.8. Let K be a field of characteristic $\neq 2, 3$ and let

$$A \coloneqq K[X_1, X_2]/(X_1^3 - X_2^2)\ .$$

(1) Show that A is regular in $(X_1 - 1, X_2 - 1)$. (Hint: prove that MA_M is principal, where $M \coloneqq (X_1 - 1, X_2 - 1)$.)

(2) Show that A is not regular in the origin (X_1, X_2).

CHAPTER 9

Dedekind domains

When we started studying prime ideals in Chapter 2 we observed that they are the correct generalization of prime elements to general rings. Since then we have seen that prime ideals have a geometric interpretation and are fundamental everywhere in commutative algebra. The question whether any ideal can be (uniquely) factorized into a product of prime ideals—which was part of the motivation in the introduction of Chapter 2—did not play any role. We now come back to this question. Since unique factorization into prime elements only works in a special class of rings—the unique factorization domains—it should not be surprising that on the ideal level this also only works in a special class of rings: this will be the Dedekind domains. We now have all the tools ready to study and characterize this beautiful and important class of rings.

9.1. Characterizations of Dedekind domains

Definition 9.1.1. A **Dedekind domain** is an integral domain A such that any non-zero ideal I in A admits a factorization

$$I = P_1 \cdots P_n \tag{9.1}$$

into a product of prime ideals P_i which is unique up to the order of the factors.[1]

Example 9.1.2. Any principal ideal domain is a Dedekind domain: we can simply pass from elements to the principal ideals they generate and since a principal ideal domain is a unique factorization domain (Lemma 1.5.24), it follows that we have unique factorization of ideals into prime ideals. In particular, $\mathbb{Z}$ and $K[X]$ are Dedekind domains.

Away from principal ideal domains it is rather difficult to prove that a given ring is a Dedekind domain (or not) by direct means, i.e. using only Definition 9.1.1. We need to establish basic properties and alternative characterizations of Dedekind domains to make progress. It'll take a bit of time to put everything together but in the end we'll have uncovered the beauty of Dedekind domains (Theorem 9.1.17) along with a large class of examples (Theorem 9.1.19). There are many different routes to get there but I like the following approach that I learned from [12].

First, there is an interesting local property:

Lemma 9.1.3. *If A is a Dedekind domain, then for any $P \in \operatorname{Spec}(A)$ the maximal ideal in A_P is principal.*

[1] The ideal $I = A$ can be written as the empty product of prime ideals. Hence, only proper non-zero ideals are really relevant in the definition. We need to exclude the zero ideal because $(0) = (0)^2 = \ldots$, so we can't have unique factorizations for it.

Proof. The claim is obvious for $P = 0$, so let $P \neq 0$. Then $P \neq P^2$ by uniqueness of factorizations. We can thus find $x \in P$ with $x \notin P^2$. We claim that

$$PA_P = (x)A_P \ . \tag{9.2}$$

The inclusion $PA_P \supseteq (x)A_P$ is obvious. To prove the converse, let $y \in P$. Let $(x, y) = P_1 \cdots P_n$ be the prime ideal factorization of (x, y). Since $P_1 \cdots P_n = (x, y) \subseteq P$ and P is prime, there is P_i with $P_i \subseteq P$. Suppose there is another $j \neq i$ with $P_j \subseteq P$. Then $x \in P_1 \cdots P_n \subseteq P_iP_j$. Since $P_iP_j \subseteq P^2$, it follows that $x \in P^2$, which contradicts the choice of x. Hence, $P_j \not\subseteq P$ and therefore $P_jA_P = A_P$ for $j \neq i$ and therefore localizing the factorization $(x, y) = P_1 \cdots P_n$ in P yields

$$(x, y)A_P = P_iA_P \ . \tag{9.3}$$

In particular, $(x, y)A_P$ is a prime ideal in A_P. Since y was arbitrary, this conclusion also applies to y^2, i.e. $(x, y^2)A_P$ is a prime ideal in A_P. Then $y^2 \in (x, y^2)A_P$ implies $y \in (x, y^2)A_P$ and we can write $y = ax + by^2$ for some $a, b \in A_P$, i.e. $(1 - by)y = ax \in (x)A_P$. Since y is contained in a prime ideal of A_P and A_P is local, it follows that y is contained in the maximal ideal of A_P and therefore $1 - by$ is a unit in A_P by Lemma 2.6.8. Hence, $y \in (x)A_P$. This proves that $PA_P = (x)A_P$, i.e. the maximal ideal of A_P is principal. □

When taking $y = 0$ in (9.3) we deduce that

$$PA_P = (x)A_P = P_iA_P \ , \tag{9.4}$$

where $(x) = P_1 \cdots P_n$ is the prime ideal factorization and P_i is the unique prime with $P_i \subseteq P$. The description of ideals in a localization (Corollary 4.2.11) implies that $P = P_i$. We thus conclude: any non-zero prime ideal in a Dedekind domain occurs as a factor in the factorization of a principal ideal. The following general lemma now tells us that any prime ideal in a Dedekind domain is finitely generated.

Lemma 9.1.4. *Let A be a ring and let I, J be ideals in A such that $IJ = (x)$ for a non-zero-divisor $x \in A$. Then I and J are finitely generated.*

For the proof of Lemma 9.1.4 we'll need another general lemma which says that we can prove finite generation of a module by checking this on an open cover of the prime spectrum.

Lemma 9.1.5. *Let A be a ring and let $f_1, \ldots, f_n \in A$ such that $(f_1, \ldots, f_n) = A$. If V is an A-module such that V_{f_i} is a finitely generated A_{f_i}-module for each i, then V is finitely generated as an A-module.*

Proof. For each i let $\boldsymbol{v}_i'$ be a finite generating set of V_{f_i} as an A_{f_i}-module. Without loss of generality we can assume that $\boldsymbol{v}_i'$ is in the image of the localization map $V \to V_{f_i}$. We can then take a preimage $\boldsymbol{v}_i \subseteq V$ of $\boldsymbol{v}_i'$ under this map. The union $\boldsymbol{v}$ of the $\boldsymbol{v}_i$ is a finite set. Consider the map $\varphi\colon A^{\boldsymbol{v}} \to V$ mapping the standard basis element e_v to v for $v \in \boldsymbol{v}$. By construction, the localized map $\varphi_{f_i}\colon A_{f_i}^{\boldsymbol{v}} \to V_{f_i}$ is surjective. We want to show that the localized map $\varphi_P\colon A_P^{\boldsymbol{v}} \to V_P$ is surjective for any $P \in \mathrm{Spec}(A)$ because this implies that φ itself is surjective by Corollary 4.4.7 and therefore V is finitely generated as an A-module. By assumption, we have $(f_1, \ldots, f_n) = A$. Hence, taking zero sets we get

$$\emptyset = \mathrm{V}(A) = \mathrm{V}(f_1, \ldots, f_n) = \bigcap_{i=1}^{n} \mathrm{V}(f_i) \ , \tag{9.5}$$

and taking complements we deduce that

$$\operatorname{Spec}(A) = \bigcup_{i=1}^{n} \operatorname{Spec}(A) \setminus \mathrm{V}(f_i) = \bigcup_{i=1}^{n} \mathrm{D}(f_i) \ , \tag{9.6}$$

i.e. the $\mathrm{D}(f_i)$ form an open cover of $\operatorname{Spec}(A)$. In particular, any $P \in \operatorname{Spec}(A)$ is contained in some $\mathrm{D}(f_i)$. Now, recall from Exercise 4.2.12 that $\mathrm{D}(f_i) \simeq \operatorname{Spec}(A_{f_i})$. We can thus view P as a prime ideal of A_{f_i}. Since φ_{f_i} is surjective and localization is exact by Lemma 4.3.4, it follows that also φ_P is surjective. □

