Rings and fields II

Abdennour Kitouni, Anatoliy Malyarenko, Sergei Silvestrov

December 3, 2015

Abstract

Contents of the lecture.

- ☞ The field of quotients of an integral domain.
- ☞ Rings of polynomials.
- ☞ Factorisation of polynomials over a field.

The field of quotients

Every subring of an integral domain is itself an integral domain. Since fields are integral domains, it follows that every subring of a field is an integral domain. The converse of this is true, and it is much more interesting: *Every integral domain is a subring of a field*.

The proof of the next theorem is a straightforward generalisation of the usual construction of the field of rational numbers $\mathbb{Q}$ from the integral domain of integers $\mathbb{Z}$.

Theorem 1. *If R is an integral domain, then there is a field F (the **field of quotients**) containing R as a subring. Moreover, F can be chosen so that, for each $f \in F$, there are $a, b \in R$ with $b \neq 0$ and $f = ab^{-1}$.*

Proof. Let $X = \{\,(a,b) \in R\colon\ b \neq 0\,\}$ and define a relation $\equiv$ on X by $(a,b) \equiv (c,d)$ if $ad = bc$. We claim that $\equiv$ is an equivalence relation. Verifications of reflexivity and symmetry are straightforward; here is the proof of transitivity. If $(a,b) \equiv (c,d)$ and $(c,d) \equiv (e,f)$, then $ad = bc$ and $cf = de$. But $ad = bc$ gives $adf = b(cf) = bde$. Cancelling d, which is nonzero, gives $af = be$; that is, $(a,b) \equiv (e,f)$.

Denote the equivalence class of (a,b) by $[a,b]$, define F as the set of all equivalence classes $[a,b]$, and equip F with the following addition and multiplication:

$$[a,b] + [c,d] = [ad + bc, bd]$$

and

$$[a,b][c,d] = [ac, bd].$$

First, since $b \neq 0$ and $d \neq 0$, we have $bd \neq 0$, because R is an integral domain, and so the formulas make sense. Let us show that addition is well-defined. If $[a,b] = [a',b']$ (that is, $ab' = a'b$) and $[c,d] = [c',d']$ (that is, $cd' = c'd$), then we must show that $[ad+bc, bd] = [a'd' + b'c', b'd']$. But this is true:

$$(ad+bc)b'd' = ab'dd' + bb'cd' = a'bdd' + bb'c'd = (a'd' + b'c')bd.$$

A similar argument shows that multiplication is well-defined.

The verification that F is a commutative ring is now routine: The zero element is $[0,1]$, the one is $[1,1]$, and the additive inverse of $[a,b]$ is $[-a,b]$. It is easy to see that the family $R' = \{\,[a,1]\colon\ a \in R\,\}$ is a subring of F, and we identify $a \in R$ with $[a,1] \in R'$.

To see that F is a field, observe that if $[a,b] \neq [0,1]$, then $a \neq 0$, and the inverse of $[a,b]$ is $[b,a]$.

Finally, if $b \neq 0$, then $[1,b] = [b,1]^{-1}$, and so $[a,b] = [a,1][b,1]^{-1}$. □

The set $\mathbf{R}[\mathbf{x}]$

Definition 1. Let R be a ring. A **sequence** σ in R is a function $\sigma\colon\ \mathbb{Z}^+ \cup \{0\} \mapsto R$.

In what follows, we write a_i instead of $\sigma(i)$, and denote a sequence by $\sigma = (a_0, a_1, \ldots, a_m, \ldots)$.

Definition 2. A sequence σ belongs to the **set** $\mathbf{R}[\mathbf{x}]$ if it has the form

$$\sigma = (a_0, a_1, \ldots, a_m, 0, 0, \ldots).$$

Definition 3. An **indeterminate** x is the following element of the set $R[x]$:

$$x = (0,1,0,0,\dots).$$

Ring operations in $\mathbf{R[x]}$

Define addition and multiplication in the set $R[x]$ as follows: if $\sigma = (a_0, a_1, \dots)$ and $\tau = (b_0, b_1, \dots)$, then

$$\sigma + \tau = (a_0 + b_0, a_1 + b_1, \dots)$$

and

$$\sigma\tau = (c_0, c_1, \dots),$$

where $c_k = \sum_{i=0}^{k} a_i b_{k-i}$.

