Factorization in Integral Domains II

1 Statement of the main theorem

Throughout these notes, unless otherwise specified, R is a UFD with field of quotients F. The main examples will be $R = \mathbb{Z}$, $F = \mathbb{Q}$, and $R = K[y]$ for a field K and an indeterminate (variable) y, with $F = K(y)$.

The basic example of the type of result we have in mind is the following (often done in high school math courses):

Theorem 1.1 (Rational roots test). *Let $f = a_n x^n + \cdots + a_0 \in \mathbb{Z}[x]$ be a polynomial of degree $n \geq 1$ with integer coefficients and nonzero constant term a_0, and let $r/s \in \mathbb{Q}$ be a rational root of f such that the fraction r/s is in lowest terms, i.e.* $\gcd(r, s) = 1$. *Then r divides the constant term a_0 and s divides the leading coefficient a_n.*

In particular, if f is monic, then a rational root of f must be an integer dividing a_0.

Proof. Since r/s is a root of f,

$$0 = f(r/s) = a_n \left(\frac{r}{s}\right)^n + a_{n-1} \left(\frac{r}{s}\right)^{n-1} + \cdots + a_0.$$

Clearing denominators by multiplying both sides by s^n gives

$$a_n r^n + a_{n-1} r^{n-1} s + \cdots + a_0 s^n = 0.$$

Moving the last term over to the right hand side gives

$$\begin{aligned} -a_0 s^n &= a_n r^n + a_{n-1} r^{n-1} s + \cdots + a_1 r s^{n-1} \\ &= r(a_n r^{n-1} + a_{n-1} r^{n-2} s + \cdots + a_1 s^{n-1}). \end{aligned}$$

Hence $r | a_0 q^n$. Since r and s are relatively prime, r and s^n are relatively prime, and thus $r | a_0$. The argument that $s | a_n$ is similar. $\square$

Clearly, the same statement is true (with the same proof) in case R is any UFD with field of quotients F. Our main goal in these notes will be to prove the following, which as we shall see is a generalization of the rational roots test:

Theorem 1.2. *Let $f \in R[x]$ be a polynomial of degree $n \geq 1$. Then f is a product of two polynomials in $F[x]$ of degrees d and e respectively with $0 < d, e < n$ if and only if there exist polynomials $g, h \in R[x]$ of degrees d and e respectively with $0 < d, e < n$ such that $f = gh$.*

We will prove the theorem later. Here we just make a few remarks.

Remark 1.3. (1) In the proof of the theorem, the factors $g, h \in R[x]$ of f will turn out to be multiples of the factors of f viewed as an element of $F[x]$.

(2) Clearly, if there exist polynomials $g, h \in R[x]$ of degrees d and e respectively with $0 < d, e < n$ such that $f = gh$, then the same is true in $F[x]$. Hence the $\Longleftarrow$ direction of the theorem is trivial.

(3) Since a (nonconstant) polynomial in $F[x]$ is reducible $\iff$ it is a product of two polynomials of smaller degrees, we see that we have shown:

Corollary 1.4. *Let $f \in R[x]$ be a polynomial of degree $n \geq 1$. If there do not exist polynomials $g, h \in R[x]$ of degrees d and e respectively with $0 < d, e < n$ such that $f = gh$, then f is irreducible in $F[x]$.* $\square$

Equivalently, if f is reducible in $F[x]$, then f factors into a product of polynomials of smaller degree in $R[x]$. However, if R is an integral domain which is **not** a UFD, then it is possible for a polynomial $f \in R[x]$ to be reducible in $F[x]$ but irreducible in $R[x]$.

(4) Conversely, if $f \in R[x]$ is irreducible in $F[x]$ but reducible in $R[x]$, then since f cannot factor as a product of polynomials of smaller degrees in $R[x]$, it must be the case that $f = cg$, where $c \in R$ and c is not a unit. A typical example is the polynomial $11x^2 - 22 \in \mathbb{Z}[x]$, which is irreducible in $\mathbb{Q}[x]$ since it is a nonzero rational number times $x^2 - 2$. But in $\mathbb{Z}[x]$, $11x^2 - 22 = 11(x^2 - 2)$ and this is a nontrivial factorization since neither factor is a unit in $\mathbb{Z}[x]$.

(5) The relation of Theorem 1.2 to the Rational Roots Test is the following: the proof of Theorem 1.2 will show that, if r/s is a root of f in lowest terms, so that $x - r/s$ divides f in $\mathbb{Q}[x]$, then in fact we will prove that $sx - r$ divides f in $\mathbb{Z}[x]$, and hence s divides the leading coefficient and r divides the constant term.