Proof of Lemma 9.1.4. By assumption, we can write $x = \sum_{i=1}^{n} x_i y_i$ with $x_i \in I$ and $y_i \in J$. On the other hand, $x_i y_i \in IJ = (x)$, so

$$x_i y_i = a_i x \tag{9.7}$$

for some $a_i \in A$. We get

$$\left(\sum_{i=1}^{n} a_i \right) x = \sum_{i=1}^{n} a_i x = \sum_{i=1}^{n} x_i y_i = x \tag{9.8}$$

and because x is a non-zero-divisor by assumption, this implies

$$\sum_{i=1}^{n} a_i = 1 \ . \tag{9.9}$$

Hence, to prove that I and J are finitely generated it is by Lemma 9.1.5 sufficient to prove that I_{a_i} and J_{a_i} are finitely generated ideals in A_{a_i} for all i. We claim that

$$I_{a_i} = A_{a_i} x_i \quad \text{and} \quad J_{a_i} = A_{a_i} y_i \ , \tag{9.10}$$

which proves in particular that I_{a_i} and J_{a_i} are finitely generated. The inclusions $A_{a_i} x_i \subseteq I_{a_i}$ and $A_{a_i} y_i \subseteq J_{a_i}$ are clear. To prove the converse, we first note that x_i and y_i are non-zero-divisors in A_{a_i}. Namely, an equation $x_i \frac{a}{a_i^m} = 0$ implies $x_i a_i^k a = 0$ for some $k \in \mathbb{N}$; multiplication with y_i and using (9.7) yields $0 = y_i x_i a_i^k a = x a_i^{k+1} a$, hence $a_i^{k+1} a = 0$ because x is a non-zero-divisor and now multiplication by a_i^{-m-k-1} yields $\frac{a}{a_i^m} = 0$. The argument for y_i is similar. An element of I_{a_i} is of the form $\frac{z}{a_i^m}$ for some $z \in I$ and $m \in \mathbb{N}$. Then $z y_i \in IJ = (x)$, so $z y_i = ax$ for some $a \in A$ and therefore $z y_i = \frac{a}{a_i} x_i y_i$ using (9.7). We get $y_i(z - \frac{a}{a_i} x_i) = 0$ and because y_i is a non-zero-divisor we deduce that $z = \frac{a}{a_i} x_i \in A_{a_i} x_i$. Similarly, one shows that $J_{a_i} \subseteq A_{a_i} y_i$. □

We have now proven that any prime ideal in a Dedekind domain is finitely generated. What's that good for? Well, this is enough to ensure that *all* ideals are finitely generated! This is again a general fact.

Lemma 9.1.6. *A ring is noetherian if and only if all prime ideals are finitely generated.*

The idea of the proof of Lemma 9.1.6 is to show that if there is a non-finitely generated ideal, then there is a maximal such ideal; ideals which are maximal with respect to some property often tend to be prime; this is true for non-finitely generated ideals and we arrive at a contradiction because we know there is no non-finitely generated prime ideal. To make this argument work, we first need to establish the meta lemma on maximality implying prime. I learned the following from [12] again. Recall that for an ideal I in a ring A and an element $a \in A$ we define

$$(I\colon a) \coloneqq \{x \in A \mid xa \in I\} \ . \tag{9.11}$$

Definition 9.1.7. A family $\mathcal{F}$ of ideals in a ring A is called an **Oka family** if $A \in \mathcal{F}$ and whenever I is an ideal in A such that $(I\colon a) \in \mathcal{F}$ and $(I, a) \in \mathcal{F}$ for some $a \in A$, also $I \in \mathcal{F}$.

Example 9.1.8. The family $\mathcal{F}$ of finitely generated ideals in a ring A is an Oka family. Clearly, $A \in \mathcal{F}$. Let I be an ideal in A such that $(I\colon a) \in \mathcal{F}$ and $(I, a) \in \mathcal{F}$ for some a. These conditions mean that we can choose a finite set of generators

$$(I\colon a) = (a_1, \ldots, a_n) \quad \text{and} \quad (I, a) = (a, b_1, \ldots, b_m) \tag{9.12}$$

with $b_i \in I$. We claim that

$$I = (aa_1, \ldots, aa_n, b_1, \ldots, b_m) \ , \tag{9.13}$$

which implies that $I \in \mathcal{F}$. Obviously, the ideal on the right hand side is contained in I. Conversely, let $x \in I$. Then $x \in (I, a)$, hence $x = ya + \sum_{i=1}^{m} y_i b_i$ for some $y, y_i \in A$. It follows that $ya = x - \sum_{i=1}^{m} y_i b_i \in I$, i.e. $y \in (I\colon a)$, and therefore $x \in (aa_1, \ldots, aa_n, b_1, \ldots, b_m)$.

Here is the promised meta lemma:

Lemma 9.1.9. *If $\mathcal{F}$ is an Oka family of ideals in a ring A, then any ideal I in A which is maximal among ideals* not *contained in $\mathcal{F}$ is a prime ideal.*

Proof. Suppose that I is not prime. Since $A \in \mathcal{F}$ and $I \notin \mathcal{F}$, we have $I \neq A$. Hence, since I is not prime, there are $a, b \in A \setminus I$ with $ab \in I$. Then $(I, a) \supsetneq I$ and $(I\colon a) \supsetneq I$. Consequently, $(I, a) \in \mathcal{F}$ and $(I\colon a) \in \mathcal{F}$ due to the maximality of I among ideals not contained in $\mathcal{F}$. But now the Oka family property implies that $I \in \mathcal{F}$, which is a contradiction. □

Proof of Lemma 9.1.6. If A is noetherian, then all ideals are finitely generated, so in particular the prime ideals are finitely generated. Conversely, assume that all prime ideals are finitely generated. Let Σ be the set of non-finitely generated ideals in A. Suppose that $\Sigma \neq \emptyset$. Let $(I_\lambda)_{\lambda \in \Lambda}$ be a chain in Σ. Then $I \coloneqq \bigcup_{\lambda \in \Lambda} I_\lambda$ is an ideal. This ideal is not finitely generated since if $I = (a_1, \ldots, a_n)$, then all the a_i would be contained in some I_λ and consequently the a_i would generate I_λ, which contradicts $I_\lambda \in \Sigma$. This shows that $I \in \Sigma$ and therefore any chain in Σ has an upper bound in Σ. We can thus apply Zorn's lemma to deduce that there is a maximal element M in Σ. By construction, the ideal M is maximal among ideals not contained in the Oka family of finitely generated ideals (Example 9.1.8) and now it follows from Lemma 9.1.9 that M is prime. But this means there is a prime ideal which is not finitely generated, which contradicts the assumption. □

We finally conclude:

Corollary 9.1.10. *A Dedekind domain is noetherian.*

And now we gain pace:

Corollary 9.1.11. *A Dedekind domain is regular and of dimension at most 1.*

Proof. Let P be a prime ideal of A. We need to show that A_P is regular and that $\dim(A_P) \leq 1$. If $P = 0$, this is obviously true, so assume $P \neq 0$. We know from Corollary 9.1.10 and Lemma 9.1.3 that A_P is a noetherian local ring whose maximal ideal is generated by a single element, say x. Now, it follows from Lemma 8.5.11 (system of parameters) that $\dim(A_P) \leq 1$. In fact, since P is non-zero and A_P is an integral domain, we must have $\dim(A_P) = 1$ and $\{x\}$ is a system of parameters for A_P. But of course $\{x\}$ is also a regular sequence (x is neither a unit nor a zero-divisor), hence it follows from Theorem 8.7.5 that A_P is a regular local ring. □

Remark 9.1.12. Note that a Dedekind domain is 1-dimensional if and only if it is not a field (the latter being the exceptional 0-dimensional case). In fact, many authors exclude the field case in the definition of a Dedekind domain but I don't like this because I want to say e.g. that any principal ideal domain is a Dedekind domain without excluding fields all the time. I am following Bourbaki [3] with this convention. But in the end no one cares about the field case in this context anyways.

