Commutative Algebra and Algebraic Geometry

Andrew Hubery

Contents

Part I

Commutative Rings

1 Rings and ideals

We recall that an (associative, unital) **ring** R is given by an abelian group under addition, with identity element 0 (called the zero element), together with an associative multiplication, with identity element 1 (called the identity or unit), satisfying the distributivity rule

$$x(y+z) = xy + xz \quad \text{and} \quad (x+y)z = xz + yz \quad \text{for all } x, y, z \in R.$$

A **commutative ring** is a ring R for which the multiplication is commutative

$$xy = yx \quad \text{for all } x, y \in R.$$

In this course, nearly all rings will be commutative, so we will just use the term ring.

Examples. The integers $\mathbb{Z}$ and all fields K provide examples of rings. If R is a ring, then we can form the ring of polynomials over R, denoted $R[X]$ and formal power series, denoted $R[\![X]\!]$. Also, given a family of rings R_i, we can form their direct product $\prod_i R_i$, which is a ring via pointwise addition and multiplication, so $(r_i) + (s_i) = (r_i + s_i)$ and $(r_i)(s_i) := (r_i s_i)$.

If $1 = 0$ in a ring R, then $R = \{0\}$ and we call R the **trivial ring**.

A **ring homomorphism** $f\colon R \to S$ is an additive group homomorphism which preserves the identity

$$f(1_R) = 1_S$$

and respects multiplication

$$f(xy) = f(x)f(y) \quad \text{for all } x, y \in R.$$

In particular, the identity map $\mathrm{id}_R\colon R \to R$ is a ring homomorphism, and if $f\colon R \to S$ and $g\colon S \to T$ are ring homomorphisms, then so is the composition $gf\colon R \to T$. Moreover, composition is associative, so $h(gf) = (hg)f$ whenever this makes sense.

Remark. It follows that there is a category having objects the commutative rings and morphisms the ring homomorphisms.

1.1 Subrings and factor rings

A **subring** $S \leq R$ is an additive subgroup containing the identity element and closed under multiplication

$$1 \in S, \quad \text{and} \quad xy \in S \quad \text{for all } x, y \in S,$$

in which case the inclusion map $S \to R$ is a ring homomorphism.

An **ideal** $I \lhd R$ is an additive subgroup closed under multiplication by elements of R

$$rx \in I \quad \text{for all } r \in R \text{ and } x \in I.$$

In this case there is a unique multiplication on the factor group R/I such that the canonical map $R \twoheadrightarrow R/I$ is a ring homomorphism.

Lemma 1.1. *Let $f\colon R \to S$ be a ring homomorphism. Then the **image** of f*

$$\mathrm{Im}(f) := \{f(x) \in S : x \in R\}$$

*is a subring of S, and the **kernel** of f is an ideal of R,*

$$\mathrm{Ker}(f) := \{x \in R : f(x) = 0\}.$$

Lemma 1.2 (Factor Lemma)**.** *Let $f\colon R \to S$ be a ring homomorphism, and $I \lhd R$ an ideal. If $I \leq \mathrm{Ker}(f)$, then there is a unique ring homomorphism*

$$\bar{f}\colon R/I \to S, \quad r + I \mapsto f(r),$$

satisfying $f = \bar{f}\pi$. Thus f factors through the canonical map π. We call $\bar{f}$ the homomorphism induced by f.

Corollary 1.3 (First Isomorphism Theorem)**.** *A ring homomorphism $f\colon R \to S$ yields an isomorphism*

$$R/\mathrm{Ker}(f) \xrightarrow{\sim} \mathrm{Im}(f), \quad r + \mathrm{Ker}(f) \mapsto f(r).$$

1.2 Algebras

Let k be a ring. A **k-algebra** is a ring R together with a ring homomorphism $i\colon k \to R$, called the **structure map**. In the special case when i is injective, we often identify k with its image, and so view k as a subring of R.

Let R and S be two k-algebras, with structure maps $i_R\colon k \to R$ and $i_S\colon k \to S$. A **k-algebra homomorphism** is a ring homomorphism $f\colon R \to S$ which commutes with the structure maps, so $fi_R = i_S$. In fact, there is a category having objects the k-algebras and morphisms the k-algebra homomorphisms.

Let R be a k-algebra. If $I \lhd R$ is an ideal, then R/I becomes a k-algebra via the composition $k \to R \to R/I$, in which case $R \to R/I$ is a k-algebra homomorphism. Similarly, a k-subalgebra of R is a subring S containing the image of k, in which case the inclusion map $S \to R$ is a k-algebra homomorphism.

For example, let k be a field. The only ideals of k are $\{0\}$ and k, so if R is a non-trivial ring, then every ring homomorphism $k \to R$ is injective. It follows that every k-algebra is a k-vector space and every k-algebra homomorphism is k-linear.

On the other hand, there is a unique ring homomorphism $\mathbb{Z} \to R$ for each ring R. Thus every ring R is a $\mathbb{Z}$-algebra, and every ring homomorphism is necessarily a $\mathbb{Z}$-algebra homomorphism. For this (and other) reasons, it is common always to work in the 'relative' setting by fixing a base ring k and considering only k-algebras.

1.3 Operations on subrings and ideals

Given a family of subalgebras S_i of an algebra R, their intersection is again a subalgebra

$$\bigcap_i S_i := \{x \in R : x \in S_i \text{ for all } i\}.$$

We therefore define for each subset $X \subset R$ the **subalgebra generated** by X, denoted $k[X]$, which is the smallest subalgebra containing X

$$k[X] := \bigcap_{X \subset S} S.$$

The elements of $k[X]$ are k-linear sums of products of elements of X. Warning: this is not to be confused with the polynomial ring.

A subalgebra S of R is called **finitely generated** if $S = k[X]$ for some finite set X.[1]

Similarly, given a family of ideals I_i in a ring R, their intersection is again an ideal

$$\bigcap_i I_i := \{x \in R : x \in I_i \text{ for all } i\}.$$

We therefore define for each subset $X \subset R$ the **ideal generated** by X, denoted (X), which is the smallest ideal containing X

$$(X) := \bigcap_{X \subset I} I.$$

The elements of (X) are those which can be expressed in the form $r_1x_1 + \cdots + r_nx_n$ for some $r_i \in R$ and $x_i \in X$. An ideal I is called **finitely generated** if $I = (x_1, \ldots, x_n)$, and **principal** if $I = (x)$.

For example, the **unit ideal** $R = (1)$ and the **trivial ideal**, or **zero ideal**, $0 = (0)$ are both principal. An ideal I is called **proper** provided $I \neq R$, or equivalently $R/I \neq 0$.

We define the sum of ideals I_i to be the ideal generated by the union $\bigcup_i I_i$, so

$$\sum_i I_i = \Big\{ \sum_i x_i : x_i \in I_i \text{ almost all zero}^{2} \Big\}.$$

Similarly we define the product of two ideals to be the ideal generated by all products xy with $x \in I$ and $y \in J$, so

$$IJ = \{x_1y_1 + \cdots + x_ny_n : n \geq 1,\ x_i \in I,\ y_i \in J\}.$$

The product of ideals is associative, so $I(JK) = (IJ)K$. We can therefore define finite products inductively.

---

[1] Determining whether or not an algebra is finitely generated is a difficult problem. For example, Hilbert's Fourteenth Problem asks whether, given a subgroup $G \leq \mathrm{GL}_n(k)$, the algebra of invariants

$$k[X_1, \ldots, X_n]^G := \{f(X) \in k[X_1, \ldots, X_n] : f(g \cdot X) = f(X) \text{ for all } g \in G\}$$

is finitely generated, where $g \cdot X_i = \sum_j g_{ij}X_j$. A counter-example was given by Nagata in 1959.

[2] The phrase 'almost all zero' means that only finitely many are non-zero.

Note that if I_i is generated by a subset X_i, then $\sum_i I_i$ is generated by $\bigcup_i X_i$ and $I_1 I_2$ is generated by $\{x_1 x_2 : x_i \in X_i\}$.

The sum and product of ideals satisfy the **distributive law**

$$I(J+K) = IJ + IK,$$

but in general the sum and intersection only satisfy the **modular law**

$$I \cap (J+K) = I \cap J + I \cap K \quad \text{provided either } J \subset I \text{ or } K \subset I.$$

We clearly also have

$$(I+J)(I \cap J) \subset IJ \subset I \cap J,$$

but these inclusions are usually strict. We say that I and J are **coprime** provided $I + J = R$, in which case we have $IJ = I \cap J$.

Examples. Let $R = \mathbb{Z}$, $I = (a)$ and $J = (b)$. Then

$$I + J = (\gcd(a,b)), \quad I \cap J = (\operatorname{lcm}(a,b)) \quad \text{and} \quad IJ = (ab).$$

Thus I and J are coprime if and only if a and b are coprime.

Let $R = K[X_1, \ldots, X_n]$ and $I = (X_1, \ldots, X_n)$. Then $f \in I^r$ if and only if f has no terms of degree less than r. In particular, $f \in I$ if and only if f has no constant term.

Proposition 1.4 (Chinese Remainder Theorem)**.** *Let $I_1, \ldots, I_n$ be ideals in a ring R, and consider the direct product $S := (R/I_1) \times \cdots \times (R/I_n)$. Let $f\colon R \to S$ be the ring homomorphism $r \mapsto (r + I_1, \ldots, r + I_n)$.*

(1) f is injective if and only if $I_1 \cap \cdots \cap I_n = (0)$.

(2) f is surjective if and only if the I_i are pairwise coprime.

(3) If the I_i are pairwise coprime, then $I_1 \cap \cdots \cap I_n = I_1 I_2 \cdots I_n$.

Proof. (1) The kernel of f is clearly $\bigcap_i I_i$.

(2) Suppose that f is surjective. Then the element $e_i \in S$ having a 1 in place i and zeros elsewhere is contained in the image of f, say $e_i = f(y)$. Then $y - 1 = x \in I_i$, and $y \in I_j$ for all $j \neq i$. Thus $1 = x + y \in I_i + I_j$, so the ideals I_i and I_j are coprime.

Conversely, let the I_i be pairwise coprime. Since $I_1 + I_i = R$ we can find elements $a_i \in I_1$ and $b_i \in I_i$ such that $a_i + b_i = 1$. Thus

$$1 = \prod_i (a_i + b_i) = x_1 + y_1, \quad \text{where} \quad y_1 = \prod_i b_i.$$

Note that $x_1 \in I_1$ and $y_1 \in \prod_{i>1} I_i$, so $f(y_1) = e_1$. Similarly, for each i we can find y_i such that $f(y_i) = e_i$. Thus, given $r_i \in R$, set $r := \sum_i r_i y_i$. Then

$$f(r) = \sum_i f(r_i) e_i = (r_1 + I_1, \ldots, r_n + I_n),$$

so f is surjective.

(3) The proof of (2) shows that if the ideals I_i are pairwise coprime, then so are the ideals I_1 and $\prod_{i>1} I_i$. Thus $I_1(I_2 \cdots I_n) = I_1 \cap (I_2 \cdots I_n)$, and by induction this equals $I_1 \cap \cdots \cap I_n$. □

2 Maximal, prime and radical ideals

Prime ideals are one of the central objects of study in commutative algebra. They generalise the notion of prime numbers (or prime elements in integral domains), and also the notion of points in algebraic geometry.

An ideal $I \lhd R$ is called

maximal	if I is proper and $I \subset J \lhd R$ implies $J = I$ or $J = R$.
prime	if I is proper and $xy \in I$ implies either $x \in I$ or $y \in I$.
radical	if $x^n \in I$ implies $x \in I$.

We also need the following terminology. In a non-trivial ring R we call an element x

a unit	if there exists y with $xy = 1$.
a zero-divisor	if there exists $y \neq 0$ with $xy = 0$.
nilpotent	if there exists $n > 0$ with $x^n = 0$.

and we call the ring R

a field	if every non-zero element is a unit.
an integral domain	if it has no zero-divisors other than 0.
reduced	if it has no nilpotent elements other than 0.

We can now relate the concepts of prime, maximal and radical ideals to the concepts of fields, integral domains and reduced rings.

Lemma 2.1. *Let R be a non-trivial ring and $I \lhd R$ a proper ideal.*

(1a) I is maximal if and only if R/I is a field.

(1b) R is a field if and only if (0) is a maximal ideal.

(2a) I is prime if and only R/I is an integral domain.

(2b) R is an integral domain if and only if (0) is a prime ideal.

(3a) I is radical if and only if R/I is reduced.

(3b) R is reduced if and only if (0) is radical.

Proof. Left as an exercise. □

In particular, each maximal ideal is prime, and each prime ideal is radical.

One useful property of prime ideals is that the preimage of a prime ideal under a ring homomorphism is again a prime ideal. This is not true for maximal ideals.

Lemma 2.2. *Let $f\colon R \to S$ be a ring homomorphism and $Q \lhd S$ a prime ideal. Then $P := f^{-1}(Q)$ is a prime ideal of R.*

Proof. Suppose $xy \in P$. Then $f(x)f(y) = f(xy) \in Q$, and Q prime implies either $f(x) \in Q$ or $f(y) \in Q$. Thus either $x \in P$ or $y \in P$, so P is prime. □

An alternative approach is as follows. We know S/Q is an integral domain, and the induced homomorphism $\bar{f}\colon R/P \to S/Q$ is injective, so R/P is also an integral domain, whence P is a prime ideal.

Proposition 2.3. *(1) Let $I_1, \dots, I_n$ be ideals, and P a prime ideal containing $\bigcap_i I_i$. Then $P \supset I_i$ for some i. Moreover, if $P = \bigcap_i I_i$, then $P = I_i$ for some i.*

(2) (Prime Avoidance) *Let $P_1, \dots, P_n$ be prime ideals, and I an ideal contained in $\bigcup_i P_i$. Then $I \subset P_i$ for some i.*

Proof. (1) We prove the contrapositive, so suppose that $I_i \not\subset P$ for all i. Take $x_i \in I_i - P$ and set $x := x_1 \cdots x_n$. Then $x \in \bigcap_i I_i$ but since P is prime, we must have $x \notin P$. Hence $\bigcap_i I_i \not\subset P$.

For the second part, if also $P \subset \bigcap_i I_i$, then $P \subset I_i$ for all i. Since $I_i \subset P$ for some i, we have $P = I_i$ for this i.

(2) We again prove the contrapositive, so suppose $I \not\subset P_i$ for all i. By induction on n we can find $x_i \in I$ with $x_i \notin \bigcup_{j \neq i} P_j$. If for some i we have $x_i \notin P_i$, then we are done. Assume therefore that $x_i \in P_i$ for all i. Set

$$y := x_1 \cdots x_{n-1} + x_n \in I.$$

For all $i < n$ we have $x_n \notin P_i$, so $y \notin P_i$; also $x_i \notin P_n$, so $x_1 \cdots x_{n-1} \notin P_n$ since P_n is prime, and hence $y \notin P_n$. Thus $I \not\subset \bigcup_i P_i$. □

2.1 Existence of maximal ideals

Since prime and maximal ideals play such an important role, we should show that they always exist.

Proposition 2.4. *Let R be a non-trivial ring. Then every proper ideal is contained in a maximal ideal.*

To prove this, we need to apply Zorn's Lemma.[3]

Let $(\mathcal{S}, \leq)$ be a **poset**, so $\mathcal{S}$ is a set and $\leq$ is a partial order on $\mathcal{S}$. A **chain** in $\mathcal{S}$ is a totally ordered subset $\mathcal{C} \subset \mathcal{S}$; that is, for all $x, y \in \mathcal{C}$ either $x \leq y$ or $y \leq x$. An **upper bound** for a subset $\mathcal{C} \subset \mathcal{S}$ is an element $m \in \mathcal{S}$ such that $x \leq m$ for all $x \in \mathcal{C}$. An element $m \in \mathcal{S}$ is called **maximal** if $m \leq x$ implies $x = m$.

Zorn's Lemma. Let $(\mathcal{S}, \leq)$ be a non-empty poset in which every chain has an upper bound. Then S has a maximal element.

We can now prove our proposition concerning maximal ideals.

Proof of Proposition. Let $\pi \colon R \to R/I$ be the canonical map. If M is a maximal ideal in R/I, then $\pi^{-1}(M)$ is a maximal ideal in R containing I. We may therefore assume that $I = 0$.

Let $\mathcal{S}$ be the set of proper ideals of R, together with the partial order coming from inclusion (so $I \leq J$ in $\mathcal{S}$ if $I \subset J$). Note that $\mathcal{S}$ is non-empty since it contains 0.

[3] Zorn's Lemma is often used in algebra as a replacement for induction, and so is usually assumed to hold. It is equivalent in Zermelo-Fraenkel Set Theory to the Axiom of Choice. Note that, in addition to implying useful results in algebra, it also leads to more 'surprising' results, such as the Banach-Tarski Paradox.

Now let $\mathcal{C}$ be a chain in $\mathcal{S}$. We claim that $X := \bigcup_{I \in \mathcal{C}} I$ is an upper bound for $\mathcal{C}$ in $\mathcal{S}$. Clearly $I \subset X$ for all $I \in \mathcal{C}$, so X is an upper bound for $\mathcal{C}$. We therefore just need to check that $X \in \mathcal{S}$, so is a proper ideal.

• X is an ideal of R. If $x, y \in X$ and $r \in R$, then there exists some ideals $I_1, I_2 \in \mathcal{C}$ with $x \in I_1$ and $y \in I_2$. Since $\mathcal{C}$ is a chain we have either $I_1 \subset I_2$ or $I_2 \subset I_1$. Let I be the bigger of I_1 and I_2, so $x, y \in I$. Now $I \in \mathcal{C}$ is an ideal, so $x + y \in I$ and $rx \in I$, whence $x + y, rx \in X$.

• X is a proper ideal. If $X = R$, then $1 \in X$, so $1 \in I$ for some $I \in \mathcal{C}$, a contradiction since I is proper.

We can now apply Zorn's Lemma to get a maximal element $M \in \mathcal{S}$, and M is necessarily a maximal ideal of R. □

2.2 The radical of an ideal

The **radical of an ideal** I is the set

$$\operatorname{rad}(I) := \{x \in R : x^n \in I \text{ for some } n > 0\}.$$

The **nilradical** of R is

$$\operatorname{nil}(R) := \operatorname{rad}(0) = \{x \in R : x \text{ is nilpotent}\}.$$

Lemma 2.5. *The radical of an ideal is a radical ideal. Under the canonical homomorphism $\pi\colon R \to R/I$ we have*

$$\operatorname{rad}(I) = \pi^{-1}(\operatorname{nil}(R/I)).$$

Proof. Suppose $x, y \in \operatorname{rad}(I)$, say with $x^m, y^n \in I$. Using the binomial expansion

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

we see that $(x + y)^N \in I$ for all $N \geq m + n - 1$. Thus $x + y \in \operatorname{rad}(I)$. Clearly $(rx)^m = r^m x^m \in I$ for all $r \in R$, so $\operatorname{rad}(I)$ is an ideal. Now, if $z^m \in \operatorname{rad}(I)$, then $(z^m)^n \in I$ for some n, so $z^{mn} \in I$ and hence $z \in \operatorname{rad}(I)$. Thus $\operatorname{rad}(I)$ is a radical ideal.

For the second part we note that $x^m \in I$ if and only if $\pi(x)^m = 0$, and so $x \in \operatorname{rad}(I)$ if and only if $\pi(x)$ is nilpotent. □

Lemma 2.6. *We have the following properties:*

(1) If I is radical, then $\operatorname{rad}(I) = I$.

(2) $I \subset \operatorname{rad}(I)$, *and* $\operatorname{rad}(I) = R$ *if and only if* $I = R$.

(3) $I \subset J$ *implies* $\operatorname{rad}(I) \subset \operatorname{rad}(J)$.

(4) $\operatorname{rad}(IJ) = \operatorname{rad}(I \cap J) = \operatorname{rad}(I) \cap \operatorname{rad}(J)$.

(5) $\operatorname{rad}(I + J) = \operatorname{rad}(\operatorname{rad}(I) + \operatorname{rad}(J))$.

Proof. (1) Clear.

(2) Clearly $I \subset \mathrm{rad}(I)$. Also, $\mathrm{rad}(I) = R$ implies $1 \in \mathrm{rad}(I)$, so $1 = 1^n \in I$, and hence $I = R$.

(3) Clear.

(4) Since $IJ \subset I \cap J$, we have $\mathrm{rad}(IJ) \subset \mathrm{rad}(I \cap J)$. Also, if $x \in \mathrm{rad}(I \cap J)$, then $x^n \in I \cap J$ for some n, so $x \in \mathrm{rad}(I) \cap \mathrm{rad}(J)$. Conversely, if $x \in \mathrm{rad}(I) \cap \mathrm{rad}(J)$, say with $x^m \in I$ and $x^n \in J$, then $x^{m+n} \in IJ$, so $x \in \mathrm{rad}(IJ)$. Thus $\mathrm{rad}(I) \cap \mathrm{rad}(J) \subset \mathrm{rad}(IJ)$.

(5) Since $I+J \subset \mathrm{rad}(I)+\mathrm{rad}(J)$, we get $\mathrm{rad}(I+J) \subset \mathrm{rad}(\mathrm{rad}(I)+\mathrm{rad}(J))$. Conversely, if $z^N \in \mathrm{rad}(I)+\mathrm{rad}(J)$, then we can write $z^N = x+y$ with $x^m \in I$ and $y^n \in J$. Using the binomial expansion we see that $(x+y)^{m+n} \in I+J$, and so $z^{(m+n)N} \in I+J$. Hence $\mathrm{rad}(\mathrm{rad}(I)+\mathrm{rad}(J)) \subset \mathrm{rad}(I+J)$. □

Note that by (4), $\mathrm{rad}(I^n) = \mathrm{rad}(I)$, so if I is radical, then $\mathrm{rad}(I^n) = I$.

Theorem 2.7. *The radical of an ideal I is the intersection of all prime ideals containing I. Conversely, each intersection of prime ideals is radical.*[4]

In particular, $\mathrm{nil}(R)$ *is the intersection of all prime ideals of R.*

Proof. If P is a prime ideal containing I, then $\mathrm{rad}(I) \subset \mathrm{rad}(P) = P$.

Conversely, suppose that x is not in $\mathrm{rad}(I)$. It is enough to show that there is a prime ideal containing I but not containing x.

Let $\mathcal{S}$ be the set of all ideals containing I but not containing any power of x, partially ordered by inclusion. Now $\mathcal{S}$ is non-empty, since it contains I, and every chain $\mathcal{C}$ in $\mathcal{S}$ has an upper bound, given by the union $\bigcup_{J \in \mathcal{C}} J$. We can therefore apply Zorn's Lemma to $\mathcal{S}$ to deduce the existence of a maximal element P.

We wish to show that P is a prime ideal, so suppose that $a, b \notin P$. Then $P + (a)$ is strictly bigger than P, whence $x^m \in P + (a)$ for some m. Similarly $x^n \in P + (b)$ for some n, so $x^{m+n} \in (P + (a))(P + (b)) \subset Q + (ab)$. Thus $ab \notin P$, and P is prime.

Applying this to the zero ideal yields the result about the nilradical.

Finally, consider an intersection of (possibly infinitely many) prime ideals $I = \bigcap_i P_i$. If $x^n \in I$, then for each i we have $x^n \in P_i$, so $x \in P_i$. Hence $x \in I$, so I is radical. □

[4] We observe that this lemma also applies to the unit ideal, and to the zero ring, using the convention that an empty intersection of ideals is the unit ideal.

3 Localisation

Recall that we can construct $\mathbb{Q}$ from $\mathbb{Z}$ as the set of equivalence classes in $\mathbb{Z}\times(\mathbb{Z}-\{0\})$ via the equivalence relation

$$(r,a)\sim(s,b) \quad \text{provided} \quad sa=rb.$$

As usual, we write $\frac{r}{a}$ for the equivalence class of (r,a).

More generally, let R be any integral domain. Then the **field of fractions**, or **quotient field**, $\mathrm{Quot}(R)$ is the set of equivalence classes in $R\times(R-\{0\})$ via the equivalence relation

$$(r,a)\sim(s,b) \quad \text{provided} \quad sa=rb.$$

The addition and multiplication in $\mathrm{Quot}(R)$ are given by the operations

$$\frac{r}{a}+\frac{s}{b}:=\frac{sa+rb}{ab} \qquad \text{and} \qquad \frac{r}{a}\cdot\frac{s}{b}:=\frac{rs}{ab}.$$

where $\frac{r}{a}$ denotes the equivalence class of (r,a). The unit is $\frac{1}{1}$ and the zero is $\frac{0}{1}$.

Lemma 3.1. *Let R be an integral domain. Then the addition and multiplication in $\mathrm{Quot}(R)$ are well-defined, $\mathrm{Quot}(R)$ is a field, and the canonical map*

$$R\to\mathrm{Quot}(R), \quad r\mapsto\tfrac{r}{1}$$

is an injective ring homomorphism. In particular, we may identitfy R with its image in $\mathrm{Quot}(R)$.

Moreover, if $f\colon R\to K$ is an injective ring homomorphism from R to a field K, then there is a unique ring homomorphism $\hat{f}\colon\mathrm{Quot}(R)\to K$ extending f.

Warning. The definition above does not work for arbitrary rings, since the relation will not be transitive. We can overcome this hurdle as follows.

Let R be a ring. A **multiplicatively closed subset** $\Sigma\subset R$ is a subset containing 1 and closed under products, so $a,b\in\Sigma$ implies $ab\in\Sigma$. Given a multiplicatively closed subset $\Sigma\subset R$, we define an equivalence relation on $R\times\Sigma$ via

$$(r,a)\sim(s,b) \quad \text{provided there exists } c\in\Sigma \text{ with } (sa-rb)c=0.$$

We again denote the equivalence class of (r,a) by $\frac{r}{a}$, and can define the sum and product of equivalence classes via

$$\frac{r}{a}+\frac{s}{b}:=\frac{as+br}{ab} \qquad \text{and} \qquad \frac{r}{a}\cdot\frac{s}{b}:=\frac{rs}{ab}.$$

We denote the set of equivalence classes by R_Σ or $R[\Sigma^{-1}]$ and call it the **ring of fractions**, or **localisation**, of R with respect to Σ.

Lemma 3.2. *Let R be a ring and Σ a multiplicatively closed subset of R. Then the addition and multiplication in R_Σ are well-defined, and we have a (usually non-injective) ring homomorphism*

$$\sigma\colon R\to R_\Sigma, \quad r\mapsto\tfrac{r}{1}.$$

Note that $R_\Sigma = 0$ if and only if $\frac{1}{1} = \frac{0}{1}$, if and only if $0 \in \Sigma$.

Proposition 3.3 (Universal Property)**.** *Let R be a ring and $\Sigma \subset R$ a multiplicatively closed subset. If $f\colon R \to S$ is a ring homomorphism such that $f(a)$ is invertible for all $a \in \Sigma$, then there exists a unique ring homomorphism $\hat{f}\colon R_\Sigma \to S$ with $f = \hat{f}\sigma$.*

Proof. Existence. We define[5] $\hat{f}(\frac{r}{a}) := f(r)f(a)^{-1}$. This is well-defined, since if $\frac{r}{a} = \frac{r'}{a'}$, then there exists $b \in \Sigma$ with $(ar' - a'r)b = 0$. Applying f we get

$$f(a)f(r')f(b) = f(a')f(r)f(b),$$

and since $f(a), f(a'), f(b)$ are all invertible in S we get $f(r)f(a)^{-1} = f(r')f(a')^{-1}$.

Now $\hat{f}(\frac{r}{1}) = f(r)f(1)^{-1} = f(r)$, so $\hat{f}\sigma = f$. In particular, $\hat{f}(\frac{1}{1}) = 1$. Moreover

$$\begin{aligned}\hat{f}\Big(\frac{r}{a} + \frac{r'}{a'}\Big) = \hat{f}\Big(\frac{ar' + a'r}{aa'}\Big) &= f(ar' + a'r)f(aa')^{-1} \\ &= f(r)f(a)^{-1} + f(r')f(a')^{-1} = \hat{f}\Big(\frac{r}{a}\Big) + \hat{f}\Big(\frac{r'}{a'}\Big),\end{aligned}$$

and

$$\hat{f}\Big(\frac{r}{a} \cdot \frac{r'}{a'}\Big) = \hat{f}\Big(\frac{rr'}{aa'}\Big) = f(rr')f(aa')^{-1} = f(r)f(a)^{-1} \cdot f(r')f(a')^{-1} = \hat{f}\Big(\frac{r}{a}\Big)\hat{f}\Big(\frac{r'}{a'}\Big),$$

so $\hat{f}$ is a ring homomorphism.

Uniqueness. Suppose that $g\colon R_\Sigma \to S$ also satisfies $g\sigma = f$. Then for $a \in \Sigma$ we have

$$f(a)g\Big(\frac{r}{a}\Big) = g\Big(\frac{a}{1}\Big)g\Big(\frac{r}{a}\Big) = g\Big(\frac{a}{1} \cdot \frac{r}{a}\Big) = g\Big(\frac{r}{1}\Big) = f(r).$$

Since $f(a)$ is invertible we have $g(\frac{r}{a}) = f(r)f(a)^{-1}$, so $g = \hat{f}$. □

Lemma 3.4. *Let $f\colon R \to S$ be a ring homomorphism such that $f(a)$ is invertible in S for all $a \in \Sigma$, and let $\hat{f}\colon R_\Sigma \to S$ be the extension of f. Then*

(1) $\hat{f}$ is injective if and only if $f(r) = 0$ in S implies $ar = 0$ in R for some $a \in \Sigma$.

(2) $\hat{f}$ is surjective if and only if every element of S can be expressed as $f(r)f(a)^{-1}$ for some $r \in R$ and $a \in \Sigma$.

Proof. (1) Since $\hat{f}(\frac{r}{a}) = f(r)f(a)^{-1}$, we see that the kernel of $\hat{f}$ is generated by all $\frac{r}{1}$ with $f(r) = 0$. So $\hat{f}$ is injective if and only if $f(r) = 0$ implies $\frac{r}{1} = 0$ in R_Σ, which is equivalent to $ar = 0$ in R for some $a \in \Sigma$.

(2) Again, $\hat{f}(\frac{r}{a}) = f(r)f(a)^{-1}$, so this is clear. □

One can sometimes use this lemma to construct the localisation, as we do, for example, in Proposition 3.6.

[5] Note the change of notation. Here, $\frac{r}{a}$ is a symbol representing an equivalence class in R_Σ, whereas $f(a)$ is invertible in S, with inverse $f(a)^{-1}$.

Examples. An important example is when $\Sigma = R - P$ for some prime ideal P, in which case the ring of fractions $R_P := R_\Sigma$ is called the **local ring at** P. In general, a **local ring** is a ring having a unique maximal ideal; we will prove this for R_P in Corollary 3.8.

We can also localise at an element. Given $a \in R$, set $\Sigma := \{a^n : n \geq 0\}$ and write R_a or $R[\frac{1}{a}]$ for the localisation R_Σ. We will see in the exercises that $R_a \cong R[X]/(aX-1)$.

As a special case, consider the prime ideal $(p) \lhd \mathbb{Z}$. Then $\mathbb{Z}_{(p)}$ is the subring of $\mathbb{Q}$ consisting of all fractions having denominator coprime to p. On the other hand, $\mathbb{Z}_p = \mathbb{Z}[\frac{1}{p}]$ is the subring[6] of $\mathbb{Q}$ consisting of all fractions having denominator a power of p.

3.1 Ideals in a localisation

Fix a ring R and a multiplicatively closed subset $\Sigma \subset R$. For an ideal I of R we set

$$I_\Sigma := \{\tfrac{x}{a} : x \in I, a \in \Sigma\}.$$

Note that $\frac{x}{a}$ may be equivalent to some $\frac{r}{b}$ with $r \notin I$. Part (2) of the lemma below describes precisely those classes lying in I_Σ.

Lemma 3.5. *The following hold.*

(1) I_Σ *is an ideal of* R_Σ*. If* I *is generated by a subset* X*, then* I_Σ *is generated by* $\{\frac{x}{1} : x \in X\}$*.*

(2) We have $\frac{r}{b} \in I_\Sigma$ *if and only if there exists* $a \in \Sigma$ *with* $ar \in I$*.*

(3) $I_\Sigma = R_\Sigma$ *if and only if* $I \cap \Sigma$ *is non-empty.*

(4) The map $I \mapsto I_\Sigma$ *commutes with forming radicals and sums, as well as finite products and intersections.*

Proof. (1) Let $r \in R$, $x, x' \in I$ and $a, a', b \in \Sigma$. Then

$$\frac{r}{b} \cdot \frac{x}{a} + \frac{x'}{a'} = \frac{abx' + a'rx}{aa'b} \in I_\Sigma,$$

so I_Σ is an ideal of R_Σ.

If I is generated by X, then every element in I can be written as a finite sum $y = r_1x_1 + \cdots + r_nx_n$ with $x_i \in X$, and hence

$$\frac{y}{a} = \frac{r_1x_1 + \cdots + r_nx_n}{a} = \frac{r_1}{a} \cdot \frac{x_1}{1} + \cdots + \frac{r_n}{a} \cdot \frac{x_n}{1}.$$

It follows that I_Σ is generated by $\{\frac{x}{1} : x \in X\}$.

(2) Suppose $\frac{r}{b} \in I_\Sigma$, say $\frac{r}{b} = \frac{x}{a}$ for some $x \in I$ and $a \in \Sigma$. Then $(ar - bx)c = 0$ for some $c \in \Sigma$, so $acr = bcx \in I$ and $ac \in \Sigma$. Conversely, if $ar = x \in I$, then $ab \in \Sigma$, and so $\frac{r}{b} = \frac{x}{ab} \in I_\Sigma$.

[6] This notation should not be confused with the cyclic group of order p, which it is better to write as $\mathbb{Z}/(p)$, or with the ring of p-adic integers.

(3) We have $I_\Sigma = R_\Sigma$ if and only if $\frac{1}{1} \in I_\Sigma$, which by (2) is if and only if there exists $a \in \Sigma$ with $a \in I$. Hence I_Σ is the unit ideal if and only if $I \cap \Sigma \neq \emptyset$.

(4) The sum $\sum_i I_i$ is generated by $\{x : x \in \bigcup_i I_i\}$, so $\left(\sum_i I_i\right)_\Sigma$ is generated by $\{\frac{x}{1} : x \in \bigcup_i I_i\}$. On the other hand, $(I_i)_\Sigma$ is generated by $\{\frac{x}{1} : x \in I_i\}$, so the sum $\sum_i (I_i)_\Sigma$ is generated by $\{\frac{x}{1} : x \in \bigcup_i I_i\}$.

The product IJ is generated by $\{xy : x \in I,\ y \in J\}$, so $(IJ)_\Sigma$ is generated by $\{\frac{xy}{1} : x \in I,\ y \in J\}$. On the other hand, I_Σ is generated by $\{\frac{x}{1} : x \in I\}$, so $I_\Sigma J_\Sigma$ is generated by $\{\frac{x}{1}\frac{y}{1} = \frac{xy}{1} : x \in I,\ y \in J\}$. The result for finite products now follows by induction.

We always have that $\left(\bigcap_i I_i\right)_\Sigma \subset \bigcap_i (I_i)_\Sigma$. For, if $x \in \bigcap_i I_i$ and $a \in \Sigma$, then $\frac{x}{a} \in \bigcap_i (I_i)_\Sigma$. If the intersection is finite, then we have equality. For, let $\frac{r}{a} \in (I_1)_\Sigma \cap \cdots \cap (I_n)_\Sigma$. Then by (2) there exists $b_i \in \Sigma$ with $b_i r \in I_i$. Set $b = b_1 \cdots b_n \in \Sigma$. Then $br \in \bigcap_i I_i$, and so $\frac{r}{a} \in \left(\bigcap_i I_i\right)_\Sigma$.

Finally, if $x \in \operatorname{rad}(I)$, say $x^n \in I$, then $(\frac{x}{a})^n = \frac{x^n}{a^n} \in I_\Sigma$, so $\frac{x}{a} \in \operatorname{rad}(I_\Sigma)$. Conversely, let $\frac{r}{a} \in \operatorname{rad}(I_\Sigma)$, say $(\frac{r}{a})^n = \frac{r^n}{a^n} \in I_\Sigma$. Then by (2) there exists $b \in \Sigma$ with $br^n \in I$, so $(br)^n \in I$ and hence $br \in \operatorname{rad}(I)$. Thus $\frac{r}{a} \in \operatorname{rad}(I)_\Sigma$. □

The next result is a special case of the fact that 'localisation is exact', Proposition 8.6.

Proposition 3.6. *Let $I \lhd R$ and $\Sigma \subset R$ a multiplicatively closed subset. Write $\bar{r}$ for the image of r in $\bar{R} := R/I$ and set $\bar{\Sigma} := \{\bar{a} : a \in \Sigma\}$. Then $\bar{\Sigma}$ is a multiplicatively closed subset of $\bar{R}$ and there is a natural isomorphism*

$$R_\Sigma / I_\Sigma \cong \bar{R}_{\bar{\Sigma}}.$$

Proof. Consider the composition $R \to R_\Sigma \to R_\Sigma/I_\Sigma$. Since I is contained in the kernel, the Factor Lemma gives

$$f \colon R/I \to R_\Sigma/I_\Sigma, \quad \bar{r} \mapsto \tfrac{r}{1} + I_\Sigma.$$

Next we observe that $f(\bar{a})$ is invertible[7] for all $\bar{a} \in \bar{\Sigma}$. For, $f(\bar{a}) = \frac{a}{1} + I_\Sigma$, and so has inverse $\frac{1}{a} + I_\Sigma$. So, by the universal property for localisations, Proposition 3.3, there exists a unique ring homomorphism

$$\hat{f} \colon \bar{R}_{\bar{\Sigma}} \to R_\Sigma/I_\Sigma, \quad \frac{\bar{r}}{\bar{a}} \mapsto f(\bar{r}) f(\bar{a})^{-1} = \frac{r}{a} + I_\Sigma.$$

We now apply Lemma 3.4. Every element in R_Σ/I_Σ is of the form $\frac{r}{a} + I_\Sigma$, so $\hat{f}$ is surjective. On the other hand, suppose $f(\bar{r}) = 0$. Then $\frac{r}{1} \in I_\Sigma$, so $ar \in I$ for some $a \in \Sigma$, and hence $\bar{a}\bar{r} = 0$. Thus $\hat{f}$ is injective. □

Theorem 3.7. *The ring homomorphism $\sigma \colon R \to R_\Sigma$, $r \mapsto \frac{r}{1}$, induces a bijection*

$$\{\textit{ideals } J \lhd R_\Sigma\} \overset{1\text{-}1}{\longleftrightarrow} \{\textit{ideals } I \lhd R,\ \textit{no element of } \Sigma \textit{ is a zero divisor in } R/I\}$$

sending $J \mapsto \sigma^{-1}(J)$ and $I \mapsto I_\Sigma$.

This restricts to a bijection

$$\{\textit{prime ideals } Q \lhd R_\Sigma\} \overset{1\text{-}1}{\longleftrightarrow} \{\textit{prime ideals } P \lhd R,\ P \cap \Sigma = \emptyset\}.$$

[7] It may happen, though, that $R_\Sigma/I_\Sigma = 0$.

Proof. If $J \lhd R_\Sigma$, then $\sigma^{-1}(J) \lhd R$ (the preimage of an ideal under a ring homomorphism is always an ideal). Observe that $\sigma^{-1}(J) = \{x \in R : \frac{x}{1} \in J\}$, so clearly $\sigma^{-1}(J)_\Sigma \subset J$. On the other hand, if $\frac{x}{a} \in J$, then $\frac{x}{1} = \frac{a}{1}\frac{x}{a} \in J$, so $x \in \sigma^{-1}(J)$ and $\frac{x}{a} \in \sigma^{-1}(J)_\Sigma$. Thus $(\sigma^{-1}(J))_\Sigma = J$.

This shows that the map $J \mapsto \sigma^{-1}(J)$ from ideals of R_Σ to ideals of R is injective, with inverse $I \mapsto I_\Sigma$. So, we just need to characterise its image.

Note that I is in the image if and only if $I = \sigma^{-1}(I_\Sigma)$. For, such an ideal is clearly in the image, whereas if $I = \sigma^{-1}(J)$, then $I_\Sigma = \sigma^{-1}(J)_\Sigma = J$, and so $\sigma^{-1}(I_\Sigma) = \sigma^{-1}(J) = I$.

We always have $I \subset \sigma^{-1}(I_\Sigma)$, so $I \neq \sigma^{-1}(I_\Sigma)$ if and only if there exists $x \notin I$ such that $\frac{x}{1} \in I_\Sigma$, which is if and only if there exists $x \notin I$ and $a \in \Sigma$ with $ax \in I$. This is if and only if there exists $a \in \Sigma$ and $\bar{x} \neq 0$ in R/I with $\bar{a}\bar{x} = 0$, equivalently some element of Σ is a zero divisor in R/I.

For the second part, we first observe that if $P \lhd R$ is prime, then R/P is an integral domain, so Σ contains a zero divisor in R/P if and only if $\Sigma \cap P \neq \emptyset$. So, if $Q \lhd R_\Sigma$ is prime, then $\sigma^{-1}(Q) \lhd R$ is a prime (Lemma 2.2) not intersecting Σ.

Conversely, let $P \lhd R$ be a prime not intersecting Σ. Then R/P is an integral domain, and $\bar{\Sigma} \subset R/P$ does not contain zero, so by Proposition 3.6

$$R_\Sigma/P_\Sigma \cong (R/P)_{\bar{\Sigma}} \subset \mathrm{Quot}(R/P).$$

Since $\mathrm{Quot}(R/P)$ is a field, it follows that R_Σ/P_Σ is an integral domain, and hence $P_\Sigma \lhd R_\Sigma$ is prime. □

Corollary 3.8. *Let $P \lhd R$ be a prime ideal. Then the prime ideals of R_P are in bijection with the prime ideals of R contained in P. In particular, R_P has a unique maximal ideal, P_P, and hence is a local ring.*

Proof. We take $\Sigma := R - P$. Clearly a prime ideal $P' \lhd R$ does not intersect Σ if and only if $P' \subset P$. Now apply the theorem. □

Let $P \subset Q$ be prime ideals of R. Then the prime ideals of R/P are in bijection with the prime ideals of R containing P (Third Isomorphism Theorem), whereas the prime ideals of R_Q are in bijection with the prime ideals of R contained in Q. Now set $\Sigma := R - Q$, so that $\bar{\Sigma} = (R/P) - (Q/P)$. Using that localisation and taking factor rings commute, Proposition 3.6, we get

$$(R_Q)/(P_Q) \cong (R/P)_{Q/P},$$

and the prime ideals of this ring are in bijection with the prime ideals of R lying between P and Q.

As a special case, when $P = Q$, we obtain the **residue field at** P,

$$\kappa(P) := (R_P)/(P_P) = \mathrm{Quot}(R/P).$$

4 Modules

Recall that a **module** M over a ring R is given by an abelian group under addition, with identity element 0, together with an associative, unital and bilinear action

$$R \times M \to M, \quad (r, m) \mapsto rm,$$

so, for all $r, s \in R$ and $m, n \in M$, we have

$$r(sm) = (rs)m, \quad 1m = m, \quad r(m+n) = rm + rn \quad \text{and} \quad (r+s)m = rm + sm.$$

Equivalently, we observe that the set $\mathrm{End}(M)$ of endomorphisms of an abelian group M is a non-commutative ring, via

$$(f+g)(m) := f(m) + g(m) \quad \text{and} \quad (fg)(m) := f(g(m)).$$

Then M is an R-module if and only if there is a ring homomorphism $\rho\colon R \to \mathrm{End}(M)$. For, if M is an R-module, then we define $\rho(r)$ to be the map $m \mapsto rm$, whereas if ρ is a ring homomorphism, then we define the action via $rm := \rho(r)(m)$.

A map $f\colon M \to N$ between R-modules is called an R-module homomorphism, or an R-linear map, if it is an additive group homomorphism which respects the action

$$f(rm) = rf(m) \quad \text{for all } r \in R \text{ and } m \in M.$$

The composition of R-module homomorphisms is again an R-module homomorphism, and the identity map $1_M\colon M \to M$ is also an R-module homomorphism.

Remark. In fact there is a category having objects the R-modules and morphisms the R-module homomorphisms.

Examples. A $\mathbb{Z}$-module is just an abelian group, and a $\mathbb{Z}$-module homomorphism is just a group homomorphism of abelian groups. Similarly if K is a field, then a K-module is just a K-vector space, and a K-module homomorphism is just a K-linear map between vector spaces.

