A month of commutative algebra

Contents

1. A BRIEF REVIEW OF BASIC CONCEPTS

1.1. THE MAIN CHARACTER

DEFINITION 1.1.1. A *commutative ring* is a set A equipped with two operations $+$ and $\cdot$ such that

1. $(A, +)$ is an abelian group[1]
2. $\cdot$ is associative;
3. $\cdot$ is distributive over addition: $a \cdot (b + c) = a \cdot b + a \cdot c$ and $(b + c) \cdot a = b \cdot a + c \cdot a$; and
4. there is a multiplicative identity element $1_A \in A$.

Unless otherwise stated, every ring in these notes is commutative; in fact, A will by default denote a commutative ring. The multiplication symbol will usually be elided, writing ab instead of $a \cdot b$. The symbol 0_A denotes the additive identity of A. These subscripts will be suppressed if the context is clear.

EXERCISE 1.1.2. Prove that the multiplicative identity of a ring is unique. (HINT: Follow the proof for identity uniqueness in groups.)

EXERCISE 1.1.3. Prove that $0 \cdot x = 0$ for every element $x \in A$. Then prove that, if $0 = 1$, then $A = \{0\}$. This is called the *trivial ring* or *zero ring*.

1.2. QUOTIENTS, IDEALS, AND HOMOMORPHISMS

A subset $B \subseteq A$ is a *subring* if it is itself a commutative ring under the same operations as A.

If $S \subseteq A$, we can try to form a new, simpler, ring by collapsing all the elements of S to zero: if $x - y \in S$, then they become identical in the new ring.[2] For this, we want S to be closed under addition and multiplication. (It doesn't mean much to say that you're identifying x with 0 if you're not also identifying, say, $x + x$ with 0). We form the new set

$$A/S = \{x + S : x \in A\}$$

and define the operations

$$(x + S) \oplus (y + S) = (x + y) + S \qquad \text{and} \qquad (x + S) \odot (y + S) = xy + S.$$

EXERCISE 1.2.1. Show that $\oplus$ is well-defined when S is closed under addition.

The problem is that $\odot$ is not always well-defined. Pick elements $x, y \in A$ and $s_1, s_2 \in S$. We want to check whether

$$(x + S) \odot (y + S) = \big((x + s_1) + S\big) \odot \big((y + s_2) + S\big)$$

This is simple: the left side is $xy + S$; the right side is

$$(x + s_1)(y + s_2) + S = (xy + xs_2 + ys_1 + s_1 s_2) + S.$$

So for $\odot$ to be well-defined, we need $xs_2 + ys_1 \in S$. In particular, choosing $x = 1_A$ shows that $ys_1 \in S$ for every $y \in A$ and $s_1 \in S$.

This is a necessary condition for A/S to be well-defined. It turns out that it's also sufficient.

DEFINITION 1.2.2. A subset $\mathfrak{a} \subseteq A$ is called an *ideal* if it is closed under addition and *absorbative* under multiplication; that is,

[1] You might wonder why we assume that $+$ is commutative. It turns out that if $\cdot$ is distributive (property 3), then $+$ is necessarily commutative, even if $\cdot$ isn't. (To prove this, expand out $(a + b) \cdot (a + b)$ in two different ways.)

[2] This is an abbreviated story about quotient rings which is perhaps too brief if you haven't already seen quotient groups.

1. $a + b \in \mathfrak{a}$ whenever $a, b \in \mathfrak{a}$ and
2. $ax \in \mathfrak{a}$ whenever $a \in \mathfrak{a}$ and $x \in A$.

What we've shown is that if ${}^{A}\!/_{S}$ is well-defined, S is an ideal. It's not hard to check the converse.

EXERCISE 1.2.3. Show that $\oplus$ and $\odot$ are well-defined whenever S is an ideal.

From this point on, we'll generally use lower-case Gothic letters to denote ideals.
Ideals are intimately connected with ring homomorphisms.

DEFINITION 1.2.4. A map $\varphi\colon A \to B$ between two rings is a *homomorphism* if
1. $\varphi(x + y) = \varphi(x) + \varphi(y)$ for every $x, y \in A$,
2. $\varphi(xy) = \varphi(x)\varphi(y)$ for every $x, y \in A$, and
3. $\varphi(1_A) = \varphi(1_B)$.

A bijective homomorphism is called an *isomorphism*.

Unless otherwise noted, any map between two rings will be assumed to be a homomorphism.

EXERCISE 1.2.5. Suppose $\varphi\colon A \to B$. Prove that the kernel of φ is an ideal of A and ${}^{A}\!/_{\ker(\varphi)} \cong \operatorname{im}(\varphi)$. (HINT: Follow the proof of the first isomorphism theorem for groups.)

EXERCISE 1.2.6. Suppose $\varphi\colon A \to B$ and $\mathfrak{a}$ is an ideal of A.
(i) Prove that $\varphi(\mathfrak{a})$ is an ideal of $\operatorname{im}(\varphi)$.
(ii) Find an example to show that $\varphi(\mathfrak{a})$ might *not* be an ideal of B.
(iii) Prove that $\varphi^{-1}(\mathfrak{b})$ is an ideal of A whenever $\mathfrak{b}$ is an ideal of B.
(iv) Prove that the correspondence $\mathfrak{b} = \varphi^{-1}(\bar{\mathfrak{b}})$ is an order-preserving bijection between the ideals $\mathfrak{b}$ in A that contain $\mathfrak{a}$ and the ideals $\bar{\mathfrak{b}}$ in B.

EXERCISE 1.2.7. Suppose that $\mathfrak{a}$ and $\mathfrak{b}$ are two ideals of A. Show that $\mathfrak{a} \cap \mathfrak{b}$ is also an ideal of A. Show by example that $\mathfrak{a} \cup \mathfrak{b}$ may not be. Also show by example that if $\mathfrak{a}$ and $\mathfrak{b}$ are prime ideals, $\mathfrak{a} \cap \mathfrak{b}$ may not be.

DEFINITION 1.2.8. If $S \subseteq A$, the ideal *generated* by S is the smallest ideal of A that contains S; it is denoted (S). If $S = \{x_1, \ldots, x_n\}$, then we may write $(x_1, \ldots, x_n)$ instead of S. An ideal generated by a single element is called *principal*.

EXERCISE 1.2.9. Prove that $(S) = \bigcap_{\mathfrak{a} \supseteq S} \mathfrak{a}$.

EXERCISE 1.2.10. Prove that every ideal in $\mathbb{Z}$ is principal. (HINT: Consider the least positive element of the ideal.)

1.3. A FEW SPECIAL RINGS

Here are some rings with a bit more structure.

DEFINITION 1.3.1. An element $x \in A$ is called a *zero-divisor* if there is a nonzero element $y \in A$ such that $xy = 0$; it is called a *unit* if there is an element $y \in A$ so that $xy = 1$. The set of all units in A is denoted $A^{\times}$.

DEFINITION 1.3.2. A nonzero ring A is called an *integral domain* (or simply a *domain*) if it has no zero-divisors. It is called a *field* if every nonzero element is a unit.

Fields can be characterized in several other ways:

PROPOSITION 1.3.3. *The following are equivalent:*

1. *A is a field.*
2. *The only ideals of A are $\{0\}$ and A.*
3. *Every homomorphism $A \to B$ is injective whenever B is not the zero ring.*

Proof sketch. (1) $\Rightarrow$ (2). Prove that $(x) = A$ whenever x is a unit.
(2) $\Rightarrow$ (3). The kernel of the homomorphism is an ideal of A.
(3) $\Rightarrow$ (1). Pick a nonzero element $x \in A$ and consider the natural map $A \to A/(x)$. □

EXERCISE 1.3.4. The *cancellation property* says that $y = z$ whenever $xy = xz$ and $x \neq 0$. Prove that a ring has the cancellation property if and only if it is an integral domain.

DEFINITION 1.3.5. A *principal ideal domain* is an integral domain in which every ideal is principal.

EXERCISE 1.3.6. A nonzero non-unit element $x \in a$ is called *prime* if the ideal (x) is prime. It is called *irreducible* if whenever $x = yz$, either y or z is a unit.

(i) Show that every prime element is irreducible.
(ii) Show that every irreducible element in a principal ideal domain is prime.
(iii) Show that 3 is an irreducible element of $\mathbb{Z}[\sqrt{-5}]$. (That is, the ring $\{a + b\sqrt{-5} : a, b \in \mathbb{Z}\} \subseteq \mathbb{C}$.) (HINT: Suppose that $3 = yz$ where y and z are not units, and take the complex norms of each side. If $x \in \mathbb{Z}[\sqrt{-5}]$ is not 0 and not a unit, then what can $|x|$ be?)
(iv) On the other hand, $(2 + \sqrt{5})(2 - \sqrt{5}) = 9 \in (3)$, so 3 is not prime. Lesson: Irreducible elements are not always prime.

1.4. A FEW SPECIAL IDEALS

Just as we single out some nicer rings, we can single out some nicer ideals.

DEFINITION 1.4.1. A proper ideal $\mathfrak{a}$ is called *maximal* if there is no ideal strictly between $\mathfrak{a}$ and the whole ring; it is called *prime* if, whenever $xy \in \mathfrak{a}$, either $x \in \mathfrak{a}$ or $y \in \mathfrak{a}$.

It turns out that every ring has an instance of each.

THEOREM 1.4.2. *Every ring has a maximal ideal.*
Proof sketch. Apply Zorn's Lemma. □

In fact, we can say more:

PROPOSITION 1.4.3. *Every ideal $\mathfrak{a}$ that is not the whole ring is contained in a maximal ideal.*
Proof. We know that $A/\mathfrak{a}$ contains a maximal ideal $\mathfrak{m}$; by part (iv) of Exercise 1.2.6, its preimage in A is a maximal ideal that contains $\mathfrak{a}$. □

PROPOSITION 1.4.4. *If x is not a unit, then x is contained in a maximal ideal.*
Proof. If x is not a unit, then $(x) \neq A$. Apply Proposition 1.4.3. □

This will be clear in a bit. First, let's show one reason why these types of ideals are notable.

PROPOSITION 1.4.5.
1. *$A/\mathfrak{a}$ is an integral domain if and only if $\mathfrak{a}$ is prime.*
2. *$A/\mathfrak{a}$ is a field if and only if $\mathfrak{a}$ is maximal.*

Proof sketch. (1) $x + \mathfrak{a}$ is a zero-divisor in $A/\mathfrak{a}$ if and only there is an element $y \notin \mathfrak{a}$ so that $xy \in \mathfrak{a}$.
(2) Use Exercise 1.2.6 and Proposition 1.3.3. □

COROLLARY 1.4.6. *Every maximal ideal is prime.*
Proof. Every field is an integral domain. □

EXERCISE 1.4.7. Let $\varphi\colon A \to B$ be a ring homomorphism, $\mathfrak{a}$ be an ideal of A, and $\mathfrak{b}$ be an ideal of B.

(i) Prove that if $\mathfrak{b}$ is prime, then $\varphi^{-1}(\mathfrak{b})$ is a prime ideal of A.
(ii) Show by example that $\varphi(\mathfrak{a})$ might not be prime even if $\mathfrak{a}$ is. (Actually, it might not even be an ideal!)
(iii) Prove that, if $\mathfrak{a}$ is a prime ideal that contains $\ker\varphi$ and φ is surjective, then $\varphi(a)$ is a prime ideal. (So φ is an order-preserving bijection between the prime ideals that contain $\ker\varphi$ and the prime ideals of B.)

EXERCISE 1.4.8. Let A be a ring in which every element satisfies an equality $x^n = x$ for some positive integer n which may depend on x. Prove that every prime ideal in A is maximal. (HINT: What happens when you quotient?)

1.5. OPERATIONS ON IDEALS

DEFINITION 1.5.1. If $\mathfrak{a}$ and $\mathfrak{b}$ are ideals of A, we define their *sum* to be the set

$$\mathfrak{a} + \mathfrak{b} = \{a + b : a \in \mathfrak{a} \text{ and } b \in \mathfrak{b}\}$$

and their product $\mathfrak{ab}$ to be the ideal generated by the set of products ab with $a \in \mathfrak{a}$ and $b \in \mathfrak{b}$.