Armed with all these properties that we have now established, we can unveil an exciting local information—a "measure" or "valuation"—of a Dedekind domain A in a non-zero prime P. Let $M_P \coloneqq PA_P$ be the maximal ideal of A_P. Since A_P is regular, we have $\dim_{A_P/M_P}(M_P/M_P^2) = 1$. This means we can find $\pi_P \in M_P$ with $\pi_P \notin M_P^2$. Such an element is called a **uniformizer** in P. It follows from Corollary 3.8.9 (Nakayama's lemma) that π_P generates M_P. Hence, any element $x \in A_P$ is of the form

$$x = u\pi_P^n \tag{9.14}$$

for some unit u and some $n \in \mathbb{N}$. This representation is in fact unique: suppose that $x = u_1\pi_P^{n_1} = u_2\pi_P^{n_2}$. Without loss of generality, we can assume that $n_1 \geq n_2$. Then $\pi_P^{n_1-n_2} = u_2u_1^{-1}$ is a unit, and this can only happen when $n_1 - n_2 = 0$, i.e. $n_1 = n_2$, and then $u_2u_1^{-1} = 1$, i.e. $u_2 = u_1$. We now define

$$v_P(x) \coloneqq \max\{n \in \mathbb{N} \mid x \in M_P^n\} \in \mathbb{N} \cup \{\infty\} , \tag{9.15}$$

i.e. $v_P(x)$ is the unique exponent n in the representation $x = u\pi_P^n$. Thinking of A_P as polynomial functions locally around P, we can think of $v_P(x)$ as the "vanishing order" of x in P. Note that $v_P(x) = \infty$ if and only if $x = 0$ and $v_P(x) = 0$ if and only if x is a unit. Let K be the fraction field of A and note that K is also the fraction field of A_P since $A \subseteq A_P \subseteq K$. We extend v_P to all of K by setting

$$v_P\left(\frac{x}{y}\right) \coloneqq v_P(x) - v_P(y) . \tag{9.16}$$

You can easily check that this is well-defined. The resulting map

$$v_P \colon K \to \mathbb{Z} \cup \{\infty\} \tag{9.17}$$

is called the **valuation** of A in P (also called **P-adic valuation**). The restriction

$$v_P \colon K^\times \to \mathbb{Z} \tag{9.18}$$

of v_P to non-zero elements is surjective and multiplicative, i.e. it is a group morphism. Moreover, we obviously have the following property:

$$v_P(x+y) \geq \min\{v_P(x), v_P(y)\} . \tag{9.19}$$

Example 9.1.13. Consider the Dedekind domain $\mathbb{Z}$ and let $P = (p)$ be a non-zero prime ideal. Then

$$\mathbb{Z}_{(p)} = \left\{ \frac{a}{b} \mid a, b \in \mathbb{Z}, p \nmid b \right\} . \tag{9.20}$$

A uniformizer in $\mathbb{Z}_{(p)}$ is given by p, and for $a \in \mathbb{Z}$ the valuation $v_p(a)$ is equal to the exponent of p in the prime factorization of a. This description of the valuations is more generally true for any principal ideal domain (being unique factorization domains).

The following definition captures our observations above:

Definition 9.1.14. A **discrete valuation** on a field K is a surjective group morphism

$$v \colon K^\times \to \mathbb{Z} \tag{9.21}$$

such that

$$v(x+y) \geq \min\{v(x), v(y)\} \tag{9.22}$$

for all $x, y \in K$.

The "discrete" in "discrete valuation" refers to the value group $\mathbb{Z}$—one also studies more general valuations taking values in a totally ordered abelian group but we won't need that here. Given a discrete valuation $v \colon K^\times \to \mathbb{Z}$ we formally define $v(0) \coloneqq \infty$ to get a map $v \colon K \to \mathbb{Z} \cup \{\infty\}$. It is straightforward to check that

$$A \coloneqq \{x \in K \mid v(x) \geq 0\} \tag{9.23}$$

is a subring of K. This ring is called the **valuation ring** of v. More abstractly, a **discrete valuation ring** is an integral domain A which is the valuation ring of a discrete valuation on its field of fractions. The following properties are all easy to verify.

Lemma 9.1.15. *Let $v \colon K^\times \to \mathbb{Z}$ be a discrete valuation on a field K and let A be the corresponding valuation ring. Show the following:*

(1) *$A^\times = \{x \in K \mid v(x) = 0\}$.*
(2) *A is local with maximal ideal $M \coloneqq \{x \in A \mid v(x) \geq 1\}$.*
(3) *An element $\pi \in A$ with $v(\pi) = 1$ is called a **uniformizer**. Fixing such an element, every $x \in A$ can be uniquely written as $x = u\pi^n$ with $u \in A^\times$ and $n \in \mathbb{N}$.*
(4) *For $x \in A$ the valuation $v(x)$ is the largest number n such that π^n divides x.*
(5) *$M^n = (\pi^n) = \{x \in A \mid v(x) \geq n\}$ and every non-zero ideal in A is of this form.*
(6) *The fraction field of A is equal to K.*
(7) *If $x \in K$, then $x \in A$ or $x^{-1} \in A$.*

Proof. Left for you as Exercise 9.1.21. □

The localizations of a Dedekind domain in the non-zero primes are of course discrete valuation rings—this is what motivated the definition of a discrete valuation ring after all. In this particular case we know from Corollary 9.1.11 that they are noetherian, 1-dimensional, and regular. This is in fact an alternative characterization of discrete valuation rings in general—and there are some more characterizations!

Theorem 9.1.16. *For an integral domain A the following are equivalent:*

(1) *A is a discrete valuation ring.*

(2) *A is a local principal ideal domain which is not a field.*
(3) *A is local, noetherian, 1-dimensional, and normal.*
(4) *A is local, noetherian, and the maximal ideal is principal.*
(5) *A is local, noetherian, 1-dimensional, and regular.*

Proof.

(1) ⇒ (2): We know from Lemma 9.1.15 that in a discrete valuation ring A all non-zero ideals are of the form (π^n) for some $n \in \mathbb{N}$. In particular, A is a principal ideal domain. Moreover, since the valuation map defining A is surjective by Definition 9.1.14 and the valuation of units is equal to 0, it follows that A has a non-zero element which is not a unit, i.e. A is not a field.

(2) ⇒ (3): This follows immediately from the properties of principal ideal domains we have already established, namely noetherian by Example 7.2.3, 1-dimensional by Example 8.2.3, normal by Example 5.1.5.

(3) ⇒ (4): Let M be the maximal ideal of A and let K be the fraction field of A. Since A is local, noetherian, and 1-dimensional, we know from Lemma 8.5.11 (system of parameters) that there is $a \in M$ and $N \in \mathbb{N}$ with $M^N \subseteq (a)$. We choose N minimal with $M^N \subseteq (a)$. If $N = 1$, then $M = (a)$ and we are done. Suppose that $N > 1$. Then we can choose $b \in M^{N-1}$ with $b \notin (a)$, i.e. $x := \frac{b}{a} \in K \setminus A$. Note that $x^{-1} = \frac{a}{b} \in A$ and $xM \subseteq A$ since $bM \subseteq M^N \subseteq (a)$. Suppose that $xM \subseteq M$. Then M is stable under multiplication by x, so M is a (finitely generated) $A[x]$-module and as such it is faithful because K is an integral domain. But this is a contradiction because $x \in K \setminus A$ and A is normal, so x is not integral and therefore there cannot exist a finitely generated faithful $A[x]$-module by Theorem 5.1.7. We conclude that $M = (a)$, so M is principal.

(4) ⇒ (5): This is exactly the argument we used in the proof of Corollary 9.1.11.

(5) ⇒ (1): We can define a discrete valuation exactly as we did in (9.15). □

We now have a very good *local* understanding of Dedekind domains. We can use this to derive a *global* characterization.

Theorem 9.1.17. *For an integral domain A which is not a field the following are equivalent:*

(1) *A is a Dedekind domain.*
(2) *A is noetherian and A_P is a discrete valuation ring for all $0 \neq P \in \operatorname{Spec}(A)$.*
(3) *A is noetherian, 1-dimensional, and normal.*
(4) *A is noetherian, 1-dimensional, and regular.*

Proof. If A is a Dedekind domain, we know from Corollary 9.1.10 that A is noetherian and from the discussion leading to the notion of discrete valuation rings we know that A_P is a discrete valuation ring for any non-zero prime ideal P in A. This proves that (1) implies (2). The equivalences of (2), (3), and (4) follow immediately from the fact that normality is a local property (Theorem 9.1.16) and regularity is a local property by definition. It remains to prove that (2) implies (1). Let I be a non-zero proper ideal in A. We need to prove that I has a unique factorization into prime ideals. Let's first prove existence. Since $(0) \in \operatorname{Spec}(A)$, $(0) \neq I$, and $\dim(A) = 1$, it follows that $\dim(A/I) = 0$. By assumption A is noetherian, hence also A/I is noetherian, and now Theorem 7.6.4 implies that A/I is artinian and has only finitely many prime ideals, i.e. there are only finitely many prime ideals $P_1, \dots, P_n$ in A above I. Since $(0) \neq I$, all the P_i must be maximal ideals and therefore each localization A_{P_i} is a discrete valuation ring by assumption. In Lemma 9.1.15 we have seen that in a discrete valuation ring any non-zero ideal is a unique power of the maximal ideal, hence

