CHAPTER 10

Unique Factorization

In this chapter, we are interested in the question of factorization in integral domains. More precisely, we are interested in whether factorization exists, and if so, whether it is unique. To do this, we first of all need to generalize the notion of divisibility.

DEFINITION. Suppose that R is a commutative ring with unity. Suppose further that $a, b \in R$ and $a \neq 0$. Then we say that a divides b, denoted by $a \mid b$, if there exists $c \in R$ such that $b = ac$. In this case, we also say that a is a divisor of b, or that b is a multiple of a.

EXAMPLES. (1) We are already familiar with divisibility in the ring $\mathbb{Z}$ and in polynomial rings of the type $R[X]$, where R is a commutative ring with unity.

(2) Suppose that $a, b, c \in R$. If $a \mid b$ and $b \mid c$, then $a \mid c$. To see this, note that if $a \mid b$ and $b \mid c$, then there exist $m, n \in R$ such that $b = am$ and $c = bn$, so that $c = amn$.

10.1. A Simple Case

Throughout this section, F denotes a field. We shall discuss the case of polynomial rings of the form $F[X]$.

PROPOSITION 10.1. *Suppose that $a(X), b(X) \in F[X]$, not both zero. Then there exists a unique monic polynomial $d(X) \in F[X]$ such that*

(i) *$d(X) \mid a(X)$ and $d(X) \mid b(X)$;*
(ii) *if $c(X) \in F[X]$ and $c(X) \mid a(X)$ and $c(X) \mid b(X)$, then $c(X) \mid d(X)$; and*
(iii) *there exist polynomials $s(X), t(X) \in F[X]$ such that $d(X) = a(X)s(X) + b(X)t(X)$.*

PROOF. Consider the set

$$I = \{a(X)u(X) + b(X)v(X) : u(X), v(X) \in F[X]\}.$$

It is not difficult to check that I is an ideal in $F[X]$. It follows from Theorem 9.1 that $I = \langle d(X) \rangle$ for some $d(X) \in F[X]$. Clearly we may assume without loss of generality that $d(X)$ is monic. Since $d(X) \in I$, there exist $s(X), t(X) \in F[X]$ such that

$$d(X) = a(X)s(X) + b(X)t(X).$$

Suppose that $c(X) \in F[X]$ and $c(X) \mid a(X)$ and $c(X) \mid b(X)$. Then there exist $f(X), g(X) \in F[X]$ such that $a(X) = c(X)f(X)$ and $b(X) = c(X)g(X)$. It follows that

$$d(X) = c(X)(f(X)s(X) + g(X)t(X)),$$

so that $c(X) \mid d(X)$. Note also that $a(X), b(X) \in I$ (why?), so that $d(X) \mid a(X)$ and $d(X) \mid b(X)$. Finally, the uniqueness of $d(X)$ follows from (i) and (ii), since if $d_1(X)$ and $d_2(X)$ both satisfy the requirements of $d(X)$, then we must have $d_1(X) \mid d_2(X)$ and $d_2(X) \mid d_1(X)$. Since both $d_1(X)$ and $d_2(X)$ are monic, it follows that $d_1(X) = d_2(X)$. ◯

DEFINITION. The polynomial $d(X)$ in Proposition 10.1 is known as the greatest common divisor of the polynomials $a(X)$ and $b(X)$, and denoted by $d(X) = (a(X), b(X))$.

PROPOSITION 10.2. *Suppose that $a(X), b(X) \in F[X]$, and that $p(X) \in F[X]$ is irreducible. If $p(X) \mid a(X)b(X)$, then $p(X) \mid a(X)$ or $p(X) \mid b(X)$.*

