6.966 Algebra and Computation September 14, 1998

Lecture 2

Lecturer: Madhu Sudan *Scribe: Sofya Raskhodnikova*

1 Primality vs. Irreducibility

Definition 1 *$a \in R$ [1] is irreducible if a is not a unit and $x \cdot y = a$ implies that x or y is a unit.*

Definition 2 *$p \in R$ is prime if $\forall a, b \in R \;\; p \mid ab \Rightarrow p \mid a$ or $p \mid b$.*

The notions of integral domain and Unique Factorization Domain (UFD) do not coincide.

Lemma 1 *Integral domain R is a UFD iff*

1. *Every element in R is expressible as a product of finitely many irreducible elements.*
2. *Every irreducible element is a prime.*

Proof Let $x \in R$ and $x = p_1 p_2 \ldots p_s = q_1 q_2 \ldots q_t$ where $p_1, \ldots, p_s, q_1, \ldots, q_t$ are irreducible. Equivalently,

$$p_1 p_2 \ldots p_s - q_1 q_2 \ldots q_t = 0. \tag{1}$$

By the second condition $p_1, \ldots, p_s, q_1, \ldots, q_t$ are primes. Since $p_1 \mid (q_1 \cdot \cdots \cdot q_t)$, by the definition of primality $\exists i \in \{1, \ldots, t\}$ such that $p_1 \mid q_i$. It follows that $q_i = p_1 \alpha$, and α is a unit. Factoring out p_1 in equation (1), we get

$$p_1(p_2 \ldots p_s - \alpha q_1 q_2 \ldots q_{i-1} q_{i+1} \ldots q_t) = 0. \tag{2}$$

Using the fact that R is an integral domain, we arrive at

$$p_2 \ldots p_s - \alpha q_1 q_2 \ldots q_{i-1} q_{i+1} \ldots q_t = 0. \tag{3}$$

We repeat the process for the remaining p_i's, showing that each of them is one of the q_i's multiplied by a unit, and that $s = t$. This proves that R is a UFD, as required. The other direction is left as an exercise to the reader. ■

2 Polynomial Rings Over Rings

2.1 Definitions and Basic Properties

Given a ring R we can create a new ring by introducing an indeterminate x. The set of all polynomials in x with coefficients in R forms a ring which is denoted $R[x]$. The degree of a nonzero polynomial **deg(p)** is the largest integer i such that the coefficient p_i of x^i is not zero. Each polynomial of degree n can also be represented as $n+1$-tuple:

$$p(x) = \sum_{i=0}^{n} p_i x^i \text{ is equivalent to } p = (p_0, p_1, \ldots, p_n), \quad p_i \in R \; \forall i.$$

We can also think of a polynomial in $R[x]$ as a list of finitely many elements from R, followed by infinitely many zeroes: $p = (p_0, p_1, \ldots, p_n, 0, \ldots)$.

[1] Throughout the lecture, R denotes a ring and F denotes a field.

Addition and multiplication of polynomials over a ring can be defined naturally. Let $p, q \in R[x]$, $p = (p_0, p_1, \ldots, p_n)$, $q = (q_0, q_1, \ldots, q_m)$, $n \leq m$.

$$\text{Addition:} \qquad p + q = (p_0 + q_0, p_1 + q_1, \ldots, p_m + q_m)$$

$$\text{Multiplication:} \qquad p \cdot q = (C_0, C_1, \ldots, C_{m+n}), \text{ where } C_k = \sum_{i=0}^{k} p_i \cdot q_{k-i}$$

We can check that R[x] satisfies all the axioms of a ring.

Theorem 2 (Inheritance properties) *R[x] inherits the following important properties from the ring R:*

1. *$R[x]$ is commutative $\Leftrightarrow$ R is commutative.*
2. *$R[x]$ is an integral domain $\Leftrightarrow$ R is an integral domain.*
3. *$R[x]$ is a UFD $\Leftrightarrow$ R is a UFD.*

Proof

1. Easy.
2. The forward direction is very easy. Suppose that $R[x]$ is an integral domain. Consider $p, q \in R$ such that $pq = 0$. Obviously, p and q are also constant polynomials in the integral domain $R[x]$. Consequently, either $p = 0$ or $q = 0$, and R is an integral domain.