$$IA_{P_i} = (P_i A_{P_i})^{\nu_i} = P_i^{\nu_i} A_{P_i} \tag{9.24}$$

for a unique $\nu_i \in \mathbb{N}$. Since $I \subseteq P_i$, the ideal IA_{P_i} is a proper ideal in A_{P_i} and therefore $\nu_i \geq 1$. Because A/I is artinian, we have by Exercise 7.6.5 a product decomposition

$$A/I \simeq \prod_{i=1}^{n} (A/I)_{P_i/I} \simeq \prod_{i=1}^{n} A_{P_i}/I_{P_i} = \prod_{i=1}^{n} A_{P_i}/IA_{P_i} = \prod_{i=1}^{n} A_{P_i}/P_i^{\nu_i} A_{P_i} \tag{9.25}$$

$$\simeq \prod_{i=1}^{n} A/P_i^{\nu_i} \overset{(*)}{\simeq} A/\bigcap_{i=1}^{n} P_i^{\nu_i} \ , \tag{9.26}$$

where the isomorphism $(*)$ is the Chinese remainder theorem (Exercise 1.3.25). Hence,

$$I = \bigcap_{i=1}^{n} P_i^{\nu_i} = \prod_{i=1}^{n} P_i^{\nu_i} \ , \tag{9.27}$$

where the second equality is again the Chinese remainder theorem. This proves existence of a factorization.

To prove uniqueness, let

$$I = Q_1^{\mu_1} \cdots Q_m^{\mu_m} \tag{9.28}$$

be a factorization into pairwise distinct Q_i, each occurring with exponent $\mu_i \geq 1$. Then $I \subseteq Q_i$ for all i, hence $Q_i = P_{\sigma(i)}$ for an injective map $\sigma \colon \{1, \ldots, m\} \to \{1, \ldots, n\}$. Since $\dim(A) = 1$, all the Q_i are maximal and as they are pairwise distinct, localizing at Q_i yields

$$IA_{Q_i} = Q_i^{\mu_i} A_{Q_i} = P_{\sigma(i)}^{\mu_i} A_{P_{\sigma(i)}} \ , \tag{9.29}$$

hence $\mu_i = \nu_{\sigma(i)}$ by uniqueness of the factorization (9.24) in a discrete valuation ring. We must have $m = n$ since otherwise there is a P_i which is not among the Q_j and then localizing the factorization (9.28) of I in P_i yields the whole ring while localizing the factorization (9.27) in P_i yields $P_i^{\nu_i}$ which is a proper ideal. □

REMARK 9.1.18. An important fact we have shown in the proof of Theorem 9.1.17 is that in a Dedekind domain A the prime ideals occurring in the factorization of a non-zero proper ideal I are precisely the prime ideals in A lying above I.

The conclusion is: “Dedekind domains are precisely regular irreducible curves”. Isn’t that beautiful? But things get even better: many of the rings of interest in algebraic number theory turn out to be Dedekind domains, i.e. they are regular irreducible curves as well!

THEOREM 9.1.19. *The integral closure $\mathcal{O}_L$ of $\mathbb{Z}$ in a number field L is a Dedekind domain.*

PROOF. We set $A \coloneqq \mathcal{O}_L$. By construction, A is normal. Moreover, $\mathbb{Z} \subseteq A$ is integral, so $\dim(A) = \dim(\mathbb{Z}) = 1$ by Lemma 8.2.4. All that remains to be done in light of Theorem 9.1.17 is to show that A is noetherian. We will proceed in 3 steps.

First, we prove that $|A/pA| < \infty$ for any prime number $p \in \mathbb{Z}$. Note that A/pA is a vector space over $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$. It is therefore enough to show that

$\dim_{\mathbb{F}_p}(A/pA) < \infty$. Let $\overline{a}_1, \ldots, \overline{a}_n \in A/pA$ be linearly independent over $\mathbb{F}_p$. We claim that preimages $a_1, \ldots, a_n \in A \subseteq L$ of the $\overline{a}_i$ are linearly independent over $\mathbb{Q}$. Suppose that $\alpha_1 a_1 + \ldots + \alpha_n a_n = 0$ for some $\alpha_i \in \mathbb{Q}$ which are not all zero. After multiplication by the least common multiple of the denominators of the α_i we can assume that $\alpha_i \in \mathbb{Z}$ for all i. Moreover, after possibly dividing by an appropriate power of p we can assume that not all of the α_i are divisible by p. Then $\overline{\alpha}_1 \overline{a}_1 + \ldots + \overline{\alpha}_n \overline{a}_n = 0$ is a non-trivial relation, which is a contradiction. Hence,

$$\dim_{\mathbb{F}_p}(A/pA) \leq \dim_{\mathbb{Q}}(L) < \infty \ . \tag{9.30}$$

Next, we show that more generally we have $|A/mA| < \infty$ for all $0 \neq m \in \mathbb{Z}$. The first step proves this for m a prime number. Let $m = m_1 m_2$ be a product. Then we have a short exact sequence

$$0 \longrightarrow A/m_1A \xrightarrow{\cdot m_2} A/(m_1m_2)A \longrightarrow A/m_2A \longrightarrow 0 \tag{9.31}$$

of abelian groups from which we get

$$|A/(m_1m_2)A| = |A/m_1A| \cdot |A/m_2A| \ . \tag{9.32}$$

The claim now follows by induction on the number of prime factors of an integer.

In the last step, we prove that any ideal I in A is finitely generated. We can assume that $I \neq 0$. We cannot have $I \cap \mathbb{Z} = 0$ because $\mathbb{Z} = \mathbb{Z}/(I \cap \mathbb{Z}) \subseteq A/I$ is integral, hence $\dim(A/I) = \dim(\mathbb{Z}) = 1$ but this is a contradiction to $\dim(A) = 1$ because we can always extend a maximal chain of prime ideals above I by the zero ideal. Hence, $I \cap \mathbb{Z} \neq 0$, which means there is $0 \neq m \in \mathbb{Z}$ with $m \in I$. By the second step we get $|I/mI| \leq |A/mA| < \infty$. Consequently, we have $I/mI = (A/mA) \cdot \{a_1, \ldots, a_n\}$ for certain $a_i \in I$ and therefore $I = A \cdot \{m, a_1, \ldots, a_m\}$ is finitely generated. □

In particular, for any prime number $p \in \mathbb{Z}$ we have a unique factorization

$$p\mathcal{O}_L = P_1^{\nu_1} \cdots P_n^{\nu_n} \tag{9.33}$$

into prime ideals P_i of $\mathcal{O}_L$, which are precisely the prime ideals in $\mathcal{O}_L$ lying above p. Understanding these decompositions is one of the central themes of number theory!

Remark 9.1.20. One can more generally prove that if A is a 1-dimensional noetherian integral domain with fraction field K and L is a finite extension field of K, then any ring B between A and L (e.g. the integral closure of A in L) is noetherian. This is the statement of the **Krull–Akizuki theorem**.

Exercises.

Exercise 9.1.21. Prove Lemma 9.1.15.

9.2. Fractional ideals and the ideal class group

Every principal ideal domain is a Dedekind domain—but of course there are Dedekind domains which are not principal ideal domains: for example the ring $\mathbb{Z}[\sqrt{-5}]$ is not a principal ideal domain (see the introduction of Chapter 2) but it is a Dedekind domain by Theorem 9.1.19 because it is the integral closure of $\mathbb{Z}$ in $\mathbb{Q}(\sqrt{-5})$ by Exercise 5.2.5.

The ideal class group measures "how far away a Dedekind domain is from being a principal ideal domain" and thus provides much more information than the simple conclusion that a given Dedekind domain is not a principal ideal domain. The basic idea is to turn the collection of ideals into a group under multiplication. But to this end, one needs a notion of an inverse of an ideal. This leads to the notion of fractional ideals. Throughout, we let A be an integral domain with fraction field K.

DEFINITION 9.2.1. A **fractional ideal** of A is an A-submodule I of K such that $aI \subseteq A$ for some non-zero $a \in A$.

Note that even though one uses the word "ideal", a fractional ideal is not an ideal in the usual sense because it is not necessarily contained in the ring but in the fraction field. In fact, the ideals (in the usual sense) are precisely the fractional ideals contained in A. In the context of fractional ideals one sometimes says **integral ideal** to refer to an ideal in the usual sense. Also note that in the notation of the definition the set $J \coloneqq aI$ is an ideal in A and

$$I = \frac{1}{a} J \; , \tag{9.34}$$

so fractional ideals are precisely those subsets of K of the form $\frac{1}{a}J$ for some $0 \neq a \in A$ and J an ideal in A. This explains the "fractional" in fractional ideal.