A $K[X]$-module is a vector space equipped with a linear endomorphism.

We denote the set of all R-module homomorphisms $f\colon M \to N$ by $\mathrm{Hom}_R(M, N)$. This is an abelian group under addition, and composition is bilinear.In fact it is even an R-module via the action

$$(rf)(m) := rf(m) \quad \text{for all } r \in R,\ f \in \mathrm{Hom}_R(M, N) \text{ and } m \in M.$$

Moreover, composition is R-bilinear, in the sense that

$$f(rg + g') = r(fg) + fg' \quad \text{and} \quad (rf + f')g = r(fg) + f'g$$

for all $f, f'\colon M \to N$, $g, g'\colon L \to M$ and $r \in R$. It follows that

Lemma 4.1. *For each R-module M we have an R-module isomorphism*

$$\mathrm{Hom}_R(R, M) \to M, \quad f \mapsto f(1).$$

Given a homomorphism $f\colon M \to N$ and a module X, we get induced homomorphisms

$$f_* = \mathrm{Hom}_R(X, f)\colon \mathrm{Hom}_R(X, M) \to \mathrm{Hom}_R(X, N), \quad g \mapsto fg$$

and

$$f^* = \mathrm{Hom}_R(f, X)\colon \mathrm{Hom}_R(N, X) \to \mathrm{Hom}_R(M, X), \quad g \mapsto gf.$$

Remark. This says that $\mathrm{Hom}_R(X, -)$ is a covariant functor from the category of R-modules to itself, and that $\mathrm{Hom}_R(-, X)$ is a contravariant functor.

One special case of this is duality for vector spaces. If K is a field, then $D(V) := \mathrm{Hom}_K(V, K)$ is the dual space to V, and f^* is the dual, or transpose, of f.

Modules over algebras

Suppose that R is a k-algebra. Then every R-module is necessarily a k-module, and every R-linear map is k-linear. Thus an R-module can be thought of as a k-module M together with an associative, unital, and k-bilinear map $R \times M \to M$.

Alternatively, it is easy to check that $\mathrm{End}_k(M)$ is a non-commutative ring via composition, in which case giving an R-module structure on M is the same as giving a k-algebra homomorphism $R \to \mathrm{End}_k(M)$.

This point of view is useful when R is a K-algebra over some field K. Then an R-module is a K-vector space M together with an R-action, or equivalently a K-algebra homomorphism $R \to \mathrm{End}_K(M)$.

4.1 Submodules and factor modules

A **submodule** $U \leq M$ of an R-module M is a subgroup closed under the action, so

$$ru \in U \quad \text{for all } r \in R \text{ and } u \in U,$$

in which case the inclusion map $U \rightarrowtail M$ is a module homomorphism.

For each module M, we always have the **trivial**, or **zero**, submodule, and the submodule M itself. A submodule $U \neq M$ is called **proper**.

The **regular module** is the ring R, viewed as a module over itself. In this case the submodules of the regular module are precisely the ideals of R.

If $U \leq M$ is a submodule, then the factor group M/U becomes a module, called the **factor module**, via the action

$$r(m + U) := rm + U \quad \text{for all } r \in R \text{ and } u \in U,$$

in which case the canonical map $\pi\colon M \twoheadrightarrow M/U$ is a module homomorphism.

Let $f\colon M \to N$ be a module homomorphism. Then the **kernel** of f

$$\mathrm{Ker}(f) := \{m \in M : f(m) = 0\}$$

is a submodule of M, the **image** of f

$$\mathrm{Im}(f) := \{f(m) \in N : m \in M\}$$

is a submodule of N, and the **cokernel** of f is

$$\operatorname{Coker}(f) := N/\operatorname{Im}(f).$$

Lemma 4.2 (Factor Lemma)**.** *Let $f\colon M \to N$ be a module homomorphism, and $U \leq M$ a submodule. If $U \leq \operatorname{Ker}(f)$, then there is a unique module homomorphism*

$$\bar{f}\colon M/U \to N, \quad m + U \mapsto f(m),$$

satisfying $f = \bar{f}\pi$. We call $\bar{f}$ the homomorphism induced by f.

Corollary 4.3 (First Isomorphism Theorem)**.** *A module homomorphism $f\colon M \to N$ yields an isomorphism*

$$M/\operatorname{Ker}(f) \xrightarrow{\sim} \operatorname{Im}(f), \quad m + \operatorname{Ker}(f) \mapsto f(m).$$

4.2 Operations on modules

Most of the operations on ideals can be extended to modules.

Given a family of submodules U_i in a module M, we can form their intersection

$$\bigcap_i U_i := \{m \in M : m \in U_i \text{ for all i}\}.$$

In particular, given a subset $X \subset M$, we can form the smallest submodule containing X, namely $\bigcap_{X \subset U} U$, and called the submodule **generated** by X. Its elements are those which can be expressed in the form $m = r_1x_1 + \cdots + r_nx_n$ for some $r_i \in R$ and $x_i \in X$. A module M is called **finitely generated** if it is generated by a finite set, and **cyclic** if it is generated by a single element.

We define the sum $\sum_i U_i$ to be the submodule generated by the union $\bigcup_i U_i$,

$$\sum_i U_i := \Big\{ \sum_i u_i : um_i \in U_i \text{ almost all zero} \Big\}.$$

If U_i is generated by X_i, then $\sum_i U_i$ is generated by $\bigcup_i X_i$.

We say that the sum is direct, written $\bigoplus_i U_i$, provided that $U_i \cap \sum_{j \neq i} U_j = 0$ for all i. This is equivalent to saying that every $m \in \sum_i U_i$ can be written uniquely as $m = \sum_i u_i$ with $u_i \in U_i$ almost all zero.

We cannot multiply submodules together, but we can define the product IU of an ideal I with a submodule U; this is the submodule generated by all products rm with $r \in I$ and $m \in U$.

Given ideals I and J, and submodules U, V and W, we have the distributive law

$$I(U + V) = IU + IV \quad \text{and} \quad (I + J)U = IU + JU,$$

associativity

$$(IJ)U = I(JU),$$

and the modular law

$$U \cap (V + W) = U \cap V + U \cap W \quad \text{provided either } V \subset U \text{ or } W \subset U.$$

Let M_i be a family of R-modules. The **direct product** $\prod_i M_i$ is the Cartesian product equipped with componentwise addition and action

$$(m_i) + (n_i) := (m_i + n_i) \quad \text{and} \quad r(m_i) := (rm_i).$$

The **coproduct** (or outer direct sum) $\coprod_i M_i$ is the submodule of $\prod_i M_i$ generated by the union of the (images of the) M_i. Thus an element (m_i) lies in $\coprod_i M_i$ if and only if the $m_i \in M_i$ are almost all zero. In particular, the direct product and coproduct agree for all finite families of modules.

Remark. Note that $\prod_i M_i$ is a product in the categorical sense, and similarly $\coprod_i M_i$ is a coproduct in the categorical sense. Since finite products and coproducts agree we see that the category of R-modules is an R-linear, abelian category.

In the special case when each M_i equals a fixed module M, then we use the notation M^I for the direct product, and $M^{(I)}$ for the coproduct (where I is the indexing set for the family). When $|I| = n$ is finite, we just write M^n.

For example, as vector spaces we have

$$K[X] \cong K^{(\mathbb{N})} \quad \text{and} \quad K[[X]] \cong K^{\mathbb{N}}.$$

A module which is isomorphic to some $R^{(I)}$ is called **free**. For example, if K is a field, then we can use Zorn's Lemma to prove that every vector space has a basis. This is equivalent to the statement that every K-module is free.

Lemma 4.4. *For each module M there exists an epimorphism $F \twoheadrightarrow M$ from a free module F. Moreover, M is finitely generated if and only if we can take $F = R^n$ for some n.*

4.3 The Cayley-Hamilton Theorem

Proposition 4.5. *Let M be a finitely generated R-module, $f \in \mathrm{End}_R(M)$ an endomorphism of M, and $I \lhd R$ an ideal. If $f(M) \subset IM$, then f satisfies an equation of the form*

$$f^n + r_{n-1}f^{n-1} + \cdots + r_1 f + r_0 \mathrm{id}_M = 0 \quad \textit{with} \quad r_i \in I.$$

Proof. We first remark that since f commutes with the action of R, we can endow M with the structure of an $R[T]$-module via $(rT^n) \cdot m := rf^n(m)$. Then, via the usual matrix multiplication rules, the (non-commutative) ring $\mathbb{M}_n(R[T])$ acts on the direct sum M^n.

Let M be generated by $m_1, \ldots, m_n$. Our assumption that $f(M) \subset IM$ gives

$$f(m_i) = \sum_j \rho(x_{ij}) m_j \quad \text{for some } x_{ij} \in I.$$

Consider the matrix $X = (\delta_{ij}T - x_{ij})$, an $n \times n$ matrix with coefficients in $R[T]$. By construction we know that $X \cdot v = 0$, where v is the column vector $v = (m_1, \ldots, m_n)^t$.

Next, Cramer's Rule gives $\mathrm{adj}(X)X = \det(X)I_n$, where $\mathrm{adj}(X)$ is the adjugate matrix of X and I_n is the identity matrix. Also, $\det(X) = \chi(T)$ for some monic polynomial $\chi \in R[T]$ of degree n, and since $x_{ij} \in I$ we must have

$$\chi(T) = T^n + r_{n-1}T^{n-1} + \cdots + r_1 T + r_0, \quad r_i \in I.$$

Now $\chi(T) \cdot m = \chi(f)(m)$ for all $m \in M$, so

$$(\chi(f)(m_i))^t = \chi(T)I_n \cdot v = \mathrm{adj}(X)Xv = 0.$$

Thus $\chi(f)(m_i) = 0$ for all i, and since the m_i generate we conclude that $\chi(f) = 0$. $\square$

The **Jacobson radical** of a ring R is given by

$$\mathrm{Jac}(R) := \bigcap_{\text{max ideals}} I,$$

where the intersection is over all maximal ideals of R. In particular, $\mathrm{Jac}(R)$ is a radical ideal. There is an alternative characterisation as

$$\mathrm{Jac}(R) = \{x \in R : 1 + rx \text{ is invertible for all } r \in R\}.$$

Proposition 4.6 (Nakayama's Lemma)**.** *Let M be a finitely-generated module and $I \subset \mathrm{Jac}(R)$ an ideal of R. If $IM = M$, then $M = 0$.*

Proof. Putting $f = 1_M$ in the previous proposition yields elements $r_i \in I$ such that $1 + r_{n-1} + \cdots + r_0$ acts as zero on M. Set $x = r_{n-1} + \cdots + r_0 \in I$. Then $(1 + x)m = 0$ for all $m \in M$, and since $1 + x$ is a unit, we deduce that $m = 0$. Hence $M = 0$. $\square$

5 Exact sequences

A sequence of R-modules and R-module homomorphisms

$$\cdots \longrightarrow M_{n+1} \xrightarrow{d_{n+1}} M_n \xrightarrow{d_n} M_{n-1} \longrightarrow \cdots$$

is said to be **exact in position** n provided $\mathrm{Ker}(d_n) = \mathrm{Im}(d_{n+1})$, and is called **exact** if it is exact in each position.

For example, the sequence

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

is exact if and only if f is injective; dually the sequence

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

is exact if and only if g is surjective.

A **short exact sequence** is an exact sequence of the form

$$0 \to L \xrightarrow{f} M \xrightarrow{g} N \to 0,$$

which is equivalent to saying that f is injective, g is surjective and $\mathrm{Ker}(g) = \mathrm{Im}(f)$.

Recall that for each module homomorphism $f\colon M \to N$ and each module X we obtain two module homomorphisms

$$f_*\colon \mathrm{Hom}_R(X, M) \to \mathrm{Hom}_R(X, N) \quad \text{and} \quad f^*\colon \mathrm{Hom}_R(N, X) \to \mathrm{Hom}_R(M, X).$$

Proposition 5.1. *(1) The sequence*

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

is exact if and only if, for each X, the following sequence is exact

$$0 \to \mathrm{Hom}(X, L) \xrightarrow{f_*} \mathrm{Hom}(X, M) \xrightarrow{g_*} \mathrm{Hom}(X, N).$$

(2) Dually, the sequence

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

is exact if and only if, for each X, the following sequence is exact

$$0 \to \mathrm{Hom}(N, X) \xrightarrow{g^*} \mathrm{Hom}(M, X) \xrightarrow{f^*} \mathrm{Hom}(L, X).$$

Proof. Exercise. □

Remark. For the first sequence, this says in particular that if $h\colon X \to M$ satisfies $gh = 0$, then there exists a unique $\bar{h}\colon X \to L$ with $h = \bar{h}f$. Thus h factors uniquely through f; in other words, f is a kernel for g.

Also, the 'only if' part of the proposition can be rephrased by saying that both the (covariant) endofunctor $\mathrm{Hom}(X, -)$ and the (contravariant) endofunctor $\mathrm{Hom}(-, X)$ on the category of R-modules are left-exact.

Proposition 5.2 (Snake Lemma)**.** *Consider a commutative diagram of R-modules and R-module homomorphisms*

$$\begin{array}{ccccccccc} & & L & \xrightarrow{f} & M & \xrightarrow{g} & N & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle \lambda} & & \downarrow{\scriptstyle \mu} & & \downarrow{\scriptstyle \nu} & & \\ 0 & \longrightarrow & L' & \xrightarrow{f'} & M' & \xrightarrow{g'} & N' & & \end{array}$$

and suppose that both rows are exact. Then there exists an exact sequence

$$\mathrm{Ker}(\lambda) \to \mathrm{Ker}(\mu) \to \mathrm{Ker}(\nu) \xrightarrow{\delta} \mathrm{Coker}(\lambda) \to \mathrm{Coker}(\mu) \to \mathrm{Coker}(\nu).$$

If f is injective, then so too is $\mathrm{Ker}(\lambda) \to \mathrm{Ker}(\mu)$*. If g' is surjective, then so too is* $\mathrm{Coker}(\mu) \to \mathrm{Coker}(\nu)$*.*

Proof. The maps between the kernels are given by the restrictions of f and g. For example, if $m \in \mathrm{Ker}(\mu)$, then $\nu g(m) = g'\mu(m) = 0$, so $g(m) \in \mathrm{Ker}(\nu)$. This restriction is clearly R-linear. Moreover, we have exactness at $\mathrm{Ker}(\mu)$. For, if $m \in \mathrm{Ker}(\mu)$ satisfies $g(m) = 0$, then $m = f(l)$ for some $l \in L$, in which case $f'\lambda(l) = \mu f(l) = 0$. Since f' is injective, we see that $\lambda(l) = 0$, so that $l \in \mathrm{Ker}(\lambda)$. It is also clear that if f is injective, then its restriction to $\mathrm{Ker}(\lambda)$ is injective.

Dually, the maps between cokernels are induced by f' and g'. For example, if $\pi\colon M' \to \mathrm{Coker}(\mu)$ is the natural map, then $\pi f'\lambda = \pi\mu f = 0$, so $\mathrm{Im}(\lambda) \leq \mathrm{Ker}(\pi f')$ and $\pi f'$ factors through $\mathrm{Coker}(\lambda)$ by the Factor Lemma. This gives the required map $\mathrm{Coker}(\lambda) \to \mathrm{Coker}(\mu)$. Again, the induced map is R-linear, we have exactness at $\mathrm{Coker}(\mu)$, and if g' is surjective, then the induced map on cokernels is surjective.

We now define the connecting homomorphism δ. Given $n \in \mathrm{Ker}(\nu)$, we can write $n = g(m)$ for some $m \in M$, in which case $g'\mu(m) = \nu g(m) = 0$. Thus there exists a (unique) $l' \in L'$ such that $\mu(m) = f'(l')$. We set $\delta(n) := l' + \mathrm{Im}(\lambda)$.

We first need to check that δ is well-defined, so suppose we have two pairs (m_1, l_1') and (m_2, l_2') in $M \times L'$ satisfying $g(m_i) = n$ and $f'(l_i') = \mu(m_i)$. We need to check that $l_1' + \mathrm{Im}(\lambda) = l_2' + \mathrm{Im}(\lambda)$, equivalently that $l_1' - l_2' \in \mathrm{Im}(\lambda)$.

Set $m := m_1 - m_2$ and $l' := l_1' - l_2'$. Then $g(m) = 0$, so $m = f(l)$ for some $l \in L$, in which case

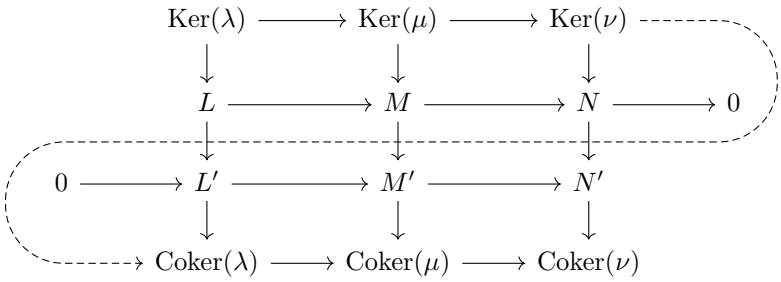
$$f'(l') = \mu(m) = \mu f(l) = f'\lambda(l).$$

Since f' is injective it follows that $l' = \lambda(l) \in \mathrm{Im}(\lambda)$ as required.

We check exactness at $\mathrm{Ker}(\nu)$. Suppose $n = g(m)$ for some $m \in \mathrm{Ker}(\mu)$. Then we have the pair $(m, 0)$, and hence $\delta(n) = 0$. Conversely, suppose that $\delta(n) = 0$, so we have some pair $(m, \lambda(l))$. Then we could instead take the pair $(m - f(l), 0)$, which satisfies $m - f(l) \in \mathrm{Ker}(\mu)$. Thus n lies in the image of $\mathrm{Ker}(\mu)$.

The proof of exactness at $\mathrm{Coker}(\lambda)$ is dual. □

The name comes from the connecting morphism snaking its way through the diagram



5.1 Projective and injective modules

A module P is called **projective** provided that for every short exact sequence

$$0 \to L \to M \to N \to 0,$$

the corresponding sequence

$$0 \to \mathrm{Hom}(P, L) \to \mathrm{Hom}(P, M) \to \mathrm{Hom}(P, N) \to 0$$

is again exact. In other words, the functor $\mathrm{Hom}(P, -)$ is **exact**.

Dually, a module I is called **injective** provided that for every short exact sequence

$$0 \to L \to M \to N \to 0,$$

the corresponding sequence

$$0 \to \mathrm{Hom}(N, I) \to \mathrm{Hom}(M, I) \to \mathrm{Hom}(L, I) \to 0$$

is again exact. In other words, the functor $\mathrm{Hom}(-, I)$ is **exact**.

Remark. We have already seen that the functor $\mathrm{Hom}(P, -)$ is left-exact, so P will be projective if and only if it preserves epimorphisms: if $g\colon M \twoheadrightarrow N$ is surjective, then $g_*\colon \mathrm{Hom}(P, M) \twoheadrightarrow \mathrm{Hom}(P, N)$ is surjective.

Dually, I is injective provided $f^*\colon \mathrm{Hom}(M, I) \twoheadrightarrow \mathrm{Hom}(L, I)$ is surjective whenever $f\colon L \rightarrowtail M$ is injective.

Note that exact functors preserve all exact sequences, not just short exact sequences. So, if we have an exact sequence

$$\cdots \longrightarrow M_{n+1} \xrightarrow{d_{n+1}} M_n \xrightarrow{d_n} M_{n-1} \longrightarrow \cdots$$

and if P is projective, then we obtain another exact sequence

$$\cdots \longrightarrow \mathrm{Hom}_R(P, M_{n+1}) \xrightarrow{(d_{n+1})_*} \mathrm{Hom}_R(P, M_n) \xrightarrow{(d_n)_*} M_{n-1} \longrightarrow \cdots$$

Part II

Tensor Products

Tensor products of modules arise in many situations. They allow one to 'extend scalars' along a ring homomorphism $R \to S$, and thus construct an S-module from each R-module; they occur as coproducts of commutative rings, which in geometric terms corresponds to taking products of varieties or schemes; as a special case, we can construct the fibre of a map between varieties or schemes.

Tensor products also arise when one studies bilinear maps from a pair of modules to another, or more generally multilinear maps. One can then take quotients and submodules of the tensor product to obtain the symmetric product, or exterior product, or divided power.

Many texts introduce tensor products in the setting of bilinear maps; here we will instead construct them as a left adjoint to the hom functor.

6 Adjoint functors

Let R and S be two rings. A **functor** G from the category of S-modules to the category of R-modules is given by

- an R-module $G(N)$ for each S-module N, and
- a linear map $G\colon \mathrm{Hom}_S(N, N') \to \mathrm{Hom}_R(G(N), G(N'))$,

such that $G(\mathrm{id}_N) = \mathrm{id}_{G(N)}$, and respecting composition $G(gg') = G(g)G(g')$.

If G' is another such functor, then a **natural transformation** $\alpha\colon G \Rightarrow G'$ consists of an R-linear map $\alpha_N\colon G(N) \to G'(N)$ for all S-modules N, such that for each S-linear $g\colon N \to N'$, the following square commutes

$$\begin{array}{ccc} G(N) & \xrightarrow{\alpha_N} & G'(N) \\ \Big\downarrow{\scriptstyle G(g)} & & \Big\downarrow{\scriptstyle G'(g)} \\ G(N') & \xrightarrow{\alpha_{N'}} & G'(N'). \end{array}$$

We say that α is a **natural isomorphism** provided each α_N is an isomorphism, in which case the collection of maps α_N^{-1} yields a natural transformation $\alpha^{-1}\colon G' \Rightarrow G$.

An **adjoint pair** of functors (F, G) consists of a functor F from R-modules to S-modules, a functor G from S-modules to R-modules, toegther with a collection of isomorphisms (of abelian groups)

$$\Phi_{M,N}\colon \mathrm{Hom}_S(F(M), N) \xrightarrow{\sim} \mathrm{Hom}_R(M, G(N))$$

for all R-modules M and S-modules N. These must satisfy the following naturalness axiom: for all R-linear $f\colon M \to M'$ and S-linear $g\colon N \to N'$, the following square

commutes

$$\begin{array}{ccc} \mathrm{Hom}_S(F(M'),N) & \xrightarrow{\Phi_{M',N}} & \mathrm{Hom}_R(M',G(N)) \\ \Big\downarrow{\scriptstyle g_*\circ F(f)^*} & & \Big\downarrow{\scriptstyle G(g)_*\circ f^*} \\ \mathrm{Hom}_S(F(M),N') & \xrightarrow{\Phi_{M,N'}} & \mathrm{Hom}_R(M,G(N')). \end{array}$$

In other words, for all S-linear $\theta\colon F(M') \to N$, we have $\Phi(g\theta F(f)) = G(g)\Phi(\theta)f$ as R-linear maps $M \to G(N')$.

We say that F is a **left adjoint** to G, and G is a **right adjoint** to F.

Given an adjoint pair, we define

$$\eta_M := \Phi(\mathrm{id}_{F(M)})\colon M \to GF(M) \quad \text{and} \quad \varepsilon_N := \Phi^{-1}(\mathrm{id}_{G(N)})\colon FG(N) \to N$$

Lemma 6.1. *Given an adjoint pair (F,G), the collections of maps η_M and ε_N yield natural transformations*

$$\eta\colon \mathrm{id}_R \Rightarrow GF \quad \textit{and} \quad \varepsilon\colon FG \Rightarrow \mathrm{id}_S,$$

*called, respectively, the **unit** and **counit** of the adjunction.*

Proof. Exercise. □

Note that for all $g\colon F(M) \to N$ we have

$$\Phi(g) = \Phi(g \cdot \mathrm{id}_{F(M)}) = G(g)\Phi(\mathrm{id}_{F(M)}) = G(g)\eta_M.$$

Similarly, for all $f\colon M \to G(N)$ we have $\Phi(\varepsilon_N F(f)) = f$.

A left adjoint comes equipped with a lot of structure, so it is not surprising that any given functor G may not have an adjoint, but if it does, then it is essentially unique. What is maybe surprising is their ubiquity: adjoint functors arise everywhere.[8]

Lemma 6.2. *If (F,G) and (F',G') be two adjoint pairs. Given a natural isomorphism $\beta\colon G \Rightarrow G'$, there is a natural isomorphism $\alpha\colon F' \Rightarrow F$ given by $\Phi'(\alpha_M) := \beta_{F(M)}\eta_M$.*

Proof. Consider the morphism $\beta_{F(M)}\eta_M\colon M \to G'F(M)$. Since Φ' is an isomorphism, there is a unique $\alpha_M\colon F'(M) \to F(M)$ for which $\beta_{F(M)}\eta_M = \Phi'(\alpha_M)$. These fit together to form a natural transformation. For, given $f\colon M \to M'$, we use that η and β are natural transformations to check

$$\Phi'(\alpha_M F'(f)) = \beta_{F(M)}\eta_M f = \beta_{F(M)}GF(f)\eta_{M'} = G'F(f)\beta_{F(M')}\eta_{M'} = \Phi'(F(f)\alpha_{M'}).$$

Again, Φ' is an isomorphism, so $\alpha_M F'(f) = F(f)\alpha_{M'}$ as required.

Analogously, there is a natural transformation $\bar{\alpha}\colon F \Rightarrow F'$, where $\Phi(\bar{\alpha}_M) = \beta^{-1}_{F'(M)}\eta'_M$. Finally, we check that these maps are mutually inverse. We have

$$\begin{aligned} \Phi'(\bar{\alpha}_M\alpha_M) &= G'(\bar{\alpha}_M)\beta_{F(M)}\eta_M = \beta_{F'(M)}G(\bar{\alpha}_M)\eta_M \\ &= \beta_{F'(M)}\Phi(\bar{\alpha}_M) = \eta'_M = \Phi'(\mathrm{id}_{F(M)}). \end{aligned}$$

Again, since Φ' is an isomorphism, we deduce that $\bar{\alpha}_M\alpha_M = \mathrm{id}_{F(M)}$. Similarly $\alpha_M\bar{\alpha}_M = \mathrm{id}_{F'(M)}$, finishing the proof. □

[8] Saunders Mac Lane, Categories for the Working Mathematician.

7 Tensor products

We will define the tensor product as a left adjoint functor.

We first introduce the concept of a bimodule. Let R and S be rings. Then an **S-R-bimodule** is an abelian group X which is simultaneously an S-module via $(s,x) \mapsto sx$, and an R-module via $(x,r) \mapsto xr$, and such that these two actions commute, so $(sx)r = s(xr)$.

It is convenient here to write the actions on different sides, so X will be a right R-module. Since R is commutative, every left R-module M is naturally a right R-module via $m \cdot r := rm$.

Now, if N is any S-module, then it is straightforward to check that $\mathrm{Hom}_S(X,N)$ is naturally an R-module via $rh\colon X \to N$, $x \mapsto h(xr)$. Moreover, if $g\colon N \to N'$, then $G(g) = g_*\colon \mathrm{Hom}_S(X,N) \to \mathrm{Hom}_S(X,N')$ is R-linear. Thus we can regard $G := \mathrm{Hom}_S(X,-)$ as a functor from S-modules to R-modules.

We will construct the tensor product $F(M) := X \otimes_R M$ as a left adjoint to G.

7.1 The construction

Given an R-module M, choose a generating set m_i. This gives an epimorphism $R^{(I)} \twoheadrightarrow M$, $(r_i) \mapsto \sum_i r_i m_i$. Let U be the kernel, so consisting of those (r_i) for which $\sum_i r_i m_i = 0$.

Next let $\bar{U} \leq X^{(I)}$ be the S-submodule generated by all (xr_i) for $x \in X$ and $(r_i) \in U$. We define the **tensor product** $X \otimes_R M$ to be the quotient $X^{(I)}/\bar{U}$, so that $X \otimes_R M$ is naturally an S-module.

We now describe certain canonical elements in $X \otimes_R M$. Given $x \in X$ and $m \in M$, we can write $m = \sum_i r_i m_i$, and define $x \otimes m$ to be the image of $\sum_i (xr_i) \in X^{(I)}$. This is well-defined, since if $m = \sum_i r_i' m_i$ is another expression, then $\sum_i (r_i - r_i')m_i = 0$, so $(r_i, r_i') \in U$, and hence $(x(r_i - r_i')) \in \hat{U}$.

We observe that $x \otimes (rm + m') = xr \otimes m + x \otimes m'$, and similarly $(sx + x') \otimes m = s(x \otimes m) + x' \otimes m$. It follows that, for a fixed $m \in M$, the assignment $x \mapsto x \otimes m$ is an S-linear map $X \to X \otimes_R M$. Moreover, these fit together to yield an R-linear map $\eta_M\colon M \to \mathrm{Hom}_S(X, X \otimes_R M)$. This will in fact be the unit of our adjunction.

Finally, note that every element of $X \otimes_R M$ is a finite sum of such elements $x \otimes m$. In particular, two maps $X \otimes_R M \to N$ agree if and only if they agree on all $x \otimes m$.

7.2 The hom space isomorphisms

We need to define the isomorphisms

$$\Phi\colon \mathrm{Hom}_S(X \otimes_R M, N) \xrightarrow{\sim} \mathrm{Hom}_R(M, \mathrm{Hom}_S(X,N)).$$

Actually, the inverse map is easier to describe: it sends $\theta\colon M \to \mathrm{Hom}_S(X,N)$ to the map $\hat{\theta}(x \otimes m) := \theta(m)(x)$.

Using Proposition 5.1 there is an exact sequence for all S-modules N

$$0 \to \mathrm{Hom}_S(X \otimes_R M, N) \to \mathrm{Hom}_S(X^{(I)}, N) \to \mathrm{Hom}_S(\bar{U}, N).$$

In other words, we can identify homomorphisms $X \otimes_R M \to N$ with tuples $\theta_i \colon X \to N$ such that (θ_i) vanishes on $\bar{U}$, equivalently $\sum_i \theta_i(xr_i) = 0$ for all $x \in X$ and $(r_i) \in U$.

On the other hand, we also have an exact sequence

$$0 \to \mathrm{Hom}_R(M, \mathrm{Hom}_S(X,N)) \to \mathrm{Hom}_R(R^{(I)}, \mathrm{Hom}_S(X,N)) \to \mathrm{Hom}_R(U, \mathrm{Hom}_S(X,N))$$

In other words, we can identify homomorphisms $M \to \mathrm{Hom}_S(X,N)$ with tuples $\phi_i \colon R \to \mathrm{Hom}_S(X,N)$ such that $\sum_i \phi_i(r_i) = 0$ for all $(r_i) \in U$.

As usual, we have the isomorphism (of abelian groups)

$$\mathrm{Hom}_R(R, \mathrm{Hom}_S(X,N)) \xrightarrow{\sim} \mathrm{Hom}_S(X,N), \quad \phi \mapsto \phi(1).$$

Note that $\phi(r)(x) = (r\phi(1))(x) = \phi(1)(xr)$. This extends to an isomorphism

$$\mathrm{Hom}_R(R^{(I)}, \mathrm{Hom}_S(X,N)) \xrightarrow{\sim} \mathrm{Hom}_S(X^{(I)}, N),$$

sending (ϕ_i) to the morphism $(x_i) \mapsto \sum_i \phi_i(1)(x_i)$. Thus $(\phi_i)((r_i)) = (\phi_i(r_i)) \colon X \to N$ is the map $x \mapsto \sum_i \phi_i(r_i)(x) = \sum_i \phi_i(1)(xr_i)$. Thus

$$(\phi_i)(U) = 0 \quad \Leftrightarrow \quad \sum_i \phi_i(r_i)(x) = 0 \text{ for all } (r_i) \in U,\, x \in X \quad \Leftrightarrow \quad (\phi_i(1))(\bar{U}) = 0.$$

We conclude that the isomorphism

$$\mathrm{Hom}_R(R^{(I)}, \mathrm{Hom}_S(X,N)) \xrightarrow{\sim} \mathrm{Hom}_S(X^{(I)}, N)$$

restricts to an isomorphism on their respective subgroups

$$\mathrm{Hom}_R(M, \mathrm{Hom}_S(X,N)) \xrightarrow{\sim} \mathrm{Hom}_S(X \otimes_R M, N).$$

Note that this indeed sends θ to the map $\hat{\theta} \colon x \otimes m \mapsto \theta(m)(x)$ as claimed.

7.3 The adjointness property

We need that $F = X \otimes_R -$ is a functor, so we need to define $F(f)$ for morphisms $f \colon M \to M'$. Given $m \in M$, consider the map $X \to X \otimes_R M'$, $x \mapsto f(m)$. As above, these determine an R-linear map $M \to \mathrm{Hom}_S(X, X \otimes_R M')$, and under Φ this corresponds to an S-linear map $F(f) \colon X \otimes_R M \to X \otimes_R M'$, $x \otimes m \mapsto x \otimes f(m)$.

It is then clear that $F(f+f') = F(f) + F(f')$ and $F(f'f) = F(f')F(f)$. For, we just need to compare their actions on elements of the form $x \otimes m$. Thus F is indeed a functor.

We also need to check that Φ satisfies the naturalness axiom. Actually we will do this for the inverse map $\theta \mapsto \hat{\theta}$. Take maps $f \colon M \to M'$, $g \colon N \to N'$, and $\theta \colon M' \to \mathrm{Hom}_S(X,N)$. We need to show that $g\hat{\theta}F(f) = \widehat{g_*\theta f}$ as maps $X \otimes_R M \to N'$. Again, we compare their actions on elements of the form $x \otimes m$. We have

$$g\hat{\theta}F(f)(x \otimes m) = g\hat{\theta}(x \otimes f(m)) = g\big(\theta(f(m))(x)\big) = \big(g_*\theta(f(m))\big)(x) = \widehat{g_*\theta f}(x \otimes m).$$

This completes the proof of the following theorem.

Theorem 7.1. *Given an S-R-bimodule X, the hom functor $G(N) := \mathrm{Hom}_S(X, N)$ has a left adjoint, given by the tensor functor $F(M) := X \otimes_R M$. In particular, we have the natural isomorphism*

$$\mathrm{Hom}_R(M, \mathrm{Hom}_S(X, N)) \xrightarrow{\sim} \mathrm{Hom}_S(X \otimes_R M, N), \quad \theta \mapsto \big(\hat{\theta}\colon x \otimes m \mapsto \theta(m)(x)\big).$$

It follows that the unit of the adjunction is given by

$$\eta_M\colon M \to \mathrm{Hom}_S(X, X \otimes_R M), \quad m \mapsto (x \mapsto x \otimes m)$$

and the counit is given by

$$\varepsilon_N\colon X \otimes_R \mathrm{Hom}_S(X, N) \to N, \quad x \otimes f \mapsto f(x).$$

7.4 Examples

Suppose $R = S = K$ for some field K. Let V and W be two vector spaces, say with bases v_i and w_j respectively. Then $V \otimes_K W$ is the vector space with basis $v_i \otimes w_j$. In particular, $\dim(V \otimes_K W) = (\dim V)(\dim W)$.

For, we know that $W \cong K^{(J)}$ is free, so the construction gives $V^{(J)} \xrightarrow{\sim} V \otimes_K W$. Now $V^{(J)}$ has basis $v_{i,j}$ having v_i in component j, and this is sent to $v_i \otimes w_j$ in $V \otimes W$.

Next take $R = S = \mathbb{Z}$, and consider the abelian groups $M = \mathbb{Z}/m\mathbb{Z}$ and $N = \mathbb{Z}/n\mathbb{Z}$. Then $M \otimes_{\mathbb{Z}} N \cong \mathbb{Z}/d\mathbb{Z}$ where $d = \gcd(m, n)$.

For, we write N as the quotient of $\mathbb{Z}$ by the submodule $n\mathbb{Z}$. Then the construction gives $M \otimes_{\mathbb{Z}} N$ as the quotient of M by the submodule nM. Now $d = am + bn$ and $n = dn'$, so $nM = dM$ and the Third Isomorphism Theorem gives $M/dM \cong \mathbb{Z}/d\mathbb{Z}$.

7.5 Basic properties

We now prove some basic isomorphisms for tensor products. The first three can all be regarded as instances of the uniqueness result for adjoint pairs, Lemma 6.2. They hold for all rings R, not just for commutative rings.

When the ring R is commutative, then every left R-module M is naturally a right R-module via $m \cdot r := rm$, and hence M is even an R-bimodule. We can therefore define both $M \otimes_R N$ and $N \otimes_R M$ for all R-modules M and N. The fourth property then says that these are isomorphic as R-modules.

Proposition 7.2. *We have the following natural isomorphisms.*

(1) $M \otimes_R R \xrightarrow{\sim} M$, *sending* $m \otimes r \mapsto mr$.

(2) $\big(\coprod_i X_i\big) \otimes_R M \xrightarrow{\sim} \coprod_i (X_i \otimes M)$, *sending* $(x_i) \otimes m \mapsto (x_i \otimes m)$.

(3) $(Y \otimes_S X) \otimes_R M \xrightarrow{\sim} Y \otimes_S (X \otimes_R M)$, *sending* $(y \otimes x) \otimes m \mapsto y \otimes (x \otimes m)$.

(4) For R commutative we have $M \otimes_R N \xrightarrow{\sim} N \otimes_R M$, *sending* $m \otimes n \mapsto n \otimes m$.

Proof. (1) Since R is free our construction immediately gives an isomorphism $M \xrightarrow{\sim} M \otimes_R R$, $m \mapsto m \otimes 1$. The inverse is $m \otimes r \mapsto mr$.

Alternatively, we have the natural isomorphism $\mathrm{Hom}_R(R,N) \xrightarrow{\sim} N$, $f \mapsto f(1)$. So we obtain a natural isomorphism between their adjoints as in Lemma 6.2, which in this case yields $M \xrightarrow{\sim} M \otimes_R R$, $m \mapsto m \otimes 1$.

(2) In this case both functors are left adjoint to $G\colon N \mapsto \prod_i \mathrm{Hom}_R(X_i, N)$. Note that a map $M \to G(N)$ is given by a collection of maps (θ_i) where $\theta_i\colon M \to \mathrm{Hom}_R(X_i, N)$.

Under the adjunction

$$\mathrm{Hom}_R(M, G(N)) \cong \mathrm{Hom}_R(M, \mathrm{Hom}_R(\coprod_i X_i, N)) \cong \mathrm{Hom}_R\left(\left(\coprod_i X_i\right) \otimes_R M, N\right)$$

we see that (θ_i) corresponds to the map $\left(\coprod_i X_i\right) \otimes_R M \to N$, $(x_i) \otimes m \mapsto \sum_i \theta_i(m)(x_i)$. This is well-defined, since only finitely many x_i are non-zero.

Under the other adjunction

$$\begin{aligned}\mathrm{Hom}_R(M, G(N)) &\cong \prod_i \mathrm{Hom}_R(M, \mathrm{Hom}_R(X_i, N)) \\ &\cong \prod_i \mathrm{Hom}_R(X_i \otimes_R M, N) \cong \mathrm{Hom}_R\left(\coprod_i (X_i \otimes_R M), N\right)\end{aligned}$$

we see that (θ_i) corresponds to the map $\coprod_i (X_i \otimes_R M) \to N$, $(x_i \otimes m_i) \mapsto \sum_i \theta_i(m_i)(x_i)$.

The isomorphism is therefore given by taking $N := \coprod_i (X_i \otimes_R M)$ and tracing through to determine where the identity map is sent. In this case, the identity corresponds to the tuple (θ_i), where $\theta_i(m)(x_i) := \iota_i(x_i \otimes m)$, which then yields the map

$$\left(\coprod_i X_i\right) \otimes_R M \xrightarrow{\sim} \coprod_i (X_i \otimes_R M), \quad (x_i) \otimes m \mapsto \sum_i \iota_i(x_i \otimes m) = (x_i \otimes m).$$

The inverse map $(x_i \otimes m_i) \mapsto \sum_i \iota_i(x_i) \otimes m_i$ is constructed similarly.

(3) Here we take a T-S-bimodule Y, and observe that both functors are left adjoint to $G\colon N \mapsto \mathrm{Hom}_S(X, \mathrm{Hom}_T(Y, N))$. The corresponding natural isomorphism thus corresponds to the unit

$$\eta\colon M \to G(Y \otimes_S (X \otimes_R M)), \quad \eta(m)(x)(y) := y \otimes (x \otimes m),$$

so is given by $(y \otimes x) \otimes m \mapsto y \otimes (x \otimes m)$.

(4) For fixed $m \in M$ we have the linear map $\theta_m\colon N \mapsto M \otimes_R N$, $n \mapsto m \otimes n$. Moreover, the map $m \mapsto \theta_m$ is clearly linear, and using that R is commutative we check that it is even R-linear:

$$(r\theta_m)(n) := \theta_m(rn) = m \otimes rn = rm \otimes n = \theta_{rm}(n).$$

We thus obtain a linear map $N \otimes_R M \to M \otimes_R N$, $n \otimes m \mapsto m \otimes n$.

Analogously we have the map $m \otimes n \mapsto n \otimes m$, and these are mutually inverse. □

We can rephrase these basic properties as saying that the category of R-modules is a **symmetric monoidal category** with respect to the tensor product.[9]

[9] Another important situation where we have a symmetric monoidal category is when we take a group G, a field K, and consider the category of K-representations of G. Thus an object is a K-vector space V together with an action of G on V; equivalently a group homomorphism $G \to \mathrm{Aut}_K(V)$. Then $V \otimes_K W$ is a G-representation by taking the diagonal action $g \cdot v \otimes w := (gv) \otimes (gw)$. In this case we have the adjunction $\mathrm{Hom}_G(V \otimes_K W, X) \cong \mathrm{Hom}_G(V, \mathrm{Hom}_K(W, X))$, where $\mathrm{Hom}_K(W, X)$ is a G-representation via $(g \cdot \theta)(w) := g\theta(g^{-1}w)$.

Lemma 7.3. *Let $f\colon M \to M'$ and $g\colon N \to N'$ be homomorphisms of R-modules. Then there is an R-linear map $f \otimes g\colon M \otimes N \to M' \otimes N'$, $m \otimes n \mapsto f(m) \otimes f(n)$.*

Proof. Since $M \otimes_R -$ is a functor we know that there is a map $\mathrm{id}_M \otimes g\colon M \otimes_R N \to M \otimes_R N'$, $m \otimes n \mapsto m \otimes g(n)$. Using that the tensor product is symmetric, we also get a map $f \otimes \mathrm{id}_{N'}\colon M \otimes_R N' \to M' \otimes_R N'$, $m \otimes n' \mapsto f(m) \otimes n'$. The composition of these is then the required map $f \otimes g$. □

7.6 Right exactness and flat modules

Proposition 7.4. *Let X be an S-R-bimodule. Then the tensor functor $X \otimes_R -$ is right exact. In other words, it sends an exact sequence of R-modules*

$$L \xrightarrow{f} M \xrightarrow{g} N \longrightarrow 0$$

to the exact sequence of S-modules

$$X \otimes_R L \xrightarrow{1\otimes f} X \otimes_R M \xrightarrow{1\otimes g} X \otimes_R N \longrightarrow 0.$$

Proof. Let Y be any S-module. Then the sequence

$$0 \to \mathrm{Hom}_S(X \otimes_R N, Y) \to \mathrm{Hom}_S(X \otimes_R M, Y) \to \mathrm{Hom}_S(X \otimes_R L, Y)$$

is isomorphic to the sequence

$$0 \to \mathrm{Hom}_R(N, \mathrm{Hom}_S(X, Y)) \to \mathrm{Hom}_R(M, \mathrm{Hom}_S(X, Y)) \to \mathrm{Hom}_R(L, \mathrm{Hom}_S(X, Y))$$

which we know is exact by Proposition 5.1. Since this holds for all S-modules Y, the same proposition tells us that the we have an exact sequence

$$X \otimes_R L \to X \otimes_R M \to X \otimes_R N \to 0.$$ □

Remark. The same proof shows that if (F, G) is an adjoint pair of functors, then F is right exact, and the dual proof gives that G is left exact.

We say that an R-module X is **flat** provided the functor $X \otimes_R -$ is exact. Thus X is a flat R-module if and only if it preserves monomorphisms: $f\colon L \rightarrowtail M$ injective implies $1 \otimes f\colon X \otimes_R L \rightarrowtail X \otimes_R M$ injective.

We say that X is **faithfully flat** if a sequence $0 \to L \to M \to N \to 0$ is exact if and only if $0 \to X \otimes_R L \to X \otimes_R M \to X \otimes_R N \to 0$ is exact.

Lemma 7.5. *(1) Let X_i be a family of R-modules. Then each X_i is flat if and only if their coproduct $\coprod_i X_i$ is flat.*

(2) The regular module R is faithfully flat.

