FACTORIZATION IN COMMUTATIVE RINGS

HUA-CHIEH LI

In this note, our ring is always a commutative ring. In other words, suppose that R is a ring. Then there exist two binary operations $+$ and $\cdot$ such that:

(1) $(R, +)$ is an abelian group;
(2) $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ for all a, b, $c \in R$;
(3) $a \cdot (b + c) = a \cdot b + a \cdot c$ for all a, b, $c \in R$;
(4) $a \cdot b = b \cdot a$ for all a, $b \in R$.

Moreover, we say R is an *integral domain* if R satisfies the following extra conditions:

- there exists $1 \in R$ such that $1 \cdot a = a \cdot 1 = a$ for all $a \in R$;
- if $a \neq 0$ and $b \neq 0$ in R, then $a \cdot b \neq 0$.

1. Euclidean Domain

Let $\mathbb{N}$ be the set of nonnegative integers and R a ring. We say that R is a *Euclidean Ring* if there is a function $\phi : R \setminus \{0\} \to \mathbb{N}$ such that: if a, $b \in R$ and $b \neq 0$, then there exist q, $r \in R$ such that $a = bq + r$ with either $r = 0$ or $\phi(r) < \phi(b)$.

A Euclidean ring which is an integral domain is called a *Euclidean domain.*

Example 1.1. The Ring $\mathbb{Z}$ of integers with $\phi(n) = |n|$ is a Euclidean domain.

Proof. For $x \in \mathbb{Q}$, denote $[x]$ the greatest integer less than or equal to x. Given a, $b \in \mathbb{Z}$, we claim that there exist q, $r \in \mathbb{Z}$ such that $a = bq + r$ with $r = 0$ or $|r| < |b|$.

We first consider the case that $b > 0$. Let $q = [a/b]$ and $r = a - b[a/b]$. Then $a = bq + r$. It remains to show that $0 \leq r < b$. We have that

$$\frac{a}{b} - 1 < \left[\frac{a}{b}\right] \leq \frac{a}{b}\ .$$

Multiplying all terms of this inequality by $-b$, we obtain

$$b - a > -b\left[\frac{a}{b}\right] \geq -a$$

and hence

$$0 \leq a - b\left[\frac{a}{b}\right] < b\,,$$

which is precisely $0 \leq r < b$ as desired.

For the case $b < 0$, use the similar argument above for a and $-b$. We find that there exist q and $r \in \mathbb{Z}$ such that $a = (-b)q + r$ with $r = 0$ or $r < |b| = -b$; so $-q$ and r have the desired properties. □

Example 1.2. If F is a field, then the ring of polynomials in one variable $F[x]$ is a Euclidean domain with $\phi(f) = \deg(f)$.

Proof. Given $f,\ g \in F[x]$ with $g \neq 0$, if $\deg(f) < \deg(g)$, then let $q = 0$ and $r = f$. If $\deg(f) \geq \deg(g)$, then we proceed by induction on $\deg(f)$.

If $\deg(f) = 0$, then $\deg(g) = 0$. Thus f and g are in F. Let $q = f \cdot g^{-1}$ and $r = 0$. We have $f = gq + r$ with $r = 0$ as desired.

Assume now that the property for Euclidean domain is true for polynomials of degree less than $n = \deg(f)$. Suppose

$$f = \sum_{i=0}^{n} a_i x^i, \quad g = \sum_{i=0}^{m} b_i x^i, \quad \text{with } a_n \neq 0, b_m \neq 0.$$

Let $f_1 = f - (a_n b_m^{-1} x^{m-n})g$. It is clear that $\deg(f_1) \leq n - 1$. By the induction hypothesis there are polynomials q_1 and r_1 such that $f_1 = gq_1 + r_1$ with $r_1 = 0$ or $\deg(r_1) < \deg(g)$. Therefore, let $q = a_n b_m^{-1} x^{n-m} + q'$ and $r = r_1$. Then

$$f = f_1 + (a_n b_m^{-1} x^{m-n})g = g(q_1 + a_n b_m^{-1} x^{m-n}) + r_1 = gq + r$$

with $r = 0$ or $\deg(r) < \deg(g)$ as desired. □

Recall that the set of complex numbers $\mathbb{C}$ consists of elements of the form $x + yi$, with $x,\ y \in \mathbb{R}$ where i satisfies $i^2 = -1$. For $\alpha = x + yi \in \mathbb{C}$, we define the norm of α by $N(\alpha) = x^2 + y^2$. Given $\alpha = x + yi$ and $\beta = u + vi$, we have that $\alpha\beta = (xu - yv) + (xv + yu)i$ and

$$N(\alpha\beta) = (xu - yv)^2 + (xv + yu)^2 = (x^2 + y^2)(u^2 + v^2) = N(\alpha)N(\beta).$$

Example 1.3. Let $\mathbb{Z}[i] = \{a + bi \mid a, b \in \mathbb{Z}\}$ be a subset of complex numbers. $\mathbb{Z}[i]$ is an integral domain called the domain of *Gaussian integers*. Moreover, $\mathbb{Z}[i]$ is a Euclidean domain with $\phi(a + bi) = N(a + bi) = a^2 + b^2$.

Proof. $\mathbb{Z}[i]$ is clearly closed under addition and substraction. Moreover, if $a+bi,\ c+di \in \mathbb{Z}[i]$, then

$$(a + bi)(c + di) = (ac - bd) + (ad + bc)i \in \mathbb{Z}[i].$$

Thus $\mathbb{Z}[i]$ is closed under multiplication and is a ring. Since $\mathbb{Z}[i]$ is contained in the complex numbers it is an integral domain.