PROOF. Suppose that $p(X) \mid a(X)b(X)$. Suppose further that $p(X) \nmid a(X)$. Since $p(X)$ is irreducible, the only divisors of $p(X)$ are non-zero elements of F or $cp(X)$, where $c \in F$ is non-zero. Clearly $cp(X) \nmid a(X)$ for any non-zero $c \in F$. Hence we must have $1 = (a(X), p(X))$. It follows from Proposition 10.1 that there exist $s(X), t(X) \in F[X]$ such that

$$1 = a(X)s(X) + p(X)t(X),$$

so that

$$b(X) = a(X)b(X)s(X) + p(X)b(X)t(X).$$

Clearly $p(X) \mid b(X)$. ◯

Using Proposition 10.2 a finite number of times, we have

PROPOSITION 10.3. *Suppose that $a_1(X), \ldots, a_k(X) \in F[X]$, and that $p(X) \in F[X]$ is irreducible. If $p(X) \mid a_1(X) \ldots a_k(X)$, then $p(X) \mid a_i(X)$ for some $i = 1, \ldots, k$.*

PROPOSITION 10.4. *Suppose that $a(X) \in F[X]$ is a non-constant monic polynomial. Then $a(X)$ is representable as a finite product of irreducible monic polynomial factors in $F[X]$, uniquely up to the order of factors.*

SKETCH OF PROOF. Existence of factorization can be demonstrated either by using the fact that every ideal in $F[X]$ is *principal*, as in Theorem 10.6, or by using the degrees of polynomials in $F[X]$, as in Theorem 10.12. Uniqueness of factorization is demonstrated by using Proposition 10.3, as in the usual proof of uniqueness of factorization in $\mathbb{N}$. ◯

10.2. Principal Ideal Domains

Note that in the previous section, the main idea in the proof of Proposition 10.1 is Theorem 9.1. In other words, we depend on the fact that any ideal in $F[X]$ is generated by a single polynomial. In this section, we shall generalize our argument to prove existence and uniqueness of factorization in integral domains in which every ideal is *principal*.

We need a few definitions.

DEFINITION. Suppose that R is a commutative ring with unity. An element $u \in R$ is said to be a unit in R if it has multiplicative inverse in R. Two elements $a, b \in R$ are said to be associates if there exists a unit $u \in R$ such that $a = ub$.

EXAMPLES. (1) In the ring $\mathbb{Z}$, 1 and -1 are the only units. Also, for every $a \in \mathbb{Z}$, a and $-a$ are associates.

(2) In $F[X]$ where F is a field, every non-zero element of F is a unit. Also the polynomials $X + 3$ and $27X + 81$ are associates in $\mathbb{R}[X]$.

DEFINITION. Suppose that R is an integral domain. A non-zero non-unit element $p \in R$ is said to be irreducible in R if the following condition is satisfied:

(U) For every $a, b \in R$ such that $p = ab$, either a or b is a unit.

DEFINITION. Suppose that R is an integral domain. A non-zero non-unit element $p \in R$ is said to be prime in R if the following condition is satisfied:

(P) For every $a, b \in R$ such that $p \mid ab$, either $p \mid a$ or $p \mid b$.

EXAMPLE. Consider the integral domain $\mathbb{Z}[X]$ and the polynomial $p(X) = 2X + 6$. Then $p(X)$ is reducible, since $p(X) = 2(X + 3)$, and neither 2 nor $(X + 3)$ is a unit. Also $p(X)$ is not prime, for let $a(X) = 4$ and $b(X) = X + 3$. Then $p(X) \mid a(X)b(X)$. But $p(X) \nmid a(X)$ and $p(X) \nmid b(X)$.

DEFINITION. Suppose that R is an integral domain.

(i) We say that R is a principal ideal domain (PID) if every ideal in R is of the form $\langle a \rangle$, where $a \in R$.

(ii) We say that R is a unique factorization domain (UFD) if every non-zero non-unit element of R can be represented as a finite product of irreducible elements of R, uniquely up to order and associates.

Our aim here is to prove that every principal ideal domain is a unique factorization domain. To prove existence of factorization, we have the following result.

THEOREM 10.5. *Suppose that R is a principal ideal domain. Then every non-zero non-unit element of R can be represented as a finite product of irreducible elements of R.*

We shall state and prove Theorem 10.5 in a slightly different form.