Elements of $\mathfrak{ab}$ are finite sums of the form $\sum_i a_i b_i$ with $a_i \in \mathfrak{a}$ and $b_i \in \mathfrak{b}$

EXERCISE 1.5.2. Prove that $\mathfrak{a} + \mathfrak{b}$ is an ideal, in fact the smallest ideal of A that contains both $\mathfrak{a}$ and $\mathfrak{b}$.

EXERCISE 1.5.3. Prove that $\mathfrak{a} \cap \mathfrak{b}$ is an ideal of A, and show that $\mathfrak{ab} \subseteq \mathfrak{a} \cap \mathfrak{b}$. Find two ideals $\mathfrak{a}, \mathfrak{b}$ of $\mathbb{Z}$ that make this inclusion strict.

Here's a rather different construction:

DEFINITION 1.5.4. The *radical* of an ideal $\mathfrak{a}$ is the set

$$\sqrt{\mathfrak{a}} = \{x \in A : x^n \in \mathfrak{a} \text{ for some } n \in \mathbb{N}\}.$$

The radical of the zero ideal is called the *nilradical* of A, denoted $\mathfrak{N}_A$, or simply $\mathfrak{N}$. Elements of the nilradical are called *nilpotent*.

Though it's difficult to tell, $\mathfrak{N}$ is a janky Gothic version of the letter N. Explicitly, an element $x \in A$ is nilpotent if there is a power $x^n = 0$.

PROPOSITION 1.5.5. *The radical of any ideal is itself an ideal.*

Proof. Let $\mathfrak{a}$ be an ideal and choose any two elements $x, y \in \sqrt{\mathfrak{a}}$. If $x^m \in \mathfrak{a}$ and $y^n \in \mathfrak{a}$, then

$$(x+y)^{m+n} = \sum_{i=0}^{m+n} \binom{m+n}{i} x^i y^{m+n-i} = 0,$$

since every term in the sum has either x^m or y^n as a factor. Moreover, for any $z \in A$, we have $(zx)^m = z^m x^m = 0$. So $\sqrt{\mathfrak{a}}$ is an ideal. □

EXERCISE 1.5.6. A *radical idea* is an ideal that's equal to its radical. In other words, $\mathfrak{a}$ is a radical ideal if whenever $x^n \in \mathfrak{a}$, we already have that $x \in \mathfrak{a}$. Prove that if $\mathfrak{a}$ and $\mathfrak{b}$ are radical ideals, then $\mathfrak{a} \cap \mathfrak{b}$ and $\mathfrak{ab}$ are also radical ideals.

EXERCISE 1.5.7. Show that every prime ideal is radical. Then show that the sum of two prime ideals is radical. Find an example to show that the sum of two (non-prime) radical ideals is not always radical. (HINT: Look in the ring $\mathbb{Z}[x]$.)

It turns out that the nilradical is intimately connected with prime ideals.

THEOREM 1.5.8. *The nilradical of a ring A is equal to the intersection of all prime ideals in A.*

Proof. Let $\cap\mathfrak{p}$ denote the intersection of all prime ideals. One direction is easy: If $x \in \mathfrak{N}$, then $x^n = 0 \in \mathfrak{p}$ for all prime ideals $\mathfrak{p}$. A simple argument shows that this means x is contained in every prime ideal, and therefore in $\cap\mathfrak{p}$.

To show the reverse containment, that $\cap\mathfrak{p} \subseteq \mathfrak{N}$, we proceed by complements: Choose an element $x \notin \cap\mathfrak{p}$; we want to find a specific prime ideal that does not contain x. To do this, apply Zorn's Lemma to the set of ideals

$$\{\mathfrak{a} \subseteq A : x^n \notin \mathfrak{a} \text{ for any } n \geq 1\}$$

ordered by containment. (This set is nonempty because it contains (0), and the union of a chain is an upper bound.) So we get a maximal ideal $\mathfrak{q}$ in this set.

We now prove that $\mathfrak{q}$ is prime. Choose any $y, z \in A \setminus \mathfrak{q}$. Since $\mathfrak{q}+(y)$ and $\mathfrak{q}+(z)$ strictly contain $\mathfrak{q}$, there are powers m and n such that $x^m \in \mathfrak{q} + (y)$ and $x^n \in \mathfrak{q} + (z)$. Therefore

$$x^{m+n} \in \big(\mathfrak{q} + (y)\big)\big(\mathfrak{q} + (z)\big) = \mathfrak{q} + (yz),$$

which means that $\mathfrak{q} + (yz)$ strictly contains $\mathfrak{q}$; so $yz \notin \mathfrak{q}$, and $\mathfrak{q}$ is prime. In sum, if $x \notin \mathfrak{N}$, then $x \notin \cap\mathfrak{p}$. □

Remark. If, in the second half of the proof, we replace $\{x^n\}_{n\geq 1}$ by any multiplicatively closed set S not containing 0, then it becomes a proof of the statement: *There is a maximal ideal that does not contain S, and this ideal is prime.*

COROLLARY 1.5.9. *The radical $\sqrt{\mathfrak{a}}$ is equal to the intersection of all prime ideals containing $\mathfrak{a}$.*

Proof. Consider the quotient map $\varphi\colon A \to A/\mathfrak{a}$ and apply Theorem 1.5.8 to the image. Then apply part (iv) of Exercise 1.4.7. □

By analogy to the nilradical, ring theorists define the intersection of all maximal ideals; this is called the *Jacobson radical* of A, denoted $\mathfrak{R}_A$. Its element-side characterization is this:

PROPOSITION 1.5.10. *An element $x \in A$ is an element of the Jacobson radical if and only if $1+xy$ is a unit for every $y \in A$.*

Proof. ($\Rightarrow$) Assume that $x \in \mathfrak{R}_A$ and suppose for the sake of contradiction that $1 + xy$ is not a unit. There is a maximal ideal $\mathfrak{m}$ that contains $1 + xy$. But $x \in \mathbb{R}_A \subseteq \mathfrak{m}$, so $xy \in \mathfrak{m}$, which implies that $1 \in \mathfrak{m}$, a contradiction.

($\Leftarrow$) Suppose that x is an element of A not contained in the maximal ideal $\mathfrak{m}$. Then $(x, \mathfrak{m}) = A$, so there are elements $y \in A$ and $z \in \mathfrak{m}$ so that $z + xy = 1$. Therefore $1 - xy \in \mathfrak{m}$, so $1 - xy$ is not a unit. □

EXERCISE 1.5.11. Let $A[\![x]\!]$ denote the ring of formal power series over A. (That is, elements look like $f = \sum_{n=0}^{\infty} a_n x^n$.)

(i) Prove that $f = \sum_{n=0}^{\infty} a_n x^n$ is a unit in $A[\![x]\!]$ if and only if a_0 is a unit of A.
(ii) Prove that $f = \sum_{n=0}^{\infty} a_n x^n$ is an element of the Jacobson radical of $A[\![x]\!]$ if and only if a_0 is an element of the Jacobson radical of A.

One last thing, just for kicks: A ring is called *local* if it has only one maximal ideal. If k is a field, then $k[\![x]\!]$ is local: Since a proper ideal can't contain a unit, the constant term of every element in a maximal ideal of $k[\![x]\!]$ is 0. But set is itself the ideal (x), so (x) is the only maximal ideal. (Take a minute to unpack that argument.)

EXERCISE 1.5.12. Prove that a ring is local if and only if its Jacobson radical is maximal.

PROPOSITION 1.5.13. *Let A be a ring and $\mathfrak{m}$ be an ideal of A. The ring A is local and $\mathfrak{m}$ is its unique maximal ideal if and only if every element in $A \setminus \mathfrak{m}$ is a unit.*

Proof. ($\Rightarrow$) If $x \in A \setminus \mathfrak{m}$ is not a unit, then x is contained in a maximal ideal that is manifestly not $\mathfrak{m}$. So every element in $A \setminus \mathfrak{m}$ must be a unit.

($\Leftarrow$) If every element in $A \setminus \mathfrak{m}$ is a unit, then any proper ideal of A is contained in $\mathfrak{m}$, so $\mathfrak{m}$ is uniquely maximal. □

2. TWO QUICK THINGS

2.1. THE CHINESE REMAINDER THEOREM

It's an unfortunate name, but it's known the world over and basically the only one used.[3]

We already know the group theory version: If $\gcd(m, n) = 1$, then $\mathbb{Z}/mn\mathbb{Z} \cong \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$. It's not too hard to prove that the same statement holds when considering these groups as rings. But we want to move to something more general: If we have two ideals $\mathfrak{a}, \mathfrak{b} \subseteq A$, when is it true that $A/\mathfrak{ab} \cong A/\mathfrak{a} \times B/\mathfrak{b}$?

Let's start by considering the map $A \to A\mathfrak{a} \times A/\mathfrak{b}$ defined by

$$\varphi(x) = (x + \mathfrak{a},\, x + \mathfrak{b}).$$

LEMMA 2.1.1. *If A is a commutative ring and $\mathfrak{a}$ and $\mathfrak{b}$ are ideals of A, then:*

(i) $\ker(\varphi) = \mathfrak{a} \cap \mathfrak{b}$.
(ii) *φ is surjective if and only if $\mathfrak{a} + \mathfrak{b} = (1)$.*
(iii) *If condition* (ii) *holds, then $\mathfrak{ab} = \mathfrak{a} \cap \mathfrak{b}$.*

Proof. Statement (i) is clear: if $x \in \ker(\varphi)$, then $x \bmod \mathfrak{a}_i = 0$ for every i; in other words, $x \in \mathfrak{a}_i$ for every i.

For (ii), first suppose that $\mathfrak{a} + \mathfrak{b} = (1)$. Then

$$\bigcup_{x \in A} \varphi(x) = \bigcup_{\substack{a \in \mathfrak{a} \\ b \in \mathfrak{b}}} \varphi(a + b) = \bigcup_{\substack{a \in \mathfrak{a} \\ b \in \mathfrak{b}}} (b + \mathfrak{a}, a + \mathfrak{b}) = A/\mathfrak{a} \times A/\mathfrak{b},$$

so φ is surjective. On the other hand, suppose that φ is surjective and $\mathfrak{a} + \mathfrak{b} = \mathfrak{c}$. There is some $x \in A$ so that $\varphi(x) = (0 + \mathfrak{a},\, 1 + \mathfrak{b})$. We see that $x \in \mathfrak{a} \subseteq \mathfrak{c}$ and $x = 1 + b$ for some $b \in \mathfrak{b} \subseteq c$, which implies that $1 \in \mathfrak{c}$. So $\mathfrak{a} + \mathfrak{b} = (1)$.

Finally, suppose that $\mathfrak{a} + \mathfrak{b} = (1)$. Then there are $a \in \mathfrak{a}$ and $b \in \mathfrak{b}$ such that $a + b = 1$. For every $c \in \mathfrak{a} \cap \mathfrak{b}$, we have

$$c = c(a + b) = ca + cb \in \mathfrak{ab},$$

so $\mathfrak{a} \cap \mathfrak{b} \subseteq \mathfrak{ab}$. The reverse inclusion is Exercise 1.5.3. □

COROLLARY 2.1.2. *$A/\mathfrak{ab} \cong A/\mathfrak{a} \times A/\mathfrak{b}$ if and only if $\mathfrak{a} + \mathfrak{b} = (1)$.*

If $\mathfrak{a} + \mathfrak{b} = (1)$, then $\mathfrak{a}$ and $\mathfrak{b}$ are called *comaximal*. They're sometimes called *coprime*, but beware—but this can be a bit confusing. In the ring $k[x, y]$, for example, the ideals (x) and y feel "coprime," since they don't share any common factors, but $(x) + (y) \neq (1)$.

Of course, Corollary 2.1.2 extends to more than two ideals:

THEOREM 2.1.3 (Chinese remainder theorem). *If $\mathfrak{a}_1, \ldots, \mathfrak{a}_n$ are ideals of A and $\mathfrak{a}_i + \mathfrak{a}_j = (1)$ for each pair of indices i, j, then*

$$A/\mathfrak{a}_1 \cdots \mathfrak{a}_n \cong A/\mathfrak{a}_1 \times \cdots \times A/\mathfrak{a}_n.$$

[3] See here for a discussion of the history of the name.

