MSM203a: Polynomials and rings
Chapter 3: Integral domains and fields

Richard Kaye

Autumn 2013*

Note: These printed handouts are intended to supplement the material provided in lectures. They are not sufficient on their own.

This section of the notes is about the nice rings and (where possible) how they can be made even nicer. *In this chapter we make a blanket assumption that all rings considered are commutative rings with 1.* In many cases the theorems require the additional assumption that the ring is a domain or a field. Where needed this will be stated.

1 Polynomial division

Definition 1.1. In a commutative ring R with 1, given $n \geqslant 1$ in $\mathbb{N}$ and $a \in R$ we write na for $a + a + \cdots + a$ (where there are n as here), n for $1 + 1 + \cdots + 1$ (where there are n 1s here), and a^n for $a \cdot a \cdot \cdots \cdot a$ (with n as). Also, $0a$ is 0, a^0 is 1, and $(-n)a$ is $-(na)$, and a^{-n} is $(a^{-1})^n$ when a^{-1} exists.

We have seen how to add, subtract and multiply polynomials. The next result says how one can divide polynomials.

Theorem 1.2 (Polynomial Division)**.** Let $p(X)$, $s(X)$ be polynomials over a commutative ring with 1, R, and let $n = \deg p(X)$ and $l = \deg s(X)$. Suppose $s(X) = s_l X^l + \cdots + s_0$ is not identically zero and $s_l \neq 0$. Then there are polynomials $q(X), r(X)$ such that $s_l^{n-l} p(X) = q(X)s(X) + r(X)$ and the degree of $r(X)$ is less than that of $s(X)$.

Proof. Take $q(X), r(X)$ and k such that

$$s_l^{n-k} p(X) = q(X)s(X) + r(X),$$

and $\deg r(X) = k$ is as small as possible. That this can always be done is clear from $1p(X) = 0(X)s(X) + p(X)$, where $0(X)$ is the zero polynomial and s_l^0 is the 1 of R. Thus we can find such $q(X), r(X)$ with $\deg r(X) = n$. So there is some least $k \leqslant n$ for which there are $q(X), r(X)$ satisfying the equation above with $\deg r(X) = k$. (Remember: we defined the degree of the zero polynomial to be -1, so $k = -1$ if the above equation can be satisfied with $r(x) = 0$.)

We have to show that for this least k we have $k < l$.

Assume instead that $k \geqslant l$. Write $r(X) = r_k X^k + \cdots + r_0$, where r_k is nonzero. Then we may write

$$s_l r(X) = r_k s(X) X^{k-l} + (s_l r(X) - r_k s(X) X^{k-l}).$$

*Version 2.1b of 2013-11-05

The polynomial $s_l r(X) - r_k s(X) X^{k-l}$ has degree at most k, since we assumed $k \geqslant l$ and $s(X)$ has degree l. But in fact coefficient of X^k in it is $s_l r_k - r_k s_l = 0$. So the polynomial $s_l r(X) - r_k S(X) X^{k-l}$ has in fact degree strictly less than k. This means

$$s_l s_l^{n-k} p(X) = s_l q(X) s(X) + r_k s(X) X^{k-l} + (s_l r(X) - r_k S(X) X^{k-l})$$

so

$$s_l^{n-k+1} p(X) = (s_l q(X) + r_k X^{k-l}) s(X) + (s_l r(X) - r_k s(X) X^{k-l}),$$

contradicting our choice of $q(X), r(X), a, k$. So $k < l$ as required. □

In practice, the quotient and remainder given by the above theorem are calculated by the usual algorithm for polynomial devision. In fact the above argument is just a proof built about the usual algorithm. At the key step in the argument we replaced $r(x)$ by $s_l r(X) - r_k s(X) X^{k-l}$. This is because we had already found $s_l^{n-k} p(X) = q(X)s(X) + r(X)$ so we need to look at the terms of highest degree in $s(x)$ and $r(x)$. This is the term of highest degree in $r_k s(X) X^{k-l}$ so $r_k X^{k-l}$ is the next term we write down in the quotient. The proof is complicated slightly by the fact that the leading term of $s(X)$ may not be 1, but the details are similar.