(3) Each projective module is flat.

Proof. (1) Consider a map $L \to M$. Then each $X_i \otimes L \to X_i \otimes M$ is injective if and only if $\coprod_i (X_i \otimes_R L) \to \coprod_i (X_i \otimes_R M)$ is injective. The result now follows from the natural isomorphism $\coprod_i (X_i \otimes_R L) \cong \left(\coprod_i X_i\right) \otimes_R L$.

(2) This follows immediately from the natural isomorphism $R \otimes_R L \cong L$.

(3) The first two parts show that every direct summand of a free module is flat. □

8 Extension of scalars

Let $\theta\colon R \to S$ be a ring homomorphism. Then S is naturally an R-module, via $s \cdot r := s\theta(r)$, and so we can regard S as an S-R-bimodule.

If N is an S-module, then $N \cong \mathrm{Hom}_S(S, N)$ is naturally an R-module, via $r{\cdot}n := \theta(r)n$. This is called **restriction of scalars**.

Conversely, if M is an R-module, then the tensor product $S \otimes_R M$ is naturally an S-module, via $s \cdot (t \otimes m) := (st) \otimes m$. The functor $M \mapsto S \otimes_R M$ is called **extension of scalars**, or **induction**.

The tensor-hom-adjunction thus becomes

$$\mathrm{Hom}_S(S \otimes_R M, N) \cong \mathrm{Hom}_R(M, N).$$

This is also referred to as **Frobenius Reciprocity**, especially in group theory.

Explicitly, an S-linear map $g\colon S \otimes_R M \to N$ is sent to the R-linear map $M \to N$, $m \mapsto g(1_S \otimes m)$. Conversely, an R-linear map $f\colon M \to N$ is sent to the S-linear map $S \otimes_R M \to N$, $s \otimes m \mapsto sf(m)$.

Example. We have a ring homomorphism $\mathbb{R} \to \mathbb{C}$. Thus, if V is an $\mathbb{R}$-vector space, then $\mathbb{C} \otimes_{\mathbb{R}} V$ is a $\mathbb{C}$-vector space. This process is called **complexification**. Note that $\dim_{\mathbb{R}} V = \dim_{\mathbb{C}} \mathbb{C} \otimes_{\mathbb{R}} V$. On the other hand, if W is a $\mathbb{C}$-vector space, then we can restrict scalars to view it as an $\mathbb{R}$-vector space, in which case $\dim_{\mathbb{R}} W = 2 \dim_{\mathbb{C}} W$.

Lemma 8.1. *Let $I \lhd R$ be an ideal and M an R-module. Then $(R/I) \otimes_R M \cong M/IM$.*

Proof. Starting from $0 \to I \to R \to R/I \to 0$, we use that the tensor product is right exact to get an exact sequence $I \otimes_R M \to R \otimes_R M \to (R/I) \otimes_R M \to 0$. Now $R \otimes_R M \cong M$ via $r \otimes m \mapsto rm$, in which case the map $I \otimes_R M \to M$ sends $x \otimes m \mapsto xm$, so has image the submodule IM. □

Lemma 8.2. *Let $R \to S$ be a ring homomorphism, and M and N two R-modules. Then there is a canonical isomorphism*

$$(S \otimes_R M) \otimes_S (S \otimes_R N) \cong S \otimes_R (M \otimes_R N), \quad (s \otimes m) \otimes (t \otimes n) \mapsto (st) \otimes (m \otimes n).$$

Proof. Using Frobenius Reciprocity, we have the sequence of natural isomorphisms

$$\begin{aligned} \mathrm{Hom}_S\big((S \otimes_R M) \otimes_S (S \otimes_R N), X\big) &\cong \mathrm{Hom}_S\big(S \otimes_R M, \mathrm{Hom}_S(S \otimes_R N, X)\big) \\ &\cong \mathrm{Hom}_R(M, \mathrm{Hom}_R(N, X)) \cong \mathrm{Hom}_R(M \otimes_R N, X) \\ &\cong \mathrm{Hom}_S(S \otimes_R (M \otimes_R N), X). \end{aligned}$$

Explicitly, this sends an R-linear map $\phi\colon M \otimes_R N \to X$ to the S-linear map $\hat{\phi}\colon (S \otimes_R M) \otimes_S (S \otimes_R N) \to X$ such that $\hat{\phi}((s \otimes m) \otimes (t \otimes n)) = st\phi(m \otimes n)$.

The isomorphism $(S \otimes_R M) \otimes_S (S \otimes_R N) \cong S \otimes_R (M \otimes_R N)$ is therefore given by taking $X := S \otimes_R (M \otimes_R N)$ and tracing through to determine where the identity map is sent. In this case, the identity corresponds to the map $\phi\colon m \otimes n \mapsto 1_S \otimes (m \otimes n)$, and hence to the map $\hat{\phi}\colon (s \otimes m) \otimes (t \otimes n) \mapsto (st) \otimes (m \otimes n)$.

The inverse map $s \otimes (m \otimes n) \mapsto (s \otimes m) \otimes (1_S \otimes n)$ is constructed similarly. □

8.1 Localisation for modules

As for rings and ideals, we can define the localisation of a module.

Let R be a ring, $\Sigma \subset R$ a multiplicatively closed subset, and M an R-module. There is an equivalence relation on $\Sigma \times M$ such that

$$(a, m) \sim (b, n) \quad \text{provided} \quad c(an - bm) = 0 \text{ for some } c \in \Sigma.$$

We denote the equivalence class of (a, m) by m/a, and write M_Σ for the set of all equivalence classes. This is naturally an R_Σ-module via

$$\frac{m}{a} + \frac{n}{b} := \frac{an + bm}{ab} \qquad \text{and} \qquad \frac{r}{a} \cdot \frac{m}{b} := \frac{rm}{ab}.$$

Lemma 8.3. *For each R-module M, the above construction yields an R_Σ-module M_Σ. Moreover, regarding M_Σ as an R-module via restriction of scalars along $\sigma\colon R \to R_\Sigma$, the natural map $\sigma\colon M \to M_\Sigma$, $m \mapsto \frac{m}{1}$, is R-linear.* □

We now show that this construction is a special case of extension of scalars.

Proposition 8.4. *There is an isomorphism of R_Σ-modules*

$$R_\Sigma \otimes_R M \xrightarrow{\sim} M_\Sigma, \quad \tfrac{r}{a} \otimes m \mapsto \tfrac{rm}{a}.$$

Proof. Using Frobenius Reciprocity we get for each R_Σ-module X an isomorphism

$$\mathrm{Hom}_{R_\Sigma}(R_\Sigma \otimes_R M, X) \cong \mathrm{Hom}_R(M, X).$$

In particular, the R-linear map $\sigma\colon M \to M_\Sigma$, $\sigma(m) = \frac{m}{1}$, gives rise to the R_Σ-linear map $f\colon R_\Sigma \otimes_R M \to M_\Sigma$, $\frac{r}{a} \otimes m \mapsto \frac{rm}{a}$.

Clearly f is surjective. To see that it is injective, we first observe that every element in $R_\Sigma \otimes_R M$ can be written as $\frac{1}{a} \otimes m$. For, a general element is of the form $x = \frac{r_1}{a_1} \otimes m_1 + \cdots + \frac{r_n}{a_n} \otimes m_n$. Set $a = a_1 \cdots a_n \in \Sigma$ and $b_i = a_1 \cdots \hat{a}_i \cdots a_n$. Then

$$x = \tfrac{r_1 b_1}{a} \otimes m_1 + \cdots + \tfrac{r_n b_n}{a_n} \otimes m_n = \tfrac{1}{a} \otimes ((r_1 b_1 m_1 + \cdots + r_n b_n m_n).$$

Now, if $\frac{1}{a} \otimes m$ lies in the kernel of f, then $\frac{m}{a} = 0$ in M_Σ, so there exists $c \in \Sigma$ with $cm = 0$ in M. Thus $\frac{1}{a} \otimes m = \frac{1}{ac} \otimes (cm) = 0$ in $R_\Sigma \otimes_R M$. □

Corollary 8.5. *There is a natural isomorphism of R_Σ-modules*

$$M_\Sigma \otimes_{R_\Sigma} N_\Sigma \cong (M \otimes_R N)_\Sigma, \quad \tfrac{m}{a} \otimes \tfrac{n}{b} \mapsto \tfrac{m \otimes n}{ab}.$$

Proof. This follows from Lemma 8.2, using the ring homomorphism $R \to R_\Sigma$. □

One of the most important properties of localisation is that it is exact. The next result is the promised generalisation of Proposition 3.6.

We first introduce some notation. Recall that, given an R-module homomorphism $f\colon M \to N$, we get an R_Σ-module homomorphism

$$1 \otimes f\colon R_\Sigma \otimes_R M \to R_\Sigma \otimes_R N, \quad \tfrac{r}{a} \otimes m \mapsto \tfrac{r}{a} \otimes f(m).$$

Using the previous proposition we can write this as

$$f_\Sigma\colon M_\Sigma \to N_\Sigma, \quad \tfrac{m}{a} \mapsto \tfrac{f(m)}{a}.$$

Proposition 8.6. *The localisation R_Σ is flat as an R-module. Thus, given an exact sequence of R-modules*

$$0 \longrightarrow L \xrightarrow{f} M \xrightarrow{g} N \longrightarrow 0,$$

we obtain an exact sequence of R_Σ-modules

$$0 \longrightarrow L_\Sigma \xrightarrow{f_\Sigma} M_\Sigma \xrightarrow{g_\Sigma} N_\Sigma \longrightarrow 0.$$

Proof. Let $f\colon L \to M$ be an injective R-module homomorphism, and suppose l/a is in the kernel of f_Σ. Then $f(l)/a = 0$ in M_Σ, so there exists $c \in \Sigma$ with $0 = cf(l) = f(cl)$ in M. Since f is injective, we must have $cl = 0$ in L, and hence that $l/a = 0$ in L_Σ. □

As a corollary we see that localisation and quotients commute.

Corollary 8.7. *Let $L \leq M$. Then $(M/L)_\Sigma \cong (M_\Sigma)/(L_\Sigma)$.* □

Lemma 8.8. *Localisation commutes with forming arbitrary sums*

$$\left(\textstyle\sum_i U_i\right)_\Sigma = \sum_i (U_i)_\Sigma \quad \text{for submodules } U_i \leq M$$

and finite intersections $\left(\bigcap_{i=1}^n U_i\right)_\Sigma = \bigcap_{i=1}^n (U_i)_\Sigma$.
Also, if $I \lhd R$ is an ideal, then $(IM)_\Sigma = (I_\Sigma)(M_\Sigma)$. □

8.2 Local properties

Let R be a ring, $\mathfrak{p} \lhd R$ a prime ideal, and M an R-module. We call a property $\mathcal{P}$ of modules **local** if M has $\mathcal{P}$ if and only if the localisation $M_\mathfrak{p}$ has $\mathcal{P}$ for all primes $\mathfrak{p}$.

Proposition 8.9. *The following are equivalent for a module M.*

(1) $M = 0$.

(2) $M_\mathfrak{p} = 0$ *for all prime ideals* $\mathfrak{p}$.

(3) $M_\mathfrak{m} = 0$ *for all maximal ideals* $\mathfrak{m}$.

In particular, being zero is a local property.

Proof. It is clear that (1) ⇒ (2) ⇒ (3).

(3) ⇒ (1). Suppose $0 \neq m \in M$ and consider $\mathrm{Ann}(m) = \{r \in R : rm = 0\}$. This is an ideal of R, and is proper since $1m = m \neq 0$. Thus it is contained in a maximal ideal $\mathfrak{m}$. It follows that $0 \neq \frac{m}{1} \in M_\mathfrak{m}$. □

Proposition 8.10. *The following are equivalent for an R-linear map $f\colon M \to N$.*

(1) f *is injective (surjective).*

(2) $f_\mathfrak{p}$ *is injective (surjective) for all prime ideals* $\mathfrak{p}$.

(3) $f_\mathfrak{m}$ *is injective (surjective) for all maximal ideals* $\mathfrak{m}$.

In particular, being a submodule (or a factor module) is a local property.

Proof. Let $K = \mathrm{Ker}(f)$, so that $0 \to K \to M \xrightarrow{f} N$ is exact. Since localisation is exact, we deduce that $K_{\mathfrak{p}} = \mathrm{Ker}(f_{\mathfrak{p}})$ for all prime ideals $\mathfrak{p}$. The result for injectivity now follows from the previous proposition applied to the module K.

The proof for surjectivity is dual. □

Proposition 8.11. *The following are equivalent for a module M.*

(1) M is a flat R-module.

(2) $M_{\mathfrak{p}}$ is a flat $R_{\mathfrak{p}}$-module for all prime ideals $\mathfrak{p}$.

(3) $M_{\mathfrak{m}}$ is a flat $R_{\mathfrak{m}}$-module for all maximal ideals $\mathfrak{m}$.

In particular, being flat is a local property.

Proof. (3) $\Rightarrow$ (1). Let X and Y be R-modules, and $X \rightarrowtail Y$ an injective map. Then $X_{\mathfrak{m}} \rightarrowtail Y_{\mathfrak{m}}$ is injective for all $\mathfrak{m}$, by the previous proposition, and so $M_{\mathfrak{m}} \otimes_{R_{\mathfrak{m}}} X_{\mathfrak{m}} \rightarrowtail M_{\mathfrak{m}} \otimes_{R_{\mathfrak{m}}} Y_{\mathfrak{m}}$ is injective for all $\mathfrak{m}$. Using Corollary 8.5 we get that $(M \otimes_R X)_{\mathfrak{m}} \rightarrowtail (M \otimes_R Y)_{\mathfrak{m}}$ is injective for all $\mathfrak{m}$, whence $M \otimes_R X \rightarrowtail M \otimes_R Y$ is injective, and so M is a flat R-module.

(2) $\Rightarrow$ (3). Easy.

(1) $\Rightarrow$ (2). Let X and Y be $R_{\mathfrak{p}}$-modules, and $X \rightarrowtail Y$ an injective map. By restriction of scalars we can view this as a monomorphism of R-modules, and since M is flat we see that $M \otimes_R X \rightarrowtail M \otimes_R Y$ is injective. Now,

$$M_{\mathfrak{p}} \otimes_{R_{\mathfrak{p}}} X \cong \left(M \otimes_R R_{\mathfrak{p}}\right) \otimes_{R_{\mathfrak{p}}} X \cong M \otimes_R \left(R_{\mathfrak{p}} \otimes_{R_{\mathfrak{p}}} X\right) \cong M \otimes_R X.$$

Thus $M_{\mathfrak{p}} \otimes_{R_{\mathfrak{p}}} X \hookrightarrow M_{\mathfrak{p}} \otimes_{R_{\mathfrak{p}}} Y$ is injective, and so $M_{\mathfrak{p}}$ is a flat $R_{\mathfrak{p}}$-module. □

In this sense we see that flat modules are better behaved than projective modules. On the other hand, we have the following results.

Given an R-module M, we can always construct a **free presentation** of M, so an exact sequence $F_1 \to F_0 \to M \to 0$ with F_0, F_1 free R-modules. We say that M is **finitely generated** if we can take F_0 to be finitely generated; **finitely related** if we can take F_1 to be finitely generated; and **finitely presented** if we can take both F_0 and F_1 to be finitely generated.

Every finitely related flat module is projective, and if R is a local ring, then every finitely generated flat module is free. Putting these together shows that if M is finitely presented such that $M_{\mathfrak{p}}$ is free for each prime $\mathfrak{p}$ (so M is locally free), then M is projective.

Conversely, Kaplansky's Theorem states that every projective module is locally free.

Also, the Quillen-Suslin Theorem states that if R is a principal ideal domain, then every projective module over $R[X_1, \ldots, X_n]$ is free.

9 Tensor products of algebras

Let k be a commutative ring. Recall that a k-algebra is a ring R together with a ring homomorphism $k \to R$. In particular, we can regard R as a k-module, in which case we have a k-linear map $\lambda\colon R \to \mathrm{End}_k(R)$ sending r to left multiplication by r, so $\lambda_r\colon r' \mapsto rr'$.

Conversely, if R is a k-module, then a k-linear map $\lambda\colon R \to \mathrm{End}_k(R)$ yields a distributive multiplication $R \times R \to R$. If this is associative and unital, then R is a ring, and we have a ring homomorphism $k \to R$, $a \mapsto a1_R$.

Note also that by the tensor-hom-adjunction, giving λ is the same as giving a k-linear map $\mu\colon R \otimes_k R \to R$.

Proposition 9.1. *The tensor product $R \otimes_k S$ is naturally a k-algebra via*

$$(r \otimes s)(r' \otimes s') := (rr') \otimes (ss'),$$

and the maps $\iota_R\colon R \to R \otimes_k S$, $r \mapsto r \otimes 1_S$, and $\iota_S\colon S \to R \otimes_k S$, $s \mapsto 1_R \otimes s$, are both k-algebra homomorphisms.

In fact, $R \otimes_k S$ is the coproduct of R and S in the category of (commutative) k-algebras. In other words, if T is any k-algebra, then

$$\mathrm{Hom}_{k\text{-}alg}(R \otimes_k S, T) \cong \mathrm{Hom}_{k\text{-}alg}(R, T) \times \mathrm{Hom}_{k\text{-}alg}(S, T), \quad f \mapsto (f\iota_R, f\iota_S).$$

Proof. Using Lemma 7.3 we see that for fixed $r \in R$ and $s \in S$, there is a well-defined k-linear map $\lambda_{r,s}\colon R \otimes_k S \to R \otimes_k S$, $r' \otimes s' \mapsto (rr') \otimes (ss')$. Next, for fixed r, the map $s \mapsto \lambda_{r,s}$ is k-linear:

$$rr' \otimes (s_1 + as_2)s' = rr' \otimes s_1 s' + a(rr' \otimes s_2 s'),$$

as is the map $r \mapsto \lambda_{r,-}$:

$$(r_1 + ar_2)r' \otimes ss' = r_1 r' \otimes ss' + a(r_2 r' \otimes ss').$$

We therefore have a k-linear map $R \to \mathrm{Hom}_k(S, \mathrm{End}_k(R \otimes_k S))$, which by the tensor-hom-adjunction yields a k-linear map $\lambda\colon R \otimes_k S \to \mathrm{End}_k(R \otimes_k S)$. It follows that there is a well-defined distributive multiplication on $R \otimes_k S$ such that

$$(r \otimes s)(r' \otimes s') = (rr') \otimes (ss').$$

It is easy to check that this is associative with unit $1_R \otimes 1_S$, and that the k-linear maps ι_R and ι_S are ring homomorphisms.

Consider now the map

$$\mathrm{Hom}_{k\text{-alg}}(R \otimes_k S, T) \to \mathrm{Hom}_{k\text{-alg}}(R, T) \times \mathrm{Hom}_{k\text{-alg}}(S, T), \quad f \mapsto (f\iota_R, f\iota_S).$$

We construct the inverse explicitly. Given a pair of algebra homomorphisms (f_R, f_S), we obtain a k-linear map

$$f\colon R \otimes_k S \to T \otimes_k T \to T, \quad r \otimes s \mapsto f_R(r) \otimes f_S(s) \mapsto f_R(r) f_S(s).$$

This is a k-algebra homomorphism: clearly $f(1_R \otimes 1_S) = 1_T$, and

$$f(rr' \otimes ss') = f_R(rr')f_S(ss') = f_R(r)f_R(r')f_S(s)f_S(s') = f(r \otimes s)f(r' \otimes s')$$

since T is commutative. □

Example. Let R be a k-algebra. Then $R \otimes_k k[X] \cong R[X]$, the ring of polynomials in X with coefficients in R. One way to see this is to note that they satisfy the same universal property:

$$\mathrm{Hom}_{k\text{-alg}}(R \otimes_k k[X], T) \cong \mathrm{Hom}_{k\text{-alg}}(R, T) \times T \cong \mathrm{Hom}_{k\text{-alg}}(R[X], T).$$

Explicitly, we have the ring isomorphism $R[X] \to R \otimes_k k[X]$, $rX^n \mapsto r \otimes X^n$.

In particular, $k[X] \otimes_k k[Y] \cong k[X, Y]$. Common uses involve studying integer polynomials by looking at their reductions mod p, so $\mathbb{F}_p \otimes_{\mathbb{Z}} \mathbb{Z}[X] \cong \mathbb{F}_p[X]$, as well as what happens 'generically', so $\mathbb{Q} \otimes_{\mathbb{Z}} \mathbb{Z}[X] \cong \mathbb{Q}[X]$.

Now, if S is a finitely generated k-algebra, so there exists a surjective ring homomorphism $k[X_1, \ldots, X_n] \twoheadrightarrow S$, then we have a surjective ring homomorphism

$$R[X_1, \ldots, X_n] \cong R \otimes_k k[X_1, \ldots, X_n] \twoheadrightarrow R \otimes_k S$$

and so $R \otimes_k S$ is a finitely generated R-algebra. If R is a finitely generated k-algebra as well, then $R \otimes_k S$ is a finitely generated k-algebra.

9.1 Ideals in tensor algebras

Let R and S be k-algebras, and let $I \lhd R$ and $J \lhd S$ be ideals. By abuse of notation we write $I \otimes J$ for the ideal of $R \otimes_k S$ generated by all elements $x \otimes y$ with $x \in I$ and $y \in J$. In particular we have the ideals $I \otimes S$ and $R \otimes J$.

Proposition 9.2.

$$(R \otimes_k S)/(I \otimes S + R \otimes J) \cong (R/I) \otimes_k (S/J).$$

Proof. Set $L := I \otimes S + R \otimes J$ and $T := (R \otimes_k S)/L$. The algebra homomorphism $R \to R \otimes_k S \to T$ factors through R/I, and similarly for S and J. By Proposition 9.1 we have an induced algebra homomorphism $(R/I) \otimes_k (S/J) \to T$, $\bar{r} \otimes \bar{s} \mapsto r \otimes s + L$.

Conversely we have an algebra homomorphism $R \to R/I \to (R/I) \otimes_k (S/J)$, and similarly for S and J. Hence by Proposition 9.1 we have an induced algebra homomorphism $R \otimes_k S \to (R/I) \otimes_k (S/J)$, $r \otimes s \mapsto \bar{r} \otimes \bar{s}$. This factors through $T = (R \otimes_k S)/L$, giving an algebra homomorphism $T \to (A/I) \otimes_R (B/J)$, $r \otimes s + L \mapsto \bar{r} \otimes \bar{s}$.

These maps are mutually inverse, finishing the proof. □

9.2 Examples

Let L/K be a field extension and $f \in K[X]$ a (monic) polynomial. Then

$$L \otimes_K \big(K[X]/(f)\big) \cong (L \otimes_K K[X])/(1 \otimes f) \cong L[X]/(f).$$

Factorising $f = \prod_i f_i^{r_i}$ into a product of irreducible polynomials over L, we can apply the Chinese Remainder Theorem to get

$$L \otimes_K \big(K[X]/(f(X))\big) \cong L[X]/(f(X)) \cong \prod_i (L[X]/(f_i(X)^{r_i}).$$

As an explicit example we have $\mathbb{C} \cong \mathbb{R}[X]/(X^2+1)$, so

$$\mathbb{C} \otimes_{\mathbb{R}} \mathbb{C} \cong \mathbb{C}[X]/(X^2+1) \cong \mathbb{C}[X]/(X+i)(X-i) \cong \mathbb{C} \times \mathbb{C}.$$

This sends $x \otimes y$ to the pair $(xy, x\bar{y})$, where $\bar{y}$ is the complex conjugate of y.

This example generalises to give that if L/K is Galois, then $L \otimes_K L \cong L^{[L:K]}$ via $x \otimes y \mapsto (x\sigma(y))$ where σ varies over the elements of $\mathrm{Gal}(L/K)$.

For, L/K is finite and separable, so by the Primitive Element Theorem we can write $L = K(\alpha) \cong K[X]/(f(X))$, where f is the minimal polynomial of α over K. Also, L/K is normal, so $f(X)$ splits into distinct linear factors over L, with roots corresponding to the conjugates of α over K. This gives the isomorphism $L \otimes_K L \cong L^{[L:K]}$.

In general there are many more prime ideals in $R \otimes_k S$ than we can construct simply from ideals of R and S. For example, $(X-Y)$ is prime in $K[X,Y] \cong K[X] \otimes_K K[Y]$. In general, we cannot even construct all maximal ideals in this way. (Compare with the Weak Nullstellensatz later.)

As an explicit example, consider the algebra homomorphism

$$\mathbb{R}[X,Y] \to \mathbb{C}, \quad X \mapsto i, \quad Y \mapsto i.$$

This is onto, so its kernel $M = (X^2+1, X-Y)$ is a maximal ideal of $\mathbb{R}[X,Y]$.

Note that this ideal is not of the form $I \otimes \mathbb{R}[Y] + \mathbb{R}[X] \otimes J$ for ideals $I \lhd \mathbb{R}[X]$ and $J \lhd \mathbb{R}[Y]$. For, if this were the case, then I would be the kernel of the composition $\mathbb{R}[X] \to \mathbb{R}[X,Y] \to \mathbb{C}$, $X \mapsto i$, so $I = (X^2+1)$, and similarly $J = (Y^2+1)$, but then

$$\mathbb{R}[X,Y]/(I \otimes \mathbb{R}[Y] + \mathbb{R}[X] \otimes J) \cong (\mathbb{R}[X]/I) \otimes_{\mathbb{R}} (\mathbb{R}[Y]/J) \cong \mathbb{C} \otimes_{\mathbb{R}} \mathbb{C} \cong \mathbb{C}^2,$$

which is not a field.

Part III

Noetherian Rings

10 Noetherian modules

An R-module M is called **Noetherian** provided that every submodule of M is finitely generated. In particular, M will itself be a finitely generated module.

Lemma 10.1. *The following are equivalent for an R-module M.*

(1) M is Noetherian.

*(2) M satisfies the **ascending chain condition**: every chain of submodules*

$$U_1 \subset U_2 \subset U_3 \subset \cdots$$

*is **stationary**, so there exists N with $U_n = U_N$ for all $n \geq N$.*

(3) Every non-empty set of submodules of M has a maximal element.

Proof. (1) $\Rightarrow$ (2). The union $U := \bigcup_i U_i$ is again a submodule of M, so is finitely generated by (1). Each of these generators must lie in some U_i, and taking N to be the maximum of the i which occur, we have $U = U_N$. Hence $U_n = U_N$ for all $n \geq N$.

(2) $\Rightarrow$ (3). Let $\mathcal{S}$ be a non-empty set of submodules of M, and suppose that $\mathcal{S}$ has no maximal element. Since $\mathcal{S}$ is non-empty, we can take some $U_1 \in \mathcal{S}$. As U_1 is not maximal we can find U_2 in $\mathcal{S}$ strictly larger than U_1. Repeating this argument we can construct inductively a non-stationary ascending chain $U_1 \subset U_2 \subset U_3 \subset \cdots$, contradicting (2).

(3) $\Rightarrow$ (1). Let U be a submodule of M, and let $\mathcal{S}$ be the set of finitely-generated submodules of U. By (3), $\mathcal{S}$ has a maximal element $U' = (x_1, \ldots, x_r)$. If $u \in U$, then $U' + (u) = (x_1, \ldots, x_r, u)$ is a finitely-generated submodule of U containing U'. By maximality this equals U' and so $u \in U'$. Hence $U = U'$ and U is finitely generated. $\square$

Proposition 10.2. *Let $0 \to L \to M \to N \to 0$ be a short exact sequence of R-modules. Then M is Noetherian if and only if both L and N are Noetherian.*[10]

Proof. We write $\pi\colon M \to N$ for the epimorphism from the short exact sequence, and will identify L with the submodule $\mathrm{Ker}(\pi)$ of M.

Suppose first that M is Noetherian. Clearly every submodule of L is *a fortiori* a submodule of M, and hence finitely generated. Thus L is Noetherian. Similarly, if $V \leq N$ is a submodule, then its preimage $\pi^{-1}(V)$ is a submodule of M, so is finitely generated, and if $m_1, \ldots, m_n$ generate $\pi^{-1}(V)$, then $\pi(m_1), \ldots, \pi(m_n)$ generate V. Thus N is Noetherian.

Conversely, suppose that both L and N are Noetherian. Let $U \leq M$ be a submodule. Then $U \cap L$ is a submodule of L, so finitely generated, say by $m_1, \ldots, m_d$. Similarly,

[10] In other words, the Noetherian modules form a **Serre subcategory**.

$\pi(U) \leq N$ is a submodule, so finitely generated, say by $\pi(m_{d+1}), \ldots, \pi(m_n)$ for some $m_i \in U$. Now, if $u \in U$, then $\pi(u) = r_{d+1}\pi(m_{d+1}) + \cdots + r_n\pi(m_n)$ for some $r_i \in R$. Next, $v := u - (r_{d+1}m_{d+1} + \cdots + r_n m_n) \in \mathrm{Ker}(\pi)$, so $v \in U \cap L$ and $v = r_1 m_1 + \cdots + r_d m_d$ for some $r_i \in R$. Hence $u = r_1 m_1 + \cdots + r_n m_n$, so $U = (m_1, \ldots, m_n)$ is finitely-generated. Thus M is Noetherian. □

In particular, we see that if $M_1, \ldots, M_n$ are Noetherian R-modules, then the direct product $M_1 \times \cdots \times M_n$ is also Noetherian.

Proposition 10.3. *Let M be a Noetherian R-module, and $\Sigma \subset R$ a multiplicatively-closed subset. Then M_Σ is a Noetherian R_Σ-module.*

Proof. Let $V \leq M_\Sigma$ be an R_Σ-submodule, and $U := \{x \in M : x/1 \in V\}$ its preimage, which is an R-submodule of M. If $\frac{m}{a} \in V$, then $\frac{m}{1} = \frac{a}{1} \cdot \frac{m}{a} \in V$, so $m \in U$ and $\frac{m}{a} \in U_\Sigma$. Thus $U_\Sigma = V$, and so U finitely generated as an R-module implies V finitely generated as an R_Σ-module. □

As examples of non-Noetherian modules, we observe that $\mathbb{Q}$ is not finitely generated as a $\mathbb{Z}$-module. In fact, if R is a domain, then $\mathrm{Quot}(R)$ is finitely generated as an R-module if and only if R is a field.

It follows that $\mathbb{Q}/\mathbb{Z}$ is also non-Noetherian. For a prime p we have the $\mathbb{Z}$-submodule $T(p)_n := \{x \in \mathbb{Q}/\mathbb{Z} : p^n x = 0\}$. If $m < n$, then $T(p)_m < T(p)_n$ is a strict inclusion, so that $\mathbb{Q}/\mathbb{Z}$ has an infinite ascending chain $T(p)_1 < T(p)_2 < \cdots$. The union $T(p) := \bigcup_n T(p)_n$ is the subgroup of all **p-torsion** elements of $\mathbb{Q}/\mathbb{Z}$.

11 Noetherian rings

We call R a **Noetherian ring** if the regular module is Noetherian, which is if and only if every ideal of R is finitely generated.

Clearly all fields are Noetherian, as is every principal ideal domain. Hilbert's Basis Theorem below shows that all finitely generated algebras over Noetherian rings are again Noetherian; this includes every ring of the form $K[X_1, \ldots, X_n]/I$ where K is a field. Finally, all localisations of Noetherian rings are Noetherian, so most rings occurring 'in nature' are Noetherian.

We begin with a result concerning Noetherian modules over Noetherian rings.

Lemma 11.1. *Let R be a Noetherian ring and M an R-module. Then M is Noetherian if and only if M is finitely generated.*

Proof. If M is Noetherian, then it is necessarily finitely generated. Conversely, if M is finitely generated, then we have an epimorphism $R^n \twoheadrightarrow M$ for some n (Lemma 4.4). Now R^n, and hence M, is Noetherian by Proposition 10.2. □

Proposition 11.2. *Let R be a Noetherian ring.*

(1) Let $I \lhd R$ an ideal. Then R/I is a Noetherian ring.

(2) Let $\Sigma \subset R$ be a multiplicatively closed subset. Then R_Σ is a Noetherian ring.

Proof. (1) Given an ideal $\bar{J} \lhd R/I$, its preimage $J \lhd R$ is an ideal containing I. Since J is finitely generated, so too is $\bar{J} = J/I$.

(2) Given an ideal $\bar{J} \lhd R_\Sigma$, its preimage $J \lhd R$ is an ideal. Since J is finitely generated, so too is $\bar{J} = J_\Sigma$ by Lemma 3.5. □

Theorem 11.3 (Hilbert's Basis Theorem)**.** *If R is Noetherian, then so too is the polynomial ring $R[X]$.*

Proof. Let $I \lhd R[X]$ be an ideal. Set

$$J := \{a \in R : \exists\, d, r_i \text{ with } aX^d + r_1X^{d-1} + \cdots + r_d \in I\}.$$

Thus J is the set of leading coefficients of polynomials in I.

We observe that J is an ideal of R. For, given $a, b \in J$, we can find polynomials $f = aX^d + r_1X^{d-1} + \cdots + r_d$ and $g = bX^e + s_1X^{e-1} + \cdots + s_e$ in I. Then for $r \in R$ the polynomial $X^e f + rX^d g \in I$ has leading term $(a + rb)X^{d+e}$, whence $a + rb \in J$.

Since R is Noetherian, J is finitely generated, say by $a_1, \ldots, a_m$. Let $f_1, \ldots, f_m$ be polynomials in I having leading coefficients $a_1, \ldots, a_r$ respectively. Multiplying each f_i by an appropriate power of X, we may assume that $\deg(f_i) = N$ for some fixed N.

Finally, let $R[X]_N \leq R[X]$ be the (free) R-submodule generated by $1, X, \ldots, X^{N-1}$. Then $R[X]_N$ is a Noetherian R-module by Lemma 11.1, and $I_N := I \cap R[X]_N$ is an R-submodule of $R[X]_N$. Thus I_N is finitely generated, say by $f_{m+1}, \ldots, f_n$.

We prove by induction on degree that $f_1, \ldots, f_n$ generate I.

Let $f \in I$ have degree d. If $d < N$, then $f \in I_N = (f_{m+1}, \ldots, f_n)$. Assume therefore that $d \geq N$. The leading coefficient a of f lies in J, so $a = r_1a_1 + \cdots + r_ma_m$ for some $r_i \in R$. Then $g := f - (r_1f_1 + \cdots + r_mf_m)X^{d-N-1}$ also lies in I and has smaller degree than f. By induction we have $g \in (f_1, \ldots, f_n)$, whence $f \in (f_1, \ldots, f_n)$ as claimed. Thus $R[X]$ is Noetherian. □

Corollary 11.4. *Let R be Noetherian. Then every localisation of every finitely generated R-algebra is Noetherian.*

Proof. Every finitely generated R-algebra can be written as $S \cong R[X_1, \ldots, X_n]/I$. By Hilbert's Basis Theorem and induction, $R[X_1, \ldots, X_n]$ is Noetherian. Thus S and all its localisations are Noetherian by Proposition 11.2. □

Proposition 11.5. *Let R be Noetherian. Then every ideal of R contains a power of its radical. In particular,* $\mathrm{nil}(R)$ *is a nilpotent ideal.*

Proof. Let I be an ideal and $J = \mathrm{rad}(I)$ its radical. Then J is finitely generated, say by $x_1, \ldots, x_d$, and for n sufficiently large we have $x_i^n \in I$ for all i. Now, J^{nd} is generated by all monomials $x_1^{a_1} \cdots x_d^{a_d}$ with $\sum_i a_i = nd$. Since $a_i \geq n$ for some i we conclude that $J^{nd} \subset I$. □

Theorem 11.6 (Krull's Intersection Theorem)**.** *Let R be Noetherian and I a proper ideal of R. If $y \in \bigcap_n I^n$, then $y \in Iy$. In particular, if R is a domain, or if $I \subset \mathrm{Jac}(R)$, then $\bigcap_n I^n = 0$.*

As motivation for this theorem, think of the commutative ring R of differentiable functions from $\mathbb{R}$ to itself, and the sequence of ideals I_n consisting of those functions with $f^{(r)}(0) = 0$ for all $0 \leq r \leq n$. Now the Taylor series approximation of f at 0 is giving by the power series $\bar{f}(x) = \sum_{n\geq 0} \frac{1}{n!} f^{(n)}(0)x^n$, and one wants to how good this approximation is. The question thus arises as to which functions lie in the intersection $\bigcap_n I_n$; these are the so-called **flat functions**, for example $f = e^{-1/x^2}$. This then leads to the definition of **real analytic functions**, which are precisely those functions which do have such an approximation.

Algebraically, we consider the approximations of $f \in R$ by its images in R/I^n. Analogously to writing a function as a power series, we can put these approximations together and consider the **inverse limit** $\varprojlim R/I^n$, also called the **I-adic completion of** R. The kernel of the map $R \to \varprojlim R/I^n$ is the intersection $\bigcap_n I^n$, and so if this intersection is zero, then every $f \in R$ is completely determined by its image in $\varprojlim R/I^n$.

The following beautiful proof of Krull's Intersection Theorem is due to H. Perdry.[11]

Proof. A **monomial** in $R[X_1, \ldots, X_n]$ is an element of the form $X_1^{a_1} \cdots X_n^{a_n}$. A polynomial $p \in R[X_1, \ldots, X_n]$ is **homogeneous** if it is a linear combination of monomials of the same degree.

We know that $I = (x_1, \ldots, x_d)$ is finitely generated. Then I^n is generated by all monomials $x_1^{a_1} \cdots x_d^{a_d}$ with $\sum_i a_i = n$. Thus any element $y \in I^n$ is of the form $p_n(x_1, \ldots, x_n)$ for some homogeneous $p_n \in R[X_1, \ldots, X_n]$ of degree n.

Suppose now that $y \in \bigcap_n I^n$. Then for each n we have $p_n \in R[X_1, \ldots, X_d]$, homogeneous of degree n, such that $y = p_n(\underline{x})$. Let $J := (p_1, p_2, \ldots) \lhd R[X_1, \ldots, X_d]$. By Hilbert's Basis Theorem, J must be finitely generated, say by $p_1, \ldots, p_m$. Write $p_{m+1} = p_1 q_1 + \cdots + p_m q_m$ for some polynomials q_i. Write $q_i = \sum_j q_{i,j}$ where $q_{i,j}$ is homogeneous of degree j. Since p_i is homogeneous of degree i, it follows that $p_{m+1} = \sum_i p_i q_{i,m+1-i}$. We therefore replace q_i by $q_{i,m+1-i}$, so $q_i(\underline{x}) \in I$, and hence

$$y = p_{m+1}(\underline{x}) = \textstyle\sum_i p_i(\underline{x}) q_i(\underline{x}) \in yI.$$

If R is a domain, then $y = yz$ with $z \in I$ implies $y(1 - z) = 0$ and $z \neq 1$. Thus $y = 0$.

If $I \subset \mathrm{Jac}(R)$, then since (y) is a finitely generated R-module and $(y) = I(y)$, we can apply Nakayama's Lemma to deduce that $(y) = 0$. □

An example of a non-Noetherian ring is the polynomial ring in infinitely many variables $k[X_1, X_2, \ldots]$. For, we have the non-stationary ascending chain

$$(X_1) \subset (X_1, X_2) \subset (X_1, X_2, X_3) \subset \cdots$$

Let $R = K[X, T_1, T_2, \ldots]/(\{XT_n - T_{n-1}\})$. Then $R \subset K(X, Y)$ via $T_n \mapsto Y/X^n$, so is a domain. On the other hand, $Y := X^n T_n$ for all n, so $Y \in \bigcap_n (X^n)$.

Recall from Theorem 2.7 that radical ideals are precisely those that can be written as an intersection of (possibly infinitely many) prime ideals. For Noetherian rings we can strengthen this to finite intersections.

[11] Hervé Perdry, 'An elementary proof of Krull's intersection theorem', *Amer. Math. Monthly* 111 (2004), 356–357.

Proposition 11.7. *Every radical ideal in a Noetherian ring is a finite intersection of prime ideals.*

Proof. Let $\mathcal{S}$ be the set of radical ideals which cannot be written as a finite intersection of prime ideals. If $\mathcal{S}$ is non-empty, then we can take a maximal element $I \in \mathcal{S}$. Now I cannot be prime, so there exist $a, b \notin I$ with $ab \in I$. Writing $I = \bigcap_{j \in J} P_j$ as an intersection of prime ideals, we set

$$J_a := \{j \in J : a \in P_j\}, \quad J_b := \{j \in J : b \in P_j\} \quad \text{and} \quad I_a := \bigcap_{j \in J_a} P_j, \quad I_b := \bigcap_{j \in J_b} P_j.$$

Since for each j we have $ab \in I \subset P_j$, we must have either $a \in P_j$ or $b \in P_j$. Thus $J = J_a \cup J_b$ and $I = I_a \cap I_b$. Moreover, both I_a and I_b are intersections of prime ideals, so are radical ideals. Finally, $I \subset I_a$ and $a \in I_a - I$, so by maximality I_a is a finite intersection of prime ideals. Similarly for I_b, and hence also for $I = I_a \cap I_b$, a contradiction. Thus $\mathcal{S}$ is empty. $\square$

Part IV

Integral Extensions

In algebraic number theory, integral extensions are used to define **algebraic integers** inside an **number field**; that is, we take a finite field extension K of $\mathbb{Q}$, and set $\mathcal{O}_K$ to be the set of elements of K which are integral over $\mathbb{Z}$.

In geometry, integral extensions give rise to morphisms of algebraic varieties which are surjective and have **finite fibres**; that is $f\colon V \to W$ such that $f^{-1}(w)$ is finite for all $w \in W$. For example, the projection of the cuspidal cubic $y^2 = x^3$ onto the Y-axis.

12 Integral Extensions

Let $R \leq S$ be rings. An element $x \in S$ is said to be **integral over** R provided there exists a **monic** polynomial $m \in R[X]$ with $m(x) = 0$. We call S integral over R if every element in S is integral over R.

Clearly every element in R is integral over R. Also, if $x \in \mathbb{Q}$ is integral over $\mathbb{Z}$, then $x \in \mathbb{Z}$. For, suppose x satisfies the monic equation

$$x^n + p_{n-1}x^{n-1} + \cdots + p_1 x + p_0 = 0, \quad p_i \in \mathbb{Z}.$$

Write $x = r/a$ with $r, a \in \mathbb{Z}$ coprime. Clearing denominators we get

$$r^n + p_{n-1}ar^{n-1} + \cdots + p_1 a^{n-1} r + p_0 a^n = 0.$$

Hence a divides r^n, so $a = \pm 1$ is a unit and $x = \pm r \in \mathbb{Z}$.

Let $R \to S$ be a ring homomorphism. Given elements $x_i \in S$, we write $R[x_1, \ldots, x_n]$ for the image of $R[X_1, \ldots, X_n] \to S$, $X_i \mapsto x_i$. We say that S is of **finite type over** R if $S = R[x_1, \ldots, x_n]$ for some x_i. In other words, S is finitely generated as an R-algebra. We say that S is **finite over** R if S is finitely generated as an R-module. The next proposition is summarised by the slogan

$$\text{finite type} + \text{integral} = \text{finite}.$$

Proposition 12.1. *Let $R \leq S$ be rings. If $S = R[x_1, \ldots, x_n]$ with the x_i integral over R, then S is finitely generated as an R-module. Conversely, if S is finitely generated as an R-module, then S is integral over R.*

Proof. Let $m_i \in R[X]$ be a monic polynomial such that $m_i(x_i) = 0$. Let $d_i := \deg(m_i)$. Then the surjective ring homomorphism $R[X_1, \ldots, X_n] \to S$, $X_i \mapsto x_i$, factors through the ideal $J = (m_1(X_1), \ldots, m_n(X_n))$. Now $R[X_i]/J$ is finitely generated as an R-module, by the monomials $X_1^{a_1} \ldots X_n^{a_n}$ for $0 \leq a_i < d_i$ (in fact it is a free R-module with this as a basis), and so S is also finitely generated as an R-module, by the elements $x_1^{a_1} \cdots x_n^{a_n}$ for $0 \leq a_i < d_i$.

Conversely, let S be finitely generated as an R-module and let $x \in S$. The Cayley-Hamilton Theorem, Proposition 4.5, applied to $f\colon s \mapsto xs$ and $I = R$ implies that, for some $r_i \in R$, we have $f^n + r_{n-1}f^{n-1} + \cdots + r_1 f + r_0 = 0$ as an endomorphism of S. Applying this to the identity in S yields $x^n + r_{n-1}x^{n-1} + \cdots + r_x + r_0 = 0$ in S, and hence x is integral over R. □

In fact, the more precise version of the Cayley-Hamilton Theorem proves the following.

