MAS 439
COMMUTATIVE ALGEBRA AND ALGEBRAIC GEOMETRY

PAUL JOHNSON
MINOR ADAPTATION OF NOTES BY TOM BRIDGELAND

Contents

Date: September 5, 2017.

1. Introduction

Commutative algebra is the study of comuutative rings. These are sets R having two composition laws (called addition and multiplication) which behave in the same way as addition and multiplication of integers. They form perhaps the third most important class of algebraic objects in the whole of maths (after groups and vector spaces). Commutative rings come up all over pure mathematics. Here are two examples:

- *Number theory.* The Gaussian integers

$$\mathbb{Z}[i] = \{a + bi : a, b \in \mathbb{Z}\} \subset \mathbb{C}$$

form a commutative ring under the usual operations of addition and multiplication of complex numbers. Similarly, if we consider the third root of unity

$$\omega = \exp(2\pi i/3) = (-1 + \sqrt{3}\, i)/2$$

then the subset

$$\mathbb{Z}[\omega] = \{a + b\omega : a, b \in \mathbb{Z}\} \subset \mathbb{C}$$

form a commutative ring. To see this note the relations

$$\omega^3 = 1, \quad \omega^2 + \omega + 1 = 0.$$

Rings of algebraic integers like these are useful for studying Diohantine equations. For example, to prove the first case $n = 3$ of Fermat's last theorem one writes

$$z^3 = x^3 + y^3 = (x + y)(x + \omega y)(x + \omega^2 y)$$

and considers prime factorizations of both sides in the ring $\mathbb{Z}[\omega]$.[1]

- *Geometry.* We can study a topological space X (for example a subset $X \subset \mathbb{R}^n$) via its ring of functions. This is the set of all continuous functions $f\colon X \to \mathbb{R}$. These can be added and multiplied pointwise:

$$(f + g)(x) = f(x) + g(x), \quad (f \cdot g)(x) = f(x) \cdot g(x).$$

Similarly we can consider rings of differentiable or analytic functions.

Algebraic geometry is the study of systems of polynomial equations. The set of solutions to such a system is called an affine variety. We first choose a field k we wish to work over (e.g. $k = \mathbb{Q}$, $\mathbb{R}$ or $\mathbb{C}$). Given a collection of polynomials $f_1, f_2, \cdots, f_r$ in n variables $x_1, \cdots, x_n$ with coefficients in k, the corresponding affine variety is the subset

$$V(f_1, \cdots, f_r) \subset k^n$$

consisting of those points $(a_1, \cdots, a_n) \in k^n$ satisfying

$$f_1(a_1, \cdots, a_n) = f_2(a_1, \cdots, a_n) = \cdots = f_r(a_1, \cdots, a_n) = 0.$$

[1] See Hardy and Wright, An Introduction to the Theory of Numbers, Chapters 12–13.

Perhaps surprisingly, algebraic geometry occupies a central place in modern pure mathematics. The basic reason is that the study of polynomial equations has many different aspects, which relate to lots of other areas of pure mathematics. Wegive a few examples of this:

- *Number theory.* If we take the field $k = \mathbb{Q}$ then polynomial equations become *Diophantine equations.* For example Fermat's last theorem is the statement that if $n > 2$ the variety
$$x^n + y^n = 1$$
has no points over the field $k = \mathbb{Q}$ with x, y nonzero.
- *Geometry and topology.* Working over $\mathbb{C}$ we can study the geometry and topology of varieties. For example, consider the variety
$$y^2 = x^3 - x.$$
This is an example of an *elliptic curve.* Topologically it is a torus (doughnut) with 3 points removed. To get the full torus we must work with the corresponding projective variety[2].
- *Physics.* In string theory the world is supposed to have 10 dimensions. 4 of these dimensions form spacetime, the other 6 are supposed to be curled up very small, and account for the properties of the fundamental forces. For string theory to work properly these curled up dimensions must have a special shape: they should be *Calabi-Yau threefolds*[3]. Examples of Calabi-Yau threefolds are most easily described using polynomial equations. For example, again working over $\mathbb{C}$, the variety
$$x_1^5 + x_2^5 + x_3^5 + x_4^5 + x_1x_2x_3x_4 = 0$$
is a Calabi-Yau threefold known as the *quintic threefold.* Note that it has $4 - 1 = 3$ complex dimensions, and hence 6 real dimensions.
- *Algebra.* The basic calculational tool in algebraic geometry is commutative algebra. Every affine variety X has an associated *co-ordinate ring*
$$k[V] = k[x_1, \cdots, x_n]/(f_1, \cdots, f_r),$$
which is a commutative ring. It is obtained by quotienting the polynomial ring $k[x_1, \cdots, x_n]$ by the ideal generated by the polynomials defining V.

This course will focus on the relationship between commutative algebra and algebraic geometry . We aim to introduce the basic commutative algebra needed to study affine varieties, and to build up an intuition for the relationship between the geometrical properties of an affine variety V and the algebraic properties of the associated co-ordinate ring $k[V]$.

[2] See Frances Kirwan, 'Complex algebraic curves', Chapter 5.

[3] See Brian Greene, 'The Elegant Universe'.

2. Definition of a ring

The definition of a ring is an abstraction of the properties of addition and multiplication of integers.

Definition 2.1. A *ring* is a set R equipped with two binary operations,

$$+\colon R\times R\to R,\quad \cdot\colon R\times R\to R;$$

called addition and multiplication, satisfying the following conditions:

(a) addition $(a,b)\mapsto a+b$ makes R into an abelian group;

(b) multiplication $(a,b)\mapsto a\cdot b$ makes R into a monoid;

(c) these two operations are related by the *distributive law*:

$$(a+b)\cdot c = a\cdot c+b\cdot c,\quad a\cdot(b+c)=a\cdot b+a\cdot c.$$

We now explain the first two conditions in more detail.

Condition (a) means that

(i) addition is associative and commutative, that is

$$a+(b+c)=(a+b)+c,\quad a+b=b+a;$$

(ii) there is a *zero element* $0_R\in R$ such that for all $a\in R$,

$$0_R+a=a=a+0_R;$$

(iii) for every element $a\in R$ there is an additive inverse $-a\in R$ such that

$$a+(-a)=0_R=(-a)+a.$$

Condition (b) means that

(i) multiplication is associative

$$a\cdot(b\cdot c)=(a\cdot b)\cdot c;$$

(ii) there is an *identity element* $1_R\in R$ such that for all $a\in R$

$$1_R\cdot a=a=a\cdot 1_R.$$

There are many basic examples of rings. It is important to become familiar with a good number of these.

Examples 2.2. (a) The set of integers $\mathbb{Z}$ forms a ring when equipped with the usual operations of addition and multiplication.

(b) Let $n>0$ be a positive integer. Let $\mathbb{Z}/n$ denote the set of integers modulo n. More precisely, the elements of $\mathbb{Z}/n$ are equivalence classes of integers, under the equivalence relation

$$a\sim b \iff n\mid(b-a),\quad a,b\in\mathbb{Z}.$$

The equivalence classes are called *residue classes* modulo n; we denote the equivalence class containing $a \in \mathbb{Z}$ by the symbol $[a]$. Thus

$$\mathbb{Z}/n = \{[0], [1], \cdots, [n-1]\}.$$

The set $\mathbb{Z}/n$ forms a commutative ring under the operations induced by the usual operations of addition and multiplication in $\mathbb{Z}$. Thus to add or multiply two residue classes in $\mathbb{Z}/n$ we choose integer representatives in the two classes, add or multiply these, and then consider the corresponding residue class modulo n. For example in $\mathbb{Z}/10$ we have

$$[5] + [7] = [12] = [2], \quad [5] \cdot [7] = [35] = [5].$$

The zero and identity elements are $0_{\mathbb{Z}/n} = [0]$ and $1_{\mathbb{Z}/n} = [1]$ respectively.

(c) Let X be a set and write

$$R = \mathrm{Fun}(X, \mathbb{R})$$

for the set of all functions $X \to \mathbb{R}$. This set forms a ring under pointwise operations:

$$(f+g)(t) = f(t) + g(t); \quad (f \cdot g)(t) = f(t) \cdot g(t).$$

The zero and identity elements are the constant functions defined by

$$0_R(x) = 0, \quad 1_R(x) = 1$$

for all points $x \in X$.

(d) Any set with a single element is a ring in a unique way: there is only one possible way to define the addition and multiplication operations. A ring with a single element is called *trivial.* Note that trivial rings are not all equal: e.g. the set

$$R = \{n \in \mathbb{Z} : 16 < n < 18\}$$

is a trivial ring, with addition and multiplication $17 + 17 = 17 = 17 \cdot 17$, and similarly the set

$$S = \{\text{David Cameron}\}$$

of current UK prime ministers forms a trivial ring. But these rings are definitely not the same. What is true is that all trivial rings are isomorphic (see Section 4).

Non-examples 2.3. (a) The non-negative integers $\mathbb{Z}_{\geqslant 0}$ do not form a ring under the usual operations, since there are no additive inverses.

(b) The even integers $2\mathbb{Z}$ do not form a ring under the usual operations, since there is no identity element.

(c) The integers $\mathbb{Z}$ equipped with the usual multiplication, but with addition

$$(a, b) \mapsto a + b + 1$$

is not a ring, since the distributive law fails.

Remarks 2.4. (a) The usual argument shows that the elements 0_R and 1_R are unique with the given properties. If we make other choices $0'_R$ and $1'_R$ the axioms imply

$$0'_R = 0'_R + 0_R = 0_R, \quad 1'_R = 1'_R \cdot 1_R = 1_R.$$

(b) The distributive law implies that for any element $a \in R$

$$0_R \cdot a + 0_R \cdot a = (0_R + 0_R) \cdot a = 0_R \cdot a.$$

Adding $-(0_R \cdot a)$ to both sides we conclude that $0_R \cdot a = 0_R$.

(c) Given an element $a \in R$ the distributive law implies that

$$(-1_R) \cdot a + 1_R \cdot a = (1_R + (-1_R)) \cdot a = 0_R \cdot a = 0_R.$$

It follows that $(-1_R) \cdot a = -a$.

(c) Another consequence of the distributive law is the following identity:

$$(a_1 + \cdots + a_r) \cdot (b_1 + \cdots + b_s) = a_1 \cdot b_1 + a_1 \cdot b_2 + \cdots + a_r \cdot b_s,$$

or in summation notation

$$\sum_{i=1}^{r} a_i \cdot \sum_{j=1}^{s} b_j = \sum_{i=1}^{r} \sum_{j=1}^{s} a_i \cdot b_j.$$

To prove this, first verify the identity for $r = 1$ using induction on s, and for $s = 1$ using induction on r. Finally, prove the general statement using induction on either r or s. The details are left as a good exercise for the reader.

(d) In a trivial ring we have the relation $0_R = 1_R$. Conversely, if this relation is satisfied in a ring R, then for all elements $a \in R$

$$a = 1_R \cdot a = 0_R \cdot a = 0_R.$$

Thus R consists of a single element and is therefore a trivial ring.

We often just write 0 for the zero element 0_R, and similarly 1 for the unit element 1_R. We use the notation

$$a - b := a + (-b).$$

We also often write ab instead of $a \cdot b$.

Definition 2.5. A ring R is *commutative* if the condition

$$a \cdot b = b \cdot a$$

holds for all elements $a, b \in R$.

Unless otherwise specified, *all rings will be assumed to be commutative.*

3. Further examples and basic properties

We give some more examples of rings.

Examples 3.1. (a) Let R be a commutative ring; we could take $R = \mathbb{R}$ for example. The set of polynomials in one variable with coefficients in R forms a commutative ring $R[x]$. The elements are finite sums of the form

$$f(x) = \sum_{i=0}^{d} a_i x^i = a_0 + a_1 x + a_2 x^2 + \cdots + a_d x^d$$

with $a_0, a_1, \cdots a_d \in R$. If $a_d \neq 0$ we call the number $d \geqslant 0$ the degree of the polynomial $f(x)$ (we consider the zero polynomial to have degree 0). Polynomials of degree 0 are called *constant polynomials.* The addition and multiplication laws are

$$\big(\sum_{i=0}^{n} a_i x^i\big) + \big(\sum_{i=0}^{n} b_i x^i\big) = \sum_{i=0}^{n} (a_i + b_i) x^i.$$

$$\big(\sum_{i=0}^{m} a_i x^i\big) \cdot \big(\sum_{i=0}^{n} b_j x^j\big) = \sum_{k=0}^{mn} c_k x^k \text{ where } c_k = \sum_{i=0}^{k} a_i \cdot b_{k-i}.$$

The zero element is the zero polynomial 0_R. The identity is the constant polynomial 1_R.

(b) The set $\mathrm{M}_{2\times 2}(\mathbb{R})$ of real 2×2 real matrices

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad a, b, c, d \in \mathbb{R},$$

forms a non-commutative ring under the usual operations of matrix addition and multiplication

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} + \begin{pmatrix} e & f \\ g & h \end{pmatrix} = \begin{pmatrix} a+e & b+f \\ c+g & d+h \end{pmatrix}.$$

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} + \begin{pmatrix} e & f \\ g & h \end{pmatrix} = \begin{pmatrix} ae+bg & af+bh \\ ce+dg & cf+dh \end{pmatrix}.$$

The zero and unit are

$$0 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad 1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

More generally, for any ring R, and any integer $n > 0$, there is a ring $\mathrm{M}_{n\times n}(R)$ of $n \times n$ matrices with entries in R.

Let R be a commutative ring and $a \in R$ some element. For any $n > 0$ we use the shorthands

$$n \cdot a = \overbrace{a + a + \cdots + a}^{n}, \qquad a^n = \overbrace{a \cdot a \cdot \cdots \cdot a}^{n}.$$

Of course we then have relations

$$(m+n)\cdot a = m\cdot a + n\cdot a, \quad a^m \cdot a^n = a^{m+n}$$

for all $m, n > 0$. We can extend this to $m, n \geqslant 0$ by defining

$$0\cdot a = 0_R, \quad a^0 = 1_R.$$

Note that it is possible for a nonzero element $a \in R$ to satisfy $n \cdot a = 0_R$ for $n > 0$. This relation holds for example for any element $a \in \mathbb{Z}/n$.

Lemma 3.2 (Binomial expansion)**.** *Let R be a commutative ring and take elements $a, b \in R$. For any $n > 0$ there is a relation*

$$(a+b)^n = \sum_{i=0}^{n} \binom{n}{i} \cdot a^i \cdot b^{n-i}.$$

Proof. For $0 < i < n$ the binomial coefficients satisfy

$$\binom{n}{i} = \binom{n-1}{i-1} + \binom{n-1}{i}.$$

The result then follows from the distributive law and induction on n. □

Note that the commutativity assumption is essential for this result. In a non-commutative ring one has

$$(a+b)^2 = a^2 + a\cdot b + b\cdot a + b^2$$

but one cannot simplify any further.

Definition 3.3. Let R be a commutative ring. An element $a \in R$ is called

(a) a *unit* if it has a multiplicative inverse, that is if there exists another element $a^{-1} \in R$ satisfying

$$a \cdot a^{-1} = 1_R = a^{-1} \cdot a.$$

(b) a *zero divisor* if $a \neq 0_R$ but there is a $0_R \neq b \in R$ with $a \cdot b = 0_R$;

(c) *nilpotent* if $a \neq 0_R$ but $a^n = 0_R$ for some $n > 1$.

Examples 3.4. (a) An element $f \in \mathrm{Fun}(X, \mathbb{R})$ is a unit precisely if $f(x) \neq 0$ for all $x \in X$.

(b) An element $f \in \mathrm{Fun}(X, \mathbb{R})$ is a zero-divisor precisely if $f(x) = 0$ for some $x \in X$.

(c) The element $[2] \in \mathbb{Z}/8\mathbb{Z}$ is nilpotent.

Definition 3.5. A non-trivial commutative ring R is said to be

(a) a *field* if every nonzero element is a unit;

(b) an *integral domain* if it has no zero divisors;

(c) *reduced* if it is has no nilpotent elements.

It is the convention that trivial rings are not considered to be fields or integral domains, but they are considered to be reduced.

Lemma 3.6. *There are implications*

$$R \text{ a field} \implies R \text{ an integral domain} \implies R \text{ reduced.}$$

Proof. For the first implication let us assume that R is a field, and suppose for a contradiction that $0_R \neq a \in R$ is a zero divisor. Thus there is an element $0_R \neq b \in R$ with $a \cdot b = 0_R$. Since R is a field there is an element $a^{-1} \in R$ with $a^{-1} \cdot a = 1_R$. Multiplying both sides of the relation $a \cdot b = 0_R$ on the left by a^{-1} then gives $b = 0_R$, a contradiction. Hence R is an integral domain.