THEOREM 10.6. *Suppose that R is a principal ideal domain, and that the non-zero non-unit element $y \in R$ cannot be represented as a finite product of irreducible elements of R. Then*

(i) *there exist two infinite sequences $y_1, y_2, y_3, \ldots$ and $q_1, q_2, q_3, \ldots$ of non-zero non-unit elements in R such that*

$$y = y_1 q_1 = y_1 y_2 q_2 = y_1 y_2 y_3 q_3 = \ldots;$$

(ii) *for every $k \in \mathbb{N}$, the principal ideals $\langle q_k \rangle$ satisfy $\langle q_k \rangle \subsetneqq \langle q_{k+1} \rangle$;*

(iii) *the union*

$$I = \bigcup_{k=1}^{\infty} \langle q_k \rangle$$

is an ideal in R; and

(iv) *there exists $j \in \mathbb{N}$ such that $\langle q_j \rangle = \langle q_{j+1} \rangle = \ldots = I$.*

PROOF OF THEOREM 10.5. Simply note that (iv) contradicts (ii). ◯

PROOF OF THEOREM 10.6. (i) Clearly y cannot be irreducible. It follows that there exist non-zero non-unit $y_1, q_1 \in R$ such that $y = y_1 q_1$. Also at least one of y_1, q_1 cannot be represented as a finite product of irreducible elements of R. Without loss of generality, we assume that q_1 cannot be represented as a finite product of irreducible elements of R. Clearly q_1 cannot be irreducible. It follows that there exist non-zero non-unit $y_2, q_2 \in R$ such that $q_1 = y_2 q_2$. Also at least one of y_2, q_2 cannot be represented as a finite product of irreducible elements of R. Without loss of generality, we assume that q_2 cannot be represented as a finite product of irreducible elements of R. And so on.

(ii) Since $q_k = y_{k+1} q_{k+1}$, we clearly have $q_k \in \langle q_{k+1} \rangle$, so that $\langle q_k \rangle \subseteq \langle q_{k+1} \rangle$. Suppose on the contrary that $\langle q_k \rangle = \langle q_{k+1} \rangle$. Then $q_{k+1} = u q_k$ for some $u \in R$, so that $y_{k+1} u = 1$, whence y_{k+1} is a unit, a contradiction.

(iii) Suppose that $a, b \in I$. Then there exist $k_1, k_2 \in \mathbb{N}$ such that $a \in \langle q_{k_1} \rangle$ and $b \in \langle q_{k_2} \rangle$. Let $k = \max\{k_1, k_2\}$. Then $a, b \in \langle q_k \rangle$. Hence we have $a - b \in \langle q_k \rangle \subseteq I$. Also, for any $x \in R$, we have $xa, ax \in \langle q_k \rangle \subseteq I$.

(iv) Since R is a principal ideal domain, there exists $q \in R$ such that

$$\langle q \rangle = I = \bigcup_{k=1}^{\infty} \langle q_k \rangle.$$

Note that $q \in I$. It follows that $q \in \langle q_j \rangle$ for some $j \in \mathbb{N}$, so that $\langle q \rangle \subseteq \langle q_j \rangle$. On the other hand, we clearly have $\langle q_k \rangle \subseteq \langle q \rangle$ for every $k \in \mathbb{N}$. Hence for every $k \geqslant j$, we have $\langle q \rangle \subseteq \langle q_j \rangle \subseteq \langle q_k \rangle \subseteq \langle q \rangle$, so that $\langle q_k \rangle = \langle q \rangle$. ◯

To prove uniqueness of factorization, we need the following three results which are generalizations of Propositions 10.1, 10.2 and 10.3.