2 Tests for irreducibility

We now explain how Theorem 1.2 above (or more precisely Corollary 1.4) leads to tests for irreducibility in $F[x]$. Applying these tests is a little like applying tests for convergence in one variable calculus: it is an art, not a science, to see which test (if any) will work, and sometimes more than one test will do the job. We begin with some notation:

Let R be any ring, not necessarily a UFD or even an integral domain, and let I be an ideal in R. Then we have the homomorphism $\pi\colon R \to R/I$ defined by $\pi(a) = a + I$ ("reduction mod I"). For brevity, we denote the image $\pi(a)$ of the element $a \in R$, i.e. the coset $a + I$, by $\bar{a}$. Similarly, there is a homomorphism, which we will also denote by π, from $R[x]$ to $(R/I)[x]$, defined as follows: if $f = \sum_{i=0}^{n} a_i x^i \in R[x]$, then

$$\pi(f) = \sum_{i=0}^{n} \bar{a}_i x^i \in (R/I)[x].$$

Again for the sake of brevity, we abbreviate $\pi(f)$ by $\bar{f}$ and refer to it as the "reduction of f mod I." The statement that π is a homomorphism means that $\overline{fg} = \bar{f}\bar{g}$. Note that $\bar{f} = 0 \iff$ all of the coefficients of f lie in I. Furthermore, if $f = \sum_{i=0}^{n} a_i x^i$ has degree n, then either $\deg \bar{f} \leq n$ or $\bar{f} = 0$, and $\deg \bar{f} = n \iff$ the leading coefficient a_n does not lie in I. We also have:

Lemma 2.1. *Let R be an integral domain and let $f = \sum_{i=0}^{n} a_i x^i \in R[x]$ with $a_n \notin I$. If $f = gh$ with $\deg g = d$ and $\deg h = e$, then $\deg \bar{g} = d = \deg g$ and $\deg \bar{h} = e = \deg h$.*

Proof. Since R is an integral domain, $n = \deg f = \deg g + \deg h = d + e$. Moreover, $\deg \bar{g} \leq d$ and $\deg \bar{h} \leq e$. But

$$d + e = n = \deg f = \deg \bar{f} = \deg(\bar{g}\bar{h}) \leq \deg \bar{g} + \deg \bar{h} \leq d + e.$$

The only way that equality can hold at the ends is if all inequalities that arise are actually equalities. In particular we must have $\deg \bar{g} = d$ and $\deg \bar{h} = e$. □

Returning to our standing assumption that R is a UFD, we then have:

Theorem 2.2. *Let $f = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0 \in R[x]$ be a polynomial of degree $n \geq 1$ and let I be an ideal in R. Suppose that $a_n \notin I$. If $\bar{f}$ is not a product of two polynomials in $(R/I)[x]$ of degrees d and e respectively with $0 < d, e < n$, then f is irreducible in $F[x]$.*

Proof. Suppose instead that f is reducible in $F[x]$. By Corollary 1.4, there exist $g, h \in R[x]$ such that $f = gh$, where $\deg g = d < n$ and $\deg h = e < n$. Then $\bar{f} = \bar{g}\bar{h}$, where, by Lemma 2.1, $\deg \bar{g} = d = \deg g$ and $\deg \bar{h} = e = \deg h$. But this contradicts the assumption of the theorem. $\square$

Remark 2.3. (1) Typically we will apply Theorem 2.2 in the case where I is a maximal ideal and hence R/I is a field, for example $R = \mathbb{Z}$ and $I = (p)$ where p is prime. In this case, the theorem says that, if the leading coefficient $a_n \notin I$ and $\bar{f}$ is irreducible in $(R/I)[x]$, then f is irreducible in $F[x]$.

For example, it is easy to check that $x^4 + x^3 + x^2 + x + 1$ is irreducible in $\mathbb{F}_2[x]$: it has no roots in $\mathbb{F}_2$, and so would have to be a product of two irreducible degree 2 polynomials in $\mathbb{F}_2[x]$. But there is only one irreducible degree 2 polynomial in $\mathbb{F}_2[x]$, namely $x^2 + x + 1$, so that we would have to have $(x^2 + x + 1)^2 = x^4 + x^3 + x^2 + x + 1$. Since the characteristic of $\mathbb{F}_2$ is 2,

$$(x^2 + x + 1)^2 = x^4 + x^2 + 1 \neq x^4 + x^3 + x^2 + x + 1.$$

Hence $x^4 + x^3 + x^2 + x + 1$ is irreducible in $\mathbb{F}_2[x]$. Then, for example,

$$117x^4 - 1235x^3 + 39x^2 + 333x - 5$$

is irreducible in $\mathbb{Q}[x]$, since it is a polynomial with integer coefficients whose reduction mod 2 is irreducible.

(2) To see why we need to make some assumptions about the leading coefficient of f, or equivalently that $\deg \bar{f} = \deg f$, consider the polynomial $f = (2x+1)(3x+1) = 6x^2 + 5x + 1$. Taking $I = (3)$, we see that $\bar{f} = 2x + 1$ is irreducible in $\mathbb{F}_3[x]$, since it is linear. But clearly f is reducible in $\mathbb{Z}[x]$ and in $\mathbb{Q}[x]$. The problem is that, mod 3, the factor $3x + 1$ has become a unit and so does not contribute to the factorization of the reduction mod 3.

(3) By (1) above, if $f \in \mathbb{Z}[x]$, say with f monic, and if there exists a prime p such that the reduction mod p of f is irreducible in $\mathbb{F}_p[x]$, then f is irreducible in $\mathbb{Q}[x]$. One can ask if, conversely, f is irreducible in $\mathbb{Q}[x]$, then does there always exist a prime p such that the reduction mod p of f is irreducible in $\mathbb{F}_p[x]$? Perhaps somewhat surprisingly, the answer is **no**: there exist monic polynomials $f \in \mathbb{Z}[x]$ such that f is irreducible in $\mathbb{Q}[x]$ but such that the reduction mod p of f is reducible in $\mathbb{F}_p[x]$ for every prime p. An example is given on the homework. Nevertheless, reducing mod p is a basic tool for studying the irreducibility of polynomials and there is an effective procedure (which can be implemented on a computer) for deciding when a polynomial $f \in \mathbb{Z}[x]$ is irreducible in $\mathbb{Q}[x]$.