For the second implication note that if $a \in R$ is a nilpotent element, we can find a minimal integer $n > 1$ such that $a^n = 0_R$. Then $a \neq 0_R$ and $a^{n-1} \neq 0_R$ but $a \cdot a^{n-1} = 0_R$. Thus a is a zero divisor. So if a ring contains no zero divisors, it contains no nilpotent elements either. □

Examples 3.7. (a) The set of rational numbers

$$\mathbb{Q} = \big\{ a/b : a, b \in \mathbb{Z} \text{ with } b \neq 0 \big\}$$

forms a field under the usual operations of addition and multiplication. Similarly, the real numbers $\mathbb{R}$, and the complex numbers

$$\mathbb{C} = \{a + bi : a, b \in \mathbb{R}\},$$

become fields when equipped with the usual operations.

(b) The integers $\mathbb{Z}$ are an integral domain but not a field.

(c) Suppose that X is a set with more than one element. Then the ring $\mathrm{Fun}(X, \mathbb{R})$ is reduced, but not an integral domain.

Lemma 3.8. *Let $n > 0$ be a positive integer. The residue class ring $\mathbb{Z}/n$ is a field if and only if it is an integral domain. This happens precisely when n is prime.*

Proof. Consider first the case when n is composite, and write $n = ab$ with $1 < a, b < n$. Then

$$[0] \neq [a] \in \mathbb{Z}/n \text{ and } [0] \neq [b] \in \mathbb{Z}/n \text{ but } [a] \cdot [b] = [0] \in \mathbb{Z}/n.$$

Hence $\mathbb{Z}/n$ is not an integral domain, and also therefore not a field.

Now consider the case when $n = p$ is prime. We claim that $\mathbb{Z}/p$ is a field, and hence also an integral domain. Take a nonzero element $[0] \neq [a] \in \mathbb{Z}/p$. We must show that there is a $[b] \in \mathbb{Z}/p$ such that $[a] \cdot [b] = [1]$ in $\mathbb{Z}/p$. The class $[a]$ is represented by an integer $a \in \mathbb{Z}$, and the assumption that $[a] \neq [0] \in \mathbb{Z}/p$ means that a is coprime to p. Using the Euclidean algorithm we can therefore find integers $b, c \in \mathbb{Z}$ such that

$$a \cdot b + p \cdot c = 1.$$

This implies that $[a] \cdot [b] = [1]$ in the ring $\mathbb{Z}/p$. □

Example 3.9. Let $n > 0$ be a positive integer. The ring $\mathbb{Z}/n$ is reduced presisely if n is square-free. This means that either $n = 1$ or $n = p_1 \cdot p_2 \cdot \cdots \cdot p_k$ can be written as a product of distinct primes. For example, $12 = 2^2 \cdot 3$ is not square-free, and the ring $\mathbb{Z}/12$ is not reduced, since the element $[6] \in \mathbb{Z}/12$ is nilpotent.

4. Homomorphisms of rings

We now consider maps of rings which preserve the relevant structure.

Definition 4.1. Let R and S be rings. A *homomorphism* $f\colon R \to S$ is a map of sets such that

(a) f preserves addition and multiplication: for all elements $a, b \in R$

$$f(a+b) = f(a) + f(b), \quad f(a \cdot b) = f(a) \cdot f(b);$$

(b) f preserves the identities: $f(1_R) = 1_S$.

We often refer to f as simply a *map of rings.*

Remark 4.2. Note that f defines a homomorphism of abelian groups

$$f\colon (R, +) \to (S, +).$$

It therefore automatically preserves the zero elements. Indeed

$$f(0_R) + f(0_R) = f(0_R + 0_R) = f(0_R),$$

so adding $-f(0_R)$ to both sides gives $f(0_R) = 0_S$. It also preserves additive inverses:

$$f(a) + f(-a) = f(a - a) = f(0_R) = 0_S,$$

so adding $-f(a)$ to both sides gives $f(-a) = -f(a)$.

Examples 4.3. (a) The inclusion maps $\mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$ are all ring homomorphisms.

(b) The quotient map $\mathbb{Z} \to \mathbb{Z}/n$ sending an integer $i \in \mathbb{Z}$ to the corresponding residue class $[i] \in \mathbb{Z}/n$ is a ring homomorphism.

(c) Complex conjugation $z \mapsto \bar{z}$ defines a ring homomorphism $\mathbb{C} \to \mathbb{C}$.

(d) For any point $x \in X$ the evaluation map $\mathrm{ev}_x\colon \mathrm{Fun}(X, \mathbb{R}) \to \mathbb{R}$ defined by

$$\mathrm{ev}_x(f) = f(x).$$

is a ring homomorphism.

(e) Let S be a trivial ring. Then for any ring R there is a unique ring homomorphism $R \to S$. There is no choice in defining the map of sets, and it is easily checked to be a ring homomorphism.

The following is a very important example:

Lemma 4.4. *Let R be a ring. Then there is a unique ring homomorphism $f\colon \mathbb{Z} \to R$. It is defined by the formulae*

$$f(n) = \begin{cases} \underbrace{1_R + \cdots + 1_R}_{n} & \textit{if } n > 0 \\ 0_R & \textit{if } n = 0 \\ -(\overbrace{1_R + \cdots + 1_R}^{|n|}) & \textit{if } n < 0 \end{cases}$$

Proof. Firstly we must show that the map $f\colon \mathbb{Z} \to R$ defined by the given formulae is indeed a ring homomorphism. Certainly we have $f(1_{\mathbb{Z}}) = 1_R$. We must also check the relations

$$f(n+m) = f(n) + f(m), \quad f(n \cdot m) = f(n) \cdot f(m)$$

for all integers $n, m \in \mathbb{Z}$. These are obvious if $n, m \geqslant 0$. One must then separately treat the cases when one or both of n and m are negative. We leave the details of this to the reader.

For the uniqueness statement suppose that $g\colon \mathbb{Z} \to R$ is a ring homomorphism. Since $g(1_Z) = 1_R$ and g preserves addition we see that for any $n > 0$

$$g(n) = g(1_{\mathbb{Z}} + \cdots + 1_{\mathbb{Z}}) = 1_R + \cdots + 1_R = f(n).$$

Since g also takes additive inverses to additive inverses we must have

$$g(-n) = -g(n) = -f(n) = f(-n).$$

Thus we see that $g = f$ which proves the uniqueness of f. □

Non-examples 4.5.

(a) The inverse map $-1\colon \mathbb{Z} \to \mathbb{Z}$ is not a ring homomorphism since it does not preserve multiplication.

(b) The determinant map

$$\det\colon \mathrm{M}_{2\times 2}(\mathbb{R}) \to \mathbb{R}$$

is not a ring homomorphism since it fails to preserve addition.

(c) The map $\mathbb{Z}/6 \to \mathbb{Z}/6$ given by $[a] \mapsto [4a]$ preserves addition and mulitplication, but is not a ring homomorphism, because it does not preserve the identity.

(d) The map $\mathbb{R} \to \mathbb{R}$ given by $t \mapsto 0$ is not a ring homomorphism because it does not preserve the identity.

Lemma 4.6.

(a) *If $f\colon R \to S$ and $g\colon S \to T$ are ring homomorphisms then so is the composite map $g \circ f\colon R \to T$.*

(b) *If $f\colon R \to S$ is a bijective ring homomorphism then the inverse map $f^{-1}\colon S \to R$ is also a ring homomorphism.*

Proof. Part (a) is immediate from the definitions and is left to the reader.

For part (b) note first that the assumption $f(1_R) = 1_S$ ensures that $f^{-1}(1_S) = 1_R$. To prove that f^{-1} preserves addition we must show that

$$f^{-1}(s_1 + s_2) = f^{-1}(s_1) + f^{-1}(s_2) \tag{1}$$

for all $s_1, s_2 \in S$. Since f is a bijection we can take $r_1, r_2 \in R$ such that $f(r_i) = s_i$. Thus $f^{-1}(s_i) = r_i$. Since f preserves addition we have $f(r_1 + r_2) = s_1 + s_2$. Hence $f^{-1}(s_1 + s_2) = r_1 + r_2$. This proves (1). A similar argument works for multiplication. □

Definition 4.7. An *isomorphism of rings* is a bijective ring homomorphism. Two rings are called *isomorphic* if they are related by a ring isomorphism.

We write $R \cong S$ to mean that the rings R and S are isomorphic. It follows from Lemma 4.6 that this is an equivalence relation. Isomorphic rings are 'sort of the same': they consist of the same elements with the same operations, but are somehow 'labelled differently'.

Examples 4.8. (a) Suppose the set $X = \{x\}$ consists of a single element. Then the evaluation homomorphism

$$\mathrm{ev}_x\colon \mathrm{Fun}(X, \mathbb{R}) \to \mathbb{R}$$

is a ring isomorphism.

(b) All rings with a single element are isomorphic. In particular, the two rings of Example 2.2 (d) are isomorphic.

(c) All rings with 2 elements are isomorphic to $R = \mathbb{Z}/2$. Indeed, by Remark 11.6(d), such a ring S must have $0_S \neq 1_S$ so that $S = \{0_S, 1_S\}$. The map

$$f\colon R \to S, \quad 0_R \mapsto 0_S, \quad 1_R \mapsto 1_S$$

is then a bijection of sets, which is easily checked to be a ring homomorphism, and hence an isomorphism of rings.

All reasonable properties of rings are invariant under isomorphism. So for example, if two rings are isomorphic and one is (commutative, an integral domain, reduced, a field) then so is the other. Proving these statements is an easy exercise, which we leave to the reader.

Example 4.9. The rings $\mathbb{Z}, \mathbb{Z}/n\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ and $\mathbb{C}$ are all non-isomorphic. To see this, first note that since isomorphic rings are related by a bijection, they must have the same cardinality ('number of elements'). Thus the only possibilities are $\mathbb{Z} \cong \mathbb{Q}$ (these sets are both countably infinite), or $\mathbb{R} \cong \mathbb{C}$ (these sets are both uncountable). Now $\mathbb{Q}$ is a field whereas $\mathbb{Z}$ is not, so certainly $\mathbb{Z}$ and $\mathbb{Q}$ are not isomorphic. To show that $\mathbb{R}$ and $\mathbb{C}$ are not isomorphic, note that $\mathbb{C}$ contains an element r satisfying $r^2 + 1 = 0$ whereas $\mathbb{R}$ does not.

Given a ring homomorphism $f\colon R \to S$ we define the image and kernel of f as follows

$$\operatorname{Im}(f) = \{b \in S : b = f(a) \text{ for some } a \in R\} \subset S$$
$$\operatorname{Ker}(f) = \{a \in R : f(a) = 0_S\} \subset R.$$

These will turn out to be the basic examples of a subring and an ideal respectively (see below).

Lemma 4.10. *Let $f\colon R \to S$ be a ring homomorphism. Then*

(a) *f is surjective precisely if* $\operatorname{Im}(f) = S$*;*

(b) *f is injective precisely if* $\operatorname{Ker}(f) = \{0_R\}$.

Proof. Part (a) is obvious.

For part (b), note that since $f(0_R) = 0_S$ we always have $0_R \in \operatorname{Ker}(f)$. Now $\operatorname{Ker}(f) = f^{-1}(0_S)$, so if f is injective then $\operatorname{Ker}(f) = \{0_R\}$. Conversely, let us assume that $\operatorname{Ker}(f) = \{0_R\}$. Suppose elements $r_1, r_2 \in R$ satisfy $f(r_1) = f(r_2)$. Then $f(r_2 - r_1) = 0_S$, so $r_2 - r_1 \in \operatorname{Ker}(f)$ and hence $r_2 - r_1 = 0_R$. Thus $r_1 = r_2$ and we have proved that f is injective. □

5. Subrings

Let R be a commutative ring. A subset $S \subset R$ is called a subring if the operations on R restrict to give a ring structure on S. More precisely we have

Definition 5.1. A *subring* of R is a subset $S \subset R$ such that

(a) S is closed under addition and multiplication:
$$a, b \in S \implies a + b \in S \text{ and } a \cdot b \in S;$$

(b) S contains additive inverses:
$$a \in S \implies -a \in S;$$

(c) S contains the identity: $1_R \in S$.

Note that a subring $S \subset R$ is itself a ring when equipped with the operations induced from R. Moreover the inclusion map $i\colon S \to R$ is a ring homomorphism. In fact this gives an alternative characterisation: a subring is a subset $S \subset R$ which is also a ring in such a way that the inclusion map $i\colon S \to R$ is a ring homomorphism.

Examples 5.2. (a) There is a chain of subrings
$$\mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}.$$

(b) Let X be the closed interval $X = [0, 1] \subset \mathbb{R}$. There is a subring
$$C(X) \subset \operatorname{Fun}(X, \mathbb{R})$$
consisting of continuous functions $f\colon X \to \mathbb{R}$. The same holds for any topological space X.

(d) Let $R = \mathbb{Z}[x]$. There is a subring $S \subset R$ consisting of polynomials of the form

$$\sum_{i=0}^{d} a_{2i}x^{2i} = a_0 + a_2x^2 + a_4x^4 + \cdots + a_{2d}x^{2d}.$$

Note that there is an isomorphism $R \cong S$ obtained by sending

$$\sum_{i=0}^{d} a_i x^i \mapsto \sum_{i=0}^{d} a_i x^{2i}.$$

(d) The Gaussian integers $\{a + bi : a, b \in \mathbb{Z}\} \subset \mathbb{C}$ form a subring.

Observe that a subring $S \subset R$ is always the image of a ring homomorphism, namely the inclusion map $i \colon S \to R$. Conversely we have

Lemma 5.3. *If $f \colon R \to S$ is a ring homomorphism then* $\mathrm{Im}(f) \subset S$ *is a subring.*

Proof. We must check the conditions of Definition 5.1. First suppose that $s_1, s_2 \in \mathrm{Im}(f)$. We can write $s_i = f(r_i)$ for elements $r_1, r_2 \in R$. Then

$$s_1 + s_2 = f(r_1 + r_2) \in \mathrm{Im}(f), \quad s_1 \cdot s_2 = f(r_1 \cdot r_2) \in \mathrm{Im}(f),$$

which proves condition (a). For condition (b), take $s \in \mathrm{Im}(f)$ and write $s = f(r)$. Then since f is a ring homomorphism, $f(-r) = -f(r) = -s$. Thus $-s \in \mathrm{Im}(f)$ which verifies condition (b). Condition (c) is immediate since $1_S = f(1_R)$. □

Remark 5.4. Suppose that $f \colon R \to S$ is a ring homomorphism. We can always factor f as a composite of ring homomorphisms $f = i \circ f'$ where $f' \colon R \to \mathrm{Im}(f)$ and $i \colon \mathrm{Im}(f) \to S$ is the inclusion of $\mathrm{Im}(f)$.

$$\begin{tikzcd}
R \arrow[rr, "f"] \arrow[dr, "f'"'] & & S . \\
& \mathrm{Im}(f) \arrow[ur, "i"'] &
\end{tikzcd}$$

Note that the homomorphism f' is surjective, by definition.

If $S_1, S_2 \subset R$ are two subrings of a ring R, then it is easy to check that the intersection $S_1 \cap S_2 \subset R$ is also a subring. More generally we have

Lemma 5.5. *Suppose given a set J, and for each element $j \in J$, a subring $S_j \subset R$. Then the intersection $S = \bigcap_{j \in J} S_j \subset R$ is a subring.*

Proof. This is immediate from the definitions. Let us check for example that the intersection S is closed under addition. Suppose given elements $a, b \in S$. This means precisely that $a, b \in S_j$ for all $j \in J$. Then since the subsets $S_j \subset R$ are subrings we have $a + b \in S_j$ for all $j \in J$. This then implies that $a + b \in S$. □

Definition 5.6. Let R be a ring and $T \subset R$ a subset. The subring of R *generated by* T is the intersection of all subrings $S \subset R$ containing T. It is denoted $\langle T \rangle \subset R$.

Note that $\langle T \rangle \subset R$ is indeed a subring by Lemma 5.5. In fact it is the smallest subring of R containing T: if $S \subset R$ is any other subring containing T then by definition we have $\langle T \rangle \subset S$. To describe $\langle T \rangle \subset R$ explicitly we first define the subset

$$\hat{T} = \{t_1 \cdot t_2 \cdots t_n : n \geqslant 0,\ t_i \in T\} \subset R, \tag{2}$$

consisting of all finite (possibly empty) products of elements of T. Note that $1_R \in \hat{T}$ by definition, since we interpret the empty product as meaning 1_R.