It is clear that the norm defines a map from $\mathbb{Z}[i]$ to $\mathbb{N}$. Let $\alpha = a + bi,\ \beta = c + di \in \mathbb{Z}[i]$ and suppose that $\beta \neq 0$. Consider

$$\frac{\alpha}{\beta} = \frac{a + bi}{c + di} = \frac{ac + bd}{c^2 + d^2} + \frac{bc - ad}{c^2 + d^2} i = s + ti.$$

Choose integers $m,\ n \in \mathbb{Z}$ such that $|s - m| \leq 1/2$ and $|t - n| \leq 1/2$. Set $\delta = m + ni$ and $\gamma = \alpha - \beta\delta$. Then $\delta,\ \gamma \in \mathbb{Z}[i]$ and either $\gamma = 0$ or

$$\phi(\gamma) = \phi(\beta(\frac{\alpha}{\beta} - \delta)) = \phi(\beta)\phi(\frac{\alpha}{\beta} - \delta)) = \phi(\beta)((s - m)^2 + (t - n)^2) \leq \frac{1}{2}\phi(\beta) < \phi(\beta).$$

□

Exercise 1. *Let $\omega = (-1 + \sqrt{-3})/2$ and $\mathbb{Z}[\omega] = \{a + b\omega \mid a, b \in \mathbb{Z}\}$. Show that $\mathbb{Z}[\omega]$ is a Euclidean domain.*

Example 1.4. Let $\theta = (1 + \sqrt{-19})/2$ and $\mathbb{Z}[\theta] = \{a + b\theta \mid a, b \in \mathbb{Z}\}$. $\mathbb{Z}[\theta]$ is an integral domain but is not a Euclidean domain.

Proof. $\mathbb{Z}[\theta]$ is clearly closed under addition and substraction. Moreover, $\theta^2 = \theta - 5$. Hence, if $a + b\theta$, $c + d\theta \in \mathbb{Z}[\theta]$, then

$$(a + b\theta)(c + d\theta) = ac + (ad + bc)\theta + bd\theta^2 = (ac - 5bd) + (ad + bc + bd)\theta \in \mathbb{Z}[\theta].$$

Thus $\mathbb{Z}[\theta]$ is closed under multiplication and is a ring. Since $\mathbb{Z}[\theta]$ is contained in the complex numbers it is an integral domain.

Suppose that $\mathbb{Z}[\theta]$ is a Euclidean domain with $\phi : \mathbb{Z}[\theta] \setminus \{0\} \to \mathbb{N}$ satisfies the Euclidean domain property. Let $\alpha \in \mathbb{Z}[\theta]$ be an element such that

$$\phi(\alpha) = \min\{\phi(\lambda) \mid \lambda \neq 0, 1, -1, \lambda \in \mathbb{Z}[\theta]\}.$$

By the Euclidean domain property, there exist δ, $\gamma \in \mathbb{Z}[\theta]$ such that $2 = \alpha\delta + \gamma$ with $\gamma = 0$ or $\phi(\gamma) < \phi(\alpha)$. However, by the definition of α, this implies that $\gamma = 0$, 1 or -1. In other words, $\alpha\delta = 1$, 2 or 3.

Recall that if $\beta = a + b\theta \in \mathbb{Z}[\theta]$, then $N(\beta) = a^2 + ab + 5b^2 \in \mathbb{N}$. Moreover, suppose $\beta \neq 0$, 1 or -1. If $a = 0$ then $N(\beta) = 5b^2 \geq 5$ and if $b = 0$ then $N(\beta) = a^2 \geq 4$. If $ab > 0$, then

$$N(\beta) = a^2 + ab + 5b^2 = (a - b)^2 + 4b^2 + 3ab \geq 4b^2 + 3ab \geq 7$$

and if $ab < 0$, then

$$N(\beta) = a^2 + ab + 5b^2 = (a + b)^2 + 4b^2 - ab \geq 4b^2 - ab \geq 5.$$

In conclusion, if $\beta \in \mathbb{Z}[\theta] \setminus \{0, 1, -1\}$ then $N(\beta) \in \mathbb{N}$ and $N(\beta) \geq 4$.

Since $N(\alpha\delta) = 1$, 4 or 9, and $N(\alpha\delta) = N(\alpha)N(\delta)$, we have that $N(\alpha) \mid 1$, $N(\alpha) \mid 4$ or $N(\alpha) \mid 9$. The discussion above shows that $N(\alpha) \neq 1, 2, 3$. Hence we have that $N(\alpha) = 4$ or $N(\alpha) = 9$.

The Euclidean domain property shows that there exist δ' and $\gamma' \in \mathbb{Z}[\theta]$ such that $\theta = \alpha\delta' + \gamma'$ with either $\gamma' = 0$ or $\phi(\gamma') < \phi(\alpha)$. Again, the definition of α implies that $\alpha\delta' = \theta$, $\theta - 1$ or $\theta + 1$. Taking norms, we have $N(\alpha) | N(\theta)$, $N(\alpha) | N(\theta - 1)$ or $N(\alpha) | N(\theta + 1)$. However, $N(\theta) = 5$, $N(\theta - 1) = 5$ and $N(\theta + 1) = 7$. Neither one of them can be divided by 4 or 9. We get a contradiction. Hence $\mathbb{Z}[\theta]$ is not a Euclidean domain. □

Definition 1.5. A nonzero element a of a ring R is said to *divide* an element $b \in R$ (notation: $a \mid b$) if there exists $x \in R$ such that $b = ax$. Elements a, b of R are said to be *associates* (notation: $a \approx b$) if $a \mid b$ and $b \mid a$.

Let S be a nonempty subset of R. An element $d \in R$ is a *greatest common divisor* of S provided:

(1) $d \mid a$ for all $a \in S$;
(2) if $c \mid a$ for all $a \in S$, then $c \mid d$.

In general, greatest common divisors do not always exist. For example, in the ring $2\mathbb{Z}$ of even integers, 2 has no divisor at all, whence $2, 4$ has no greatest common divisor. Even when a greatest common divisor exists, it need not be unique. However, any two greatest common divisors of S are clearly associates by property (2). Furthermore any associate of a greatest common divisor of S is easily seen to be a greatest common divisor of S.

In the following we provide some basic properties of greatest common divisor.

Lemma 1.6. *Let R be a ring and $a, b, c \in R$. Suppose that d is a greatest common divisor of a, b.*

(1) *Suppose that $c = aq + b$ for some $q \in R$. Then d is a greatest common divisor of a, c.*

(2) *Suppose that d' is a greatest common divisor of d, c. Then d' is a greatest common divisor of a, b, c.*

Proof. (proof of (1)) We first show that d divides a and c. We know d divides a by definition. Since $d \mid a$ and $d \mid b$, we have $a = dx$ and $b = dy$ for some $x, y \in R$. Hence $c = dxq + dy = d(xq + y)$. This shows that $d \mid c$.