EXAMPLE 9.2.2. For any $x \in K$ the subset $xA \subseteq K$ is a fractional ideal. The fractional ideals of this form are called **principal**.

The set of non-zero fractional ideals clearly forms a commutative monoid under the obvious multiplication with identity element being A. This brings us to the following definition.

DEFINITION 9.2.3. A fractional ideal I is called **invertible** if there is a fractional ideal J such that $IJ = A$.

EXAMPLE 9.2.4. Any non-zero principal fractional ideal xA is invertible because $(xA) \cdot (x^{-1}A) = A$.

By definition, the invertible fractional ideals are precisely the invertible elements in the monoid of non-zero fractional ideals. It follows that that if a fractional ideal I is invertible, then the fractional ideal J with $IJ = A$ is uniquely determined: it is the inverse of I. There's a more explicit description of the inverse.

LEMMA 9.2.5. *If I and J are fractional ideals, then so is*

$$(I \colon J) \coloneqq \{x \in K \mid xJ \subseteq I\} \; . \tag{9.35}$$

In particular,

$$I^{-1} \coloneqq (A \colon I) = \{x \in K \mid xI \subseteq A\} \tag{9.36}$$

is fractional.

PROOF. First, suppose that I and J are ideals, i.e. $I, J \subseteq A$. Let $0 \neq a \in J$. If $x \in (I \colon J)$, then $xJ \subseteq I$ by definition, hence $xa \in I$. This shows that $a(I \colon J) \subseteq A$, i.e. $(I \colon J)$ is fractional. Next, let I and J be general fractional ideals. Let $0 \neq a, b \in A$ with $aI \subseteq A$ and $bJ \subseteq A$. Then

$$(abI \colon abJ) = \{x \in K \mid xabJ \subseteq abI\} = \{x \in K \mid xJ \subseteq I\} = (I \colon J) \; . \tag{9.37}$$

Since $abI, abJ \subseteq A$, it follows from above that $(abI \colon abJ) = (I \colon J)$ is fractional. □

LEMMA 9.2.6. *A fractional ideal I is invertible if and only if $II^{-1} = A$.*

PROOF. Suppose that $IJ = A$. Then $aI \subseteq A$ for any $a \in J$, so $J \subseteq (A \colon I) = I^{-1}$. It follows that $A = IJ \subseteq II^{-1} \subseteq A$ and therefore $II^{-1} = A$. □

Note that even though I^{-1} is a well-defined fractional ideal for *any* fractional ideal I, only for I *invertible* it is the actual inverse of I. The set of invertible fractional ideals of A forms an abelian group I_A under multiplication, called the **ideal group** of A. There is a special subgroup of the ideal group, namely the set P_A of non-zero principal fractional ideals. The quotient

$$\mathrm{Cl}_A \coloneqq I_A/P_A \tag{9.38}$$

is called the (ideal) **class group** of A and its order h_A is called the **class number** of A. By definition, we have

$$I = J \text{ in } \mathrm{Cl}_A \quad \text{if and only if} \quad aI = bJ \text{ for some } 0 \neq a, b \in A\ . \tag{9.39}$$

Moreover, we have an exact sequence

$$1 \longrightarrow A^\times \longrightarrow K^\times \longrightarrow I_A \longrightarrow \mathrm{Cl}_A \longrightarrow 1 \tag{9.40}$$

of (multiplicative) abelian groups.

In general, one cannot say much about the class group because it's too complicated because one basically needs to understand the whole ideal theory of the ring. The first problem already is: which fractional ideals are actually invertible? We have the following restriction.

Lemma 9.2.7. *If I is invertible, then I is finitely generated.*

Proof. Since $II^{-1} = A$, we can write $1 = \sum_{i=1}^n a_i b_i$ with $a_i \in I$ and $b_i \in I^{-1}$. We certainly have $A \cdot \{a_1, \ldots, a_n\} \subseteq I$. The converse holds as well because for $a \in I$ we have $a = \sum_{i=1}^n (ab_i)a_i$ and $ab_i \in II^{-1} = A$. □

The following theorem gives a characterization of the nicest possible situation.

Theorem 9.2.8. *All non-zero fractional ideals are invertible if and only if A is a Dedekind domain.*

Proof. Suppose that A is a Dedekind domain. It is enough to show that all non-zero (non-fractional) ideals are invertible because a general non-zero fractional ideal is of the form $\frac{1}{a}I$ with $0 \neq a \in A$ and I a non-zero ideal in A, and we have $(\frac{1}{a}I)(aI^{-1}) = II^{-1}$, so if I is invertible, then $\frac{1}{a}I$ is invertible as well. Furthermore, since every non-zero ideal is a product of prime ideals, it is enough to show that every non-zero prime ideal is invertible. Let P be a non-zero prime ideal. Choose $0 \neq a \in P$ and let $(a) = P_1 \cdots P_n$ be the factorization of (a) into prime ideals. Since (a) is invertible, all the P_i are invertible as well (if a product xy in a commutative monoid is equal to a invertible element u, then x and y are invertible as well with $x^{-1} = u^{-1}y$ and $y^{-1} = u^{-1}x$). As P is a prime above (a), we know from Remark 9.1.18 that P is one of the P_i and therefore P is invertible.

Now, suppose that all non-zero fractional ideals are invertible. Since any ideal is a fractional ideal and invertible ideals are finitely generated by Lemma 9.2.7, it follows that all ideals are finitely generated, i.e. A is noetherian. In light of Theorem 9.1.17 and Theorem 9.1.16 it is therefore enough to show that for any non-zero prime ideal P the maximal ideal M_P of the localization A_P is principal. Note that any ideal in A_P is of the form IA_P for I an ideal in A. Since I is invertible, we get

$$(IA_P)(I^{-1}A_P) = (II^{-1})A_P = AA_P = A_P\ , \tag{9.41}$$

i.e. IA_P is an invertible ideal in A_P. In particular, M_P is invertible. Since $P \neq 0$, the localization A_P is not a field, hence $\dim(A) \geq 1$ and therefore $M_P/M_P^2 \neq 0$, i.e.

$M_P \neq M_P^2$. We can thus choose $a \in M_P$ with $a \notin M_P^2$. We have $M_P M_P^{-1} = A_P$, hence $aM_P^{-1} \subseteq A_P$. Moreover, since $a \notin M_P^2$, it follows that $aM_P^{-1} \nsubseteq M_P$. This implies $aM_P^{-1} = A_P$ and therefore $aA_P = M_P$, i.e. M_P is principal. □

For *Dedekind domains* the ideal class group thus involves all of the non-zero ideals and serves as a measure of how far the ideals are from being principal. This is made more precise by the following fact.

Theorem 9.2.9. *For a Dedekind domain A the following are equivalent:*

(1) *A is a unique factorization domain.*
(2) *A is a principal ideal domain.*
(3) *Cl_A is trivial, i.e. $I_A = P_A$.*

Proof.
(1) ⇒ (2): Since every non-zero ideal factorizes into a product of prime ideals, it is enough to show that every prime ideal is principal. Let P be a non-zero prime ideal. Choose $0 \neq p \in P$. Then $(p) \subseteq P$. Since A is a unique factorization domain, we can write $p = up_1^{\nu_1} \cdots p_n^{\nu_n}$ with a unit u and prime elements p_i. It follows that

$$(p) = P_1^{\nu_1} \cdots P_n^{\nu_n} \subseteq P \, , \tag{9.42}$$

where $P_i \coloneqq (p_i)$. Since P is a prime ideal above (p), it follows from Remark 9.1.18 that P is equal to one of the P_i, hence P is principal.

(2) ⇒ (3): Let I be an invertible ideal. Then aI is an ideal in A for some $0 \neq a \in A$, hence $aI = bA$ for some $b \in A$ since A is a principal ideal domain. It follows that $I = \frac{b}{a}A$, i.e. I is principal.

(3) ⇒ (1): We show that A is a principal ideal domain, then we know from Lemma 1.5.24 that A is a unique factorization domain. Let I be a non-zero ideal in A. Then I is invertible by Theorem 9.2.8 (we are assuming that A is a Dedekind domain). Hence, $I \in I_A = P_A$, i.e. I is principal. We thus have $I = aA$ for some $a \in K$. Since $I \subseteq A$, we must have $a \in A$, i.e. $I = (a)$ is a principal ideal. □

Remark 9.2.10. We have now proven that the ideal class group of a Dedekind domain A is generated by the ideal classes of all non-zero prime ideals. Unfortunately, this is as much as one can say in general and class groups can get arbitrarily complicated: one can show that every abelian group is the class group of some Dedekind domain!