Lemma 5.7. *The subring $\langle T \rangle \subset R$ consists of those elements of R which can be written as finite (possibly empty) sums of the form*

$$r = \pm(p_1 + p_2 + \cdots + p_k) \text{ with } k \geqslant 0 \text{ and } p_i \in \hat{T}. \tag{3}$$

Here, when $k = 0$, we interpret the empty sum as meaning 0_R.

Proof. Let $X \subset R$ be the set of all elements of the form (3). It is easy to see that any subring $S \subset R$ containing T also contains X, since by definition S is closed under addition, multiplication and additive inverses. In particular $X \subset \langle T \rangle$. On the other hand, we claim that the subset $X \subset R$ is itself a subring. Since $T \subset X$ it then follows from Definition 5.6 that $\langle T \rangle \subset X$, and hence that $\langle T \rangle = X$.

To prove the claim note that X is clearly closed under addition and additive inverses and contains 1_R. To see that it is closed under multiplication we use

$$(p_1 + \cdots + p_k) \cdot (q_1 + \cdots + q_l) = \sum_{i,j} p_i \cdot q_j.$$

Since $p_i \cdot q_j \in \hat{T}$ for all i, j, the sum on the right also lies in X. □

Definition 5.8. We say that a ring R is *generated by* a subset $T \subset R$ if $R = \langle T \rangle$. We say that a ring R is *finitely-generated* if it is generated by a finite subset.

Examples 5.9.

(a) Let R be a ring and consider the empty subset $T = \emptyset \subset R$. Then $\hat{T} = \{1_R\}$ and $\langle T \rangle = \{n \cdot 1_R : n \in \mathbb{Z}\}$. Thus the subring generated by $\emptyset$ is the image of the unique ring homomorphism $f\colon \mathbb{Z} \to R$ of Lemma 4.4.

(b) The residue class ring $\mathbb{Z}/n$ and the integers $\mathbb{Z}$ are generated by the empty subset $\emptyset$.

(c) The polynomial ring $\mathbb{Z}[x]$ is generated by the single element set $\{x\} \subset \mathbb{Z}[x]$.

(d) Take $R = \mathbb{Z}[x]$ and consider the one-element subset $T = \{x^2\} \subset R$. The subring $S = \langle T \rangle \subset R$ is the subring of Example 5.2(d).

Example 5.10. The ring $\mathbb{Q}$ is not finitely-generated. Indeed, given any finite subset $T \subset \mathbb{Q}$ there is a finite set of prime numbers $\mathcal{P}$ such that every element of T can be written as a fraction a/b, with the integer b being a product of primes from $\mathcal{P}$. The subset $S \subset \mathbb{Q}$ consisting of all rational numbers which can be written in this form is easily seen to be a subring of $\mathbb{Q}$. Hence $\langle T \rangle \subset S$ and since $S \subsetneq \mathbb{Q}$ is a proper subring, it follows that $\mathbb{Q}$ is not generated by T.

6. Ideals

Let R be a commutative ring. Besides subrings, there is another very important class of subsets of R known as ideals.

Definition 6.1. An *ideal* in R is a non-empty subset $I \subset R$ such that

(a) I is closed under addition: $a, b \in I \implies a + b \in I$;

(b) I is preserved by multiplication by arbitrary elements of R:

$$a \in R,\ b \in I \implies a \cdot b \in I.$$

There are two trivial examples: the zero ideal $\{0\} \subset R$, and the whole ring R itself. An ideal is called *proper* if $I \neq R$.

Remarks 6.2. (a) If $I \subset R$ is an ideal, then $(I, +)$ is a subgroup of $(R, +)$, since I is closed under addition, and

$$a \in I \implies -a = (-1_R) \cdot a \in I.$$

(b) If an ideal $I \subset R$ contains 1_R then $I = R$ because for all $r \in R$ we have $r = r \cdot 1_R \in I$

In particular, Remark 6.2(b) shows that if a subset $S \subset R$ is both a subring and an ideal, then $S = R$.

Examples 6.3. (a) Take an integer $n \in \mathbb{Z}$. The subset

$$n\mathbb{Z} = \{m \in \mathbb{Z} : m = n \cdot k \text{ for some } k \in \mathbb{Z}\} \subset \mathbb{Z}$$

is an ideal.

(b) Let X be a set and consider the ring $\mathrm{Fun}(X, \mathbb{R})$. For any point $x \in X$ there is an ideal

$$I_x = \{f \in \mathrm{Fun}(X, \mathbb{R}) : f(x) = 0\}$$

consisting of functions vanishing at the point x.

Lemma 6.4. *Any ideal in $\mathbb{Z}$ is of the form $n\mathbb{Z}$.*

Proof. The zero ideal $\{0\} = 0\mathbb{Z}$ so it is enough to consider non-zero ideals $I \subset \mathbb{Z}$. Any such ideal contains a positive integer, since it is an additive subgroup. Let n be the smallest such. Then $n\mathbb{Z} \subset I$ and in fact equality holds, because if $r \in I$ is not divisible by n then by adding multiples of n we can find $0 < r < n$ with $r \in I$, contradicting the minimality of n. □

The basic examples of ideals are kernels of ring homomorphisms.

Lemma 6.5. *If $f\colon R \to S$ is a ring homomorphism then* $\operatorname{Ker}(f) \subset R$ *is an ideal.*

Proof. Certainly $\operatorname{Ker}(f)$ is non-empty since $0_R \in \operatorname{Ker}(f)$. Suppose $a_1, a_2 \in \operatorname{Ker}(f)$. Then

$$f(a_1 + a_2) = f(a_1) + f(a_2) = 0_S + 0_S = 0_S$$

so $a_1 + a_2 \in \operatorname{Ker}(f)$. Finally, suppose $r \in R$ and $a \in \operatorname{Ker}(f)$. Then

$$f(r \cdot a) = f(r) \cdot f(a) = f(r) \cdot 0_S = 0_S$$

so that $r \cdot a \in \operatorname{Ker}(f)$ also. □

If $I_1, I_2 \subset R$ are ideals, then so is the intersection $I_1 \cap I_2 \subset R$. More generally

Lemma 6.6. *Suppose given a set J, and for each element $j \in J$ an ideal $I_j \subset R$. Then the intersection $I = \bigcap_{j \in J} I_j \subset R$ is an ideal.*

Proof. Note that the intersection I is definitely non-empty since every ideal I_j contains the zero element 0_R. The proof is now very similar to that of Lemma 5.5. Let us show for example that I is closed under multiplication by arbitrary elements of R. Suppose then that $i \in I$ and $r \in R$. Then $i \in I_j$ for all $j \in J$, so since each $I_j \subset R$ is an ideal, we have $r \cdot i \in I_j$ for all $j \in J$. This implies that $r \cdot i \in I$ which is what we wanted to prove. □

Definition 6.7. Let R be a commutative ring, and $T \subset R$ a subset. The *ideal generated by* T is the intersection of all ideals $I \subset R$ containing T. It is denoted $(T) \subset R$.

Note that the subset $(T) \subset R$ is indeed an ideal by Lemma 6.6. In fact it is the smallest ideal of R containing T: if I is any ideal in R containing T then by definition we have $(T) \subset I$.

Lemma 6.8. *The ideal (T) consists of all finite sums of the form*

$$r = r_1 \cdot t_1 + \cdots + r_k t_k, \quad k \geqslant 1, \quad r_j \in R, \quad t_j \in T.$$

Proof. Let X be the set of such elements. Any ideal $I \subset R$ containing T must also contain X since it is closed under addition and multiplication by elements of R. It follows from Definition 6.7 that $X \subset (T)$. On the other hand the subset X is clearly an ideal, so by Definition 6.7 again we get $(T) \subset X$. Hence $(T) = X$, which is what we wanted to show. □

Definition 6.9. We say that an ideal $I \subset R$ is *generated by* a subset $T \subset R$ if $(T) = I$. An ideal is said to be *finitely-generated* if it is generated by a finite set of elements.

If $T = \{t_1, \cdots, t_k\}$ is a finite set we write $(t_1, \cdots, t_k) \subset R$ for the ideal $(T) \subset R$ generated by T. It consists of all elements $r \in R$ of the form

$$r = r_1 \cdot t_1 + \cdots + r_k \cdot t_k \text{ with } r_1, \cdots, r_k \in R.$$

An ideal generated by a single element is of the form

$$I = (t) = \{r \cdot t : r \in R\}.$$

Such ideals are said to be *principal.*

Remarks 6.10. (a) In any ring R, the zero ideal $\{0\} = (0)$ and the full ring $R = (1)$ are both principal ideals.

(b) Lemma 6.4 shows that all ideals in the ring $\mathbb{Z}$ are principal. Integral domains with this property are called *principal ideal domains.*

Examples 6.11. (a) The principal ideal $(x) \subset \mathbb{Z}[x]$ consists of all polynomials with zero constant term.

(b) The ideal $(3, 5) \subset \mathbb{Z}$ is the whole of $\mathbb{Z}$.

(c) The ideal generated by the empty set is always the zero ideal (0).

(c) Consider the subset $T = \{2x, x^2\} \subset \mathbb{Z}[x]$. The subring $\langle T\rangle$ consists of polynomials with even coefficients in odd orders, i.e. those of the form

$$a_0 + 2a_1x + a_2x^2 + 2a_3x^3 + \cdots, \quad a_i \in \mathbb{Z}.$$

The ideal (T) consists of polynomials with constant term 0 and even linear term, i.e. those of the form

$$2a_1x + a_2x^2 + a_3x^3 + \cdots, \quad a_i \in \mathbb{Z}.$$

If $I_1, I_2 \subset R$ are ideals, we can form their sum

$$I_1 + I_2 = \{r \in R : r = i_1 + i_2 \text{ with } i_1 \in I_1 \text{ and } i_2 \in I_2\}.$$

It is easy to see that this is an ideal, and in fact is the ideal generated by the subset $I_1 \cup I_2 \subset R$. Note that

$$(t_1) + (t_2) = (t_1, t_2)$$

for any elements $t_1, t_2 \in R$.

7. Quotient rings

We now come to the very important concept of a quotient ring. It generalises the construction of the ring of residue classes $\mathbb{Z}/n$. Let R be a commutative ring and $I \subset R$ an ideal. We shall construct another commutative ring R/I which is called the quotient of the ring R by the ideal I.

Define an equivalence relation $\sim$ on the elements of R by defining

$$a_1 \sim a_2 \iff a_2 - a_1 \in I, \quad a_1, a_2 \in R.$$

First of all note that this is indeed an equivalence relation. It is reflexive because $a - a = 0_R \in I$. It is symmetric because $a_2 - a_1 = -(a_1 - a_2)$ and I is closed under additive inverses. It is transitive because I is closed under addition:

$$a_2 - a_1 \in I, \quad a_3 - a_2 \in I \implies a_3 - a_1 = (a_3 - a_2) + (a_2 - a_1) \in I.$$

We write R/I for the set of equivalence classes for this relation. The equivalence class containing an element $a \in R$ will be denoted $[a]$. Thus

$$a_1 \sim a_2 \iff [a_1] = [a_2] \iff a_2 - a_1 \in I.$$

Note that so far we only used the fact that I is a subgroup of $(R, +)$.

Remark 7.1. An alternative piece of notation for the equivalence class $[a]$ is $a + I$. This makes sense since an element $a' \in R$ lies in this equivalence class precisely if it is of the form $a + i$ for some element $i \in I$.

We claim that we can define addition and multiplication on the set R/I of equivalence classes by the rules

$$[a] + [b] = [a + b], \quad [a] \cdot [b] = [a \cdot b].$$

It is not immediately clear that this works though, because a given element of R/I can be represented in the form $[a]$ in many different ways.

Lemma 7.2. *These rules are well-defined and define the structure of a commutative ring on the set R/I. The zero and identity elements in R/I are the equivalence classes $0_{R/I} = [0_R]$ and $1_{R/I} = [1_R]$ respectively.*

Proof. Suppose that we take elements $a', b' \in R$ such that $[a] = [a']$ and $[b'] = [b]$. By definition, this translates into the condition that

$$a' - a = i \in I, \quad b' - b = j \in I.$$

We need to check that if we follow the rule above with these representatives then we get the same answers for the sum and product of the two equivalence classes. Now

$$a' + b' = a + b + i + j.$$

Since I is closed under addition we have $i + j \in I$ so this means that $[a + b] = [a' + b']$ and the sum is indeed a well-defined equivalence class. Similarly

$$a' \cdot b' = (a + i) \cdot (b + j) = a \cdot b + a \cdot j + i \cdot b + i \cdot j.$$

Since I is closed under addition and arbitrary products, this implies that $[a \cdot b] = [a' \cdot b']$ and the multiplication in R/I is also well-defined.

The required properties of addition and multiplication in R/I now follow easily from the corresponding statements for R. For example, to check associativity of multiplication we observe that

$$([a] \cdot [b]) \cdot [c] = [a \cdot b] \cdot [c] = [(a \cdot b) \cdot c] = [a \cdot (b \cdot c)] = [a] \cdot [b \cdot c] = [a] \cdot ([b] \cdot [c]),$$

where we used the definition of the multiplication in R/I and the associativity of the multiplication in R. □

Examples 7.3. (a) Take $R = \mathbb{R}[x]$ and $I = (x^2)$. Then every element of R is equivalent to one of the form $a + bx$ and no two such are equivalent. So

$$R/I = \{[a + bx] : a, b \in \mathbb{R}\}.$$

The element $[x]$ satisfies $[x]^2 = [x^2] = [0]$.

(b) Take $R = \mathbb{R}[x]$ and $I = (x^2 + 1)$. Using polynomial division we can write any $f \in \mathbb{R}[x]$ in the form

$$f(x) = q(x) \cdot (x^2 + 1) + r(x)$$

with $r(x) = ax + b$ having degree $\leqslant 1$. It follows that $f(x) \sim r(x)$ so as in (a) every element of R is equivalent to one of the form $a + bx$. No two of these can be equivalent because any non-zero element of the principal ideal (x^2+1) has degree $\geqslant 2$. So again

$$R/I = \{[a + bx] : a, b \in \mathbb{R}\}.$$

This time the element $[x]$ satisfies $[x]^2 = [x^2] = [-1]$. It's not hard to see that $R/I \cong \mathbb{C}$.

(c) Take $R = (\mathbb{Z}/2)[x]$ and $I = (x^2 + x + 1)$. Then as above we can write any $f(x) \in R$ in the form

$$f(x) = q(x) \cdot (x^2 + x + 1) + r(x)$$

with $\deg r(x) < \deg(x^2) = 2$. Then $f(x) \sim r(x)$ so we conclude that R/I consists of equivalence classes

$$R/I = \{[a + bx] : a, b \in \mathbb{Z}/2\}.$$

It follows that R/I has precisely 4 elements: $R/I = \{[0], [1],][x], [x+1]\}$. Note that we have relations

$$[x] \cdot [x + 1] = [x^2 + x] = [-1] = [1],$$

so every nonzero element of R/I is a unit. Thus R/I is a field. In field theory one learns that for every prime power $q = p^n$ there is a unique (up to isomorphism) field $\mathbb{F}_q$ with q elements, and that these are all the fields with finitely many elements. When $q = p$ is prime we have $\mathbb{F}_p = \mathbb{Z}/p$, but of course this doesn't work when $q = p^n$ with $n > 1$. We just constructed the field $\mathbb{F}_4$.

8. The isomorphism theorem

Let R be a commutative ring and $I \subset R$ an ideal. There is an obvious ring homomorphism $p \colon R \to R/I$ sending an element $r \in R$ to the corresponding equivalence class $[r] \in R/I$. This is usually called the *quotient map*. Note that its kernel is precisely the ideal I. Indeed

$$p(r) = 0_{R/I} \iff [r] = [0_R] \iff r \in I.$$

The quotient map has the following universal property.

Proposition 8.1 (Universal property of quotient rings). *Suppose that $f\colon R \to S$ is a ring homomorphism such that*

$$r \in I \implies f(r) = 0.$$

Then there is a unique ring homomorphism $\bar{f}\colon R/I \to S$ such that $f = \bar{f} \circ p$.

$$\begin{tikzcd} R \arrow[rr, "f"] \arrow[dr, "p"'] & & S . \\ & R/I \arrow[ur, dotted, "\bar{f}"'] & \end{tikzcd}$$

Proof. First we check the uniqueness claim. The property $f = \bar{f} \circ p$ means that

$$\bar{f}([r]) = \bar{f}(p(r)) = f(r) \text{ for all } r \in R, \tag{4}$$

which is enough to specify $\bar{f}$ uniquely.