The next method is the so-called *Eisenstein criterion*:

Theorem 2.4 (Eisenstein criterion). *Let $f = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0 \in R[x]$ be a polynomial of degree $n \geq 1$. Let M be a maximal ideal in R. Suppose that*

1. *The leading coefficient a_n of f does not lie in M;*

2. *For $i < n$, $a_i \in M$;*

3. *$a_0 \notin M^2$, in particular there do not exist $b, c \in M$ such that $a_0 = bc$.*

Then f is not the product of two polynomials of strictly smaller degree in $R[x]$ and hence f is irreducible as an element of $F[x]$.

Proof. Suppose that $f = gh$ where $g, h \in R[x]$, $\deg g = d < n$ and $\deg h = e < n$. Then $\bar{f} = \bar{g}\bar{h}$, where, by Lemma 2.1, $\deg \bar{g} = d = \deg g$ and $\deg \bar{h} = e = \deg h$. But $\bar{f} = \bar{a}_n x^n$. so we must have $\bar{g} = r_1 x^d$ and $\bar{h} = r_2 x^e$ for some $r_1, r_2 \in R/M$. Thus $g = b_d x^d + \cdots + b_0$ and $h = c_e x^e + \cdots + c_0$, with $b_i, c_j \in M$ for $i < d$ and $j < e$. In particular, since $d > 0$ and $e > 0$, both of the constant terms $b_0, c_0 \in M$. But then the constant term of $f = gh$ is $b_0 c_0 \in M^2$, contradicting (iii). $\square$

Remark 2.5. (1) A very minor modification of the proof shows that it is enough to assume that M is a prime ideal.

(2) If $M = (r)$ is a principal ideal, then $M^2 = (r^2)$. Thus, for $R = \mathbb{Z}$ and $I = (p)$, where p is a prime number, the conditions read: p does not divide a_n, p divides a_i for all $i < n$, and p^2 does not divide a_0.

Example 2.6. Using the Eisenstein criterion with $p = 2$, we see that $x^n - 2$ is irreducible for all $n > 0$. More generally, if p is a prime number, then $x^n - p$ is irreducible for all $n > 0$, as is $x^n - pk$ where k is any integer such that p does not divide k.

For another example,

$$f = 55x^5 - 45x^4 + 105x^3 + 900x^2 - 405x + 75$$

satisfies the Eisenstein criterion for $p = 3$, hence is irreducible in $\mathbb{Q}[x]$. Note that f is **not** irreducible in $\mathbb{Z}[x]$, since

$$f = 5(11x^5 - 9x^4 + 21x^3 + 180x^2 - 81x + 15).$$

3 Cyclotomic polynomials

An n^{th} root of unity ζ in a field F is an element $\zeta \in F$ such that $\zeta^n = 1$, i.e. a root of the polynomial $x^n - 1$ in F. We let $\mu_n(F)$ be the set of all such, i.e.

$$\mu_n(F) = \{\zeta \in F : \zeta^n = 1\}.$$

Lemma 3.1. *The set $\mu_n(F)$ is a finite cyclic subgroup of F^* (under multiplication) of order dividing n.*

Proof. There are at most n roots of the polynomial $x^n - 1$ in F, and hence $\mu_n(F)$ is finite. It is a subgroup of F^* (under multiplication): if ζ_1 and ζ_2 are n^{th} roots of unity, then $\zeta_1^n = \zeta_2^n = 1$, and thus $(\zeta_1\zeta_2)^n = \zeta_1^n\zeta_2^n = 1$ as well. Thus $\mu_n(F)$ is closed under multiplication. Since $1^n = 1$, $1 \in \mu_n(F)$. Finally, if ζ is an n^{th} root of unity, then $(\zeta^{-1})^n = (\zeta^n)^{-1} = 1^{-1} = 1$. Then $\mu_n(F)$ is a finite subgroup of F^*, hence it is a cyclic group. Since a generator ζ satisfies $\zeta^n = 1$, the order of ζ, and hence of $\mu_n(F)$, divides n. $\square$

For example, for $F = \mathbb{C}$, the group $\mu_n(\mathbb{C}) = \mu_n$ of (complex) n^{th} roots of unity is a cyclic subgroup of $\mathbb{C}^*$ (under multiplication) of order n, and a generator is $e^{2\pi i/n}$. On the other hand, if $F = \mathbb{R}$, then $\mu_n(\mathbb{R}) = \{1\}$ if n is odd and $\{\pm 1\}$ if n is even, and a similar statement holds for $F = \mathbb{Q}$. If the characteristic of F is 0, or does not divide n, then by a homework problem $x^n - 1$ has distinct roots, and so there is some algebraic extension E of F for which the number of n^{th} roots of unity in E is exactly n. On the other hand, if the characteristic of F is p, then $x^p - 1 = (x-1)^p$, and the only p^{th} root of unity in every extension field of F is 1. **For the rest of this section, we take $F = \mathbb{C}$ and thus $\mu_n(\mathbb{C}) = \mu_n$ as we have previously defined it.**