Lemma 12.2. *Let $R \leq S$ be integral, and suppose $x \in \mathfrak{a}S$ for some ideal $\mathfrak{a} \lhd R$. Then x satisfies some monic polynomial $m \in R[X]$, all of whose coefficients except the first lie in $\mathfrak{a}$.*

Proof. Write $x = \sum_{i=1}^{m} a_i s_i$ with $a_i \in \mathfrak{a}$ and $s_i \in S$. As above, $S' := R[s_1, \ldots, s_m]$ is finite over R, and $x \in \mathfrak{a}S'$. The Cayley-Hamilton Theorem applied to multiplication by x yields a monic polynomial $m \in R[X]$ with $m(x) = 0$ and with all but the first coefficient lying in $\mathfrak{a}$. □

Lemma 12.3 (Transitivity of Integral Extensions)**.** *Let $R \leq S \leq T$ be rings. Then $R \leq T$ is integral if and only if both $R \leq S$ and $S \leq T$ are integral.*

Proof. It is clear that $R \leq T$ being integral implies that both $R \leq S$ and $S \leq T$ are integral. Suppose conversely that $R \leq S$ and $S \leq T$ are both integral. Given $x \in T$ there exist $y_i \in S$ with $x^n + s_{n-1}x^{n-1} + \cdots + s_1 x + s_0 = 0$. Now consider $S' := R[s_0, \ldots, s_{n-1}]$. This algebra is finitely generated by integral elements, so is finitely generated as an R-module. Also, $S'[x]$ is a finitely-generated S'-module. It follows that $S'[x]$ is finitely generated as an R-module, and hence is integral over R. Thus x is integral over R. □

Proposition 12.4. *Let $R \leq S$ be integral.*

(1) If $J \lhd S$ and $I = J \cap R$, then $R/I \leq S/J$ is integral.

(2) If $\Sigma \subset R$ is multiplicatively closed, then $R_\Sigma \leq S_\Sigma$ is integral.

Proof. (1) Given $\bar{x} \in S/J$, take a representative $x \in S$. Then x is integral over R, say $m(x) = 0$ with $m \in R[X]$ monic. Then $\bar{m} \in (R/I)[X]$ remains monic and $\bar{m}(\bar{x}) = m(x) + J$ is zero. Hence $\bar{x}$ is integral over R/I.

(2) Given $x/a \in S_\Sigma$ we know that x is integral over R, say $m(x) = 0$ with $m \in R[X]$ monic. Write $m = X^n + r_{n-1}X^{n-1} + \cdots + r_1 X + r_0$, and define

$$\widetilde{m} := X^n + \tfrac{r_{n-1}}{a} X^{n-1} + \cdots + \tfrac{r_1}{a^{n-1}} X + \tfrac{r_0}{a^n} \in R_\Sigma[X].$$

Then $\widetilde{m}(x/a) = m(x)/a^n = 0$ and so x/a is integral over R_Σ. □

Lemma 12.5. *Given $R \leq S$, define*

$$S^{\mathrm{int}/R} := \{x \in S : x \text{ is integral over } R\}.$$

Then $S^{\mathrm{int}/R}$ is a subring of S and $R \leq S^{\mathrm{int}/R}$ is integral.

Proof. Let $x, y \in S^{\mathrm{int}/R}$. The algebra $R[x, y]$ is finitely generated by integral elements, so is integral. In particular, $x \pm y, xy$ are integral over R, so contained in $S^{\mathrm{int}/R}$. Thus $S^{\mathrm{int}/R}$ is closed under all ring operations, and obviously contains R. □

We call $S^{\mathrm{int}/R}$ the **integral closure** of R in S. Clearly $S^{\mathrm{int}/R} = S$ if and only if S is integral over R. We say that R is **integrally closed in** S if $S^{\mathrm{int}/R} = R$. Note that by transitivity, Lemma 12.3, the ring $S^{\mathrm{int}/R}$ is integrally closed in S.

As a special case, an **integrally closed domain** is a domain which is integrally closed in its quotient field. We saw earlier that $\mathbb{Z}$ is integrally closed, and in fact every unique factorisation domain is integrally closed.

A Noetherian domain which is integrally closed is often called **normal**. In geometry, this is related to resolution of singularities of irreducible curves; in number theory the algebraic integers in a number field K is defined to be the integral closure of $\mathbb{Z}$ in K.

Let $R \leq S$ be domains, with quotient fields $K \leq L$. If $x \in S$ is integral over R, then it is the root of some monic polynomial $m \in R[X]$. On the other hand, the map $K[X] \to L$, $X \mapsto x$, has kernel a non-zero prime ideal, so is generated by a monic irreducible polynomial $\min_{x/K} \in K[X]$, called the **minimal polynomial** of x over K. One may then ask how these polynomials are related.

Proposition 12.6. *Let $R \leq S$ be domains, with quotient fields $K \leq L$, and assume that R is integrally closed. If $x \in S$ is integral over R, then $\mu := \min_{x/K} \in R[X]$ and generates the kernel of $R[X] \to S$, $X \mapsto x$.*

Moreover, if $R \leq S$ is integral and $x \in \mathfrak{p}S$ for some prime ideal $\mathfrak{p} \lhd R$, then every coefficient of $\min_{x/K}$ but the first lies in $\mathfrak{p}$.

Proof. Let $m \in R[X]$ be a monic polynomial having x as a root. By definition, $\mu := \min_{x/K}$ generates the kernel of the map $K[X] \to L$, $X \mapsto x$. Since $m(x) = 0$ we must have that μ divides m over K, so there exists $g \in K[X]$ with $m = g\mu$.

Now let M be an algebraically-closed field containing L. We know that μ splits over M, say as $\mu = (X - x_1)(X - x_2) \cdots (X - x_n)$ with $x = x_1$. Then $m(x_i) = g(x_i)\mu(x_i) = 0$ for all i, so each x_i is integral over R. Expanding our expression for μ we see that every coefficient is a sum of products of the x_i, and hence every coefficient of μ lies in K and is integral over R. Since R is integrally closed we conclude that $\mu \in R[X]$.

Now let $f \in R[X]$ be any polynomial having x as a root. Since μ is monic we can apply the Division Algorithm to get $f = q\mu + r$ with $q, r \in R[X]$ and $\deg(r) < \deg(\mu)$. On the other hand, μ divides f over K, and hence also r. Thus by comparing degrees we get $r = 0$, and hence μ divides f over R.

Finally, let $R \leq S$ be integral and $x \in \mathfrak{p}S$. As in Lemma 12.2, we obtain a monic polynomial $f \in R[X]$, say of degree d, with all but the first coefficient lying in $\mathfrak{p}$. Now $f = g\mu$ for some $g \in R[X]$, so $\bar{g}\bar{\mu} = \bar{f} = X^d \in (R/\mathfrak{p})[X]$. Thus μ is a monic polynomial such that $\bar{\mu} = X^r$. Hence every coefficient of μ except the first lies in $\mathfrak{p}$. □

12.1 Lying-Over, Going-Up and Going-Down

We now discuss existence of primes lying over a given prime.

Theorem 12.7 (Lying-Over)**.** *Let $R \leq S$ be rings and $\mathfrak{p} \lhd R$ a prime ideal.*

(1) There exists a prime ideal $P \lhd S$ with $P \cap R = \mathfrak{p}$ if and only if $\mathfrak{p}S \cap R \subset \mathfrak{p}$.

(2) If $R \leq S$ is integral, then there exists such a prime P.

Proof. (1) If we have such a P, then $\mathfrak{p}S \subset P$, and hence $\mathfrak{p}S \cap R \subset P \cap R = \mathfrak{p}$. Conversely, suppose $\mathfrak{p}S \cap R \subset \mathfrak{p}$. Localise with respect to $\Sigma = R - \mathfrak{p}$ to get $R_\mathfrak{p} \leq S_\mathfrak{p}$

(necessarily still a subring). Now, by Lemma 3.5, we have $\mathfrak{p}S_{\mathfrak{p}} = (\mathfrak{p}S)_{\mathfrak{p}}$, and is a proper ideal of $S_{\mathfrak{p}}$. It is thus contained in a maximal ideal, which by Theorem 3.7 we can write as $P_{\mathfrak{p}}$ for some prime $P \lhd S$ with $P \cap \Sigma = \emptyset$, so $P \cap R \subset \mathfrak{p}$. On the other hand, $\mathfrak{p}_{\mathfrak{p}} \subset P_{\mathfrak{p}}$, so $\mathfrak{p} \subset P$, and hence $P \cap R = \mathfrak{p}$.

(2) Let $x \in \mathfrak{p}S$. By Lemma 12.2 we have $m(x) = 0$ for some polynomial $m = X^n + p_{n-1}X^{n-1} + \cdots + p_0$ with $p_i \in R$. If moreover $x \in R$, then $x^n = -(p_{n-1}x^{n-1} + \cdots + p_0) \in \mathfrak{p}$, and so $x \in \mathfrak{p}$. This shows that $\mathfrak{p}S \cap R \subset \mathfrak{p}$ as required. □

More generally, we are interested in lifting chains of prime ideals, so given rings $R \leq S$ and a chain of primes in R, can we lift this to a chain of primes in S. This separates into two problems: lifting ascending chains and lifting descending chains, and covered by the Going-Up and Going-Down Theorems.

Pictorially we can think of these properties as

$$\begin{array}{ccccc} S\colon & P & \subset & \exists Q \\ | & | & & | \\ R\colon & \mathfrak{p} & \subset & \mathfrak{q} \end{array} \qquad\qquad \begin{array}{ccccc} S\colon & \exists P & \subset & Q \\ | & | & & | \\ R\colon & \mathfrak{p} & \subset & \mathfrak{q} \end{array}$$

Going Up Going Down

Explicitly, we say that a ring extension $R \leq S$ has the **Going-Up Property** if, given a prime $P \lhd S$ and a prime $\mathfrak{q} \lhd R$ containing $\mathfrak{p} := R \cap P$, then there is a prime $Q \lhd S$ lying over $\mathfrak{q}$ and containing P.

Similarly, we say that $R \leq S$ has the **Going-Down Property** if, given a prime $Q \lhd S$ and a prime $\mathfrak{p} \lhd R$ contained in $\mathfrak{q} := R \cap Q$, then there exists a prime $P \lhd S$ lying over $\mathfrak{p}$ and contained in Q.

Theorem 12.8 (Going-Up)**.** *Every integral extension has the Going-Up Property.*

Proof. Let $R \leq S$ be integral. By Proposition 12.4 we know that $R/\mathfrak{p} \leq S/P$ is integral and $\mathfrak{q}/\mathfrak{p} \lhd R/\mathfrak{p}$ is prime, so there exists a prime in S/P lying over $\mathfrak{q}/\mathfrak{p}$. Take Q to be its preimage in S. □

Theorem 12.9 (Going-Down)**.** *Let $R \leq S$ be an integral extension of domains, and assume R is integrally closed. Then $R \leq S$ has the Going-Down Property.*

Proof. Consider the localisation of S at Q, giving $R \leq S_Q$. We claim that $\mathfrak{p}S_Q \cap R \subset \mathfrak{p}$, so by the Lying-Over Theorem there exists a prime ideal in S_Q lying over $\mathfrak{p}$, and we can take P to be its preimage in S.

Take $0 \neq x \in \mathfrak{p}S_Q \cap R$. Since $\mathfrak{p}S_Q = (\mathfrak{p}S)_Q$ we can write $x = s/a$ with $s \in \mathfrak{p}S$ and $a \in S - Q$. Since $s \in \mathfrak{p}S$ is integral over R and R is integrally closed we know from Proposition 12.6 that the minimal polynomial m of s over $K := \mathrm{Quot}(R)$ is of the form $m = X^n + r_{n-1}X^{n-1} + \cdots + r_0$ with $r_i \in \mathfrak{p}$.

On the other hand, $x^{-1} \in K$, so the minimal polynomial of $a = s/x$ over K is $\widetilde{m} = X^n + \tilde{r}_{n-1}X^{n-1} + \cdots + \tilde{r}_0$ where $\tilde{r}_i = r_i/x^{n-i}$. Since a is also integral over R, we similarly have $\tilde{r}_i \in R$.

If $x \notin \mathfrak{p}$, then from $x^{n-i}\tilde{r}_i = r_i \in \mathfrak{p}$ we deduce that $\tilde{r}_i \in \mathfrak{p}$ for all i. Thus $a^n = -(\tilde{r}_0 + \tilde{r}_1 a + \cdots + \tilde{r}_{n-1}a^{n-1})$ lies in $\mathfrak{p}S \subset Q$, whence $a \in Q$, a contradiction. □

We now show some further properties about prime ideals in integral extensions.

Theorem 12.10. *Let $R \leq S$ be an integral extension.*

(1) If we have primes $P \subset Q$ in S with $P \cap R = Q \cap R$, then $P = Q$.

(2) A prime $P \lhd S$ is maximal if and only if $P \cap R$ is a maximal ideal of R.

Proof. Using Proposition 12.4 we pass to the factor rings $R/(R \cap P) \leq S/P$. We may therefore assume that R and S are domains, and that $P = 0$.

(1) Given $x \in Q$ we can write $x^n + \cdots + r_1 x + r_0 = 0$ for some $r_i \in R$. Then $r_0 \in R \cap Q = 0$ so either $x = 0$, or else $x^{n-1} + \cdots + r_2 x + r_1 = 0$ and $x = 0$ by induction. Thus $Q = 0$.

(2) Let $\mathfrak{q} \lhd R$ be a prime, and $Q \lhd S$ a prime lying over $\mathfrak{q}$, which exists by the Lying-Over Theorem. Now $\mathfrak{q} \neq 0$ implies $Q \neq 0$, whereas $Q \neq 0$ implies $\mathfrak{q} \neq 0$ by (1). Thus R is a field if and only if S is a field. □

We also have the following useful corollary of the Lying Over Theorem.

Corollary 12.11. *Let $R \leq S$ be a ring extension, and $\mathfrak{p} \lhd R$ a prime ideal. Then the set of primes of S lying over $\mathfrak{p}$ is in bijection with the set of primes of $S \otimes_R \kappa(\mathfrak{p})$.*

Proof. Exercise. □

12.2 Noether Normalisation

The Noether Normalisation Lemma provides a basic structure theorem for finitely-generated algebras over a field, and is the basis of many proofs in algebraic geometry.

We shall need the following lemma.

Lemma 12.12. *Let K be an infinite field and $f \in K[X_1, \ldots, X_n]$ a non-zero polynomial. Then there exist $\lambda_i \in K$ such that $f(\lambda_1, \ldots, \lambda_n) \neq 0$.*

Proof. The proof is by induction on n. The case $n = 0$ is trivial, and the case $n = 1$ follows from the fact that $K[X_1]$ is a principal ideal domain.

Let $f \in K[X_1, \ldots, X_n]$ be non-zero, and write $f = g_r X_n^r + \cdots + g_1 X_n + g_0$ with $g_i \in K[X_1, \ldots, X_{n-1}]$ and $g_r \neq 0$. By induction $g_r(\lambda_1, \ldots, \lambda_{n-1}) \neq 0$ for some $\lambda_i \in K$. Then $f(\lambda_1, \ldots, \lambda_{n-1}, X_n)$ is a non-zero polynomial in X_n, and hence there exists $\lambda_n \in K$ with $f(\lambda_1, \ldots, \lambda_n) \neq 0$. □

We can now prove the Normalisation Lemma. For a geometric interpretation see Theorem 14.13.

Theorem 12.13 (Noether Normalisation Lemma)**.** *Let $R = K[x_1, \ldots, x_n]$ with K a field. Then there exists a polynomial subalgebra $S \leq R$ over which R is finite; that is, $S \cong K[X_1, \ldots, X_r]$ for some $r \leq n$, and R is a finite S-algebra.*

In fact, if K is infinite, then after renumbering we may take $X_i = x_i + \sum_{j>r} \lambda_{ij} x_j$ to be homogeneous linear combinations of the x_i.

Proof. We will prove this when K is an infinite field; for finite fields see the exercises.

The proof goes by induction on n, the case $n = 0$ being trivial.

Consider $R = K[x_1, \ldots, x_n]$ and consider the surjective algebra homomorphism

$$K[X_1, \ldots, X_n] \to R = K[x_1, \ldots, x_n], \quad X_i \mapsto x_i.$$

We say that the x_i are **algebraically independent** if this is an isomorphism, in which case we take $S = R$. Otherwise there exists some non-zero f in the kernel, say of degree d. We claim that, after a homogeneous linear transformation, we may assume that f is a monic polynomial of degree d in one of the variables.

Let $F \in K[X_1, \ldots, X_n]$ be the homogeneous part of f of degree d. After renumbering we may assume that F involves $T := X_n$. Since F is homogeneous, the inhomogeneous polynomial $T - 1$ does not divide F, and so $F(X_1, \ldots, X_{n-1}, 1)$ is a non-zero polynomial. Since K is infinite we can apply the lemma to obtain $\lambda_i \in K$ with $\alpha := F(\lambda_1, \ldots, \lambda_{n-1}, 1) \neq 0$. Note that $f(\lambda_1, \ldots, \lambda_{n-1}, T)$ has degree d, and T^d occurs with coefficient α. Thus, setting $Z_i := X_i - \lambda_i T$ and

$$g(Z_1, \ldots, Z_{n-1}, T) := \tfrac{1}{\alpha} f(Z_1 + \lambda_1 T, \ldots, Z_{n-1} + \lambda_{n-1} T, T),$$

we see that g has degree d, and is monic of degree d in T.

Set $t := x_n$ and $z_i := x_i - \lambda_i t$. Then $R = K[z_1, \ldots, z_{n-1}, t]$, and t is a root of the monic polynomial $g(z_1, \ldots, z_{n-1}, T)$. Thus t is integral over $\bar{R} := K[z_1, \ldots, z_{n-1}]$, so R is finite over $\bar{R}$ by Proposition 12.1.

By induction we can find $S \leq \bar{R}$ such that $S \cong K[X_1, \ldots, X_r]$ is a polynomial algebra, $r \leq n$, and $\bar{R}$ is finite over S. It follows from Lemma 12.3 that $S \leq R$ is also finite.

Moreover, after renumbering, we can write $X_i = z_i + \sum_{r<j<n} \mu_{ij} z_j$ for some scalars $\mu_{ij} \in K$. Setting $\mu_{in} := \lambda_i + \sum_{r<j<n} \mu_{ij} \lambda_j$, it follows that $X_i = x_i + \sum_{j>r} \mu_{ij} x_j$ as required. □

Example. It is instructive to see an explicit example, so take $R = K[X, Y]/(XY - 1)$. We take $f = XY - 1$, so $d = 2$ and $F = XY$. We work with $T = Y$, though we could equally work with $T = X$. Now, $F(1, 1) = 1$, so we set $Z = X - Y$. Then $g(Z, Y) = f(Z + Y, Y) = Y^2 + ZY - 1$. Hence $R = K[Y, Z]/(Y^2 + ZY - 1)$, which is integral over $S = K[Z]$.

Theorem 12.14 (Weak Nullstellensatz)**.** *Let K be a field and R a finitely generated K-algebra. If R is a field, then $\dim_K R$ is finite.*

In particular, if K is algebraically closed, then every maximal ideal in $K[X_1, \ldots, X_n]$ is of the form $(X_1 - a_1, \ldots, X_n - a_n)$ for some $a_i \in K$.

Proof. Using the Noether Normalisation Lemma we have a polynomial subalgebra $S = K[X_1, \ldots, X_r]$ of R over which R is finite. If R is a field, then so too is S, by Theorem 12.10 (2). If $r \geq 1$, then (X_1) is a non-zero ideal in S, a contradiction. Hence $r = 0$ and $S = K$. Thus R is finitely generated as a K-module; i.e. $\dim_K R < \infty$.

Now take $R = K[X_1, \ldots, X_n]/\mathfrak{m}$ for some maximal ideal $\mathfrak{m}$. Then R is a field, and finitely generated as a K-algebra, so finite dimensional over K. Thus if K is algebraically closed, then $R = K$. Let a_i be the image of X_i in K. Then $\mathfrak{n} := (X_1 - a_i, \ldots, X_n - a_n) \subset \mathfrak{m}$, and since $\mathfrak{n}$ is also maximal, we must have $\mathfrak{n} = \mathfrak{m}$. □

To relate this to our earlier work on tensor products of algebras, we recall that $K[X,Y] \cong K[X] \otimes_K K[Y]$, in which case $(X-a, Y-b) = (X-a) \otimes K[Y] + K[X] \otimes (Y-b)$. Such an ideal is always maximal, but as we saw earlier, there will in general be other maximal ideals.

12.3 Krull dimension

Let R be a ring. The **Krull dimension** of R, denoted simply $\dim R$, is the supremum of the lengths of chains of prime ideals

$$\dim R := \sup\{n : \text{there exists a chain } \mathfrak{p}_0 \subset \mathfrak{p}_1 \subset \cdots \subset \mathfrak{p}_n \text{ of prime ideals of } R\}.$$

Note that $\dim K = 0$ for all fields K, and $\dim R = 1$ for all principal ideal domains R.

We say that a chain of primes $\mathfrak{p}_0 \subset \mathfrak{p}_1 \subset \cdots \subset \mathfrak{p}_m$ is **saturated** if $\mathfrak{p}_0$ is minimal, $\mathfrak{p}_m$ is maximal, and we cannot insert any prime $\mathfrak{p}_i \subset \mathfrak{q} \subset \mathfrak{p}_{i+1}$. We say that a ring R is **catenary** provided every saturated chain of primes has length $\dim R$.

More generally, if $I \lhd R$ is an ideal, then we define the **height** of I to be

$$\mathrm{ht}(I) := \sup\{n : \text{there exists a chain of primes } \mathfrak{p}_0 \subset \mathfrak{p}_1 \subset \cdots \subset \mathfrak{p}_n \text{ with } \mathfrak{p}_n \subset I\}.$$

Proposition 12.15. *Let $R \leq S$ be an integral extension. Then $\dim R = \dim S$.*

Proof. Let $\mathfrak{p}_0 \subset \mathfrak{p}_1 \subset \cdots \subset \mathfrak{p}_n$ be a chain of primes in R. By the Lying-Over Theorem there is a prime $P_0 \lhd S$ lying over $\mathfrak{p}_0$. Then, by the Going-Up Theorem, we can lift this to a chain of primes $P_0 \subset P_1 \subset \cdots \subset P_n$ in S such that $P_i \cap R = \mathfrak{p}_i$. Thus $\dim S \geq \dim R$.

On the other hand, let $P_0 \subset P_1 \subset \cdots \subset P_n$ be a chain of primes in S. Set $\mathfrak{p}_i := P_i \cap R$. Then each inclusion $\mathfrak{p}_i \subset \mathfrak{p}_{i+1}$ is strict, by Theorem 12.10 (1), so we have a chain of primes $\mathfrak{p}_0 \subset \mathfrak{p}_1 \subset \cdots \subset fp_n$ in R. Thus $\dim S \leq \dim R$. □

Proposition 12.16. *The polynomial ring $K[X_1, \ldots, X_n]$ is catenary of dimension n.*

Proof. Take a height one prime $\mathfrak{p} \lhd R_n := K[X_1, \ldots, X_n]$. Since R_n is a unique factorisation domain, we know that $\mathfrak{p} = (f)$ is principal, generated by an irreducible polynomial. Following the proof of the Noether Normalisation Lemma we may assume that f is monic in X_n, yielding a finite algebra homomorphism $R_{n-1} \to R_n/(f)$. This is moreover injective, since if h lies in the kernel, then $h = fg$ for some $g \in R$. Comparing coefficients of X_n we deduce that $g = h = 0$. By induction R_{n-1} is catenary of dimension $n-1$, so every staurated chain of primes in R_n containing $\mathfrak{p}$ has length n. Thus R_n is catenary of dimension n. □

Lemma 12.17. *Let S be a domain, and $a \in S$ non-zero. Then $\dim S_a = \dim S$.*[12]

Proof. Apply the Noether Normalisation Lemma to obtain a finite extension $R := K[X_1, \ldots, X_n] \rightarrowtail S$, so $\dim S = n$. Since a is integral, it satisfies some monic polynomial μ, and since S is a domain we may assume that the constant term $f \in R$ is

[12] Alternatively, we can interpret the Noether Normalisation Lemma as saying that for a domain S, finitely generated over a field K, the Krull dimension equals the **transcendence degree** over K, so $\dim S = \text{tr. deg}_K \operatorname{Quot}(S)$. Since $\operatorname{Quot}(S_a) = \operatorname{Quot}(S)$, it follows immediately that $\dim S_a = \dim S$.

non-zero. Then S_{af} is integral over $R_f \cong R[T]/(fT-1)$. The minimal primes in this ring correspond to the irreducible factors of $fT-1$, and hence $\dim S_{af} = \dim R_f = n$ by the theorem. On the other hand, using Theorem 3.7 describing primes in a localisation, we must have $\dim S_{af} \leq \dim S_a \leq \dim S$. □

Theorem 12.18. *Let S be a domain, finitely generated over a field K, and $a \in S$ non-zero and not a unit. If $\mathfrak{q}$ is a minimal prime over a, then* $\operatorname{ht}(\mathfrak{q}) = 1$ *and* $\dim S/\mathfrak{q} = \dim S - 1$. *In particular, S is catenary.*

Proof. Set $n := \dim S$. We first show that $\dim S/\mathfrak{q} = n-1$.

Consider the minimal decomposition into primes $\operatorname{rad}(a) = \mathfrak{q} \cap \mathfrak{q}_2 \cdots \cap \mathfrak{q}_r$. Taking $b \in \mathfrak{q}_2 \cap \cdots \cap \mathfrak{q}_r$, $b \notin \mathfrak{q}$, we have $\operatorname{rad}(a) = \mathfrak{q}_b$ in S_b by Lemma 3.5. Also, by the lemma $\dim S_b = n$ and $\dim S_b/\mathfrak{q}_b = \dim S/\mathfrak{q}$. Thus we may replace S by S_b, and hence assume that $\operatorname{rad}(a) = \mathfrak{q}$ is prime.

Applying the Noether Normalisation Lemma we obtain a finite extension $R \leq S$, where $R = K[X_1, \ldots, X_n]$. By Proposition 12.6 the minimal polynomial μ of a has coefficients in R, say with constant term f. We claim that $\operatorname{rad}(a) \cap R = \operatorname{rad}(f)$, in which case $S/\operatorname{rad}(a)$ is finite over $R/\operatorname{rad}(f)$. Since $\operatorname{rad}(f)$ is prime, f must be a power of an irreducible polynomial, so $\operatorname{ht}\operatorname{rad}(f) = 1$ and $\dim R/\operatorname{rad}(f) = n-1$ by the proposition.

Since $\mu(a) = 0$ we must have $f \in (a) \cap R$, so $\operatorname{rad}(f) \subset \operatorname{rad}(a) \cap R$. Conversely, given $x \in \operatorname{rad}(a) \cap R$, we have $x^n = fg$ for some g, say with minimal polynomial ν of degree s. Using the substitution $fg = x^n$, we see that $0 = f^s\nu(g)$ is a polynomial in f, with coefficients in R and constant term x^{sn}. This is therefore divisible by μ, and hence $x^{sn} \in (a)$. This proves the claim.

By induction on dimension, $S/\mathfrak{q}$ is catenary, so every saturated chain of primes in S containing $\mathfrak{q}$ has length n. Now every height one prime of S arises in this way. For, if $\mathfrak{q} \lhd S$ is any height one prime, then it is minimal over every non-zero $a \in \mathfrak{q}$. We conclude that S is catenary. □

Part of this theorem generalises to give Krull's Hauptidealsatz: if S is Noetherian and $a \in S$ not a unit, then every minimal prime over a has height at most one.

Part V
Affine Geometry

We fix once and for all an algebraically-closed field K. For convenience we set

$$R_n := K[X_1, \ldots, X_n]$$

to be the polynomial algebra in n variables over K.

13 Affine varieties

Consider the set $\mathbb{A}^n := K^n$ of n-tuples $\underline{a} := (a_1, \ldots, a_n)$ of elements of K, and called **affine n-space**. It is common to denote this by $\mathbb{A}^n$ in order to differentiate it from the vector space K^n. In particular, we regard all points of $\mathbb{A}^n$ equally, whereas in the vector space K^n the zero is uniquely determined. Later we will also endow $\mathbb{A}^n$ with a topology.

We can regard polynomials $f \in R_n := K[X_1, \ldots, X_n]$ as **functions**

$$f \colon \mathbb{A}^n \to K, \quad \underline{a} \mapsto f(\underline{a}).$$

Observe that $(fg)(\underline{a}) = f(\underline{a})g(\underline{a})$ and $(f+g)(\underline{a}) = f(\underline{a}) + g(\underline{a})$. Also, $X_i(\underline{a}) = a_i$, so we call X_i the i-th **co-ordinate function**.

Now, given a set of polynomials $S \subset K[X_1, \ldots, X_n]$, we can consider the common zeros of all the polynomials in S, yielding the subset

$$\mathbb{V}(S) := \{\underline{a} \in \mathbb{A}^n : f(\underline{a}) = 0 \text{ for all } f \in S\}.$$

Note that if $I := (S)$ is the ideal generated by S, then $\mathbb{V}(S) = \mathbb{V}(I)$.

We define an **affine variety** to be a subset $V \subset \mathbb{A}^n$ of the form $V = \mathbb{V}(I)$ for some ideal $I \lhd R_n$. If $V = \mathbb{V}(f)$ for some non-constant polynomial $f \in R_n$, then we call V a **hypersurface**. If $n = 2$, then we also call a hypersurface a **plane curve**.

We have some easy properties of the map $\mathbb{V}$.

Proposition 13.1. *The map $\mathbb{V}$ from ideals of R_n to subsets of $\mathbb{A}^n$ satisfies:*

(1) $\mathbb{V}(0) = \mathbb{A}^n$ *and* $\mathbb{V}(R_n) = \emptyset$.

(2) $I \subset J$ *implies* $\mathbb{V}(I) \supset \mathbb{V}(J)$.

(3) $\mathbb{V}(I) = \mathbb{V}(\mathrm{rad}(I))$.

(4) $\mathbb{V}(IJ) = \mathbb{V}(I \cap J) = \mathbb{V}(I) \cup \mathbb{V}(J)$.

(5) $\mathbb{V}(\sum_i I_i) = \bigcap_i \mathbb{V}(I_i)$.

Proof. (1) This is immediate.

(2) If $\underline{a} \in \mathbb{V}(J)$, then $f(\underline{a}) = 0$ for all $f \in I$, so $\underline{a} \in \mathbb{V}(I)$.

(3) Since $I \subset \mathrm{rad}(I)$ we have $\mathbb{V}(I) \supset \mathbb{V}(\mathrm{rad}(I))$. On the other hand, if $\underline{a} \in \mathbb{V}(I)$ and $f \in \mathrm{rad}(I)$, then $f^n \in I$ for some n, so $f(\underline{a})^n = f^n(\underline{a}) = 0$. Hence $f(\underline{a}) = 0$, and so $\underline{a} \in \mathbb{V}(\mathrm{rad}(I))$.

(4) We have $\mathrm{rad}(IJ) = \mathrm{rad}(I \cap J) = \mathrm{rad}(I) \cap \mathrm{rad}(J)$ by Lemma 2.6, so $\mathbb{V}(IJ) = \mathbb{V}(I \cap J) \supset \mathbb{V}(I) \cup \mathbb{V}(J)$, by (2) and (3). Conversely, take $\underline{a} \in \mathbb{V}(IJ) - \mathbb{V}(I)$, so there exists $f \in I$ with $f(\underline{a}) \neq 0$. Then for all $g \in J$ we have $fg \in IJ$, so $f(\underline{a})g(\underline{a}) = 0$ and hence $g(\underline{a}) = 0$. Thus $\underline{a} \in \mathbb{V}(J)$.

(5) Again, $I_j \subset \sum_i I_i$ for all j, so $\bigcap_j \mathbb{V}(I_j) \supset \mathbb{V}(\sum_i I_i)$ by (2). Conversely, let $\underline{a} \in \bigcap_i \mathbb{V}(I_i)$. Then, for each $f \in \sum_i I_i$, we can write $f = f_1 + \cdots + f_n$ with $f_r \in I_{i_r}$, so $f_r(\underline{a}) = 0$, and hence $f(\underline{a}) = 0$. Thus $\underline{a} \in \mathbb{V}(\sum_i I_i)$. □

As a consequence we see that finite unions and arbitrary intersections of affine varieties are again affine varieties. Also, since R_n is Noetherian, we know that every ideal is finitely generated, and if $I = (f_1, \ldots, f_r)$, then $\mathbb{V}(I) = \mathbb{V}(f_1) \cap \cdots \cap \mathbb{V}(f_r)$. Thus every affine variety is a finite intersection of hypersurfaces.

The converse to (2) does not hold in general, just for radical ideals, as we shall see in the Nullstellensatz. To investigate this, consider for each subset $U \subset \mathbb{A}^n$ the functions which vanish on U

$$\mathbb{I}(U) := \{f \in R_n : f(\underline{a}) = 0 \text{ for all } \underline{a} \in U\}.$$

It is clear that $\mathbb{I}(U)$ is always an ideal of R_n. In fact it is a radical ideal, since if $f^r \in \mathbb{I}(U)$, then $f(\underline{a})^r = 0$ for all $\underline{a} \in U$, so $f(\underline{a}) = 0$ and $f \in \mathbb{I}(U)$.

The map $\mathbb{I}$ from subsets of $\mathbb{A}^n$ to radical ideals of R_n satisfies properties that are dual to those for $\mathbb{V}$. In particular, if $U \subset V$, then $\mathbb{I}(U) \supset \mathbb{I}(V)$. In fact, we have the following result.

Lemma 13.2. *The maps $\mathbb{V}$ and $\mathbb{I}$ are inclusion reversing maps between the posets of ideals of R_n and subsets of $\mathbb{A}^n$. Moreover, they satisfy $\mathbb{V}\mathbb{I}(U) \supset U$ and $\mathbb{I}\mathbb{V}(I) \supset I$ for all subsets $U \subset \mathbb{A}^n$ and all ideals $I \lhd R_n$.*

It follows that $\mathbb{V}\mathbb{I}\mathbb{V} = \mathbb{V}$ and $\mathbb{I}\mathbb{V}\mathbb{I} = \mathbb{I}$, and hence $\mathbb{V}$ and $\mathbb{I}$ induce inverse bijections on their respective images.[13]

Proof. Take $U \subset \mathbb{A}^n$. We need to show that $f(\underline{a}) = 0$ for all $\underline{a} \in U$ and all $f \in \mathbb{I}(U)$. This is then clear.

Similarly. given $I \lhd R_n$, we need to show that $f(\underline{a}) = 0$ for all $f \in I$ and all $\underline{a} \in \mathbb{V}(I)$. Again, this is clear.

We now have that $\mathbb{V}\big(\mathbb{I}\mathbb{V}(I)\big) \subset \mathbb{V}(I) \subset \mathbb{V}\mathbb{I}\big(\mathbb{V}(I)\big)$, so $\mathbb{V}\mathbb{I}\mathbb{V} = \mathbb{V}$. Similarly $\mathbb{I}\mathbb{V}\mathbb{I} = \mathbb{I}$. □

By definition the image of $\mathbb{V}$ is the set of all affine subvarieties, whereas the Nullstellensatz proves that the image of $\mathbb{I}$ is precisely the set of radical ideals.

[13] A pair of inclusion reversing bijections between posets is often called a **Galois Correspondence**, after the famous situation from Galois Theory involving the intermediate fields of a finite separable extension and subgroups of the Galois group.

13.1 The Nullstellensatz

Lemma 13.3. *Let $I \lhd R_n$ be a radical ideal. Then $I = \bigcap_{I \subset \mathfrak{m}} \mathfrak{m}$ is the intersection of all maximal ideals containing I.*[14]

Proof. Clearly I is contained in the intersection. For the converse we need to show that if $f \in R_n - I$, then there exists a maximal ideal containing I but not containing f.

Now, ideals containing I are in bijection with ideals of the quotient $\bar{R} := R_n/I$, and those not containing f are then in bijection with ideals of the localisation $\bar{R}_{\bar{f}}$. Now, since I is radical, $\bar{f}$ is not nilpotent in $\bar{R}$, and hence $\bar{R}_{\bar{f}} \neq 0$. It therefore contains a maximal ideal $\bar{\mathfrak{m}}$, corresponding to a prime ideal $\mathfrak{m} \lhd R_n$, say.

We need to show that $\mathfrak{m}$ is maximal. Now $\bar{R}_{\bar{f}} = \bar{R}[T]/(\bar{f}T - 1)$ is finitely generated over K, so by the Weak Nullstellensatz we must have $\bar{R}_{\bar{f}}/\bar{\mathfrak{m}} = K$. Thus $K \rightarrowtail R_n/\mathfrak{m} \rightarrowtail \bar{R}_{\bar{f}}/\bar{\mathfrak{m}} = K$, and hence $R_n/\mathfrak{m} = K$. Thus $\mathfrak{m}$ is a maximal ideal, as required. □

Theorem 13.4 (Nullstellensatz)**.** *The maps $\mathbb{V}$ and $\mathbb{I}$ give mutually inverse, inclusion-reversing bijections between the radical ideals of R_n and the affine subvarieties of $\mathbb{A}^n$. In particular there is a bijection between the points of $\mathbb{A}^n$ and the maximal ideals of R_n, given by $\underline{a} \mapsto \mathfrak{m}_{\underline{a}} := (X_1 - a_1, \ldots, X_n - a_n)$.*

Proof. By the Weak Nullstellensatz, Theorem 12.14 we know that every maximal ideal of R_n is of the form $\mathfrak{m}_{\underline{a}} = (X_1 - a_1, \ldots, X_n - a_n)$, and clearly $\mathbb{V}(\mathfrak{m}_{\underline{a}}) = \{\underline{a}\}$. On the other hand, $\mathfrak{m}_{\underline{a}} \subset \mathbb{I}(\{\underline{a}\}) \neq R_n$, so that $\mathfrak{m}_{\underline{a}} = \mathbb{I}(\{\underline{a}\})$. Thus $\mathbb{V}$ and $\mathbb{I}$ give mutually inverse bijections between the maximal ideals of R_n and the points of $\mathbb{A}^n$.

For the first statement, using that we have a Galois Correspondnece, it is enough to show that the image of $\mathbb{I}$ is precisely the set of radical ideals. So, let I be a radical ideal. A maximal ideal $\mathfrak{m}_{\underline{a}}$ contains I if and only if $\underline{a} \in \mathbb{V}(I)$. It follows that

$$I \subset \mathbb{IV}(I) \subset \bigcap_{\underline{a} \in \mathbb{V}(I)} \mathfrak{m}_{\underline{a}} = I,$$

using the previous lemma. Thus $I = \mathbb{IV}(I)$ as required. □

We note that $\mathbb{IV}(S) = \mathrm{rad}(S)$ sends a subset $S \subset R_n$ to the radical of the ideal (S).

13.2 Zariski topology

We have seen that the collection of affine subvarieties of $\mathbb{A}^n$ is closed under finite unions and arbitrary intersections. It follows that the affine subvarieties form the closed subsets of a topology on $\mathbb{A}^n$, called the **Zariski topology**. In particular, if $U \subset \mathbb{A}^n$, then $\mathbb{VI}(U) = \bar{U}$ is the Zariski closure of U.

The open sets are the complements of the closed sets. In particular, the complement of the hypersurface $\mathbb{V}(f)$ is the **distinguished open set**

$$D(f) := \{\underline{a} : f(\underline{a}) \neq 0\}.$$

[14] A ring satisfying this property is called a **Jacobson ring**. In fact, if R is Jacobson, then so too is any finitely generated R-algebra. Note that not all Noetherian rings are Jacobson, and not every Jacobson ring is Noetherian.

Note also that $D(f) \cap D(g) = D(fg)$. That every affine variety (closed set) is a finite intersection of hypersurfaces is dual to the statement that every open set is a finite union of distinguished open sets. Thus the distinguished open sets form a **base** for the Zariski topology, which is furthermore closed under intersections.

Example. (1) A proper closed subset of $\mathbb{A}^1$ is of the form $\mathbb{V}(f)$ for some polynomial $f \in K[X]$, and is therefore a finite set of points. Conversely, each finite set of points is of this form, so is closed. Thus the Zariski topology on $\mathbb{A}^1$ is just the cofinite topology.

(2) The topology on $\mathbb{A}^2$ is not the propduct topology coming from $\mathbb{A}^2 \cong \mathbb{A}^1 \times \mathbb{A}^1$. Since the topology on $\mathbb{A}^1$ is the cofinite topology, the proper closed sets of the product topology on $\mathbb{A}^2$ are finite unions of horizontal and vertical lines, and points. Clearly the Zariski topology has many more closed sets, including diagonal lines $\mathbb{V}(X - Y)$.

Recall that a (non-empty) topological space is called

Noetherian	if it satisfies the ascending chain condition on open sets (equivalently the descending chain condition on closed sets).
quasi-compact	if every open cover has a finite subcover.
irreducible	if, whenever $X = V_1 \cup V_2$ for some closed subsets V_i, then $X = V_1$ or $X = V_2$.

We will see in the exercises that a space is Noetherian if and only if every subset (with the induced topology) is quasi-compact. We also recall that X is irreducible if and only if every pair of non-empty open subsets intersect non-trivially, or equivalently if every non-empty open subset is dense. Finally the **irreducible components** of X are the maximal closed irreducible subsets.

Lemma 13.5. *Affine space $\mathbb{A}^n$ is Noetherian. In particular, every subset of $\mathbb{A}^n$ is quasi-compact.*

Proof. Suppose we have a descending chain of closed sets $V_1 \supset V_2 \supset V_3 \supset \cdots$. By the Nullstellensatz we obtain an ascending chain of radical ideals $I_1 \subset I_2 \subset I_3 \subset \cdots$, where $I_i := \mathbb{I}(V_i)$. Since R_n is a Noetherian ring this chain is stationary, and since $V_i = \mathbb{V}(I_i)$ we see that the chain of closed sets is stationary. □

Proposition 13.6. *Let $V \subset \mathbb{A}^n$ be an affine variety. Then V is irreducible if and only if $\mathbb{I}(V) \lhd R_n$ is a prime ideal. In particular, $\mathbb{A}^n$ itself is irreducible.*

More generally, writing $\mathbb{I}(V) = \mathfrak{p}_1 \cap \cdots \cap \mathfrak{p}_r$ with $\mathfrak{p}_i$ prime ideals and $\mathfrak{p}_i \not\subset \mathfrak{p}_j$ for all $i \neq j$, then the irreducible components of V are the $V_i := \mathbb{V}(\mathfrak{p}_i)$.

Proof. Let $V, V_i \subset \mathbb{A}^n$ be affine varieties, with corresponding radical ideals $I, I_i \lhd R_n$.

We know that $V = V_1 \cup V_2$ if and only if $I = I_1 \cap I_2$. Thus V is irreducible if and only if, whenever $I = I_1 \cap I_2$, we have $I = I_1$ or $I = I_2$. Such an ideal I is called **irreducible**, and we therefore need to show that the irreducible radical ideals are precisely the prime ideals. Now, every prime ideal is irreducible by Proposition 2.3. Conversely, if I is radical and irreducible, then by Proposition 11.7 we can write $I = \mathfrak{p}_1 \cap \cdots \cap \mathfrak{p}_r$ as a finite intersection of prime ideals, and by irreducibility $I = \mathfrak{p}_i$ for some i.

In general, write $I = \mathfrak{p}_1 \cap \cdots \cap \mathfrak{p}_r$ as a finite intersection of primes. We may assume that this expression is minimal, so $\mathfrak{p}_i \not\subset \mathfrak{p}_j$ for all $i \neq j$. Clearly each $\mathbb{V}(\mathfrak{p}_i) \subset V$ is closed and irreducible and $V = \mathbb{V}(\mathfrak{p}_1) \cup \cdots \cup \mathbb{V}(\mathfrak{p}_r)$. Moreover, $\mathbb{V}(\mathfrak{p}_i) \not\subset \mathbb{V}(\mathfrak{p}_j)$ for all $i \neq j$. Conversely, if $U \subset V$ is closed and irreducible, then $\mathbb{I}(U) = \mathfrak{p}$ is prime and contains I. By Proposition 2.3 again we have $\mathfrak{p} \supset \mathfrak{p}_i$ for some i, and hence $U \subset \mathbb{V}(\mathfrak{p}_i)$. It folllows that the $\mathbb{V}(\mathfrak{p}_i)$ are precisely the irreducible components of V. □

Let X be a Noetherian topological space. Then we define $\dim X$ to be the supremum of lengths of chains of closed irreducible subsets. We say that X is catenary provided every saturated chain of closed irreducible subsets has length $\dim X$.