Proof sketch. This is by induction using Corollary 2.1.2. We need only show that if every pair of ideals is comaximal, then $\mathfrak{a}_1$ and $\mathfrak{a}_2 \cdots \mathfrak{a}_n$ are. There are $x_i \in \mathfrak{a}_1$ and $y_i \in \mathfrak{a}_i$ for each $2 \leq i \leq n$ such that $x_i + y_i = 1$. Then

$$1 = (x_2 + y_2) \cdots (x_n + y_n) \in \mathfrak{a}_1 + \mathfrak{a}_2 \cdots \mathfrak{a}_n,$$

so $\mathfrak{a}_1$ and $\mathfrak{a}_2 \cdots \mathfrak{a}_n$ are comaximal. □

One of the easiest ways to find comaximal ideals is to simply use maximal ones: If $\mathfrak{m}_1$ and $\mathfrak{m}_2$ are maximal, then $\mathfrak{m}_1 + \mathfrak{m}_2$ is always the whole ring.

COROLLARY 2.1.4. *If $\mathfrak{m}_1, \ldots, \mathfrak{m}_n$ are maximal ideals of A, then*

$$A/\mathfrak{m}_1 \cdots \mathfrak{m}_n \cong A/\mathfrak{m}_1 \times \cdots \times A/\mathfrak{m}_n.$$

EXERCISE 2.1.5. Prove that, in a principal ideal domain, every nonzero prime ideal is maximal.

So in a PID, we can apply the Chinese Remainder theorem to arbitrary prime ideals.

2.2. SPECTRUM OF A RING

Here's something else which it's not quite clear why we're defining. There's usually the vague motivation that "boy, prime ideals sure are important, aren't they?" So, there you go. Anyway. To the definition.

DEFINITION 2.2.1. The *spectrum* of a ring A, denoted $\mathrm{Spec}(A)$, is the set of prime ideals of A.

EXERCISE 2.2.2. For each subset $E \subseteq A$, let $V(E)$ denote the set of prime ideals of A which contain E. Prove that

(i) $V(E) = \mathrm{Spec}(A)$ if and only if $E \subseteq \mathfrak{N}_A$.
(ii) If $\mathfrak{a}$ is the ideal generated by E, then $V(E) = V(\mathfrak{a}) = V(\sqrt{\mathfrak{a}})$.
(iii) $V(0) = \mathrm{Spec}(A)$ and $V(A) = \emptyset$.
(iv) If $(E_i)_{i \in I}$ is an arbitrary collection of subsets of A, then

$$\bigcap_{i \in I} V(E_i) = V\left(\bigcup_{i \in I} E_i \right)$$

(v) If $\mathfrak{a}$ and $\mathfrak{b}$ are ideals of A, then $V(\mathfrak{a} \cap \mathfrak{b}) = V(\mathfrak{a}\mathfrak{b}) = V(\mathfrak{a}) \cup V(\mathfrak{b})$.
(vi) Use parts (iii)–(v) to deduce that the sets $\{V(E) : E \subseteq A\}$ satisfy the axioms for the closed sets of a topology on $\mathrm{Spec}(A)$.

The topology generated by the closed sets $V(E)$ is called the *Zariski topology* of $\mathrm{Spec}(A)$.

DEFINITION 2.2.3. A topological space X is called *irreducible* if whenever $X = U \cup V$ and both U and V are closed, then either $U = X$ or $V = X$. Equivalently, X is irreducible if every two nonempty open sets intersect.

PROPOSITION 2.2.4. *$\mathrm{Spec}(A)$ is irreducible if and only if $\mathfrak{N}_A$ is prime.*

Proof. Every open set in $\mathrm{Spec}(A)$ has the form $V(E)^c$ for some set $E \subseteq A$. So $\mathrm{Spec}(A)$ is irreducible if and only if $V(E)^c \cap V(F)^c \neq \emptyset$ for every pair of sets $E, F \subseteq A$ such that $V(E), V(F) \not\subseteq \mathfrak{N}_A$.

($\Leftarrow$) Suppose that $\mathfrak{N}_A = \mathfrak{p}$. For any two sets $E, F \subseteq A$ such that $V(E)^c, V(F)^c \not\subseteq \mathfrak{N}_A$, we have

$$\mathfrak{p} \in V(E)^c \cap V(F)^c,$$

so $\mathrm{Spec}(A)$ is not irreducible.

($\Rightarrow$) Assume that $\mathrm{Spec}(A)$ is irreducible and suppose for the sake of contradiction that the nilradical is not a prime ideal; so there are $x, y \in A \setminus \mathfrak{N}_A$ such that $xy \in \mathfrak{N}_A$. Since $x, y \notin \mathfrak{N}_A$, the sets $V(\{x\})^c$ and $V(\{y\})^c$ are not empty; by assumption, there is a prime ideal $\mathfrak{p}$ in their intersection. This prime ideal $\mathfrak{p}$ contains neither x nor y, but $xy \in \mathfrak{N}_A \subseteq \mathfrak{p}$, which contradicts the primality of $\mathfrak{p}$. So if $\mathrm{Spec}(A)$ is irreducible, $\mathfrak{N}_A$ is prime. □

A maximal ideal $\mathfrak{m}$ is called a "closed point" in $\operatorname{Spec}(A)$, since $V(\mathfrak{m}) = \{\mathfrak{m}\}$, so the point $\mathfrak{m} \in \operatorname{Spec}(A)$ forms a closed set.

3. MODULES

3.1. THE BASICS

The whole idea is encaptulated in the analogy

$$\text{module} : \text{ring} \;::\; \text{vector space} : \text{field}$$

To wit:

DEFINITION 3.1.1. An *A-module* is an additive abelian group M on which A acts linearly: There is a map $A \times M \to M$ mapping $(a, x) \mapsto ax$ such that

1. $a(x + y) = ax + ay$,
2. $(a + b)x = ax + bx$,
3. $a(bx) = (ab)x$, and
4. $1x = x$

for all $a, b \in A$ and $x, y \in M$.

If you're keen for a more abstract viewpoint, you can consider an A-module as a pair (M, μ), where $\mu\colon A \to \operatorname{End}(M)$, the set of group homomorphisms $M \to M$. But you needn't be keen.

Here are several ready examples of a module:

EXAMPLE 3.1.2.

1. If A is a field, then an A-module is a vector space over A.
2. Any ideal of A is an A-module.
3. Every group is a $\mathbb{Z}$-module by setting $nx = \underbrace{x + \cdots + x}_{n}$ if $n > 0$ and its negative if $n < 0$ (and $0x = 0$).
4. The set A^n is the *free module* with n generators over A.

$\Diamond$

You can do basically all the same initial things with modules that you can with groups and rings. So here come the definitions.

DEFINITION 3.1.3. If M and N are two A-modules, a map $\varphi\colon M \to N$ is called an *A-module homomorphism*, or *A-linear* if $\varphi(x + y) = \varphi(x) + \varphi(y)$ and $\varphi(ax) = a\varphi(x)$ for all $a \in A$ and $x, y \in M$. (In other words, φ is a group homomorphism $M \to N$ that commutes with the actions of A.) As usual, φ is an *isomorphism* if it is also a bijection.

DEFINITION 3.1.4. A *submodule* of the A-module M is a subset that is itself an A-module.

EXERCISE 3.1.5. Suppose that N is a submodule of M. Show that the quotient group M/N is an A-module under the action $a(x + N) = ax + N$. (In particular, prove that this is well-defined and, in fact, an action.) This module is also denoted M/N and is called the *quotient module* of M by N.

EXERCISE 3.1.6. Prove that if $\varphi\colon M \to N$ is an A-module homomorphism, then $\ker(\varphi)$ is a submodule of M and $M/\ker(\varphi) \cong \operatorname{im}(\varphi)$. In particular, if φ is surjective, then $M/\ker(\varphi) \cong N$.

DEFINITION 3.1.7. If $\varphi\colon M \to N$ is A-linear, then the *cokernel* of φ is $\operatorname{coker}(\varphi) = N/\operatorname{im}(\varphi)$.

Just like with groups and fractions, we can cancel denominators:

PROPOSITION 3.1.8. *If $L \subseteq N \subseteq N$ are A-modules, then*

$$M/N = \frac{M/L}{N/L}.$$

Proof. Ther is a simple map from $\varphi\colon M/L \to M/N$ by sending $x + L \mapsto x + N$ (which is well-defined since $L \subseteq N$). This map is surjective and its kernel is N/L, so applying the result of Exercise 3.1.6 completes the proof. □

And, of course, there are two easy ways to build bigger modules—we just extend the definitions for groups.

DEFINITION 3.1.9. If $\{M_i\}_{i\in I}$ is a collection of A-modules, then the *direct product* $\prod_{i\in I} M_i$ is the module on the set $\{(x_i)_{i\in I} : x_i \in M_i\}$ with the action induced componentwise: $a(x_i)_{i\in I} = a(x_i)_{i\in I}$. The *direct sum* $\bigoplus_{i\in I} M_i$ is the submodule of $\prod_{i\in I}$ of families $(x_i)_{i\in I}$ where all but finitely many x_i are 0.

If I is a finite set, then $\prod_{i\in I} M_i = \bigoplus_{i\in I} M_i$, but equality does not hold if I is infinite. The notation A^n denotes the direct product $\prod_{i=1}^n A$. The direct product and sum $\prod_{i\in I} A$ and $\bigoplus_{i\in I} A$ are sometimes denoted A^I and $A^{(I)}$, respectively.

3.2. FINITELY GENERATED MODULES

DEFINITION 3.2.1. If $\{N_i\}_{i\in I}$ is a collection of submodules of M, then $\sum_{i\in I} N_i$ is the collection of finite sums $\sum_j x_j$ with each x_j a member of some N_i.

It's straightforward to check that $\sum_{i\in I} N_i$ and $\bigcap_{i\in I} N_i$ are always submodules of M.

DEFINITION 3.2.2. Given an element $x \in M$, the set of elements $\{ax : a \in A\}$ is called the *submodule generated by* x and is denoted Ax or (x). If $M = \sum_{i\in I} Ax_i$, then the set $\{x_i\}_{i\in I}$ *generates* M. We say that M is *finitely generated* if there is a finite generating set.

In other words, if M is finitely generated, then there are elements $x_1, \ldots, x_n \in M$ such that every element in M can be expressed as a linear combination $\sum_{i=1}^n a_i x_i$.

There's one easy way to make finitely generated A-modules: take the product of A with itself finitely many times; it is generated by the n elements

$$e_i = (\underbrace{0, \ldots, 0}_{n-1}, 1, 0, \ldots, 0).$$

The set $\{e_i\}$ is called the *standard basis* of A^n. Moreover, every quotient A^n is finitely generated. The converse is (perhaps a little surprisingly) true, too.

PROPOSITION 3.2.3. *An A-module M is finitely generated if and only if it is (isomorphic to) a quotient of A^n for some n.*

Proof. ($\Leftarrow$) If $e_1, \ldots, e_n$ generate A^n, then the elements $e_i + N$ generate A^n/N.

($\Rightarrow$) Suppose $u_1, \ldots, u_n$ generate M and define the module homomorphism $\varphi\colon A^n \to M$ by $\varphi(e_i) = u_i$. Then φ is surjective, so $M \cong A^n/\ker(\varphi)$. □

EXERCISE 3.2.4. Prove that the quotient of any finitely generated module is itself finitely generated.

Finitely generated A-modules are nice: We have a concrete finite set which we can manipulate. But unlike in vector spaces, there's not always a generating set with unique representation. For example, any finite group G is a $\mathbb{Z}$-module, but Lagrange's theorem guarantees that $nx = (n+|G|)x$ for every $n \in \mathbb{Z}$.

DEFINITION 3.2.5. A *free basis* for a A-module M is a generating set $\{x_i\}_{i\in I}$ such that every element in M can be written uniquely as a finite linear combination $\sum_j a_j x_j$. A module that has a free basis is called *free*.

EXERCISE 3.2.6. Show that M has a free basis if and only if it is isomorphic to $\bigoplus_{i\in I} A$ for some set I.

So every finitely generated A-module with a free basis is isomorphic to A^n for some n.