The situation for the ring of integers $\mathcal{O}_L$ in a number field L is somewhat better. Here one can show that $\mathcal{O}_L$ is indeed a *finite* group. The proof is based on geometric arguments using Minkowski's "geometry of numbers". Moreover, one can show that there is a constant C such that any ideal is equivalent in the class group to an ideal whose "ideal norm" is bounded by C. Hence, one just needs to focus on such ideals. The point is that there are only finitely many. One then factorizes these ideals into prime ideals and gets in this way a small set of generators of the class group. One then needs to determine the relations (9.39) between these generators and ends up with the complete structure of the class group. All this can be done algorithmically.

Just to give an idea in case of our favorite example $\mathbb{Z}[\sqrt{-5}]$. We know this is not a principal ideal domain, so the class group is not trivial. It is not too difficult to show that the prime ideal $P \coloneqq (2, 1 + \sqrt{-5})$ is not principal and thus gives a non-trivial element (of order 2) in the class group. What is more difficult to show is that P already generates the class group, so

$$\mathrm{Cl}_{\mathbb{Z}[\sqrt{-5}]} \simeq \mathbb{Z}/2\mathbb{Z}\ . \tag{9.43}$$

That one can—in principle—compute a class group of a ring of integers doesn't mean that one understands much about class groups of rings of integers in general. For example, it is not even known if among the real quadratic fields $\mathbb{Q}(\sqrt{d})$ with $d > 1$ square-free there are *infinitely* many with class number 1, i.e. where the ring of integers is a principal ideal domain—it is an old conjecture by Gauss that there are infinitely many!

Exercises.

EXERCISE 9.2.11. Show that if A is noetherian, then the fractional ideals of A are precisely the finitely generated A-submodules of K.

CHAPTER 10

Primary decomposition

In the last chapter we have seen that the (unique) factorization of ideals into prime ideals is only possible in a very special class of rings: the Dedekind domains—which are noetherian, normal, 1-dimensional integral domains. In particular, in dimension > 1 we never have factorizations of arbitrary ideals. Is there anything we can do in general?

It is always good to start from geometry. Recall from Proposition 2.7.7 that every topological space X is the union of its irreducible components X_λ, which are by definition the maximal irreducible closed subsets. Applied to the closed subset $\mathrm{V}(I)$ of $\mathrm{Spec}(A)$ defined by an ideal I in a ring A, we get

$$\mathrm{V}(I) = \bigcup_\lambda \mathrm{V}(P_\lambda) \ , \tag{10.1}$$

where the P_λ are the minimal prime ideals in A lying above I. Taking the ideal operator I from (2.64), we get

$$\sqrt{I} = \mathrm{I}(\mathrm{V}(I)) = \bigcap_{P \in \mathrm{V}(I)} P = \bigcap_\lambda P_\lambda \ , \tag{10.2}$$

which is exactly the statement from Corollary 2.6.5. This means we can write any radical ideal in any ring as the intersection of the minimal prime ideals above that ideal.

But what if the ideal I is not radical? Let us assume that A is noetherian. Then we know from Exercise 7.3.9 that the intersection (10.2) is finite. By (2.58) finite intersections of ideals correspond to finite unions of closed subsets of $\mathrm{Spec}(A)$ under the V-operator. This motivates that we want to find an expression of I as an intersection of ideals which yields (10.2) when taking the radical—such an expression is thus a refinement on the algebraic side of the decomposition into irreducible components on the geometric side (which only sees the radical). But how can we refine the decomposition (10.2) to get such a decomposition of I, i.e. which kind of ideals do we have to admit? If A is a Dedekind domain, then the powers of the minimal primes above I are exactly what we need: if $I = P_1^{\nu_1} \cdots P_n^{\nu_n}$ is the factorization of I into powers of pairwise distinct prime ideals P_i, then since A is 1-dimensional, the P_i are pairwise coprime, so by Exercise 1.3.25 (Chinese remainer theorem) we have

$$I = P_1^{\nu_1} \cdots P_n^{\nu_n} = P_1^{\nu_1} \cap \ldots \cap P_n^{\nu_n} \ . \tag{10.3}$$

However, in general powers of the minimal primes are not enough. Consider for example the ideal $I := (X_1, X_2^2)$ in $K[X_1, X_2]$. The only prime ideal above I is $P := (X_1, X_2)$ and

$$P^2 = (X_1, X_2)^2 = (X_1^2, X_1 X_2, Y_1^2) \subsetneq I \subsetneq P \ , \tag{10.4}$$

hence I is not an intersection of powers of prime ideals.

The idea of this chapter is to filter out for any ideal I some special ideals called **primary ideals** whose radicals are prime ideals (called **associated primes** of I) and to write I as an intersection of primary ideals. Such an expression is called a **primary decomposition** and we will show that this always exists—but in general it won't be unique. All the minimal primes will be among the associated primes but there may be more: the so-called **embedded components**, which carry their name because you can't see them geometrically. To give an idea, the ideal (X_1, X_2^2) from above will be primary and this is already the primary decomposition—we cannot do better.

While primary decomposition is a useful tool in computational applications, I find it is not so much in theory—here, the associated primes are actually more important than primary decompositions!

10.1. Generalities on ideal decompositions

As motivated above, we want to express an ideal I as an intersection $I = \bigcap_{i=1}^{n} Q_i$ of certain ideals Q_i. We call such an expression a **decomposition** of I. Similar to factorizations into prime numbers, we want a decomposition of an "atomic nature". This leads to the following concept.

Definition 10.1.1. An ideal I is called **irreducible** if there is no non-trivial decomposition of I, i.e. if $I = I_1 \cap I_2$, then $I = I_1$ or $I = I_2$.

Similar to the existence of a factorization of an element into irreducible elements in a noetherian ring, we have:

Lemma 10.1.2. *In a noetherian ring every ideal admits a decomposition into irreducible ideals.*

Proof. Let A be a noetherian ring and suppose the claim is not true in A. Let Σ be the set of ideals which do not admit a decomposition into irreducible ideals. Then Σ is non-empty by assumption. Since A is noetherian, Σ has a maximal element I. The ideal I is not irreducible, so $I = I_1 \cap I_2$ with $I \subsetneq I_1$ and $I \subsetneq I_2$. But then $I_1, I_2 \notin \Sigma$, so I_1, I_2 admit decompositions into irreducible ideals and then so does I—a contradiction. □

10.2. Primary ideals

We are now going to define the "atoms" we want to consider for ideal decompositions.

Definition 10.2.1. An ideal Q in a ring A is called **primary** if $A/Q \neq 0$ and every zero-divisor in A/Q is already nilpotent.

Lemma 10.2.2. *An ideal Q is primary if and only $Q \neq A$ and $xy \in Q$ for some $x, y \in A$ implies that $x \in Q$ or $y^n \in Q$ for some $n > 0$.*

Proof. The condition $Q \neq A$ is equivalent to $A/Q \neq 0$. Let Q be primary. If $xy \in Q$, then $\overline{xy} = 0 \in A/Q$. If $x \notin Q$, then $\overline{x} \neq 0 \in A/Q$, hence $\overline{y} \in A/Q$ is a zero-divisor. Since Q is primary, this means that $y^n \in Q$ for some n. The converse statement is proven analogously. □

Lemma 10.2.3. *If A is noetherian, then irreducible ideals are primary.*

Proof. By passing to the quotient, it is enough to prove the claim for the zero ideal. So, suppose (0) is irreducible. We need to prove that (0) is primary. Let $xy = 0$ and suppose that $x \neq 0$. We need to show that $y^n = 0$ for some $n > 0$. Consider the ideal chain $\operatorname{Ann}(y) \subseteq \operatorname{Ann}(y^2) \subseteq \ldots$. Since A is noetherian, there is n such that $\operatorname{Ann}(y^n) = \operatorname{Ann}(y^{n+1})$. Let $a \in (x) \cap (y^n)$. Then $a = bx$ for some $b \in A$, hence $ay = bxy = 0$. Moreover, $a = cy^n$ for some $c \in A$, hence $0 = ay = cy^{n+1}$, i.e. $c \in \operatorname{Ann}(y^{n+1}) = \operatorname{Ann}(y^n)$ and therefore $a = cy^n = 0$. We have thus shown $(0) = (x) \cap (y^n)$. Since (0) is irreducible and $(x) \neq 0$, we must have $(y^n) = 0$, i.e. $y^n = 0$. □

Lemma 10.2.4. *If Q is primary, then $\sqrt{Q}$ is a prime ideal. It is the unique minimal prime above Q and it is called the **associated prime** of Q.*

Proof. Let $xy \in \sqrt{Q}$. Then $x^m y^m = (xy)^m \in Q$ for some $m \in \mathbb{N}$. Since Q is primary, it follows that $x^m \in Q$ or $y^{mn} \in Q$ for some $n > 0$. Hence, $x \in \sqrt{Q}$ or $y \in \sqrt{Q}$. This proves that $\sqrt{Q}$ is prime. By Corollary 2.6.5 it is then clear that $\sqrt{Q}$ is the unique minimal prime above. □

Definition 10.2.5. The primary ideals whose associated prime is equal to a given prime P are called the **P-primary ideals** in A.