Theorem 13.7. *Let $V \subset \mathbb{A}^n$ be an affine variety. Then $\dim V$ equals the Krull dimension of $R_n/\mathbb{I}(V)$. In particular, $\dim \mathbb{A}^n = n$.*

Moreover, if V is irreducible, then it is catenary, and if $f \in R_n/\mathbb{I}(V)$ is non-zero and not a unit, then every irreducible component of $\mathbb{V}(F) \cap V$ has dimension $\dim V - 1$.

Finally, if $U \subset V$ is open and dense, then $\dim U = \dim V$.

Proof. By the previous result and the Nullstellensatz we get an inclusion-reversing bijection between the irreducible subvarieties of V and the prime ideals of R_n containing $\mathbb{I}(V)$, and hence to the prime ideals of $R_n/\mathbb{I}(V)$. Thus $\dim V = \dim R_n/\mathbb{I}(V)$, and in particular $\dim \mathbb{A}^n = \dim R_n = n$ by Proposition 12.16. The second statement is now the geometric interpretation of Theorem 12.18.

Finally, suppose V is irreducible and $f \in K[V]$ is non-zero. Then $\dim D(f) = \dim V$ by Lemma 12.17. Since every open is a finite union of such distinguished opens, we must have $\dim U = \dim V$ for all non-empty opens $U \subset V$. In general, if $V = V_1 \cup \cdots \cup V_r$ is the decomposition into irreducible components, then $\dim V = \max\{\dim V_i\}$, and if $U \subset V$ is open and dense, then $U \cap V_i$ is open and dense in V_i, so $\dim U = \dim V$. □

13.3 Rational curves

Consider a degree two polynomial in $K[X, Y]$. If the characteristic is not 2, then we can write the homogeneous part as $(X - aY)(X - bY)$, and so by a homogeneous change of variables we have either Y^2 or XY. then, by considering $X + c$ and $Y + d$, we see that there are essentially only three quadratics to consider: the hyperbola $XY - a$, the parabola $Y^2 - X$, and a pair of lines $Y^2 - a$. (When $a = 0$ we also have the degenerate cases of crossing lines XY, and a doubled line Y^2.)

The hyperbola and the parabola are irreducible **conics** (as is the circle $X^2 + Y^2 - a$, which can be transformed into a hyperbola using $(X + iY)(X - iY)$). A conic C has the nice property that we can parameterise (almost all of) its points; more prcecisely, there exist non-constant rational functions $p, q \in K(T)$ such that $(p(t), q(t)) \in C$ whenever this point is defined. An irreducible curve with this property is called **rational**.

To see why this holds for a conic $C = \mathbb{V}(f)$, fix a point $(\bar{x}, \bar{y}) \in C$. For each t consider the line through $(\bar{x}, \bar{y})$ with gradient t

$$L_t := \{(x, y) : y - \bar{y} = t(x - \bar{x})\} = \mathbb{V}((Y - tX) - (\bar{y} - t\bar{x})).$$

To find the points in the intersection $L_t \cap C$, we make the substitution $Y = tX + (\bar{y} - t\bar{x})$ and then solve $f(X, tX + (\bar{y} - t\bar{x})) = 0$. This is a polynomial in X of degree at most two, and is not identically zero (else $L_t \subset C$ and f is divisible by $Y - tX - \bar{y} + t\bar{x}$).

We know that $\bar{x}$ is one solution, so there exists at most one other solution x_t, which we can write as a rational function of t. We then get $y_t = tx_t + (\bar{y} - t\bar{x})$, and $L_t \cap C = \{(\bar{x}, \bar{y}), (x_t, y_t)\}$. The map $t \mapsto (x_t, y_t)$ is defined on a Zariski open set of $\mathbb{A}^1$, so for all but finitely many t.

Conversely, given a point $(x, y) \in C$ with $x \neq \bar{x}$, then the line joining these two points must have some gradient $t \in K$, and so $(x, y) = (x_t, y_t)$. Hence the map $C \setminus \{x = \bar{x}\} \to \mathbb{A}^1$, $(x, y) \mapsto \frac{y-\bar{y}}{x-\bar{x}}$, is a rational function, and inverse to the first map.

The points of C which are missed are those of the form $(\bar{x}, y)$. The same reasoning applied to the vertical line

$$L_\infty := \{(x, y) : x = \bar{x}\} = \mathbb{V}(X - \bar{x})$$

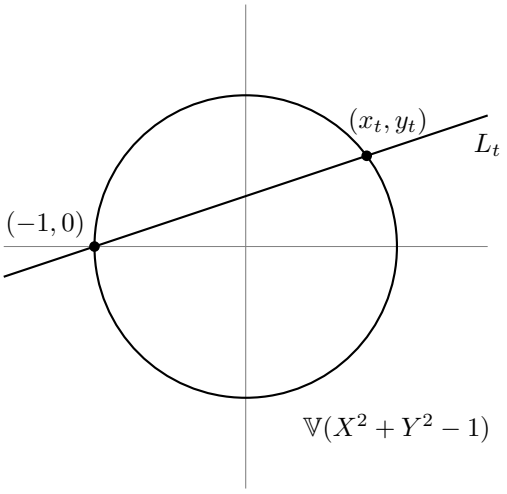
shows that there are at most two such points. Thus we have parameterised all but at most two points of C.

An important observation is that if the coefficients of f lie in some subfield $k \subset K$ and $\bar{x}, \bar{y}$ are chosen to also lie in k, then any other point defined over k will give a gradient $t \in k$. In other words, we can restrict our parameterisation to k to obtain (almost) all the k-rational points.

For example, let $C = \mathbb{V}(X^2 + Y^2 - 1)$ be the circle and take $(\bar{x}, \bar{y}) = (-1, 0)$. Then $(x_t, y_t) = \left(\frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2}\right)$. This gives a parameterisation of all points of C except $(-1, 0)$.

Since both the curve C and the point $(-1, 0)$ are defined over $\mathbb{Q}$, if we restrict to $t \in \mathbb{Q}$, then we get a parameterisation of all rational points on the circle, except $(-1, 0)$. In fact, writing $t = b/a$ with $a, b \in \mathbb{Z}$, we get a parameterisation of all Pythagorean triples

$$(x, y, z) = s(a^2 - b^2, 2ab, a^2 + b^2), \quad a, b, s \in \mathbb{Z}.$$



In fact, we will see later that if we 'complete' the picture by adding a 'point at infinity', then there is essentially only one conic. For, we make our function f homogeneous of degree two by adding in a new variable Z. Then the hyperbola becomes $XY - Z^2$ and the parabola becomes $Y^2 - XZ$

In other words, all irreducible conics in the projective plane $\mathbb{P}^2$ are isomorphic to the projective line $\mathbb{P}^1$.

The **twisted cubic** curve

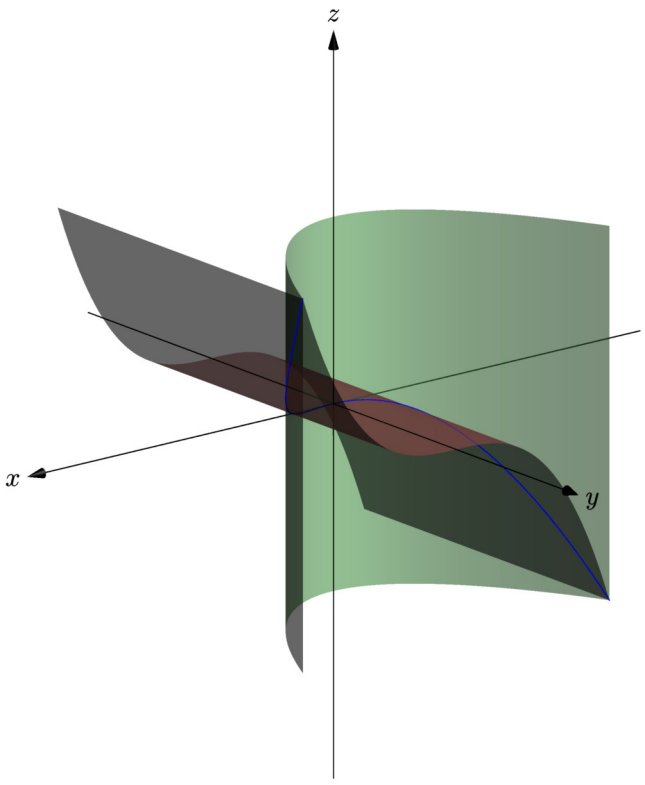
$$V = \{(t, t^2, t^3) : t \in K\} \subset \mathbb{A}^3.$$

is also an irreducible rational curve. Note that it is clearly rational, so we just need to show that it is an irreducible variety. We do this by computing the ideal $\mathbb{I}(V)$. Clearly $I := (Y - X^2, Z - X^3)$ is contained in $\mathbb{I}(V)$, and since $K[X, Y, Z]/I \cong K[X]$ we see that I is prime. Conversely, if $(x, y, z) \in \mathbb{V}(I)$, then $y = x^2$ and $z = x^3$, so $(x, y, z) = (x, x^2, x^3) \in V$. Thus $V = \mathbb{V}(I)$ is an irreducible variety, and $I = \mathbb{I}(V)$.

Consider instead the rational curve

$$W := \{(t^3, t^5, t^7) : t \in K\} \subset \mathbb{A}^3.$$

In this case $\mathbb{I}(W) = (Y^2 - XZ, X^4 - YZ, Z^2 - X^3Y)$ and $W = \mathbb{V}\mathbb{I}(W)$, so that W is indeed a variety. Note that we need three generators to describe $\mathbb{I}(W)$, rather than the two that one might expect. In fact one usually needs more generators to describe the ideal than the height of the ideal (which equals the codimension of the variety).



The twisted cubic as an intersection of two surfaces.

We next consider a plane curve of the form $V = \mathbb{V}(Y^2 - f(X))$ where $f \in K[X]$ has degree three. Again, by shifting $X \mapsto X - a$ and rescaling we have essentially three

cases: the **cuspidal cubic** Y^2-X^3, the **nodal cubic** $Y^2-X^2(X-1)$, and an **elliptic curve** $Y^2-X(X-1)(X-\lambda)$ with $\lambda \neq 0,1$. (If $\lambda = 0,1$, then we have one root with multiplicity two, so a nodal cubic.)

Now, a cuspidal or a nodal cubic is also rational. For, we can use as our base point the 'singular' point on the curve. This is $(0,0)$ for both the cubics $\mathbb{V}(Y^2-X^3)$ and $\mathbb{V}(Y^2-X^2(X-1))$. After making the substitution $Y = tX + (\bar{y} - t\bar{x})$, the polynomial $Y^2 - f(X)$ becomes a cubic in X having $\bar{x}$ as a double root. Hence we can solve for the third root uniquely.

For example, the cuspidal cubic $\mathbb{V}(Y^2-X^3)$ has the rational parameterisation

$$(x_t, y_t) = (t^2, t^3), \quad t \in K.$$

This is even a bijection, and also a bijection on every subfield of K.

Similarly, the nodal cubic $\mathbb{V}(Y^2-X^2(X-1))$ has the rational parameterisation

$$(x_t, y_t) = (t^2+1, t(t^2+1)), \quad t \in K.$$

This is almost a bijection: it is onto, and only fails to be injective since $t = \pm i$ both map to $(0,0)$. In particular, for any subfield $k \subset K$ we obtain all the k-rational points except for $(0,0)$ itself by taking $t \in k \setminus \{\pm i\}$.

On the other hand, no elliptic curve is rational, which has important consequences for elliptic key cryptography. To see this, consider the elliptic curve

$$E := \mathbb{V}\big(Y^2 - X(X-1)(X-\lambda)\big) \subset \mathbb{A}^2 \quad \text{with } \lambda \neq 0,1 \text{ and } \operatorname{char}(K) \neq 2.$$

We have to show that if $p, q \in K(T)$ satisfy $q^2 = p(p-1)(p-\lambda)$ in $K(T)$, then $p, q \in K$ are constant. Write $p = a/b$ and $q = c/d$ with $a, b, c, d \in K[T]$ such that a and b are coprime, as are c and d. Clearing denominators then gives

$$c^2 b^3 = d^2 a(a-b)(a-\lambda b).$$

Using that c and d are coprime we get

$$b^3 = ud^2 \quad \text{and} \quad uc^2 = a(a-b)(a-\lambda b).$$

Then any irreducible factor of u must divide both a and b, so a, b coprime implies that u is a unit. Moreover, since a, $a-b$ and $a-\lambda b$ are pairwise coprime, these three must also all be squares in $K[T]$.

We now prove that $a, b \in K$ are constant, whence $c, d \in K$ are also constant. Our hypothesis is that a and b are coprime, and that four distinct linear combinations of a and b are squares in $K[T]$, namely $a, b, a-b, a-\lambda b$.

Write $a = x^2$ and $b = y^2$, and also $\lambda = \mu^2$. Then x and y are again coprime, and

$$a - b = x^2 - y^2 = (x-y)(x+y) \quad \text{and} \quad a - \lambda b = (x - \mu y)(x + \mu y).$$

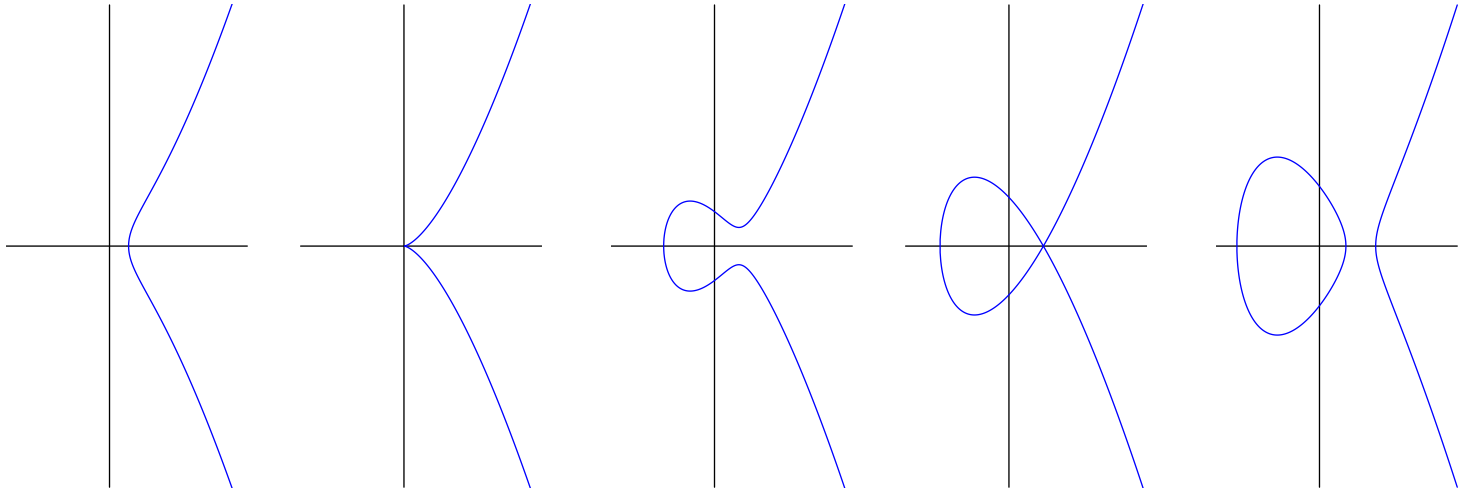
Since $a-b$ and $a-\lambda b$ are also squares, we deduce that

$$x - y, \quad x + y, \quad x - \mu y, \quad x + \mu y$$

are again all squares in $K[T]$. We are in the same situation as we started with, but with $\deg(x) = \frac{1}{2}\deg(a)$ and $\deg(y) = \frac{1}{2}\deg(b)$. We deduce that $\deg(a) = \deg(b) = 0$, and so $a, b \in K$ are constant.[15]

[15] This is Fermat's method of infinite descent.

Some examples of plane cubic curves



$y^2 = x^3 + 1.5x - 1$ $y^2 = x^3$ $y^2 = x^3 - 1.5x + 1$ $y^2 = x^3 - 3x + 2$ $y^2 = x^3 - 4.5x + 3$

In general, a plane curve of the form $V = \mathbb{V}(Y^2 - f(X))$ is called an elliptic curve if $\deg(f) = 3, 4$. The reason for this is that we can again pass to the projective plane by making our equation homogeneous. If $\deg f = 4$, then by shifting X we may assume that f has no constant term. Making the equation homogeneous gives $Y^2Z^2 - Z^4 f(X/Z)$, and then substituting $(1, Y/Z, Z)$ we recover an equation of the form $Y^2 = g(Z)$ with g of degree three.

A plane curve of the form $Y^2 - f(X)$ with $\deg f \geq 5$ is called **hyperelliptic**, and has **genus** $g := \lfloor \frac{\deg(f)-1}{2} \rfloor$. Thus the elliptic curves are precisely those of genus one.

We have seen that rational curves allow one to parameterise their points, even over non-algebraically closed fields. Historically rational curves also played an important role in evaluating integrals. An abelian integral is one of the form

$$\int_\gamma \phi(x, y)\, dx$$

where $\gamma\colon [0, 1] \to \mathbb{C}$ is a continuous path, $\phi(X, Y)$ is a rational function, and x and y are related by some polynomial relation $f(x, y) = 0$. If the curve $\mathbb{V}(f)$ is rational, say via $t \mapsto (p(t), q(t))$, then we can make the substitution to get

$$\int_\gamma \phi(x, y)\, dx = \int_c^d \phi(p(t), q(t))p'(t)\, dt.$$

This is now a (probably rather complicated) rational function of t. Using partial fractions, however, we can solve this in terms of elementary functions.[16]

For example, to calculate the arc-length of an ellipse

$$x^2 + (y/a)^2 = 1, \quad 0 < a < 1,$$

between the points with x co-ordinates c and d, one has to evaluate the integral

$$\text{arc-length} = \int_c^d \sqrt{1 + (y')^2}\, dx.$$

Now $y' = -(a^2x)/y$, and so

$$\text{arc-length} = \int_c^d \sqrt{\frac{1 - e^2x^2}{1 - x^2}}\, dx = \int_c^d \frac{\sqrt{(1 - x^2)(1 - e^2x^2)}}{1 - x^2}\, dx,$$

where $e = \sqrt{1 - a^2}$ is the eccentricity of the ellipse. This is an abelian integral, where

$$\phi(X, Y) = \frac{Y}{1 - X^2} \quad \text{and} \quad f(X, Y) = Y^2 - (1 - X^2)(1 - e^2X^2).$$

We are therefore left trying to understand the curve $\mathbb{V}(f)$, which is elliptic and hence not rational. This helps explain why the arc length of an ellipse can not be expressed in terms of elementary functions.

Another class of examples is given by determinantal varieties. Identifying $\mathbb{A}^{n^2}$ with the set $\mathbb{M}_n(K)$ of $n \times n$ matrices, we can define the subvariety consisting of all matrices

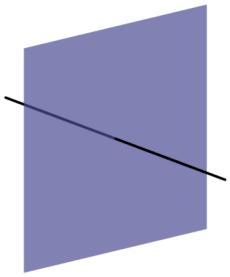
[16] A function is **elementary** if it can be written in terms of powers, roots, exponentials and logarithms, together with the field operations.

having rank at most r. For, a matrix has rank at most r if and only if all its $(r+1)$-minors vanish. Since these are clearly (homogeneous) polynomials of degree $r+1$, we see that this set is indeed closed, and hence an affine subvariety.

These varieties include $\mathbb{V}(\det)$, the set of all singular matrices. We can also define $\mathrm{SL}_n(K) = \mathbb{V}(\det - 1)$ as an affine variety, and even the sets of orthogonal, special orthogonal and symplectic matrices. On the other hand, $\mathrm{GL}_n(K) = D(\det)$ is a distinguished open subset, and hence is homeomorphic to an affine subvariety of $\mathbb{A}^{n^2+1}$ as we shall see in the example following Lemma 14.9.

Finally consider the union of a plane and a line, say given as

$$V = \mathbb{V}(X, Z) \cup \mathbb{V}(Y) = \mathbb{V}((X, Z) \cap (Y)) = \mathbb{V}(XY, ZY).$$



This has two irreducible components, namely the plane $\mathbb{V}(Y)$ and the line $\mathbb{V}(X, Z)$, corresponding to the expression $\mathbb{I}(V) = (X, Z) \cap (Y)$ as an intersection of prime ideals. The dimensions of these irreducible components are 2 and 1 respectively, so $\dim V = 2$.

14 Regular maps

14.1 Regular functions on affine varieties

Let $V \subset \mathbb{A}^m$ be an affine variety, with corresponding radical ideal $\mathbb{I}(V)$. Recall that we can regard polynomials in R_m as functions $\mathbb{A}^m \to K$. Each such polynomial restricts to a function $V \to K$, and the ideal $\mathbb{I}(V)$ is then the set of functions which are identically zero on V. We therefore define the **co-ordinate algebra** of V to be $K[V] := R_m/\mathbb{I}(V)$, and call an element of $K[V]$ a **regular function** on V. Thus a regular function on V is the restriction to V of a polynomial function on $\mathbb{A}^m$, and $K[V]$ is a finitely-generated reduced K-algebra.

The term 'co-ordinate algebra' comes from the fact that $K[V]$ is generated by the images of the X_i, which are the co-ordinate functions on $\mathbb{A}^n$.

Lemma 14.1. *The closed subsets of V correspond bijectively to the radical ideals of $K[V]$.*

Proof. Writing $I = \mathbb{I}(V)$, so that $V = \mathbb{V}(I)$, we see that the closed subsets of V are those of the form $\mathbb{V}(J)$ for some radical ideal $J \subset I$, equivalently a radical ideal of $K[V] = R_m/I$. □

14.2 Regular maps between affine varieties

Identifying $\mathbb{A}^1 = K$ we see that regular functions on V give maps $V \to \mathbb{A}^1$. In general we define a **regular map** $\phi\colon V \to \mathbb{A}^n$ to be an n-tuple of regular functions on V, so $\phi = (f_1, \ldots, f_n)$ with $f_i \in K[V]$. If $W \subset \mathbb{A}^n$ is an affine variety, then a regular map $V \to W$ is the same as a regular map $V \to \mathbb{A}^n$ having image contained in W.

Suppose we have a regular map $\phi = (f_1, \ldots, f_n)\colon V \to \mathbb{A}^n$. Then we obtain a K-algebra homomorphism

$$\phi^*\colon R_n = K[Y_1, \ldots, Y_n] \to K[V], \quad Y_i \mapsto f_i.$$

Now $\phi^*(g)$ is obtained by substituting $Y_i \mapsto f_i$, so $\phi^*(g) = g(f_1, \ldots, f_n)$. As a function on V this sends $\underline{a} \in V$ to $g(f_1(\underline{a}), \ldots, f_n(\underline{a})) = g(\phi(\underline{a}))$. Thus the composite $g\phi$ is a regular function on V, given by $\phi^*(g)$.

Note the change of direction: we have $\phi\colon V \to \mathbb{A}^n$ but $\phi^*\colon R_n \to K[V]$.

Theorem 14.2. *Let V be an affine variety. The map $\phi \mapsto \phi^*$ gives a bijection between regular maps $V \to \mathbb{A}^n$ and K-algebra homomorphisms $R_n \to K[V]$. More generally, if $W \subset \mathbb{A}^n$ is an affine variety, then we get an induced bijection between regular maps $V \to W$ and K-algebra homomorphisms $K[W] \to K[V]$.*

Proof. Write $R_n = K[Y_1, \ldots, Y_n]$, so the Y_i are the co-ordinate functions on $\mathbb{A}^n$. If $\phi\colon V \to \mathbb{A}^n$ is regular, then $\phi^*\colon R_n \to K[V]$ is the algebra homomorphism $g \mapsto g\phi$. Conversely, given $\theta\colon R_n \to K[V]$, we obtain a regular map $(f_1, \ldots, f_n)\colon V \to \mathbb{A}^n$, where $f_i = \theta(Y_i)$. It is clear that these maps are mutually inverse.

For the second part, consider $\phi\colon V \to \mathbb{A}^n$. Then $\phi(V) \subset W$ if and only if $g(\phi(\underline{a})) = 0$ for all $\underline{a} \in V$ and $g \in \mathbb{I}(W)$. This is equivalent to $\phi^*(g) = 0$ for all $g \in \mathbb{I}(W)$, and hence to $\mathbb{I}(W) \subset \mathrm{Ker}(\phi^*)$. Now apply the Factor Lemma. □

We can describe the regular map associated to an algebra homomorphism more explicitly. Recall that points $\underline{a} \in V$ correspond to maximal ideals $\mathfrak{m}_{\underline{a}} \lhd K[V]$, equivalently algebra homomorphisms $\mathrm{ev}_{\underline{a}}\colon K[V] \to K$.

Lemma 14.3. *Let $\theta\colon K[W] \to K[V]$, and $\phi\colon V \to W$ the corresponding regular map. Then $\phi(\underline{a})$ corresponds to the maximal ideal $\theta^{-1}(\mathfrak{m}_{\underline{a}})$, equivalently the algebra homomorphism $\mathrm{ev}_{\underline{a}}\,\theta$.* □

Corollary 14.4. *Let $\phi\colon U \to V$ and $\psi\colon V \to W$ be regular. Then their composition is again regular, corresponding to $(\psi\phi)^* = \phi^*\psi^*$.*

Proof. Consider the algebra homomorphism $\phi^*\psi^*\colon K[W] \to K[U]$. If $\underline{a} \in U$, then

$$\mathrm{ev}_{\underline{a}}\,\phi^*\,\psi^* = \mathrm{ev}_{\phi(\underline{a})}\,\psi^* = \mathrm{ev}_{\psi\phi(\underline{a})}.$$

Thus the corresponding regular map is precisely the composition $\psi\phi$. □

Corollary 14.5. *(1) If V is an affine variety, then a map $\phi\colon V \to \mathbb{A}^n$ is regular if and only if $Y_i\,\phi \in K[V]$ for all i, where Y_i is the i-th co-ordinate function on $\mathbb{A}^n$.*

(2) Every regular map is continuous with respect to the Zariski topology. □

Remark. It follows that the affine varieties together with the regular maps form a category. Accordingly a regular map $\phi\colon V \to W$ is an isomorphism provided there exists a regular map $\psi\colon W \to V$ such that $\psi\phi = \mathrm{id}_V$ and $\phi\psi = \mathrm{id}_W$. It follows that two affine varieties are isomorphic if and only if their co-ordinate algebras are isomorphic. Note however that a bijective regular map need not be an isomorphism.

We have already seen an example of this. Recall that we have the bijective regular map $\mathbb{A}^1 \to \mathbb{V}(Y^2 - X^3)$, $t \mapsto (t^2, t^3)$. This corresponds to the algebra homomorphism $K[X,Y]/(Y^2 - X^3) \to K[T]$, $X \mapsto T^2$, $Y \mapsto T^3$, which is injective but not an isomorphism.

In fact, there is a duality between the categories of affine varieties and affine algebras, so finitely generated, reduced K-algebras. This sends a variety V to its co-ordinate algebra $K[V]$, and a regular map $\phi\colon V \to W$ to the algebra homomorphism $\phi^*\colon K[W] \to K[V]$.

Conversely, given an affine algebra A, we fix an isomorphism $A \cong R_m/I$ for some radical ideal I, and send A to $\mathbb{V}(I) \subset \mathbb{A}^m$. If $B = R_n/J$, then an algebra homomorphism $\theta\colon A \to B$ yields a regular map $\mathbb{V}(J) \to \mathbb{V}(I)$.

These constructions yield functors, and then the duality.

A regular map $\phi\colon V \to W$ is **dominant** provided its image is dense in W.

Corollary 14.6. *Let $\phi\colon V \to W$ be a regular map of affine varieties, with corresponding algebra homomorphism $\phi^*\colon K[W] \to K[V]$.*

(1) ϕ^* is injective if and only if ϕ is dominant.

(2) If ϕ^* is surjective, then ϕ is injective.

14.3 Products of affine varieties

Recall that we have an isomorphism

$$R_m \otimes_K R_n \cong R_{m+n}, \quad X_i \otimes 1 \mapsto Z_i, \quad 1 \otimes Y_i \mapsto Z_{m+i}.$$

Under this identification, the function $f \otimes g$ acts on $\mathbb{A}^{m+n}$ as

$$(f \otimes g)(c_1, \ldots, c_{m+n}) = f(c_1, \ldots, c_m)g(c_{m+1}, \ldots, c_{m+n}).$$

In other words, this is compatible with the set-theoretic product

$$\mathbb{A}^m \times \mathbb{A}^n \cong \mathbb{A}^{m+n}, \quad (\underline{a}, \underline{b}) \mapsto (a_1, \ldots, a_m, b_1, \ldots, b_n),$$

so that $(f \otimes g)(\underline{a}, \underline{b}) = f(\underline{a})g(\underline{b})$.

In general, given subsets $V \subset \mathbb{A}^m$ and $W \subset \mathbb{A}^n$, we will define $V \times W$ to be the corresponding subset of $\mathbb{A}^{m+n}$, equipped with the subspace topology. Thus, although $V \times W$ is the set-theoretic product of V and W, the topology on $V \times W$ is not the product topology.

Theorem 14.7. *Let $V \subset \mathbb{A}^m$ and $W \subset \mathbb{A}^n$ be affine varieties. Then $V \times W \subset \mathbb{A}^{m+n}$ is again an affine variety, and $K[V \times W] \cong K[V] \otimes_K K[W]$.*

Proof. Write $K[V] = R_m/I$ and $K[W] = R_n/J$ for radical ideals I and J, and set $L := I \otimes R_n + R_m \otimes J$. Then $K[V] \otimes K[W] = R_{m+n}/L$ by Proposition 9.2, so we need to show that $V \times W = \mathbb{V}(L)$ and that $K[V] \otimes K[W]$ is reduced.

Clearly $\mathbb{V}(I \otimes R_n) = V \times \mathbb{A}^n$, and similarly $\mathbb{V}(R_m \otimes J) = \mathbb{A}^m \times W$, so $\mathbb{V}(L)$ is their intersection, which is $V \times W$.

Next, take K-bases $\{f_i\}$ for $K[V]$ and $\{g_j\}$ for $K[W]$. Then the $f_i \otimes g_j$ form a K-basis for the tensor product $K[V] \otimes K[W]$. Suppose therefore that $\sum_{i,j} \lambda_{ij} f_i \otimes g_j$ is nilpotent, where $\lambda_{ij} \in K$ are almost all zero. Given $\underline{a} \in V$, the function $\sum_{i,j} \lambda_{ij} f_i(\underline{a}) g_j \in K[W]$ must also be nilpotent, and hence zero. Since the g_j are linearly independent, we get that $\sum_i \lambda_{ij} f_i(\underline{a}) = 0$ for all j. Thus $\sum_i \lambda_{ij} f_i \in K[V]$ vanishes on every point of V, so is zero. Since the f_i are linearly independent, we deduce that $\lambda_{ij} = 0$ for all i, j. □

We now prove that the subvariety $V \times W \subset \mathbb{A}^{m+n}$ is the product of V and W in the category of affine varieties.

Proposition 14.8. *Let V_1, V_2 be affine varieties. Then the projection maps $\pi_i \colon V_1 \times V_2 \to V_i$ are regular. Moreover, given an affine variety U and regular maps $\phi_i \colon U \to V_i$, we get a unique regular map $\phi \colon U \to V_1 \times V_2$ such that $\pi_i \phi = \phi_i$.*

Proof. The projection maps π_i are clearly regular; in fact π_1^* is the natural algebra homomorphism $K[V_1] \to K[V_1] \otimes_K K[V_2]$, $f \mapsto f \otimes 1$, and similarly for π_2^*.

Suppose now that we are given an affine variety U and regular maps $\phi_i \colon U \to V_i$, corresponding to algebra homomorphisms $\phi_i^* \colon K[V_i] \to K[U]$. Then by Proposition 9.1 there is a unique algebra homomorphism $\phi^* \colon K[V_1] \otimes_K K[V_2] \to K[U]$ satisfying $\phi^* \pi_i^* = \phi_i^*$. □

Remark. Alternatively, consider the set-theoretic map $\phi(\underline{a}) = (\phi_1(\underline{a}), \phi_2(\underline{a}))$. This is clearly the unique map such that $\pi_i \phi = \phi_i$. To see that it is regular, we observe that the co-ordinate functions on $V_i \times V_2$ are just those from V_i, and so ϕ_i both regular implies ϕ is regular.

Since the category of affine varieties is dual to the category of affine K-algebras, the direct product of varieties must correspond to the coproduct of their co-ordinate algebras. We already knew that the coproduct of $K[V]$ and $K[W]$ *in the category of all K-algebras* was given by their tensor product, and that this was finitely generated. Thus the main part of both of these results was to show that the tensor product is reduced. In fact this holds over any perfect field K.[17]

Lemma 14.9. *Let V and W be affine varieties. Then the projection map $\pi \colon V \times W \to V$ is open, so sends open sets to open sets.*

Proof. It is enough to show that this image of a distinguished open set is again open, so let $F \in K[V] \otimes K[W]$. Then

$$\pi(D(F)) = \{\underline{a} \in V : \exists\, \underline{b} \in W \text{ with } F(\underline{a}, \underline{b}) \neq 0\}.$$

[17] Since K is a field, the tensor product $- \otimes_K S$ is exact. Now, every element of $R \otimes_K S$ lies in some subalgebra $R' \otimes_K S$ with $R' \leq R$ finitely generated. Since R' is reduced and Noetherian, $0 = \bigcap_i \mathfrak{p}_i$ is a finite intersection of prime ideals, so $R' \rightarrowtail \prod_i \kappa(\mathfrak{p}_i)$, where $\kappa(\mathfrak{p}_i) = \mathrm{Quot}(R/\mathfrak{p}_i)$, and hence $R' \otimes S \rightarrowtail \prod_i \kappa(\mathfrak{p}_i) \otimes_K S$. Applying the same argument to S, it is enough to prove that $L \otimes_K M$ is reduced for all finitely generated field extensions L/K and M/K. Now apply MacLane's Criterion (see for example S. Lang, Algebra VIII §4 or P. M. Cohn, Algebra Volume 3, Theorem 5.11).

Consider the algebra homomorphism

$$\mathrm{id} \otimes \mathrm{ev}_{\underline{b}} \colon K[V] \otimes K[W] \to K[V] \otimes_K K \cong K[V].$$

This sends F to $F(-, \underline{b})$, so each $F(-, \underline{b})$ is regular on V, and clearly $\pi(D(F))$ is the union of the distinguished opens $D(F(-, \underline{b}))$, hence is open. □

Example. Let V be an affine variety, and $f \in K[V]$. Then the projection $V \times \mathbb{A}^1 \to V$ sends the affine variety $\mathbb{V}(fT-1)$ bijectively onto the distinguished open $D(f)$. Here T is the co-ordinate function on $\mathbb{A}^1$. Since the projection map is continuous and open, so too is its restriction, yielding a homeomorphism $\mathbb{V}(fT-1) \cong D(f)$. Moreover, the co-ordinate algebra of $\mathbb{V}(fT-1)$ is the localisation $K[V]_f$, using that $R_a \cong R[T]/(aT-1)$.

For example, the punctured line $\mathbb{A}^1 - \{0\}$ is homeomorphic to the hyperbola $\mathbb{V}(XY-1)$, and the subset of invertible matrices $\mathrm{GL}_n(K) = D(\det) \subset \mathbb{A}^{n^2}$ is homeomorphic to the affine variety $\mathbb{V}(T \det -1)$.

Proposition 14.10. *Let $\phi \colon V \to W$ be a dominant map of affine varieties, so $\phi^* \colon K[W] \to K[V]$ is injective. If ϕ^* is integral, then ϕ is surjective with finite fibres.*

Proof. Let $\underline{b} \in W$, so that $\mathfrak{m}_{\underline{b}} \lhd K[W]$ is a maximal ideal. By Corollary 12.11 we know that the primes lying over $\mathfrak{m}_{\underline{b}}$ are in bijection with the primes of $K[V]/\mathfrak{m}_{\underline{b}}K[V]$. Writing $\mathrm{rad}(\mathfrak{m}_{\underline{b}}K[V]) = \mathfrak{p}_1 \cap \cdots \cap \mathfrak{p}_d$ as a finite intersection of primes, we see that the primes lying over $\mathfrak{m}_{\underline{b}}$ are precisely the primes containing some $\mathfrak{p}_i$.

Now, since ϕ^* is integral, there is a prime lying over $\mathfrak{m}_{\underline{b}}$, and each such prime is necessarily maximal by Theorem 12.10. Writing $\mathfrak{p}_i = \mathfrak{m}_{\underline{a}_i}$, we have $\phi^{-1}(\underline{b}) = \{\underline{a}_1, \ldots, \underline{a}_d\}$. Thus ϕ is surjective with finite fibres. □

Theorem 14.11. *Let V and W be affine varieties. Then* $\dim(V \times W) = \dim V + \dim W$.

Proof. By the Noether Normalisation Lemma we have finite extensions $R_m \rightarrowtail K[V]$ and $R_n \rightarrowtail K[W]$, and we wish to show that $\phi^* \colon R_m \otimes R_n \rightarrowtail K[V] \otimes K[W]$ is finite. The result will then follow from Proposition 12.15.

By the previous proposition the maps $V \to \mathbb{A}^m$ and $W \to \mathbb{A}^n$ are both surjective, hence the map $\phi \colon V \times W \to \mathbb{A}^{m+n}$ is also surjective, so ϕ^* is injective by Corollary 14.6. Moreover, $K[V] \otimes K[W]$ is finitely generated over K, so ϕ^* is of finite type.

Finally, every element of $K[V] \otimes K[W]$ is a sum of elements of the form $x \otimes y = (x \otimes 1)(1 \otimes y)$, and the set of integral elements forms a subalgebra by Lemma 12.5. We therefore only need to show that $x \otimes 1$ (and similarly $1 \otimes y$) is integral over R_{m+n}. This is clear: for example, if x satisfies the integral equation

$$x^r + a_{r-1}x^{r-1} + \cdots + a_0 = 0, \quad a_i \in R_m,$$

then $x \otimes 1$ satisfies the integral equation

$$(x \otimes 1)^r + (a_{r-1} \otimes 1)(x \otimes 1)^{r-1} + \cdots + (a_0 \otimes 1) = 0, \quad a_i \otimes 1 \in R_m \otimes R_n. \qquad \square$$

Theorem 14.12. *Let V and W be irreducible affine varieties. Then $V \times W$ is again irreducible.*

Proof. Exercise. □

14.4 Noether Normalisation revisited

We now wish to give a geometric interpretation of the Noether Normalisation Lemma. We recall that that finite algebra extensions give rise to maps of varieties which are surjective with finite fibres.

Theorem 14.13 (Geometric version of NNL)**.** *Let $V \subset \mathbb{A}^n$ be an affine variety. Then there exists a linear map $\pi\colon \mathbb{A}^n \to \mathbb{A}^d$ such that $\pi|_V\colon V \to \mathbb{A}^d$ is surjective and has finite fibres. Moreover, after applying a linear automorphism of $\mathbb{A}^n$, we may assume that π is the projection onto the first d co-ordinates.*

Proof. We have $K[V] = R_n/\mathbb{I}(V)$, so has generators $x_i := X_i + \mathbb{I}(V)$. Applying the Noether Normalisation Lemma we obtain, after renumbering, a finite extension $R_d \rightarrowtail K[V]$, $Y_i \mapsto x_i + \sum_{j>d} \lambda_{ij} x_j$.

Consider therefore the algebra homomorphism $R_d \to R_n$, $Y_i \mapsto X_i + \sum_{j>d} \lambda_{ij} X_j$. This corresponds to a regular map $\pi\colon \mathbb{A}^n \to \mathbb{A}^d$, sending $\underline{a} \mapsto \underline{b}$, where $b_i = a_i + \sum_j \lambda_{ij} a_j$. Thus π is linear. Moreover, the restriction $\pi|_V$ corresponds to the composition $R_d \to R_n \to K[V]$, which is our original monomorphism, and hence integral. It is therefore surjective with finite fibres by the previous theorem.

Finally, consider the linear automorphism ϕ of $\mathbb{A}^n$ given by sending $\underline{a}$ to the point $(\pi(\underline{a}), a_{d+1}, \ldots, a_n)$. Then clearly $\pi = p\phi$, where $p\colon \mathbb{A}^n \to \mathbb{A}^d$ is the projection onto the first d co-ordinates. In terms of the co-ordinate algebras we have $p^*\colon R_d \to R_n$, $Y_i \mapsto X_i$, and

$$\phi^*\colon R_n \to R_n, \quad X_i \mapsto \begin{cases} X_i + \sum_{j>d} \lambda_{ij} X_j & \text{if } 1 \leq i \leq d; \\ X_i & \text{if } d < i \leq n. \end{cases}$$

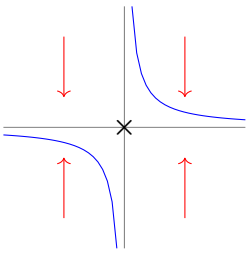
□

Example. Take $V = \mathbb{V}(XY - 1) \subset \mathbb{A}^2$, so $K[V] = K[x, y] = K[X, Y]/(XY - 1)$. The monomorphism $K[T] \to K[V]$, $T \mapsto x$, is not integral. Geometrically the projection $p\colon \mathbb{A}^2 \to \mathbb{A}^1$, $(a, b) \mapsto a$, is not surjective when restricted to V.

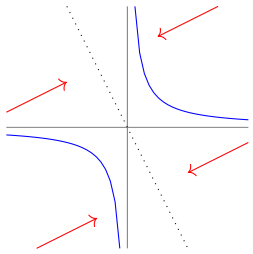
On the other hand, the monomorphism $K[T] \to K[V]$, $T \mapsto x - \varepsilon y$, is integral for all $0 \neq \varepsilon \in K$. For, if $z = x - \varepsilon y$, then $K[V] = K[y, z]/(\varepsilon y^2 + yz - 1)$. Geometrically, the projection $\pi_\varepsilon\colon (a, b) \mapsto a - \varepsilon b$, is surjective with finite fibres when restricted to V. In fact, the fibre over the point $2t$ is

$$\pi_\varepsilon^{-1}(2t) = \left\{(a, 1/a) : a = t \pm \sqrt{t^2 + \varepsilon}\right\}.$$

This consists of two points if $t^2 + \varepsilon \neq 0$, and a single point otherwise.

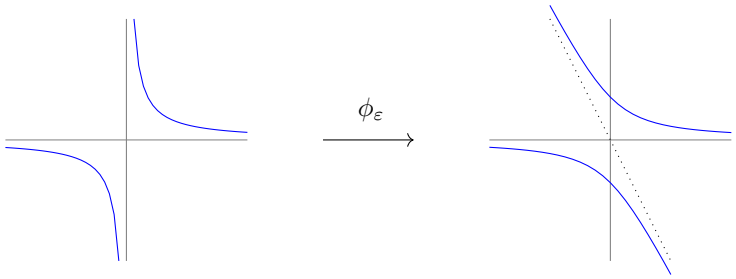


$(a, b) \mapsto a$



$(a, b) \mapsto a - \varepsilon b$

Alternatively we can first apply the horizontal shear $\phi_\varepsilon = (X - \varepsilon Y, Y)$ on $\mathbb{A}^2$. This sends V to $V_\varepsilon := \mathbb{V}(\varepsilon Y^2 + XY - 1)$, and $\pi_\varepsilon = p\phi_\varepsilon$ is the projection of V_ε onto the X-axis.



Consider instead the plane cubic $V = \mathbb{V}\big(Y^2 - X(X-1)(X-\lambda)\big) \subset \mathbb{A}^2$, so

$$K[V] = K[x, y] = K[X, Y]/(Y^2 - X(X-1)(X-\lambda)).$$

This time y is already integral over $K[x]$, and so projection onto the X-axis is a ramified covering. The fibre over a is

$$\big\{(a, b) : b = \pm\sqrt{a(a-1)(a-\lambda)}\big\}.$$

This consists of two points unless $a = 0, 1, \lambda$, in which case we get a unique point.

14.5 Automorphisms of affine space

We saw earlier that the parabola is a rational curve, in the sense that we can parameterise its points using rational functions in a single variable. More generally, we can do this for the plane curve $V = \mathbb{V}(Y - X^r) \subset \mathbb{A}^2$ via

$$\mathbb{A}^1 \to V, \quad t \mapsto (t, t^r).$$

This map is clearly bijective, and its inverse is given by the restriction to V of the projection onto the first co-ordinate

$$\mathbb{A}^2 \to \mathbb{A}^1, \quad (a, b) \mapsto a.$$

We deduce that $V \cong \mathbb{A}^1$ as affine varieties.