Suppose $e \in R$ such that $e \mid a$ and $e \mid c$. Then there exist $u, v \in R$ such that $a = eu$ and $c = ev$. Hence $b = c - aq = e(v - uq)$. This shows that $e \mid b$. Since e divides a and b, by the definition of greatest common divisors, we have $e \mid d$. □

Exercise 2. *Prove (2) of Lemma 1.6.*

Example 1.7 (The Euclidean Algorithm)**.** Let $a, b \in \mathbb{Z}$. By Example 1.1, there exist q_1, $r_1 \in \mathbb{Z}$ such that

$$a = bq_1 + r_1, \quad 0 \le r_1 < |b|\,.$$

If $r_1 > 0$, there exist q_2, $r_2 \in \mathbb{Z}$ such that

$$b = r_1 q_2 + r_2, \quad 0 \le r_2 < r_1.$$

If $r_2 > 0$, there exist q_3, $r_3 \in \mathbb{Z}$ such that

$$r_1 = r_2 q_3 + r_3, \quad 0 \le r_3 < r_2.$$

Continue this process. Then $r_n = 0$ for some $n \in \mathbb{N}$. If $n > 1$ then r_{n-1} is a greatest common divisor of a, b. If $n = 1$, then b is a greatest common divisor of a, b.

Proof. Note that $r_1 > r_2 > \dots$. If $r_n \neq 0$ for all $n \in \mathbb{N}$, then $r_1, r_2, r_3, \dots$ is an infinite, strictly decreasing sequence of positive integers, which is impossible. So $r_n = 0$ for some n.

If $r_1 = 0$, then $a = bq_1$. So $b \mid a$ and of course $b \mid b$. If c divides both a and b, then of course $c \mid b$. Hence b is a greatest common divisor of a, b.

Now suppose $r_n = 0$ for $n > 1$. Then $r_{n-2} = r_{n-1} q_n$ (we set $r_0 = b$). By the argument above, we have that r_{n-1} is a greatest common divisor of r_{n-2}, r_{n-1}. However, $r_{n-3} = r_{n-2} q_{n-1} + r_{n-1}$ (we set $r_{-1} = a$). By Lemma 1.6 (1), we have r_{n-1} is a greatest common divisor of r_{n-2}, r_{n-3}. Continue this argument inductively. We have that r_{n-1} is a greatest common divisor of a, b. □

Exercise 3. *Suppose R is a Euclidean domain and $a_1, \dots, a_n \in R$. Show that there exists a greatest common divisor of $a_1, \dots, a_n$.*

2. Principle Ideal Domain

Given a ring R, a subring I of R is an *ideal* provided $rx \in I$ for $r \in R$, $x \in I$. A *principal ideal ring* is a ring in which every ideal is principle. In other words, for every ideal I of R, there exists $x \in I$ such that if $\lambda \in I$, $\lambda = rx$ for some $r \in R$. A principle ideal ring which is an integral domain is called a *principle ideal domain*

Example 2.1. $\mathbb{Z}$ is a principle ideal domain.

Proof. Given a nonzero ideal I of $\mathbb{Z}$. Consider $n \in \mathbb{Z}$ such that

$$|n| = \min\{|x| \,:\, x \in I \setminus \{0\}\}.$$

Given $a \in I$, by Example 1.1, there exist $h, r \in \mathbb{Z}$ such that $a = nh + r$ with either $r = 0$ or $|r| < |n|$. Since $r = a - nh \in I$, by the definition of n, we conclude that $r = 0$ and hence $a = nh$. In other words, $I = (n)$. □

Using similar argument we can show the following:

Theorem 2.2. *Every Euclidean ring is a principle ideal ring.*

Exercise 4. *Prove Theorem 2.2.*

From Theorem 2.2, the polynomial ring $F[x]$ in Example 1.2 and the Gaussian integers $\mathbb{Z}[i]$ in Example 1.3 are principle ideal domains.

In general, to prove a ring is a principle ideal ring is not easy. We can imitate the proof of Theorem 2.2 to show certain rings are principle ideal rings.

Theorem 2.3. *Let R be a ring. Suppose that there is a function $\phi : R \setminus \{0\} \to \mathbb{N}$ such that given $\alpha, \beta \in R$, $\beta \neq 0$, if β does not divide α then there exist $\gamma, \delta \in R$ such that $\alpha\gamma - \beta\delta \neq 0$ and*

$$\phi(\alpha\gamma - \beta\delta) < \phi(\beta).$$

Then R is a principle ideal ring.

Proof. Let I be a nonzero ideal of R. Let $\beta \in I$ be an element with the property that

$$\phi(\beta) = \min\{\phi(x) : x \in I \setminus \{0\}\}.$$

We claim that $I = (\beta)$. Given $\alpha \in I$, suppose that β does not divide α. By the hypothesis, there exist $\delta, \gamma \in R$ such that $\alpha\gamma - \beta\delta \neq 0$ and $\phi(\alpha\gamma - \beta\delta) < \phi(\beta)$. Since $\alpha\gamma - \beta\delta \in I$ and $\alpha\gamma - \beta\delta \neq 0$, this contradicts the assumption of β. Therefore β divides every element of I. □

Example 2.4. Let $\theta = (1 + \sqrt{-19})/2$ and $\mathbb{Z}[\theta] = \{a + b\theta \,|\, a, b \in \mathbb{Z}\}$. $\mathbb{Z}[\theta]$ is a principle ideal domain.

Proof. Let $\phi(\alpha) = N(\alpha)$ for all $\alpha \in \mathbb{Z}[\theta] \setminus \{0\}$. We will show that $\mathbb{Z}[\theta]$ satisfies the condition in Theorem 2.3.

Given $\alpha, \beta \in \mathbb{Z}[\theta]$ with $\beta \neq 0$, if β does not divide α then a case by case consideration will lead to elements $\gamma, \delta \in \mathbb{Z}[\theta]$ such that

$$0 < N\left(\frac{\alpha}{\beta}\gamma - \delta\right) < 1,$$

whence $\alpha\gamma - \beta\delta \neq 0$ and $N(\alpha\gamma - \beta\delta) < N(\beta)$.