   The other direction is also not too hard. Suppose R is an integral domain. Consider two polynomials $p(x), q(x) \in R[x]$ such that $p(x)q(x) = 0$. To prove that $R[x]$ is an integral domain, we have to show that either $p(x) = 0$ or $q(x) = 0$.

   For contradiction, suppose that both $p(x)$ and $q(x)$ are nonzero. Let p_n be the highest nonzero coefficient of $p(x)$, and q_m be the highest nonzero coefficient of $q(x)$. Then the coefficient of x^{n+m} of the product polynomial is equal to $p_n q_m$. Since the product polynomial is zero, $p_n q_m = 0$. Since R is an integral domain, either $p_n = 0$ or $q_m = 0$. Contradiction. Either $p(x)$ or $q(x)$ has to be zero.
3. This property is known as Gauss's Lemma. We will not prove it in class - but see appendix to this lecture appendix for a proof.

■

Having defined polynomials over a ring, we in fact created polynomials in more than one variable. By definition, $(R[x])[y]$ represents polynomials in y over the ring R[x]. It is the ring of polynomials of the form $\sum_{i,j} c_{ij} x^i y^j$, where $c_{ij} \in R$. We denote this ring by $R[x, y]$. Notice, that we can also define $R[x, y]$ as $(R[y])[x]$.

$$(R[x])[y] = R[x, y] = (R[y])[x].$$

Later we will use the two alternative definitions of $R[x, y]$ to obtain proofs of some nice facts. Notice that $R[x, y]$ inherits the properties of the ring R.

Theorem 3 (Division Theorem) *Given $a(x), b(x) \in F[x]$ $\exists!$ quotient $q(x)$, remainder $r(x) \in F[x]$ such that $a(x) = q(x) \cdot b(x) + r(x)$ and $deg(r) < deg(b)$.*

Proof **(Division Algorithm)**

Let $a^{(1)}(x)$ and $b(x)$ be two polynomials in $F[x]$.

$$\begin{aligned}
&a^{(1)}(x) = a^{(1)}_{n_1} x^{n_1} + a^{(1)}_{n_1-1} x^{n_1-1} + \cdots + a^{(1)}_1,\\
&b(x) = b_m x^m + b_{m-1} x^{m-1} + \cdots + b_1,\\
&\text{and } n_1 \geq m.
\end{aligned}$$

At the i^{th} step of the algorithm, we multiply $b(x)$ by $T^{(i)}(x) = \frac{a^{(i)}_{n_i}}{b_m} x^{n_i - m}$, and subtract the result from $a^{(i)}(x)$ to obtain a new polynomial $a^{(i+1)}(x)$.

$$a^{(i+1)}(x) = a^{(i)}(x) - \frac{a^{(i)}_{n_i}}{b_m} x^{n_i-m} \cdot b(x).$$

Let $n_{(i+1)}$ denote the degree of $a^{(i+1)}(x)$. Since $n_{i+1} \leq n_i$, the algorithm will eventually find $a^{(k)}$ with $n_k < m$. Set $r(x) = a^{(k)}(x)$, and $q(x) = \sum_{i=1}^{k-1} T^{(i)}(x)$.

Notice that we can divide by b_m because we are working with polynomials over a field. This algorithm is not useful for bivariate polynomials, but later we will define division in more general settings. The proof of uniqueness is left as an exercise to the reader. ■

2.2 GCD

Usually, concepts defined for integers have analogues for univariate polynomials. We show how familiar terminology can be applied to polynomials. We say that $a(x)$ is a **divisor** of $b(x)$ ($a(x) \mid b(x)$) if $b(x) = a(x) \cdot q(x)$ for some $q(x)$. If $g(x) \mid a(x)$ and $g(x) \mid b(x)$, $g(x)$ is called a **common divisor** of $a(x)$ and $b(x)$. Finally, $g(x)$ is a **greatest common divisor** (GCD) of $a(x)$ and $b(x)$, denoted $\gcd(a(x), b(x))$, if $g(x)$ has the largest degree among all common divisors of $a(x)$ and $b(x)$.