In fact, Abhyankar and Moh (1975) and Suzuki (1974) proved the following theorem.

Theorem 14.14. *Suppose K has characteristic zero. A curve $V \subset \mathbb{A}^2$ is isomorphic to $\mathbb{A}^1$ if and only if there exists an automorphism of $\mathbb{A}^2$ taking V to the X-axis, i.e. the line $\mathbb{V}(Y)$.*

It follows that $\mathbb{A}^2$ has a rather interesting automorphism group, since it clearly contains all **affine linear transformations**

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} p \\ q \end{pmatrix}, \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2(K), \quad (p, q) \in \mathbb{A}^2,$$

as well as all transformations of the form

$$(x, y) \mapsto (x, y + x^r), \quad r \geq 1.$$

In fact, these two types of automorphism generate the whole automorphism group for $\mathbb{A}^2$. This was first proved by Jung (1942) in characteristic zero, and by van der Kulk (1954) in arbitrary characteristic.

For general n we still have the subgroup $\mathcal{A}_n$ of affine linear transformations

$$\underline{x} \mapsto A\underline{x} + \underline{b}, \quad A \in \mathrm{GL}_n(K), \quad \underline{b} \in \mathbb{A}^n$$

and the subgroup $\mathcal{J}_n$ of **triangular automorphisms**, or **Jonquière transformations**,

$$(x_1, \ldots, x_n) \mapsto (x_1 + \phi_1, \ldots, x_n + \phi_n), \quad \phi_i \in K[X_1, \ldots, X_{i-1}].$$

Together these generate the subgroup of **tame** automorphisms of $\mathbb{A}^n$; all other automorphisms are called **wild**.

Then for $n = 2$ we have that every automorphism is tame, and the full automorphism group is the amalgamated product of $\mathcal{A}_n$ and $\mathcal{J}_n$.

For $n = 3$ it is known that there are wild automorphisms in characteristic zero, Shestakov and Umirbaev (2003).

One should compare this to the case for projective space, where every automorphism is linear (Corollary 17.6).

15 Local definition of regular maps

Recall that bijective regular maps are not necessarily isomorphisms. For example, the map $\mathbb{A}^1 \to V = \mathbb{V}(Y^2 - X^3)$, $t \mapsto (t^2, t^3)$. This corresponds to the algebra homomorphism $K[V] \to K[T]$, $x \mapsto T^2$, $y \mapsto T^3$, so the image does not contain T. Thus the inverse map $V \to \mathbb{A}^1$, $(x, y) \mapsto x/y$, is not regular.

On the other hand, suppose $V \subset \mathbb{A}^n$ is affine and $f \in R_n$ is nowhere zero on V. Then the Nullstellensatz tells us that $\mathrm{rad}(f) = K[V]$, so $f \in K[V]$ is a unit, and hence $1/f$ is regular on V. In other words, $1/f$ is a regular function on all varieties $V \subset \mathbb{A}^n$ contained in the distinguished open $D(f)$. It therefore makes sense to say that $1/f$ is a regular function on the open set $D(f)$. This is also compatible with the fact that there is a homeomorphism $D(f) \cong \mathbb{V}(fT - 1) \subset \mathbb{A}^{n+1}$ and that $\mathbb{V}(fT - 1)$ has co-ordinate algebra the localisation $R_n[\frac{1}{f}]$.

We are therefore interested in enlarging the category of affine varieties to include all open subsets of affine varieties, which we call **quasi-affine varieties**, and the regular maps between them. This will also be useful if we want to examine what is happening locally near a point, so on an open neighbourhood of a point.

15.1 Regular maps on quasi-affine varieties

Let V be an affine variety, and $U \subset V$ a locally closed subset, so the intersection of an open subset and a closed subset.

A function $\phi\colon U \to K$ is said to be **regular at a point** $\underline{a} \in U$ provided there exist $f, g \in K[V]$ with $\phi(\underline{b}) = f(\underline{b})/g(\underline{b})$ on some open subset $U' \subset D(g) \cap U$ such that $\underline{a} \in U'$. In other words, $\phi = f/g$ on some open neighbourhood of $\underline{a}$.

A function $\phi\colon U \to K$ is **regular** if it is regular at each point of U. We denote the set of regular functions by $\mathcal{O}(U)$, and call U together with $\mathcal{O}(U)$ a quasi-affine variety.

Lemma 15.1. *Let V be an affine variety, and $U \subset V$ a locally closed subset. Then $\mathcal{O}(U)$ is a ring, and restriction yields a ring homomorphism $K[V] \to \mathcal{O}(U)$.*

Proof. Take $\phi_1, \phi_2 \in \mathcal{O}(U)$, and $\underline{a} \in U$. Then we can find open neighbourhoods U_i of $\underline{a}$ and $f_i, g_i \in K[V]$ with $\phi_i = f_i/g_i$ on U_i. It follows that $U := U_1 \cap U_2$ is again an open neighbourhood of $\underline{a}$ with $\phi_1 + \phi_2 = (f_1 g_2 + f_2 g_1)/g_1 g_2$ and $\phi_1 \phi_2 = f_1 f_2 / g_1 g_2$ in U. Thus $\mathcal{O}(U)$ is a ring, and clearly restriction yields a ring homomorphism $K[V] \to \mathcal{O}(U)$. □

Theorem 15.2. *Let V be an affine variety, and $g \in K[V]$. Then $\mathcal{O}(D(g)) = K[V][1/g]$. In particular, setting $g = 1$ gives $\mathcal{O}(V) = K[V]$.*

Proof. We know that restriction yields an algebra homomorphism $K[V] \to \mathcal{O}(D(g))$, and since the function $\underline{a} \mapsto 1/g(\underline{a})$ is regular on $D(g)$, we have an induced algebra homomorphism $K[V][1/g] \to \mathcal{O}(D(g))$. If $f \in K[V]$ is zero on $D(g)$, then fg is zero on all of V, and hence $fg = 0$ in $K[V]$, so the map is injective by Lemma 3.4. It remains to prove surjectivity, so let ϕ be regular on $D(g)$.

For each $\underline{a} \in D(fg)$ we can find $p', q' \in K[V]$ with $q'(\underline{a}) \neq 0$ and $\phi = p'/q'$ on an open neighbourhood U of $\underline{a}$. Thus U is open in $D(q') \cap D(g) = D(q'g)$, and so we may assume that $U = D(q'gh)$ for some $h \in K[V]$ with $h(\underline{a}) \neq 0$. Writing $p = p'gh$ and $q = q'gh$, we have thus shown that for each $\underline{a} \in D(g)$ there exist $p, q \in K[V]$ with $\phi = p/q$ on $D(q) \subset D(g)$.

Doing this for each point of $D(g)$, we obtain an open cover of $D(g)$, which by quasi-compactness must contain a finite subcover. We therefore have $D(g) = D(q_1) \cup \cdots \cup D(q_n)$, and $\phi = p_i/q_i$ on $D(q_i)$.

Next, on $D(q_i) \cap D(q_j) = D(q_i q_j)$ we have $p_i/q_i = p_j/q_j$, equivalently $p_i q_j = p_j q_i$. Now, $V = D(q_i q_j) \cup \mathbb{V}(q_i q_j)$, from which it follows that $(p_i q_j - p_j q_i)(q_i q_j) = 0$ on all of V, and hence that $p_i q_i q_j^2 = p_j q_j q_i^2$ in $K[V]$. Since $D(q_i^2) = D(q_i)$, we can thus replace p_i by $p_i q_i$ and q_i by q_i^2.

We now have $D(g) = D(q_1) \cup \cdots \cup D(q_n)$, $\phi = p_i/q_i$ on $D(q_i)$, and $p_i q_j = p_j q_i \in K[V]$ for all i, j. Observe that taking complementary sets gives $\mathbb{V}(g) = \mathbb{V}(q_1, \ldots, q_n)$, so by the Nullstellensatz $\operatorname{rad}(g) = \operatorname{rad}(q_1, \ldots, q_n)$. In particular $g^d \in (q_1, \ldots, q_n)$ for some d, say $g^d = \sum_j q_j r_j$. Set $f := \sum_j p_j r_j$. Then

$$f q_i = \sum_j p_j q_i r_j = \sum_j p_i q_j r_j = p_i g^d,$$

so that $p_i/q_i = f/g^d$ on $D(q_i)$. This holds for all i, so that $\phi = f/g^d \in K[V][1/g]$. □

Recall from Lemma 14.9 that the projection map $V \times \mathbb{A}^1 \to V$ is open and restricts to a homeomorphism $\mathbb{V}(gT - 1) \to D(g)$.

Corollary 15.3. *Let V be an affine variety and $g \in K[V]$. A function $D(g) \to K$ is regular if and only if the composition $\mathbb{V}(gT - 1) \to D(g) \to K$ is regular.*

Proof. This follows since $K[\mathbb{V}(gT-1)] = K[V][1/g] = \mathcal{O}(D(g))$. □

Since restrictions of regular functions are clearly regular, we see that if U is an open subset of an affine variety V, say $U = \bigcup_i D(g_i)$ for some $g_i \in K[V]$, then $\phi\colon U \to K$ is regular if and only if its restriction to $D(g_i)$ lies in $K[V][1/g_i]$ for each i.

Let U be a quasi-affine variety. A map $\phi = (\phi_1, \ldots, \phi_n)\colon U \to \mathbb{A}^n$ is **regular** provided each $\phi_i\colon U \to K$ is regular.

Corollary 15.4. *Regular maps are continuous, and the composition of two regular maps is again regular.*

Proof. Let $\phi\colon D(f) \to \mathbb{A}^m$ be regular, where $f \in K[\bar{U}]$. The composition $\mathbb{V}(fT-1) \to D(f) \to \mathbb{A}^m$ is regular, so continuous, and hence ϕ is continuous. In general, a regular map $\phi\colon U \to \mathbb{A}^m$ is locally of this form, and hence continuous.

Now suppose $\mathrm{Im}(f) \subset V$ and $\psi\colon V \to \mathbb{A}^n$ is regular. Take $g_0 \in K[\bar{V}]$ such that $D(g_0) \subset V$. Then $\phi^{-1}(D(g_0))$ is open, so contains some $D(f_0)$ with $f \in K[\bar{U}]$. Write $\phi = \frac{1}{f_0}(f_1, \ldots, f_m)$ on $D(f_0)$, and $\psi = \frac{1}{g_0}(g_1, \ldots, g_n)$ on $D(g_0)$. Set $h_i := g_i(\frac{f_1}{f_0}, \ldots, \frac{f_m}{f_0})$. Then $h_i \in K[\bar{U}][\frac{1}{f_0}]$, h_0 is nowhere zero on $D(f_0)$, and $\psi\phi = \frac{1}{h_0}(h_1, \ldots, h_n)$ on $D(f_0) \subset D(h_0)$. Thus the composition $\psi\phi$ is locally regular, and hence regular. □

Corollary 15.5. *Let U_i be quasi-affine varieties. Then each regular map $\phi\colon U_1 \to U_2$ yields an algebra homomorphism $\phi^*\colon \mathcal{O}(U_2) \to \mathcal{O}(U_1)$. If $\psi\colon U_2 \to U_3$ is regular, then $(\psi\phi)^* = \phi^*\psi^*$.*

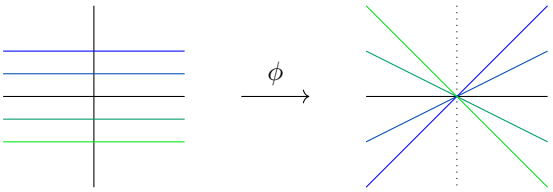
Proof. Let $\theta \in \mathcal{O}(U_2)$, so a regular function $U_2 \to K$. The composition $\theta\phi\colon U_1 \to K$ is regular, so an element of $\mathcal{O}(U_1)$, and the map $\theta \mapsto \theta\phi$ is an algebra homomorphism $\phi^*\colon \mathcal{O}(U_2) \to \mathcal{O}(U_1)$. If $\psi\colon U_2 \to U_3$ is regular, then clearly $(\psi\phi)^* = \phi^*\psi^*$, sending $\theta \in \mathcal{O}(U_3)$ to $\phi\psi\theta \in \mathcal{O}(U_1)$. □

Remark. It follows that the quasi-affine varieties, together with the regular maps between them, form a category, containing the affine varieties as a full subcategory; that is, there are no 'new' regular maps between affine varieties, Theorem 15.2. Moreover, we have an isomorphism in this larger category between the distinguished open $D(g)$ and the affine variety $\mathbb{V}(gT-1)$.

It is no longer true, however, that quasi-affine varieties are determined by their regular functions, or that algebra homomorphisms between their rings of regular functions determine regular maps. For example, let $U = \mathbb{A}^2 - \{(0,0)\}$ be the punctured plane, so an open set, but not distinguished open. An easy computation shows that the inclusion $U \rightarrowtail \mathbb{A}^2$ yields an isomorphism $K[X,Y] \xrightarrow{\sim} \mathcal{O}(U)$. The inverse algebra homomorphism, however, does not yield a regular map $\mathbb{A}^2 \to U$. For, if that were the case, then the composition $\mathbb{A}^2 \to U \rightarrowtail \mathbb{A}^2$ would correspond to the identity on the co-ordinate algebras, and hence would be the identity map, a contradiction since it is not surjective. This example also shows that U is not isomorphic to an affine variety.

15.2 Example

Consider the map $\phi\colon \mathbb{A}^2 \to \mathbb{A}^2$, $(a,b) \mapsto (a,ab)$



which collapses the Y-axis, i.e. the line $\mathbb{V}(X)$, to the point $(0,0)$.

The restriction $D(X) \to \mathbb{A}^2$ induces an isomorphism $D(X) \cong D(X)$. On co-ordinate algebras this is the isomorphism $K[X^{\pm 1}, Y] \xrightarrow{\sim} K[X^{\pm 1}, Y]$, $(X,Y) \mapsto (X, Y/X)$. We thus have a regular map restricting to an isomorphism on dense open subsets, but which is not even a bijection.

In general we say that two irreducible affine varieties are **birationally equivalent** provided they have isomorphic non-empty open subsets. This is a much coarser notion of equivalence than being isomorphic. Hironaka proved in characteristic zero, every variety is birational to a smooth projective variety, and the aim of the 'minimal model program' is to describe simple representatives for the birational equivalence classes.

This example is also interesting for another reason. We know that localisation is flat, so the inclusion $\psi\colon D(X) \rightarrowtail \mathbb{A}^2$ is flat. Also, the composition $\phi\psi$ decomposes as $D(X) \xrightarrow{\sim} D(X) \rightarrowtail \mathbb{A}^2$, so this is also flat. However, the morphism ϕ itself is not flat.

In other words, we have algebra homomorphisms ϕ^* and ψ^* such that both ψ^* and $\psi^*\phi^*$ are flat, but ϕ^* is not flat.

Geometrically, flat morphisms $\phi\colon X \to Y$ are open maps such that the fibre dimensions $\dim \phi^{-1}(y)$ are locally constant on Y, and hence correspond to the intuitive idea of a 'continuous family of varieties'. For the map ϕ above we see that $\phi^{-1}(x,y)$ is the point $(x, y/x)$ if $x \neq 0$, but is the whole Y-axis if $(x,y) = (0,0)$. On the other hand, ϕ is 'generically flat', since it is flat (in fact an isomorphism) on the open set $D(X)$.

Finally, this example also shows that the image of a regular map is not necessarily a variety, or even locally closed. For, the image of the map ϕ above is

$$\{(x,y) : x \neq 0\} \cup \{(0,0)\},$$

so the union of the distinguished open set $D(X)$ with the closed point $\{(0,0)\}$.

In a sense, this is about as bad as it can get. More precisely, Chevalley's Theorem states that the image of a regular map is always constructible, so a finite union of locally closed sets. (It follows that the image of every constructible set is constructible.)

We have seen that a quasi-affine variety is no longer determined by its ring of regular functions. It is therefore an interesting question as to how we can distinguish affine varieties from quasi-affine varieties. One can do this by considering **cohomology of quasi-coherent sheaves**. (Quasi-coherent sheaves are geometric analogues of finitely presented modules.)

Theorem 15.6 (Serre)**.** *Let X be a quasi-affine variety (or more generally a Noetherian scheme). Then X is affine if and only if $H^i(X, \mathcal{F}) = 0$ for all $i > 0$ and all quasi-coherent sheaves $\mathcal{F}$ on X.*

Part VI

Projective Geometry

We have seen that for affine varieties, the projection map $X \times Y \to Y$ is open, but not closed. For example, if $f \in K[V]$, then the closed subset $\mathbb{V}(fT-1) \subset V \times \mathbb{A}^1$ is sent homeomorphically to the distinguished open $D(f) \subset V$.

We say that a topological space X is **complete** if the projection map $X \times Y \to Y$ is closed for all Y. One of the main theorems about projective varieties is that they are all complete. As a consequence, many results about affine varieties can be quite complicated, but their analogues for projective varieties become much nicer.

Note that the definition of completeness depends on the topology of the product space $X \times Y$. If all products have the product topology, then being complete is equivalent to being quasi-compact. For varieties, however, the topology on the product is in general much finer than the product topology. Thus, even though varieties are always quasi-compact, the quasi-affine varieties are in general not complete.

16 Projective varieties

We begin with the construction of projective space $\mathbb{P}^n$. As a set, the points of $\mathbb{P}^n$ correspond to the lines in $\mathbb{A}^{n+1}$ through the origin. Observe that a line through the origin is determined by any non-zero point on the line, and that any two non-zero points lie on the same line if and only if they differ by a non-zero scalar multiple. Thus $\mathbb{P}^n := (\mathbb{A}^{n+1} - \{0\})/K^\times$ is a quotient set, and we endow it with the quotient topology. We write $[\underline{a}]$ for the equivalence class containing $\underline{a} \neq 0$.

In other words, writing $\pi\colon \mathbb{A}^{n+1} - \{0\} \to \mathbb{P}^n$ for the quotient map, then a subset $U \subset \mathbb{P}^n$ is open (respectively closed) provided that $\pi^{-1}(U)$ is open (respectively closed) in $\mathbb{A}^{n+1} - \{0\}$. Observe that U is open if and only if $\pi^{-1}(U)$ is open as a subset of $\mathbb{A}^{n+1}$, whereas U is closed if and only if $\widetilde{U} = \pi^{-1}(U) \cup \{0\} \subset \mathbb{A}^{n+1}$ is closed.

In general, a **cone** is a subset $W \subset \mathbb{A}^{n+1}$ closed under scalar multiplication, so $w \in W$ and $\lambda \in K$ implies $\lambda w \in W$. If $U \subset \mathbb{P}^n$, then the set $\widetilde{U} = \pi^{-1}(U) \cup \{0\}$ is always a cone.

16.1 Homogeneous ideals and cones

A **projective variety** is a closed subset of some projective space, so a subset of the form

$$\mathbb{V}_+(I) := \big(\mathbb{V}(I) - \{0\}\big)/K^\times \quad \text{where } \mathbb{V}(I) \text{ is a cone in } \mathbb{A}^{n+1}.$$

More generally, a **quasi-projective variety** is a locally-closed subset of some projective space.

We therefore need to describe which ideals $I \lhd R_{n+1}$ give rise to cones in $\mathbb{A}^{n+1}$.

Let $f \in R_{n+1} = K[X_0, \ldots, X_n]$. As usual we can write $f = f_0 + f_1 + \cdots + f_d$ with f_i homogeneous of degree i, called the homogeneous parts of f. We say that an ideal $I \lhd R_{n+1}$ is **homogeneous** provided $f \in I$ implies $f_i \in I$ for all i.

Lemma 16.1. *An ideal $I \lhd R_{n+1}$ is homogeneous if and only if it is generated by homogeneous polynomials. Sums, intersections, finite products, and radicals of homogeneous ideals are again homogeneous.*

Proof. Let $I \lhd R_{n+1}$ be a homogeneous ideal, and take any set of generators f_i for I. If we decompose $f_i = \sum_j f_{ij}$ into its homogeneous parts, then $f_{ij} \in I$ for all i, j since I is homogeneous. Thus I is generated by the homogeneous polynomials f_{ij}.

Conversely, suppose I is generate by homogeneous polynomials f_i, and let $g \in I$. Write $g = \sum_i a_i f_i$ as a finite sum and decompose each a_i into its homogeneous parts a_{ij}. Then $g = \sum_{i,j} a_{ij} f_i$, and each $a_{ij} f_i$ is homogeneous and lies in I. Thus the homogeneous parts of g all lie in I, so I is a homogeneous ideal.

Clearly intersections of homogeneous ideals are again homogeneous. If each I_i is homogeneous, generated by a set of homogeneous polynomials X_i, then the sum $\sum_i I_i$ is generated by $\bigcup_i X_i$ and the product $I_i I_j$ is generated by $\{fg : f \in X_i, g \in X_j\}$. These are again sets of homogeneous polynomials, so both $\sum_i I_i$ and $I_i I_j$ are homogeneous.

Finally, let I be homogeneous and take $f \in \operatorname{rad}(I)$. Decompose $f = f_0 + \cdots + f_d$ into its homogeneous parts. Then $f^N \in I$ for some N, and this has f_d^N as its homogeneous part of maximal degree. Since I is homogeneous we know $f_d^N \in I$, whence $f_d \in \operatorname{rad}(I)$. Since $f - f_d \in \operatorname{rad}(I)$, the result now follows by induction on degree. □

For an ideal $I \lhd R_{n+1}$ we define $I_h \subset I$ to be the ideal generated by all homogeneous polynomials in I.

Lemma 16.2. *Let $\mathfrak{p} \lhd R_{n+1}$ be a prime ideal. Then $\mathfrak{p}_h$ is also prime. More generally, if I is radical, then I_h is radical. In particular, if I is homogeneous and radical, then each minimal prime $\mathfrak{p} \supset I$ is homogeneous.*

Proof. Let f, g be homogeneous with $fg \in \mathfrak{p}_h$ but $g \notin \mathfrak{p}_h$. Then also $fg \in \mathfrak{p}$ but $g \notin \mathfrak{p}$, so $\mathfrak{p}$ prime implies $f \in \mathfrak{p}$, and hence $f \in \mathfrak{p}_h$. In other words, $\mathfrak{p}_h$ is prime with respect to homogeneous elements.

In general, if $fg \in \mathfrak{p}_h$ but $g \notin \mathfrak{p}_h$, then we decompose $f = f_0 + \cdots + f_d$ and $g = g_0 + \cdots + g_e$ in terms of their homogeneous parts. By removing those parts which are in $\mathfrak{p}_h$ we may assume that $g_e \notin \mathfrak{p}_h$, so also $g_e \notin \mathfrak{p}$. Now, since $\mathfrak{p}_h$ is homogeneous, it contains all the homogeneous parts of fg, and in particular it contains the highest degree homogeneous part, namely $f_d g_e$. Thus as above $f_d \in \mathfrak{p}_h$. Replacing f by $f - f_d$, we then see by induction that $f - f_d \in \mathfrak{p}_h$, and hence also $f \in \mathfrak{p}_h$. Thus $\mathfrak{p}_h$ is prime.

Let $I = \bigcap_i \mathfrak{p}_i$ be radical. We know that each $\mathfrak{p}_{i,h}$ is prime and homogeneous, so $\bigcap_i \mathfrak{p}_{i,h} \subset I$ is homogeneous, and hence is contained in I_h. Conversely, $I_h \subset \mathfrak{p}_i$ is homogeneous, so is contained in $\mathfrak{p}_{i,h}$, and hence $I_h \subset \bigcap_i \mathfrak{p}_{i,h}$.

Finally, let I be homogeneous and radical, say with minimal decomposition $I = \mathfrak{p}_1 \cap \cdots \cap \mathfrak{p}_r$. Then $I = I_h = \mathfrak{p}_{1,h} \cap \cdots \cap \mathfrak{p}_{r,h}$. Each $\mathfrak{p}_{i,h}$ is prime, so by Proposition 2.3 we must have $\mathfrak{p}_j \subset \mathfrak{p}_{i,h} \subset \mathfrak{p}_i$ for some j. Minimality then gives $i = j$, so $\mathfrak{p}_i = \mathfrak{p}_{i,h}$ is homogeneous. □

Geometrically we have the following interpretation.

Proposition 16.3. *If $I \lhd R_{n+1}$ is homogeneous, then $\mathbb{V}(I)$ is a cone. Conversely, if $V \subset \mathbb{A}^{n+1}$ is a cone, then $\mathbb{I}(V)$ is homogeneous, and each irreducible component of V is again a cone.*

Proof. Let I be generate by homogeneous polynomials f_i, say of degree d_i. If $\underline{a} \in \mathbb{V}(I)$ and $\lambda \in K$, then $f_i(\lambda \underline{a}) = \lambda^{d_i} f_i(\underline{a}) = 0$, whence $\lambda \underline{a} \in \mathbb{V}(I)$. Thus $\mathbb{V}(I)$ is a cone.

Conversely, let V be a cone. Let $f \in \mathbb{I}(V)$ and write $f = f_0 + \cdots + f_d$ as a sum of its homogeneous parts. Take $\underline{a} \in V$. Then $\lambda \underline{a} \in V$ for all $\lambda \in K$, so

$$0 = f(\lambda \underline{a}) = f_0 + \lambda f_1(\underline{a}) + \cdots + \lambda^d f_d(\underline{a}).$$

In particular, given distinct $\lambda_0, \lambda_1, \ldots, \lambda_d \in K$, set

$$M := \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \lambda_0 & \lambda_1 & \cdots & \lambda_d \\ \vdots & \vdots & & \vdots \\ \lambda_0^d & \lambda_1^d & \cdots & \lambda_d^d \end{pmatrix}.$$

Then $\big(f_0(\underline{a})\ f_1(\underline{a})\ \cdots\ f_d(\underline{a})\big)M = 0$. Since M is a van der Monde matrix, its determinant is non-zero:

$$\det(M) = \prod_{i>j} (\lambda_i - \lambda_j).$$

Thus $f_i(\underline{a}) = 0$ for all i. This holds for all $\underline{a} \in V$, so $f_i \in \mathbb{I}(V)$ and $\mathbb{I}(V)$ is homogeneous.

Finally, the irreducible components of a cone V are the $\mathbb{V}(\mathfrak{p})$ for the minimal primes $\mathfrak{p} \supset \mathbb{I}(V)$. Each of these primes is homogeneous, so each irreducible component is a cone. □

We next interpret these results in terms of projective varieties.

Theorem 16.4 (Projective Nullstellensatz)**.** *The map $I \mapsto \mathbb{V}_+(I)$ gives an inclusion-reversing bijection between the proper homogeneous radical ideals $I \lhd R_{n+1}$ and the closed subsets of $\mathbb{P}^n$. The empty set corresponds to the* ***irrelevant ideal*** *$(X_0, \ldots, X_n)$. The point $[\underline{a}]$ corresponds to the ideal $(\{a_i X_j - a_j X_i\})$.*

Proof. The map $U \mapsto \tilde{U} := \pi^{-1}(U) \cup \{0\}$ is an inclusion-preserving bijection between the closed subsets of $\mathbb{P}^n$ and the closed, non-empty cones in $\mathbb{A}^{n+1}$. By the previous proposition and the Nullstellensatz, there is an inclusion-reversing bijection between the closed, non-empty cones in $\mathbb{A}^{n+1}$ and the proper homogeneous radical ideals $I \lhd R_{n+1}$.

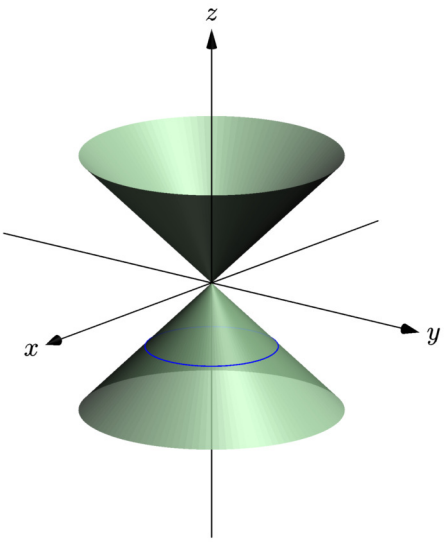
Note that the point $[\underline{a}]$ corresponds to the line through $\underline{a}$ and the origin, and hence to the homogeneous prime ideal $(\{a_i X_j - a_j X_i\})$. □

Corollary 16.5. *Let I be a homogeneous radical ideal. Then the irreducible components of $\mathbb{V}_+(I)$ are precisely the $\mathbb{V}_+(\mathfrak{p}_i)$ where $I = \mathfrak{p}_1 \cap \cdots \cap \mathfrak{p}_r$ is the minimal decomposition of I into primes.*

Proof. This is now immediate, since the primes in the minimal decomposition are all homogeneous. □

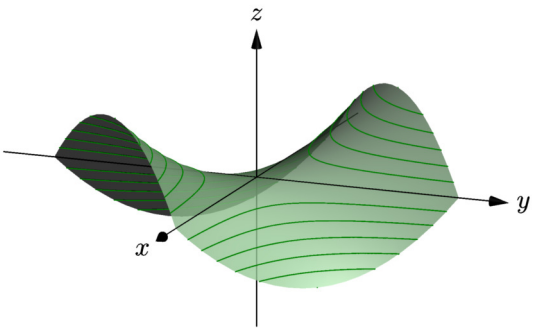
16.2 Examples

(1) Consider the unit circle in the $Z = -1$ plane, so $\mathbb{V}(X^2 + Y^2 - 1, Z + 1)$. Then the smallest cone containing this is $\mathbb{V}(X^2 + Y^2 - Z^2)$.



(2) One can also consider the largest cone contained in a variety. The largest cone contained in $\mathbb{V}(I)$ is $\mathbb{V}(I')$, where I' is the ideal generated by all homogeneous parts of all polynomials in I. Even if V is irreducible, the largest cone contained in V will in general not be irreducible.

Consider the saddle $\mathbb{V}(XY - Z)$. The largest cone this contains is $\mathbb{V}(XY, Z) = \mathbb{V}(X, Z) \cup \mathbb{V}(Y, Z)$, so the union of the X- and Y-axes.



16.3 Open affine covering of projective space

We have so far described the relationship between projective varieties and cones in affine space. An alternative approach which is very useful in practice is to view projective varieties as being made by gluing together affine varieties. This also allows us to construct a projective closure of an arbitrary affine variety. (Note however that this is not 'intrinsic' and will depend on the embedding into affine space.)

Theorem 16.6. *For $0 \leq i \leq n$ set $U_i := \{[a_0, \ldots, a_n] : a_i \neq 0\} \subset \mathbb{P}^n$. Then*

(1) the U_i form an open cover of $\mathbb{P}^n$, and each U_i is dense in $\mathbb{P}^n$.

(2) there is a homeomorphism

$$\phi_i \colon \mathbb{A}^n \to U_i, \quad (b_1, \ldots, b_n) \mapsto [b_1, \ldots, b_i, 1, b_{i+1}, \ldots, b_n].$$

*(3) the **transition functions** are isomorphisms of open subsets of $\mathbb{A}^n$*

$$\phi_j^{-1}\phi_i \colon \phi_i^{-1}(U_i \cap U_j) \to \phi_j^{-1}(U_i \cap U_j).$$

Proof. Since every point $[\underline{a}] \in \mathbb{P}^n$ satisfies $a_i \neq 0$ for some i, we see that the U_i cover $\mathbb{P}^n$. Also, the preimage of U_i in $\mathbb{A}^{n+1}$ is the distinguished open $D(X_i)$. Since this is open and dense in $\mathbb{A}^{n+1}$, it follows that each U_i is open and dense in $\mathbb{P}^n$.

Inside $\mathbb{A}^{n+1}$ we have the hyperplane $H_i := \mathbb{V}(X_i - 1)$, and an isomorphism

$$\iota_i \colon \mathbb{A}^n \to H_i, \quad (b_1, \ldots, b_n) \mapsto (b_1, \ldots, b_i, 1, b_{i+1}, \ldots, b_n).$$

Thus $\phi_i = \pi\iota_i$ will be a homeomorphism provided the restriction $\pi \colon H_i \to U_i$ is a homeomorphism. Now, the restriction is clearly bijective and continuous. For the converse we first observe that there is a regular map $\sigma_i \colon D(X_i) \to H_i$ given by $(a_0, \ldots, a_n) \mapsto (a_0/a_i, \ldots, a_n/a_i)$. In particular, σ_i is continuous. So, if $U \subset H_i$ is open, then $\pi^{-1}(\pi(U)) = \sigma_i^{-1}(U)$ is also open, and hence $\pi(U)$ is open as required.

Finally, since the ι_i are isomorphisms of affine varieties, we see that the transition functions will be regular isomorphisms if and only if the corresponding map $H_i \cap D(X_j) \to H_j \cap D(X_i)$ is a regular isomorphism. This is now clear, since it sends $\underline{a}$ with $a_i = 1$ and $a_j \neq 0$ to the point $(a_0/a_j, \ldots, a_n/a_j)$. □

We can therefore view $\mathbb{P}^n$ as being $n+1$ copies of affine space $\mathbb{A}^n$, glued together via the transition functions. We call the decomposition $\mathbb{P}^n = U_0 \cup \cdots \cup U_n$ the **standard open affine cover**. Given a subvariety $W \subset \mathbb{P}^n$, the sets $W_i := W \cap U_i$ form an open cover of W. On the other hand, each W_i is closed in U_i, so $W_i \cong \phi_i^{-1}(W)$ is an affine subvariety of $\mathbb{A}^n$.

We can also go in the opposite direction. Starting from an affine variety $V \subset \mathbb{A}^n$, we can identifying this with a subset of U_0 via ϕ_0, and then take its Zariski closure inside $\mathbb{P}^n$. The resulting projective variety is often called **the projective closure** of V, but note that this depends on the embedding $V \rightarrowtail \mathbb{A}^n$, so isomorphic affine varieties can have non-isomorphic projective closures.

We describe this process in terms of cones in $\mathbb{A}^{n+1}$. The regular map $\iota_0 \colon \mathbb{A}^n \xrightarrow{\sim} H_0 \rightarrowtail \mathbb{A}^{n+1}$ corresponds to the surjective algebra homomorphism $\iota_0^* \colon R_{n+1} \twoheadrightarrow R_n$, $X_0 \mapsto 1$, $X_i \mapsto X_i$ for $i \geq 1$. If $I \lhd R_n$, then $(\iota_0^*)^{-1}(I) = (I, X_0 - 1) \lhd R_{n+1}$ is the ideal generated by I and $X_0 - 1$. Thus if $V = \mathbb{V}(I) \subset \mathbb{A}^n$, then the smallest cone containing $\iota_0(V)$ is $\hat{V} = \mathbb{V}(\hat{I})$, where $\hat{I} = (I, X_0 - 1)_h$ is the ideal generated by all homogeneous elements in $(I, X_0 - 1)$.

Lemma 16.7. *Let $\mathfrak{p} \lhd R_n$ be prime. Then $\hat{\mathfrak{p}} \lhd R_{n+1}$ is prime.*

Let $I \lhd R_n$ be radical, with minimal decomposition into primes $I = \mathfrak{p}_1 \cap \cdots \cap \mathfrak{p}_r$. Then $\hat{I}$ is radical and has minimal decomposition $\hat{I} = \hat{\mathfrak{p}}_1 \cap \cdots \cap \hat{\mathfrak{p}}_r$.

Proof. We have the surjective algebra homomorphism $\iota_0^*\colon R_{n+1} \twoheadrightarrow R_n$, and $(\iota_0^*)^{-1}(I) = (I, X_0 - 1)$. Thus if $\mathfrak{p} \supset I$ is a minimal prime, then $(\mathfrak{p}, X_0 - 1) \supset (I, X_0 - 1)$ is a minimal prime. Now $\hat{I} = (I, X_0 - 1)_h$, so the result follows from Lemma 16.2. □

Suppose $f \in I$, and decompose it homogeneous parts $f = f_0 + \cdots + f_d$. Then

$$\hat{f} := f_0 X_0^d + \cdots + f_{d-1} X_0 + f_d = f_0(X_0^d - 1) + \cdots + f_{d-1}(X_0 - 1) + f$$

is homogeneous, and $\iota_0^*(\hat{f}) = f$. Alternatively we can write $\hat{f} = f(\frac{X_1}{X_0}, \ldots, \frac{X_d}{X_0}) X_0^d$.

Lemma 16.8. *The ideal $\hat{I}$ is generated by all $\hat{f}$ for $f \in I$.*

Proof. Let $f \in I$. Then $\hat{f}$ is homogeneous $\iota_0^*(\hat{f}) = f$, so $\hat{f} \in \hat{I}$. Conversely, let $g \in \hat{I}$ be homogeneous. Then $f := \iota_0^*(g) \in I$, and writing $g = g_0 X_0^d + \cdots + g_e X_0^{d-e} \in R_n[X_0]$ we see that $f = g_0 + \cdots + g_e$ and $g = \hat{f} X_0^{d-e}$. □

Note. It is not true that if I is generated by polynomials f_i, then $\hat{I}$ is generated by $\hat{f}_i$.

Proposition 16.9. *Let $I \lhd R_n$ be an ideal, $\hat{I} \lhd R_{n+1} = R_n[X_0]$ the corresponding homogeneous ideal, and set $R := R_n/I$ and $\hat{R} := R_n[X_0]/\hat{I}$. Then, after a homogeneous change of variables in R_n, we may assume that $K[X_1, \ldots, X_d] \rightarrowtail R$ and $K[X_0, X_1 \ldots, X_d] \rightarrowtail \hat{R}$ are both integral.*

In other words, given $V \subset \mathbb{A}^n$, then there is a linear map $p\colon \mathbb{A}^n \to \mathbb{A}^d$, lifting to a linear map $\hat{p} = (\mathrm{id}, p)\colon \mathbb{A}^{n+1} \to \mathbb{A}^{d+1}$, whose respective restrictions to V and $\hat{V}$ are both surjective with finite fibres.

Proof. The idea is to show that the proof of the Noether Normalisation Lemma goes through in parallel for R and $\hat{R}$. The main claim is therefore that if $J := I \cap R_{n-1}$, then $\hat{J} = \hat{I} \cap R_{n-1}[X_0]$. To see this, take $p \in J$. Then clearly $\hat{p} \in \hat{I} \cap R_{n-1}[X_0]$. Conversely, take $q \in \hat{I} \cap R_{n-1}[X_0]$ and set $p := \iota_0^*(q)$. Then $p \in J$ and $q = \hat{p} X_0^d \in \hat{J}$.

So, as in the proof of the Noether Normalisation Lemma, we take some non-zero $f \in I$ of degree r, say $f = f_0 + \cdots + f_r$ as a sum of its homogeneous parts. After a homogeneous linear transformation we may assume f_r, and hence also f, is monic of degree r in X_n. Setting $J := I \cap R_{n-1}$ and $S := R_{n-1}/J$, we see that $S \rightarrowtail R$ is finite.

Now $\hat{f} \in \hat{I}$ is also monic of degree r in X_n, so if $\hat{S} := R_{n-1}[X_0]/\hat{J}$, then $\hat{S} \rightarrowtail \hat{R}$ is finite. The proof now follows by induction. □

Theorem 16.10. *Let $W \subset \mathbb{P}^n$ be a projective variety, and $\pi^{-1}(W) \subset \mathbb{A}^{n+1}$ the corresponding cone. Then* $\dim W = \dim \pi^{-1}(W) - 1$.

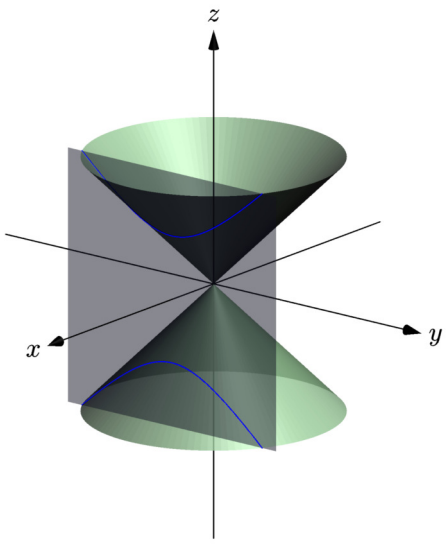
In particular, if $V \subset \mathbb{A}^n$ is an affine variety and $W \subset \mathbb{P}^n$ is its projective closure of V, then $\dim V = \dim W$.

Proof. Given W, set $V_i := \phi_i^{-1}(W) \cong W \cap U_i$, let W_i be the closure of $W \cap U_i$, and $\hat{V}_i$ the corresponding cone. Then W is the union of the W_i, and $\pi^{-1}(W)$ is the union of the cones $\hat{V}_i$. Thus $\dim W = \max\{\dim W_i\}$ and $\dim \pi^{-1}(W) = \max\{\dim \hat{V}_i\}$.

Now $V_i \cong W \cap U_i$ is an open subset of W_i, so $\dim V_i \leq \dim W_i$. The projective Nullstellensatz tells us that $\dim W_i$ is the maximum length of a chain of non-zero closed and irreducible cones in $\hat{V}_i$, so $\dim W_i \leq \dim \hat{V}_i - 1$. Finally, the previous proposition gives $\dim V_i = \dim \hat{V}_i - 1$. □

16.4 Examples

(1) Consider the projective variety $W = \mathbb{V}_+(X^2 + Y^2 - Z^2) \subset \mathbb{P}^2$. The corresponding cone in $\mathbb{A}^3$ is



The affine pieces of W are the $V_i = W \cap U_i$ for $i = 0, 1, 2$. These are

$$V_0 = \mathbb{V}(1 + Y^2 - Z^2), \quad V_1 = \mathbb{V}(X^2 + 1 - Z^2), \quad V_2 = \mathbb{V}(X^2 + Y^2 - 1).$$

Thus V_0 and V_1 are hyperbolae, whereas V_2 is a circle.

(2) After a homogeneous change of variables we obtain $W = \mathbb{V}_+(X^2 - YZ)$. In this case

$$V_0 = \mathbb{V}(1 - YZ), \quad V_1 = \mathbb{V}(X^2 - Y), \quad V_2 = \mathbb{V}(X^2 - Z),$$

so V_0 is a hyperbola, whereas V_1 and V_2 are parabolae.

(3) We know that if $I \lhd R_n$, then $\hat{I} \lhd R_{n+1}$ is generated by all $\hat{f}$ for $f \in I$. In particular, if $V = \mathbb{V}(f)$ is a hypersurface, then $\bar{V} = \mathbb{V}_+(\hat{f})$ is a (projective) hypersurface.

We mentioned earlier that if I is generated by f_i, then $\hat{I}$ is in general not generated by $\hat{f}_i$. As a counter-example, consider the twisted cubic

$$V := \{(t, t^2, t^3) : t \in K\} = \mathbb{V}(I) \quad \text{where } I = (X^2 - Y, X^3 - Z) \text{ is prime.}$$

Note that $I = (X^2 - Y, XY - Z, XZ - Y^2)$. Thus $\hat{I}$ contains

$$J := (X^2 - WY, XY - WZ, Y^2 - XZ),$$

but that $(X^2 - WY, X^3 - W^2Z)$ is strictly contained in J (consider homogeneous polynomials of degree 2).

In fact, we can prove $\hat{I} = J$. Set

$$W := \{[s^3, s^2t, st^2, t^3] : [s, t] \in \mathbb{P}^1\}.$$

We observe that $\iota(t,t^2,t^3) = [1,t,t^2,t^3]$, so $W = \iota(V) \cup \{[0,0,0,1]\}$. Also, $W = \mathbb{V}_+(J)$. For, let $[w,x,y,z] \in \mathbb{V}_+(J)$. If $w \neq 0$, then we may assume $w = 1$, and hence $y = x^2$ and $z = xy = x^3$, so we have the point $[1,x,x^2,x^3] \in \iota(V)$. If $w = 0$, then $x^2 = 0$ and $y^2 = xz = 0$, so we must have the point $[0,0,0,1]$.

Thus W is closed and contains the projective closure $\bar{V}$. To get equality we just need to show that $[0,0,0,1]$ lies in the closure of $\iota(V)$. Consider

$$W \cap U_3 = \mathbb{V}(X^2 - WY, XY - W, Y^2 - X) = \mathbb{V}(Y^2 - X, Y^3 - W).$$

This is again a twisted cubic, so irreducible. Thus every open neighbourhood of $[0,0,0,1]$ is dense, so contains some point $[s^3,s^2,s,1]$ with $s \neq 0$, and every such point lies in $W \cap U_0$.