EXERCISE 3.2.7. Show that every module is isomorphic to a quotient of some free module. (HINT: Consider $\bigoplus_{x\in M} A$.)

⚠ Not every submodule of a free module is free! If $A = \mathbb{C}[x, y]$, then A is of course a free module over itself. But the ideal (x, y), which is a submodule of A, is not free: It is not generated by a single element, and no two elements are linearly independent: If $a, b \in (x, y)$, then $ba - ab = 0$. When A is a field, however, modules become vector spaces, and every subspace of a vector space has a (free) basis.

⚠ In fact, it gets worse: Not every submodule of a finitely generated module is free! This time, set $A = \mathbb{C}[x_1, x_2, \dots]$; again A is a module over itself with one generator. But the ideal $(x_i)_{i\in\mathbb{N}}$ is not generated by any finite set.

3.3. ASIDE: SOME (NOT-QUITE-)LINEAR ALGEBRA

This section goes into more detail than is really necessary; you can simply skim it if you want. The highlights are the definition of the adjugate matrix (Definition 3.3.4) and its central property (Proposition 3.3.5).

We denote the set of $n\times n$ matrices with entries in A by $A^{n\times n}$. These matrices can be multiplied just as matrices over a field can; $A^{n\times n}$ is both a ring and an A-module. As an A-module, $A^{n\times n}$ is isomorphic to A^{n^2}, but as they are not isomorphic as rings. The determinant of a matrix, too, can be defined for matrices with ring elements.

As usual, we denote by I_n the matrix with 1's on the diagonal and 0's elsewhere.

DEFINITION 3.3.1. The *determinant* of a matrix $(a_{i,j}) \in A^{n\times n}$ is

$$\det(a_{i,j}) = \sum_{\sigma\in S_n} \prod_{i=1}^{n} (-1)^{\operatorname{sgn}(\sigma)} a_{i,\sigma(i)},$$

where $\operatorname{sgn}(\sigma)$ is the *sign* of σ, which is 1 is σ can be written as a product of an even number of transpositions and -1 otherwise.

You may be familiar with a rather more abstract characterization of the determinant; it extends to rings, as well. Here, we think of the determinant as a function of its columns, so as a function $(A^n)^n \to A$.

PROPOSITION 3.3.2. *There is a unique function $f\colon (A^n)^n \to A$ that is*

1. ***multilinear**: For any $\mathbf{a}_1, \dots, \mathbf{a}_n, \mathbf{b} \in A^n$ and $a'_i \in A$, we have*

$$f(\mathbf{a}_1, \dots, \mathbf{a}_{i-1}, a\mathbf{a}_i + \mathbf{a}'_i, \mathbf{a}_{i+1}, \dots, \mathbf{a}_n) = af(\mathbf{a}_1, \dots, \mathbf{a}_n) + f(\mathbf{a}_1, \dots, \mathbf{a}_{i-1}, \mathbf{a}'_i, \mathbf{a}_{i+1}, \dots, \mathbf{a}_n);$$

2. ***alternating**: If $\mathbf{a}_i = \mathbf{a}_j$ and $i \neq j$, then $f(\mathbf{a}_1, \dots, \mathbf{a}_n) = 0$; and*
3. *satisfies $f(e_1, \dots, e_n) = 1$, where $\{e_i\}$ are the standard basis of A^n.*

Proof sketch. You can first check that Definition 3.3.1 satisfies all three properties. To prove that such a function is unique, it suffices to show that the value of $f(\mathbf{a}_1, \ldots, \mathbf{a}_n)$ is completely determined by the three axioms in the theorem. By multilinearity, we have

$$f(\mathbf{a}_1, \ldots, \mathbf{a}_n) = \sum_{r \in \{1,\ldots,n\}^n} f(a_{r_1,1} e_{r_1}, \ldots, a_{r_n,n} e_{r_n}) = \sum_{r \in \{1,\ldots,n\}^n} f(e_{r_1}, \ldots, e_{r_n}) \prod_{i=1}^{n} a_{r_i,i}.$$

So it suffices to show that the value of f is determined when each column consists of exactly one 1 with the remaining entries 0. If $r_i = r_j$ for some $i \neq j$, then the value is 0 since f is alternating. Otherwise, r is a permutation. A consequence of f being alternating is that if the position of two columns are switched, then the value of f is negated. If r is a permutation, then the columns of (e_{r_i}) may be switched to create the identity matrix $I_n = (e_i)$, and the number of switches has the same parity as the sign of the permutation r. So $f(e_{r_i}) = (-1)^{\operatorname{sgn}(r)}$. □

This makes it pretty easy to show that the determinant is multiplicative:

PROPOSITION 3.3.3. *If* □ *and* ■ *are any two matrices in* $A^{n\times n}$,[4] *then* $\det(\square\,\blacksquare) = \det(\square)\det(\blacksquare)$.

Proof sketch. If we fix ■, this is a multilinear and alternating map $(A^n)^n \to A$ that takes the value $\det(\blacksquare)$ at the identity. So $\det(\square\,\blacksquare) = \det(\square)\det(\blacksquare)$. □

You may be more familiar with the definition of a determinant in terms of cofactor expansion: If □ is a matrix, the (i,j)th *minor* of □, denoted $\square_{i,j}$, is the determinant of the matrix obtained from □ by deleting the ith row and jth column. The (i,j)th *cofactor* of □ is $(-1)^{i+j}\square_{i,j}$. In these terms, the determinant of $\square = (a_{i,j})$ is

$$\det(\square) = \sum_{j=1}^{n} (-1)^{i+j} a_{i,j} \square_{i,j}$$

for any $1 \leq i \leq n$. You can check, if you wish, that this definition is independent of the choice of i and satisfies the three axioms of Proposition 3.3.2, so it is in fact the same function as the one in Definition 3.3.1. Cofactors are a bit tedious to calculate by hand, but they're useful theoretically:

DEFINITION 3.3.4. The *cofactor matrix* of $\square \in A^{n\times n}$ is the $n \times n$ matrix whose (i,j)th entry is the cofactor $(-1)^{i+j}\square_{i,j}$. The *adjugate* of □, denoted $\operatorname{adj}(\square)$, is the transpose of its cofactor matrix.

Why is this useful? Well:

PROPOSITION 3.3.5. *For any* $\square \in A^{n\times n}$, *we have*

$$\operatorname{adj}(\square)\,\square \;=\; \det(\square)\, I_n \;=\; \square\,\operatorname{adj}(\square).$$

Proof sketch. Expand out: The (i,i)th entry of $\operatorname{adj}(\square)\,\square$ is

$$\sum_{j=1}^{n} (-1)^{i+j} a_{i,j} \square_{i,j} = \det(\square).$$

If $i \neq j$, then the (i,j)th entry of $\operatorname{adj}(\square)\,\square$ is the determinant of the matrix where the ith column of □ is replaced by the jth column; since this matrix has two identical columnns, its determinant is 0. The reasoning for $\square\,\operatorname{adj}(\square)$ is the same, but for rows instead of columns. □

This has a nice corollary that extends what we know from linear algebra.

PROPOSITION 3.3.6. *A matrix* $\square \in A^{n\times n}$ *has an inverse if and only if* $\det(\square)$ *is a unit in* A.

Proof. (⇐) If $\det(\square) = u \in A^\times$, then $u^{-1}\operatorname{adj}(\square)$ is an inverse to □.

(⇒) If □ has an inverse, then $1 = \det(\square\,\square^{-1}) = \det(\square)\det(\square^{-1})$. □

[4] What? I can't use M or N or A or B, and I'm not going to use something funky like Z. So □ it is.

3.4. BACK TO FINITELY GENERATED MODULES

PROPOSITION 3.4.1. *Let M be a finitely generated A-module and $\mathfrak{a}$ an ideal of A. Every endomorphism $\varphi\colon M \to M$ that satisfies $\varphi(M) \subseteq \mathfrak{a}M$ satisfies an equation of the form*

$$\varphi^n + a_1\varphi^{n-1} + \cdots + a_{n-1}\varphi + a_n = 0$$

where each $a_i \in \mathfrak{a}$ and n is bounded above by the size of a generating set.

It's good to view this as a sort of Cayley-Hamilton–type theorem (which says that any linear transformation from a vector space to itself is a "root" of its characteristic polynomial). In fact, this theorem (which is stronger than Proposition 3.4.1) is true for all commutative rings. It's harder to prove and we won't need it, so these notes don't have a proof. But Wikipedia has a collection of proofs that work for all commutative rings which you can browse if you're interested.[5]

Proof of Proposition 3.4.1. Suppose that $x_1, \ldots, x_n$ generate M and write

$$\varphi(x_i) = \sum_{j=1}^{n} a_{i,j}x_j$$

with each $a_{i,j} \in \mathfrak{a}$ (which is possible since $\mathfrak{a}\{x_1, \ldots, x_n\} = \mathfrak{a}M$). Let $\mathbf{x} = (x_1, \ldots, x_n)$ and $\square = (a_{i,j}) \in A^{n\times n}$. We can rewrite the values of φ as a matrix equality:$\varphi I_n\mathbf{x} = \square\,\mathbf{x}$, or

$$(\varphi I_n - \square)\,\mathbf{x} = 0.$$

Multiplying both sides of the equation by the adjugate of $\varphi I_n - \square$, we find that $\det(\varphi I_n - \square)x_i = 0$ for each i; in other words, $\det(\varphi I_n - \square)$ is the zero map. Expressing it as a polynomial finishes the proof. □

Here are some consequences.

COROLLARY 3.4.2. *If M is a finitely generated A-module and $\mathfrak{a}$ is an ideal such that $\mathfrak{a}M = M$, then there is an element $a \in \mathfrak{a}$ such that $(1 + a)M = 0$.*

Proof. Applying Proposition 3.4.1 with φ as the identity map, we get an element $x = 1+a_1+\cdots a_n$ such that $xM = 0$. So take $a = x - 1$. □

COROLLARY 3.4.3. *If M is a finitely generated A-module and $\varphi\colon M \to M$ is surjective, then φ is an isomorphism.*

Proof. We can consider M as an $A[x]$-module where $x \cdot m = \varphi(m)$. Since φ is surjective, we have $(x)M = M$, so we can apply Corollary 3.4.2 to get a polynomial $f(x) \in A[x]$ such that $\big(1 + xf(x)\big)M = 0$. If $m \in \ker\varphi$, then $xf(x) \cdot m = 0$, so

$$m = \big(1 + xf(x)\big) \cdot m = 0,$$

which shows that φ is injective. □

COROLLARY 3.4.4. *If $A^m \cong A^n$, then $m = n$.*

Proof. Suppose that $m \geq n$ and consider an isomorphism $\psi\colon A^n \to A^m$ and the surjection $\varphi\colon A^m \to A^n$ mapping e_i in A^m to e^i in A^n if $i \leq m$ and to 0 otherwise. The composite $\psi \circ \varphi$ is a surjective endomorphism of A^m; by Corollary 3.4.3, the composite map is an isomorphism. Since $\psi \circ \varphi$ and ψ are both bijections, φ must be, too. And the map that we chose is only surjective if $m = n$. □

[5] They also have this charming proof of the Cayley-Hamilton theorem for matrices over $\mathbb{C}$: The set of matrices in $\mathbb{C}^{n\times n}$ which have no repeated eigenvalues is dense in $\mathbb{C}^{n\times n}$, so the set of diagonalizable matrices (which is a superset) is also dense in $\mathbb{C}^{n\times n}$. If $p(x)$ is the characteristic polynomial of a diagonal matrix $\square$, it's easy to check that $p(\square)v = p(\lambda)v = 0$ for any eigenvector with eigenvalue λ. There is a basis of eigenvectors of any diagonalizable matrix, so $p(\square)$ must be the zero matrix. And since the map $\square \mapsto p(\square)$ is a continuous, $p(\square) = 0$ for every matrix.

Phew! It sure would be a nasty time if A^m and A^n could be isomorphic when $m \neq n$.

I'm told the next theorem is very important, though it's not clear why.