Since 1 is always an n^{th} root of unity, $x - 1$ divides $x^n - 1$, and the set of nontrivial n^{th} roots of unity is the set of roots of $\dfrac{x^n - 1}{x - 1} = x^{n-1} + x^{n-2} + \cdots + x + 1$ (geometric series). In general, this polynomial is reducible. For example, with $n = 4$, and $F = \mathbb{Q}$, say,

$$x^4 - 1 = (x^2 - 1)(x^2 + 1) = (x - 1)(x + 1)(x^2 + 1).$$

Here, the root 1 of $x - 1$ has order 1, the root -1 of $x + 1$ has order 2, and the two roots $\pm i$ of $x^2 + 1$ have order 4. For another example,

$$x^6 - 1 = (x^3 - 1)(x^3 + 1) = (x - 1)(x^2 + x + 1)(x + 1)(x^2 - x + 1).$$

As before 1 has order 1 in μ_6, -1 has order 2, the two roots of x^2+x+1 have order 3, and the two roots of x^2-x+1 have order 6. Note that, if $d|n$, then $\mu_d \leq \mu_n$ and the roots of x^d-1 are roots of x^n-1. In fact, if $n=kd$, then as before

$$x^n - 1 = x^{kd} - 1 = (x^d - 1)(x^{k(d-1)} + x^{k(d-2)} + \cdots + x^k + 1).$$

In general, we refer to an element ζ of μ_n of order n as a *primitive* n^{th} root of unity. Since a primitive n^{th} root of unity is the same thing as a generator of μ_n, there are exactly $\varphi(n)$ primitive n^{th} roots of unity; explicitly, they are exactly of the form $e^{2\pi i a/n}$, where $0 \leq a \leq n-1$ and $\gcd(a,n)=1$.

In case n is prime, we have the following:

Theorem 3.2. *Let p be a prime number. Then the cyclotomic polynomial*

$$\Phi_p = \frac{x^p - 1}{x - 1} = x^{p-1} + x^{p-2} + \cdots + x + 1$$

is irreducible in $\mathbb{Q}[x]$.

Proof. The trick is to consider, not Φ_p, but rather $\Phi_p(x+1)$. Clearly, $\Phi_p(x)$ is irreducible if and only if $\Phi_p(x+1)$ is irreducible (because a factorization $\Phi_p = gh$ gives a factorization $\Phi_p(x+1) = g(x+1)h(x+1)$, and conversely a factorization $\Phi_p(x+1) = ab$ gives a factorization $\Phi_p(x) = a(x-1)b(x-1)$.) To see that $\Phi_p(x+1)$ is irreducible, use:

$$\begin{aligned}\Phi_p(x+1) = \frac{(x+1)^p - 1}{(x+1)-1} &= \frac{x^p + \binom{p}{1}x^{p-1} + \binom{p}{2}x^{p-2} + \cdots + \binom{p}{p-1}x + 1 - 1}{x} \\ &= x^{p-1} + \binom{p}{1}x^{p-2} + \binom{p}{2}x^{p-3} + \cdots + \binom{p}{p-1}.\end{aligned}$$

As we have seen (homework on the Frobenius homomorphism), if p is prime, p divides each binomial coefficient $\binom{p}{k}$ for $1 \leq k \leq p-1$, but p^2 does not divide $\binom{p}{p-1} = p$. Hence $\Phi_p(x+1)$ satisfies the hypotheses of the Eisenstein criterion. $\square$

Corollary 3.3. *Let p be a prime number. Then $[\mathbb{Q}(e^{2\pi i/p}) : \mathbb{Q}] = p-1$.* $\square$

In case n is not necessarily prime, we define the n^{th} cyclotomic polynomial $\Phi_n \in \mathbb{C}[x]$ by:

$$\Phi_n = \prod_{\substack{\zeta \in \mu_n \\ \zeta \text{ is primitive}}} (x - \zeta).$$

For example, $\Phi_1 = x - 1$, $\Phi_4 = x^2 + 1$, and $\Phi_6 = x^2 - x + 1$. If p is a prime, then every p^{th} root of unity is primitive except for 1, and hence, consistent with our previous notation, $\Phi_p = \dfrac{x^p - 1}{x - 1} = x^{p-1} + x^{p-2} + \cdots + x + 1$. Clearly, $\deg \Phi_n = \varphi(n)$, and

$$x^n - 1 = \prod_{d|n} \Phi_d,$$

reflecting the fact that $\sum_{d|n} \varphi(d) = n$. We then have the following theorem, which we shall not prove:

Theorem 3.4. *The polynomial $\Phi_n \in \mathbb{Z}[x]$ and Φ_n is irreducible in $\mathbb{Q}[x]$. Hence, the irreducible factors of $x^n - 1$ are exactly the polynomials Φ_d for d dividing n.* $\square$

Corollary 3.5. *For every $n \in \mathbb{N}$, $[\mathbb{Q}(e^{2\pi i/n}) : \mathbb{Q}] = \varphi(n)$.* $\square$

For example, we have seen that

$$x^6 - 1 = (x-1)(x+1)(x^2 + x + 1)(x^2 - x + 1) = \Phi_1\Phi_2\Phi_3\Phi_6.$$

Moreover, $e^{2\pi i/6} = e^{\pi i/3} = -e^{4\pi i/3} = \frac{1}{2} + \frac{1}{2}\sqrt{-3}$, hence

$$[\mathbb{Q}(e^{2\pi i/6}) : \mathbb{Q}] = [\mathbb{Q}(\sqrt{-3}) : \mathbb{Q}] = 2 = \varphi(6).$$

4 Proofs

We turn now to Theorem 1.2, discussed earlier and give its proof. Recall the following basic property of a UFD (Proposition 1.13 of the last handout): Let $r \in R$ with $r \neq 0$. Then r is an irreducible element of $R \iff$ the principal ideal (r) is a prime ideal of R.