To prove existence, let us define $\bar{f}$ by the rule (4). We must check that it defines a ring homomorphism. Certainly $\bar{f}$ preserves addition since

$$\bar{f}([a] + [b]) = \bar{f}([a + b]) = f(a + b) = f(a) + f(b) = \bar{f}(a) + \bar{f}(b).$$

A similar argument works for multiplication. Finally

$$\bar{f}(1_{R/I}) = \bar{f}([1_R]) = f(1_R) = 1_S$$

which completes the proof. □

Putting this together with Remark 5.4 we get the following result:

Theorem 8.2 (Isomorphism theorem). *Any ring homomorphism $f\colon R \to S$ can be written uniquely in the form $f = i \circ \bar{f}' \circ p$*

$$\begin{tikzcd} R \arrow[r, "f"] \arrow[d, "p"'] & S & . \\ R/\operatorname{Ker}(f) \arrow[r, "\bar{f}'"'] & \operatorname{Im}(f) \arrow[u, "i"'] & \end{tikzcd}$$

where $p\colon R \to R/\operatorname{Ker}(f)$ is the quotient map, $i\colon \operatorname{Im}(f) \to S$ is the inclusion map, and $\bar{f}'$ is an isomorphism.

Proof. By Proposition 8.1 we can factor f as the quotient map p followed by the induced homomorphism

$$\bar{f}\colon R/\operatorname{Ker}(f) \to S.$$

Note that $\bar{f}$ is injective, since

$$\bar{f}([r]) = 0 \iff f(r) = 0 \iff r \in \operatorname{Ker}(f) \iff r \sim 0_R.$$

Note also that $\text{Im}(\bar{f}) = \text{Im}(f) \subset S$. The result then follows by factoring $\bar{f}$ via its image, as in Remark 5.4. □

Example 8.3. Consider the ring homomorphism $f\colon \mathbb{R}[x] \to \mathbb{C}$ obtained by sending

$$a_0 + a_1 x + \cdots + a_d x^d \mapsto a_0 + a_1 i + \cdots + a_d i^d.$$

The image is the whole of $\mathbb{C}$ since $a + bx$ maps to $a + bi$.

We claim that the kernel of f is the principal ideal $(x^2 + 1)$. Certainly $(x^2 + 1) \subset \text{Ker}(f)$. Suppose for a contradiction that the inclusion is strict, and take a polynomial

$$g \in \text{Ker}(f) \setminus (x^2 + 1).$$

Subtracting multiples of $x^2 + 1$ we can assume that $\deg(g) < 2$, so that $g = ax + b$ with $a, b \in \mathbb{R}$. But then $f(g) = a + bi$, and since we assumed that $g \in \text{Ker}(f)$ this implies that $a = b = 0$. But then $g = 0$, contradicting the assumption $g \notin (x^2 + 1)$.

The isomorphism theorem shows that we get an isomorphism of rings

$$\mathbb{R}[x]/(x^2 + 1) \cong \mathbb{C}.$$

Thus we can construct the complex numbers from $\mathbb{R}$ by adjoining a symbol x and imposing the condition $x^2 + 1 = 0$.

Next we want to study ideals in quotient rings. First of all, we note the following general result:

Lemma 8.4. *Suppose $f\colon R \to S$ is a ring homomorphism and $J \subset I$ is an ideal. Then $f^{-1}(J) \subset R$ is an ideal.*

Proof. Suppose $a, b \in f^{-1}(J)$. Thus $f(a), f(b) \in J$. Then

$$f(a + b) = f(a) + f(b) \in J$$

so $a + b \in f^{-1}(J)$. Next suppose $a \in f^{-1}(J)$ and $r \in R$. Then

$$f(r \cdot a) = f(r) \cdot f(a) \in J$$

so $r \cdot a \in f^{-1}(J)$. Finally, note that $f^{-1}(J)$ is certainly non-empty: since $f(0_R) = 0_S$ and $0_S \in J$ we have $0_R \in f^{-1}(J)$. □

Let $I \subset R$ be an ideal and consider the quotient map $p\colon R \to R/I$. For any ideal $K \subset R/I$ there is an ideal

$$p^{-1}(K) = \{r \in R : [r] \in K\} \subset R,$$

and this ideal contains I since $I = p^{-1}(\{0\}) \subset p^{-1}(K)$.

Conversely, suppose $J \subset R$ is an ideal containing I and consider

$$J/I = \{[r] \in R/I : r \in J\}.$$

Note first that this is well-defined subset of R/I: if $[r_1] = [r_2] \in R/I$ then $r_2 - r_1 \in I$ and hence, since $I \subset J$, we have $r_2 \in J \iff r_1 \in J$.

We now show that $J/I \subset R/I$ is an ideal. Firstly, J/I is non-empty because it contains $[0] \in R/I$. Secondly, J/I is closed under addition because if $[a], [b] \in J/I$, then $a, b \in J$ and so, since J is an ideal, $a + b \in J$ and therefore $[a + b] \in J/I$. Finally, J/I is closed under multiplication by arbitrary elements of R/I, because if $[r] \in R/I$ and $[a] \in J/I$, then $a \in J$ so, since J is an ideal, $r \cdot a \in J$ and therefore $[r] \cdot [a] = [r \cdot a] \in J/I$.

Proposition 8.5. *There is a bijection*

$$\big\{\textit{Ideals in } R \textit{ containing } I\big\} \longrightarrow \big\{\textit{Ideals in } R/I\big\}$$

sending an ideal $J \subset R$ containing I to the ideal $J/I \subset R/I$. The inverse map is given by $K \mapsto p^{-1}(K)$.

Proof. It will be enough to show that the composite of the two given maps in either order is the identity. Firstly, if $J \subset R$ is an ideal containing I then by definition $p^{-1}(J/I) = J$. For the other composition, suppose $K \subset R/I$ is an ideal, and set $J = p^{-1}(K)$. We must show that $K = J/I$. Suppose $[a] \in K$. Then $a \in p^{-1}(K) = J$ so also $[a] \in J/I$. Conversely, suppose $[a] \in J/I$. Then $a \in J = p^{-1}(K)$ so $[a] \in K$. □

The following is sometimes called the third isomorphism theorem.

Proposition 8.6. *Let R be a commutative ring. Given ideals $I \subset J \subset R$, there is an isomorphism of rings*

$$R/J \to (R/I)/(J/I).$$

Proof. There is a map of sets $f\colon R/J \to R/I$ sending $[r] \in R/J$ to $[r] \in R/I$ for all $r \in R$. This is well-defined because $r_2 - r_1 \in J \implies r_2 - r_1 \in I$. It is easy to check that f is a ring homomorphism, since all operations are induced from those in R, and it is clearly surjective. An element $[r] \in R/J$ satisfies $f([r]) = 0$ precisely if $r \in J$. Thus the kernel of f is the ideal J/I considered above. The result therefore follows from the isomorphism theorem. □

Example 8.7. Take $R = \mathbb{Z}$ and $I = (8)$ and set $S = \mathbb{Z}/8$. The ideals in R containing I are $I \subset (4) \subset (2) \subset (1) = R$. Thus the ideals in S are $([0]) \subset ([4]) \subset ([2]) \subset ([1]) = S$. Let us put $J = (2) \subset R$. Then $J/I = ([2]) \subset S$ and

$$S/([2]) = (R/I)/(J/I) \cong R/J \cong \mathbb{Z}/2.$$

9. Maximal, prime and radical ideals

In this section we consider ideals with certain special properties. We start with the following simple observation.

Lemma 9.1. *A commutative ring R is a field precisely if its only proper ideal is $\{0\}$.*

Proof. If R is a field and $\{0\} \neq I \subset R$ is a nonzero ideal, then taking an element $0_R \neq r \in I$ we have that $1_R = r \cdot r^{-1} \in I$ which implies that $I = R$. Conversely, suppose that the only proper ideal of R is the zero ideal. Then for any nonzero element $0_R \neq r \in R$ the principal ideal (r) must be the full ring R. But this implies that $1_R = a \cdot r$ for some $a \in R$, which shows that r has a multiplicative inverse. This proves that R is a field. □

Remark 9.2. Suppose $f\colon R \to S$ is a ring homomorphism, and R is a field. Then exactly one of the following happens

(a) $\mathrm{Ker}(f) = (0)$, which is to say that f is injective.

(b) $\mathrm{Ker}(f) = R$. Then $f(1_R) = 1_S = 0_S$ and hence S is a trivial ring.

So ring homomorphisms from fields to non-trivial rings are always injective.

Definition 9.3. A proper ideal $I \subset R$ is said to be

(a) *maximal* if there are no proper ideals of R strictly containing I;

(b) *prime* if for all elements $a, b \in R$
$$a \cdot b \in I \implies a \in I \text{ or } b \in I;$$

(c) *radical* if for all elements $a \in R$
$$a^n \in I \text{ for some } n \geqslant 1 \implies a \in I.$$

As a matter of convention, the non-proper ideal R itself is considered to be radical but not prime or maximal.

Example 9.4. Let $p > 1$ be a prime number. Then the ideal $(p) \subset \mathbb{Z}$ is prime because
$$p \mid ab \implies p \mid a \text{ or } p \mid b.$$
Conversely if $n > 1$ is a composite number then (n) is not a prime ideal. To see this write $n = a \cdot b$ with $1 < a, b > n$ and then
$$a \notin (n) \text{ and } b \notin (n) \text{ but } a \cdot b \in (n).$$
Thus for any integer $n > 0$ the ideal (n) is prime precisely if n is a prime number.

The following result explains the importance of the above special classes of ideals.

Lemma 9.5. *An ideal $I \subset R$ is*

(a) *maximal if and only if R/I is a field;*

(b) *prime if and only if R/I is an integral domain;*

(c) *radical if and only if R/I is reduced.*

Proof. Part (a) follows immediately from Proposition 8.5 together with Lemma 9.1.

For part (b), note that R/I is an integral domain precisely if
$$[a] \cdot [b] = 0_{R/I} \implies [a] = 0_{R/I} \text{ or } [b] = [0_{R/I}].$$
This easily translates into the condition that I be prime. Similary for the last part. □

Lemma 4.6 shows that there are implications

$$I \text{ maximal} \implies I \text{ prime} \implies I \text{ radical.}$$

Example 9.6. Fix a positive integer $n > 0$, and consider the ideal $(n) \subset \mathbb{Z}$. Lemma 3.8 and Lemma 9.5 show that we have implications

$$(n) \text{ is prime} \iff (n) \text{ is maximal} \iff n \text{ is prime.}$$

The ideal (0) is prime but not maximal.

Example 9.7. Let $n > 0$ be a positive integer. Example 3.9 and Lemma 9.5 show that the ideal $(n) \subset \mathbb{Z}$ is radical precisely if n is square-free. Thus for example, if we take $n = 12 = 2^2 \cdot 3$, then n is not square-free and the ideal (n) is not radical because $6 \notin (12)$ but $6^2 \in (12)$.

Lemma 9.8. *Let $f\colon R \to S$ be a ring homomorphism. If I is a prime (respectively radical) ideal then so is $f^{-1}(I)$.*

Proof. Suppose that $a \cdot b \in f^{-1}(I)$. Then $f(a \cdot b) = f(a) \cdot f(b) \in I$. Since I is prime this implies either $f(a) \in I$ or $f(b) \in I$. But then either $a \in f^{-1}(I)$ or $b \in f^{-1}(I)$.

Similarly, suppose $a^n \in f^{-1}(I)$. Then $f(a)^n \in I$ so since I is radical, $f(a) \in I$ and hence $a \in f^{-1}(I)$. □

The corresponding statement for maximal ideals is false.

Example 9.9. Let $f\colon \mathbb{Z} \to \mathbb{Q}$ be the inclusion map. Then the ideal $(0) \subset \mathbb{Q}$ is maximal, but the ideal $f^{-1}(0) = (0) \subset \mathbb{Z}$ is not.

10. Algebras

Let k be a commutative ring.

Definition 10.1. A *k-algebra* is a ring R together with a ring homomorphism

$$\phi\colon k \to R.$$

The map ϕ is called the *structure map* of the algebra.

We often abuse terminology and say that R is a k-algebra, when we hope that the structure map is clear from the context.

Definition 10.2. Suppose that $\phi_1\colon k \to R_1$ and $\phi_2\colon k \to R_2$ are k-algebras. A homomorphism of k-algebras $f\colon R_1 \to R_2$ is a ring homomorphism satisfying the extra condition $f \circ \phi_1 = \phi_2$.

We represent the condition for an algebra homomorphism f by the commutative diagram

$$\begin{tikzcd}
R \arrow[rr, "f"] & & S . \\
& k \arrow[ul, "\phi_1"] \arrow[ur, "\phi_2"'] &
\end{tikzcd}$$

In words we say that f commutes with the structure maps.

Examples 10.3. (a) The ring $\mathrm{Fun}(X, \mathbb{R})$ is an $\mathbb{R}$-algebra where the structure map $\phi\colon \mathbb{R} \to \mathrm{Fun}(X, \mathbb{R})$ sends an element $r \in \mathbb{R}$ to the corresponding constant function defined by setting $f(x) = r$ for all $x \in X$.

(b) If k is any commutative ring, the polynomial algebra $k[x]$ is a k-algebra. The structure map is the inclusion $k \subset k[x]$ sending an element of k to the corresponding constant polynomial.

(c) Recall from Lemma 4.4 that if R is any commutative ring there is a unique homomorphism $\phi\colon \mathbb{Z} \to R$. It follows that a $\mathbb{Z}$-algebra is the same thing as a ring. Similarly, a homomorphism of $\mathbb{Z}$-algebras is the same thing as a homomorphism of rings.

(d) We can view $\mathbb{C}$ as a $\mathbb{C}$-algebra via the identity map, and as an $\mathbb{R}$-algebra via the inclusion $\mathbb{R} \subset \mathbb{C}$. Then complex conjugation defines an $\mathbb{R}$-algebra isomorphism $\mathbb{C} \to \mathbb{C}$ which is not a homomorphism of $\mathbb{C}$-algebras.

We often consider the case when the ring k is a field. By Remark 9.2, assuming the algebra is non-trivial, the structure map $\phi\colon k \to R$ is injective. The image of ϕ is then an isomorphic copy of k lying inside R. We often suppress the structure map and view k as being a subring of R.

Suppose k is a field and R is a k-algebra with structure map $\phi\colon k \to R$. We can consider R to be an abelian group using the addition operation in R. Moreover, we can define scalar multiplication by elements $\lambda \in k$ by the rule

$$\lambda \cdot r = \phi(\lambda) \cdot r,$$

using the structure map and the multiplication operation in R.

Lemma 10.4. *With these operations R becomes a vector space over k.*

Proof. The required axioms are

$$\lambda \cdot (r_1 + r_2) = \lambda \cdot r_1 + \lambda \cdot r_2, \quad (\lambda_1 + \lambda_2) \cdot r = \lambda_1 \cdot r + \lambda_2 \cdot r$$

$$\lambda_1 \cdot (\lambda_2 \cdot v) = (\lambda_1 \cdot \lambda_2) \cdot v, \quad 1_k \cdot r = r.$$

These are easily checked using the distributive law in R and the fact that the structure map ϕ is a ring homomorphism. □

Note that multiplication in R defines a map

$$\cdot\colon R \times R \to R, \quad (r_1, r_2) \mapsto r_1 \cdot r_2.$$

which is bilinear. This means that the map is linear in each variable, i.e. if we fix $s \in R$ then the maps $R \to R$ given by $r \mapsto r \cdot s$ and $r \mapsto s \cdot r$ are both linear maps. Explicitly this means that for $\lambda_1, \lambda_2, \lambda_3 \in k$ and $r_1, r_2, r_3 \in R$ we have

$$(\lambda_1 \cdot r_1 + \lambda_2 \cdot r_2) \cdot r_3 = \lambda_1 \cdot (r_1 \cdot r_3) + \lambda_2 \cdot (r_2 \cdot r_3),$$

$$r_1 \cdot (\lambda_2 \cdot r_2 + \lambda_3 \cdot r_3) = \lambda_2 \cdot (r_1 \cdot r_2) + \lambda_3 \cdot (r_1 \cdot r_3).$$

Remark 10.5. Conversely, one can easily show that a vector space R over k equipped with a bilinear, associative multiplication law, which has an identity element 1_R, is naturally a k-algebra. The structure map $\phi\colon k \to R$ is defined by the rule $\phi(\lambda) = \lambda \cdot 1_R$.

Definition 10.6. A k-algebra is called *finite-dimensional* if it is a finite-dimensional as a vector space over k. Its dimension is then the dimension of this vector space.