Theorem 2. *If R is a ring then $R[x]$ is a ring that contains R as a subring.*

Proof. Verification of the axioms in the definition of a ring is routine. The subset $\{\,(a,0,0,\dots) : a \in R\,\}$ is a subring of $R[x]$ that we identify with R. □

Theorem 3. *If $\sigma = (a_0, a_1, \dots, a_m, 0, 0, \dots)$, then*

$$\sigma = a_0 + a_1 x + \cdots + a_m x^m.$$

Polynomials

Definition 4. Any element $f(x) = a_0 + a_1 x + \cdots + a_m x^m$ of the ring $R[x]$ is called a **polynomial**.

Definition 5. Let $f(x) = a_0 + a_1 x + \cdots + a_m x^m$ be a polynomial. We call a_m the **leading coefficient** of f, we call a_0 the **constant term**, we call m the **degree** of f. A **constant polynomial** is either the zero polynomial or a polynomial of degree 0. Polynomials of degree 1, namely, $a + bx$ with $b \neq 0$, are called **linear**, polynomials of degree 2 are **quadratic**, degree 3's are **cubic**, then **quartics**, **quintics**, and so on. A polynomial with $a_m = 1$ is **monic**.

The **zero polynomial** 0 does not have a degree because it has no nonzero coefficients. We choose not to assign a degree to 0 because it is often a genuinely different case that must be dealt with separately.

Definition 6. Let F be a field. The field of quotients of $F[x]$, denoted by $F(x)$, is called the **field of rational functions** over F.

The evaluation homomorphisms

Definition 7. Let F be a subfield of a field E, let $\gamma \in E$, and let $x \in F[x]$ be an indeterminate. The map $\varphi_\gamma \colon F[x] \to E$ defined by

$$\varphi_\gamma(a_0 + a_1 x + \cdots + a_m x^m) = a_0 + a_1\gamma + \cdots + a_m\gamma^m$$

is called an **evaluation at** γ.

Theorem 4. *φ_γ is a homomorphism of $F[x]$ into E. Also, $\varphi_\gamma(x) = \gamma$ and φ_γ maps F isomorphically by the identity map, that is, $\varphi_\gamma(a) = a$ for $a \in F$.*

Definition 8. Let F be a subfield of a field E, let $f \in F[x]$, and let $\gamma \in E$. Let $f(\gamma)$ be the following element in E:

$$f(\gamma) = \varphi_\gamma(f(x)) \in E.$$

γ is a **zero** or a **root** of $f(x)$, if $f(\gamma) = 0$.

The division algorithm

We are now going to see that, when F is a field, virtually all the familiar theorems proved for $\mathbb{Z}$ have polynomial analogs in $F[x]$; moreover, we shall see that the familiar proofs can be translated into proofs here.

The division algorithm for polynomials with coefficients in a field says that long division is possible.

Theorem 5. *Assume that F is a field and that $f(x)$, $g(x) \in F[x]$ with $f(x) \neq 0$. Then there are unique polynomials $q(x)$, $r(x) \in F[x]$ with*

$$g(x) = q(x)f(x) + r(x)$$

and either $r(x) = 0$ or the degree of $r(x)$ is less that the degree of $f(x)$.

Proof. We first prove the existence of such q and r. If $g = qf$ for some q, define the remainder $r = 0$, and we are done. Otherwise consider all (necessarily nonzero) polynomials of the form $g - qf$ as q varies over $F[x]$. Let $r = g - qf$ has the least degree among all such polynomials.

Since $g = qf + r$, it suffices to show that the degree of r is less than the degree of f. Write $f(x) = a_n x^n + \cdots + a_1 x + a_0$ and $r(x) = b_m x^m + \cdots + 1_1 x + b_0$. Now $a_n \neq 0$ implies that a_n is a unit, because F is a field, and so a_n^{-1} exists in F. If the degree of r is greater than or equal to the degree of f, define

$$h(x) = r(x) - \frac{b_m x^m}{a_n x^n} f(x).$$

Note that either $h = 0$ or the degree of h is less than the degree of r. If $h = 0$, then $r = [(b_m x^m)/(a_n x^n)]f$ and

$$\begin{aligned} g &= qf + r \\ &= qf + \frac{b_m x^m}{a_m x^m} f \\ &= \left[q + \frac{b_m x^m}{a_n x^n} \right] f, \end{aligned}$$

contradicting our assumption. If $h \neq 0$, then the degree of h is less than the degree of r and

$$g - qf = r = h + \frac{b_m x^m}{a_n x^n} f.$$

Thus, $g - [q + (b_m x^m)/(a_n x^n)]f = h$, contradicting r being a polynomial of least degree having this form. Therefore, the degree of r is less that the degree of f.