THEOREM 10.7. *Suppose that R is a principal ideal domain, and that $a, b \in R$, not both zero. Then there exists an element $d \in R$, unique up to associates, such that*

(i) *$d \mid a$ and $d \mid b$;*

(ii) *if $c \in R$ and $c \mid a$ and $c \mid b$, then $c \mid d$; and*

(iii) *there exist $s, t \in R$ such that $d = as + bt$.*

PROOF. Consider the set

$$I = \{au + bv : u, v \in R\}.$$

It is not difficult to check that I is an ideal in R. Since R is a principal ideal domain, there exists $d \in R$ such that $I = \langle d \rangle$. Since $d \in I$, there exist $s, t \in R$ such that $d = as + bt$. Suppose that $c \in R$ and $c \mid a$ and $c \mid b$. Then there exist $f, g \in R$ such that $a = cf$ and $b = cg$. It follows that

$d = c(fs + gt)$, so that $c \mid d$. Note also that $a, b \in I$, so that $d \mid a$ and $d \mid b$. Finally, the uniqueness of d up to assiciates follows from (i) and (ii), since if d_1 and d_2 both satisfy the requirements of d, then we must have $d_1 \mid d_2$ and $d_2 \mid d_1$, so that d_1 and d_2 are associates. ◯

DEFINITION. The element d in Theorem 10.7 is known as a greatest common divisor of the elements a and b, and denoted by $d = (a, b)$.

THEOREM 10.8. *Suppose that R is a principal ideal domain. Suppose further that $a, b \in R$, and that $p \in R$ is irreducible. If $p \mid ab$, then $p \mid a$ or $p \mid b$.*

PROOF. Suppose that $p \mid ab$. Suppose further that $p \nmid a$. Since p is irreducible, the only divisors of p are units or associates of p. Clearly no associate of p will divide a. Hence we must have $1 = (a, p)$. It follows from Theorem 10.7 that there exist $s, t \in R$ such that $1 = as + pt$, so that $b = abs + pbt$. Clearly $p \mid b$. ◯

Using Theorem 10.8 a finite number of times, we have

THEOREM 10.9. *Suppose that R is a principal ideal domain. Suppose further that $a_1, \ldots, a_k \in R$, and that $p \in R$ is irreducible. If $p \mid a_1 \ldots a_k$, then $p \mid a_i$ for some $i = 1, \ldots, k$.*

THEOREM 10.10. *Suppose that R is a principal ideal domain. Then the representation of any non-zero non-unit element of R as a finite product of irreducible elements of R is unique up to order and associates.*

PROOF. Suppose that

$$p_1 \ldots p_k = q_1 \ldots q_m,$$

where $p_1, \ldots, p_k, q_1, \ldots, q_m$ are all irreducible in R. Since $p_1 \mid q_1 \ldots q_m$, it follows from Theorem 10.9 that $p_1 \mid q_i$ for some $i = 1, \ldots, m$. Without loss of generality, we may assume that $p_1 \mid q_1$. Since q_1 is irreducible, it follows that p_1 and q_1 are associates. Hence

$$p_2 \ldots p_k = u q_2 \ldots q_m$$

for some unit $u \in R$. We now continue by induction, noting that $p \nmid u$ for any irreducible $p \in R$ and unit $u \in R$. ◯

10.3. The Simple Case Again

Let us briefly return to the case of $F[X]$ where F is a field. In Section 10.1, Proposition 10.2 is proved using Proposition 10.1, while the proof of Proposition 10.1 hinges on the fact that $F[X]$ is a principal ideal domain. Let us now give an alternative proof of Proposition 10.2 without using Proposition 10.1 or the fact that $F[X]$ is a principal ideal domain.

Recall Proposition 8.6, which we restate below in the notation of this chapter.

PROPOSITION 10.11. *Suppose that $a(X), b(X) \in F[X]$, and that $a(X) \neq 0$. Then there exist unique polynomials $q(X), r(X) \in F[X]$ such that*

(i) *$b(X) = a(X)q(X) + r(X)$; and*
(ii) *either $r(X) = 0$ or $\deg r(X) < \deg a(X)$.*