To be completely explicit, if R is a k-algebra with structure map $\phi\colon k \to R$, then R is finite-dimensional precisely if there exists a finite basis for R as a vector space over k. Such a basis is a collection of elements $\{r_1, \cdots, r_n\}$ in terms of which every element $r \in R$ can be uniquely written in the form

$$r = \phi(\lambda_1) \cdot r_1 + \cdots + \phi(\lambda_n) \cdot r_n$$

for some elements $\lambda_1, \cdots, \lambda_n \in k$. The dimension of R is then the number n of elements of this basis.

Remark 10.7. Note that if a k-algebra R has a basis $\{r_1, \cdots, r_n\}$ then the multiplication on R is completely determined by the products of the basis elements: once $r_i \cdot r_j \in R$ are given, all other products are determined by the bilinearity property.

Examples 10.8.

(a) The complex numbers $\mathbb{C}$ form a two-dimensional algebra over the real numbers $\mathbb{R}$. A basis is given by $1, i$, with 1 being the identity. The multiplication is completely determined by bilinearity and the condition $i^2 = -1$:

$$(a + bi) \cdot (c + di) = ac - bd + (ad + bc)i.$$

(b) Similarly, the quaternions form a (non-commutative) algebra over $\mathbb{R}$ of dimension 4 with basis $\{1, i, j, k\}$. The multiplication is completely determined by the rules

$$i^2 = j^2 = k^2 = ijk = -1,$$

because these imply that $ij = k = -ji$, $jk = i = -kj$ and $ki = j = -ik$.

(c) The ring $\mathbb{C}[x]/(x^n)$ is an algebra over $\mathbb{C}$ of dimension n, with basis

$$\{1, [x], [x^2], \cdots, [x^{n-1}]\}.$$

The multiplication is determined by the rule $[x^i] \cdot [x^j] = [x^{i+j}]$.

Let R be a k-algebra with structure map $\phi\colon k \to R$.

Definition 10.9. A *subalgebra* of R is a subring $S \subset R$ which contains the image of the structure map ϕ. Note that S is then a k-algebra via the induced ring homomorphism $\phi\colon k \to S$.

It follows from Lemma 5.5 that the intersection of any collection of subalgebras of R is also a subalgebra.

Definition 10.10. Let $T \subset R$ a subset. The *subalgebra of R generated by T* is the intersection of all subalgebras of R containing T. We denote it by $k[T] \subset R$.

Note that $k[T] \subset R$ is just the subring of R generated by $\operatorname{Im} \phi$ and T. To give an explicit description first recall the definition (2) of the subset $\hat{T} \subset R$.

Lemma 10.11. *The elements of $k[T] \subset R$ are precisely the k-linear combinations of the elements of $\hat{T}$, that is the sums of the form*

$$r = \lambda_1 \cdot p_1 + \cdots + \lambda_n \cdot p_n, \quad n \geqslant 0, \quad \lambda_i \in k, \quad p_i \in \hat{T}.$$

Proof. This follows in the same way as Lemma 5.7: any subalgebra of R containing the subset T must also contain all these elements, on the other hand these elements do indeed form a subalgebra of R containing T. □

Example 10.12. Consider the subset T of the ring $R = \mathbb{C}[x]$ consisting of the single polynomial x. Then

(a) The subring of R generated by T is then the subring $\mathbb{Z}[x] \subset \mathbb{C}[x]$ of polynomials with integer coefficients.

(b) The ideal in R generated by T is the principal ideal $(x) \subset \mathbb{C}[x]$ consisting of polynomials with zero constant term.

(c) The $\mathbb{C}$-subalgebra of R generated by T is the full algebra $R = \mathbb{C}[x]$.

Definition 10.13. A k-algebra R is said to be *generated by* a subset $T \subset R$ if the subalgebra $k[T] \subset R$ generated by T is equal to R itself. A k-algebra is said to be *finitely generated* if there is a finite subset $T \subset R$ such that R is generated by T.

Remark 10.14. Note that being finitely-generated as a k-algebra is different from being finitely-generated as a ring. Thus $\mathbb{C}[x]$ is finitely-generated as a $\mathbb{C}$-algebra, but not as a ring.

Remark 10.15. The condition that a k-algebra be finitely-generated is much weaker than the condition that it should be finite-dimensional. A k-algebra R is finite-dimensional if we can find elements $r_1, \cdots, r_n$ such that all elements of R are linear combinations of the element r_i; it is finitely-generated if we can find elements $r_1, \cdots, r_n$ such that all elements of R are linear combinations of *products of* the elements r_i. Thus, for example, the k-algebra $k[x]$ is finitely-generated, but not finite-dimensional.

11. Polynomial algebras

Let k be a commutative ring. As noted before, the polynomial ring $k[x]$ is a k-algebra as noted before, with the structure map $\phi\colon k \to k[x]$ sending an element $a \in k$ to the corresponding constant polynomial. We usually suppress ϕ from the notation, and confuse an element $a \in k$ with the corresponding constant polynomial.

Remark 11.1. When we write an element of the algebra $k[x]$ in the form

$$f(x) = a_0 + a_1 x + \cdots + a_n x^n, \quad a_i \in k \tag{5}$$

we can mean two slightly different things.

Firstly (5) can be thought of as a formal expression: the ring $k[x]$ is defined to be the set of all such expressions. In this way of looking at (5) the symbols x^i are just formal place-holders: we could equally well define $k[x]$ to consist of all sequences of elements of k of the form $(a_0, a_1, a_2, \cdots)$ such that $a_i = 0$ for $i \gg 0$. We could then define addition and mulitplication for such sequences in the same way as in Example 3.1(a).

The second way to look at (5) is as an equality in the ring $k[x]$. From this point-of-view the symbol x denotes a particular element of the ring $k[x]$ (corresponding to the string $(0, 1, 0, \cdots)$), and the symbol a_i denotes the constant polynomial associated to an element $a_i \in k$ (corresponding to the string $(a_i, 0, 0, \cdots)$). Thus the right-hand side of (5) is a sum of terms, each being a constant polynomial multiplied by a power of the element $x \in k[x]$.

Proposition 11.2 (Universal property of the polynomial algebra)**.** *Let $\phi\colon k \to R$ be a k-algebra. Then for every element $r \in R$ there is a unique homomorphism of k-algebras $f\colon k[x] \to R$ such that $f(x) = r$.*

Proof. By the definition of an algebra homomorphism, any such map f must take the constant polynomial $a \in k \subset k[x]$ to $\phi(a) \in R$. By assumption we also have $f(x) = r$. Since f preserves addition and multiplication it follows that

$$f(a_0 + a_1x + \cdots a_2x^2 + \cdots + a_dx^d) = a_0 + a_1r + \cdots a_2r^2 + \cdots + a_dr^d.$$

Thus f is obtained by substituting $x = r$ in a given polynomial. This proves that there can be at most one such homomorphism f. Conversely, it is easy to check that this f does indeed define an algebra homomorphism. □

Remark 11.3. The algebra $k[x]$ with its element x is essentially unique with the universal property of Proposition 11.2. To see this, suppose there were some other k-algebra S with an element $s \in S$ for which the universal property of Proposition11.2 also holds, meaning that for each k-algebra R there is a unique homomorphism of k-algebras $g\colon S \to R$ such that $g(s) = r$. Applying the univesral property for $k[x]$ gives an algebra homomorphism $f\colon k[x] \to S$ sending x to s. Applying the universal property for S gives an algebra homomorphism $g\colon S \to k[x]$ sending s to x. Now the composite $g \circ f$ is an algebra homomorphism $k[x] \to k[x]$ sending x to x, and hence, by the uniqueness part of the universal property of $k[x]$, must be the identity map. Similarly, the uniqueness part of the universal property of S shows that the composite $f \circ g$ is the identity map on S. This proves that f and g are mutually inverse isomorphisms of k-algebras.

More generally we can define the polynomial ring $k[x_1, \cdots, x_n]$ in n variables. A *monomial* in n variables is an expression of the form

$$x_1^{i_1} \cdots x_n^{i_n}, \quad i_1, \cdots, i_n \geqslant 0.$$

The elements of $k[x_1,\cdots,x_n]$ are defined to be finite combinations of monomials of the form

$$\sum_{i_1,\cdots,i_n\geqslant 0} a_{i_1,\cdots,i_n} x_1^{i_1}\cdots x_n^{i_n}$$

with coefficients $a_{i_1,\cdots,i_n}\in k$. Addition is given by the rule

$$\sum_{i_1,\cdots,i_n\geqslant 0} a_{i_1,\cdots,i_n} x_1^{i_1}\cdots x_n^{i_n} + \sum_{i_1,\cdots,i_n\geqslant 0} b_{i_1,\cdots,i_n} x_1^{i_1}\cdots x_n^{i_n}$$

$$= \sum_{i_1,\cdots,i_n\geqslant 0} (a_{i_1,\cdots,i_n}+b_{i_1,\cdots,i_n}) x_1^{i_1}\cdots x_n^{i_n},$$

and multiplication is defined by

$$\sum_{i_1,\cdots,i_n\geqslant 0} a_{i_1,\cdots,i_n} x_1^{i_1}\cdots x_n^{i_n} \cdot \sum_{i_1,\cdots,i_n\geqslant 0} b_{i_1,\cdots,i_n} x_1^{i_1}\cdots x_n^{i_n}$$

$$= \sum_{i_1,\cdots,i_n\geqslant 0} c_{i_1,\cdots,i_n} x_1^{i_1}\cdots x_n^{i_n}$$

where

$$c_{i_1,\cdots,i_n} = \sum_{0\leqslant j_1\leqslant i_1}\sum_{0\leqslant j_2\leqslant i_2}\cdots\sum_{0\leqslant j_n\leqslant i_n} a_{j_1,\cdots,j_n}\cdot b_{i_1-j_1,\cdots,i_n-j_n}.$$

The ring $k[x_1,\cdots,x_n]$ is a k-algebra via the ring homomorphism sending an element $a\in k$ to the corresponding constant polynomial. It has the following universal property, which is proved in the same way as Proposition 11.2.

Proposition 11.4 (Universal property of the polynomial algebra)**.** *Let $\phi\colon k\to R$ be a k-algebra. Then for every n-tuple of elements $r_1,\cdots,r_n\in R$ there is a unique homomorphism of k-algebras*

$$f\colon k[x_1,\cdots,x_n]\to R$$

such that $f(x_i)=r_i$.

We can alternatively define multi-variable polynomial algebras step-by-step by adjoining indeterminates one at a time:

Lemma 11.5. *For any $n\geqslant 1$ there is an isomorphism of k-algebras*

$$k[x_1,\cdots,x_{n-1}][x_n]\cong k[x_1,\cdots,x_n].$$

Proof. This can be proved directly by constructing a map in the obvious way. Alternatively one can use the universal property: both sides have the property that there is a unique map to any k-algebra R sending the elements x_i to any given n-tuple of elements $r_i\in R$. By the uniqueness argument of Remark 11.3 the two sides must be isomorphic. □

Remark 11.6. Suppose that R is a k-algebra generated by a finite set $\{r_1, \cdots, r_n\} \subset R$. By the universal property of polynomial rings there is a unique algebra homomorphism

$$f\colon k[x_1, \cdots, x_n] \to R$$

sending x_i to the element $r_i \in R$. This map is surjective since its image is a k-subalgebra which contains the elements r_i, and by assumption, the smallest such subalgebra is R itself. Hence by the isomorphism theorem there is an isomorphism

$$R \cong k[x_1, \cdots, x_n]/I, \quad I = \operatorname{Ker}(f).$$

Thus we conclude that all finitely generated k-algebras are quotients of polynomial algebras. This is one reason for the importance of polynomial rings.

12. Noetherian rings

We now consider a very important special class of rings.

Definition 12.1. A ring R is Noetherian if every ascending chain of ideals

$$I_1 \subset I_2 \subset I_3 \subset \cdots$$

eventually stabilizes, i.e. there exists some $N \geqslant 0$ such that $I_N = I_{N+1} = ...$

Lemma 12.2. *A ring is Noetherian if and only if every ideal is finitely-generated.*

Proof. Suppose every ideal is finitely-generated and consider an ascending chain $I_1 \subset I_2 \subset \cdots$. The union $I = \bigcup_{n \geqslant 1} I_n$ is an ideal[4]. Indeed, if $a, b \in I$ then there exists n such that $a, b \in I_n$ and then $a + b \in I_n$ so also $a + b \in I$. Similarly if $r \in R$ and $a \in I$ then since $a \in I_n$ for some n we have $r \cdot a \in I_n$ and hence $r \cdot a \in I$.

By assumption we can find a finite set of generators $(f_1, \cdots, f_r)$ for I. Hence we can find N large enough so that $f_j \in I_N$ for all $1 \leqslant j \leqslant r$. But then $I \subset I_N$. But $I_n \subset I$ for all $n \geqslant N$. Hence $I = I_n$ for all $n \geqslant N$.

Conversely, consider an ideal $I \subset R$ and assume that I is not finitely-generated. Take an element $f_1 \in I$ and set $I_1 = (f_1) \subset I$. Then I_1 is finitely-generated so $I_1 \subsetneq I$ and we can take $f_2 \in I \setminus I_1$ and consider $I_2 = (f_1, f_2)$. Again I_2 is finitely-generated so $I_2 \subsetneq I$ and we can take $f_3 \in I \setminus I_2$. And so on. We get an infinite chain

$$I_1 \subsetneq I_2 \subsetneq I_3 \subsetneq \cdots,$$

and so R is not Noetherian. □

Example 12.3.

(a) Any field k is Noetherian since the only ideals are (0) and R and these are both principal.

(b) The ring $\mathbb{Z}$ is Noetherian; in fact any ideal is generated by a single element.

(c) If R is Noetherian and $I \subset R$ is an ideal then R/I is noetherian. Indeed, ideals in R/I correspond bijectively to ideals in R containing I.

[4] In general unions of ideals are not ideals e.g. in $\mathbb{Z}$ we have $3, 5 \in (3) \cup (5)$ but $3 + 5 \notin (3) \cup (5)$.

The following important result is known as Hilbert's basis theorem.

Theorem 12.4. *If R is a Noetherian ring then so is $R[x]$.*

Proof. Let $I \subset R[x]$ be an ideal. We must prove that I is finitely-generated. For each $n \geqslant 0$ define $J_n \subset R$ to be the subset of leading terms of degree n polynomials in I:

$$J_n = \big\{ a \in R : \exists f \in I \text{ of the form } f(x) = ax^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \big\}.$$

Note that each J_n is an ideal because I is. Also $J_n \subset J_{n+1}$ because if $f(x) \in I$ has leading term $a \cdot x^n$ then $x \cdot f(x)$ has leading term $a \cdot x^{n+1}$.

Since R is Noetherian each of the ideals J_m for $1 \leqslant m \leqslant N$ has a finite set of generators $t_{m,1}, \cdots, t_{m,k(m)}$. We can then take polynomials $g_{m,i} \in I$ with leading term $t_{m,i}x^m$. Again, since R is Noetherian there is some $N > 0$ such that $J_N = J_{N+1} = \cdots$. We claim that the union

$$\bigcup_{m=0}^{N} \{g_{m,1}, \cdots, g_{m,k(m)}\}$$

is a finite generating set for the ideal I.

Suppose $f \in I$. We must show that f is a linear combination of the above generators. Suppose f has leading term ax^d. Then $a \in J_d$ so is of the form

$$a = \sum_{i=1}^{k(m)} r_i \cdot t_{m,i}, \quad r_i \in R,$$

where we take $m = d$ if $d \leqslant N$ and $m = N$ for $d > N$. Then the polynomial

$$x^{d-m} \cdot \sum_{i=1}^{k(m)} r_i \cdot g_{m,i}$$

has degree d and the same leading coefficient as f. Hence we can subtract it and reduce to the case where f hasr degree $< d$. The result follows by induction on the degree of f. □

Applying this result repeatedly gives

Theorem 12.5. *If k is a Noetherian commutative ring then every finitely-generated algebra over k is noetherian.*

Proof. By repeatedly applying the basis theorem and using Lemma 11.5 we see that $k[x_1, \cdots, x_n]$ is Noetherian. By Example 12.3(b) any ring of the form $k[x_1, \cdots, x_n]/I$ is Noetherian. But by Remark 11.6 any finitely-generated k-algebra is of this form. □

Not all rings are Noetherian however:

Example 12.6. The ring R of continuous functions $f\colon \mathbb{R} \to \mathbb{R}$ is not noetherian. Indeed, consider the ideal $I \subset R$ consisting of functions of bounded support: those for which there exists some $R > 0$ with

$$|x| > R \implies f(x) = 0. \tag{6}$$

If I is finitely generated it would follow that there is some universal constant $R > 0$ such that for any $f \in I$ the relation (6) holds. But that is plainly untrue: there are functions of bounded support which are nonzero on any given bounded subset of $\mathbb{R}$.