Lemma 4 *Let $S = \{p(x) \in F[x]$ such that $\exists u, v \in F[x]$ with $p = ua + vb\}$. The smallest degree polynomial $g \in S$ is*

1. *a common divisor of a and b;*
2. *$gcd(a, b)$;*
3. *unique up to multiplication by units.*

Proof By definition, $g = ua + vb$ for some $u, v \in F[x]$.

1. Suppose, for contradiction, that g does not divide a. Let r be the remainder

   $$r = a - qg = a - q[ua + vb] = (1 - qu)a + (-qv)b.$$

   Then $r \in S$, and has a smaller degree than g. Contradiction.

2. Let $g\prime$ be a common divisor of a and b. Then $g\prime \mid a$ and $g\prime \mid b$, and consequently $g\prime \mid f = ua + vb \ \ \forall u, v$. By definition of g, $g\prime \mid g$. Since any common divisor of a and b divides g, g is $\gcd(a, b)$.

3. Suppose $g\prime = \gcd(a, b)$. By (2) $g \mid g\prime$, i.e., $g = \alpha \cdot g\prime$. Since, by definition of gcd, deg(g)=deg($g\prime$), α is a unit.

■

2.3 Evaluation of polynomials

Polynomials (of low degree) are very interesting to coding theory and complexity theory. For these applications, we need the notion of evaluation of polynomials. Given $p(x)$, define function $EV_p : F \to F$ as $EV_p(\alpha) = \sum_{i=0}^{n} p_i \alpha^i$. $EV_p(\alpha)$ is usually denoted by $p(\alpha)$. The following properties of the evaluation map can be verified syntactically.

$$\begin{aligned} \forall p, q, \alpha \quad & EV_p(\alpha) + EV_q(\alpha) = EV_{p+q}(\alpha) \\ & EV_p(\alpha) \cdot EV_q(\alpha) = EV_{pq}(\alpha) \end{aligned}$$

Theorem 5 *$p(\alpha) = 0$ iff $(x - \alpha) \mid p(x)$*

Proof By Division theorem $\exists q, r$ such that $p(x) = q(x)(x - \alpha) + r(x)$ with deg(r) <deg$(x - \alpha) = 1$, or equivalently with $r \in F$.

$$\begin{aligned} & 0 = p(\alpha) = q(\alpha)[\alpha - \alpha] + r \Rightarrow \\ & r = 0 \Rightarrow \\ & p(x) = q(x)(x - \alpha) \end{aligned}$$

■

Definition 3 *α is a <u>root</u> of $p(x)$ if $p(\alpha) = 0$.*

Lemma 6 *If $\alpha_1, \ldots, \alpha_{n+1}$ are distinct roots of $p(x)$, then $\Pi_{i=1}^{n+1}(x - \alpha_i) \mid p(x)$.*

Proof By theorem 5, $p(x) = (x - \alpha_1)q_1(x)$. Evaluating $p(x)$ at α_2, we get $0 = (\alpha_1 - \alpha_2)q_1(\alpha_2)$. Note that $\alpha_1 - \alpha_2 \neq 0$ because the roots are distinct. Since we are working with polynomials over a field (hence an integral domain), $q_1(\alpha_2) = 0$. By theorem 5, $(x - \alpha_2) \mid q_1(x)$. As before, $q_1(x) = (x - \alpha_2)q_2(x)$, and $p = (x - \alpha_1)(x - \alpha_2)q_2(x)$. We repeat the above derivation until we show that $\Pi_{i=1}^{n+1}(x - \alpha_i) \mid p(x)$. ■

Theorem 7 *A nonzero polynomial $p(x) \in F[x]$ of degree n has at most n roots.*

Proof Suppose, for contradiction, that $\alpha_1, \alpha_2, \ldots, \alpha_{n+1}$ are distinct roots of $p(x)$. Then by lemma 6, $\Pi_{i=1}^{n+1}(x - \alpha_i) \mid p(x)$. But a polynomial of degree $n + 1$ cannot divide $p(x)$. Contradiction. ■

Exercise Prove the analogue of the previous theorem for bivariate polynomials.

$$\begin{aligned} \text{Given} \quad & p(x, y) \in F[x, y], \\ \text{show} \quad & p(x, q(x)) = 0 \Rightarrow y - q(x) \mid p(x, y). \end{aligned}$$