THEOREM 3.4.5 (Nakayama's lemma). *If M is a finitely generated A-module and $\mathfrak{a} \subseteq \mathfrak{R}_A$ such that $\mathfrak{a}M = M$, then M is the zero module.*

Proof. The element $1 + a$ from Corollary 3.4.2 is a unit by Proposition 1.5.10, so $M = (1+a)M = 0$. □

In fact, you don't really need all this machinery to prove Nakayama's lemma:

An alternate proof of Theorem 3.4.5. Suppose that $u_1, \ldots, u_n$ form a mimimal set of generators for M; since $u_n \in M = \mathfrak{a}M$, we have $u_n = a_1u_1 + \cdots + a_nu_n$ for some $a_i \in \mathfrak{a}$. Subtracting a_nu_n from both sides, we get

$$(1 - a_n)u_n = a_1u_1 + \cdots + a_nu_n.$$

But $1 - a_n$ is a unit, so u_n is generated by $u_1, \ldots, u_{n-1}$, a contradiction. □

COROLLARY 3.4.6. *If M is a finitely generated A-module, N a submodule of M, and $\mathfrak{a} \subseteq \mathfrak{R}_A$ an ideal of A such that $M = \mathfrak{a}M + N$, then $M = N$.*

The benefit of this corollary is that it provides a way to check if a set generates a module. Setting $N = (x_1, \ldots, x_n)$ and $\mathfrak{a} = \mathfrak{R}_A$, we need only prove that $M = \mathfrak{R}_A M + (x_1, \ldots, x_n)$ to show that $M = N$. Or, in other words, it suffices to show that $\{x_i + \mathfrak{a}M\}$ generate $M/\mathfrak{a}M$.

Proof of Corollary 3.4.6. In the module M/N, we have $M/N = (\mathfrak{a}M + N)/N = \mathfrak{a}(M/N)$; by Nakayama's lemma, $M/N = 0$, so $N = M$. □

This tool is made even nicer if A is a local ring; then $A/\mathfrak{m}A$ is a field and $M/\mathfrak{m}M$ is a finite-dimensional vector space over $A/\mathfrak{m}A$. If the set $\{x_i + \mathfrak{m}M\}$ forms a basis of this vector space, then the set $\{x_i\}$ generates M.

3.5. WHAT ARE EXACT SEQUENCES, EXACTLY?

I'm glad you asked.

DEFINITION 3.5.1. A sequence of modules and homomorphisms

$$\cdots \xrightarrow{f_{i-1}} M_i \xrightarrow{f_i} M_{i+1} \xrightarrow{f_{i+1}} \cdots$$

(which may be either finite or infinite) is called *exact* at M_i if $\operatorname{im}(f_{i-1}) = \ker(f_i)$. We call the whole sequence exact if it is exact at every x_i.

What a funky definition. The thing is, it turns out to be quite useful. We can gesture at some reasons for that by showing how it can package up some notions we already have. For example: The sequence

$$N \xrightarrow{f} M \to 0$$

is exact if and only if f is surjective. (We don't need to specify the map $M \to 0$, since there's only one.) And the sequence

$$0 \to N \xrightarrow{f} M$$

is exact if and only if f is injective. (Again there's only one homomorphism $0 \to N$.)

EXERCISE 3.5.2. Verify these last two statements about sur- and injectivity.

Also, note that a sequence that's exact at M_i satisfies $f_i \circ f_{i-1} = 0$; but this is not a sufficient condition:

EXERCISE 3.5.3. Suppose that $f_1 \colon M_1 \to M_2$ and $f_2 \colon M_2 \to M_3$. Prove that $f_2 \circ f_1 = 0$ if and only if $\operatorname{im}(f_1) \subseteq \ker(f_2)$.

Now let's move up: a *short exact sequence* has the form

$$0 \to M_1 \xrightarrow{f} M_2 \xrightarrow{g} M_2 \to 0.$$

In this case, f is injective, g is surjective, and $\operatorname{im}(f) = \ker(g)$, so $M_2/\operatorname{im}(f) \cong M_3$. Similarly, if $N \subseteq M$, then the sequence

$$0 \to N \hookrightarrow M \to M/N \to 0$$

is always exact (when we choose the usual projection map $M \to M/N$).

EXERCISE 3.5.4. Verify this.

In fact, any long exact sequence

$$\cdots \xrightarrow{f_{i-1}} M_i \xrightarrow{f_i} M_{i+1} \xrightarrow{f_{i+1}} \cdots$$

can be broken into short exact sequences of the form

$$0 \to \operatorname{im}(f_{i-1}) \xrightarrow{f_{i-1}} M_i \xrightarrow{f_i} \operatorname{im}(f_i) \to 0.$$

Moreover, if you have a set of short exact sequences

$$0 \to N_{i-1} \to M_i \to N_i \to 0,$$

then they can be stitched together into a long exact sequence

$$\cdots \to M_i \to M_{i+1} \to \cdots.$$

So that's an intro to short exact sequences. Now that you've graduated from training, you can start working with two-dimensional diagrams. There are a lot of fun things you can prove with them. For example:

PROPOSITION 3.5.5. *Suppose that*

$$\begin{array}{ccccccccc} 0 & \longrightarrow & M_1 & \xrightarrow{f_1} & M_2 & \xrightarrow{f_2} & M_3 & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle h_1} & & \downarrow{\scriptstyle h_2} & & \downarrow{\scriptstyle h_3} & & \\ 0 & \longrightarrow & N_1 & \xrightarrow[g_1]{} & N_2 & \xrightarrow[g_2]{} & N_3 & \longrightarrow & 0 \end{array}$$

is a commutative diagram of A-modules with both rows exact.[6] *Then there is an exact sequence*

$$0 \to \ker(h_1) \xrightarrow{\bar{f}_1} \ker(h_2) \xrightarrow{\bar{f}_2} \ker(f_3) \xrightarrow{u} \operatorname{coker}(h_1) \xrightarrow{\bar{g}_1} \operatorname{coker}(h_2) \xrightarrow{\bar{g}_2} \operatorname{coker}(h_3) \to 0$$

Strictly speaking, g_1 doesn't map from $\operatorname{coker}(h_1)$ to $\operatorname{coker}(h_2)$, but it does induce a map $\bar{g}_1 \colon x + \operatorname{im}(h_1) \mapsto g_1(x) + \operatorname{im}(h_2)$. This doesn't look well-defined, but it is: if $y \in \operatorname{im}(h_1)$, then there's an element $z \in M_1$ such that $h_1(z) = y$. By the commutativity of the diagram,

$$g_1(y) = g_1 \circ h_1(y) = h_2 \circ f_1(z) \in \operatorname{im}(h_2).$$

So if $x_1 - x_2 \in \operatorname{im}(h_1)$, then $g_1(x_1) + \operatorname{im}(h_2) = g_1(x_2) + \operatorname{im}(h_2)$ and the map $\bar{g}_1$ is well-defined. The same, of course, holds for $\bar{g}_2$.

This is the most basic form of a *diagram chase*, the fun-filled type of argument that pervades this area of math, consisting of running around the diagram to prove what you want to prove. Here's a chance for you to try it out yourself.

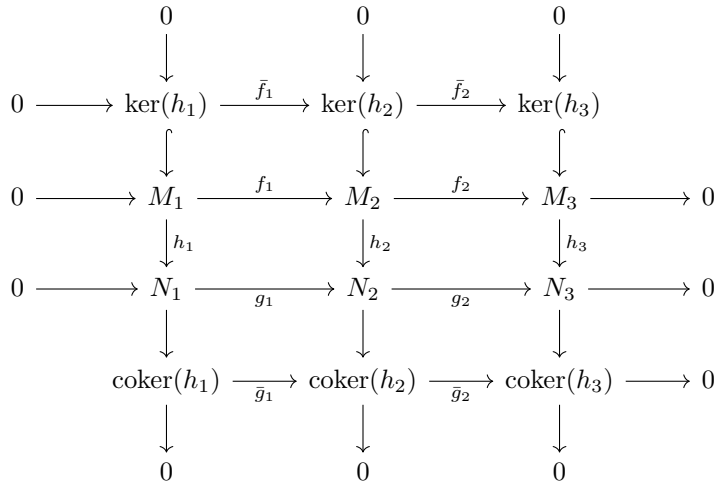
The maps $\bar{f}_1$ and $\bar{f}_2$ are just the restrictions of f_1 and f_2 to the kernels. It's not clear that these are well-defined, either.

[6] A diagram is called *commutative* if any way of traversing it gives the same result. In this case, that means that $g_1 \circ h_1 = h_2 \circ f_1$ and $g_2 \circ h_2 = h_3 \circ f_2$.

EXERCISE 3.5.6. Prove that, in the diagram of Proposition 3.5.5, we have $f_1(x) \in \ker(h_2)$ whenever $x \in \ker(h_1)$.

Let's do some more chasing.

LEMMA 3.5.7. *Under the assumptions of Proposition 3.5.5, the following diagram is commutative and exact in every row and column.*



Boy, that's sure a big 'un, isn't it? When you go through the proof of this lemma, it's useful to trace the path that the argument follows on the diagram.

Most of a proof of Lemma 3.5.7. That the columns are exact is easy to check. So you can do that. For the rest of the proof, "exactness" will refer to exactness in the row sequence.

Since f_1 is injective, so is $\bar{f}_1$, so the diagram is exact at $\ker(h_1)$. To show that it is exact at $\ker(h_2)$, we need to show that any $x \in \ker(\bar{f}_2)$ has a preimage in $\ker(h_2)$; that $\operatorname{im}(\bar{f}_1) \subseteq \ker(\bar{f}_2)$ follows from the fact that $\bar{f}_i$ is the restriction of f_i. So pick an element $x \in \ker(\bar{f}_2)$. Since $x \in \ker(f_2)$, there is an element $y \in M_1$ such that $f_1(y) = x$; then

$$g_1 \circ h_1(y) = h_2 \circ f_1(y) = h_2(x) = 0,$$

so $h_1(y) \in \ker(g_1)$. But g_1 is an injection, so $h_1(y) = 0$; in other words, $y \in \ker(h_1)$, so $\bar{f}_1(y) = x$. So it's exact at $\ker(h_2)$.

You can verify that the diagram is exact at $\operatorname{coker}(h_3)$. For the same reason as before, $\operatorname{im}(\bar{g}_1) \subseteq \ker(\bar{g}_2)$. To prove the opposite inclusion, let $x \in \ker(\bar{g}_2)$ and let p_i denote the map $N_i \to \operatorname{coker}(h_i)$. There is a $y \in N_2$ such $p_2(y) = x$, and

$$p_3 \circ g_2(y) = \bar{g}_2 \circ p_2(y) = 0,$$

so $g_2(y) \in \ker(p_3)$. So there is a $z \in M_3$ with $h_3(z) = g_2(y)$; since f_2 is surjective, there is an element $w \in M_2$ with $f_2(w) = z$. By commutativity,

$$g_2 \circ h_2(w) = h_3 \circ f_2(w) = g_2(y).$$

Therefore $y - h_2(w) \in \ker(g_2)$, so there is an element $u \in N_1$ such that $g_1(y) = y - h_2(w)$; for this element,

$$g_1 \circ p_1(u) = p_2 \circ g_1(u) = p_2\big(y - h_2(w)\big) = x + 0$$

since the column is exact. So $x \in \operatorname{im}(\bar{g}_1)$, which finishes the proof. □

The full proof of Proposition 3.5.5 consists of defining the map $\ker(h_3) \to \operatorname{coker}(h_1)$ and then doing more diagram chasing like we did in the lemma. You can see a fully detailed proof here; alternatively, see this proof, which appeared in the 1980 American romantic comedy-drama film *It's My Turn*.

It's possible to go quite far in this vein, building bigger diagrams with longer sequences and proving theorems about them. But instead of going down a rabbit hole, we'll develop what we need as we need it.