One can similarly show that the ring of polynomials in infinitely many variables is non-Noetherian. The Noetherian condition can be thought of as a type of finiteness condition: it asks that the ring should not be too big in a certain sense.

13. Algebraic subsets

Fix a field k, for example $k = \mathbb{Q}, \mathbb{R}$ or $\mathbb{C}$. Define the affine space

$$\mathbb{A}^n_k = \{(a_1, \cdots, a_n) \in k^n\}.$$

Note that this is just the set k^n; the funny notation is traditional in algebraic geometry.

Given a polynomial $f = f(x_1, \cdots, x_n) \in k[x_1, \cdots, x_n]$ and a point

$$p = (a_1, \cdots, a_n) \in \mathbb{A}^n_k$$

we have an element $f(p) \in k$ obtained by substituting $x_i = a_i$ in the polynomial f. In other words, we view elements of the ring $k[x_1, \cdots, x_n]$ as functions $f\colon \mathbb{A}^n_k \to k$, and $f(p)$ is the evaluation of the function f at the point p.

Consider the maps

$$\begin{tikzcd}
\text{Ideals } I \subset k[x_1, \cdots, x_n] \arrow[r, bend left, "V"] & \text{Subsets } V \subset \mathbb{A}^n_k \arrow[l, bend left, "I"]
\end{tikzcd}$$

defined by

$$V(I) = \{p \in \mathbb{A}^n_k : f(p) = 0 \text{ for all } f \in I\},$$

$$I(V) = \{f \in k[x_1, \cdots, x_n] : f(p) = 0 \text{ for all } p \in V\}.$$

Note that these maps are order-reversing:

$$I_1 \subset I_2 \implies V(I_1) \supset V(I_2); \quad V_1 \subset V_2 \implies I(V_1) \supset I(V_2).$$

We now make the very important

Definition 13.1. A subset $X \subset \mathbb{A}^n_k$ is said to be *algebraic* if it is of the form $X = V(I)$ for some ideal $I \subset k[x_1, \cdots, x_n]$.

Since the ring $k[x_1, \cdots, x_n]$ is Noetherian, any ideal $I \subset k[x_1, \cdots, x_n]$ is finitely-generated and therefore of the form $I = (f_1, \cdots, f_r)$ for some finite set of polymonials $f_1, \cdots, f_r$. Then

$$V(I) = V(f_1, \cdots, f_r) = \{(a_1, \cdots, a_n) \in \mathbb{A}^n_k : f_i(a_1, \cdots, a_n) = 0 \text{ for } 1 \leqslant i \leqslant r\}.$$

Thus, a subset $X \subset \mathbb{A}^n_k$ is algebraic precisely if it is the vanishing locus of a finite set of polynomials.

Examples 13.2. (a) The empty-set $\emptyset \subset \mathbb{A}^n_k$ is an algebraic subset; it is the vanishing locus of the non-proper ideal $k[x_1, \cdots, x_n]$.

(b) The set $\mathbb{A}^n_k$ itself is an algebraic subset; it is the vanishing locus of the zero ideal $(0) \subset k[x_1, \cdots, x_n]$.

(b) The circle

$$V = \{(a, b) \in \mathbb{A}^2_{\mathbb{R}} : a^2 + b^2 = 1\} \subset \mathbb{A}^2_{\mathbb{R}}$$

is an algebraic subset. Indeed, $V = V(I)$ where $I \subset \mathbb{R}[x, y]$ is the principal ideal generated by the polynomial

$$f = x^2 + y^2 - 1 \in \mathbb{R}[x, y].$$

(c) The union of the two co-ordinate axes

$$V = \{(a, b) \in \mathbb{A}^2_{\mathbb{R}} : ab = 0\} \subset \mathbb{A}^2_{\mathbb{R}}$$

is an algebraic subset; being equal to $V(I)$ where

$$I = (xy) \subset \mathbb{R}[x, y].$$

(d) Any point $(a, b) \in \mathbb{A}^2_{\mathbb{R}}$ is an algebraic subset. It is the vanishing locus of the non-principal ideal

$$I = (x - a, y - b) \subset \mathbb{R}[x, y].$$

(e) The subset

$$V = \{a \in \mathbb{R} : a \geqslant 0\} \subset \mathbb{A}^1_{\mathbb{R}}$$

is not algebraic. Indeed, any polynomial in $\mathbb{R}[x]$ has only finitely many roots, so any proper algebraic subset of $\mathbb{A}^1_{\mathbb{R}}$ consists of at most finitely many points.

Finite intersections and arbitrary unions of algebraic subsets are also algebraic:

Proposition 13.3. (a) *Given an arbitrary collection of ideals $I_s \subset R$ indexed by a set S we have*

$$V\big(\sum_{s\in S} I_s\big) = V\big(\bigcap_{s\in S} V(I_s)\big).$$

(b) *Given two ideals $I_1, I_2 \subset R$ we have*

$$V(I_1 \cap I_2) = V(I_1) \cup V(I_2).$$

Proof. For part (a) note that by definition the ideal $\sum_{s\in S} I_s$ consists of all finite sums of elements of $\bigcup_{s\in S} I_s$. The statement $P \in V(I_s)$ for all $s \in S$ is equivalent to the statement that for any $f \in \bigcup_{s\in S} I_s$ we have $f(P) = 0$. But then the same is true for any element $f \in \sum_{s\in S} I_s$.

For part (b) note that there are obvious inclusions $V(I_j) \subset V(I_1 \cap I_2)$ which gives an inclusion

$$V(I_1 \cap I_2) \supset V(I_1) \cup V(I_2).$$

For the reverse inclusion suppose that $p \notin V(I_1) \cup V(I_2)$. Then we can find $f_1 \in I_1$ and $f_2 \in I_2$ with $f_1(p) \neq 0$ and $f_2(p) \neq 0$. But then $f_1 \cdot f_2 \in I_1 \cap I_2$ and $f_1 \cdot f_2(p) \neq 0$ so that $p \notin V(I_1) \cap V(I_2)$. □

Example 13.4. Consider the ideals $I_1 = (x, y) \subset \mathbb{C}[x, y, z]$ and $I_2 = (z) \subset \mathbb{C}[x, y, z]$. Thus $V(I_1)$ is the z-axis, and $V(I_2)$ is the (x, y)-plane. We claim that

$$I_1 + I_2 = (x, y, z), \quad I_1 \cap I_2 = (xz, yz).$$

For the second statement suppose that $f \in I_1 \cap I_2$. Then since $f \in I_1$, all monomials appearing with nonzero coefficient in $f(x, y, z)$ must contain either an x or a y. On the other hand, since $f \in I_2$, any monomial appearing with nonzero coefficient in $f(x, y, z)$ must contain a z. Hence all monomials appearing in $f(x, y, z)$ contain either xz or yz, and it follows that $f \in (xz, yz)$.

Finally note that $V(I_1 + I_2) = \{(0, 0, 0)\}$ and $V(I_1 \cap I_2)$ is the union of the z-axis with the (x, y)-plane.

Remark 13.5. Digression on the definition of a topological space. Recall that $\mathbb{R}^n$ has a distance function on it: given two points $x, y \in \mathbb{R}^n$ we define

$$d(x, y) = \left(\sum_{i=1}^{n} (y_i - x_i)^2 \right)^{\frac{1}{2}}.$$

A subset $U \subset \mathbb{R}^n$ is open if for any $x \in U$ there exists an $r > 0$ such that

$$x \in B_r(x) = \{y \in \mathbb{R}^n : d(x, y) < r\} \subset U.$$

A subset $F \subset \mathbb{R}^n$ is called closed if it is the complement of an open subset. One can make the same definitions in any metric space X.

Open subsets of a mteric space X (e.g. $X = \mathbb{R}^n$) have the following properties:

(a) The empty set $\emptyset \subset X$ and the set X itself are both open subsets;

(b) If $U_1, U_2 \subset X$ are open subsets then so is the intersection $U_1 \cap U_2 \subset X$;

(c) If $U_i \subset X$ is an arbitrary collection of open subsets of X indexed by some set I then the (possibly infinite) union $\bigcup_{i\in I} U_i \subset X$ is also an open subset.

A set X with a collection of subsets (called open subsets) satisfying these conditions is called a *topological space*. There are some rather extreme examples: you could take just $\emptyset, X$ as your open sets, or you could take all subsets to be open. A metric space always gives rise to a topological space by defining open subsets as above. But different metrics (equivalent ones) can give rise to the same topological space. And some topological spaces do not come from a metric space at all.

Remark 13.6. A topological space X is called Hausdorff if given any two distinct points $x \neq y \in X$, there exist open subsets $U, V \subset X$ such that

$$x \in U, y \in V \text{ and } U \cap V = \emptyset.$$

The topological space arising from a metric space (X, d) is always Hausdorff: given $x, y \in X$ we can take U and V to be open balls at x and y with radius $r < d(x, y)/2$. But for example, if a set X has more than one element, then the topology whose only open subsets are $\emptyset$ and X is never Hausdorff.

Proposition 13.3 shows that the algebraic subsets form the closed subsets of a topology on $\mathbb{A}^n_k$. This is called the Zariski topology. It's a rather weird topology, and almost never comes from a metric structure.

Example 13.7. Consider the case $n = 1$. Any polynomial $f \in k[x]$ of degree d has at most d roots. Hence the algebraic subsets of $\mathbb{A}^1_k = k$ are just the finite sets of points. Note that if k is infinite then any two non-empty open subsets intersect. Thus this topology is certainly not Hausdorff in general.

14. Nullstellensatz

Consider again the correspondences (8.5). We would like to say that V and I define inverse bijections, but this is of course not true as they are currently defined.

Example 14.1. (a) Let $I = (x^2) \subset \mathbb{C}[x]$ and $J = (x) \subset \mathbb{C}[x]$. Then $V(I) = V(J) = \{0\} \subset \mathbb{A}^1_{\mathbb{C}}$.

(b) Let $I = (x^2 + 1) \subset \mathbb{R}[x]$ and $J = \mathbb{R}[x]$. Then $V(I) = V(J) = \emptyset$.

(c) Let $k = \mathbb{Z}/p$ for some prime number $p > 0$ and take $I = x^p - x$. Then $V(I) = V(0) = \mathbb{A}^1_k$ because by Fermat's little theorem, any element $a \in \mathbb{Z}$ satisfies $a^p \equiv a$ modulo p.

The problem with Example (a) is that the ideal I is not radical. Recall the definition of the radical of an ideal

$$\sqrt{I} = \{r \in R : r^n \in I \text{ for some } n \geqslant 1\}.$$

Note that $V(I) = V(\sqrt{I})$ because $f^n(p) = 0 \implies f(p) = 0$. Moreover, for the same reason, any ideal of the form $I(V)$ is radical. Thus we can restrict our maps V and I

to correspondences

$$\begin{tikzcd}\text{Radical ideals } I \subset R \arrow[r, bend left, "V"] & \text{Algebraic subsets } V \subset \mathbb{A}^n_k \arrow[l, bend left, "I"]\end{tikzcd}$$

The basic problem in Example (b) is that we considered a polynomial x^2+1 which had no real roots. Of course this would not have worked over $\mathbb{C}$, where the polynomial x^2+1 has the two roots $\pm i$. Example (c) shows a different problem when the field k is finite: there are nonzero polynomials which vanish at every point of $\mathbb{A}^1_k$. To eliminate both of these problems we restrict to a special class of fields.

Definition 14.2. A field k is said to be *algebraically closed* if every non-constant polynomial $f \in k[x]$ has a root in k.

Using polynomial division it follows that if k is algebraically closed then any polynomial $f \in k[x]$ can be written in the form

$$f(x) = c \cdot (x - a_1) \cdots (x - a_k),$$

with $c \in k$ and $a_i \in k$.

Example 14.3.
(a) The fields $\mathbb{Q}$ and $\mathbb{R}$ are not algebraically closed because $f(x) = x^2 + 1$ has no roots.
(b) The field $\mathbb{C}$ is algebraically closed: this is called the fundamental theorem of algebra.
(c) A finite field k is never algebraically closed because the polynomial

$$f(x) = 1 + \prod_{a \in k} (x - a)$$

has no roots. Thus algebraically closed fields have infinitely many elements.

Remark 14.4. It can be proved that any field k can be embedded as a subfield $k \subset \bar{k}$ of a minimal algebraically closed field known as the *algebraic closure* of k. It is defined by formally adjoining roots of all polynomials in $k[t]$. Thus for example $\mathbb{C}$ is the algebraic closure of $\mathbb{R}$, and the algebraic closure of $\mathbb{Q}$ is the field $\bar{\mathbb{Q}} \subset \mathbb{C}$ of algebraic numbers.

From now on we always assume that our field k is algebraically closed. The reader will not lose much by taking $k = \mathbb{C}$. The following important result is called Hilbert's Nullstellensatz (the German means 'Zeroes theorem').

Theorem 14.5. *Assume that k is algebraically closed. Then*

$$I(V(I)) = \sqrt{I}.$$

In particular, if I is a radical ideal then $I(V(I)) = I$.

We defer the proof to later in the course.

Example 14.6. Consider the ideal

$$I = (x^2 + y^2 - 1, y - 1) \subset \mathbb{C}[x, y].$$

The algebraic set $V = V(I)$ is the intersection of the circle of unit radius centered at the origin with the line $y = 1$. It therefore consists of a single point $(0, 1)$. The element $x \in \mathbb{C}[x, y]$ lies in the ideal $I(X)$ since it vanishes on the set V. Thus by the Nullstellensatz we must have $x \in \sqrt{I}$. In fact

$$x^2 = x^2 + y^2 - 1 - (y + 1) \cdot (y - 1) \in I.$$

But we claim that $x \notin I$. Indeed, if $x \in I$ we can write

$$x = p(x, y) \cdot (x^2 + y^2 - 1) + q(x, y) \cdot (y - 1)$$

for some polynomials $p, q \in \mathbb{C}[x, y]$. Applying the homomorphism $\mathbb{C}[x, y] \to \mathbb{C}[x]$ obtained by setting $y = 1$ gives $x = p(x, 1) \cdot x^2$ which gives a contradiction.

The following important corollary is absolutely vital for algebraic geometry.

Corollary 14.7. *The two maps I, V give mutually-inverse order-reversing bijections*

$$\begin{tikzcd}
\textit{Radical ideals } I \subset R \arrow[r, bend left, "V"] & \textit{Algebraic subsets } V \subset \mathbb{A}^n_k \arrow[l, bend left, "I"]
\end{tikzcd}$$

Proof. The Theorem shows that if I a radical ideal then $I(V(I)) = I$. Now suppose that $V \subset \mathbb{A}^n_k$ is an algebraic subset. By definition we can write $V = V(I)$ for some ideal, which we can take to be radical since $V(I) = V(\sqrt{I})$. But now

$$V(I(V)) = V(I(V(I)) = V(I) = V$$

so the result is proved. □

It follows that we get a bijection between maximal ideals of R and mimimal algebraic subsets of $\mathbb{A}^n_k$. Note that any point $p = (a_1, \cdots, a_n) \in \mathbb{A}^n_k$ is an algebraic subset since it is the vanishing locus of the ideal

$$I(p) = (x_1 - a_1, \cdots, x_n - a_n).$$

Hence we get a bijection

$$\begin{tikzcd}
\text{Maximal ideals } I \subset R \arrow[r, bend left, "V"] & \text{Points } p \in \mathbb{A}^n_k \arrow[l, bend left, "I"]
\end{tikzcd}.$$

15. Affine varieties

We have given geometric interpretations of radical and maximal ideals in polynomial rings above. The following definition allows a similar interpretation of prime ideals.

Definition 15.1. We say that an algebraic set $V \subset \mathbb{A}^n_k$ is irreducible if it cannot be written in the form $V = V_1 \cup V_2$ with $V_i \subsetneq V$ being proper algebraic subsets.

Example 15.2. (a) The algebraic set $V = V(xz, yz) \subset \mathbb{A}^3_{\mathbb{C}}$ is not irreducible since it is the union $V = V(z) \cup V(x, y)$ which are both proper algebraic subsets.

(b) The algebraic set $V = V(xy) \subset \mathbb{A}^2_{\mathbb{C}}$ is not irreducible since it is the union $V = V(x) \cup V(y)$ which are both proper algebraic subsets.

Lemma 15.3. *An algebraic set $V \subset \mathbb{A}^n_k$ is irreducible iff the ideal $I(V) \subset R$ is prime.*

Proof. Suppose first that $I = I(V)$ is prime. Suppose $V = V_1 \cup V_2$ with $V_i \subsetneq V$. Put $I_i = I(V_i)$. Then $V(I) = V(I_1 \cap I_2)$. Since I and $I_1 \cap I_2$ are both radical, the Nullstellensatz shows that $I = I_1 \cap I_2$ and with $I \subsetneq I_i$. Take $f_i \in I_i \setminus I$. Then $f_1 \cdot f_2 \in I_1 \cap I_2 = I$ contradicting the statement that I is prime.