Theorem 8 (Interpolation Theorem) *Given $x_1, y_1, \ldots, x_n, y_n \in F$ with $x_i \neq x_j$ for $i \neq j$, $\exists$ polynomial $p^{(n)} \in F[x]$ of degree less than or equal to $n - 1$, such that $p^{(n)}(x_i) = y_i$.*

Proof We prove this theorem by giving an inductive construction of the polynomial $p^{(n)}(x)$. If $n = 1$, then $p^{(1)}(x) = y_1$. Suppose we found $p^{(i)}(x)$ of degree less than or equal to $i - 1$, such that $p^{(i)}(x_j) = y_j$ for $j = 0, \ldots, i$, and we are looking for $p^{(i+1)}(x)$. If $p^{(i)}(x_{j+1}) = y_{j+1}$, we are done. Otherwise, let

$$d^{(i+1)}(x) = p^{(i+1)}(x) - p^{(i)}(x).$$

Our goal is to find a suitable $d^{(i+1)}(x)$ and, consequently, $p^{(i+1)}(x)$. For all $j = 0, \ldots, i$, $p^{(i+1)}(x_j) = p^{(i)}(x_j)$, and hence $d^{(i+1)}(x_j) = 0$. The degree of $d^{(i+1)}(x)$ is less than or equal to i. The polynomial

$q^{(i+1)}(x) = \Pi_{j=1}^{i}(x - x_j)$ satisfies the above conditions. However, in general $q^{(i+1)}(x_{j+1}) \neq d^{(i+1)}(x_{j+1})$. Since $d^{(i+1)}(x_{j+1}) = y_{j+1} - p^{(i)}(x_{j+1})$, we define

$$d^{(i+1)}(x) = \frac{y_{j+1} - p^{(i)}(x_{j+1})}{q^{(i+1)}(x_{j+1})} q^{(i+1)}(x).$$

And finally,

$$p^{(i+1)}(x) = p^{(i)}(x) + d^{(i+1)}(x) = p^{(i)}(x) + \frac{y_{j+1} - p^{(i)}(x_{j+1})}{q^{(i+1)}(x_{j+1})} q^{(i+1)}(x).$$

By construction, $p^{(i)}(x)$ satisfies the required conditions. ■

The proof of the interpolation theorem yields the following algorithm for the interpolation problem:

Interpolate($\{(x_1, y_1), \ldots, (x_n, y_n)\}$); /* x_i's distinct */
    Initialize
        $p^{(0)}(x) \leftarrow \mathbf{0}$.
        $q^{(0)}(x) \leftarrow \mathbf{1}$.
    Iterate for $i = 1$ to n;
        $\lambda_i \leftarrow \frac{y_i - p^{(i)}(x_i)}{q^{(i-1)}(x_i)}$;
        $p^{(i)}(x) \leftarrow p^{(i-1)}(x) + \lambda_i q^{(i-1)}(x)$;
        $q^{(i)}(x) \leftarrow q^{(i-1)}(x) \cdot (x - x_i)$;
    Output($p^{(n)}$).

Appendix

We work towards Gauss's Lemma by first showing that the ring of polynomials over a field F is a unique factorization domain.

Lemma 9 *For any field F, $F[x]$ is a UFD.*

Proof We apply Lemma 1. To do so we need to verify that every element $p \in F[x]$ has a finite factorization into irreducibles. This is easily established: the number of irreducible factors is at most the degree of p. Thus to prove the lemma it suffices to show that every irreducible polynomial p is also a prime. I.e., if $p|ab$ then $p|a$ or $p|b$.

Assume $p \not| a$ and $p \not| b$. Then $\gcd(p, a) = 1$ and $\gcd(p, b) = 1$. (Since the common divisor must divide p and p is irreducible.) By Lemma 4, we know that we can express the gcd's as a linear combination of the arguments. I.e., there exist u_1, u_2, v_1 and u_2 such that

$$u_1 p + v_1 a = 1 \text{ and } u_2 p + v_2 b = 1.$$

Multiplying the above we get

$$1 = (u_1 p + v_1 a)(u_2 p + v_2 b) = (u_1 u_2 p + u_1 v_2 b + u_2 v_1 a)p + (v_1 v_2)ab.$$