For a UFD R, we have already defined the gcd of two elements $r, s \in R$, not both 0, and have noted that it always exists and is unique up to multiplying by a unit. More generally, if $r_1, \dots, r_n \in R$, where the r_i are not all 0, then we define the gcd of $r_1, \dots, r_n$ to be an element d of R such that $d|r_i$ for all i, and if e is any other element of R such that $e|r_i$ for all i, then $e|d$. As in the case $i = 2$, the gcd of $r_1, \dots, r_n$ exists and is unique up to multiplication by a unit. Since not all of the r_i are 0, a gcd of the r_i is also nonzero. We denote a gcd of $r_1, \dots, r_n$ by $\gcd(r_1, \dots, r_n)$. In fact, we can define the gcd of n elements inductively: once the gcd of $n-1$ nonzero elements has been defined, if $r_1, \dots, r_n \in R$ are such that not all of $r_1, \dots, r_{n-1}$ are 0, and $d_{n-1} = \gcd(r_1, \dots, r_{n-1})$, then it is easy to see that

$\gcd(r_1, \dots, r_n) = \gcd(d_{n-1}, r_n)$. Similarly, we say that $r_1, \dots, r_n \in R$ are *relatively prime* if $\gcd(r_1, \dots, r_n) = 1$, or equivalently if $d|r_i$ for all $i \implies d$ is a unit. There are the following straightforward properties of a gcd:

Lemma 4.1. *Let R be a UFD and let $r_1, \dots, r_n \in R$, not all 0.*

(i) *If d is a gcd of $r_1, \dots, r_n$, then $r_1/d, \dots, r_n/d$ are relatively prime, i.e.*

$$\gcd(r_1/d, \dots, r_n/d) = 1.$$

(ii) *If $c \in R$, $c \neq 0$, then*

$$\gcd(cr_1, \dots, cr_n) = c \gcd(r_1, \dots, r_n).$$

Proof. To see (i), if $e|(r_i/d)$ for every i, then $de|r_i$ for every i, hence de divides d, so that e divides 1 and hence e is a unit. To see (ii), let d be a gcd of $r_1, \dots, r_n$ and let $d' = \gcd(cr_1, \dots, cr_n)$. Since d is a gcd of $r_1, \dots, r_n$, $d|r_i$ for every $i \implies cd$ divides cr_i for every $i \implies cd$ divides d'. Thus $d' = ce$ for some $e \in R$. Then ce divides $cr_i \implies e$ divides r_i for every $i \implies e|d$. Thus $d' = ce$ divides cd, and since cd divides d', $d' = cd$ up to multiplication by a unit. $\square$

Definition 4.2. Let $f = \sum_{i=0}^{n} a_i x^i \in R[x]$ with $f \neq 0$. Then the *content* $c(f)$ is the gcd of the coefficients of f:

$$c(f) = \gcd(a_n, \dots, a_0).$$

It is well defined up to a unit. The polynomial f is a *primitive* polynomial $\iff$ the coefficients of f are relatively prime $\iff$ $c(f)$ is a unit. By Lemma 4.1(i), every nonzero $f \in R[x]$ is of the form $c(f)f_0$, where $f_0 \in R[x]$ is primitive. If $r \in R$ and $f \in R[x]$ with $f \neq 0$, $r \neq 0$, then $c(rf) = rc(f)$, by Lemma 4.1(ii).

We now recall the statement of Theorem 1.2:

Let $f \in R[x]$ be a polynomial of degree $n \geq 1$. Then f is a product of two polynomials in $F[x]$ of degrees d and e respectively with $0 < d, e < n$ if and only if there exist polynomials $g, h \in R[x]$ of degrees d and e respectively with $0 < d, e < n$ such that $f = gh$.

As we noted earlier, the $\Longleftarrow$ direction is trivial. The proof of the $\implies$ direction is based on the following lemmas:

Lemma 4.3. *Suppose that f and g are two primitive polynomials in $R[x]$, and that there exists a nonzero $\alpha \in F$ such that $f = \alpha g$. Then $\alpha \in R$ and α is a unit, i.e. $\alpha \in R^*$.*

Proof. Write $\alpha = r/s$, with $r, s \in R$. Then $sf = rg$. Thus $c(sf) = sc(f) = s$ up to multiplying by a unit in R. Likewise $c(rg) = r$ up to multiplying by a unit in R. Since $sf = rg$ and content is well-defined up to multiplying by a unit in R, $r = us$ for some $u \in R^*$ and hence $r/s = \alpha = u$ is an element of R^*. □