Now I is prime, and is the kernel of the epimorphism

$$R[X,Y,Z] \twoheadrightarrow K[X], \quad (X,Y,Z) \mapsto (X,X^2,X^3).$$

Similarly J is prime, and is the kernel of the algebra homomorphism

$$K[W,X,Y,Z] \to K[S,T], \quad (W,X,Y,Z) \mapsto (S^3,S^2T,ST^2,T^3).$$

For, using $X^3 - W^2Z, Y^3 - WZ^2 \in J$, we see that $\hat{R} := K[W,X,Y,Z]/J$ is finite over $K[W,Z]$. Moreover, using $X^2 - WY, Y^2 - XZ \in J$, we see that, as a $K[W,Z]$-module, $\hat{R}$ is spanned by $1,X,Y$. These are sent to $1,S^2T,ST^2$, which for degree reasons are linearly independent over $K[S^3,T^3]$. Thus $\hat{R}$ is free as a $K[W,Z]$-module, with basis $1,X,Y$, and $\hat{R} \to K[W,Z]$ is injective.

Finally, $\hat{I}$ and J are both prime, and determine the same cone, or projective variety, so $J = \hat{I}$.

(4) In the proof of Proposition 16.9 we emphasised the point that when applying the Noether Normalisation Lemma to V and $\hat{V}$, we started with a polynomial $f \in I$ of degree r which was monic of the same degree r in one of the variables, X_n say, in order to ensure that $\hat{f}$ was again monic of degree r in X_n.

We provide an example of what can go wrong by again looking at the twisted cubic

$$V = \mathbb{V}(X^2 - Y, X^3 - Z) \quad \text{and} \quad \hat{V} = \mathbb{V}(X^2 - WY, Y^2 - XZ, XY - WZ).$$

Since $K[V] \cong K[X]$, the projection $p\colon V \to \mathbb{A}^1$, $(x,y,z) \mapsto x$ is an isomorphism, but the polynomials $Y - X^2$ and $Z - X^3$, although monic in Y and Z, do not satisfy the criterion on degrees. This is reflected in the fact that when we take the corresponding cone $\hat{V}$, the projection $\hat{p}\colon \hat{V} \to \mathbb{A}^2$, $(w,x,y,z) \mapsto (w,x)$, is not surjective. In fact, if $(w,x,y,z) \in \hat{V}$, then $w = 0$ implies $x = 0$, so the point $(0,1)$ does not lie in the image.

Instead, we should consider $X^3 - Z, Y^3 - Z^2$ as a monic polynomials over $K[Z]$, so $K[V]$ is finite over $K[Z]$. These do satisfy the degree criterion, and the projections $p\colon V \to \mathbb{A}^1$, $(x,y,z) \mapsto z$ and $\hat{p}\colon \hat{V} \to \mathbb{A}^2$, $(w,x,y,z) \mapsto (w,z)$, are both surjective with finite fibres.

17 Regular maps

Let $F_0, \dots, F_n \in R_{m+1}$ be homogeneous of the same degree. Set $U := \bigcup_i D(F_i)$, an open cone in $\mathbb{A}^{m+1}$. Then $f := [F_0, \dots, F_n]$ determines a map $U \to \mathbb{P}^n$ which is constant on $K^\times$-orbits, and so descends to a map $\pi(U) \to \mathbb{P}^n$.

If $V \subset \mathbb{P}^m$ is locally-closed, then a map $f\colon V \to \mathbb{P}^n$ will be called **regular** if it is given locally by homogeneous polynomials of the same degree.

Unlike for affine varieties, it is usually not possible to represent a regular map 'globally' by the same homogeneous polynomials. For example, we have the regular map

$$f\colon \mathbb{V}_+(X^2 - YZ) \to \mathbb{P}^1, \quad f = \begin{cases} [X, Y] & \text{on } D_+(Y) := \pi(D(Y)) \\ [Z, X] & \text{on } D_+(Z) \end{cases}$$

Note that $[x, y] = [z, x]$ for all $[x, y, z] \in \mathbb{V}_+(X^2 - YZ) \cap D_+(YZ)$, so f is well-defined.

We first show that this notion of regularity agrees with our previous version for quasi-affine varieties.

Lemma 17.1. *Let $V \subset \mathbb{A}^m$ be locally closed, and $f\colon V \to \mathbb{A}^n$ and $g\colon \iota_0(V) \to \mathbb{P}^n$ two maps such that $\iota_0 g = \iota_0 f$. Then f is regular (as a map between quasi-affine varieties) if and only if g is regular (as a map between quasi-projective varieties).*

Proof. Suppose first that f is regular. Locally, on some open $V' \subset V$, we can write $f = \frac{1}{f_0}(f_1, \dots, f_n)$ for some $f_i \in R_m$, where $V' \subset D(f_0)$. Take $d \geq \max\{\deg f_i\}$ and set $F_i := \hat{f}_i X_0^{d - \deg f_i}$. Then the $F_i \in R_{m+1}$ are homogeneous of the same degree, and $g = [F_0, \dots, F_n]$ on $\iota_0(V')$. For, $F_i(1, \underline{a}) = f_i(\underline{a})$ and $f_0(1, \underline{a}) \neq 0$ on V'.

Suppose instead that $g = [F_0, \dots, F_n]$ on $\iota_0(V')$ for some open $V' \subset V$. Thus the $F_i \in R_{m+1}$ are homogeneous of the same degree. The image of g is contained in the standard open affine U_0 of $\mathbb{P}^n$, so F_0 is never zero on $\iota_0(V')$. Setting $f_i := F_i(1, X_1, \dots, X_n)$, we see that f_0 is nowhere zero on V', and $f = \frac{1}{f_0}(f_1, \dots, f_n)$ on V'. Thus f is regular on V. □

Lemma 17.2. *Each regular map is continuous, and the composition of two regular maps is again regular.*

Proof. Consider a regular map $f = [F_0, \dots, F_n]\colon U \to \mathbb{P}^n$, where $F_i \in R_{m+1}$ are homogeneous of the same degree and $U \subset \bigcup_i D_+(F_i) \subset \mathbb{P}^m$. If $G \in R_{n+1}$ is homogeneous, then $f^{-1}(D_+(G))$ is $U \cap D_+(G(F_0, \dots, F_n))$ is open. Since the $D_+(G)$ form a basis for the Zariski topology on $\mathbb{P}^n$, we see that f is continuous.

In general, a regular map $f\colon U \to \mathbb{P}^n$ is locally of the above form, and hence f is continuous.

Now suppose $g\colon V \to \mathbb{P}^r$ is regular, where $\mathrm{Im}(f) \subset V \subset \mathbb{P}^n$. Locally we have an open $V' \subset V$ such that $g = [G_0, \dots, G_r]$ on V'. Then $f^{-1}(V')$ is open, so have an open $U' \subset f^{-1}(V')$ such that $f = [F_0, \dots, F_n]$ on U'. Now $gf = [H_0, \dots, H_r]$ on U', where $H_i = G_i(F_0, \dots, F_n)$, so gf is regular. □

We therefore have a category of quasi-projective varieties and the regular maps between them. Moreover, the category of quasi-affine varieties forms a full subcategory by Lemma 17.1.

17.1 Regular maps from projective space

The next proposition shows that in certain cases, every regular map can be given 'globally' by the same homogeneous polynomials.

Proposition 17.3. *Let $I \lhd R_{m+1}$ be a homogeneous prime ideal. Set $V = \mathbb{V}_+(I) \subset \mathbb{P}^m$ and $K[V] := R_{m+1}/I$, and assume that $K[V]$ is a unique factorisation domain. If $f\colon V \to \mathbb{P}^n$ is regular, then there exist $F_i \in K[V]$, homogeneous of the same degree, such that $f = [F_0, \dots, F_n]$ on all of V. In particular, $(X_0, \dots, X_n) \subset \operatorname{rad}(F_0, \dots, F_n)$.*

Proof. By definition there is an open cover $V = \bigcup_i V_i$ and $F_{ip} \in K[V]$, homogeneous of degree d_i, such that $f = [F_{i0}, \dots, F_{in}]$ on V_i. In particular, $V_i \subset \bigcup_p D_+(F_{ip})$. Since $K[V]$ is a UFD, we may further assume that the $F_{i0}, \dots, F_{in}$ are coprime.

On $V_i \cap V_j$ we have $[F_{i0}, \dots, F_{in}] = [F_{j0}, \dots, F_{jn}]$, so $F_{ip}F_{jq} = F_{iq}F_{jp}$. Thus the closed set $\mathbb{V}_+(\{F_{ip}F_{jq} - F_{iq}F_{jp}\})$ contains the open set $V_i \cap V_j$, which we know is dense since V is irreducible. It follows from the projective Nullstellensatz that $(\{F_{ip}F_{jq} - F_{iq}F_{jp}\}) = 0$, so $F_{ip}F_{jq} = F_{iq}F_{jp}$ in $K[V]$ for all p, q.

Since $K[V]$ is a UFD, we deduce that F_{ip} and F_{jp} differ only up to scalar, so there exists $\lambda \in K^\times$ with $F_{jp} = \lambda F_{ip}$ for all p. Hence $f = [F_{i0}, \dots, F_{in}]$ on $V_i \cup V_j$. Doing this for all j, we see that $f = [F_{i0}, \dots, F_{in}]$ on all of V.

In particular, $V = D_+(F_{i0}) \cup \cdots \cup D_+(F_{in})$, so $\mathbb{V}_+(F_{i0}, \dots, F_{in}) = \emptyset$, and hence $(X_0, \dots, X_n) \subset \operatorname{rad}(F_{i0}, \dots, F_{in})$. □

We now have several important corollaries.

Corollary 17.4. *Every regular map $\mathbb{P}^m \to \mathbb{A}^n$ is constant.*

Proof. Let $f\colon \mathbb{P}^m \to \mathbb{P}^n$ be regular. Since $K[\mathbb{P}^m] = R_{m+1}$, the previous proposition applies to give $f = [F_0, \dots, F_n]$ for some $F_i \in R_{m+1}$, homogeneous of the same degree. If the image of f is contained in the standard open affine U_0, then F_0 is nowhere vanishing on $\mathbb{P}^n$, and hence also on $\mathbb{A}^{n+1} - \{0\}$. Thus F_0 is a constant, so has degree zero, and hence each F_i is constant. □

In a similar vein, we have the following.

Corollary 17.5. *There is no non-constant morphism $f\colon \mathbb{P}^m \to \mathbb{P}^n$ when $m > n$.*

Proof. Let $f\colon \mathbb{P}^m \to \mathbb{P}^n$ be a non-constant regular map. Then we can write $f = [F_0, \dots, F_n]$ for non-constant $F_i \in R_{m+1}$, homogeneous of the same degree $d \geq 1$. In particular, $\mathbb{P}^m = \bigcup_i D_+(F_i)$, equivalently $\mathbb{V}(F_0, \dots, F_n) = \{0\}$ in $\mathbb{A}^{m+1}$.

We claim by induction on i that every irreducible component of $\mathbb{V}(F_0, \dots, F_i)$ has dimension at least $m - i$. Let X be an irreducible component of $\mathbb{V}(F_0, \dots, F_{i-1})$ (so $X = \mathbb{A}^{m+1}$ if $i = 0$). Then X is a cone, so $0 \in X \cap \mathbb{V}(F_i)$, and hence $F_i \in K[X]$ is not a unit. If F_i is zero in $K[X]$, then $X \cap \mathbb{V}(F_i) = X$ and we are done. Otherwise we use Theorem 13.7 to conclude that every irreducible component of $X \cap \mathbb{V}(F_i)$ has dimension $\dim X - 1 \geq m - i$.

In particular, since $\mathbb{V}(F_0, \dots, F_n) = \{0\}$ has dimension zero, we must have $n \geq m$. □

The next corollary shows that, in stark contrast to affine space, every automorphism of projective space is linear. Recall that the scalar multiples of the identity form a normal subgroup $K^\times \lhd \mathrm{GL}_{n+1}(K)$. We call the corresponding quotient group the **projective general linear group**, denoted $\mathrm{PGL}_{n+1}(K) := \mathrm{GL}_{n+1}(K)/K^\times$.

Corollary 17.6. *We have* $\mathrm{Aut}(\mathbb{P}^n) \cong \mathrm{PGL}_{n+1}(K)$.

Proof. Each $\theta \in \mathrm{GL}_{n+1}(K)$ determines an automorphism of $\mathbb{A}^{n+1}$ fixing the origin, and the composition $\pi\theta$ is constant on $K^\times$-orbits, so we have an induced automorphism of $\mathbb{P}^n$. Also, θ fixes every line through the origin in $\mathbb{A}^{n+1}$ if and only if $\theta \in K^\times$, so the induced group homomorphism $\Psi\colon \mathrm{PGL}_{n+1}(K) \to \mathrm{Aut}(\mathbb{P}^n)$ is injective.

Now take any $f \in \mathrm{Aut}(\mathbb{P}^n)$, say with inverse g. By the proposition we can write $f = [F_0, \dots, F_n]$ and $g = [G_0, \dots, G_n]$ with F_i homogeneous of degree d, and G_i homogeneous of degree e. Note that $gf = [H_0, \dots, H_n]$, where $H_i = G_i(F_0, \dots, F_n)$ is homogeneous of degree de. Also, $gf = \mathrm{id} = [X_0, \dots, X_n]$. Thus $X_i H_j = X_j H_i$ for all i, j, so by unique factorisation we must have $H_i = X_i H$ for some homogeneous H of degree $de - 1$. It follows that H is nowhere zero on $\mathbb{P}^n$, so is a constant. Hence $de = 1$, so $d = 1$ and f is given by homogeneous linear polynomials. It follows that Ψ is surjective, so an isomorphism. □

Finally we prove an extension theorem for regular maps.

Corollary 17.7. *Let $U \subset \mathbb{P}^1$ be non-empty open. Then any regular map $f\colon U \to \mathbb{P}^n$ can be extended uniquely to a regular map $f\colon \mathbb{P}^1 \to \mathbb{P}^n$.*[18]

Proof. Locally we can write $f = [F_0, \dots, F_n]$ for some $F_i \in K[S,T]$, homogeneous of the same degree. Since $K[S,T]$ is a UFD we may further assume that the F_i are coprime. Now $F_i(a,b) = 0$ if and only if $bS - aT$ divides F_i, so since the F_i are coprime we must have necessarily $\bigcup_i D_+(F_i) = \mathbb{P}^1$. Thus $g := [F_0, \dots, F_n]\colon \mathbb{P}^1 \to \mathbb{P}^n$ is regular.

We now have that $f, g\colon U \to \mathbb{P}^n$ are regular and agree on an open subset of U. Thus they agree on all of U, and hence $g\colon \mathbb{P}^1 \to \mathbb{P}^n$ is a regular map extending f. □

17.2 Examples

(1) Consider the regular map $f\colon \mathbb{V}_+(X^2 - YZ) \to \mathbb{P}^1$ given by $f = [X, Y]$ on $D_+(Y)$ and $f = [Z, X]$ on $D_+(Z)$. We claim that f is not given 'globally' by any pair $[F, G]$ of homogeneous polynomials $F, G \in R := K[X,Y,Z]/(X^2 - YZ)$ of the same degree.

For, the algebra homomorphism $R \to K[S,T]$, $(X,Y,Z) \mapsto (ST, S^2, T^2)$, is injective, so R is a domain (but not a UFD), and $\mathbb{V}(X^2 - YZ)$ is irreducible. Now $[F,G] = [Z,X]$ on the dense open $D_+(Z)$ implies that $FX = GZ$ in R. Consider $F(S, S^2, 1)$. If this is not constant, then it has a root a, and $G(a, a^2, 1) = aF(a, a^2, 1) = 0$, so $[F,G]$ is not defined at $[a, a^2, 1]$. Otherwise $[F,G] = [Z^{r+1}, XZ^r]$, so is not defined at $[0,1,0]$.

(2) The next example shows that isomorphic affine varieties can have non-isomorphic projective closures.

[18] This is a special case of the following more general result: if C is a smooth curve and $x \in C$, then any regular map $f\colon (C - \{x\}) \to \mathbb{P}^n$ can be extended uniquely to a regular map $f\colon C \to \mathbb{P}^n$.

There is an isomorphism $\mathbb{A}^1 \to \mathbb{V}(Z - Y^3)$, $t \mapsto (t, t^3)$. This gives a regular map $[1, T, T^3] \colon \mathbb{A}^1 \to \mathbb{P}^2$, which by Corollary 17.7 extends uniquely to the regular map $[S^3, S^2T, T^3] \colon \mathbb{P}^1 \to \mathbb{P}^2$, with image the projective closure $\mathbb{V}_+(X^2Z - Y^3)$.

This regular map is bijective, but is not an isomorphism. For, it restricts to the regular map $\mathbb{A}^1 \to \mathbb{V}(X^2 - Y^3)$, $s \mapsto (s^3, s^2)$, from the affine line to the cuspidal cubic.

(3) The previous example also shows that a regular map on an affine curve will in general not extend to its projective closure. For, we have the regular map $[1, Y] \colon \mathbb{V}(Z - Y^3) \to \mathbb{P}^1$, but this does not extend to a regular map $\mathbb{V}_+(X^2Z - Y^3) \to \mathbb{P}^1$.

18 Segre varieties and products

Recall that in the category of affine spaces we have the isomorphism $\mathbb{A}^m \times \mathbb{A}^n \cong \mathbb{A}^{m+n}$. The same cannot be true for projective space, since there are no non-constant maps $\mathbb{P}^m \to \mathbb{P}^n$ for $m > n$ by Corollary 17.5.

Consider affine space $\mathbb{A}^{(m+1)(n+1)}$, which we identify with the sapce of all matrices of size $(m+1) \times (n+1)$. The co-ordinate functions can thus be written as Z_{ip} for $0 \leq i \leq m$ and $0 \leq p \leq n$. We can then consider the homogeneous ideal $I := (\{Z_{ip}Z_{jq} - Z_{iq}Z_{ip}\}) \leq K[\{Z_{ip}\}]$. In terms of matrices, $\mathbb{V}(I)$ consists of those matrices of rank at most one.

We define the **Segre variety** as

$$\Sigma_{m,n} := \mathbb{V}_+(I) \subset \mathbb{P}^{(m+1)(n+1)-1}.$$

Lemma 18.1. *There are regular maps $\pi_1 \colon \Sigma_{m,n} \to \mathbb{P}^m$ and $\pi_2 \colon \Sigma_{m,n} \to \mathbb{P}^n$, given on the standard open affine $D(Z_{ip})$ by*

$$\pi_1 := [Z_{0p}, \ldots, Z_{mp}] \quad \text{and} \quad \pi_2 := [Z_{i0}, \ldots, Z_{in}].$$

Moreover, these induce a bijection of sets $(\pi_1, \pi_2) \colon \Sigma_{m,n} \to \mathbb{P}^m \times \mathbb{P}^n$, with inverse given by $\sigma([\underline{a}], [\underline{b}]) := [\ldots, a_i b_p, \ldots]$.

Proof. We need to check that the maps are well-defined. This is clear: we have

$$D(Z_{ip}) \cap D(Z_{jq}) = D(Z_{ip}Z_{jq}) = D(Z_{iq}Z_{jp}) = D(Z_{iq}) \cap D(Z_{jp}),$$

and on this open we have

$$[Z_{0p}, \ldots, Z_{mp}] = [Z_{0p}Z_{jq}, \ldots, Z_{mp}Z_{jq}] = [Z_{0q}Z_{jp}, \ldots, Z_{mq}Z_{jp}] = [Z_{0q}, \ldots, Z_{mq}].$$

For the final part, we know that some $a_i b_p$ is non-zero, so σ defines a map of sets. It is then easy to check that it is inverse to (π_1, π_2). □

Theorem 18.2. *The Segre variety $\Sigma_{m,n}$ is a product of $\mathbb{P}^m$ and $\mathbb{P}^n$ in the category of quasi-projective varieties.*

Proof. Let $V \subset \mathbb{P}^N$ be any quasi-projective variety, and consider regular maps $f \colon U \to \mathbb{P}^m$ and $g \colon U \to \mathbb{P}^n$. There is a unique set-theoretic map $h \colon V \to \Sigma_{m,n}$ such that $(\pi_1, \pi_2)h = (f, g)$, so we just need to check that h is regular.

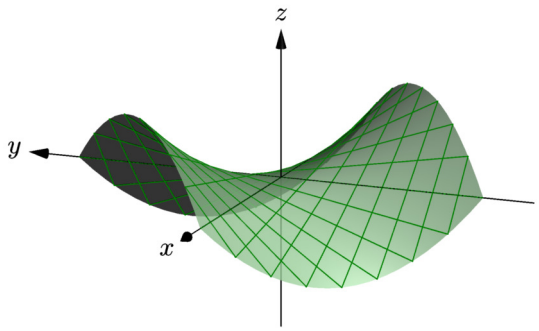
Locally, on some open $U \subset V$, we can write $f = [F_0, \ldots, F_m]$ and $g = [G_0, \ldots, G_n]$ for $F_i, G_j \in R_{N+1}$ homogeneous of degrees d, e. Then necessarily $h = [\ldots, F_i G_p, \ldots]$ on U, so is regular as required. □

Corollary 18.3. *The category of quasi-projective varieties has finite products. Moreover* $\mathbb{A}^m \times \mathbb{A}^n \cong \mathbb{A}^{m+n}$.

Proof. Let $X \subset \mathbb{P}^m$ and $Y \subset \mathbb{P}^n$ be quasi-projective varieties. Then their product is given by the locally closed subset $\pi_1^{-1}(X) \cap \pi_2^{-1}(Y) \subset \Sigma_{m,n}$.

If we identify as usual $\mathbb{A}^m = D(X_0) \subset \mathbb{P}^m$, and similarly for $\mathbb{A}^n$, then $\pi_1^{-1}(\mathbb{A}^m) \cap \pi_2^{-1}(\mathbb{A}^n) \subset \Sigma_{m,n}$ is contained in the standard open affine $D(Z_{00})$, so is an affine variety. Now products are unique up to isomorphism, and hence $\pi_1^{-1}(\mathbb{A}^m) \cap \pi_2^{-1}(\mathbb{A}^n) \cong \mathbb{A}^{m+n}$. □

The following picture shows the affine piece $D(W) \cong \mathbb{V}(XY - Z)$ the Segre variety $\Sigma_{1,1} \cong \mathbb{V}_+(WX - YZ)$, as well as the images of the grid of lines $\{[\underline{a}]\} \times \mathbb{P}^1$ and $\mathbb{P}^1 \times \{[\underline{b}]\}$.



18.1 Zariski topology on the product

We know that there is a bijection $\Sigma_{m,n} \cong \mathbb{P}^m \times \mathbb{P}^n$, and hence we can endow the Cartesian product with the Zariski topology coming from $\Sigma_{m,n}$. We can describe this directly in terms of bihomogeneous polynomials.

We say that a polynomial $f \in K[\{X_i, Y_p\}]$ is **bihomogeneous of degree** (d, e) provided that it is homogeneous of degree d when viewed as a polynomial in the X_i, and is homogeneous of degree e when viewed as a polynomial in the Y_p. This is if and only if

$$f(\lambda \underline{a}, \mu \underline{b}) = \lambda^d \mu^e f(\underline{a}, \underline{b}) \quad \text{for all } \lambda, \mu \in K \text{ and } (\underline{a}, \underline{b}) \in \mathbb{A}^{m+1} \times \mathbb{A}^{n+1}.$$

Proposition 18.4. *The closed subsets of* $\mathbb{P}^m \times \mathbb{P}^n$ *are precisely those given by the vanishing of bihomogeneous polynomials*

$$\{([\underline{a}], [\underline{b}]) : f_1(\underline{a}, \underline{b}) = \cdots = f_r(\underline{a}, \underline{b}) = 0\}.$$

Proof. We know from the projective Nullstellensatz that there is a bijection between proper homogeneous ideals of $K[\{Z_{ip}\}]$ containing I and the closed subsets of $\Sigma_{m,n}$.

We also have an algebra homomorphism $\theta\colon K[\{Z_{ip}\}] \to K[\{X_i, Y_p\}]$, $Z_{ip} \mapsto X_i Y_p$, with kernel containing I. Note that $\theta(g)([\underline{a}], [\underline{b}]) = g\sigma([\underline{a}], [\underline{b}])$.

Observe that $f \in K[\{X_i, Y_p\}]$ is bihomogeneous of degree (d, d) if and only if $f = \theta(g)$ for some $g \in K[\{Z_{ip}\}]$ homogeneous of degree d. For, it is enough to show it for monomials, where we have $\theta(Z_{i_1 p_1} \cdots Z_{i_d p_d}) = X_{i_1} \cdots X_{i_d} Y_{p_1} \cdots Y_{p_d}$.

Thus $\mathbb{V}_+(g) \cap \Sigma_{m,n}$ corresponds to the subset

$$\{([\underline{a}], [\underline{b}]) \in \mathbb{P}^m \times \mathbb{P}^n : \theta(g)(\underline{a}, \underline{b}) = 0\}.$$

Conversely, given a bihomogenous polynomial f of degree (d, e), we may assume that $d \geq e$ and thus construct the bihomogeneous polynomials $f_0, \ldots, f_n$, where $f_p := f Y_p^{d-e}$. Clearly the zero locus of f is the same as the common zeros of $f_0, \ldots, f_n$, and hence corresponds to the projective variety $\mathbb{V}_+(g_0, \ldots, g_n)$, where $\theta(g_p) = f_p$.

The result follows by taking finite intersections of such sets. □

We can therefore simplify notation and just write $\mathbb{P}^m \times \mathbb{P}^n$ for the product, where the topology is understood to be coming from bihomogeneous polynomials.

As an example, we see that $X_0^2 Y_1 + X_1^2 Y_0$ determines a closed subset of $\mathbb{P}^1 \times \mathbb{P}^1$, but the corresponding subset of $\Sigma_{1,1}$ is given by the two homogeneous polynomials

$$Z_{00} Z_{01} + Z_{10}^2 \quad \text{and} \quad Z_{01}^2 + Z_{10} Z_{11}.$$

Proposition 18.5. *Let $U \subset \mathbb{P}^m$ and $V \subset \mathbb{P}^n$ be locally closed subsets. Then the closure of $U \times V$ is $\bar{U} \times \bar{V}$. In particular, $\mathbb{P}^m \times \mathbb{P}^n$ is the closure of $\mathbb{A}^m \times \mathbb{A}^n$, hence is irreducible of dimension $m + n$.*

Proof. We know that $\bar{U} \times \bar{V}$ is closed, so we just need to show the reverse inclusion. Let f be a bihomogeneous polynomial vanishing on $U \times V$. Fixing $[\underline{a}] \in U$, we see that $f(\underline{a}, Y)$ is homogeneous and vanishes on V, hence vanishes on $\bar{V}$. Thus f vanishes on $U \times \bar{V}$. Now take $[\underline{b}] \in \bar{V}$. Then $f(X, \underline{b})$ vanishes on U, and hence on $\bar{U}$. Thus f vanishes on $\bar{U} \times \bar{V}$.

Now $\mathbb{P}^m \times \mathbb{P}^n$ is the closure of $\mathbb{A}^m \times \mathbb{A}^n$, which we know is irreducible of dimension $m + n$. Thus $\mathbb{P}^m \times \mathbb{P}^n$ is irreducible by Lemma 16.7 and has dimension $m + n$ by Theorem 16.10. □

18.2 Separatedness

Recall that a topological space X is **separated** provided the diagonal $\Delta_X \subset X \times X$ is closed. For many applications, this is an appropriate replacement of Hausdorff, and the two notions are equivalent when the product has the product topology.

Lemma 18.6. *The map $\Delta\colon \mathbb{P}^n \to \mathbb{P}^n \times \mathbb{P}^n$ is regular, with image given by the vanishing of the bihomogeneous polynomials $X_i Y_j - X_j Y_i$. It follows that all quasi-projective varieties are separated.*

Proof. The map $\Delta\colon \mathbb{P}^n \to \Sigma_{n,n}$ is given by $[\ldots, X_i X_j, \ldots]$, so is regular. Next observe that $[\underline{a}] = [\underline{b}] \in \mathbb{P}^n$ if and only if $a_i b_j = a_j b_i$ for all i, j. Thus the diagonal is given by

the vanishing of the bihomogeneous polynomials $X_iY_j - X_jY_i$ for all i, j, and hence is closed.

If $X \subset \mathbb{P}^n$ is locally closed, then $X \times X = \pi_1^{-1}(X) \cap \pi_2^{-1}(X) \subset \mathbb{P}^n \times \mathbb{P}^n$ is again locally closed. Thus $\Delta_X = \Delta \cap (X \times X)$ is closed in $X \times X$. □

Inside $\Sigma_{n,n}$ we have $\Delta = \mathbb{V}_+(\{Z_{ij} - Z_{ji}\})$. Identifying the cone over $\Sigma_{n,n}$ as matrices of rank at most one, we see that the preimage of the diagonal consists of the symmetric matrices of rank at most one.

Lemma 18.7. *Let $f\colon X \to Z$ and $g\colon Y \to Z$ be regular maps between quasi-projective varieties. Then the **fibred product** is closed in $X \times Y$*

$$X \times_Z Y := \{(x, y) : f(x) = g(y) \in Z\} \subset X \times Y.$$

Proof. We have regular maps $f\pi_1\colon X \times Y \to Z$ and $g\pi_2\colon X \times Y \to Z$, and thus a corresponding regular map $(f, g)\colon X \times Y \to Z \times Z$. Then $X \times_Z Y$ is just the preimage of the diagonal Δ_Z. □

As special cases we see that the **graph** of a regular map $f\colon X \to Z$ is closed in $X \times Z$

$$\Gamma_f := \{(x, f(x)) : x \in X\} = X \times_Z Z \subset X \times Z$$

and if $f, g\colon X \to Z$ are regular, then their **equaliser** is closed in X

$$\mathrm{Eq}\,(f, g) := \{x \in X : f(x) = g(x)\} = \Delta^{-1}(X \times_Z X) \subset X.$$

19 Completeness of projective varieties

We say that a topologicial space is **complete** provided that, for each Y, the projection $X \times Y \to Y$ is closed. Again,for many applications this is an appropriate replacement of compactness, and the two notions are equivalent when all products have the product topology.

For example, a continuous map $f\colon X \to Z$ is **universally closed** provided, for each map $Y \to Z$, the projection $X \times_Z Y \to Y$ is closed. We know from Proposition D.7 that if X is complete and Z is separated, then every continuous map $f\colon X \to Z$ is universally closed.

Theorem 19.1. *Every projective variety is complete.*

Proof. We begin by showing that the projection map $\pi\colon \mathbb{P}^m \times \mathbb{P}^n \to \mathbb{P}^n$ is closed, so let $V \subset \mathbb{P}^m \times \mathbb{P}^n$ be closed, say given by the bihomogeneous polynomials $f_1, \ldots, f_r$, where f_i has degree (d_i, e_i).

Now $[\underline{b}] \notin \pi(V)$ provided $\{[\underline{a}] \in \mathbb{P}^m : ([\underline{a}], [\underline{b}]) \in V\} = \emptyset$. The right hand side equals $\mathbb{V}_+(f_1(X, \underline{b}), \ldots, f_r(X, \underline{b}))$, so by the projective Nullstellensatz the condition becomes $(X_0, \ldots, X_m) \subset \mathrm{rad}(f_1(X, \underline{b}), \ldots, f_r(X, \underline{b}))$.

We therefore introduce the sets

$$U_N := \{[\underline{b}] \in \mathbb{P}^m : (X_0, \ldots, X_m)^N \subset (f_1(X, \underline{b}), \ldots, f_r(X, \underline{b}))\}.$$

Then clearly $\pi(V)^c = \bigcup_N U_N$, so if each U_N is open, then $\pi(V)^c$ is open, and hence $\pi(V)$ is closed as required.

It remains to show that each U_N is open. The monomials $M_\alpha := X_0^{\alpha_0} \cdots X_m^{\alpha_m}$ form a basis for polynomial algebra, and M_α is homogeneous of degree $\alpha_0 + \cdots + \alpha_m$. Set $D_t = \binom{m+t}{t}$ to be the number of M_α of degree t.

We have $(X_0, \ldots, X_m)^N \subset (f_1(X, \underline{b}), \ldots, f_r(X, \underline{b}))$ if and only if every homogeneous polynomial of degree N can be written in the form $\sum_i f_i(X, \underline{b}) g_i(X)$ for some homogeneous polynomials $g_i(X)$ of degree $N - d_i$. Thus, for each monomial M_α of degree $N - d_i$, we can write $f_i M_\alpha$ as a linear combination of the monomials M_β of degree N, with coefficients $h_\beta \in K[\{Y_p\}]$. We arrange the h_β into a column vector, and then form the matrix S by taking all such columns coming from the various products $f_i M_\alpha$. Finally, let $J \lhd K[\{Y_p\}]$ be the (homogeneous) ideal generated by all the minors of S of size D_N.

It follows that $[\underline{b}] \in U_N$ if and only if the matrix $S(\underline{b})$ has rank D_N, if and only if $[\underline{b}] \notin \mathbb{V}(J_N)$. This proves that each U_N is open, as required.

In general, let $X \subset \mathbb{P}^m$ be closed and $Y \subset \mathbb{P}^n$ be locally closed. We wish to show to that the projection $\pi \colon X \times Y \to Y$ is closed. Let $V \subset X \times Y$ be closed, and note that its closure $\bar{V}$ lies in $X \times \bar{Y}$, and that $V = \bar{V} \cap (X \times Y)$. It follows that the image $\pi(\bar{V}) \subset \mathbb{P}^n$ is closed, and that $\pi(V) = \pi(\bar{V}) \cap Y$ is closed in Y. Hence X is complete. □

We can now generalise Corollary 17.4 to all projective varieties.

Corollary 19.2. *Every regular map from an irreducible projective variety to an affine variety is constant.*

Proof. Let X be an irreducible projective variety, and $f \colon X \to \mathbb{A}^1$ a regular map. Then the composition is a regular map $X \to \mathbb{P}^1$, so must have closed image. Since it is not surjective, the image is a finite union of points, and since X is irreducible the image must be a single point.

In general, given a regular map $f \colon X \to \mathbb{A}^n$, each of the co-ordinates $f_i \colon X \to \mathbb{A}^1$ are constant, so f is itself constant. □

Finally, projective varieties are characterised by being complete.

Corollary 19.3. *Let $X \subset \mathbb{P}^n$ be locally closed. Then X is complete if and only if $X \subset \mathbb{P}^n$ is closed.*

Proof. We know that if X is closed, then it is complete. Conversely, the diagonal $\Delta \subset X \times \mathbb{P}^n$ is closed, since it is the graph of the inclusion map. Hence its image $\pi_2(\Delta) = X$ is closed in $\mathbb{P}^n$. □

It follows that every complete affine variety is a finite union of points.

Part VII

Appendix

A Unique Factorisation Domains

We say that an element $a \in R$ **divides** b, written $a|b$, if there exists $x \in R$ such that $b = ax$. Equivalently, $b \in (a)$, or $(b) \subset (a)$. Note that 1 divides every other element, and each element divides 0.

If R is an integral domain, then $a|b$ and $b|a$ if and only if there exists a unit $u \in R^\times$ such that $b = au$, in which case we call a and b **associates**. For, there exist $u, v \in R$ such that $b = au$ and $a = bv$. If $b = 0$ then $a = 0$. Otherwise, since $b = buv$, we have $uv = 1$, so that $u, v \in R^\times$ are units.

Let R be an integral domain and $a \in R$ non-zero and not a unit. We call a

prime	if $a\|xy$ implies $a\|x$ or $a\|y$.
irreducible	if $a = xy$ implies x is a unit or y is a unit.

Proposition A.1. *Let R be an integral domain and $a \in R$ non-zero and not a unit.*

(1) a is prime if and only if (a) is a prime ideal.

(2) a prime implies a irreducible. The converse holds if R is a principal ideal domain, in which case (a) is a maximal ideal.

Proof. (1) Let a be prime and suppose that $xy \in (a)$. Then $a|xy$, whence $a|x$ or $a|y$. In other words, $x \in (a)$ or $y \in (a)$, so that (a) is a prime ideal. The converse is similar.

(2) Let a be prime and suppose that $a = xy$. Without loss of generality $a|x$, so that $x = ab$ for some b. Now $a = xy = aby$, so $by = 1$ and y is a unit. Thus a is irreducible.

Now suppose that R is a principal ideal domain and let a be irreducible. Suppose that $(a) \subset (x)$. Then $a = xy$ for some y, and since a is irreducible, either x is a unit, in which case $(x) = R$, or else y is a unit, in which case $(a) = (x)$. Hence (a) is a maximal ideal. □

An integral domain R is called a **unique factorisation domain** if every element can be written uniquely as a product of irreducibles, in the following sense

existence	each non-zero $a \in R$ can be written as $a = ux_1 \cdots x_m$ with u a unit and x_i irreducible.
uniqueness	given two such expressions $a = ux_1 \cdots x_m$ and $a = vy_1 \cdots y_n$, then $m = n$ and (after re-ordering) x_i and y_i are associates.

Clearly if R is a unique factorisation domain, then every irreducible element is prime. For, if a is irreducible and $xy \in (a)$, then $xy = ab$ for some b. Since a is irreducible and factorisations are unique, a must occur in the factorisation of either x or y, whence $x \in (a)$ or $y \in (a)$.

Theorem A.2. *Every principal ideal domain is a unique factorisation domain.*[19]

[19] In fact, if R is a Noetherian integral domain, then R is a unique factorisation domain if and only

Proof. Let R be a principal ideal domain. We first show that every increasing sequence of ideals stabilises (so that R is **Noetherian**).

Suppose we have an increasing sequence of ideals $I_1 \subset I_2 \subset \cdots$. Then the union $I := \bigcup_i I_i$ is again an ideal, and since R is a principal ideal domain we can write $I_i = (a_i)$ and $I = (a)$. Now, $a \in \bigcup_i I_i$, so $a \in I_i$ for some i. Therefore $I \subset I_i$, so $I = I_i$, and hence $I = I_n$ for all $n \geq i$.

Now take $a_1 \in R$ non-zero and not a unit, and suppose for contradiction that a_1 cannot be written as a product of irreducibles. Then a_1 is not irreducible, so we can write $a_1 = a_2 a_2'$ with neither a_2 nor a_2' a unit. If both a_2 and a_2' can be expressed as a product of irreducibles, then the same would be true of a_1, so we may assume that a_2 cannot be written as a product of irreducibles. Repeating the argument yields an increasing sequence of ideals $(a_1) \subset (a_2) \subset \cdots$. Also, by construction, $(a_{i-1}) \neq (a_i)$, since $a_{i-1} = a_i a_i'$ and a_i' is not a unit. Therefore this sequence of ideals does not stabilise, contradicting the above result.

To see that this expression is unique, let $a = x_1 \cdots x_m = y_1 \cdots y_n$ with each x_i and y_j irreducible. Since (x_1) is a prime ideal (in fact maximal), $R/(x_1)$ is an integral domain (in fact a field) and $\bar{y}_1 \cdots \bar{y}_n = \bar{a} = 0$ in $R/(x_1)$. Thus, after re-ordering, $\bar{y}_1 = 0$. Hence $y_1 \in (x_1)$, say $y_1 = u_1 x_1$. Since both x_1 and y_1 are irreducible, u_1 must be a unit. Therefore $(x_1) = (y_1)$ and $x_2 \cdots x_m = u_1 y_2 \cdots y_n$. Since $y_2' := u_1 y_2$ is irreducible and $(y_2') = (y_2)$, the result follows by induction on $m + n$. □

For the remainder of this section, R will denote a unique factorisation domain and K its quotient field. Our aim is to prove that the polynomial ring $R[X]$ is again a unique factorisation domain.

Lemma A.3. *Any two elements in R have a **greatest common divisor**, and this is unique up to associates.*

Proof. Given a and b, write $ab = u x_1^{m_1} \cdots x_n^{m_n}$ for some unit u and pairwise non-associate irreducible elements x_i (so $(x_i) \neq (x_j)$ for $i \neq j$). We can now write $a = u' x_1^{r_1} \cdots x_n^{r_n}$ and $b = u'' x_1^{s_1} \cdots x_n^{s_n}$ for some units u', u''. Note that $m_i = r_i + s_i$. Set $\gcd(a, b) := x_1^{l_1} \cdots x_n^{l_n}$, where $l_i := \min(r_i, s_i)$.

Clearly $\gcd(a, b)$ divides both a and b, and any other element which divides both a and b must divide $\gcd(a, b)$ by unique factorisation. □

Given a non-zero polynomial $f = a_n X^n + \cdots + a_1 X + a_0 \in R[X]$, we define its **content** $\mathrm{cont}(f)$ to be the greatest common divisor of the coefficients a_i. We call f **primitive** if $\mathrm{cont}(f)$ is a unit. Note that, if $0 \neq d \in R$, then $\mathrm{cont}(df) = d \cdot \mathrm{cont}(f)$.

More generally, let $f \in K[X]$ be non-zero. By clearing denominators, there exists $0 \neq d \in R$ such that $df \in R[X]$. We therefore define $\mathrm{cont}(f) := \mathrm{cont}(df)/d \in K$. To see that this is well-defined let $0 \neq d' \in R$ also satisfy $d'f \in R[X]$. Then

$$d' \cdot \mathrm{cont}(df) = \mathrm{cont}(dd'f) = d \cdot \mathrm{cont}(d'f),$$

so that $\mathrm{cont}(df)/d = \mathrm{cont}(d'f)/d'$. It follows as before that if $d \in K^\times$ and $f \in K[X]$, then $\mathrm{cont}(df) = d \cdot \mathrm{cont}(f)$.

---

if all irreducible elements are prime. The proof is the same, but using the Noetherian property to deduce that the ascending chain of ideals stabilises. This will be an exercise later.

Lemma A.4. *Let $f, g \in K[X]$ be non-zero.*

(1) $f/\mathrm{cont}(f) \in R[X]$ *and is primitive. Conversely, if* $c \in K^\times$ *is such that* $f/c \in R[X]$ *is primitive, then* $c = \mathrm{cont}(f)$ *(up to a unit of* R*).*

(2) $\mathrm{cont}(f) \in R$ *if and only if* $f \in R[X]$.

(3) $\mathrm{cont}(fg) = \mathrm{cont}(f)\mathrm{cont}(g)$.

Proof. 1. Suppose first that $f \in R[X]$ and has coefficients a_i. Since $\mathrm{cont}(f) = \gcd(a_i)$ we know that $a_i/\mathrm{cont}(f) \in R$ and that these elements are coprime. Thus $f/\mathrm{cont}(f) \in R[X]$ is primitive.

Now let $f \in K[X]$. Taking $0 \neq d \in R$ such that $df \in R[X]$ we see that $f/\mathrm{cont}(f) = df/\mathrm{cont}(df) \in R[X]$ is primitive.

Finally, let $c \in K^\times$ be such that $f/c \in R[X]$ is primitive. Then $1 = \mathrm{cont}(f/c) = \mathrm{cont}(f)/c$, so that $c = \mathrm{cont}(f)$.

2. By (1) we can write $f = \mathrm{cont}(f)f'$ for some $f' \in R[X]$ primitive, so $\mathrm{cont}(f) \in R$ implies $f \in R[X]$. The converse is immediate.

3. Set $c := \mathrm{cont}(f)$ and $d := \mathrm{cont}(g)$. By (1) we can write $f = cf'$ and $g = dg'$ for some $f', g' \in R[X]$ primitive. Then $fg = cdf'g'$ and $f'g' \in R[X]$, so if we can show that $f'g'$ is primitive, then $\mathrm{cont}(fg) = cd$ as required.

Let $p \in R$ be prime and consider the quotient ring $(R/(p))[X]$. Since $R/(p)$ is an integral domain, so too is $(R/(p))[X]$. Since f' and g' are primitive, we know that p does not divide every coefficient of f' or g', so $\overline{f'}$ and $\overline{g'}$ are non-zero in $(R/(p))[X]$. Thus $\overline{f'g'} = \overline{f'} \cdot \overline{g'}$ is non-zero, so p does not divide $\mathrm{cont}(f'g')$.