In other words 1 can be expressed as a linear combination of p and ab, thus implying that $\gcd(p, ab) = 1$ contradicting the fact that $p|ab$. ■

Lemma 10 (Gauss) *$R[x]$ is a UFD $\Leftrightarrow$ R is a UFD.*

Proof It is easy to see that if R is not a UFD, then $R[x] \supset R$ is not a UFD. The harder direction is the other way around. The main idea behind this direction is to use the fact that the ring $\tilde{R}[x]$ is a UFD, where $\tilde{R}$ is the field of fractions of R. [2]

Again we use Lemma 1. To do so we need to verify first that any $p(x) \in R[x]$ can be factored into finitely many irreducibles. This time we need to be slightly careful, since some factors may be elements of R. Let $p_0, \ldots, p_n$ be the coefficients of $p(x)$. Then notice that $a \in R$ divides $p(x)$ if and only if $a|p_i$ for every $i \in \{0, \ldots, n\}$. Thus we first write $p(x) = a \cdot q(x)$ where $a \in R$ and the coefficients of $q(x)$ have no non-trivial common factors. Since R is a UFD, a is expressible as a product of finitely many irreducibles. Divisors of $q(x)$, on the other hand have positive degree in x and thus q factors into at most $\deg_x q$ irreducibles. Thus to establish the lemma it suffices to show that every irreducible element of $R[x]$ is prime. This is established below.

Claim 11 *If $p(x)$ is irreducible in $R[x]$, then:*

1. *$p(x)$ is irreducible (and hence prime) in $\tilde{R}[x]$.*
2. *$p(x)$ is prime in $R[x]$.*

Proof

1. Say $p(x) = \tilde{a}(x)\tilde{b}(x)$ for $\tilde{a}(x), \tilde{b}(x) \in \tilde{R}(x)$. Then by taking the least common multiple of the denominators of the coefficients of $\tilde{a}(x)$ and $\tilde{b}(x)$ and separating them out, we get $cp(x) = a(x)b(x)$ where $a(x), b(x) \in R[x]$, $c \in R$. If c is a unit in R, then we are done, since we find that p is reducible in $R[x]$. If not, then we need to find some contradiction. Let $c = c_1 c_2 \ldots c_t$ be a prime factorization of c. We may assume that c_i does not divide some coefficient a_j of $a(x)$ and b_k of $b(x)$ (else we could remove common factors from both sides). Further, let j and k be the least indices with this property: i.e., $c_i|a_{j'}$ for $j' < j$ and $c_i|b_{k'}$ for $k' < k$. Then, consider the coefficient of x^{j+k} in the equation $cp(x) = a(x)b(x)$. On the LHS we have a multiple of c_i and on the right we have the summation $\sum_{l=0}^{j+k} a_l b_{j+k-l}$. All terms in this summation, except the term $a_j b_k$ are divisible by c_i. Rearranging terms, we get that $a_j b_k$ is divisible by c_i which gives a contradiction.

2. Let $p(x)$ be irreducible in $R[x]$. Then we have that the gcd of the coefficients of $p(x)$ is 1. Now suppose $p(x)|a(x)b(x)$. Then since $p(x)$ is prime in $\tilde{R}[x]$ we have that $p(x)|a(x)$ or $p(x)|b(x)$ in $\tilde{R}[x]$. Say it is the former. Then we can find $f(x) \in R[x]$ and $c \in R$ such that $ca(x) = f(x)p(x)$, and furthermore c does not have a common divisor with all coefficients of f. As in Part (1), we can show that every prime factor of c either divides every coefficient of f or divides every coefficient of p. In either case we have a contradiction - hence c must be a unit and thus $a(x) = (c^{-1}f(x))p(x)$ and thus $p(x)|a(x)$ over $R[x]$.

■■

[2] This idea of relying on the properties of $\tilde{R}$ rather than R is a standard method, and can be quite powerful. This method is often used to argue about $R = F[x_1, \ldots, x_n]$, the ring of polynomials in $x_1, \ldots, x_n$ over a field F. In such cases $\tilde{R}$ is also denoted $F(x_1, \ldots, x_n)$ — note the subtle distinction in the notation: $F(\cdots)$ vs. $F[\cdots]$.