Write

$$\frac{\alpha}{\beta} = s + t\theta, \quad \text{with } s, t \in \mathbb{Q}.$$

(1) $t \in \mathbb{Z}$: In this case, $s \notin \mathbb{Z}$. Let $n \in \mathbb{Z}$ such that $|s - n| \leq 1/2$ and take $\gamma = 1$, $\delta = n + t\theta$. Now,

$$0 < N\left(\frac{\alpha}{\beta}\gamma - \delta\right) = N(s - n) \leq \frac{1}{4} < 1.$$

(2) $s \in \mathbb{Z}$:

(a) $5t \in \mathbb{Z}$: Let $m \in \mathbb{Z}$ such that $|t - m| \leq 1/2$. In fact, because $5t \in \mathbb{Z}$, we have $|t - m| \leq 2/5$. Take $\gamma = 1$ and $\delta = s + m\theta$. Now

$$0 < N\left(\frac{\alpha}{\beta}\gamma - \delta\right) = N((t - m)\theta) \leq \frac{4}{25} \times 5 < 1.$$

(b) $5t \notin \mathbb{Z}$: Consider

$$(s + t\theta)(1 - \theta) = s - s\theta + t\theta - t\theta^2 = s - s\theta + t\theta - t\theta + 5t = s + 5t - s\theta.$$

Let $n \in \mathbb{Z}$ such that $|s + 5t - n| \leq 1/2$ and take $\gamma = 1 - \theta$, $\delta = n - s\theta$. Now,

$$0 < N\left(\frac{\alpha}{\beta}\gamma - \delta\right) = N(s + 5t - n) \leq \frac{1}{4} < 1.$$

(3) $s,\ t \notin \mathbb{Z}$:

(a) $2s,\ 2t \in \mathbb{Z}$: Consider

$$(s + t\theta)\theta = s\theta + t\theta - 5t = -5t + (s + t)\theta.$$

Since $s + t \in \mathbb{Z}$, letting $n \in \mathbb{Z}$ such that $|-5t - n| \leq 1/2$, we can take $\gamma = \theta$ and $\delta = n + (s + t)\theta$. Now,

$$0 < N\left(\frac{\alpha}{\beta}\gamma - \delta\right) = N(-5t - n) \leq \frac{1}{4} < 1.$$

(b) $2s \notin \mathbb{Z}$ and $2t \in \mathbb{Z}$: Let $n \in \mathbb{Z}$ such that $|2s - n| \leq 1/2$. Take $\gamma = 2$ and $\delta = n + 2t\theta$. Now,

$$0 < N\left(\frac{\alpha}{\beta}\gamma - \delta\right) = N(2s - n) \leq \frac{1}{4} < 1.$$

(c) $2s \in \mathbb{Z}$ and $2t \notin \mathbb{Z}$: When $10t \in \mathbb{Z}$, let $m \in \mathbb{Z}$ such that $|2t - m| \leq 1/2$. In fact, because $5 \times 2t \in \mathbb{Z}$, we have $|2t - m| \leq 2/5$. Take $\gamma = 2$ and $\delta = 2s + m\theta$. Now

$$0 < N\left(\frac{\alpha}{\beta}\gamma - \delta\right) = N((2t - m)\theta) \leq \frac{4}{25} \times 5 < 1.$$

If $10t \notin \mathbb{Z}$, then consider

$$(s + t\theta)(2 - 2\theta) = 2s - 2s\theta + 2t\theta - 2t\theta^2 = 2s + 10t - 2s\theta.$$

Let $n \in \mathbb{Z}$ such that $|2s + 10t - n| \leq 1/2$ (note that $10t \notin \mathbb{Z}$) and take $\gamma = 2 - 2\theta$, $\delta = n - 2s\theta$. Now,

$$0 < N\left(\frac{\alpha}{\beta}\gamma - \delta\right) = N(2s + 10t - n) \leq \frac{1}{4} < 1.$$

(d) $2s \notin \mathbb{Z}$ and $2t \notin \mathbb{Z}$: Let $m \in \mathbb{Z}$ such that $|t-m| \leq 1/2$. If $|t-m| \leq 1/3$, letting $n \in \mathbb{Z}$ such that $|s-n| \leq 1/2$, then we can take $\gamma = 1$ and $\delta = n + m\theta$. Now,

$$0 < N\left(\frac{\alpha}{\beta}\gamma - \delta\right) = N((s-n) + (t-m)\theta) \leq \frac{1}{4} + \frac{1}{6} + \frac{1}{9} \times 5 = \frac{35}{36} < 1.$$

If $1/3 < |t-m| < 1/2$, then $2/3 < |2t-2m| < 1$. Let $m' \in \mathbb{Z}$ such that $|2t-m'| \leq 1/2$. Then we have $|2t-m'| < 1/3$. Let $n' \in \mathbb{Z}$ such that $|2s-n'| \leq 1/2$. Take $\gamma = 2$ and $\delta = n' + m'\theta$. Now,

$$0 < N\left(\frac{\alpha}{\beta}\gamma - \delta\right) = N((2s-n') + (2t-m')\theta) < \frac{1}{4} + \frac{1}{6} + \frac{1}{9} \times 5 = \frac{35}{36} < 1.$$

□

Remark 2.5. The converse of Theorem 2.2 is false since $\mathbb{Z}[\theta]$ is a principle ideal domain that is not a Euclidean domain (Example 1.4).

Example 2.6. Let $\mathbb{Z}[x]$ be the ring of polynomials over $\mathbb{Z}$. Then $\mathbb{Z}[x]$ is an integral domain but is not a principle ideal domain.

Proof. Considering the leading coefficients of $f(x)$ and $g(x)$, we can easily conclude that if $f(x) \neq 0$ and $g(x) \neq 0$ in $\mathbb{Z}[x]$, then $f(x)g(x) \neq 0$.

To show that $\mathbb{Z}[x]$ is not a principle ideal domain, we consider the ideal I generated by 2 and x (i.e. $I = (2, x)$). We first claim that $I \neq \mathbb{Z}[x]$. Otherwise there exist $u(x), v(x) \in \mathbb{Z}[x]$ such that $1 = 2u(x) + xv(x)$. Substitute $x = 0$ into the identity. We have that $1 = 2u(0)$ which is absurd because $u(0) \in \mathbb{Z}$.