Conversely, suppose that V is irreducible and take elements $f_i \in k[x_1, \cdots, x_n] \setminus I$ such that $f_1 \cdot f_2 \in I$. Consider the algebraic subsets $V_i = V \cap V(f_i) = V(I + (f_i))$. Then $V_i \subsetneq V$. But if $p \in V$ then $f_1 \cdot f_2(p) = 0$ so either $p \in V_1$ or $p \in V_2$. Hence $V = V_1 \cup V_2$. □

Thus we get a final bijection

$$\begin{tikzcd}
\text{Prime ideals } I \subset R \arrow[r, bend left, "V"] & \text{Irreducible algebraic subsets } V \subset \mathbb{A}^n_k \arrow[l, bend left, "I"]
\end{tikzcd}$$

It is fairly obvious that we can break any given algebraic set into irreducible pieces. More precisely we have

Lemma 15.4. *Every algebraic subset $X \subset \mathbb{A}^n_k$ has a unique (up to reordering) decomposition*

$$X = X_1 \cup X_2 \cup \cdots \cup X_r$$

with each $X_i \subset \mathbb{A}^n_k$ an irreducible subset, and $X_i \not\subset X_j$ for $i \neq j$.

Proof. Let us call an algebraic subset $X \subset \mathbb{A}^n_k$ good if it can be written as a finite union of irreducible algebraic subsets. Suppose $X_0 \subset \mathbb{A}^n_k$ is bad. Then in particular X_0 is not irreducible, so we can write it as a union $X_0 = X_1 \cup Y$ with $X_1, Y \subsetneq X$. If X_1 and Y are both good then obviously X_0 would also be good. Hence without loss of generality X_1 must be bad. We can repeat the argument and obtain a bad subset $X_2 \subsetneq X_1$. In this way we obtain an infinite descending chain

$$\cdots \subsetneq X_3 \subsetneq X_2 \subsetneq X_1 \subsetneq X_0$$

of bad subsets of X_0. But note that the ring $k[x_1, \cdots, x_n]$ is Noetherian. So by the order-reversing bijection of Corollary 14.7 there can be no infinite descending chains of algebraic subsets of $\mathbb{A}^n_k$. This gives a contradiction. We have therefore shown that any algebraic subset X is good, and hence has a decomposition as in the statement. We can easily ensure the minimality condition $X_i \not\subseteq X_j$ for $i \neq j$ by simply discarding any unnecessary factors.

For uniqueness, suppose we have

$$X = X_1 \cup \cdots \cup X_n = Y_1 \cup \cdots \cup Y_m$$

both satisfying the minimality condition. Then we have

$$Y_1 = (Y_1 \cap X_1) \cup \cdots \cup (Y_1 \cap X_n).$$

Since Y_1 is irreducible we must have $Y_1 \cap X_i = Y_1$ for some i. Relabelling we can therefore assume that $Y_1 \subset X_1$. The same argument shows that we must also have $X_1 \subset Y_i$ for some i. Since this implies that $Y_1 \subset Y_i$ the minimality condition implies that $i = 1$ and hence $X_1 = Y_1$. Repeating this argument gives the result. □

Example 15.5. Consider $I = (x^2 - yz, xz - x) \subset \mathbb{C}[x, y, z]$ and the corresponding algebraic subset $V = V(I) \subset \mathbb{A}^3_{\mathbb{C}}$. Suppose that $(x, y, z) \in V$. Then the second equation shows that either $x = 0$ or $z = 1$. In the first case we must have either $y = 0$ or $z = 0$. In the second case we must have $y = x^2$. It is then easy to see that

$$V = V(x, y) \cup V(x, z) \cup V(y - x^2, z - 1).$$

Let us label the subsets appearing on the right hand side V_1, V_2 and V_3 respectively. Note that the minimality condition $V_i \not\subseteq V_j$ for $i \neq j$ is satisfied because the points $(0, 0, 2)$, $(0, 1, 0)$ and $(1, 1, 1)$ are contained only in V_1, V_2 and V_3 respectively.

To prove that the subsets V_i are irreducible it is enough to show that the defining ideals are prime. For the first two this follows as in the last example. For the last we note that

$$\mathbb{C}[x, y, z]/(y - x^2, z - 1) \cong \mathbb{C}[x]$$

via the unique $\mathbb{C}$-algebra homomorphism sending $(x, y, z) \mapsto (x, x^2, 1)$. We have now proved that the irreducible components of V are the subsets V_1, V_2 and V_3.

Example 15.6. Consider $I = (xz - y^2, x^3 - yz) \subset \mathbb{C}[x, y, z]$ and the corresponding algebraic subset $V = V(I) \subset \mathbb{A}^3_{\mathbb{C}}$. Suppose that $(x, y, z) \in V$. Multipltying the first equation by z and the second by y and subtracting gives $xz^2 = x^3y$. Hence either $x = 0$ or $z^2 = x^2y$. In the first case the first equation then implies that $y = 0$. Conversely, if $x = y = 0$ then both defining equations are satisfed. Thus we can write

$$V = V(x, y) \cup V(xz - y^2, x^3 - yz, z^2 - x^2y).$$

Let us call the two subsets appearing on the right hand side V_1 and V_2 respectively. Note that we don't have $V_1 \subset V_2$ or $V_2 \subset V_1$. Indeed $(0, 0, 1) \in V_1 \setminus V_2$ and $(1, 1, 1) \in V_2 \setminus V_1$.

The subset V_1 is the line $x = y = 0$. This is easily seen to be irreducible geometrically, because it is a copy of the affine line $\mathbb{A}^1_{\mathbb{C}}$, so the only algebraic subsets are finite sets of points. Algebraically we note that $\mathbb{C}[x, y, z]/(x, y) \cong \mathbb{C}[z]$. This shows firstly that the ideal (x, y) is radical, which ensures that $I(V_1) = (x, y)$, and secondly that $(x, y) = I(V_1)$ is prime, and hence that the subset V_1 is irreducible.

We will show later that the other subset $V_2 \subset V$ is also irreducible (see Example 17.4). We will then have proved that the irreducible components of V are V_1 and V_2.

One last definition

Definition 15.7. An *affine variety* is an irreducible algebraic subset $X \subset \mathbb{A}^n_k$.

16. Co-ordinate ring

Let $X \subset \mathbb{A}^n_k$ be an algebraic set. The co-ordinate ring of X is defined to be the quotient ring

$$k[X] = k[x_1, \cdots, x_n]/I(X).$$

Note that $k[X]$ is always reduced, since the ideal $I(X)$ is radical.

Remark 16.1. An algebraic subset $X \subset \mathbb{A}^n_k$ is a variety iff the co-ordinate ring $k[X]$ is an integral domain.

We can think of the elements of $k[X]$ as functions on the algebraic set X.

Definition 16.2. Let $X \subset \mathbb{A}^n_k$ be an algebraic set. A function $f\colon X \to k$ is called *polynomial* if there is a polynomial $g \in k[x_1, \cdots, x_n]$ such that for all points $p = (a_1, \cdots, a_n) \in X$ we have $f(p) = g(a_1, \cdots, a_n)$.

Lemma 16.3. *The set of polynomial functions on an algebraic subset $X \subset \mathbb{A}^n_k$ forms a k-algebra under pointwise operations. This algebra is isomorphic to the co-ordinate ring $k[X]$.*

Proof. Let $\mathrm{Fun}_{\mathrm{poly}}(X)$ denote the ring of polynomial functions $X \to k$ with pointwise operations and give it the obvious k-algebra structure, where the structure map sends an element $\lambda \in k$ to the corresponding constant function $X \to k$. There is then a k-algebra homomorphism $k[x_1, \cdots x_n] \to \mathrm{Fun}_{\mathrm{poly}}(X)$ sending a polynomial to the induced polynomial function on X. This homomorphism is surjective, and by definition its kernel is $I(X)$. The result therfore follows from the isomorphism theorem. □

For any algebraic set $X \subset \mathbb{A}^n_k$ we have a correspondence

$$\begin{tikzcd}
\text{Ideals } I \subset k[X] \arrow[r, bend left, "V"] & \text{Algebraic subsets } Y \subset X \arrow[l, bend left, "I"] .
\end{tikzcd}$$

Indeed, the two sides can be identified with

$$\begin{tikzcd}\text{Ideals } I(X) \subset I \subset k[x_1, \cdots, x_n] \arrow[r, bend left, "V"] & \text{Algebraic subsets } Y \subset X \subset \mathbb{A}^n_k \arrow[l, bend left, "I"]\end{tikzcd}.$$

Applying the Nullstellensatz we get

Theorem 16.4. *Let k be an algebraically closed field, and $X \subset \mathbb{A}^n_k$ an algebraic subset. Then the two maps I, V give mutually-inverse order-reversing bijections*

$$\begin{tikzcd}\textit{Radical ideals } I \subset k[X] \arrow[r, bend left, "V"] & \textit{Algebraic subsets } V \subset X \arrow[l, bend left, "I"]\end{tikzcd}$$

In particular, the points of X (which are minimal algebraic subsets) correspond to maximal ideals of $k[X]$. Similarly, irreducible algebraic subsets $V \subset X$ correspond to prime ideals in the ring $k[X]$.

Example 16.5. Consider the algebra $R = \mathbb{C}[x, y]/(xy^2 - x)$. To understand it geometrically we would like to view it as a co-ordinate ring of some algebraic subset. Write $I = (xy^2 - x) \subset \mathbb{C}[x, y]$. The obvious thing to do is take $V = V(I) \subset \mathbb{A}^2_{\mathbb{C}}$. Then by the Nullstellensatz $I(V) = \sqrt{I}$ and the co-ordinate ring $\mathbb{C}[V]$ is $\mathbb{C}[x, y]/\sqrt{I}$. So the first thing is to check that the ideal I is radical, so that $I = \sqrt{I}$.

The easiest way to do that is to use the fact that all polynomial rings over fields are unique factorisation domains. For any $f \in \mathbb{C}[x, y]$ of positive degree we can write f in the form

$$f = g_1 \cdot g_2 \cdots g_k,$$

with the polynomials $g_i \in \mathbb{C}[x, y]$ being *irreducible*: that is, not the product of lower-order polynomials. The important fact is that this decomposition of f is basically unique: the factors g_i appearing are unique up to reordering and multiplication by scalar factors. It is immediate from this that $f \in I$ precisely if it is divisible by all three of the polynomials x, $y - 1$ and $y + 1$. It is then clear that $f^n \in I \implies f \in I$ and hence I is radical.

The maximal ideals of R are in bijection with the points of the set V which consists of the y-axis and the lines $y = \pm 1$. They are therefore of the form

$$(x, y - b), \quad (x - a, y + 1), \quad (x - a, y - 1) \text{ with } a, b \in \mathbb{C}.$$

The prime ideals which are not maximal correspond to irreducible subsets of V which are not points. There are three such, namely

$$(x), \quad (y - 1), \quad (y + 1).$$

These correspond to the y-axis, and the two lines $y = \pm 1$.

We equip an algebraic subset $X \subset \mathbb{A}^n_k$ with the topology induced from the Zariski topology on $\mathbb{A}^n_k$. This means that the closed subsets $Y \subset X$ are precisely the algebraic subsets $Y \subset \mathbb{A}^n_k$ which are contained in X.

17. Polynomial maps

We now start to consider maps between algebraic subsets.

Definition 17.1. Suppose $X \subset \mathbb{A}^m_k$ and $Y \subset \mathbb{A}^n_k$ are algebraic subsets. A map $\phi\colon X \to Y$ is called *polynomial* if there are polynomials

$$\phi_1 \in k[x_1, \cdots, x_m], \quad \phi_n \in k[x_1, \cdots, x_m],$$

such that

$$\phi(a_1, \cdots, a_m) = \big(\phi_1(a_1, \cdots, a_m), \cdots, \phi_n(a_1, \cdots, a_m)\big).$$

Note that we could equivalently view the ϕ_i as being polynomial functions on X or as elements of the co-ordinate ring $k[X]$.

Examples 17.2. (a) Cuspidal cubic. Consider the polynomial map $\phi\colon \mathbb{A}^1_{\mathbb{C}} \to \mathbb{A}^2_{\mathbb{C}}$ given by $t \mapsto (t^2, t^3)$. Its image is contained in the algebraic subset $V = V(y^2 - x^3) \subset \mathbb{A}^2_{\mathbb{C}}$. We thus get an induced map

$$\phi\colon \mathbb{A}^1_{\mathbb{C}} \to V.$$

This polynomial map is a bijection: indeed, given any point $(x, y) \in \mathbb{A}^2_{\mathbb{C}}$ satisfying $y^2 = x^3$ we can take $t = \pm\sqrt{x}$ and choose the sign uniquely so that $t^3 = y$.

(b) Consider the polynomial map $\phi\colon \mathbb{A}^2_{\mathbb{C}} \to \mathbb{A}^2_{\mathbb{C}}$ given by $(s, t) \mapsto (s^2, st, t^2)$. The image is the algebraic subset $V = (y^2 - xz) \subset \mathbb{A}^3_{\mathbb{C}}$. Indeed, given any $(x, y, z) \in V$ we can take $s = \pm\sqrt{x}$ and $t = \pm\sqrt{z}$ and then, changing one of the signs if necessary, we have $st = y$. We thus get an induced polynomial map

$$\phi\colon \mathbb{A}^2_{\mathbb{C}} \to V$$

in which the inverse image of every point except the origin consists of 2 points $\pm(s, t)$. The inverse image of the origin $(0, 0)$ is just $(0, 0)$.

Note that if $f\colon Y \to k$ is a polynomial function on Y then the composition $f \circ \phi$ is a polynomial function on X. This defines a k-algebra homomorphism $\phi^*\colon k[Y] \to k[X]$.

Proposition 17.3. *Suppose $X \subset \mathbb{A}^m_k$ and $Y \subset \mathbb{A}^n_k$ are algebraic subsets. Sending a regular map $\phi\colon X \to Y$ to the algebra homomorphism $\phi^*\colon k[Y] \to k[X]$ defines a bijection*

$$\big\{\textit{Polynomial maps } X \to Y\big\} \to \big\{k\textit{-algebra homomorphisms } k[Y] \to k[X]\big\}.$$

Proof. First of all, a polynomial map $\phi\colon X \to Y$ is determined by the corresponding algebra homomorphism $\phi^*\colon k[Y] \to k[X]$. Indeed, if we take the co-ordinate functions $y_i \in k[Y]$ then $\phi^*(y_i) = \phi_i \in k[X]$ is the ith component of the map ϕ. Knowing all these clearly determines the map ϕ. This shows that the map $\phi \mapsto \phi^*$ is injective.

To prove that $\phi \mapsto \phi^*$ is surjective, suppose we have an algebra map $h\colon k[Y] \to k[X]$. Set $\phi_i = h(y_i)$ and consider the resulting regular map $\phi\colon X \to Y$ given by

$$\phi(a_1, \cdots, a_m) = \big(\phi_1(a_1, \cdots, a_m), \cdots, \phi_n(a_1, \cdots, a_m)\big).$$

Then $\phi^*(y_i) = \phi_i = h(y_i)$. But $k[Y]$ is generated as an k-algebra by the elements y_i. Since both h and ϕ^* are k-algebra maps it follows that they are equal. □

Example 17.4. Consider again the algebraic subset of Example 15.6

$$X = V(xz - y^2, x^3 - yz, z^2 - x^2y) \subset \mathbb{A}^3_k.$$

We claim that X is the image of the morphism

$$\phi\colon \mathbb{A}^1_k \to \mathbb{A}^3_k, \quad t \mapsto (t^3, t^4, t^5).$$

Indeed, given a point $(x, y, z) \in X$ note that $x = 0 \implies y = z = 0$. On the other hand, if $x \neq 0$, set $t = y/x \neq 0$ and $u = z/x \neq 0$. The defining relations become

$$u = t^2, \quad x = t \cdot u, \quad u^2 = xt.$$

These reduce to $u = t^2$ and $x = t^3$ and hence $(x, y, z) = (t^3, t^4, t^5)$.