3.6. MODULES OF HOMOMORPHISMS

Let's get back on the Building Bigger Modules bus. If we have two A-modules M and N, the set of A-linear maps between them, denoted $\operatorname{Hom}_A(M, N)$, is itself an A-module, with operations defined as you expect:

$$(f+g)(x) = f(x) + g(x)$$
$$(a \cdot f)(x) = a \cdot f(x).$$

That's that for building the bigger module. But we can also view $\operatorname{Hom}_A(-, N)$, for each fixed N, as a function from the colleciton of A-modules to the collection of A-modules sending $M \mapsto \operatorname{Hom}_A(M, N)$. And $\operatorname{Hom}_A(M, -)$ is another function $N \mapsto \operatorname{Hom}_A(M, N)$. Each map $f\colon N_1 \to N_2$ induces the map $\bar{f}\colon \operatorname{Hom}_A(M, N_1) \to \operatorname{Hom}_A(M, N_2)$ given by $\bar{f}(u) = f \circ u$; similarly, any map $g\colon M_1 \to M_2$ induces a map $\bar{g}\colon \operatorname{Hom}_A(M_2, N) \to \operatorname{Hom}_A(M_1, N)$ by $\bar{g}(u) = u \circ g$. Under these rules, $\operatorname{Hom}_A(-, N)$ and $\operatorname{Hom}_A(M, -)$ are so-called *functors*: They map modules to modules and maps to maps in a way that distributes over composition. (If $f_1\colon N_1 \to N_2$ and $f_2\colon N_2 \to N_3$, then $\overline{f_2 \circ f_1} = \bar{f}_1 \circ \bar{f}_2$. The functor $\operatorname{Hom}_A(-, N)$ is called *contravariant* because $\overline{g_2 \circ g_1} = \bar{g}_1 \circ \bar{g}_2$.)

Anyway, here's the point: These maps are *left exact*.

PROPOSITION 3.6.1. *The sequence*

$$0 \to N_1 \xrightarrow{f} N_2 \xrightarrow{g} N_3$$

is exact if and only if the sequence

$$0 \to \operatorname{Hom}(M, N_1) \xrightarrow{\bar{f}} \operatorname{Hom}(M, N_2) \xrightarrow{\bar{g}} \operatorname{Hom}(M, N_3)$$

is exact for every A-module M.

PROPOSITION 3.6.2. *The sequence*

$$M_1 \xrightarrow{f} M_2 \xrightarrow{g} M_3 \to 0$$

is exact if and only if the sequence

$$0 \to \operatorname{Hom}(M_3, N) \xrightarrow{\bar{g}} \operatorname{Hom}(M_2, N) \xrightarrow{\bar{f}} \operatorname{Hom}(M_1, N)$$

is exact for every A-module N.

I'm running a bit behind on writing these notes, so I'll leave it to you to prove these statement. Atiyah and Macdonald say that "all four parts ... are easy exercises," so how hard could it be? (Actually though, it does mostly consist of unravelling definitions.)

Why is $\operatorname{Hom}(-, N)$ only left exact, even if f is injective? Well, $\bar{f}$ is surjective exactly when every map $M_1 \to N$ extends to a map $M_2 \to N$, and this doesn't always happen. Consider the exact sequence of $\mathbb{Z}$-modules

$$0 \to \mathbb{Z} \xrightarrow{f} \mathbb{Q} \xrightarrow{g} \mathbb{Q}/\mathbb{Z} \to 0.$$

The induced sequence is

$$0 \to \mathrm{Hom}(\mathbb{Q}/\mathbb{Z}, \mathbb{Z}) \xrightarrow{\bar{g}} \mathrm{Hom}(\mathbb{Q}, \mathbb{Z}) \xrightarrow{\bar{f}} \mathrm{Hom}(\mathbb{Z}, \mathbb{Z}) \to 0.$$

The only map in $\mathrm{Hom}(\mathbb{Q}, \mathbb{Z})$ is the zero map! Why? If $\varphi\colon \mathbb{Q} \to \mathbb{Z}$ is a group homomorphism and $\varphi(1) = n$, then

$$n = \varphi(1) = \varphi\left((n+1)\frac{1}{n+1}\right) = (n+1)\varphi\left(\frac{1}{n+1}\right),$$

so $\varphi\big(1/(n+1)\big) \notin \mathbb{Z}$. But since $\mathrm{Hom}(\mathbb{Z}, \mathbb{Z})$ contains nonzero maps (for example, the identity), the map $\bar{f}$ is not surjective.

EXERCISE 3.6.3. Use this exact sequence to give an example that shows why $\bar{g}$ in Proposition 3.6.2 is not necessarily surjective, even if g is.

3.7. TENSOR PRODUCTS

We can of course build bigger modules by taking the direct product or direct sum. These are ways to build bigger modules by setting two modules next to each other. But sometimes we want to intertwine the workings of the modules, so that the structure isn't just linear, but bi-linear. Here's what this means.

Start with two A-modules M and N, and let I be the submodule of $A^{(M\times N)} = \bigoplus_{x\in M\times N} A$ generated by the elements of the form

$$\begin{gathered}(ax, y) - a\cdot(x, y)\\ (x, ay) - a\cdot(x, y)\\ \big((x+y), z\big) - (x, z) - (y, z)\\ \big(x, (y+z)\big) - (x, y) - (x, z)\end{gathered}$$

The *tensor product* of M and N, denoted $M \otimes N$, is defined as the quotient module $A^{(M\times N)}/I$. Intuitively, we think of $M \otimes N$ as the largest bilinear module built from the pairs $(x, y) \in M \times N$. The image of (x, y) in $M \otimes N$ is denoted $x \otimes y$. By definition, we have

$$\begin{gathered}(ax) \otimes y = a(x \otimes y) = x \otimes (ay)\\ (x+y) \otimes z = x \otimes z + y \otimes z\\ x \otimes (y+z) = x \otimes y + x \otimes z.\end{gathered}$$

⚠ The elements $x \otimes y$ generate $M \otimes N$. However, unlike in the module $M \oplus N$, not every element has the form $x \otimes y$. It may not be possible, for example, to reduce $x \otimes y + z \otimes w$ to any simpler expression.

⚠ No, really, this is important. Most functions on $M \otimes N$ are defined on the generators, and there will come a time when you need to analyze one of these functions. And you will be misled unless you remember: *Elements of $M \otimes N$ are **finite linear combinations** of elements of the form $x \otimes y$.*

EXERCISE 3.7.1. Find the identity and zero element of $M \otimes N$. Prove that $x \otimes 0 = 0 \otimes 0$ and $0 \otimes y = 0 \otimes 0$ for every $x \in M$ and $y \in N$.

This is all very nice. The definition, however, is a bit difficult to work with directly. It's not even clear that sometimes, the tensor product collapses to the zero ring!

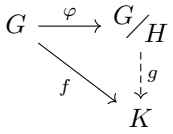
EXAMPLE 3.7.2. Set $A = \mathbb{Z}/p\mathbb{Z} \otimes \mathbb{Z}/q\mathbb{Z}$ and choose any generator $x \otimes y \in A$. We have

$$x \otimes y = p\big(x \otimes (p^{-1}y)\big) = (px) \otimes (p^{-1}y) = 0 \otimes (p^{-1}y).$$

And $0 \otimes z = 0 \otimes 0$ for any z, so $x \otimes y = 0 \otimes 0$ for every $x \in M$ and $y \in N$. So A is actually the zero ring. ◊

Also, the generating elements aren't independent: $(ax) \otimes y = x \otimes (ay)$. This is why it's useful to characterize the tensor product in another way. Before we do that, it will help to see an example with an operation that's more familiar.

Suppose G is an abelian group and H is a subgroup of G; let $\varphi\colon G \to G/H$ be the map from an element to its coset. If K is another group and $f\colon G \to K$ is a group homomorphism such that $f(H) = 0$, then f induces a well-defined homomorphism $g\colon G/H \to K$ given by $g(x + H) = f(x)$. Moreover, this map is unique, in the sense that if the following diagram commutes, then g must be defined in this way:

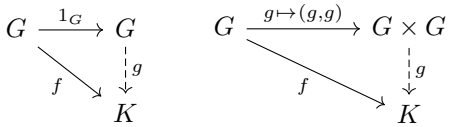


We say that the pair $(G/H, \varphi)$ is *universal* among all pairs (K, f) with the property that $f(H) = 0$. This means exactly what we said before: That for any pair (K, f) with $f(H) = 0$, there is a unique map $g\colon G/H \to K$ such that $g \circ \varphi = f$.

The somewhat surprising thing is that the group G/H is actually *characterized* by this property: If L is any group that has this same property, then $L \cong G/H$. This fact about how this group relates to other groups is enough to tell you everything about it.

To see why this is, suppose that (L, ψ) is also a universal pair. Then we can apply the universal property both ways to get maps $g_1\colon G/H \to L$ and $g_2\colon L \to G/H$ such that $g_1 \circ \varphi = \psi$ and $g_2 \circ \psi = \varphi$. Combining these, we get that $g_2 \circ g_1 \circ \varphi = \varphi$ and $g_1 \circ g_2 \circ \psi = \psi$. Now comes the tricky bit: Apply the universal property with $(K, f) = (G/H, \varphi)$. It says that there's a unique map $g\colon G/H \to G/H$ such that $g \circ \varphi = \varphi$. Certainly the identity map on G/H satisfies this criterion; but so does $g_2 \circ g_1$! Since this map is unique, $g_2 \circ g_1$ must be the identity. And similar reasoning for L instead of G/H shows that $g_1 \circ g_2$ is the identity. So $g_1 = g_2^{-1}$ are isomorphisms, and $G/H \cong L$. In fact, the uniqueness part of the universal property guarantees that g_1 is the *unique* isomorphism from G/H to L.

As this argument shows, the fact that the induced map is unique plays a big part. If we get rid of that condition, then we get rid of the isomorphism property, as well. For example, in both of the following diagrams, there always exists a map g that makes it commute:

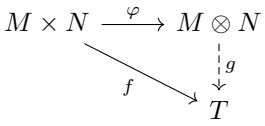


Importantly, these induced maps are not unique. And correspondingly, the groups in the upper right corner are not isomorphic.

So this gives an idea of universal properties. It turns out the tensor product has one.

THEOREM 3.7.3. *Let M and N be A-modules and define $\varphi\colon M \times N \to M \otimes_A N$ by $\varphi(x, y) = x \otimes y$. The pair $(M \otimes N, f)$ is universal among all pairs $(T, f\colon M \times N \to T)$ in which f is an A-bilinear map. In other words, if $f\colon M \times N \to T$ is an A-bilinear map, then there exists a unique A-linear map $g\colon M \otimes N \to T$ such that $g \circ \varphi = f$.*

We can sum this up in a diagram very similar to the quotient diagram:



Proof of Theorem 3.7.3. If $f\colon M \times N \to T$, then f extends to a map $f'\colon A^{(M\times N)} \to T$ by linearity. Moreover, since f is bilinear, it vanishes on all of the generators of the ideal I (see the beginning of this section). Recalling that $M^{(M\times N)}/I = M \otimes N$, this means that f' induces a well-defined homomorphism $g\colon M \otimes N \to T$ such that $g(x \otimes y) = f(x, y)$. Moreover, any $g\colon M \otimes N \to T$ for which $g \circ \varphi = f$ must satisfy this relation, which in fact determines g; so g is unique. □

EXERCISE 3.7.4. Prove that if (T, f) that is universal in the sense of Theorem 3.7.3, then $T \cong M \otimes N$. (HINT: This is very similar to the argument for groups.)

This way of looking at things is very useful; in practice, often more useful than remembering the actual construction. For example:

PROPOSITION 3.7.5. *If M is an A-module, then $A \otimes_A M \cong M$.*

Proof. The map $A \times M \to M$ defined by $(a, m) \mapsto am$ is bilinear, so (by Theorem 3.7.3) it induces a unique map $f\colon A \otimes M \to M$. And we also define $g\colon M \to A \otimes M$ by $g(m) = 1 \otimes m$. Then $f(g(m)) = m$ and $g(f(a \otimes m)) = 1 \otimes (am) = a \otimes m$. Therefore f and g are inverses, so $A \otimes M \cong M$. □

The proof hides something more general: If we want to define a map $M \otimes N \to P$ by specifying its action on the generators $x \otimes y$, we might worry that it's not well-defined, since $x \otimes y$ is really an equivalence class. However, Theorem 3.7.3 guarantees us that as long as the map is bilinear in the coordinates, it's well defined.