Now, suppose that there exists $f(x) \in \mathbb{Z}[x]$ such that $(f(x)) = I$. In other words, there exist $g(x) \in \mathbb{Z}[x]$ and $h(x) \in \mathbb{Z}[x]$ such that $2 = g(x)f(x)$ and $x = h(x)f(x)$. From $2 = g(x)f(x)$, we conclude that $f(x) \in \mathbb{Z}$. Because $I \neq \mathbb{Z}[x]$, $f(x)$ can not be a unit, whence $f(x) = \pm 2$. On the other hand, by $x = h(x)f(x)$, we have $h(x) = ax + b$ for some $a, b \in \mathbb{Z}$. Since $\pm 2a \neq 1$ for all $a \in \mathbb{Z}$, we get a contradiction. □

Exercise 5. *Suppose that R is an integral domain. Suppose further that there exists $a \in R$ such that $a \neq 0$ and a is not a unit in R. Prove that $R[x]$ the polynomial ring over R is an integral domain but is not a Euclidean domain.*

Finally we provide some basic properties of principle ideal rings.

Proposition 2.7. *Every principle ideal ring is a ring with identity.*

Proof. Since R itself is an ideal of R, $R = (a)$ for some $a \in R$. Consequently, $a \in R$, so $a = ea = ae$ for some $e \in R$. If $b \in R$, then $b = xa$ for some $x \in R$. Therefore, $be = (xa)e = x(ae) = xa = b$, whence e is the identity of R. □

Exercise 6. *Prove that every Euclidean ring is a ring with identity without using the fact that every Euclidean ring is a principle ideal ring.*

Proposition 2.8. *If R is a principle ideal ring, given $a_1, \ldots, a_n \in R$, then a greatest common divisor of $\{a_1, \ldots, a_n\}$ exists.*

Proof. Consider $I = (a_1, \ldots, a_n)$ the ideal generated by $a_1, \ldots, a_n$. Since R is a principle ideal ring, there exists $d \in R$ such that $I = (d)$. We claim that d is a greatest common divisor of $\{a_1, \ldots, a_n\}$.

First, since $a_i \in I = (d)$, there exist $r_i \in R$ such that $a_i = r_i d$ for $i = 1, \ldots, n$. Hence $d \mid a_i$ for $i = 1, \ldots, n$.

Second, since $(a_1, \ldots, a_n) = (d)$, there exist $\lambda_i \in R$ such that $d = \sum_{i=1}^{n} \lambda_i a_i$. Suppose that $c \mid a_i$ for $i = 1, \ldots, n$. There exist $\gamma_i \in R$ such that $a_i = \gamma_i c$ for $i = 1, \ldots, n$. This implies that $d = \sum_{i=1}^{n} (\lambda_i \gamma_i) c$, whence $c \mid d$. □

Recall that a ring is *Noetherian* if it satisfies the ascending chain condition on ideals. It can be proved that R is Noetherian if and only if every ideal of R is finitely generated. We do not need this fact here. However, we can show that a principle ideal ring is Noetherian.

Lemma 2.9. *If R is a principle ideal ring and*

$$I_1 \subseteq I_2 \subseteq \cdots \subseteq I_n \subseteq \cdots$$

is a chain of ideals in R, then for some $n \in \mathbb{N}$, $I_j = I_n$ for all $j \geq n$.

Proof. Let $I = \cup_{i \in \mathbb{N}} I_i$. We claim that I is an ideal of R. If $b, c \in I$, then we have $b \in I_i$ and $c \in I_j$ for some $i, j \in \mathbb{N}$. Without loss of generality, we can assume that $i \geq j$. Consequently $I_j \subseteq I_i$, and hence $b, c \in I_i$. Therefore, $b - c \in I_i \subseteq I$. Similarly, if $r \in R$ and $b \in I$, then $b \in I_i$ for some $i \in \mathbb{N}$, whence $rb \in I_i \subseteq I$. Therefore, I is an ideal of R. By hypothesis I is principle, say $I = (a)$. Since $a \in I$, we have $a \in I_n$ for some $n \in \mathbb{N}$. Hence $(a) \subseteq I_n$. Therefore, for every $j \geq n$,

$$(a) \subseteq I_n \subseteq I_j \subseteq (a),$$

whence $I_j = I_n$. □

Exercise 7. *Suppose that R is a principle ideal ring. Let $a_1, \ldots, a_n, \ldots$ be (infinitely many) elements in R. Prove that there exists a greatest common divisor of $\{a_1, \ldots, a_n, \ldots\}$.*

3. Unique Factorization Domain

3.1. **General Properties.** The *Fundamental Theorem of Arithmetic* says that any positive integer $n > 1$ can be written uniquely in the form $n = p_1^{t_1} \cdots p_r^{t_r}$, where $p_1 < \cdots < p_r$ are primes and $t_i > 0$ for all i. In this section we study those integral domains in which an analogue of the fundamental theorem of arithmetic holds.

In $\mathbb{Z}$, a prime number p has the following properties:

(1) If $p = ab$ then a or b is a unit.
(2) If $p \mid ab$ then $p \mid a$ or $p \mid b$.

For arbitrary ring, these are two different properties.

Definition 3.1. Let R be a ring with identity. An element $\pi \in R$ is *irreducible* provided that π is not a unit and if $\pi = ab$ for some $a, b \in R$ then a or b is a unit.

An element $p \in R$ is *prime* provided that p is not a unit and if $p \mid ab$ then $p \mid a$ or $p \mid b$.

Example 3.2. In the ring $\mathbb{Z}/6\mathbb{Z} = \{\bar{0}, \bar{1}, \bar{2}, \bar{3}, \bar{4}, \bar{5}\}$, $\bar{2}$ is prime but it is not irreducible.

Proof. $\bar{2}$ does not divide $\bar{1} \cdot \bar{1} = \bar{5} \cdot \bar{5} = \bar{1}$, $\bar{1} \cdot \bar{3} = \bar{3} \cdot \bar{3} = \bar{3} \cdot \bar{5} = \bar{3}$, and $\bar{1} \cdot \bar{5} = \bar{5}$. Hence $\bar{2}$ is prime. On the other hand, $\bar{2}$ is not irreducible because $\bar{2} = \bar{2} \cdot \bar{4}$ and neither $\bar{2}$ nor $\bar{4}$ are units in $\mathbb{Z}/6\mathbb{Z}$. □

Example 3.3. In the ring $\mathbb{Z}[\sqrt{10}] = \{a + b\sqrt{10} : a, b \in \mathbb{Z}\}$, 2 is irreducible but it is not prime.