To prove uniqueness of $q(x)$ and $r(x)$, assume that $g = q'f + r'$ where the degree of r' is less that the degree of f. Then

$$(q - q')f = r' - r.$$

If $r' \neq r$, then each side has a degree. But the degree of $(q - q')f$ is greater or equal than the degree of f, while the degree of $r' - r$ is less than or equal to the maximum of the degrees of r and r', which is less than the degree of f, a contradiction. Hence, $r = r'$ and $(q - q')f = 0$. As $F[x]$ is a domain and $f \neq 0$, it follows that $q - q' = 0$ and $q = q'$. □

The factor theorem

Theorem 6. *If $f(x) \in F[x]$, where F is a field, then a is a root of $f(x)$ in F if and only if $x - a$ divides $f(x)$ in $F[x]$.*

Proof. If a is a root of $f(x)$ in F, then $f(a) = 0$. The division algorithm gives

$$f(x) = q(x)(x - a) + r;$$

the remainder r is a constant because $x-a$ has degree 1. Now evaluate

$$f(a) = q(a)(a-a)+r,$$

and so $r = f(a) = 0$ and $f(x) = q(x)(x-a)$. Conversely, if $f(x) = q(x)(x-a)$, then evaluating at a gives $f(a) = q(a)(a-a) = 0$. □

Zeroes of polynomials

Theorem 7. *Let F be a field and let $f(x) \in F[x]$. If $f(x)$ has degree n, then $f(x)$ has at most n roots in F.*

Proof. We prove the statement by induction on $n \geq 0$. If $n = 0$, then $f(x)$ is a nonzero constant, and so the number of its roots in F is zero. Now let $n > 0$. If $f(x)$ has no roots in F, then we are done, for $0 \leq n$. Otherwise, we may assume that there is $a \in F$ with a a root of $f(x)$; hence, by the factor theorem,

$$f(x) = q(x)(x-a);$$

moreover, $q(x) \in F[x]$ has degree $n-1$. If there is a root $b \in F$ with $b \neq a$, then

$$0 = f(b) = q(b)(b-a).$$

Since $b-a \neq 0$, we have $q(b) = 0$ (because F is a field, hence is an integral domain), so that b is a root of $q(x)$. Now the degree of q is $n-1$, so that the inductive hypothesis says that $q(x)$ has at most $n-1$ roots in F. Therefore, $f(x)$ has at most n roots in F. □

Corollary 1. *If F is a field and G is a finite subgroup of the multiplicative group $F^\times$, then G is cyclic. In particular, if F itself is finite, then $F^\times$ is cyclic.*

Proof. Let d be a divisor of $|G|$. If there are two subgroups of G of order d, say, S and T, then $|S \cup T| > d$. But each $a \in S \cup T$ satisfies $a^d = 1$, by Lagrange's theorem, and hence it is a root of $x^d - 1$. This contradicts Theorem 7, for this polynomial has too many roots in F. Thus, G is cyclic. □

Irreducible polynomials

Definition 9. A polynomial $p(x)$ is called **irreducible** if its degree is greater than or equal to one and there is no factorisation in $F[x]$ of the form $p(x) = g(x)h(x)$ in which both factors have degree smaller than the degree of p.

For example, $p(x) = x^2 + 1$ is irreducible in $\mathbb{R}[x]$, but it factors as $(x+i)(x-i)$ in $\mathbb{C}[x]$.

Theorem 8. *Let F be a field and let $f(x)$ be a quadratic or cubic polynomial. Then $f(x)$ is irreducible in $F[x]$ if and only if $f(x)$ does not have a root in F.*

Proof. If $f(x) = g(x)h(x)$ and neither g nor h is constant, then the degree of f is equal to the sum of degrees of g and h. This implies that at least one of the factors has degree 1. □

Example: the irreducible polynomials of small degree in $\mathbb{Z}_2[x]$

Example 1. As always, the linear polynomials x and $x+1$ are irreducible.

There are four quadratics: x^2, x^2+x, x^2+1, x^2+x+1. Since each of the first three has a root in $\mathbb{Z}_2$, there is only one irreducible quadratic.

There are eight cubics, of which four are reducible because their constant term is 0. The remaining polynomials are

$$x^3+1; \quad x^3+x+1; \quad x^3+x^2+1; \quad x^3+x^2+x+1.$$

Since 1 is a root of the first and fourth, the middle two are the only irreducible cubics.