The easiest cases of the division algorithm are when $q(X)$ is a monic polynomial, i.e. has leading coefficient 1. In this case the coefficient $s_l = 1$ so the conclusion is there are $q(x)$ and $r(X)$ with $p(X) = q(X)s(X) + r(X)$. Almost as easy is the case when R is a field, for then s_l^{-1} exists and the same conclusion holds.

The problem case is when $s_l^{n-k} = 0$, which might happen in a ring that is not an integral domain. For completeness' sake we briefly look at this now.

Definition 1.3. An element $s \in R$ is *nilpotent* if $a^n = 0$ for some positive $n \in \mathbb{N}$.

Exercise 1.4. If R is an integral domain show that no nonzero element of R is nilpotent.

Exercise 1.5. If R is an commutative ring with 1 and is not the zero ring, show that 1 is not nilpotent.

Corollary 1.6. *Let $p(X)$, $s(X)$ be polynomials over a ring R, and let $n = \deg p(X)$ and $l = \deg s(X)$. Suppose $s(X) = s_l X^l + \cdots + s_0$ is not identically zero and $s_l \neq 0$. Suppose also that* either *R is an integral domain* or *(more generally) s_l is not nilpotent in R. Then there are polynomials $q(X), r(X)$ and an element $a \neq 0$ of R such that $ap(X) = q(X)s(X) + r(X)$ and the degree of $r(X)$ is less than that of $s(X)$.*

Integral domains have no zero divisors, and consequently allow one to solve equations of the form $x(y - z) = 0$ in the way one expects: either $x = 0$ or else $y = z$. (This is precisely the cancellation law for multiplication discussed earlier.)

Theorem 1.7. *Let $p(X)$ be a polynomial over an integral domain R with degree $n > 0$. Then the polynomial equation $p(x) = 0$ has at most $n = \deg(p(X))$ roots in R.*

Proof. By induction on n. For $n = 1$ the polynomial is $a + bX$ with $b \neq 0$. So it has as root any $x \in R$ satisfying $a + bx = 0$. There is at most one such x since if $a + bx_1 = a + bx_2 = 0$ then $b(x_1 - x_2) = 0$ and as $b \neq 0$ and R has no zero divisors we must have $x_1 = x_2$.

Assume the result is true for $n \geqslant 1$ and consider a polynomial $p(X)$ of degree $n + 1$. Suppose $p(X) = 0$ has a root $a \in R$ (for if it has no roots we are finished), i.e. suppose $p(a) = 0$. Then $p(X) = q(X)(X - a) + b$ for some $q(X)$ and b (a polynomial of degree less than 1, hence a constant) by Polynomial Division applied to $p(X)$ and $(X - a)$. But since $p(a) = 0$ we have $0 = p(a) = q(a)(a - a) + b = b$, so $b = 0$ so $p(X) = q(X)(X - a)$, and obviously $q(X)$ has degree n. By the induction hypothesis there are at most n roots of $q(x) = 0$. Now if $d \in R$ is any root of $p(X)$, i.e. $p(d) = 0$, we have $0 = q(d)(d - a)$ so either $d = a$ or $q(d) = 0$ since R has no zero divisors. It follows that the only roots of $p(X) = 0$ are a and the (at most n) roots of $q(X)$. □

2 Characteristic

Definition 2.1. Let R be an integral domain. The *characteristic* of R, char$R$, is the least positive $k \in \mathbb{N}$ such that a sum of k ones, $1 + 1 + \cdots + 1$, is 0. By convention, if there is no such k we write char$R = 0$.

According to this definition, the characteristic of the zero ring $\{0\}$ is 1. This is the only way this definition can come out to be 1. Some people don't allow $\{0\}$ to be an integral domain. In any case, we shall ignore this uninteresting case.

Proposition 2.2. The characteristic of an integral domain R other than $\{0\}$ is either 0 or a prime number p.