Proof. Recall that the map $\mathcal{N} : \mathbb{Z}[\sqrt{10}] \to \mathbb{Z}$ given by $\mathcal{N}(a + b\sqrt{10}) = a^2 - 10b^2$ has the properties that $\mathcal{N}(\alpha\beta) = \mathcal{N}(\alpha)\mathcal{N}(\beta)$ for all α, $\beta \in \mathbb{Z}[\sqrt{10}]$ and $\mathcal{N}(\alpha) = \pm 1$ if and only if α is a unit.

Suppose that there exist α and β in $\mathbb{Z}[\sqrt{10}]$ which are not units such that $2 = \alpha\beta$. Then we have $4 = \mathcal{N}(2) = \mathcal{N}(\alpha)\mathcal{N}(\beta)$. Since $\alpha = a + b\sqrt{10}$ is not a unit, we have $\mathcal{N}(\alpha) = a^2 - 10b^2 = \pm 2$. This shows that $a^2 \equiv \pm 2 \pmod 5$. However, neither 2 nor -2 is a quadratic residue modulo 5. We get a contradiction. Hence 2 is irreducible.

On the other hand, since $2 \cdot 3 = 6 = (4 + \sqrt{10})(4 - \sqrt{10})$, we have that $2 \mid (4 + \sqrt{10})(4 - \sqrt{10})$. Suppose that $2 \mid (4 + \sqrt{10})$ or $2 \mid (4 - \sqrt{10})$. By taking $\mathcal{N}$, we have that $4 \mid 6$ in $\mathbb{Z}$, which is absurd. Hence 2 is not prime in $\mathbb{Z}[\sqrt{10}]$. □

From examples above, we know that in general prime elements and irreducible elements are distinct. However in some cases, they are related.

Lemma 3.4. *Let R be an integral domain. Then every prime element of R is irreducible.*

Proof. Suppose that p is prime. If $p = ab$, then either $p \mid a$ or $p \mid b$; say $p \mid a$. Thus there exist $x \in R$ such that $a = px$. Therefore, $p = ab = pxb$, and hence $p(1 - xb) = 0$. Since R is an integral domain, this implies that $1 = xb$. Therefore, b is a unit. Hence p is irreducible. □

We include an important property for irreducible elements of an integral domain which is familiar for the integer ring $\mathbb{Z}$.

Lemma 3.5. *Let R be an integral domain. The only divisors of an irreducible element of R are its associates and the units of R.*

Proof. If π is irreducible and $d \mid \pi$, then because $\pi = dx$ for some $x \in R$, this implies that either d or x is a unit. The second case implies that d and π are associates. □

Exercise 8. *Let R be an integral domain. Suppose that $a, b \in R$ are associates.*

(1) *Prove that there exists an unit $u \in R$ such that $a = ub$.*
(2) *Prove that a is irreducible if and only if b is irreducible.*
(3) *Prove that a is prime if and only if b is prime.*

Definition 3.6. An integral domain R is a *unique factorization domain* provided that:

(1) Every nonzero element $a \in R$ which is not a unit can be written as $a = \alpha_1 \cdots \alpha_n$ with α_i irreducible.
(2) If $a = \alpha_1 \cdots \alpha_n = \beta_1 \cdots \beta_m$ with α_i, β_j irreducible, then $n = m$ and for some permutation σ of $\{1, 2, \ldots, n\}$, α_i and $\beta_{\sigma(i)}$ are associates for every i.

Remark 3.7. From the definition, every irreducible element in a unique factorization domain is necessary prime. Consequently, prime elements and irreducible elements coincide in a unique factorization domain by Lemma 3.4.

Example 3.8. The polynomial ring $F[x]$ over a field F is a unique factorization domain.

Proof. Because every nonzero constant is a unit, we show first that every nonconstant polynomial can be written as a product of finitely many irreducible polynomial. It is to see that polynomials of degree 1 are irreducible. assume that we have proved the result for all polynomials of degree less than n and that $\deg(f) = n$. If f is irreducible, we are done. Otherwise $f = gh$ where $1 \leq \deg(g), \deg(h) < n$. By the induction assumption both g and h can be written as products of finitely many irreducible polynomials. Thus so is f.

Next, we show that every irreducible polynomial is prime. Suppose that π is an irreducible polynomial and $\pi \,|\, fg$. Consider the ideal (f, π). Since $F[x]$ is a principle ideal domain (c.f. Theorem 2.2), we have $(f, \pi) = (d)$ for some $d \in F[x]$. $\pi \in (d)$ implies that $d \,|\, \pi$, and hence by Lemma 3.5, $(f, \pi) = (1)$ or (π). If $(f, \pi) = (\pi)$, then $\pi \,|\, f$. If $(f, \pi) = 1$, then there exist $l, h \in F[x]$ such that $l\pi + hf = 1$. Thus $l\pi g + hfg = g$. Since π divides the left-hand side of this equation, π must divide g.

Finally if $f = \pi_1 \cdots \pi_n = p_1 \cdots p_m$ with π_i, p_j irreducible, then since π_1 is prime, π_1 divides some p_j; say p_1. On the other hand, since p_1 is irreducible and π_1 is not a unit, by Lemma 3.5 π_1 and p_1 are associates; say $u\pi_1 = p_1$ for some unit u of R. Hence $\pi_2 \cdots \pi_n = (up_2) \cdots p_m$. By Exercise 8, up_2 is also irreducible, the proof of uniqueness is now completed by a routine inductive argument. □

Exercise 9. *Let R be an integral domain.*

(1) *Prove that p is a prime element in R if and only if (p) is a prime ideal of R.*
(2) *Suppose that R is a principle ideal domain. Prove that π is irreducible in R if and only if (π) is a maximal ideal of R.*
(3) *Suppose that R is a principle ideal domain. Prove that an element in R is prime if and only if it is irreducible.*
(4) *Show that $\mathbb{Z}[\sqrt{10}]$ is not a principle ideal domain.*

In general, to show a ring is a unique factorization domain we only have to show the following:

(1) using the irreducibility to show that in the specific ring every nonzero element which is not a unit can be written as a product of finitely many irreducible elements;
(2) show that in the specific ring every irreducible element is prime. Then the proof of uniqueness can be completed by a routine inductive argument as in the proof of Example 3.8.