There are 16 quartics, of which eight are reducible because their constant term is 0. Of the eight with nonzero constant term, those having an even number of nonzero coefficients have 1 as a root. There are now only four surviving polynomials $f(x)$, and each of them has no roots in $\mathbb{Z}_2$; i.e., they have no linear factors. If $f(x) = g(x)h(x)$, then both $g(x)$ and $h(x)$ must be irreducible quadratics. But there is only one irreducible quadratic, namely, x^2+x+1, and so $(x^2+x+1)^2 = x^4+x^2+1$ is reducible while the other three quartics are irreducible. The following list summarises these observations.

$$x, \quad x+1;$$

$$x^2+x+1;$$

$$x^3+x+1, \quad x^3+x^2+1;$$

$$x^4+x^3+1, \quad x^4+x+1, \quad x^4+x^3+x^2+x+1.$$

The Eisenstein criterion

Theorem 9. *Let $p \in \mathbb{Z}$ be a prime and $f(x) = a_n x^n + \cdots + a_0$ be a polynomial with integer coefficients such that $a_n \not\equiv 0 \pmod{p}$, $a_i \equiv 0 \pmod{p}$ for all $i < n$ and $a_0 \not\equiv 0 \pmod{p^2}$. Then $f(x)$ is irreducible over $\mathbb{Q}$.*

Example 2. If $f = 2x^5 - 6x^3 + 9x^2 - 15$, then the Eisenstein criterion with $p = 3$ shows that f is irreducible in $\mathbb{Q}$.

Lemma 1. *Let F be a field and let $f(x)$, $g(x) \in F[x]$. If $p(x)$ is an irreducible polynomial in $F[x]$, and $p(x)$ divides $f(x)g(x)$, then either $p(x)$ divides $f(x)$ or $p(x)$ divides $g(x)$.*

Unique factorisation

Theorem 10. *If F is a field, then every polynomial $f(x) \in F[x]$ of degree ≥ 1 is a product of a nonzero constant and monic irreducibles. Moreover, if $f(x)$ has two such factorisations*

$$f(x) = ap_1(x)\dots p_m(x) \quad \textit{and} \quad f(x) = bq_1(x)\dots q_n(x),$$

that is, a and b are nonzero constants and the p's and d's are monic irreducibles, then $a = b$, $m = n$, and the q's may be re-indexed so that $q_i = p_i$ for all i.

Proof. We prove the existence of a factorisation for a polynomial $f(x)$ by induction on the degree of f. If the degree of f is equal to 1, then $f(x) = ax + c = a(x + a^{-1}c)$.As every linear polynomial, $x + a^{-1}c$ is irreducible. Assume now that the degree of f is greater than 1. If $f(x)$ is irreducible and its leading coefficient is a, write $f(x) = a(a^{-1}f(x))$; we are done, for $a^{-1}f(x)$ is monic. If $f(x)$ is not irreducible, then $f(x) = g(x)h(x)$, where the degrees of both g and h are less than the degree of f. By the inductive hypothesis, there are factorisations $g(x) = bp_1(x)\dots p_m(x)$ and $h(x) = cq_1(x)\dots q_m(x)$, where the p's and q's are monic irreducibles. It follows that

$$f(x) = (bc)p_1(x)\dots p_m(x)q_1(x)\dots q_n(x)$$

as desired.

We now prove, by induction on $M = \max\{m, n\} \geq 1$, that if there is an equation

$$ap_1(x)\dots p_m(x) = bq_1(x)\dots q_n(x)$$

in which a and b are nonzero constants and the p's and q's are monic irreducibles, then $a = b$, $m = n$, and the q's may be re-indexed so that $q_i = p_i$ for all i. For the base step $M = 1$, the hypothesis

gives a polynomial, call it $g(x)$, with $g(x) = ap_1(x) = bq_1(x)$. Now a is the leading coefficient of $g(x)$, because $p_1(x)$ is monic; similarly, b is the leading coefficient of $g(x)$ because $q_1(x)$ is monic. Therefore, $a = b$, and cancelling gives $p_1(x) = q_1(x)$. For the inductive step, the given equation shows that $p_m(x)$ divides $q_1(x)\dots q_n(x)$. By Lemma 1, there is some i such that $p_m(x)$ divides $q_i(x)$. But $q_i(x)$, being monic irreducible, has no monic divisors other than 1 and itself, so that $q_i(x) = p_m(x)$. Re-indexing, we may assume that $q_n(x) = p_m(x)$. Cancelling this factor, we have $ap_1(x)\dots p_{m-1}(x) = bq_1(x)\dots q_{n-1}(x)$. By the inductive hypothesis, $a = b$, $m-1 = n-1$ (hence $m = n$), and after possible re-indexing, $q_i = p_i$ for all i. □