It follows that $\mathrm{cont}(f'g')$ is not divisible by any irreducible element of R, hence is a unit, and $f'g'$ is primitive. □

Lemma A.5 (Gauss's Lemma)**.** *If $f \in R[X]$ is irreducible over R, then it is irreducible over K. The converse holds when f is primitive.*

Proof. We prove the contrapositive. Suppose $f = gh \in K[X]$. Since $\mathrm{cont}(f) = \mathrm{cont}(g)\mathrm{cont}(h)$ we can factorise f over R as

$$f = \mathrm{cont}(f) \cdot (g/\mathrm{cont}(g)) \cdot (h/\mathrm{cont}(h)).$$

Conversely let $f \in R[X]$ be primitive and suppose that f is irreducible over K. Let $f = gh$ be a factorisation over R. Since f is irreducible over K we may assume without loss of generality that g is a unit in $K[X]$, so $\deg(g) = 0$ and hence $g \in R$. Therefore g divides $\mathrm{cont}(f)$, which is a unit since f is primitive. Hence g is a unit, so f is irreducible over R. □

Theorem A.6. *The polynomial ring $R[X]$ is again a unique factorisation domain. The units of $R[X]$ are the units of R. The irreducible elements of $R[X]$ are the irreducible elements of R together with the primitive irreducible polynomials.*

Proof. Since R is an integral domain, we can consider leading terms of polynomials to deduce that $R[X]$ is also an integral domain and that the units of $R[X]$ are just the

units of R. Also, by considering degrees, we see that each irreducible in R remains irreducible in $R[X]$.

Let $f \in R[X]$ be non-constant. Since $K[X]$ is a principal ideal domain, it is a unique factorisation domain, so we can write $f = g_1 \cdots g_r$ with each g_i irreducible in $K[X]$. Set $c_i := \mathrm{cont}(g_i)$, $c := c_1 \cdots c_r$ and $f_i := g_i/c_i$, so $f_i \in R[X]$ is a primitive irreducible polynomial by Gauss's Lemma and $f = cf_1 \cdots f_r$. Then $c = \mathrm{cont}(f) \in R$, so can be written as a product of irreducibles in R. Thus each polynomial can be written as a product of irreducible elements.

To see that this expression is unique, suppose that $f = cg_1 \cdots g_r$ and $f = dh_1 \cdots h_s$ with $c, d \in R$ and $g_i, h_j \in R[X]$ primitive irreducible polynomials. Then $g_i, h_j \in K[X]$ are irreducible by Gauss's Lemma, so using that $K[X]$ is a unique factorisation domain we deduce that, after reordering, $r = s$ and $h_i = u_i g_i$ for some $u_i \in K^\times$. Then $u_i = \mathrm{cont}(h_i) \in R^\times$, so g_i and h_i are associates. Finally, setting $u := u_1 \cdots u_r \in R^\times$ gives that $c = ud \in R$, so c and d are associates. Since R is a unique factorisation domain, we are done. $\square$

Category Theory

B Categories

Categories provide a convenient language and setting to deal with common ideas found in diverse areas of mathematics, including algebra and topology. They abstract from the notion of sets and maps, to objects and morphisms, so many proofs become 'arrow theoretic', rather than working with elements of sets. For example, one can no longer say that a map is injective if it is one-to-one on elements, so one introduces the notion of a monomorphism instead.

A **category** $\mathcal{C}$ is given by a collection of objects $\mathrm{Ob}(\mathcal{C})$ and a collection of morphisms $\mathrm{Hom}(A, B)$ for each pair of objects A and B, together with a composition operation

$$\mathrm{Hom}(B, C) \times \mathrm{Hom}(A, B) \to \mathrm{Hom}(A, C), \quad (g, f) \mapsto gf,$$

which is

Associative	$h(gf) = (hg)f$ for all composable morphisms f, g, h
Unital	there exists $1_A \in \mathrm{Hom}(A, A)$ for each A such that $f1_A = f = 1_B f$ for all $f \in \mathrm{Hom}(A, B)$.

Examples of categories include

(1) the category of sets, **Set**, having as objects all sets, and as morphisms all maps between sets. Also, its subcategory of finite sets, **FinSet**.

(2) the category of groups, **Gp**, having as objects all groups, and as morphisms all group homomorphisms. Also, its subcategories of abelian groups **Ab**, finite groups **FinGp**, and finite abelian groups **FinAb**.

(3) the category of modules over a fixed ring, $\mathbf{Mod}_R$, with morphisms all module homomorphisms. This generalises the category of vector spaces over a field K, $\mathbf{Vec}_K = \mathbf{Mod}_K$, as well as the category of abelian groups $\mathbf{Ab} = \mathbf{Mod}_{\mathbb{Z}}$.

(4) the category of rings, **Ring**, with morphisms all ring homomorphisms. Also, its subcategory of commutative rings **CommRing**, and its subcategory of k-algebras $\mathbf{CommAlg}_k$.

(5) the category of topological spaces, **Top**, with morphisms all continuous maps.

(6) the category arising from a poset (partially ordered set), having as objects all elements of the set, and a unique morphism in $\mathrm{Hom}(x, y)$ whenever $x \leq y$.

As is usual for functions and homomorphisms, we also write $f\colon A \to B$ instead of $f \in \mathrm{Hom}(A, B)$.

We will generally only be interested in **locally small** categories, where each $\mathrm{Hom}(A, B)$ is actually a set. A locally small category is called **small** if the collection of objects is also a set. For example, the category of sets **Set** is locally small but not small.

A **subcategory** $\mathcal{D}$ of a category $\mathcal{C}$ is given by taking subcollections

$$\mathrm{Ob}(\mathcal{D}) \subset \mathrm{Ob}(\mathcal{C}) \quad \text{and} \quad \mathrm{Hom}_{\mathcal{D}}(A,B) \subset \mathrm{Hom}_{\mathcal{C}}(A,B) \text{ for all } A,B \in \mathrm{Ob}(\mathcal{D})$$

such that the restriction of composition of morphisms makes $\mathcal{D}$ into a category. In other words, we have $1_A \in \mathrm{Hom}_{\mathcal{D}}(A,A)$ for all $A \in \mathrm{Ob}(\mathcal{D})$ and $gf \in \mathrm{Hom}_{\mathcal{D}}(A,C)$ for all $f \in \mathrm{Hom}_{\mathcal{D}}(A,B)$ and $g \in \mathrm{Hom}_{\mathcal{D}}(B,C)$.

A subcategory is called **full** if $\mathrm{Hom}_{\mathcal{D}}(A,B) = \mathrm{Hom}_{\mathcal{C}}(A,B)$ for all $A,B \in \mathrm{Ob}(\mathcal{D})$.

If $\mathcal{C}$ is a category, then we can form the **opposite** category $\mathcal{C}^{\mathrm{op}}$, having the same objects as $\mathcal{C}$ but reversing all morphisms, so $\mathrm{Hom}_{\mathcal{C}^{\mathrm{op}}}(A,B) := \mathrm{Hom}_{\mathcal{C}}(B,A)$, and hence with the opposite composition to $\mathcal{C}$; that is, if $f\colon A \to B$ in $\mathcal{C}$, then we have $f^{\mathrm{op}}\colon B \to A$ in $\mathcal{C}^{\mathrm{op}}$, and $f^{\mathrm{op}}g^{\mathrm{op}} := (gf)^{\mathrm{op}}$ whenever f and g are composable in $\mathcal{C}$.

One of the principles of category theory is to unify various constructions occuring in mathematics by providing an appropriate abstract setting in which to work. As such, we can no longer assume that the objects under consideration are sets, and so we cannot apply our naive notions of injectivitiy and surjectivity to morphisms.

To this end, we call a morphism $f\colon A \to B$

a **monomorphism**	if, for all $g,h\colon X \to A$, we have $fg = fh$ implies $g = h$.
an **epimorphism**	if, for all $g,h\colon B \to X$, we have $gf = hf$ implies $g = h$.
an **isomorphism**	if there exists $g\colon B \to A$ with $fg = 1_B$ and $gf = 1_A$.

Note that being an isomorphism is stronger than being both a monomorphism and an epimorphism.

B.1 Initial and terminal objects

An object I in a category $\mathcal{C}$ is called **initial** if there exists a unique morphism $\iota_A\colon I \to A$ for each object A in $\mathcal{C}$.

Similarly, an object T is called **terminal** if there exists a unique morphism $\pi_A\colon A \to T$ for each object A.

Note that these definitions are dual to one another, in the sense that an object I is initial in a category $\mathcal{C}$ if and only if it is terminal in the opposite category $\mathcal{C}^{\mathrm{op}}$.

In general, we say that two properties of a category are **dual to one another** if they are swapped under taking the opposite category. In particular, if a property is given by certain arrows (morphisms), then the dual property is given by reversing all arrows. For example, the notions of monomorphism and epimorphism are dual to one another.

Lemma B.1. *Initial and terminal objects, if they exist, are unique up to unique isomorphism.*

Proof. Let I and I' both be initial objects of a category $\mathcal{C}$. Since I is an initial object, there exists a unique morphism $f\colon I \to I'$. Similarly, since I' is an initial object, there exists a unique morphism $g\colon I' \to I$. Composing these gives a morphism $gf\colon I \to I$, which by uniqueness must equal the identity map 1_I. Similarly $fg = 1_{I'}$. □

A **zero object** in a category is an object which is both initial and terminal. In this case we write 0 for the zero object, and also for the unique morphisms $0 \to A$ and $A \to 0$ for each A. Note that $1_0 = 0$ in this notation.

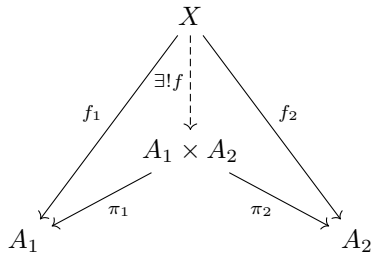
B.2 Products and coproducts

Given a family of objects A_i in a category (indexed by some set I), their **product**, if it exists, is an object $\prod_i A_i$ together with a family of morphisms $\pi_i \colon \prod_i A_i \to A_i$ such that, for each object X, the natural map

$$\mathrm{Hom}\left(X, \prod_i A_i\right) \to \prod_i \mathrm{Hom}(X, A_i), \quad f \mapsto (\pi_i f)_i,$$

is a bijection. Note that the product on the right hand side is the usual Cartesian product of sets.

In other words, whenever we have a family of morphisms $f_i \colon X \to A_i$, there exists a unique morphism $f \colon X \to \prod_i A_i$ satisfying $f_i = \pi_i f$ for all i. For the product of two objects, $A_1 \times A_2$, this can be represented by the commutative diagram

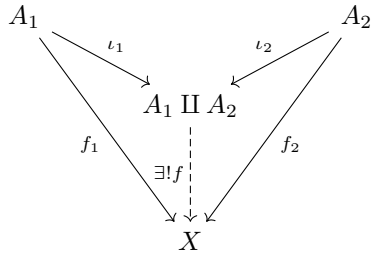


We also have the dual notion of coproducts. Given a family A_i, their **coproduct**, if it exists, is an object $\coprod_i A_i$ together with a family of morphisms $\iota_i \colon A_i \to \coprod_i A_i$ such that, for each object X, the natural map

$$\mathrm{Hom}\left(\coprod_i A_i, X\right) \to \prod_i \mathrm{Hom}(A_i, X), \quad f \mapsto (f\iota_i)_i,$$

is a bijection

In other words, whenever we have a family of morphisms $f_i \colon A_i \to X$, there exists a unique morphism $f \colon \coprod_i A_i \to X$ satisfying $f_i = f\iota_i$ for all i. For the coproduct $A_1 \amalg A_2$ we can represent this by the commutative diagram



It is important to note that, as with initial and terminal objects in any category, products and coproducts need not exist, but when they do, they are unique up to unique isomorphism. We will return to this when we look at universal properties.

Lemma B.2. *Let $\mathcal{C}$ be a category with a zero object, let A_i be a family of objects, and suppose that both the product and the coproduct exist. Then there is a unique*

morphism

$$\varphi\colon \coprod_i A_i \to \prod_i A_i$$

such that

$$\pi_i \varphi \iota_i = 1_{A_i} \quad \textit{and} \quad \pi_i \phi \iota_j = 0 \textit{otherwise.}$$

Proof. Define morphisms $f_{ji}\colon A_i \to A_j$ such that $f_{ii} = 1_{A_i}$ and $f_{ji} = 0$ otherwise.

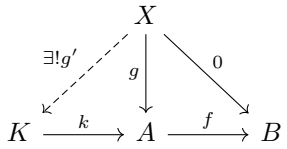
For a fixed j, the family $f_{ji}\colon A_i \to A_j$ induces a unique morphism $f_j\colon \coprod_i A_i \to A_j$ such that $f_{ji} = f_j \iota_i$ for all i. Doing this for each j, we obtain a family of morphisms f_j, which in turn induces a unique morphism $\varphi\colon \coprod_i A_i \to \prod_i A_i$ such that $f_j = \pi_j \varphi$. It follows that $f_{ji} = f_j \iota_i = \pi_j \varphi \iota_i$. □

B.3 Kernels and cokernels

Now suppose that our category has a zero object.

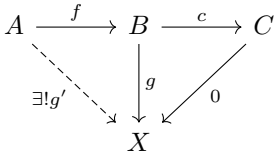
A **kernel** of a morphism $f\colon A \to B$ is an object K together with a morphism $k\colon K \to A$ such that $fk = 0$ and, whenever $g\colon X \to A$ satisfies $fg = 0$, then there exists a unique morphism $g'\colon X \to K$ with $g = kg'$.

We can represent this as a commutative diagram



Dually, we have the notion of a **cokernel** of a morphism $f\colon A \to B$. This is given by an object C and a morphism $c\colon B \to C$ such that $cf = 0$ and, whenever $g\colon B \to X$ satisfies $gf = 0$, then there exists a unique morphism $g'\colon C \to X$ with $g = g'c$.

We can represent this as a commutative diagram



Again, kernels and cokernels, if they exist, are unique up to unique isomorphism.

Lemma B.3. *Every kernel is a monomorphism, and every cokernel is an epimorphism.*

Proof. Let $k\colon K \to A$ be the kernel of a morphism $f\colon A \to B$. Given $g\colon X \to K$, set $j := kg$. Then $fj = 0$, so g is the unique map such that $kg = j$. Thus, if $kg = kh$, then $g = h$. □

Suppose now that all kernels and cokernels exist. Then, given a morphism $f\colon A \to B$ we have both a kernel $k\colon \mathrm{Ker}(f) \to A$ and a cokernel $c\colon B \to \mathrm{Coker}(f)$. Furthermore, k has a cokernel $\pi\colon A \to \mathrm{Coim}(f)$ and c has a kernel $\iota\colon \mathrm{Im}(f) \to B$.

Lemma B.4. *There is a unique morphism* $\bar{f}\colon \mathrm{Coim}(f) \to \mathrm{Im}(f)$ *such that* $f = \iota \bar{f} \pi$.

Proof. We know $fk = 0$, so since π is a cokernel of k we have $f = f'\pi$ for some $f'\colon \mathrm{Coim}(f) \to B$. Now, π is an epimorphism and $cf'\pi = cf = 0$. Thus $cf' = 0$. Finally, since ι is a kernel of c we have $f' = \iota \bar{f}$ for some $\bar{f}\colon \mathrm{Coim}(f) \to \mathrm{Im}(f)$. □

B.4 Additive and abelian categories

Two important classes of categories are those of additive categories, and more specifically, abelian categories.

An **additive category** is a (locally small) category such that

Add1 there exists a zero object.

Add2 for each finite family A_i, the product and coproduct exist and the natural map $\varphi\colon \coprod_i A_i \to \prod_i A_i$ is an isomorphism.

Add3 each set $\mathrm{Hom}(A, B)$ is equipped with the structure of an abelian group under addition[20] such that the composition of morphisms is bilinear. In this case, the zero morphism $A \to B$ is necessarily equal to the composition of the unique maps $A \to 0 \to B$.

An **abelian category** is an additive category such that

Ab1 all kernels and cokernels exist.

Ab2 for all morphisms $f\colon A \to B$, the induced morphism $\bar{f}\colon \mathrm{Coim}(f) \to \mathrm{Im}(f)$ is an isomorphism.

The motivating example for abelian categories is the category of abelian groups. More generally, the category of modules over a fixed ring is an abelian category. The category of rings and ring homomorphisms is not an abelian category, though.

C Functors

Having defined categories, we now want to consider morphisms between categories.

Let $\mathcal{C}$ and $\mathcal{D}$ be categories. A (covariant) **functor** $F\colon \mathcal{C} \to \mathcal{D}$ is given by a map $F\colon \mathrm{Ob}(\mathcal{C}) \to \mathrm{Ob}(\mathcal{D})$ and maps $F\colon \mathrm{Hom}_{\mathcal{C}}(A, B) \to \mathrm{Hom}_{\mathcal{D}}(F(A), F(B))$ for each pair of objects A, B of $\mathcal{C}$ such that

Fun1 F preserves identities, so $F(1_A) = 1_{F(A)}$ for all objects A in $\mathcal{C}$.

Fun2 F respects composition, so $F(gf) = F(g)F(f)$ for all composable morphisms f, g in $\mathcal{C}$.

[20] It is interesting to note that, under the assumptions Add1 and Add2, there is a natural addition on each hom set, giving it the structure of an abelian monoid, and for which composition is bilinear. Moreover, if Add3 holds, then this addition must equal the naturally-defined one. Thus we may replace Add3 by the axiom that every homomorphism has an additive inverse with respect to the natural addition law given by Add1 and Add2. See the exercises.

For example, the identity map on a category is clearly a functor from that category to itself (an endofunctor). More interestingly, the map sending a set to its power set defines an endofunctor on the category of sets.

As another example, since a group is given by a set together with a binary operation satisfying certain axioms, we have the **forgetful functor** from the category of groups to the category of sets, $\mathbf{Gp} \to \mathbf{Set}$ defined by sending a group to its underlying set (and hence forgetting the binary operation).

Clearly, given functors $F\colon \mathcal{C} \to \mathcal{D}$ and $G\colon \mathcal{D} \to \mathcal{E}$, we can compose them to get a functor $GF\colon \mathcal{C} \to \mathcal{E}$. It quickly follows that there is a category **Cat** having objects all small categories, and morphisms all functors between small categories.

Two categories $\mathcal{C}$ and $\mathcal{D}$ are thus called **isomorphic** if there exist functors $F\colon \mathcal{C} \to \mathcal{D}$ and $G\colon \mathcal{D} \to \mathcal{C}$ such that $GF = 1_{\mathcal{C}}$ and $FG = 1_{\mathcal{D}}$. This notion is not especially interesting[21], since for example it implies that there is a bijection between the objects in the two categories, which is an unnecessarily strong condition for the majority of applications. The more appropriate definition is that of equivalence, for which we need natural transformations.

If one is considering only additive or abelian categories, then in line with viewing morphisms as 'structure preserving maps', we often restrict ourselves to **additive functors**. These are functors such that the maps $F\colon \operatorname{Hom}(A, B) \to \operatorname{Hom}(F(A), F(B))$ are group homomorphisms. Note that a functor between additive categories is itself additive if and only if it preserves all finite products (or coproducts).

C.1 Natural transformations

Let $F, G\colon \mathcal{C} \to \mathcal{D}$ be functors. A **natural transformation** $\eta\colon F \Rightarrow G$ consists of a morphism $\eta_A\colon F(A) \to G(A)$ in $\mathcal{D}$ for each object $A \in \mathcal{C}$ such that, whenever $f\colon A \to B$ is a morphism in $\mathcal{C}$, we have a commutative diagram in $\mathcal{D}$

$$\begin{array}{ccc} F(A) & \xrightarrow{F(f)} & F(B) \\ \downarrow{\scriptstyle \eta_A} & & \downarrow{\scriptstyle \eta_B} \\ G(A) & \xrightarrow{G(f)} & G(B) \end{array}$$

Now, given natural transformations $\eta\colon F \Rightarrow G$ and $\xi\colon G \Rightarrow H$, we can compose these to get a natural transformation $\xi\eta\colon F \Rightarrow H$.

It follows that, given categories $\mathcal{C}$ and $\mathcal{D}$, there is a functor category $\operatorname{Fun}(\mathcal{C}, \mathcal{D})$ having as objects the functors from $\mathcal{C}$ to $\mathcal{D}$, and as morphisms the natural transformations between functors[22].

A **natural isomorphism** is a natural transformation such that, for each $A \in \mathcal{C}$, the morphism η_A is an isomorphism in $\mathcal{D}$. We say that two functors are **naturally isomorphic** if there exists a natural isomorphism between them.

[21] In the nLab, such a notion is deemed evil.

[22] One can also construct a 2-category having as objects the small categories, as morphisms the functors between categories, and as 2-morphisms the natural transformations between functors.

C.2 Equivalence of categories

Two categories $\mathcal{C}$ and $\mathcal{D}$ are **equivalent** provided there exist functors $F\colon \mathcal{C} \to \mathcal{D}$ and $G\colon \mathcal{D} \to \mathcal{C}$ such that $GF \cong 1_{\mathcal{C}}$ and $FG \cong 1_{\mathcal{D}}$; that is, the composition GF is naturally isomorphic to the identity functor on $\mathcal{C}$, and similarly for FG.

Given a functor $F\colon \mathcal{C} \to \mathcal{D}$, we have a map on hom sets

$$F\colon \operatorname{Hom}_{\mathcal{C}}(A, B) \to \operatorname{Hom}_{\mathcal{D}}(F(A), F(B)) \quad \text{for all objects } A, B \in \mathcal{C}.$$

We call F

faithful	if F is injective on hom sets.
full	if F is surjective on hom sets.
fully faithful	if it is both full and faithful.

We also say that F is **essentially surjective** if, for all $X \in \mathcal{D}$, there exists $A \in \mathcal{C}$ with $F(A) \cong X$.

A useful theorem in category theory is the following (which assumes the axiom of choice).

Theorem C.1. *A functor $F\colon \mathcal{C} \to \mathcal{D}$ determines an equivalence of categories if and only if it is fully faithful and essentially surjective.*

Proof. Since F is essentially surjective, for each $X \in \mathcal{D}$ we can choose some $G(X) \in \mathcal{C}$ together with an isomorphism $\eta_X\colon X \xrightarrow{\sim} FG(X)$. Next, since F is fully faithful, for each $f\colon X \to Y$ in $\mathcal{D}$ we have $\eta_Y f \eta_X^{-1}\colon FG(X) \to FG(Y)$, and hence there exists a unique $G(f)\colon G(X) \to G(Y)$ in $\mathcal{C}$. It then follows that G is a functor, and that $\eta\colon 1_{\mathcal{D}} \cong FG$ is a natural isomorphism.

Finally, given $A \in \mathcal{C}$, we have $F(A) \in \mathcal{D}$, and hence an isomorphism $\eta_{F(A)}\colon F(A) \xrightarrow{\sim} FGF(A)$. Again, since F is fully faithful, $\eta_{F(A)} = F(\theta_A)$ for some unique $\theta_A\colon A \to GF(A)$. It again follows that $\theta\colon 1_{\mathcal{C}} \cong GF$ is a natural isomorphism. □

C.3 Representable functors

Let $\mathcal{C}$ be a (locally small) category and X an object of $\mathcal{C}$. Then $h_X\colon A \mapsto \operatorname{Hom}(X, A)$ detemines a functor from $\mathcal{C}$ to the category sets **Set**. Note that if $f\colon A \to B$, then

$$h_X(f)\colon \operatorname{Hom}(X, A) \to \operatorname{Hom}(X, B), \quad g \mapsto fg.$$

A functor $F\colon \mathcal{C} \to \mathbf{Set}$ is called **representable** if it is naturally isomorphic to some h^X.

For example, the forgetful functor $\mathbf{Gp} \to \mathbf{Set}$ is represented by the group $\mathbb{Z}$. Similarly, the forgetful functor $\mathbf{Ring} \to \mathbf{Set}$ is represented by the polynomial ring $\mathbb{Z}[X]$.

Lemma C.2 (Yoneda's Lemma)**.** *Let $F\colon \mathcal{C} \to \mathbf{Set}$ be a functor and X an object of $\mathcal{C}$. Then the natural transformations $\operatorname{Hom}(h_X, F)$ are in bijection with the elements of $F(X)$.*

Proof. Let $\eta\colon h_X \Rightarrow F$ be a natural transformation. Since $\eta_X\colon \operatorname{Hom}(X, X) \to F(X)$, we can define a map

$$\operatorname{Hom}(h_X, F) \to F(X), \quad \eta \mapsto \eta_X(1_X).$$

Conversely, let $x \in F(X)$. We define $\eta_A \colon \mathrm{Hom}(X, A) \to F(A)$ by sending a morphism $f \colon X \to A$ to the element $F(f)(x)$. The maps η_A determine a natural transformation $\eta \colon h_X \to F$. For, let $g \colon A \to B$ and consider the diagram

$$\begin{array}{ccc} \mathrm{Hom}(X, A) & \xrightarrow{h_X(g)} & \mathrm{Hom}(X, B) \\ \downarrow{\scriptstyle \eta_A} & & \downarrow{\scriptstyle \eta_B} \\ F(A) & \xrightarrow{F(g)} & F(B). \end{array}$$

Let $f \colon X \to A$. Then $\eta_B h_X(g)$ sends f to the element $\eta_B(gf) = F(gf)(x)$. On the other hand, $F(g)\eta_A$ sends f to the element $F(g)F(f)(x)$. Since F is a functor, it respects composition of morphisms and so these two elements agree. This shows that the diagram commutes, and hence that η is a natural transformation.

Finally we need to show that these two maps are mutually inverse.

Given $x \in F(X)$ we obtain the natural transformation η, which in turn gives us the element $\eta_X(1_X) = F(1_X)(x) = 1_{F(X)}(x) = x$.

Conversely, if we start with a natural transformation η, we obtain the element $x := \eta_X(1_X)$, which in turn gives us the natural transformation η' such that

$$\eta'_A(f) := F(f)(x) = F(f)\eta_X(1_X) = \eta_A h_X(f)(1_X) = \eta_A(f1_X) = \eta_A(f).$$

Hence $\eta' = \eta$, finishing the proof. □

As a special case, we have that

$$\mathrm{Hom}(h_X, h_Y) \cong \mathrm{Hom}(Y, X),$$

so the natural transformations between the representable functors h_X and h_Y are in bijection with the morphisms between Y and X. Note the change of order, from $h_X \Rightarrow h_Y$ to $Y \to X$. Thus h_* determines a functor $\mathcal{C}^{\mathrm{op}} \to \mathrm{Fun}(\mathcal{C}, \mathbf{Set})$.

C.4 Contravariant functors

Up to now we have only considered **covariant** functors. A **contravariant** functor $F \colon \mathcal{C} \to \mathcal{D}$ is the same as a covariant functor $F \colon \mathcal{C}^{\mathrm{op}} \to \mathcal{D}$.

For example, for a (locally small) category $\mathcal{C}$ and an object X, the functor $h^X := \mathrm{Hom}(-, X)$ is a contravariant representable functor from $\mathcal{C}$ to the category of sets, and we have a covariant functor $h^* \colon \mathcal{C} \to \mathrm{Fun}(\mathcal{C}^{\mathrm{op}}, \mathrm{Set})$.

A **duality** between two categories $\mathcal{C}$ and $\mathcal{D}$ is a contravariant equivalence, so an equivalence $\mathcal{C}^{\mathrm{op}} \cong \mathcal{D}$, or more generally any contravariant functor yielding such an equivalence.

An important example is given by vector space duality. Consider the category $\mathbf{vec}_K$ of finite dimensional vector spaces over a fixed field K. Then duality $D := \mathrm{Hom}_K(-, K)$ yields a functor $\mathbf{vec}_K \to \mathbf{vec}_K^{\mathrm{op}}$, and the evaluation map ev yields a natural isomorphism $\mathrm{id} \cong D^2$.

As usual, the map $\mathrm{ev} \colon V \to D^2(V)$ sends a vector v to the linear map $D(V) \to K$, $\theta \mapsto \theta(v)$. To see that this is a natural transformation, one thus needs to check that if $f \colon V \to W$, then $D^2(f)(\mathrm{ev}(v)) = \mathrm{ev}(f(v))$.

C.5 Universal properties

There are close connections between initial/terminal objects, representable functors, and universal properties.

Let $F\colon \mathcal{C} \to \mathrm{Set}$ be a functor, and (X, x) a pair consisting of an object $X \in \mathcal{C}$ and an element $x \in F(X)$. We say that (X, x) satisfies the **universal property** for F provided that, for all objects $Y \in \mathcal{C}$, and all elements $y \in F(Y)$, there exists a unique morphism $f\colon X \to Y$ with $F(f)(x) = y$.

Starting from F we can also form the **coslice category** $\mathcal{C} \downarrow F$. This has objects the pairs (X, x), where $X \in \mathcal{C}$ and $x \in F(X)$, and morphisms $f\colon (X, x) \to (Y, y)$ given by those morphisms $f\colon X \to Y$ such that $F(f)(x) = y$.

Finally, by Yoneda's Lemma, each pair (X, x) determines a natural transformation $\eta_{X,x}\colon \mathrm{Hom}(X, -) \to F$.

Lemma C.3. *The following are equivalent for a functor $F\colon \mathcal{C} \to \mathrm{Set}$.*

(1) There exists some (X, x) satisfying the universal property for F.

(2) The coslice category $\mathcal{C} \downarrow F$ has an initial object (X, x).

(3) The natural transformation $\eta_{X,x}$ is a natural isomorphism.

In this case we see that the object (X, x) is unique up to unique isomorphism in the coslice category.

Proof. It is clear that (1) and (2) are equivalent. For (3) we need to describe when $\eta_{X,x}$ yields an isomorphism $\mathrm{Hom}(X, A) \to F(A)$ for all objects $A \in \mathcal{C}$. This happens precisely when, for each A and $a \in F(A)$, there exists a unique $f\colon X \to A$ such that $F(f)(x) = a$, which is thus equivalent to (1).

Finally, we have already seen that initial objects are unique up to unique isomorphism. □

We can now rephrase several of our constructions in terms of universal properties or representable functors.

The coproduct of objects A_i represents the functor $X \mapsto \prod_i \mathrm{Hom}(A_i, X)$. Dually, the product represents the contravariant functor $X \mapsto \prod_i \mathrm{Hom}(X, A_i)$.

The cokernel of $f\colon A \to B$ represents the functor $F(X) := \{x\colon B \to X : xf = 0\}$. Dually, the kernel represents the contravariant functor $X \mapsto \{x\colon X \to A : fx = 0\}$.

Topology

D Topological spaces

Let X be a non-empty set. A **topology** on X is defined by specifying a set τ of subsets of X, whose elements are called the **open sets** of the topology, satisfying the following axioms:

Top1 both X and $\emptyset$ are open.
Top2 unions of open sets are open.
Top3 a finite intersection of open sets is open.

The complement of an open set is called a **closed set**. We can therefore define a topology on X by specifying the collection of closed sets, satisfying the dual axioms:

Top1′ both X and $\emptyset$ are closed.
Top2′ intersections of closed sets are closed.
Top3′ a finite union of closed sets is closed.

Given an arbitrary subset U of X, we define its **closure** $\bar{U}$ to be the smallest closed set containing U, and its **interior** U^o to be the largest open set contained in U. We call U **dense** if $\bar{U} = X$, and **nowhere dense** if $U^o = \emptyset$.

Let σ and τ be two topologies on X. Then τ is **finer** than σ (or σ is **coarser** than τ) provided that $\sigma \subset \tau$; in other words every set which is open in σ is also open in τ.

Given topologies τ_i on X, we can define a topology $\tau := \bigcap_i \tau_i$ by taking as open sets only those subsets of X which are open in every τ_i. Thus τ is coarser than every τ_i. In particular, given any set τ' of subsets of X, there is a unique smallest topology on X containing τ', given as the intersection over all topologies containing τ'. This is called the topology **generated** by τ'.

Example. The **trivial topology** on X is given by $\tau := \{\emptyset, X\}$. At the other extreme, we have the discrete topology, where every subset of X is open (and hence also closed).

Two somewhat more interesting topologies are given by the **cofinite topology**, where the closed sets are the finite sets together with X itself, and the **cocountable topology**, where the closed sets are the countable sets together with X itself.

A **base** for a topology τ on X is given by a subset $\mathcal{B} \subset \tau$ such that every open set is a union of open sets in $\mathcal{B}$. Sufficient (but not necessary) conditions for a collection $\mathcal{B}$ to form the base of a topology are that

(1) it **covers** X, so that $X = \bigcup_{B \in \mathcal{B}} B$.

(2) it is closed under finite intersections, so that $B_1 \cap B_2 \in \mathcal{B}$ for all $B_i \in \mathcal{B}$.

If the set $\mathcal{B}$ satisfies these conditions, then it is a base of the topology it generates.

Example. A metric space has a canonical topology having as base the open balls $B_r(x) := \{y : d(x, y) < r\}$. These are not closed under finite intersections, though.

Let Y be a non-empty subset of a topological space (X, τ). The **subspace topology** on Y has open sets the intersections $U \cap Y$ for $U \in \tau$.

Let $\sim$ be an equivalence relation on a toplogical space (X, τ). Then the set of equivalence classes $X/\sim$ can be given the **quotient topology**, which is the smallest topology making the canonical map $\pi\colon X \to X/\sim$ continuous. Thus a subset V of $X/\sim$ is open precisely when $\pi^{-1}(V)$ is open in X.

D.1 Continuous maps

The appropriate maps to consider in topology are the continuous ones.

A **continuous map** $f\colon (X, \tau) \to (Y, \sigma)$ between topological spaces is given by a map $f\colon X \to Y$ such that $f^{-1}(\sigma) \subset \tau$.

A **homeomorphism** is a bijective continuous map whose inverse is also continuous.

We can now form the category having objects the topological spaces and morphisms the continuous maps. Observe that a map is a homeomorphism if and only if it is an isomorphism in the category of topological spaces.

A continuous map is **open** if it sends open sets to open sets, and is **closed** if it sends closed sets to closed sets. Clearly a continuous bijection is a homeomorphism if and only if it is open, equivalently closed.

D.2 Separation axioms

There are many levels of the notion of **separation** that can be considered in topology. The first few are given below. Let (X, τ) be a topological space and let $x \neq y$ be elements of X.

- **T0** There exists an open set U such that either $x \in U$ and $y \notin U$, or else $y \in U$ and $x \notin U$.
- **T1** There exist open sets U, V such that $x \in U$ and $y \notin U$, and also $y \in V$ and $x \notin V$.
- **T2** There exist disjoint open sets U, V with $x \in U$ and $y \in V$.
- **T2.5** There exist open sets U, V with $\bar{U} \cap \bar{V} = \emptyset$ with $x \in U$ and $y \in V$.

These notions form a strict heirarchy, so that we have strict implications

$$\text{T2.5} \Rightarrow \text{T2} \Rightarrow \text{T1} \Rightarrow \text{T0}.$$

A T0 space is also called **Kolmogorov**; a T1 space **Fréchet**; and a T2 space is usually called **Hausdorff**.

The axiom T2 is often assumed in topology, hence the frequent use of the term Hausdorff space, since this axiom implies the uniqueness of limits of sequences. For example, it is satisfied by all metric spaces.

D.3 Irreducibility and locally-closed subsets

A topological space (X, τ) is **irreducible** if, whenever $X = V_1 \cup V_2$ with both V_i closed, then $X = V_i$ for some i. Alternatively, if U_i are non-empty and open, then $U_1 \cap U_2$ is again non-empty. A third characterisation says that every non-empty open subset of X is dense.

A non-empty subset Y of X is said to be **locally closed** if it is the intersection in X of a closed set and an open set. The term comes from the convention that 'local' properties are those that happen in a **neighbourhood**, so an open set containing the point or set in question. From this perspective, if $Y = U \cap V$ with U open and V closed, then U is a neighbourhood of Y and, in the induced topology, Y is closed in U.

A **constructible** set is a finite union of locally closed sets.

We say that X is **Noetherian** if it satisfies the ascending chain condition on open sets: every chain of open subsets of X

$$U_1 \subset U_2 \subset U_3 \subset \cdots$$

is stationary, so there exists N with $U_n = U_N$ for all $n \geq N$.

D.4 Filters and quasi-compact spaces

A **filter** $\mathcal{F}$ on X is a non-empty collection of subsets of X such that

Filt1 $X \in \mathcal{F}$ but $\emptyset \notin \mathcal{F}$.
Filt2 $U, V \in \mathcal{F}$ implies $U \cap V \in \mathcal{F}$.
Filt3 $U \in \mathcal{F}$ and $U \subset V$ implies $V \in \mathcal{F}$.

For filters $\mathcal{F}$ and $\mathcal{G}$ on X we say that $\mathcal{F} \subset \mathcal{G}$ provided $U \in \mathcal{F}$ implies $U \in \mathcal{G}$. An **ultrafilter** is a filter $\mathcal{F}$ such that for every $U \subset X$ we have either $U \in \mathcal{F}$ or $U^c \in \mathcal{F}$.

Lemma D.1. *Let $\mathcal{U}$ be a non-empty collection of subsets of X such that all finite intersections are non-empty. Then the collection*

$$\{V \subset X : V \text{ contains some } U_1 \cap \cdots \cap U_n,\ U_i \in \mathcal{U}\}$$

defines a filter $\langle \mathcal{U} \rangle$. □

Lemma D.2. *Every filter is contained inside an ultrafilter, and these are precisely the maximal filters.*

Proof. Suppose we have a chain of filters $\mathcal{F}_i$, and let $\mathcal{F}$ be their union. Thus $U \in \mathcal{F}$ if and only if $U \in \mathcal{F}_i$ for some i. Then $\mathcal{F}$ is again a filter. By Zorn's Lemma we thus see that maximal filters exist, and every filter is contained in a maximal filter. If $\mathcal{F}$ is an ultrafilter, then it is clearly maximal. For, if $U \notin \mathcal{F}$, then U^c is in $\mathcal{F}$, and no filter can contain both U and U^c. Conversely, let $\mathcal{F}$ be maximal. Suppose $U, U^c \notin \mathcal{F}$. If $V \in \mathcal{F}$ satisfies $V \subset U^c$, then $U^c \in \mathcal{F}$, a contradiction. Hence $V \cap U \neq \emptyset$. It follows that the collection $\mathcal{U} := \{V \in \mathcal{F}\} \cup \{U\}$ determines a filter $\langle \mathcal{U} \rangle$ strictly containing $\mathcal{F}$, a contradiction. Thus $\mathcal{F}$ must be an ultrafilter. □

Lemma D.3. *Let $\mathcal{F}$ be an ultrafilter. Then $U_1 \cup \cdots \cup U_n \in \mathcal{F}$ if and only if $U_i \in \mathcal{F}$ for some i.*

Proof. Suppse $U, V \notin \mathcal{F}$. Then $U^c, V^c \in \mathcal{F}$, so $(U \cup V)^c = U^c \cap V^c \in \mathcal{F}$, and hence $U \cup V \notin \mathcal{F}$. The proof now follows by induction. □

We say that x is a **limit point** of an ultrafilter $\mathcal{F}$ if $\mathcal{F}$ contains every open neighbourhood of x.

We call a space X **quasi-compact** provided every open cover contains a finite subcover.[23]

Lemma D.4. *A space is quasi-compact if and only if every ultrafilter has a limit point.*

Proof. Let X be quasi-compact, and suppose that $\mathcal{F}$ is an ultrafilter without any limit points. Then for each $x \in X$ there exists some open neighbourhood $x \in U_x$ with $U_x^c \in \mathcal{F}$. Since X is quasi-compact, the open cover $X = \bigcup_x U_x$ has a finite subcover, say $U_1 \cup \cdots \cup U_n$. Now, if $V \in \mathcal{F}$ we have $V = (V \cap U_1) \cup \cdots (V \cap U_n)$, so some $V \cap U_i \in \mathcal{F}$, and hence $U_i \in \mathcal{F}$, a contradiction.

Conversely, suppose X is not quasi-compact, so there is an open cover $X = \bigcup_i U_i$ having no finite subcover. Set $\mathcal{V} := \{U_i^c\}$. Then all finite intersections are non-empty, so we get a filter $\langle \mathcal{V} \rangle$, which we can then extend to an ultrafilter $\mathcal{F}$. If $x \in X$, then $x \in U_i$ for some i, and $U_i \notin \mathcal{F}$. Thus $\mathcal{F}$ has no limit points. □

D.5 Separated and complete spaces

We will fix a subcategory of topological spaces admitting all finite products. For example, we could take all topological spaces, in which case the product $X \times Y$ has the usual product topology. Our main interest, however, will be in varieties with the Zariski topology. Then finite products also exist, but the topology on $X \times Y$ is finer than the product of the Zariski topologies.

We call a topological space X

separated if the diagonal $\Delta_X := \{(x, x) : x \in X\}$ is closed in $X \times X$.
complete if the projection map $X \times Y \to Y$ is closed for all Y.

We begin by studying separated spaces. Classically these are just the Hausdorff spaces.

Lemma D.5. *Suppose $X \times X$ has the product topology. Then X is separated if and only if it is Hausdorff.*

Proof. Let $x, y \in X$ be distinct, so $(x, y) \notin \Delta_X$. Then X is separated if and only if (x, y) has an open neighbourhood $U \times V$ not intersecting Δ_X, if and only if U and V are open neighbourhoods of x and y respectively such that $U \cap V = \emptyset$, which is if and only if X is Hausdorff. □

Let $f\colon X \to Z$ and $g\colon Y \to Z$ be continuous. Then the maps $f\pi_1, g\pi_2\colon X \times Y \to Z$ yield a map $(f, g)\colon X \times Y \to Z \times Z$, and so we can form the **fibred product**

$$X \times_Z Y := \{(x, y) : f(x) = g(y)\} \subset X \times Y$$

as the preimage of the diagonal Δ_Z.

[23] We reserve the term **compact** for quasi-compact Hausdorff spaces. The reason for doing this is that many of the nice properties for compact spaces use that the space is Hausdorff, and also that all finite products have the usual product topology.

For example, if $f\colon X \to Z$ is continuous and Z is separated, then the **graph** of f is closed in $X \times Z$

$$\Gamma_f := \{(x, f(x)) : x \in X\} = X \times_Z Z \subset X \times Z.$$

Similarly, if $f, g\colon X \to Z$, then their **equaliser** is closed in X

$$\mathrm{Eq}\,(f, g) := \{x : f(x) = g(x)\} = \Delta^{-1}(X \times_Z X) \subset X.$$

We now turn to complete spaces. Classically, these are just the quasi-compact spaces.

Theorem D.6. *Suppose $X \times Y$ has the product topology for all Y. Then X is quasi-compact if and only if it is complete.*

Proof. Suppose $V \subset X \times Y$ is closed, and take $y \notin \pi(V)$. For each $x \in X$ we know $(x, y) \notin V$, so there exists some open neighbourhood $U'_x \times U''_x$ not intersecting V. Now $X = \bigcup_x U'_x$, so we have a finite subcover, say $U'_1 \cup \cdots \cup U'_n$. Set $U'' := U''_1 \cap \cdots \cap U''_2$, an open neighbourhood of y in Y. Fix $y' \in U''$ and consider some (x, y'). Then $x \in U'_i$ for some i, so $(x, y') \in U'_i \times U''_i$ and hence $(x, y') \notin V$. This holds for all $x \in X$, so $y' \notin \pi(V)$. Thus $\pi(V)$ is closed.

Conversely, let $\mathcal{F}$ be an ultrafilter on X. Define $Y := X \sqcup \{\infty\}$ with the topology generated by the open sets $\{x\}$ for $x \in X$ and $U \cup \{\infty\}$ for $U \in \mathcal{F}$. Set $\Delta := \{(x, x)\} \subset X \times Y$. Then $\pi(\Delta) = X \subset Y$ is not closed, since its complement $\{\infty\}$ is not open. Since the projection map is closed we have $\pi(\bar{\Delta}) = \overline{\pi(\Delta)} = Y$, so $\bar{\Delta}$ contains some (x, ∞). Let U be an open neighbourhood of x, and $V \in \mathcal{F}$. Then $U \times (V \cup \{\infty\})$ is an open neighbourhood of (x, ∞), so intersects Δ non-trivially, and hence $U \cap V$ is non-empty. In particular, $U^c \notin \mathcal{F}$, so $U \in \mathcal{F}$ and hence x is a limit point of $\mathcal{F}$. □

We say that $f\colon X \to Y$ is **universally closed** if the projection $X \times_Y Z \to Z$ is closed for all $Z \to Y$. Note that putting $Z = Y$ shows that $f\colon X \to Y$ is closed.

Proposition D.7. *Let $f\colon X \to Y$ be continuous, where X is complete and Y is separated. Then f is universally closed.*

Proof. Take $g\colon Z \to Y$. Then $X \times_Y Z \subset X \times Z$ is closed, since it is the preimage of Δ_Y under the map $(f, g)\colon X \times Z \to Y \times Y$. So every closed subset of $X \times_Y Z$ is closed in $X \times Z$, and hence its image is closed in Z. □