Proof. Let $k = \mathrm{char}R > 0$, so $\overbrace{1 + 1 + \cdots + 1}^{k} = 0$ in R (k 1s). If $k = 1$ then $1 = 0$ in R so R is the zero ring. If $k > 1$ is not prime, $k = uv$ say, with $1 < u, v < k$. Then

$$\overbrace{1 + \cdots + 1}^{k} = (\overbrace{1 + \cdots + 1}^{u})(\overbrace{1 + \cdots + 1}^{v})$$

by straightforward expansion using distributivity (see the next exercise), so one of $(\overbrace{1 + \cdots + 1}^{u})$, $(\overbrace{1 + \cdots + 1}^{v})$ is 0 as R is an integral domain. But this contradicts the choice of k. □

Exercise 2.3. Let R be a commutative ring with 1 and $k \in \mathbb{N}$ positive. Prove by induction on n that

$$(\overbrace{1 + \cdots + 1}^{k}) \cdot (\overbrace{1 + \cdots + 1}^{n}) = \overbrace{1 + \cdots + 1}^{nk}$$

for all positive $n \in \mathbb{N}$.

Theorem 2.4. An integral domain R other than $\{0\}$ either contains a copy of $\mathbb{Z}$ as a subring or else contains a copy of $\mathbb{Z}/p\mathbb{Z}$ as a subring for some prime number p.

Remark 2.5. For each prime number p the ring $\mathbb{Z}/p\mathbb{Z}$ is actually a field and to commemorate the fact that it is a field we write it as $\mathbb{F}_p$.

Theorem 2.6. Let F be a field and $p = \mathrm{char}F$, and suppose $p \neq 0$. Then either F is infinite or else $|F|$ is a power of p.

Sketch proof. F is a field. The characteristic subfield of F, $\mathbb{F}_p = \{(\overbrace{1 + \cdots + 1}^{u}) : 0 \leqslant u < p\}$, is also a field. It turns out by simple axiom-checking that F is a $\mathbb{F}_p$-vector space. (You didn't study vector spaces over $\mathbb{F}_p$ last year, but all the theory works out just the same.) In particular F is either an infinite dimensional $\mathbb{F}_p$-space or else has finite dimension k. In the latter case $F = \{\lambda_1 e_1 + \cdots + \lambda_k e_k : \lambda_i \in \mathbb{F}_p\}$ for some basis $\{e_1, \ldots, e_k\} \subseteq F$, and this representation of the elements of F is unique (different λs give different elements). So there are p^k elements of F. □

We have seen how a integral domain R of characteristic $p > 1$ has as a subfield a copy of $\mathbb{F}_p$, the *characteristic subfield* formed from the set of all $\overbrace{1+1+\cdots+1}^{n}$ $(0 \leqslant n < p)$ in F, and that one of characteristic 0 contains a copy of the integers. In general this is all one can say, but a *field* F will either contain a copy of $\mathbb{F}_p$ or else a copy of the *rationals*, the characteristic subfield in the case when the characteristic is 0.

Proposition 2.7. A field of characteristic 0 contains a copy of $\mathbb{Q}$.

Proof. We already have a copy $\mathbb{Z}'$ of the integers in F, the set of elements

$$\{\overbrace{1+1+\cdots+1}^{n} : n \in \mathbb{N}^+\} \cup \{0\} \cup \{\overbrace{-1+-1+\cdots+-1}^{n} : n \in \mathbb{N}^+\}.$$

with the isomorphism $n \mapsto n' = \overbrace{1+1+\cdots+1}^{n}$ from $\mathbb{Z}$ to $\mathbb{Z}'$. For each $q = n/m \in \mathbb{Q}$ we map it to $q' = n'(m')^{-1}$ in F. This defines a copy of $\mathbb{Q}$ in F, since the map $q \mapsto q'$ just defined is easily checked to be a homomorphism, and is injective since if $n_1'(m_1')^{-1} = n_2'(m_2')^{-1}$ then $n_1'm_2' = n_2'm_1'$ so $n_1/m_1 = n_2/m_2$ in $\mathbb{Q}$. □

Remark 2.8. By all means continue to use the notation with over-braces as here. However note that Definition 1.1 gives a shorter alternative as long as you don't get muddled between elements of R and elements of $\mathbb{Z}$.

3 Units

Another way to find parts of a ring that 'looks nice' is to look at the elements that have multiplicative inverse.