Lemma 4.4. *Let $f \in F[x]$ with $f \neq 0$. Then there exists an $\alpha \in F^*$ such that $\alpha f \in R[x]$ and αf is primitive.*

Proof. Write $f = \sum_{i=0}^{n}(r_i/s_i)x^i$, where the $r_i, s_i \in R$ and, for all i, $s_i \neq 0$. If $s = s_0 \cdots s_n$, then $sf \in R[x]$, so we can write $sf = cf_0$, where $f_0 \in R[x]$ and f_0 is primitive. Then set $\alpha = s/c$, so that $\alpha f = f_0$, a primitive polynomial in $R[x]$ as desired. □

Lemma 4.5 (Gauss Lemma)**.** *Let $f, g \in R[x]$ be two primitive polynomials. Then fg is also primitive.*

Proof. We prove the contrapositive: If fg is not primitive, then one of f, g is not primitive. If fg is not primitive, then $c(fg)$ is not 0 or a unit, so there is an irreducible element $r \in R$ which divides all of the coefficients of fg. Consider the natural homomorphism from $R[x]$ to $(R/(r))[x]$, and as usual let the image of a polynomial $p \in R[x]$, i.e. the reduction of p mod (r), be denoted by $\bar{p}$. Thus, $\overline{(fg)} = 0 = \bar{f}\bar{g}$. Now $(R/(r))[x]$ is an integral domain, because (r) is a prime ideal and hence $R/(r)$ is an integral domain. Thus, since $\bar{f}\bar{g} = \overline{(fg)}$ is zero, one of $\bar{f}, \bar{g}$ is zero, say $\bar{f}$. It follows that r divides all of the coefficients of f, hence that f is not primitive as claimed. □

We just leave the following corollary of Lemma 4.4 as an exercise:

Corollary 4.6. *Let $f, g \in R[x]$ be two nonzero polynomials. Then $c(fg) = c(f)c(g)$.* □

Completion of the proof of Theorem 1.2. We may as well assume that f is primitive to begin with ($f = cf_0$ factors in $F[x] \implies f_0$ also factors in $F[x]$, and a factorization of $f_0 = gh$ in $R[x]$ gives one for f as $(cg)h$, say). Suppose that f is primitive and is a product of two polynomials g_1, h_1 in $F[x]$ of degrees $d, e < n$. Then, by Lemma 4.4, there exist $\alpha, \beta \in F^*$ such that $\alpha g_1 = g \in R[x]$ and $\beta h_1 = h \in R[x]$, where g and h are primitive. Clearly, $\deg g = \deg g_1$ and $\deg h = \deg h_1$. Then $\alpha\beta g_1 h_1 = (\alpha\beta)f = gh$.

By the Gauss Lemma, gh is primitive, and f was primitive by assumption. By Lemma 4.3, $\alpha\beta \in R$ and is a unit, say $\alpha\beta = u \in R^*$. Thus $f = u^{-1}gh$. Renaming $u^{-1}g$ by g gives a factorization of f in $R[x]$ as claimed. $\square$

The proof of Theorem 1.2 actually shows the following:

Corollary 4.7. *Let R be a UFD with quotient field F, let $f \in R[x]$ be a primitive polynomial, and let $g \in R[x]$. Then f divides g in $F[x]$ $\iff$ f divides g in $R[x]$.*

Proof. $\Longleftarrow$: This is obvious.

$\Longrightarrow$: Writing $g = c(g)g_0$, where g_0 is primitive, we see that f divides g_0 in $F[x]$ as well. Write $g_0 = fh$ for some $h \in F[x]$. By Lemma 4.4, there exists with $\alpha \in F^*$ such that $\alpha h = h_0$ is a primitive polynomial in $R[x]$. Then

$$\alpha g_0 = \alpha fh = f \cdot (\alpha h) = fh_0.$$

By Lemma 4.5, fh_0 is a primitive polynomial in $R[x]$, hence by Lemma 4.3 $\alpha \in R$ and α is a unit u. It follows that $g_0 = f \cdot (u^{-1}h_0)$, so that f divides g_0 and hence g. $\square$

For example, this gives a quick proof of the rational roots test: if $f = \sum_{i=1}^{n} a_i x^i \in R[x]$ is a polynomial of degree n and $f(r/s) = 0$, where r and s are relatively prime, then the linear polynomial $sx - r$ divides f in $F[x]$, and $sx - r$ is primitive since r and s are relatively prime. Hence $sx - r$ divides f in $R[x]$, which easily implies that $s|a_n$ and $r|a_0$.

Here is another corollary of Theorem 1.2:

Corollary 4.8. *Let R be a UFD with quotient field F and let $f \in R[x]$ be a primitive polynomial. Then f is irreducible in $F[x]$ $\iff$ f is irreducible in $R[x]$.*

Proof. $\Longrightarrow$: If f is irreducible in $F[x]$, then a factorization in $R[x]$ would have to be of the form $f = rg$ for some $r \in R$ and $g \in R[x]$. Then $c(f) = rc(g)$, and, since f is primitive, $c(f)$ is a unit. Hence r is a unit as well. Thus f is irreducible in $R[x]$.