There are several more common isomorphisms to keep in mind:

PROPOSITION 3.7.6. *If M, N, and P are A-modules, then*

1. $M \otimes N \cong N \otimes M$ *(commutativity)*
2. $(M \otimes N) \otimes P \cong M \otimes (N \otimes P)$ *(associativity)*
3. $(M \oplus N) \otimes P \cong (M \otimes P) \oplus (N \otimes P)$ *(distributivity)*

Partial proof. Of these three statements, the first two are eminently believable, so I'll only prove the last; the others are proved similarly. We define the function $f\colon (M \oplus N) \otimes P \to (M \otimes P) \oplus (N \otimes P)$ by $f\big((x, y) \otimes z\big) = (x \otimes z, y \otimes z)$; this is well-defined, since it's bilinear in both coordinates. And we can define a map g in the other direction by $(x \otimes z_1, y \otimes z_2) \mapsto (x, 0) \otimes z_1 + (0, y) \otimes z_2$, which is again well-defined. Since $f \circ g$ and $g \circ f$ are both identity mappings (check this!), the two A-modules are isomorphic. □

EXERCISE 3.7.7. Induction shows that the tensor product distributes over any finite direct sum. Extend the previous proof to show that $\bigoplus_{i\in I} M_i \otimes N \cong \bigoplus_{i\in I}(M_i \otimes P)$ for *any* index set I.

The same is not true for direct products, but we'll need some fancier tools to prove it. (The hard part, in fact, is proving that one of the tensor products is not secretly the zero ring.) Example 4.2.11 proves a counterexample.

Some last properties of tensor products:

DEFINITION 3.7.8. Suppose A and B are rings. An abelian group M is called an (A, B)-*bimodule* if it is both an A- and a B-module and these structures are compatible in the sense that $(ax)b = a(xb)$ for every $a \in A$, $b \in B$, and $x \in M$.

PROPOSITION 3.7.9. *If M is an A-module, P is a B-module, and N is an (A, B)-bimodule, then $M \otimes_A N$ and $N \otimes_B P$ are (A, B)-bimodules, and*

$$(M \otimes_A N) \otimes_B P \cong M \otimes_A (N \otimes_B P).$$

If $f\colon A \to B$ is a ring homomorphism, then B can be made into an A-module via the action $a \cdot b = f(a)b$. If M is an A-module, the tensor product $B \otimes_A M$ is of course an A-module, but we can make it into a B-module by the action $b_1 \cdot (b_2 \otimes m) = (b_1 b_2) \otimes m$. This is called *extension of scalars*. If N is a B-module, then it is also an A-module by the rule $a \cdot n = f(a)n$; regarding it as an A-module is called *restriction of scalars*. (We think of the underlying ring as the scalars by which we can multiply, just like the underlying field of a vector space; these operations either extend this scalar set or restrict it—thus the names.)

3.8. EXACTNESS PROPERTIES OF THE TENSOR PRODUCT

As Theorem 3.7.3 indicates, it's useful to think of the tensor product as having an intimate connection with bilinear mappings.

PROPOSITION 3.8.1. *For any three A-modules M, N, and P, there is a "canonical" isomorphism*

$$\mathrm{Hom}_A(M \otimes N, P) \cong \mathrm{Hom}_A\big(M, \mathrm{Hom}_A(N, P)\big).$$

Proof. Any bi-linear map $f\colon M \times N \to P$ can be thought of as a map $\bar{f}\colon M \to \mathrm{Hom}_A(N, P)$, where $\bar{f}(x) = f_x$ and $f_x(y) = f(x, y)$. Conversely, any such map induces a bilinear map $M \times N \to P$. But by Theorem 3.7.3, the set of bilinear maps is in bijection with the set of maps $M \otimes N \to P$; you can check that this bijection is an isomorphism. □

We already know something about the interactions of the Hom operator with exact sequences (recall Propositions 3.6.1 and 3.6.2). It turns out that we can bootstrap that to learn about the exactness of the tensor operation.

PROPOSITION 3.8.2. *If the sequence*

$$M_1 \xrightarrow{f} M_2 \xrightarrow{g} M_3 \to 0$$

is exact, then the sequence

$$M_1 \otimes N \xrightarrow{f \otimes 1_N} M_2 \otimes N \xrightarrow{g \otimes 1_N} M_3 \otimes N \to 0$$

is exact for any A-module N. (1_N is the identity map on N.)

Proof. Let E denote the first exact sequence. By Proposition 3.6.2, the sequence $\mathrm{Hom}\big(E, \mathrm{Hom}(N, P)\big)$ is exact for every A-module P. Using Proposition 3.8.1, this means that $\mathrm{Hom}(E \otimes N, P)$ is exact for every P; but then using Proposition 3.6.2 again shows that $E \otimes N$ is exact. □

So the tensor product is *right exact*. Unfortunately, tensoring with an A-module doesn't keep arbitrary exact sequences exact.

EXAMPLE 3.8.3. Set $A = \mathbb{Z}$ and consider the short exact sequence

$$0 \to \mathbb{Z} \xrightarrow{f} \mathbb{Z} \xrightarrow{g} \mathbb{Z}/2\mathbb{Z} \to 0,$$

where $f(x) = 2x$ and $g(x) = x \bmod 2$. Tensoring with the $\mathbb{Z}$-module $\mathbb{Z}/2\mathbb{Z}$ yields the sequence

$$0 \to \mathbb{Z}/2\mathbb{Z} \xrightarrow{\bar{f}} \mathbb{Z}/2\mathbb{Z} \xrightarrow{\bar{g}} \mathbb{Z}/2\mathbb{Z} \to 0$$

where $\bar{f}$ is the zero map and g is the identity. Since $\bar{f}$ is not injective, this sequence is not exact. ◊

But we really, really would like tensoring to be exact, so we give this property a name:

DEFINITION 3.8.4. An A-module is called *flat* if tensoring with it transforms exact sequences to exact sequences.

There's a good question to ask here: Why call them "flat"? Turns out: no good reason.

EXAMPLE 3.8.5. Any free module $A^{(I)}$ is flat, since $M \otimes A^{(I)} \cong (M \otimes A)^{(I)} \cong M^{(I)}$ (by Exercise 3.7.7). The maps of the resulting sequence are the same as the original, just in each coordinate, so this maps any exact sequence to another exact sequence. ◊

Remember that on page 14 we showed that any long exact sequence can be broken into short exact sequences. This has the consequence that a module is flat if and only if tensoring with it takes short exact sequences to short exact sequences. But Proposition 3.8.2 tells us that it will do this for all but possibly the first nontrivial map: So we only need to check that one.

PROPOSITION 3.8.6. *An A-module N is flat if and only if $f \otimes 1 \colon M_1 \otimes N \to M_2 \otimes N$ is injective whenever $f \colon M_1 \to M_2$ is.*

EXERCISE 3.8.7. Prove that $\bigoplus_{i \in I} M_i$ is flat if and only if each M_i is flat. Then prove that $A[x]$ is a flat A-module (for every ring A).

3.9. ALGEBRAS

You didn't know it, but there's sometimes a very specific meaning to the word algebra.

DEFINITION 3.9.1. An *algebra* over A is a pair (B, f) where B is a commutative ring and $f \colon A \to B$ is a ring homomorphism.

We think of an A-algebra as a sort of enriched module: The map f specifies an action $a \cdot b = f(a)b$, but you also can multiply elements of the algebra together, which is not allowed in a module. In fact, this type of structure is necessary: If you want an A-action on a ring B to be consistent with addition and multiplication in B, then the action must be given by a ring homomorphism $A \to B$.[7]

Just as every abelian group is a $\mathbb{Z}$-module, every commutative ring is a $\mathbb{Z}$-algebra via the same action.

DEFINITION 3.9.2. An A-algebra (B, f) is called *finite* if it is finitely generated as an A-module. It is called a *finitely generated A-algebra* if there are finitely many elements $x_1, \ldots, x_n \in B$ such that every element in B can be written as a polynomial in these elements with coefficients in $f(A)$.

DEFINITION 3.9.3. If (B, f) to (C, g) are A-algebras, an *A-algebra homomorphism* from (B, f) to (C, g) is a map $h \colon B \to C$ that is simultaneously a ring homomorphism and an A-module homomorphism.

EXERCISE 3.9.4. Prove that a ring homomorphism $h \colon B \to C$ is an A-algebra homomorphism from (B, f) to (C, g) if and only if $g = h \circ f$.

EXERCISE 3.9.5. Show that an A-algebra B is finitely generated if and only if there is a surjective A-algebra homomorphism $A[x_1, \ldots, x_n] \to B$ for some $n \in \mathbb{N}$.

[7] The proof of this is not hard; it comes from noting that $a \cdot b = (a \cdot 1_B)b$, so the action of each element is determined by its action on 1_B. You can then define $f(a) = a \cdot 1_B$ and check that this is a homomorphism.

We end by noting that if (B, f) and (C, g) are two A-algebras, then we may take their tensor product to form not just a module, but a ring. Multiplication on $B \otimes_A C$ is defined on the generating elements by

$$(x \otimes y) \cdot (z \otimes w) = (xz) \otimes (yw),$$

which presents no immediate problems, since multiplication is defined in B and C. The only issue is that classic one of whether this operation is actually well-defined, since, as usual, the symbol $x \otimes y$ is an equivalence class. It turns out that it is; one argument for it appears on pp. 30–31 of Atiyah and Macdonald's book.

4. RINGS OF FRACTIONS AND LOCALIZATION

4.1. A BRIEF, APPRECIATIVE LOOK AT $\mathbb{Q}$

$\mathbb{Q}$ is a nice ring. I mean, it's even a field. We want to work with $\mathbb{Z}$ for all sorts of things, but it has the pesky property that basically none of its elements have inverses. Sometimes, though, we can con $\mathbb{Z}$ into telling us something by working first in $\mathbb{Q}$, where most everything is nice, and then pushing what we get back into $\mathbb{Z}$.

For a very slightly fancy example, consider $\mathbb{Z}[x]$. Its irreducible elements are quite hard to pin down, so it seems difficult to say anything about them. Nevertheless, we can actually say something pretty strong:

PROPOSITION 4.1.1. *Any two distinct non-constant irreducible polynomials in $\mathbb{Z}[x]$ generate the whole ring.*

Proof. Here's how the scam works. Take two polynomials $f, g \in \mathbb{Z}[x]$ and consider them instead over the ring $\mathbb{Q}[x]$. This is a very nice ring: a Euclidean domain, in fact. This means that they have a greatest common divisor h which is unique up to multiplication by nonzero elements of $\mathbb{Q}$. So we can multiply by a minimal common denominator to ensure that the coefficients of h are integers. Moreover, since $\deg(f), \deg(g) \geq 1$ and $f \neq g$, the Euclidean algorithm guarantees that $\deg(h)$ is strictly less than both $\deg(f)$ and $\deg(g)$.

Now we go back to $\mathbb{Z}[x]$. The fact that h divides f in $\mathbb{Q}[x]$ translates to the fact that h divides af, for some $a \in \mathbb{N}$, in the ring $\mathbb{Z}[x]$. Suppose that $h\bar{h} = af$. Since (f) is prime, either $h \in (f)$ or $\bar{h} \in (f)$. But $\deg(h) < \deg(f)$ and f is irreducible, so $\bar{h} \in (f)$. Therefore $\deg(\bar{h}) = \deg(f)$, so h must be a constant. But its coefficients have a

Back in the ring $\mathbb{Q}[x]$, this means that f and g generate the entire ring; in particular, we can use the Euclidean algorithm to obtain a polynomial $\bar{g} \in \mathbb{Q}[x]$ such that $f + g\bar{g} = 1$. (This assumes $\deg(g) \leq \deg(f)$; but this or the reverse must be true, so assume this one.) If m is the minimal common denominator of the coefficients of $\bar{g}$, then there is an integer polynomial $\tilde{g} = m\bar{g}$ such that $mf + g\tilde{g} = m \in \mathbb{Z}$. But then m divides $g\tilde{g}$, which is impossible (see the following exercise). So $m = \pm 1$, which means that f and g generate all of $\mathbb{Z}[x]$. □

EXERCISE 4.1.2. Let A be a ring. A polynomial $f = \sum_{i=0}^{n} a_i x^i \in A[x]$ is called *primitive* if $(a_0, \ldots, a_n) = (1)$. Prove that the product of two primitive polynomials is primitive. (The converse is also true, but easy to see.) (HINT: If $(a_0, \ldots, a_n) \neq (1)$, then it is contained in a maximal ideal $\mathfrak{m}$. Consider the projection homomorphism $A[x] \to (A/\mathfrak{m})[x]$ and use the fact that $(A/\mathfrak{m})[x]$ is an integral domain.)