Suppose now that $X = X_1 \cup X_2$ with $X_1, X_2 \subset X$ proper algebraic subsets. Then we can find $g_1 \in I(X_1) \setminus I(X)$ and $g_2 \in I(X_2) \setminus I(X)$. Then $g_1 \cdot g_2 \in I(X_1 \cup X_2) = I(X)$. Consider the algebra homomorphism

$$\phi^*\colon \mathbb{C}[x, y, z] \to \mathbb{C}[t], \quad (x, y, z) \mapsto (t^3, t^4, t^5).$$

Note that $I(X) = \mathrm{Ker}(\phi^*)$. Hence $\phi^*(g_1 \cdot g_2) = 0$. But since $\mathbb{C}[t]$ is an integral domain this means that $\phi^*(g_i) = 0$ for $i = 1$ or 2. But then $g_i \in I(X)$, a contradiction.

Definition 17.5. Let $X \subset \mathbb{A}^n_k$ and $Y \subset \mathbb{A}^m_k$ be algebraic subsets. A polynomial map $X \to Y$ is said to be an *isomorphism* if there is a polynomial map $\psi\colon Y \to X$ such that $\psi \circ \phi = \mathrm{id}_X$ and $\phi \circ \psi = \mathrm{id}_Y$.

Note that $\phi\colon X \to Y$ being an isomorphism implies that $\phi^*\colon k[Y] \to k[X]$ is an isomorphism of k-algebras with inverse ψ^*. Thus isomorphic algebraic sets have isomorphic co-ordinate rings.

Example 17.6. The map $\phi\colon \mathbb{A}^1_{\mathbb{C}} \to V$ of Example 17.2(a) corresponds to the k-algebra homomorphism

$$\phi^*\colon \mathbb{C}[V] = \mathbb{C}[x, y]/(y^2 - x^3) \to k[\mathbb{A}^1_{\mathbb{C}}] = \mathbb{C}[t], \quad (x, y) \mapsto (t^2, t^3).$$

Note that this is not an isomorphism: its image is $\mathbb{C}[t^2, t^3] \subset \mathbb{C}[t]$, which does not contain the element t. Thus ϕ is also not an isomorphism: there is no inverse polynomial map $\psi\colon V \to \mathbb{A}^1_{\mathbb{C}}$. Of course there is an inverse to ϕ, as we discussed in Example 17.2(a), but it is not a polynomial map since it involves taking square-roots.

18. Field of fractions

Throughout this lecture R is an integral domain. We construct a field $K(R)$ called the *field of fractions* of R together with an injective ring homomorphism

$$q\colon R \to K(R).$$

The basic example is when $R = \mathbb{Z}$; in that case the resulting field of fractions $K(\mathbb{Z})$ is the field $\mathbb{Q}$ of rational numbers.

Consider pairs

$$(r, s) \text{ with } r \in R \text{ and } s \in R \setminus \{0\}.$$

You should think of such a pair (r, s) as standing for the fraction r/s. Two such pairs (r_1, s_1) and (r_2, s_2) are defined to be equivalent if

$$r_1 \cdot s_2 - r_2 \cdot s_1 = 0.$$

Note that for any nonzero $a \in R$ we have $(a \cdot r, a \cdot s) \sim (r, s)$. The relation $\sim$ is clearly symmetric and reflexive. To check transitivity, suppose we have

$$r_1 \cdot s_2 - r_2 \cdot s_1 = 0 = r_2 \cdot s_3 - r_3 \cdot s_2$$

so that $(r_1, s_1) \sim (r_2, s_2) \sim (r_3, s_3)$. Multiplying the first relation by s_3 and the second by s_1 gives

$$r_1 \cdot s_2 \cdot s_3 = r_3 \cdot s_1 \cdot s_2.$$

Since R is an integral domain and $s_2 \neq 0$ we can cancel s_2 and conclude that $r_1 \cdot s_3 = r_3 \cdot s_1$ and hence $(r_1, s_1) \sim (r_3, s_3)$ as required.

We denote by $K(R)$ the set of equivalence classes of such pairs (r, s). We define addition and multiplication in $K(R)$ by the rules

$$(r_1, s_1) + (r_2, s_2) = (r_1 \cdot s_2 + r_2 \cdot s_1, s_1 \cdot s_2); \quad (r_1, s_1) \cdot (r_2, s_2) = (r_1 \cdot r_2, s_1 \cdot s_2).$$

We should check that these are well-defined. Suppose then that $(r_1', s_1') \sim (r_1, s_1)$ so that $r_1' \cdot s_1 = r_1 \cdot s_1'$. Then we have

$$(r_1', s_1') + (r_2, s_2) = (r_1' \cdot s_2 + r_2 \cdot s_1', s_1' \cdot s_2) \sim (r_1' \cdot s_1 \cdot s_2 + r_2 \cdot s_1 \cdot s_1', s_1 \cdot s_1' \cdot s_2)$$

$$\sim (r_1 \cdot s_1' \cdot s_2 + r_2 \cdot s_1 \cdot s_1', s_1 \cdot s_1' \cdot s_2) \sim (r_1 \cdot s_2 + r_2 \cdot s_1, s_1 \cdot s_2) = (r_1, s_1) + (r_2, s_2)$$

which shows that addition is well-defined. Similarly, for multiplication

$$(r_1', s_1') \cdot (r_2, s_2) = (r_1' \cdot r_2, s_1' \cdot s_2) \sim (r_1' \cdot s_1 \cdot r_2, s_1 \cdot s_1' \cdot s_2)$$

$$\sim (r_1 \cdot s_1' \cdot r_2, s_1 \cdot s_1' \cdot s_2) \sim (r_1 \cdot r_2, s_1 \cdot s_2) = (r_1, s_1) \cdot (r_2, s_2)$$

We must now check that the ring axioms hold. The zero and unit are the equivalence classes containing the pairs $0 = (0, 1)$ and $1 = (1, 1)$ respectively. To help with checking the axioms, note that given pairs (r_1, s_1) and (r_2, s_2) we can find representatives (i.e. pairs in the same equivalence classes) which have the same denominator (just like any two fractions can be put over the same denominator). Indeed

$$(r_1, s_1) \sim (r_1 \cdot s_2, s_1 \cdot s_2), \quad (r_2, s_2) \sim (r_2 \cdot s_1, s_1 \cdot s_2).$$

Addition then becomes very simple:

$$(r_1, s) + (r_2, s) = (s \cdot (r_1 + r_2), s) \sim (r_1 + r_2, s).$$

As an example, let us consider the distributive law

$$a_1 \cdot (a_2 + a_3) = a_1 \cdot a_2 + a_1 \cdot a_3$$

for any $a_1, a_2, a_3 \in K(R)$. We can take representatives $a_i = (r_i, s_i)$ such that $s_2 = s_3 = s$. Then

$$a_1 \cdot (a_2 + a_3) = (r_1, s_1) \cdot [(r_2, s) + (r_3, s)] = (r_1, s_1) \cdot (r_2 + r_3, s)$$

$$= (r_1 \cdot (r_2 + r_3), s_1 \cdot s) = (r_1, s_1) \cdot (r_2, s) + (r_1, s_1) \cdot (r_3, s) = a_1 \cdot a_2 + a_1 \cdot a_3.$$

The other axioms are pretty-much immediate, and we leave them to the reader.

The resulting commutative ring is $K(R)$ is called the *field of fractions* of R. To see that it is a field note that $(r, s) \sim (0, 1)$ precsiely if $r = 0$. Hence if (r, s) represents a nonzero element of $K(R)$ then $r \neq 0$ and we have

$$(r, s) \cdot (s, r) = (r \cdot s, r \cdot s) \sim (1, 1).$$

There is an obvious ring homomorphism

$$q \colon R \to K(R), \quad r \mapsto (r, 1)$$

Note that q is always injective, becasue its kernel consists of elements $r \in R$ such that $(r, 0) \sim (0, 0)$ and as above the only such element is $r = 0$.

Remark 18.1. We can view R as a subring of a field $K(R)$. More precisely, the image of the map q is isomorphic to R and is a subring of $K(R)$. Note that the assumption that R was an integral domain was clearly necessary for this: any subring of an integral domain (in particular, of a field) is also an integral domain.

The field $K(R)$ has the following universal property:

Examples 18.2. (a) The field of fractions of $\mathbb{Z}$ is the rational numbers $\mathbb{Q}$. The localization map $\mathbb{Z} \to \mathbb{Q}$ is just the inclusion.

(b) If R is a field then the field of fractions is just R itself. More precisely, the homomorphism $q \colon R \to K(R)$ is an isomorphism. It is always injective, and when R is a field it is also surjective, because for any pair (r, s) we have $(r, s) \sim (r \cdot s^{-1}, 1) = q(r \cdot s^{-1})$.

(c) Let $R = \mathbb{R}[x]$ be the ring of real polynomials. The field of fractions is the field of rational functions, i.e. expressions of the form

$$f(x)/g(x), \quad f, g \in \mathbb{R}[x], g \neq 0.$$

Note that any such expression gives a partially-defined function $\mathbb{R} \dashrightarrow \mathbb{R}$: it is undefined at the finite set of points where $g(x) = 0$.

The field of fractions satisfies the following universal property:

Lemma 18.3. *Suppose $f\colon R \to K$ is an injective ring homomorphism with K a field.*

$$\begin{tikzcd}
R \arrow[rr, "f"] \arrow[dr, "q"'] & & K . \\
& K(R) \arrow[ur, dotted, "\bar{f}"'] &
\end{tikzcd}$$

Then there is a unique ring homomorphism $\bar{f}\colon K(R) \to S$ such that $f = \bar{f} \circ q$.

Proof. Note that for any element $(r, s) \in K(R)$ we can write

$$(r, s) = (r, 1) \cdot (1, s) = (r, 1) \cdot (s, 1)^{-1} = q(r) \cdot q(s)^{-1}.$$

Therefore if we want to have $f = \bar{f} \circ q$ there is no choice in defining $\bar{f}$: we must set

$$\bar{f}(r, s) = f(r) \cdot f(s)^{-1}. \tag{7}$$

Note that this makes sense: by assumption the map f must be injective, so $s \neq 0$ ensures that $f(s) \neq 0$. This proves uniqueness. To prove existence, we first check that (7) gives a well-defined map of sets:

$$(r_1, s_1) \sim (r_2, s_2) \iff r_1 \cdot s_2 = r_2 \cdot s_1$$

$$\implies f(r_1) \cdot f(s_2) = f(r_2) \cdot f(s_1) \implies f(r_1) \cdot f(s_1)^{-1} = f(r_2) \cdot f(s_2)^{-1}.$$

Finally we must check that $\bar{f}$ is a ring homomorphism. For this we take representatives with a common denominator, and observe that

$$\bar{f}((r_1, s) + (r_2, s)) = \bar{f}(r_1 + r_2, s) = f(r_1 + r_2) \cdot f(s)^{-1}$$

$$= f(r_1) \cdot f(s^{-1}) + f(r_2) \cdot f(s)^{-1} = \bar{f}(r_1, s) + \bar{f}(r_2, s).$$

Similarly

$$\bar{f}((r_1, s_1) \cdot (r_2, s_2)) = \bar{f}(r_1 \cdot r_2, s_1 \cdot s_2) = f(r_1) \cdot f(r_2) \cdot [f(s_1) \cdot f(s_2)]^{-1}$$

$$= \bar{f}(r_1, s_1) \cdot \bar{f}(r_2, s_2).$$

Finally $\bar{f}(1, 1) = f(1) \cdot f(1)^{-1} = 1$. This completes the proof. □

19. Function fields

Let us now consider an affine variety $X \subset \mathbb{A}^n_k$. The co-ordinate ring $k[X]$ is an integral domain, so we can consider the corresponding field of fractions.

Definition 19.1. The *function field* $k(X)$ of an affine variety $X \subset \mathbb{A}^n_k$ is the field of fractions of the co-ordinate ring $k[X]$. The elements of $k(X)$ are called *rational functions* on X.

Thus a rational function $f \in k(X)$ is an equivalence class of pairs (g, h) with $g, h \in k[X]$ and $h \neq 0$. We usually write such an element as a fraction $f = g/h$. Note that there is a partially-defined function $X \dashrightarrow k$ given by setting

$$f(a_1, \cdots, a_n) = \frac{g(a_1, \cdots, a_n)}{h(a_1, \cdots, a_n)}$$

for points $(a_1, \cdots, a_n) \in X$ such that $h(a_1, \cdots, a_n) \neq 0$.

Definition 19.2. We say that the rational function f is *regular* at a point $(a_1, \cdots, a_n) \in X$ if there is a representative (g, h) for f such that $h(a_1, \cdots, a_n) \neq 0$. The *domain of definition* of f is the set of points $U \subset X$ at which f is regular. It is denoted $\operatorname{dom}(f)$.

Note that f gives a well-defined function $f\colon \operatorname{dom}(f) \to k$. Indeed, for any point $(a_1, \cdots, a_n) \in \operatorname{dom}(f)$ we can find a representative $f = g/h$ with $h(a_1, \cdots, a_n) \neq 0$ and hence define $f(a_1, \cdots, a_n)$ as above. Moreover, if we have two such representatives, then the equivalence relation used in defining $k(V)$ ensures that they give the same value for f at the point $(a_1, \cdots, a_n)$. When we use the notation

$$f\colon X \dashrightarrow k$$

we mean that f is an element of the function field $k(X)$, and hence a well-defined function $f\colon \operatorname{dom}(f) \to k$, where $\operatorname{dom}(f) \subset X$ is the domain of definition of f.

Lemma 19.3. *The domain of definition of a rational function $f \in k(X)$ is an open subset of X.*

Proof. Take a representative $f = g/h$ of f. It is well-defined on the open subset which is the complement of the algebraic subset $V(h) \subset X$. The domain of definition of f is the union of these open subsets for all representatives of f. This is therefore an open subset of X. □

Example 19.4. Consider the algebraic subset $X = V(y^2 - x^3) \subset \mathbb{A}^2_{\mathbb{C}}$. The rational function $f = y/x$ is well-defined everywhere except at the origin $(0, 0) \in X$. Suppose we take some other representative $f = p/q$ with $p, q \in k[X]$. Thus $q(x, y) \cdot y = x \cdot p(x, y)$. Suppose $q(0, 0) \neq 0$ so that $q(x, y)$ has nonzero constant term. Then $x \cdot p(x, y)$ has a nonzero coefficient of y which is impossible. Hence no representative of f is well-defined at the origin. Hence $\operatorname{dom}(f)$ is the subset $X \setminus \{(0, 0)\}$, which is open because it is the complement of the point $\{0, 0\} = V(x, y)$.

Example 19.5. Consider the algebraic subset $X = V(xy - zw) \subset \mathbb{A}^4_{\mathbb{C}}$.

Note that the pairs (z, x) and (y, w) define the same element of $k(X)$ because $xy - zw = 0 \in k[X]$. Hence the two expressions z/x and y/w define the same element of the function field $k(X)$. This rational function $X \dashrightarrow k$ is regular on the open subset $X \setminus \{(0, y, z, 0)\}$ because at any other point of X either x or w is nonzero. Note that the particular representative z/x is only well-defined on the smaller open subset $X \setminus \{(0, y, z, w)\}$. Thus to get the full domain of definition of the rational function one may need to consider various different representatives.

A rational function $f \in k(X)$ is called *regular* if its domain of definition $\mathrm{dom}(f)$ is the whole of X.

Proposition 19.6. *Assume that k is algebraically closed. Then a rational function $f \in k(X)$ is regular iff it lies in the subring $k[X] \subset k(X)$.*

Proof. Certainly if $f \in k[X]$ then f is regular since the denominator is 1. Conversely, assume f is regular, and let $I \subset k[X]$ be the set of elements occuring as denominators in some representative for f. More precisely

$$I = \{0\} \cup \{h \in k[X] \setminus \{0\} : f = g/h \text{ with } g \in k[X]\}.$$

It is not quite clear that this is closed under addition. Suppose $f = g_1/h_1$ and$f = g_2/h_2$ so that $h_2 g_1 = h_1 g_2$. Then we claim that also $f = (g_1 + g_2)/(h_1 + h_2)$. Indeed

$$(h_1 + h_2)g_1 = h_1 g_1 + h_2 g_1 = h_1 g_1 + h_1 g_2 = h_1(g_1 + g_2).$$

Thus I is an ideal. Consider the closed subset $V(I) \subset X$. At points of this set, all denominators for f vanish, so f is not well-defined. But f is assumed to be regular, so $V(I) = \phi$. By the Nullstellensatz it follows that $\sqrt{I} = I(\emptyset) = k[X]$. It follows that $I = k[X]$ since $1 \in \sqrt{I} \implies 1^n \in I \implies 1 \in I$. Thus 1 occurs as a denominator of f, which is to say f has a representative $f = g/1$. This is precisely the statement that $f \in k[X]$. □

For this reason, polynomial maps $f \colon X \to k$ are also called *regular maps.*

Department of Pure Mathematics, The University of Sheffield
E-mail address: paul.johnson@shef.ac.uk