Definition 3.1. In a ring R, a *unit* is an element $x \in R$ which has a multiplicative inverse y such that $xy = yx = 1$.

For example, 1 is always a unit, but there may be other units besides this. As we saw before, if a multiplicative inverse of x exists then is unique. This multiplicative inverse when it exists is denoted x^{-1}.

Definition 3.2. In a ring R, the set of units of R is denoted R^* or $R^\times$. It is a group when considered as a set with the multiplication operation.

There are many nice theorems about $(\mathbb{Z}/n\mathbb{Z})^\times$ for various $n \in \mathbb{Z}$, several of them rather difficult, and all of them properly part of number theory. Some of these results can be generalised to other rings, but this is out of the scope of this course.

Unfortunately $R^\times \cup \{0\}$ is rarely a ring itself since it is not usually closed under addition.

Exercise 3.3. Show that if $a \in R$ is nilpotent then $1 - a$ and $1 + a$ are units. Hint: show

$$(1-a)(1+a+a^2+\cdots+a^{n-1}) = 1 - a^n$$

and

$$(1+a)(1-a+a^2+\cdots+(-a)^{n-1}) = 1 + a^n$$

for some n.

4 Maximal and prime ideals

This section looks in more detail at properties of ideals, and principal ideals in particular, looking to generalise the idea of 'prime' in an arbitrary ring.

To start with, we collect together some easy properties of principal ideals. Note that, in a ring R, we write $b|a$ (b divides a, or b is a factor of a) to mean there is $x \in R$ with $a = bx$.

Proposition 4.1. Let $a, b \in R$.

(a) The ideal (a) is the smallest ideal containing a. So if $a \in I \lhd R$ then $(a) \subseteq I$.

(b) We have $(a) \subseteq (b)$ iff $b|a$.

(c) The principal ideal generated by 0 is $(0) = \{0\}$.

(d) The principal ideal generated by 1 is $(1) = R$. More generally $(u) = R$ whenever u is a unit.

(e) In an integral domain, $(a) = (b)$ iff $a = b = 0$ or there is a unit $u \in R$ with $au = b$.

The following is a sort of converse to the definition of an ideal.

Definition 4.2. An ideal $I \lhd R$ is *prime* if $I \neq R$ and whenever $x, y \in R$ and $xy \in I$ then at least one of x, y is in I.

Prime ideals are exactly what we need if we want integral domains.

Theorem 4.3. Let R be a ring and $I \lhd R$ an ideal not equal to the whole of R. Then R/I is an integral domain if and only if I is prime in R.

Corollary 4.4. A ring R is an integral domain if and only if $\{0\}$ is a prime ideal.

In more advanced work one does not normally talk about prime *numbers* in a ring R but rather about prime *ideals* of R. If one really wants to say whether $a \in R$ is prime one looks at the ideal (a) it generates (ruling out the trivial cases). Thus our notion of prime ideal immediately yields a definition of prime number.

Definition 4.5. In an arbitrary ring R, an element $a \in R$ is *prime* if a is not zero nor a unit and (a) is a prime ideal.

Proposition 4.6. $a \in R$ is prime if it is nonzero, not a unit and the following holds for all $x, y \in R$:

$$a|xy \;\Rightarrow\; a|x \text{ or } a|y.$$

Example 4.7. The primes in $\mathbb{Z}$ are the familiar numbers $p \in \mathbb{Z}$ that you have up to now called 'prime', and their additive inverses. Note however that the new definition of 'prime' is not the same as the familiar one. (The familiar definition is in some sense 'wrong' and will have to be unlearnt. See also the notions of 'maximal' and 'irreducible' below.)

A similar story applies to quotients R/I that are fields.

Definition 4.8. An ideal $I \lhd R$ is *maximal* if $I \neq R$ and whenever $I \subseteq J \subseteq R$ with $J \lhd R$ we have $I = J$ or $J = R$.

Theorem 4.9. If R is a ring and $I \lhd R$ then R/I is a field if and only if I is maximal in R.

Corollary 4.10. Let R be a ring. Then R is a field if and only if $\{0\} \lhd R$ is maximal, i.e. if and only if there are no proper ideals.