Example 10.2.6. If P is prime, then P is a P-primary ideal.

Lemma 10.2.7. *If Q is an ideal in A such that $\sqrt{Q}$ is maximal, then Q is primary. In particular, the powers M^n of a maximal ideal M are M-primary.*

Proof. Let $M := \sqrt{Q}$ be a maximal ideal. Then $A/Q \neq 0$. Since

$$\sqrt{Q} = \bigcap_{\substack{P \in \operatorname{Spec}(A) \\ P \supseteq Q}} P \tag{10.5}$$

by Corollary 2.6.5, it follows that A/Q has exactly one prime ideal, namely M/Q. In particular, A/Q is local with maximal ideal M/Q. If $\overline{x}$ is a zero-divisor in A/Q, we must have $\overline{x} \in M/Q$ since elements outside the maximal ideal are units. Since $M = \sqrt{Q}$, it follows that M/Q is the nilradical in A/Q, see Theorem 2.6.4. Hence, all elements in M/Q are nilpotent. In particular, $\overline{x}$ is nilpotent. This shows that Q is primary. □

Example 10.2.8. Let A be a Dedekind domain and let P be a non-zero prime ideal. We claim the P-primary ideals are precisely the powers P^n for $n > 0$. Note that P is maximal since A is 1-dimensional, so all powers of P are primary by Lemma 10.2.7. Conversely, let Q be a P-primary ideal. Then Q is the unique minimal prime above P by Lemma 10.2.4, so it follows from Remark 9.1.18 that in the factorization of Q only P can occur, i.e. $Q = P^n$ for some $n > 0$.

Example 10.2.9. In general, primary ideals and powers of prime ideals are two different things:

(1) Let $A := K[X_1, X_2]$ and $Q := (X_1, X_2^2)$. Then $A/Q = K[X_2]/(X_2^2)$. The zero-divisors in A/Q are scalar multiples of X_2, in particular they are nilpotent. Hence, Q is primary. We have $P := \sqrt{Q} = (X_1, X_2)$. But

$$P^2 \subsetneq Q \subsetneq P \,, \tag{10.6}$$

so Q is not a power of a prime. This is exactly the example we mentioned in the introduction of this chapter.

(2) Let $A \coloneqq K[X_1, X_2, X_3]/(X_1X_2 - X_3^2)$. Since (X_1, X_3) is a prime ideal above $(X_1X_2 - X_3^2)$ in $K[X_1, X_2, X_3]$, it follows that $P \coloneqq (\overline{X}_1, \overline{X}_3)$ is a prime ideal in A. We have $\overline{X}_1\overline{X}_2 = \overline{X}_3^2 \in P^2$ but $\overline{X}_1 \notin P^2$ and $\overline{X}_2 \notin \sqrt{P^2} = P$. Hence, P^2 is not a primary ideal in A.

10.3. Properties of primary decomposition

Definition 10.3.1. A **primary decomposition** of an ideal I in a ring A is an expression

$$I = \bigcap_{i=1}^{n} Q_i \tag{10.7}$$

with primary ideals Q_i. The primary decomposition is called **minimal** if the following conditions are satisfied:

(1) all the prime ideals $\sqrt{Q_i}$ are distinct;
(2) no Q_i can be removed, i.e. $Q_i \not\supseteq \bigcap_{j \neq i} Q_j$ for all i.

Lemma 10.3.2. *Every primary decomposition can be modified to a minimal one.*

Proof. Let $I = \bigcap_{i=1}^{n} Q_i$ be a primary decomposition. Let

$$\mathcal{P} \coloneqq \{\sqrt{Q_i} \mid i = 1, \ldots, n\} . \tag{10.8}$$

For $P \in \mathcal{P}$ let

$$Q_P \coloneqq \bigcap_{i \in \mathcal{I}(P)} Q_i \, , \quad \mathcal{I}(P) \coloneqq \{i = 1, \ldots, n \mid \sqrt{Q}_i = P\} . \tag{10.9}$$

We claim that Q_P is P-primary. First of all, we have

$$\sqrt{Q_P} = \bigcap_{i \in \mathcal{I}(P)} \sqrt{Q}_i = \bigcap_{i \in \mathcal{I}(P)} P = P . \tag{10.10}$$

Let $xy \in Q_P$ and suppose that $x \notin Q_P$. We have $xy \in Q_i$ for all $i \in \mathcal{I}(P)$ and there must exist an $i \in \mathcal{I}(P)$ with $x \notin Q_i$. Since Q_i is primary, this implies $y \in \sqrt{Q}_i = P = \sqrt{Q_P}$. This proves that Q_P is P-primary.

Now, $I = \bigcap_{P \in \mathcal{P}} Q_P$ is a primary decomposition in which all components have pairwise distinct radical. By removing all superfluous components, we get a minimal primary decomposition. □

Theorem 10.3.3. *In a noetherian ring, every ideal admits a (minimal) primary decomposition.*

Proof. This follows at once from Lemma 10.1.2 and Lemma 10.2.3. □

Let I be an ideal and recall from (9.11) the construction

$$(I \colon x) = \{a \in A \mid ax \in I\} \tag{10.11}$$

for an element $x \in A$. There's another interpretation of this: if we consider A/I as an A-module and let $\overline{x}$ be the image in A/I, then

$$(I \colon x) = \{a \in A \mid ax \in I\} = \{a \in A \mid a\overline{x} = 0\} = \operatorname{Ann}_A(\overline{x}) . \tag{10.12}$$

Definition 10.3.4. An **associated prime ideal** of I is a prime ideal P which is of the form $P = \sqrt{(I \colon x)}$ for some $x \in A$. We denote by $\operatorname{Ass}(I)$ the set of such ideals.

By definition, the associated primes of I are a special selection of prime ideals above I. They are of tremendous importance. We will shortly see that all minimal primes above I are among them.

LEMMA 10.3.5. *Let Q be a P-primary ideal and let $x \in A$.*

(1) *If $x \in Q$, then $(Q\colon x) = A$;*
(2) *If $x \notin Q$, then $(Q\colon x)$ is P-primary;*
(3) *If $x \notin P$, then $(Q\colon x) = Q$.*

In particular, $\operatorname{Ass}(Q) = \{P\}$.

PROOF.

(1): This is clear.

(2): We first show that $P = \sqrt{(Q\colon x)}$. Let $y \in (Q\colon x)$. Then $yx \in Q$. Since $x \notin Q$ and Q is primary, we must have $y \in \sqrt{Q} = P$. Hence, $Q \subseteq (Q\colon x) \subseteq P$. It follows that

$$P = \sqrt{Q} \subseteq \sqrt{(Q\colon x)} \subseteq \sqrt{P} = P \tag{10.13}$$

and therefore $P = \sqrt{(Q\colon x)}$. Now, let $yz \in (Q\colon x)$. Suppose that $y \notin \sqrt{(Q\colon x)} = P$. Since $xyz \in Q$ and Q is primary, we must have $xz \in Q$. Hence, $z \in (Q\colon x)$ and this shows that $(Q\colon x)$ is primary.

(3): Let $y \in (Q\colon x)$. Then $xy \in Q$. Since $x \notin P = \sqrt{Q}$ and Q is primary, it follows that $y \in Q$.