We'd like to be able to use tricks like this for general rings. In the next section, we set up the machinery to do so.

4.2. RINGS OF FRACTIONS

For reasons that I will leave mysterious for the moment, we might sometimes want to only add inverses for specific elements of A instead of all nonzero elements. But that's no problem. Say we want to form a ring on the set

$$\left\{\frac{a}{s} : a \in A \text{ and } s \in S\right\},$$

where S is some subset of A (not necessarily an ideal or subring!). If we want this to be like a regular ring of fractions, then S will have to be multiplicative: we want $\frac{a}{s} \cdot \frac{b}{t} = \frac{ab}{st}$, so st must be an element of S. We'd also like $1_A \in S$ so that we can think of $a \in A$ as the element $\frac{a}{1}$.

If you've seen the algebraic construction of $\mathbb{Q}$ from $\mathbb{Z}$, then the rest of the construction will feel familiar. If not, no worries. Here's how it goes: Consider the set $A \times S$, and define the equivalence relation $(a, s) \sim (b, t)$ if and only if $at = bs$ (which is exactly what we would want if we think of (a, s) as the fraction $\frac{a}{s}$). We write $\frac{a}{s}$ for the equivalence class of (a, s) and equip the set $A \times S/\sim$ with the operations

$$\frac{a}{s} + \frac{b}{t} = \frac{at + bs}{st}$$
$$\frac{a}{s} \cdot \frac{b}{t} = \frac{ab}{st}.$$

Under these operations, $(A \times S)/\sim$ is a ring.

Well, maybe. In fact, $\sim$ is not always an equivalence relation. Take, for example, $A = \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2$. (We'll write the element (a, b, c) as abc to avoid drowning in parentheses.) We have

$$(110, 110) \sim (101, 101)$$
$$(101, 101) \sim (011, 001)$$
$$(110, 110) \nsim (011, 001),$$

which shows that the relation is not transitive.

If A is an integral domain, then $\sim$ is transitive: Suppose $(a, s) \sim (b, t) \sim (c, r)$, meaning that $at = bs$ and $br = ct$. Then

$$rat = rbs = cts,$$

and the cancellation property of integral domains guarantees that $ar = cs$, so $(a, s) \sim (c, r)$.

We could just work with integral domains, but full generality would be nice (and very useful, in the future). So instead we press on and define a new janky equivalence relation $\curlywedge$ on $A \times S$ by

$$(a, s) \curlywedge (b, t) \text{ if } u(at - bs) = 0 \text{ for some } u \in S.$$

If A is an integral domain, then this reduces to the one before, since you can just cancel u (as long as $0 \notin S$). But this time, it works for all rings:

LEMMA 4.2.1. *$\curlywedge$ is an equivalence relation on $A \times S$.*

Proof. Since $1 \in S$, the relation $\curlywedge$ is reflexive; multiplying both sides of the equiation by -1 shows that it's symmetric. To show symmetry, suppose that $(a, s) \curlywedge (b, t) \curlywedge (c, r)$, so that there are $u, v \in S$ such that $u(at - bs) = 0$ and $v(br - ct) = 0$. Multiply the first equation by rv and the second by su, then add; you get $(ar - cs)uvt = 0$. Since S is multiplicative $uvt \in S$, so $(a, s) \curlywedge (c, r)$. □

It is straightforward but tedious to verify that the operations defined above are well-defined on the equivalence classes of $\curlywedge$ (they don't depend on the choice of representative) and that under them, $(A \times S)/\curlywedge$ is a commutative ring with identity.

DEFINITION 4.2.2. If $S \subseteq A$ is a multiplicative set, then the commutative ring $(A \times S)/\sim$ is called the *ring of fractions* of A with respect to S and is denoted $S^{-1}A$.

If A is an integral domain and $S = A \setminus \{0\}$, then $S^{-1}A$ is a field, called the *field of fractions* of A. (So, for instance, $\mathbb{Q}$ is the field of fractions of $\mathbb{Z}$.)

⚠ We can map A into $S^{-1}A$ by $\varphi\colon a \mapsto \frac{a}{1}$. But this map is not always injective! We have $\frac{a}{1} = \frac{0}{1}$ if and only if there is a $u \in S$ such that $ua = 0$. So the map is injective if and only if S contains no zero-divisors.

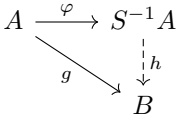
But no matter the choice of S, this map turns $S^{-1}A$ into an A-algebra.

EXERCISE 4.2.3. Show that $S^{-1}A$ is the zero ring if and only if $0 \in S$.

Before the examples, let's look at a universal property of the fraction ring; it will give one very good way to think about $S^{-1}A$ in general.

PROPOSITION 4.2.4. *Let $S \subseteq A$ be a multiplicative set. If $g\colon A \to B$ is a ring homomorphism such that $g(s)$ is a unit in B for every $s \in S$, then there is a unique ring homomorphism $h\colon S^{-1}A \to B$ such that $h \circ \varphi = g$.*

So we get this familiar picture:



Proof of Proposition 4.2.4. If such a map h exists, then $h(a/s) = g(a/s) = g(a)g(s)^{-1}$, so h is uniquely defined. To show that h always exists, we just need to show that the function $h(a/s) = g(a)g(s)^{-1}$ is well-defined on equivalence classes. If $a/s = b/t$, then there is a $u \in S$ such that $(at - bs)u = 0$. Applying g to this expression yields

$$\big(g(a)g(t) - g(b)g(s)\big)g(u) = 0.$$

Since $u \in S$, we know that $g(u)$ is a unit. Rearranging the equality shows that $g(a)g(s)^{-1} = g(b)g(t)^{-1}$. □

What Proposition 4.2.4 indicates is that it's good to think of $S^{-1}A$ what happens when you take A and artificially add some inverses to the elements of S.

As promised, some examples.

EXAMPLE 4.2.5. If S is the set of all non–zero-divisors in A, then $S^{-1}A$ is called the *total quotient ring* of A. It is the largest set S for which the map $\varphi\colon a \mapsto \frac{a}{1}$ is injective. ♢

EXAMPLE 4.2.6. If $S = \{1, n, n^2, n^3, \dots\}$, then $S^{-1}\mathbb{Z}$ is the ring of fractions whose denominator is a power of n. In general, we can set $S = \{1, x, x^2, x^3, \dots\}$; the ring $S^{-1}A$ is the "smallest" localization that makes $x \in A$ a unit. ♢

DEFINITION 4.2.7. Let $\mathfrak{p}$ be a prime ideal of A and $S = A \setminus \mathfrak{p}$.[8] The fraction ring $S^{-1}A$ is called the *localization* of A at $\mathfrak{p}$ and is denoted $A_{\mathfrak{p}}$.

Unlike flat modules, there's a good reason to call this process localization.

PROPOSITION 4.2.8. *Every localization is a local ring.*

[8] This is multiplicatively closed because $\mathfrak{p}$ is prime.

Proof. The set $\mathfrak{p}S^{-1}A$ (or equivalently, the set of fractions $\frac{a}{s}$ with $a \in \mathfrak{p}$) is an ideal in $S^{-1}A$ which we'll denote $\mathfrak{m}$. If $\frac{b}{t} \notin \mathfrak{m}$, then $b \notin \mathfrak{p}$; in other words, $b \in S$, so $\frac{b}{t}$ is a unit in $S^{-1}A$. Conversely, if $\frac{b}{t}$ is a unit, then there is some $\frac{c}{r}$ such that $\frac{b}{t} \cdot \frac{c}{r} = 1 = \frac{1}{1}$. Then, for some element $u \in S$, we have

$$u(bc - sr) = 0.$$

But $u \notin \mathfrak{p}$ and $sr \notin \mathfrak{p}$, which means that $bc \notin \mathfrak{p}$: So $b \notin \mathfrak{p}$, which means $\frac{b}{s} \notin \mathfrak{m}$. So $\mathfrak{m}$ is an ideal that contain exactly the non-units of $S^{-1}A$. By Proposition 1.5.13, A is a local ring and $\mathfrak{m}$ its unique maximal ideal. □

Local rings are nice, which is why localization is nice. Of course, we'll see reasons for this later, but here's one: If M and N are finitely-generated A-modules and A is local, then $M \otimes N$ is the zero ring if and only if $M = 0$ or $N = 0$. For local rings, then, and somewhat reasonably sized modules, we don't need to worry about the tensor product collapsing to nothing. (See exercises 2 and 3 in chapter 2 of Atiyah and Macdonald for an outline of the proof.)

Speaking of the dynamic duo, they describe the connection of localization with algebraic geometry like this:[9]

> Let k be an infinite field and set $A = k[x_1, \ldots, x_n]$; let $\mathfrak{p}$ be any ideal in A. The localization $A_\mathfrak{p}$ is the ring of all rational functions f/g where $g \notin \mathfrak{p}$. If V is the set of all $\mathbf{x} \in k^n$ such that $f(\mathbf{x}) = 0$ whenever $f \in \mathfrak{p}$ (the *variety* defined by $\mathfrak{p}$), then $A_\mathfrak{p}$ can be identified with the ring of all rational functions on k^n which are defined at almost all points of V. This is the prototype of the local rings which arise in algebraic geometry.

You can follow the same string of definitions to define $S^{-1}M$ for *any* A-module M; this is the set of fractions $\frac{x}{s}$ with $x \in M$ and $s \in S$. This new module is also naturally an $S^{-1}A$-module via the operation $\frac{a}{s} \cdot \frac{x}{t} = \frac{a \cdot x}{st}$. In fact, the module $S^{-1}M$ is the same as extending by scalars:

PROPOSITION 4.2.9. *The $S^{-1}A$ modules $S^{-1}M$ and $M \otimes_A S^{-1}A$ are isomorphic.*

(A proof is not too difficult; see Proposition 3.5 here.)

Fixing some $S \subseteq A$, the map $M \mapsto S^{-1}M$ is actually exact. Together, this means that

PROPOSITION 4.2.10. *If S is a multiplicatively closed subset of A, then $S^{-1}A$ is a flat A-module.*

This is enough to provide the counterexample for tensor product distributivity that was promised before.

EXAMPLE 4.2.11. Both $\mathbb{Q}$ and $\mathbb{Z}/p\mathbb{Z}$ are $\mathbb{Z}$-modules. You can check that $\mathbb{Q} \otimes \mathbb{Z}/n\mathbb{Z} = \{0\}$ for any nonzero integer n. On the one hand

$$\prod_p (\mathbb{Q} \otimes_\mathbb{Z} \mathbb{Z}/p\mathbb{Z}) = 0.$$

But on the other, $\mathbb{Q} \otimes_\mathbb{Z} \prod_p \mathbb{Z}/p\mathbb{Z}$ is *not* the zero ring. Here's why: Let $x \in \prod_p \mathbb{Z}/p\mathbb{Z}$ be the element with 1 in every coordinate. This element has infinite order, so

$$\langle x \rangle \otimes_\mathbb{Z} \mathbb{Q} \cong \mathbb{Z} \otimes_\mathbb{Z} \mathbb{Q} \cong \mathbb{Q}.$$

On the other hand, $\mathbb{Q}$ is the field of fractions of $\mathbb{Z}$, so it is a flat $\mathbb{Z}$-module by Proposition 4.2.10. This means that the map $\langle x \rangle \otimes_\mathbb{Z} \mathbb{Q} \to \prod_p \mathbb{Z}/p\mathbb{Z} \otimes_\mathbb{Z} \mathbb{Q}$ induced by the map $\langle x \rangle \hookrightarrow \prod_p \mathbb{Z}/p\mathbb{Z}$ is injective. In particular, $\prod_p \mathbb{Z}/p\mathbb{Z} \otimes_\mathbb{Z} \mathbb{Q}$ is nonempty. ◊

[9] Mostly. I've made some minor modifications for clarity.

It was at this point that I realized the world doesn't need another set of commutative algebra notes and my time was better spent elsewhere. These notes therefore come to an abrupt end exactly here. Ciao!