Corollary 4.11. Let R be a ring and $I \lhd R$. Then if I is maximal it is prime.

The following result requires a theorem from set theory called Zorn's Lemma, so will not be proved in this course.

Fact 4.12. If R is a ring and $I \lhd R$ is a proper ideal then there is a maximal ideal $J \lhd R$ containing I. Hence in particular any commutative ring with one, other than the zero ring, has a quotient that is a field.

As for primes, we can look at the corresponding properties of elements of the ring.

Definition 4.13. Let R be a ring and $a \in R$. Then a is *maximal* if it is nonzero, not a unit and (a) is a maximal ideal.

Exercise 4.14. Show that for $\mathbb{Z}$ the notion of being 'maximal' coincides with being 'prime'.

For general rings R, 'maximal' and 'prime' are not the same, and (even worse!) neither corresponds to the familiar school definition of 'prime number' which is

Definition 4.15. An element $a \in R$ is *irreducible* if a is nonzero and not a unit, and whenever $a = uv$ for some $u, v \in R$ then one of u, v is a unit.

The definition of irreducible elements is particularly useful in the case of the ring $F[X]$ of polynomials over a field.

Proposition 4.16. A polynomial $p(X) \in F[X]$, where F is a field, is irreducible in $F[X]$ if it has degree at least 1 and it is impossible to write $p(X) = q(X)r(X)$ for two other polynomials $q(X), r(X) \in F[X]$ both of *smaller degree* than $p(X)$.

Of all the notions of 'prime' studied here, 'irreducible' is the weakest.

Proposition 4.17. If a is prime in an integral domain R then it is irreducible.

There is no converse in general to the previous proposition, but there is a converse in an important family of cases, which is the subject of the next section.

Example 4.18. (A ring R in which 'prime', 'maximal' and 'irreducible' are all different.)

5 Principal ideal domains

After the disappointment of the last section where we found three incompatible meanings to the word 'prime', we introduce a useful family where the three notions coincide.

Definition 5.1. A ring R is a *principal ideal domain* (PID) if it is an integral domain and each ideal $I \lhd R$ is a principal ideal (a) for some $a \in R$.

Example 5.2. The ring $\mathbb{Z}$ is a PID. (Hint: if $I \lhd \mathbb{Z}$ consider the least positive element of I.)

Example 5.3. If F is a field, the ring $F[X]$ of polynomials over F is a PID. (Hint: if $I \lhd F[X]$ consider the monic polynomial of least positive degree in I.)

Theorem 5.4. If the ring R is a PID then the notions of 'irreducible', 'prime' and 'maximal' coincide. That is, in a PID an ideal (a) is maximal if and only if a is irreducible.

This gives us perhaps the most import family of examples of all. Given a field F we form the polynomial ring $R = F[X]$. By results above, $F[X]$ is a PID. Now suppose $p(X) \in F[X]$ is irreducible. Then the principal ideal $(p(X))$ is maximal and hence $F[X]/(p(X))$ is a field.

The following particular case is very familiar indeed.

Example 5.5. Let $R = \mathbb{R}$ and $p(X) = X^2+1$. Let $I = (p(X))$, the principal ideal generated by $p(X)$. Then $R[X]/(p(X))$ is the familiar ring ... [insert name here].

Be aware that this relies on working over a field F in the first place. When R is not a field, $R[X]/(p(X))$ need not be a field even if $p(X)$ is irreducible in $R[X]$.

Example 5.6. Let $R = \mathbb{Z}$ and $p(X) = X^2-2$. Let $I = (p(X))$. Then $R[X]/(p(X))$ is $\mathbb{Z}[\sqrt{2}]$.

6 The field of fractions over an integral domain

This section contains an easy but technical construction. Given an integral domain R we will show how to build a field $F = Q(R)$ that contains R. The construction is exactly like (and generalises) the construction of the rationals $\mathbb{Q}$ from the integers $\mathbb{Z}$.

Given an integral domain R, let

$$Q_0 = \{(r, s) : r, s \in R,\ \ s \neq 0\}.$$

We define an equivalence relation on Q_0 by

$$(r, s) \sim (r', s') \Leftrightarrow rs' = r's.$$

Proposition 6.1. $\sim$ is an equivalence relation on Q_0.