ALTERNATIVE PROOF OF PROPOSITION 10.2. Suppose that $p(X) \nmid a(X)$ in $F[X]$. Let

$$S = \{b(X) \in F[X] : p(X) \mid a(X)b(X),\ p(X) \nmid b(X)\}.$$

Clearly it is sufficient to show that $S = \emptyset$. Suppose, on the contrary, that $S \neq \emptyset$. Since $0 \notin S$, there exists a non-zero $c(X) \in S$ of smallest degree. In particular,

$$p(X) \mid a(X)c(X) \quad \text{and} \quad p(X) \nmid c(X).$$

Since $p(X) \nmid a(X)$, $c(X)$ must be non-constant, so that $\deg c(X) \geqslant 1$. It follows that $c(X)$ is not a unit in $F[X]$ (why?). On the other hand, since $p(X) \nmid c(X)$, $c(X)$ is not an associate of $p(X)$. Next, we can assume that $\deg c(X) < \deg p(X)$, for otherwise write $c(X) = p(X)u(X) + v(X)$, where $u(X), v(X) \in F[X]$ and where $v(X) = 0$ or $\deg v(X) < \deg p(X)$. Clearly $v(X) \neq 0$, and it is easily shown that $v(X) \in S$, contradicting the minimality of $\deg c(X)$. By Proposition 10.11, there exist $q(X), r(X) \in F[X]$ such that $p(X) = c(X)q(X) + r(X)$, where $r(X) = 0$ or $\deg r(X) < \deg c(X)$. If

$r(X) = 0$, then $p(X) = c(X)q(X)$, contradicting the assumption that $p(X)$ is irreducible in $F[X]$. Hence $r(X) \neq 0$. Note now that

$$p(X) \mid a(X)r(X) \quad \text{and} \quad \deg r(X) < \deg c(X).$$

Also, it follows from $\deg r(X) < \deg c(X) < \deg p(X)$ that $p(X) \nmid r(X)$. Hence $r(X) \in S$, contradicting the minimality of $\deg c(X)$. ◯

10.4. Euclidean Domains

Note that our alternative proof of Proposition 10.2 depends heavily on the use of the degree of polynomials. Indeed, if we think of Proposition 10.2 as the analogue of the corresponding result in $\mathbb{Z}$, then the only essential difference in the proof is that whereas we consider the magnitude of numbers in $\mathbb{Z}$, we consider the degree of polynomials in $F[X]$. It follows that if we wish to consider more general situations, we need a notion that will incorporate these two cases as special cases.

DEFINITION. Suppose that R is an integral domain. Then we say that R is a euclidean domain if there exists a function $N : R \to \mathbb{N} \cup \{0\}$ which satisfies the following two conditions:

(ED1) For every non-zero $a, b \in R$, we have $N(a) \leqslant N(b)$ whenever $a \mid b$.

(ED2) For every $a, b \in R$ with $a \neq 0$, there exist $q, r \in R$ such that $b = aq + r$, where either $r = 0$ or $N(r) < N(a)$.

REMARKS. (1) The function $N : R \to \mathbb{N} \cup \{0\}$ is sometimes called a euclidean function.

(2) Suppose that R is a euclidean domain. Suppose further that non-zero $a, b \in R$ are not associates and $a \mid b$. Then $N(a) < N(b)$.

EXAMPLES. (1) In $\mathbb{Z}$, we can take the function $N(a) = |a|$ for every $a \in \mathbb{Z}$.

(2) In $F[X]$, where F is a field, we can take the function $N(a(X)) = \deg a(X)$, in view of Proposition 10.11.

(3) One of the most important examples of a euclidean domain is the ring

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

of all gaussian integers. To show this, define a function $N : \mathbb{Z}[\mathrm{i}] \to \mathbb{N} \cup \{0\}$ by writing $N(a+b\mathrm{i}) = a^2+b^2$ for every $a, b \in \mathbb{Z}$. Suppose that $(a + b\mathrm{i}) \mid (c + d\mathrm{i})$. Then there exists $e + f\mathrm{i} \in \mathbb{Z}[\mathrm{i}]$ such that $(c + d\mathrm{i}) = (a + b\mathrm{i})(e + f\mathrm{i})$. Note then that $c = ae - bf$ and $d = af + be$, so that