□

THEOREM 10.3.6. *Let $I = \bigcap_{i=1}^n Q_i$ be a minimal primary decomposition. Then*

$$\operatorname{Ass}(I) = \{\sqrt{Q_i} \mid i = 1, \ldots, n\}\ . \tag{10.14}$$

In particular, the radicals of the primary ideals in a primary decomposition of I are independent of the primary decomposition.

PROOF. Let $P_i \coloneqq \sqrt{Q_i}$. For every $x \in A$ we have the relation

$$(I\colon x) = \left(\bigcap_{i=1}^n Q_i\colon x\right) = \bigcap_{i=1}^n (Q_i\colon x) \tag{10.15}$$

and taking the radical we get

$$\sqrt{(I\colon x)} = \bigcap_{i=1}^n \sqrt{(Q_i\colon x)} = \bigcap_{\substack{i=1\\ x\notin Q_i}}^n P_i\ , \tag{10.16}$$

where in the second equality we use Lemma 10.3.5. Suppose that $\sqrt{(I\colon x)}$ is a prime ideal P, i.e. P is an associated prime of I. We claim that $P = P_i$ for some i. Suppose that $P \not\supseteq P_i$ for all i. Then for every i there is an element $x_i \in P_i$ with $x_i \notin P$. We have $\prod_{i=1}^n x_i \in \prod_{i=1}^n P_i \subseteq \bigcap_{i=1}^n P_i$. But $\prod_{i=1}^n x_i \notin P$ since P is prime. Hence, $P \not\supseteq \bigcap_{i=1}^n P_i$. But this is a contradiction since $P \supseteq \sqrt{I} = \bigcap_{i=1}^n \sqrt{Q}_i = \bigcap_{i=1}^n P_i$. We thus have $P \supseteq P_i$ for some i. On the other hand, $P = \bigcap_{i=1}^n P_i \subseteq P_i$ by (10.16), hence $P = P_i$. This proves that $\operatorname{Ass}(I) \subseteq \{P_1, \ldots, P_n\}$. On the other hand, since the decompositon of I is minimal, there is for each i an element $x_i \in \bigcap_{j\neq i} Q_j$ with $x_i \notin Q_i$. Then (10.16) shows that

$$\sqrt{(I\colon x_i)} = \bigcap_{\substack{j=1\\ x_i\notin Q_j}}^n P_j = P_i\ , \tag{10.17}$$

hence $P_i \in \mathrm{Ass}(I)$. □

Theorem 10.3.7. *If I is admits a primary decomposition, then the minimal elements in* $\mathrm{Ass}(I)$ *are precisely the minimal prime ideals above I.*

Proof. Let $I = \bigcap_{i=1}^n Q_i$ be a minimal primary decomposition and let $P_i \coloneqq P$. If P is a prime above I, then with the same argument as in the proof of Theorem 10.3.6 we see that $P \supseteq P_i$ for some i. Hence, if P is minimal above I, then $P = P_i$ and this is a minimal element in $\mathrm{Ass}(I)$. Conversely, let P_i be minimal in $\mathrm{Ass}(I)$. Let P be a prime ideal with $I \subseteq P \subseteq P_i$. As we have just argued, P lies above an element $P' \in \mathrm{Ass}(I)$. But then $P' \subseteq P \subseteq P_i$, so $P' = P = P_i$ since P_i is minimal in $\mathrm{Ass}(I)$. This shows that P_i is a minimal prime above I. □

We conclude that if $I = \bigcap_{i=1}^n Q_i$ is a primary decomposition, then $\sqrt{I} = \bigcap_{i=1}^n \sqrt{Q_i}$ is—after removing superfluous components—precisely the decomposition of I into the minimal prime ideals above I, i.e. the decomposition into irreducible components on the geometric side. Hence, primary decomposition is an algebraic refinement of the decomposition into irreducible components—that's exactly what we wanted. This also leads to the following terminology.

Definition 10.3.8. The minimal elements in $\mathrm{Ass}(I)$ are called **isolated**, all the others are called **embedded**. A component Q_i of a minimal primary decomposition $I = \bigcap_{i=1}^n Q_i$ of I is called **isolated** if its associated prime ideal (i.e. its radical) is isolated, and otherwise it is called **embedded**.

One can prove:

Theorem 10.3.9. *The isolated components of a minimal primary decomposition of I are independent of the minimal primary decomposition.*

Example 10.3.10. Let $A \coloneqq K[X_1, X_2]$ and let $I \coloneqq (X_1^2, X_1X_2)$. Let $P_1 \coloneqq (X_1)$ and $P_2 \coloneqq (X_1, X_2)$. Since P_1 is prime, it is a P_2-primary ideal. Since P_2 is maximal, the power P_2^2 is P_2-primary by Lemma 10.2.7. We have

$$(X_1^2, X_1X_2) = (X_1) \cap (X_1^2, X_1X_2, X_2^2) \ , \tag{10.18}$$

hence

$$I = P_1 \cap P_2^2 \tag{10.19}$$

is a minimal primary decomposition. In particular, it follows that

$$\mathrm{Ass}(I) = \{P_1, P_2\} \ . \tag{10.20}$$

The component P_1 is isolated but P_2^2 is embedded. We have another primary decomposition

$$I = (X_1^2, X_1X_2) = (X_1) \cap (X_1^2, X_2) \ , \tag{10.21}$$

which shows that the embedded components are not necessarily unique.

References

[1] D. D. Anderson and S. Valdes-Leon. “Factorization in commutative rings with zero divisors.” In: *Rocky Mountain J. Math.* 26.2 (1996), pp. 439–480. URL: `https://doi.org/10.1216/rmjm/1181072068`.

[2] M. F. Atiyah and I. G. Macdonald. *Introduction to commutative algebra.* Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1969, pp. ix+128.

[3] N. Bourbaki. *Commutative algebra. Chapters 1–7.* Elements of Mathematics (Berlin). Translated from the French, Reprint of the 1972 edition. Springer-Verlag, Berlin, 1989, pp. xxiv+625.

[4] P. L. Clark. *Commutative algebra.* URL: `http://math.uga.edu/~pete/integral2015.pdf`.

[5] D. Eisenbud. *Commutative algebra.* Vol. 150. Graduate Texts in Mathematics. With a view toward algebraic geometry. Springer-Verlag, New York, 1995, pp. xvi+785. URL: `https://doi.org/10.1007/978-1-4612-5350-1`.

[6] A. Grothendieck. “Éléments de géométrie algébrique. IV. Étude locale des schémas et des morphismes de schémas. I.” In: *Inst. Hautes Études Sci. Publ. Math.* 20 (1964), p. 259. URL: `http://www.numdam.org/item/PMIHES_1964__20__5_0`.

[7] A. Grothendieck and J. A. Dieudonné. *Eléments de géométrie algébrique. I.* Vol. 166. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer-Verlag, Berlin, 1971, pp. ix+466.

[8] K. Heinrich. “Some remarks on biequidimensionality of topological spaces and Noetherian schemes.” In: *J. Commut. Algebra* 9.1 (2017), pp. 49–63. URL: `https://doi.org/10.1216/JCA-2017-9-1-49`.

[9] T. Y. Lam. *A first course in noncommutative rings.* Second. Vol. 131. Graduate Texts in Mathematics. Springer-Verlag, New York, 2001, pp. xx+385. URL: `https://doi.org/10.1007/978-1-4419-8616-0`.

[10] M. Nagata. *Local rings.* Interscience Tracts in Pure and Applied Mathematics, No. 13. Interscience Publishers a division of John Wiley & Sons New York-London, 1962, pp. xiii+234.

[11] G. Scheja and U. Storch. *Lehrbuch der Algebra. Teil 2.* Mathematische Leitfäden. [Mathematical Textbooks]. Unter Einschluss der linearen Algebra. [Including linear algebra]. B. G. Teubner, Stuttgart, 1988, p. 816. URL: `https://doi.org/10.1007/978-3-322-80092-3`.

[12] The Stacks Project Authors. *Stacks Project.* URL: `https://stacks.math.columbia.edu`.

[13] U. Thiel. “Introduction to categorical thinking and categorification.” `https://ulthiel.com/math/wp-content/uploads/lecture-notes/Tensor-Categories.pdf`. 2021.

[14] O. Zariski and P. Samuel. *Commutative algebra, Volume I.* The University Series in Higher Mathematics. With the cooperation of I. S. Cohen. D. Van Nostrand Company, Inc., Princeton, New Jersey, 1958, pp. xi+329.

Index