Theorem 3.9. *Every principle ideal domain is a unique factorization domain.*

Proof. Suppose that R is a principle ideal domain. We claim first that if $a \in R$, $a \neq 0$ and a is not a unit, then a can be written as a product of finitely many irreducible elements. If a can not be written as a product of finitely many irreducible elements, then a is not irreducible and hence $a = a_1 b_1$ for some a_1, $b_1 \in R$ which are not units. By assumption, one of the a_1 or b_1 can not be written as a product of finitely many irreducible elements; say a_1. Then $a_1 = a_2 b_2$ for some a_2, $b_2 \in R$ which are not units and a_2 can not be written as a product of finitely many irreducible elements. Continuing in this way, we construct infinitely many a_i with $a_i = a_{i+1} b_{i+1}$ where all the a_i and $b_i \in R$ are not units. Since $a = a_1 b_1$ and b_1 is not a unit, we have that $(a) \subsetneq (a_1)$. Similarly, we have $(a_i) \subsetneq (a_{i+1})$. In other words we have a nonstop ascending chain of ideals

$$(a) \subsetneq (a_1) \subsetneq \cdots \subsetneq (a_i) \subsetneq \cdots,$$

contradicting Lemma 2.9.

For the uniqueness, exercise 9 says that every irreducible element of R is prime. This completes the proof.

□

Exercise 10. *Suppose that R is a unique factorization domain. Let S be a set of primes in R such that every prime in R is associate to a prime in S and no two primes in S are associate.*

(1) *If $a \in R$, $a \neq 0$, show that we can uniquely write*

$$a = u \prod_{p \in S} p^{v_p(a)},$$

where u is a unit and $v_p(a)$ are nonnegative integers which are positive only for finitely many $p \in S$.

(2) *Prove that $v_p(ab) = v_p(a) + v_p(b)$ for all $p \in S$ and $a, b \in R$.*

(3) *Given $a_1, \ldots, a_n \in R$, prove that there exists a greatest common divisor of $a_1, \ldots, a_n$.*

By Theorem 3.9, we know that $\mathbb{Z}[i]$ and $\mathbb{Z}[\frac{1+\sqrt{-19}}{2}]$ are unique factorization domains. The converse of Theorem 3.9 is not always true. For example, we know that $\mathbb{Z}[x]$ is not a principle ideal domain (c.f. Example 2.6), but we will show later that $\mathbb{Z}[x]$ is a unique factorization domain.

3.2. **Factorization in Polynomial Rings.** In the rest of this section, we devote entirely to show that if R is a unique factorization domain, then $R[x]$, the polynomial ring over R is also a unique factorization domain.

Let F be the quotient field of R. In other words, every element of F can be written as a/b for some $a, b \in R$ with $b \neq 0$. Our strategy is using the fact that $F[x]$ is a unique factorization domain to show that $R[x]$ is a unique factorization domain.

Let $f = \sum_{i=0}^{n} a_i x^i$ be a nonzero polynomial in $R[x]$. Since R is a unique factorization domain, by Exercise 10 (3), a greatest common divisor of the coefficients $a_0, a_1, \ldots, a_n$ exists. We call it a *content* of f and denotes it by $C(f)$. Strictly speaking, $C(f)$ is ambiguous since greatest common divisors are not unique. But any two contents of are necessarily associates. We shall write $b \approx c$ whenever b and c are associates in R. If $f \in R[x]$ and $C(f)$ is a unit in R, then f is said to be *primitive*.

Lemma 3.10. *Let R be a unique factorization domain. $a \in R$ and $f, g \in R[x]$.*

(1) *$C(af) \approx aC(f)$. In particular, $f = C(f)f_1$ with f_1 primitive in $R[x]$.*
(2) (Gauss) $C(fg) \approx C(f)C(g)$. In particular, the product of primitive polynomials in $R[x]$ is also primitive.

Proof. (1) Suppose that $f = \sum_{i=0}^{n} a_i x^i$ and $d = C(f)$ which is a greatest common divisor of $a_0, a_1, \ldots, a_n$. Then $af = \sum_{i=0}^{n} aa_i x^i$ and ad is a greatest common divisor of $aa_0, aa_1, \ldots, aa_n$. On the other hand, let $b_i = a_i/d \in R$. The greatest common divisor of $b_0, b_1, \ldots, b_n$ is a unit. Hence $f = d\sum_{i=0}^{n} b_i x^i = C(f)f_1$ with $f_1 = \sum_{i=0}^{n} b_i x^i$ primitive.

(2) $f = C(f)f_1$ and $g = C(g)g_1$ with f_1, g_1 primitive, by (1). Consequently $C(fg) \approx C(f)C(g)C(f_1g_1)$. Hence it suffices to prove that if f and g are primitive then fg is primitive (i.e. $C(fg)$ is a unit). If $f = \sum_{i=0}^{n} a_i x^i$ and $g = \sum_{j=0}^{m} b_j x^j$, then $fg = \sum_{k=0}^{n+m} c_k x^k$ with $c_k = \sum_{i+j=k} a_i b_j$. If $C(fg)$ is not a unit, then since R is a unique factorization domain, there exists a prime element $p \in R$ such that $p \mid C(fg)$. That is, $p \mid c_k$ for all k. Since $C(f)$ is a unit, $p \nmid C(f)$. Hence there is an integer s such that $p \mid a_i$ for $i < s$ and $p \nmid a_s$. Similarly there is an integer t such that $p \mid b_j$ for $j < t$ and $p \nmid b_t$. Consider

$$c_{s+t} = a_0 b_{s+t} + a_1 b_{s+t-1} + \cdots + a_{s-1} b_{t+1} + a_s b_t + a_{s+1} b_{t-1} + \cdots + a_{s+t} b_0.$$

p divides every term on the right-hand side of the equation except the term $a_s b_t$. Hence $p \nmid c_{s+t}$. This is a contradiction. Therefore fg is primitive. □

Now for study the irreducible elements in $R[x]$, we first notice that if $\alpha \in R$ is irreducible in R, then α is also irreducible in $R[x]$. Indeed, if $\alpha = f_1 f_2$ for $f_1, f_2 \in R[x]$, then comparing the degrees of both side we have $f_1, f_2 \in R$. Since α is irreducible in R, either f_1 or f_2 is a unit in R and hence a unit in $R[x]$.