$$N(c + d\mathrm{i}) = (ae - bf)^2 + (af + be)^2 = (e^2 + f^2)N(a + b\mathrm{i}) \geqslant N(a + b\mathrm{i}),$$

since clearly $e^2 + f^2 \geqslant 1$ if $e + f\mathrm{i} \neq 0$. This gives (ED1). The proof of (ED2) is harder, and we shall cheat briefly to illustrate the ideas. Suppose that we try to divide $a + b\mathrm{i}$ by $c + d\mathrm{i}$ to get a main term and a remainder. Going to $\mathbb{Q}(\mathrm{i}) = \{a + b\mathrm{i} : a, b \in \mathbb{Q}\}$, we can think of

$$\frac{a + b\mathrm{i}}{c + d\mathrm{i}} = \frac{(a + b\mathrm{i})(c - d\mathrm{i})}{c^2 + d^2} = \frac{ac + bd}{c^2 + d^2} + \frac{bc - ad}{c^2 + d^2}\mathrm{i}.$$

We can write

$$\frac{ac + bd}{c^2 + d^2} = q + \frac{r}{c^2 + d^2} \quad \text{and} \quad \frac{bc - ad}{c^2 + d^2} = s + \frac{t}{c^2 + d^2},$$

where $q, r, s, t \in \mathbb{Z}$ and

$$\left|\frac{r}{c^2 + d^2}\right| \leqslant \frac{1}{2} \quad \text{and} \quad \left|\frac{t}{c^2 + d^2}\right| \leqslant \frac{1}{2}.$$

Then

$$a + b\mathrm{i} = (c + d\mathrm{i})(q + s\mathrm{i}) + (c + d\mathrm{i})\left(\frac{r}{c^2 + d^2} + \frac{t}{c^2 + d^2}\mathrm{i}\right).$$

Clearly

$$(c + d\mathrm{i})\left(\frac{r}{c^2 + d^2} + \frac{t}{c^2 + d^2}\mathrm{i}\right) \in \mathbb{Z}[\mathrm{i}].$$

(ED2) now follows from

$$N\left((c + d\mathrm{i})\left(\frac{r}{c^2 + d^2} + \frac{t}{c^2 + d^2}\mathrm{i}\right)\right) \leqslant \frac{1}{2}(c^2 + d^2) < N(c + d\mathrm{i}).$$

Our aim here is to prove that every euclidean domain is a unique factorization domain. To prove existence of factorization, we have the following result.

THEOREM 10.12. *Suppose that R is a euclidean domain. Then every non-zero non-unit element of R can be represented as a finite product of irreducible elements of R.*

PROOF. We shall prove the result by induction on $N(a)$, where $N : R \to \mathbb{N} \cup \{0\}$ is a euclidean function. In view of Remark (2) above, if $a \in R$ is not irreducible, then clearly $N(a) \geqslant 1$. It follows that every non-zero non-unit element $a \in R$ with $N(a) = 0$ is irreducible, and obviously can be represented as a finite product of irreducible elements of R. Suppose now that $k \in \mathbb{N}$, and that every non-zero non-unit element $b \in R$ with $N(b) < k$ can be represented as a finite product of irreducible elements of R. Then for every non-zero non-unit $a \in R$ satisfying $N(a) = k$, either a is irreducible or there exist non-zero non-unit $b_1, b_2 \in R$ such that $a = b_1 b_2$. Clearly $N(b_1) < k$ and $N(b_2) < k$, so that b_1 and b_2 can both be represented as a finite product of irreducible elements of R. In either case, a can be represented as a finite product of irreducible elements of R. ◯

To prove uniqueness of factorization, we need the following two results which are generalizations of Propositions 10.2 and 10.3.