$\Longleftarrow$: Conversely, if f is reducible in $F[x]$, then Theorem 1.2 implies that f is reducible in $R[x]$. $\square$

Very similar ideas can be used to prove the following:

Theorem 4.9. *Let R be a UFD with quotient field F. Then the ring $R[x]$ is a UFD. In fact, the irreducibles in $R[x]$ are exactly the $r \in R$ which are irreducible, and the primitive polynomials $f \in R[x]$ such that f is an irreducible polynomial in $F[x]$.*

Proof. There are three steps:

Step I: We claim that, if r is an irreducible element of R, then r is irreducible in $R[x]$ and that, if $f \in R[x]$ is a primitive polynomial which is irreducible in $F[x]$, then f is irreducible in $R[x]$. In other words, the elements described in the last sentence of the theorem are in fact irreducible. Clearly, if r is an irreducible element of R, then if r factors as gh with $g, h \in R[x]$, then $\deg g = \deg h = 0$, i.e. $g = s$ and $h = t$ are elements of the subring R of $R[x]$. Since r is irreducible in R, one of s, t is a unit in R and hence in $R[x]$. Thus r is irreducible in $R[x]$. Likewise, if $f \in R[x]$ is a primitive polynomial such that f is an irreducible polynomial in $F[x]$, then f is irreducible in $R[x]$ by Corollary 4.8.

Step II: We claim that every polynomial in $R[x]$ which is not zero or a unit in $R[x]$ (hence a unit in R) can be factored into a product of the elements listed in Step I. In fact, if $f \in R[x]$ is not 0 or a unit, we can write $f = c(f)f_0$, where $c(f) \in R$ and f_0 is primitive, and either $c(f)$ is not a unit or $\deg f_0 \geq 1$. If $c(f)$ is not a unit, it can be factored into a product of irreducibles in R. If $\deg f_0 \geq 1$, the f_0 can be factored in $F[x]$ into a product of irreducibles: $f_0 = g_1 \cdots g_k$, where the $g_i \in F[x]$ are irreducible. By Lemma 4.4, for each i there exists an $\alpha_i \in F^*$ such that $\alpha_i g_i = h_i \in R[x]$ and such that h_i is primitive. By the Gauss Lemma (Lemma 4.5), the product $h_1 \cdots h_k$ is also primitive. Then

$$(\alpha_1 \cdots \alpha_k)g_1 \cdots g_k = (\alpha_1 \cdots \alpha_k)f_0 = h_1 \cdots h_k.$$

Since both $h_1 \cdots h_k$ and f_0 are primitive, $\alpha_1 \cdots \alpha_k \in R$ and $\alpha_1 \cdots \alpha_k$ is a unit, by Lemma 4.3. Absorbing this factor into h_1, say, we see that f_0 is a product of primitive polynomials in $R[x]$.

Step III: Finally, we claim that the factorization is unique up to units. Suppose then that

$$f = r_1 \cdots r_a g_1 \cdots g_k = s_1 \cdots s_b h_1 \cdots h_\ell,$$

where the r_i and s_j are irreducible elements of R and the g_i, h_j are irreducible primitive polynomials in $R[x]$. Then $g_1 \cdots g_k$ and $h_1 \cdots h_\ell$ are both primitive, by the Gauss Lemma (Lemma 4.5). Hence $c(f)$, which is well-defined up to a unit, is equal to $r_1 \cdots r_a$ and also to $s_1 \cdots s_b$, i.e. $r_1 \cdots r_a = us_1 \cdots s_b$ for some unit $u \in R^*$. By unique factorization in R, $a = b$, and, after a permutation of the s_i, r_i and s_i are associates. Next, we consider the two factorizations of f in $F[x]$, and use the fact that the g_i, h_j are irreducible in $F[x]$, whereas the r_i, s_j are units. Unique factorization in $F[x]$ implies that $k = \ell$ and that, after a permutation of the h_i, for every i there exists a unit in $F[x]$, i.e. an element $\alpha_i \in F^*$, such that $g_i = \alpha_i h_i$. Since both g_i and h_i are primitive polynomials in $R[x]$, Lemma 4.3 implies that $\alpha_i \in R^*$ for every i, in other words that g_i and h_i are associates in $R[x]$. Hence the two factorizations of f are unique up to order and units. $\square$

Corollary 4.10. *Let R be a UFD. Then the ring $R[x_1, \dots, x_n]$ is a UFD. In particular, $\mathbb{Z}[x_1, \dots, x_n]$ and $F[x_1, \dots, x_n]$, where F is a field, are UFD's.*

Proof. This is immediate from Theorem 4.9 by induction. $\square$

5 Algebraic curves

We now discuss a special case which is relevant for algebraic geometry. Here $R = K[y]$ for some field K and hence $F = K(y)$. Thus $R[x] = K[x, y]$. In studying geometry, we often assume that K is algebraically closed, for example $K = \mathbb{C}$. For questions related to number theory we often take $K = \mathbb{Q}$. By Theorem 4.9, $K[x, y]$ is a UFD. To avoid confusion, we will usually write an element of $K[x, y]$ as $f(x, y)$; similarly an element of $K[x]$ or $K[y]$ will be written as $g(x)$ or $h(y)$.