Next, we compare elements in $R[x]$ and elements in $F[x]$. Suppose $f = \sum_{i=0}^{n} a_i x^i \in F[x]$. We can write $a_i = \alpha_i \beta_i^{-1}$ for some $\alpha_i, \beta_i \in R$ and $\beta_i \neq 0$. Let $\beta = \prod_{i=0}^{n} \beta_i$. We have $\beta a_i = \alpha_i \gamma_i$ for some $\gamma_i \in R$ and hence $\beta f = \sum_{i=0}^{n} \alpha_i \gamma_i x^i \in R[x]$. In other word, every $f \in F[x]$ can always be written as $f = ab^{-1} f_1$ with $a, b \in R$, $b \neq 0$ and f_1 primitive in $R[x]$.

Lemma 3.11. *Let f be a primitive polynomial in $R[x]$ and $g \in R[x]$. Then f divides g in $R[x]$ if and only if f divides g in $F[x]$.*

Proof. If $f \mid g$ in $R[x]$, then $g = fh$ for some $h \in R[x] \subseteq F[x]$. Hence $f \mid g$ in $F[x]$.

On the other hand, if $f \mid g$ in $F[x]$, then $g = fh$ for some $h \in F[x]$. Because $h = ab^{-1}h_1$ with $a, b \in R$, $b \neq 0$ and h_1 primitive in $R[x]$, we have that $bg = afh_1$. Taking contents on both side, by Lemma 3.10 we have

$$bC(g) \approx C(bg) \approx C(afh_1) \approx aC(f)C(h_1) \approx a,$$

because $C(f)$ and $C(h_1)$ are units in R. Hence $ab^{-1} \in R$. In other words, $h = ab^{-1}h_1 \in R[x]$ and hence $f \mid g$ in $R[x]$. □

Lemma 3.12. *Let f be a primitive polynomial in $R[x]$. Then f is irreducible in $R[x]$ if and only if f is irreducible in $F[x]$.*

Proof. Suppose f is irreducible in $F[x]$ and $f = gh$ with $g, h \in R[x]$. Then one of g and h is a unit in $F[x]$; say g and hence g is a constant. Thus $C(f) \approx gC(h)$. Since $C(f)$ is a unit in R, g must be a unit in R and hence in $R[x]$. Therefore, f is irreducible in $R[x]$.

Conversely, if f is irreducible in $R[x]$ and $f = gh$ with $g, h \in F[x]$. We can write $g = ab^{-1}g_1$ with $a, b \in R$, $b \neq 0$ and g_1 primitive in $R[x]$ and $h = cd^{-1}h_1$ with $c, d \in R$, $d \neq 0$ and h_1 primitive in $R[x]$. Consequently, $bdf = acg_1h_1$. Since f and g_1h_1 are primitive,

$$bd \approx bdC(f) \approx C(bdf) \approx C(acg_1h_1) \approx acC(g_1h_1) \approx ac.$$

Thus bd and ac are associates and this implies that $acb^{-1}d^{-1} = \alpha \in R$ is a unit. Hence $f = \alpha g_1 h_1$ in $R[x]$. By hypothesis, one of g_1, h_1 is a unit in $R[x]$; say g_1. Hence g_1 is a constant and so is $g = ab^{-1}g_1$. This implies that f is irreducible in $F[x]$. □

Exercise 11. *Let f be a primitive polynomial in $R[x]$. Prove that f is prime in $R[x]$ if and only if f is prime in $F[x]$.*

Theorem 3.13. *If R is a unique factorization domain, then the polynomial ring $R[x]$ is also a unique factorization domain.*

Proof. Given $f \in R[x]$, we can write f as $f = C(f)f_1$ with f_1 primitive in $R[x]$. Since $C(f) \in R$ and R is a unique factorization domain, if $C(f)$ is not a unit, we can write $C(f)$ as a product of finitely many irreducible elements in R. Theses elements are also irreducible in $R[x]$. Hence it is sufficient to show that every primitive polynomial of positive degree in $R[x]$ can be written as a product of finitely many irreducible elements in $R[x]$. Suppose f is a primitive polynomial in $R[x]$. Since $F[x]$ is a unique factorization domain (c.f. Example 3.8) which contains $R[x]$, $f = p_1 \cdots p_n$ with each p_i irreducible in $F[x]$. Writing $p_i = a_i b_i^{-1} q_i$ with $a_i, b_i \in R$, $b_i \neq 0$ and q_i primitive in $R[x]$. Clearly each q_i is irreducible in $F[x]$ and hence is irreducible in $R[x]$ by Lemma 3.12. Let $a = a_1 \cdots a_n$ and $b = b_1 \cdots b_n$. Then $bf = aq_1 \cdots q_n$. Because $C(f)$ and $C(q_1 \cdots q_n)$ are units in R, it follows that a and b are associates in R. Thus $a = bu$ with u a unit in R. Therefore $f = uq_1 \cdots q_n$ with uq_1 and $q_2, \ldots, q_n$ irreducible in $R[x]$.

To show the uniqueness, as in the proof of Theorem 3.9, we only have to show that every irreducible polynomial in $R[x]$ is prime. Suppose f is irreducible in $R[x]$. If $f \in R$, then by R is a unique factorization domain, f is prime in R. If $f \mid gh$ for some $g, h \in R[x]$, then $lf = gh$ for some $l \in R[x]$. By Lemma 3.10, we have

$$fC(l) \approx C(lf) \approx C(gh) \approx C(g)C(h).$$

This implies that $f \mid C(g)C(h)$ in R and hence $f \mid C(g)$ or $f \mid C(h)$. Therefore, $f \mid g$ or $f \mid h$ in $R[x]$. Therefore, f is prime in $R[x]$. Now suppose that f is a polynomial of positive degree in $R[x]$. f is irreducible in $R[x]$ implies that f is a primitive polynomial in $R[x]$. Lemma 3.12 says that f is irreducible in $F[x]$ and hence f is prime in $F[x]$ because $F[x]$ is a unique factorization domain. By Exercise 11, f is prime in $R[x]$. □

Corollary 3.14. *If R is a unique factorization domain, then the polynomial ring over R in n indeterminates, $R[x_1, \ldots, x_n]$ is also a unique factorization domain.*

Proof. By Theorem 3.13, $R[x_1]$ is a unique factorization domain. Since $R[x_1, \ldots, x_n] = R[x_1, \ldots, x_{n-1}][x_n]$, the proof is now completed by a routine inductive argument. □

Department of Mathematics, National Taiwan Normal University, Taipei, Taiwan, R.O.C.
E-mail address: li@math.ntnu.edu.tw