THEOREM 10.13. *Suppose that R is a euclidean domain. Suppose further that $a, b \in R$, and that $p \in R$ is irreducible. If $p \mid ab$, then $p \mid a$ or $p \mid b$.*

PROOF. Suppose that $p \nmid a$ in R. Let

$$S = \{b \in R : p \mid ab,\ p \nmid b\}.$$

Clearly it is sufficient to show that $S = \emptyset$. Suppose, on the contrary, that $S \neq \emptyset$. Since $0 \notin S$, there exists a non-zero $c \in S$ for which $N(c)$ is minimal. In particular,

$$p \mid ac \quad \text{and} \quad p \nmid c.$$

Since $p \nmid a$, it follows that c is not a unit in R. On the other hand, since $p \nmid c$, it follows that c is not an associate of p. Next, we can assume that $N(c) < N(p)$, for otherwise, by (ED2), we can write $c = pu + v$, where $u, v \in R$ and where $v = 0$ or $N(v) < N(p)$. Clearly $v \neq 0$, and it is easily shown that $v \in S$, contradicting the minimality of $N(c)$. By (ED2), there exist $q, r \in R$ such that $p = cq + r$, where $r = 0$ or $N(r) < N(c)$. If $r = 0$, then $p = cq$, contradicting the assumption that p is irreducible in R. Hence $r \neq 0$. Note now that

$$p \mid ar \quad \text{and} \quad N(r) < N(c).$$

Also, it follows from $N(r) < N(c) < N(p)$ that $p \nmid r$. Hence $r \in S$, contradicting the minimality of $N(c)$. ◯

Using Theorem 10.13 a finite number of times, we have

THEOREM 10.14. *Suppose that R is a euclidean domain. Suppose further that $a_1, \ldots, a_k \in R$, and that $p \in R$ is irreducible. If $p \mid a_1 \ldots a_k$, then $p \mid a_i$ for some $i = 1, \ldots, k$.*

We can now deduce the following theorem from Theorem 10.14 in the same way as we deduce Theorem 10.10 from Theorem 10.9.

THEOREM 10.15. *Suppose that R is a euclidean domain. Then the representation of any non-zero non-unit element of R as a finite product of irreducible elements of R is unique up to order and associates.*

10.5. A Shortcut

Indeed, the argument in Section 10.4 appears to have a great deal in common with the argument in Section 10.2. This is explained by the following important result, which gives a shortcut to the result that every euclidean domain is a unique factorization domain. We have the following generalization of Theorem 9.1.

THEOREM 10.16. *Every euclidean domain is a principal ideal domain.*

PROOF. Let R be a euclidean domain, and let I be an ideal in R. If $I = \{0\}$, then clearly $I = \langle 0 \rangle$. Suppose now that I is a non-zero ideal. Let $a \in I$ be non-zero and such that $N(a)$ is minimal. We shall show that $I = \langle a \rangle$. Clearly $\langle a \rangle \subseteq I$. To show that $I \subseteq \langle a \rangle$, note that $0 \in \langle a \rangle$. Also, if $b \in I$ is non-zero, then by (ED2), there exist $q, r \in R$ such that $b = aq + r$, where $r = 0$ or $N(r) < N(a)$. If $r \neq 0$, then clearly $r \in I$, contradicting the minimality of $N(a)$. Hence $r = 0$, so that $b = aq \in \langle a \rangle$, as required. ◯

EXAMPLE. $\mathbb{Z}[X]$ is not a euclidean domain. In fact, we can show that $\mathbb{Z}[X]$ is not a principal ideal domain. It is easy to show that

$$I = \{2u(X) + Xv(X) : u(X), v(X) \in \mathbb{Z}[X]\}$$

is an ideal in $\mathbb{Z}[X]$. Suppose that $I = \langle a(X) \rangle$, where $a(X) \in \mathbb{Z}[X]$. Then since $2 \in I$, we must have $2 = a(X)r(X)$ for some $r(X) \in \mathbb{Z}[X]$. Hence $a(X) \in \mathbb{Z}$. On the other hand, since $X \in I$, we must have $X = a(X)s(X)$ for some $s(X) \in \mathbb{Z}[X]$. It follows that we must have $a(X) = \pm 1$, so that $\langle a(X) \rangle = \mathbb{Z}[X]$. But then $5 \notin I$.