A *plane algebraic curve* is a subset C of $K^2 = K \times K$, often written as $V(f)$, defined by the vanishing of a polynomial $f(x, y) \in K[x, y]$:

$$C = V(f) = \{(a, b) \in K^2 : f(a, b) = 0\}.$$

This situation is familiar from one variable calculus, where we take $K = \mathbb{R}$ and view $f(x, y) = 0$ as defining y "implicitly" as a function of x. For example, the function $y = \sqrt{1 - x^2}$ is implicitly defined by the polynomial $f(x, y) = x^2 + y^2 - 1$. A function y which can be implicitly so defined is called an *algebraic function*. In general, however, the equation $f(x, y) = 0$ defines many different functions, at least locally: for example, $f(x, y) = x^2 + y^2 - 1$ also defines the function $y = -\sqrt{1 - x^2}$. Over $\mathbb{C}$, or fields other than $\mathbb{R}$, it is usually impossible to sort out these many different functions, and it is best to work with the geometric object C. (In complex analysis, we speak of trying to define the–not single valued–"function" $\sqrt{z}$, by taking the corresponding Riemann surface, which in this case is the plane algebraic curve $w^2 = z$.)

If $f(x, y)$ is irreducible in $K[x, y]$, we call $C = V(f)$ an *irreducible plane curve*. Since $K[x, y]$ is a UFD, an arbitrary $f(x, y)$ can be factored into its irreducible factors: $f(x, y) = f_1(x, y) \cdots f_n(x, y)$, where the $f_i(x, y)$ are irreducible elements of $K[x, y]$. It is easy to see from the definition that

$$C = V(f) = V(f_1) \cup \cdots \cup V(f_n) = C_1 \cup \cdots \cup C_n,$$

where $C_i = V(f_i)$ is defined by the vanishing of the factor $f_i(x, y)$. We call the C_i the ***irreducible components*** of C. Thus, the irreducible plane curves are the basic building blocks for all plane curves and we want to be able to decide if a given polynomial $f(x, y) \in K[x, y]$ is irreducible. Restating Theorem 4.9 gives:

Theorem 5.1. *A polynomial $f(x, y) \in K[x, y]$ is irreducible $\iff$ $f(x, y)$ is primitive in $K[y][x]$ (i.e. writing $f(x, y)$ as a polynomial $a_n(y)x^n + \cdots + a_0(y)$ in x whose coefficients are polynomials in y, the polynomials $a_n(y), \ldots, a_0(y)$ are relatively prime) and $f(x, y)$ does not factor as a product of two polynomials of strictly smaller degrees in $K(y)[x]$.* $\square$

Example 5.2. (1) Let $f(x, y) = x^2 - g(y)$, where $g(y)$ is a polynomial in y which is not a perfect square in $K[y]$, for example any polynomial which has at least one non-multiple root. We claim that $f(x, y)$ is irreducible in $K[x, y]$. Since it is clearly primitive as an element of $K[y][x]$ (the coefficient of x^2 is 1), it suffices to prove that $f(x, y)$ is irreducible as an element of $K(y)[x]$. Since $f(x, y)$ has degree two in x, it is irreducible $\iff$ it has no root in $K(y)$. By the Rational Roots Test, a root of $x^2 - g(y)$ in $K(y)$ can be written as $p(y)/q(y)$, where $p(y)$ and $q(y)$ are relatively prime polynomials and $q(y)$ divides 1, i.e. $q(y)$ is a unit in $K[y]$, which we may assume is 1. Hence a root of $x^2 - g(y)$ in $K(y)$ would be of the form $p(y) \in K[y]$, in other words $g(y) = (p(y))^2$ would be a perfect square in $K[y]$. As we assumed that this was not the case, $f(x, y)$ is irreducible in $K[x, y]$.

(2) Consider the Fermat polynomial $f(x, y) = x^n + y^n - 1 \in K[x, y]$, where we view $K[x, y]$ as $K[y][x]$. The coefficients of $f(x, y)$ (viewed as a polynomial in x) are $a_n(y) = 1$ and $a_0(y) = y^n - 1$, so the gcd of the coefficients is 1. Hence $f(x, y)$ is primitive in $R[x]$.

Theorem 5.3. *If* $\operatorname{char} F = 0$ *or if* $\operatorname{char} F = p$ *and p does not divide n, then $f(x, y) = x^n + y^n - 1$ is irreducible in $K[x, y]$.*

Proof. Note that the constant term $y^n - 1$ factors as

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

We apply the Eisenstein criterion to $f(x, y)$, with $M = (y - 1)$. Clearly M is a maximal ideal in $K[y]$ since $y - 1$ is irreducible; in fact $M = \operatorname{Ker} \operatorname{ev}_1$. We can apply the Eisenstein criterion to $f(x, y)$ since $1 \notin M$, provided that $y^n - 1 \notin M^2$, or equivalently $y^{n-1} + \cdots + 1 \notin M$. But $y^{n-1} + \cdots + 1 \in M \iff \operatorname{ev}_1(y^{n-1} + \cdots + 1) = 0$. Now $\operatorname{ev}_1(y^{n-1} + \cdots + 1) = 1 + \cdots + 1 = n$, and this is zero in $K \iff \operatorname{char} K = p$ and p divides n. $\square$

We note that if $\operatorname{char} K = p$ and, for example, if $n = p$, then $x^p + y^p - 1 = (x + y - 1)^p$ and so is reducible; a similar statement holds if we just assume that $p|n$.