To make $Q_0/{\sim}$ into a ring-like structure we define

$$(r_1, s_1)/{\sim} + (r_2, s_2)/{\sim} = (r_1s_2 + r_2s_1, s_1s_2)/{\sim}$$

and

$$(r_1, s_1)/{\sim} \cdot (r_2, s_2)/{\sim} = (r_1r_2, s_1s_2)/{\sim}.$$

Note that this relies on s_1s_2 being non-zero, so will not work if there are zero divisors.

Proposition 6.2. These are well-defined binary operations on $Q_0/{\sim}$.

Proposition 6.3. With addition and multiplication as just defined, $Q_0/{\sim}$ is a field, and contains a copy of R via the injective homomorphism

$$x \mapsto (x, 1)/{\sim}.$$

Definition 6.4. $Q_0/{\sim}$ is called *the field of fractions of* R, and is written $Q(R)$.

Example 6.5. $\mathbb{Q} = Q(\mathbb{Z})$

Example 6.6. Let F be a field. Then $F[X]$ is an integral domain. Its field of fractions is denoted $F(X)$ and is $Q(F[X])$, consisting of all so-called *rational polynomials* $p(X)/q(X)$ for which $q(X)$ is not the zero polynomial. The object $X = X/1$ is an element of $F(X)$ that is not in F. $F(X)$ is a field containing a copy of F as well as $X = X/1$.

7 Field extensions

A field extension is a pair of fields $F \subseteq K$. Galois theory, which some of you may take later, studies field extensions. We can use what we have learnt to see some of the beginnings of this subject.

Exercise 7.1. Given fields $F \subseteq K$ show that K is an F-vector space. Hence it has a dimension n over F. If this dimension is finite and $|F|$ is finite then $|K| = |F|^n$.

Now consider some $\alpha \in K \setminus F$.

Definition 7.2. $F(\alpha)$ is the smallest subfield of K containing α and F. It is the 'closure' of F, α under $+, \cdot, -, ^{-1}$.

We look here at *simple* field extensions, i.e. ones of the form $F(\alpha)$ over F. There are two possibilities for α.

Definition 7.3. α is *algebraic over F* if there is a polynomial $p(X) \in F[X]$ of positive degree such that $p(\alpha) = 0$. If there is no such polynomial, then α is *transcendental over F*.

Example 7.4. The element i in $\mathbb{C}$ is algebraic over $\mathbb{R}$.

Example 7.5. You may know the theorem (or have heard the result) that $\pi \in \mathbb{R}$ is transcendental over $\mathbb{Q}$.

Exercise 7.6. Let K be any field, and $F = Q(K[X])$, the field of 'rational polynomials' over K, i.e. the field of fractions of $K[X]$. Show that the element $X = X/1$ of F is transcendental over K.

Theorem 7.7. If α is algebraic over F then there is a unique monic polynomial $m_\alpha(X) \in F[X]$ of least positive degree such that $m_\alpha(\alpha) = 0$. Also, for this polynomial we have, for all $p(X) \in F[X]$,

$$p(\alpha) = 0 \;\Rightarrow\; \exists q(X) \in F[X]\ (p(X) = q(X)m_\alpha(X)).$$

Proof. Let $m_\alpha(X) \in F[X]$ be the polynomial of least positive degree such that $m_\alpha(\alpha) = 0$. Then the property given holds by polynomial division as shown in lectures. □

Definition 7.8. The polynomial $m_\alpha(X)$ is unique (because it is monic) and is called the *minimum polynomial of α over F.*

Proposition 7.9. If α is algebraic over F then $m_\alpha(X)$ is irreducible.

Corollary 7.10. If α is algebraic over F then $m_\alpha(X)$ is maximal, and hence $F[X]/(m_\alpha(X))$ is a field.

Theorem 7.11. If α is algebraic over F then $F(\alpha) \cong F[X]/(m_\alpha(X))$.

The other case is described by the field of fractions construction.

Theorem 7.12. If α is transcendental over F then $F(\alpha) \cong F(X) = Q(F[X])$.