10.6. Two Remarks

The first remark concerns the integral domain $\mathbb{Z}[X]$. We have shown that $\mathbb{Z}[X]$ is not a euclidean domain nor a principle ideal domain. Nevertheless, it can be shown that $\mathbb{Z}[X]$ is a unique factorization domain. However, this result naturally does not follow from Theorems 10.5 and 10.6. There is, in fact, an approach which uses the facts that $\mathbb{Z}$ is contained in $\mathbb{Q}$ and that $\mathbb{Q}[X]$ is a unique factorization domain. Let $a(X) \in \mathbb{Z}[X]$. The idea is to factorize $a(X)$ in $\mathbb{Q}[X]$ instead, and then attempt to carry this factorization back to $\mathbb{Z}[X]$.

The second remark concerns the origin of ideal theory. Consider the integral domain

$$\mathbb{Z}[\sqrt{15}] = \{a + b\sqrt{15} : a, b \in \mathbb{Z}\}.$$

Here we have two essentially different factorizations into products of irreducibles

$$10 = 2 \cdot 5 = (5 + \sqrt{15})(5 - \sqrt{15}).$$

It follows that $\mathbb{Z}[\sqrt{15}]$ is not a unique factorization domain. Indeed, non-uniqueness of factorization gives rise to many difficulties in mathematics. Perhaps the most celebrated case is *Fermat's last theorem*, that for every $n \in \mathbb{N}$ satisfying $n \geqslant 3$, there is no solution of the equation

$$x^n + y^n = z^n$$

with $x, y, z \in \mathbb{Z} \setminus \{0\}$. It is generally believed that Fermat wrongly thought that he could prove this famous assertion because he did not realize that in the ring of integers of some number fields, there is no uniqueness of factorization. In order to study this very difficult problem, Kummer introduced the notion of ideal theory and managed to prove a unique factorization theorem of ideals. However, this is rather difficult algebraic number theory, and Fermat's last theorem resisted many attempts to establish it.

We conclude by saying that Fermat's last theorem is now established, through the fundamental work of the English mathematician Sir Andrew Wiles.

Problems for Chapter 10

1. Prove that $\mathbb{Z}$ is a principal ideal domain.

2. Suppose that R is a field. Is R a principal ideal domain?

3. Suppose that F is a field, and consider the polynomial ring $F[X]$. Show that the function $N : F[X] \to \mathbb{N} \cup \{0\}$, defined for every $a(X) \in F[X]$ by $N(a(X)) = 2^{\deg a(X)}$, is a euclidean function.

4. Show that $\mathbb{Z}[\sqrt{2}] = \{a + b\sqrt{2} : a, b \in \mathbb{Z}\}$ is a euclidean domain by considering the function $N : \mathbb{Z}[\sqrt{2}] \to \mathbb{N} \cup \{0\}$, defined for every $a, b \in \mathbb{Z}$ by $N(a + b\sqrt{2}) = |a^2 - 2b^2|$.

5. Consider the ring $\mathbb{Z}[\mathrm{i}]$ of all gaussian integers, and let $N : \mathbb{Z}[\mathrm{i}] \to \mathbb{N} \cup \{0\}$ be defined for every $a, b \in \mathbb{Z}$ by $N(a + b\mathrm{i}) = a^2 + b^2$.

   (i) Show that N is multiplicative, *i.e.* $N(\alpha\beta) = N(\alpha)N(\beta)$ for every $\alpha, \beta \in \mathbb{Z}[\mathrm{i}]$.
   (ii) Show that $u \in \mathbb{Z}[\mathrm{i}]$ is a unit if and only if $N(u) = 1$.
   (iii) Show that ± 1 and $\pm\mathrm{i}$ are the only units in $\mathbb{Z}[\mathrm{i}]$.
   (iv) Show that $\alpha \in \mathbb{Z}[\mathrm{i}]$ is irreducible in $\mathbb{Z}[\mathrm{i}]$ if $N(\alpha)$ is a prime in $\mathbb{N}$.
   (v) Show that if $\alpha \in \mathbb{Z}[\mathrm{i}]$ is irreducible, then α divides precisely one prime in $\mathbb{N}$.