Chapter 1. Factorization in Rings

1.1 Definition: Let R be a commutative ring (with identity). Let $a, b \in R$. We say that a **divides** b (or a is a **divisor** or **factor** of b, or b is a **multiple** of a), and we write $a\big|b$, when $b = ar$ for some $r \in R$. We say that a and b are **associates**, and we write $a \sim b$, when $a\big|b$ and $b\big|a$.

1.2 Theorem: *Let R be a commutative ring. Let $a, b \in R$. Then*
(1) $a\big|b$ if and only if $b \in \langle a\rangle$ if and only if $\langle b\rangle \subseteq \langle a\rangle$,
(2) $a \sim b$ if and only if $\langle a\rangle = \langle b\rangle$ if and only if a and b have the same multiples and divisors,
(3) $a \sim 0$ if and only if $a = 0$ if and only if $\langle a\rangle = \{0\}$,
(4) $a \sim 1$ if and only if a is a unit if and only if $\langle a\rangle = R$.
(5) if R is an integral domain then $a \sim b$ if and only if $b = au$ for some unit $u \in R$.

Proof: The proof is left as an exercise.

1.3 Definition: Let R be a commutative ring. Let $a \in R$ be a non-zero non-unit. We say that a is **reducible** when $a = bc$ for some non-units $b, c \in R$, and otherwise we say that a is **irreducible**. Note that if a is irreducible then the divisors of a are the units and the associates of a. We say that a is **prime** when for all $b, c \in R$, if $a\big|bc$ then either $a\big|b$ or $a\big|c$.

1.4 Theorem: *Let R be a commutative ring. Let $a, b \in R$ with $a \sim b$. Then*
(1) $a = 0$ if and only if $b = 0$,
(2) a is a unit if and only if b is a unit,
(3) a is irreducible if and only if b is irreducible,
(4) a is prime if and only if b is prime.

Proof: The proof is left as an exercise.

1.5 Theorem: *Let R be an integral domain. Then every prime element in R is also irreducible.*

Proof: The proof is left as an exercise.

1.6 Exercise: Find all primes and irreducible elements in $\mathbf{Z}_{12}$.

1.7 Exercise: Use the method of the Sieve of Eratosthenes to find several irreducible elements in $\mathbf{Z}[\sqrt{3}\,i]$ and also some irreducible elements which are not prime.

1.8 Definition: Let R be a ring. An ideal P in R is called **prime** when $P \neq R$ and for all $a, b \in R$, if $ab \in P$ then either $a \in P$ or $b \in P$. An ideal M in R is called **maximal** when $M \neq R$ and for all ideals A in R, if $M \subseteq A$ then either $A = M$ or $A = R$.

1.9 Example: Show that the maximal ideals in $\mathbf{Z}$ are the ideals of the form $\langle p \rangle$ with p prime, and the prime ideals in $\mathbf{Z}$ are the ideals of the form $\langle p \rangle$ with $p = 0$ or p prime.

1.10 Theorem: *Let R be a commutative ring and let $a \in R$. Then*
(1) a is prime if and only if $\langle a \rangle$ is a non-zero prime ideal, and
(2) if R is an integral domain then a is irreducible if and only if $\langle a \rangle$ is maximal amongst non-zero principal ideals.

Proof: The proof is left as an exercise.

1.11 Theorem: *Let R be a commutative ring. Let P be an ideal in R. Then P is prime if and only if R/P is an integral domain.*

Proof: Suppose that P is prime. Since $P \neq R$ we have $1 \notin P$ (since $\langle 1 \rangle = R$) and so $1 + P \neq 0 + P \in P/R$. Since R is commutative, so is R/P. Finally, note that R/P has no zero divisors because for $a, b \in R$ we have

$$\begin{aligned}(a+P)(b+P) = (0+P) \rightarrow ab + P = 0 + P \rightarrow ab \in P \rightarrow a \in P \text{ or } b \in P \\ \rightarrow a + P = 0 + P \text{ or } b + P = 0 + P.\end{aligned}$$

Conversely, suppose that R/P is an integral domain. Since $1 + P \neq 0 + P \in R/P$, it follows that $1 \notin P$ and so $P \neq R$. Let $a, b \in R$ with $ab \in P$. Then we have $ab + P = 0 + P$, and so $(a+P)(b+P) = 0 + P$. Since R/P has no zero divisors, this implies that either $a + P = 0 + P$ or $b + P = 0 + P$, and so either $a \in P$ or $b \in P$.

1.12 Theorem: *Let R be a commutative ring. Let M be an ideal in R. Then M is maximal if and only if R/M is a field.*

Proof: Suppose M is maximal. Since $M \neq R$ we have $1 \notin M$ and so $1 + M \neq 0 + M \in R/M$. Since R is commutative, so is R/M. Let $a + M$ be a nonzero element in R/M. We must show that $a + M$ is a unit. Since $a + M \neq 0 + M$ we have $a \notin M$. Since $a \notin M$ we have $M \subsetneqq M + \langle a \rangle$. Since M is maximal, we must have $M + \langle a \rangle = R$. In particular, $1 \in M + \langle a \rangle$, say $1 = m + ar$ with $r \in R$. Then $1 + M = ar + M = (a+M)(r+M)$ and so $r + M$ is the inverse of $a + M$.

Conversely, suppose that R/M is a field. Since $1 + M \neq 0 + M$ in R/M, we have $1 \notin M$ so $M \neq R$. Let I be an ideal with $M \subseteq I \subseteq R$. Suppose $I \neq M$. Choose $a \in I$ with $a \notin M$. Since $a \notin M$ we have $a + M \neq 0 + M$ in R/M. Since R/M is a field, $a + M$ has an inverse, say $(a+M)(b+M) = 1 + M$. Then $ab + M = 1 + M$ so we have $1 - ab \in M$. Since $M \subseteq I$ we have $1 - ab \in I$. Since $a \in I$ we have $ab \in I$, so $1 \in I$ and hence $I = R$.

1.13 Example: Since $\mathbf{Q}[x]/\langle x^2 - 2 \rangle \cong \mathbf{Q}[\sqrt{2}]$, which is a field, it follows that $\langle x^2 - 2 \rangle$ is maximal (and prime). In $\mathbf{R}[x]$, however, we have $(x^2 - 2) = (x - \sqrt{2})(x + \sqrt{2})$, and so the ideal $\langle x^2 - 2 \rangle$ is not maximal because $\langle x^2 - 2 \rangle \subsetneqq \langle x - \sqrt{2} \rangle \subsetneqq \mathbf{R}[x]$ and it is not prime because $(x - \sqrt{2})(x + \sqrt{2}) \in \langle x^2 - 2 \rangle$ but $(x - \sqrt{2}) \notin \langle x^2 - 2 \rangle$ and $(x + \sqrt{2}) \notin \langle x^2 - 2 \rangle$.

1.14 Example: In $\mathbf{Z}[x]$, we have $\langle x \rangle = \big\{ f \in \mathbf{Z}[x] \big| f(0) = 0 \big\}$. The ideal $\langle x \rangle$ is prime because for $f, g \in \mathbf{Z}[x]$, if $fg \in \langle x \rangle$ then $f(0)g(0) = 0$ and so either $f(0) = 0$ or $g(0) = 0$. But the ideal $\langle x \rangle$ is not maximal since $\langle x \rangle \subsetneqq \langle 2, x \rangle = \big\{ f \in \mathbf{Z}[x] \big| f(0) \text{ is even} \big\} \subsetneqq \mathbf{Z}[x]$.

1.15 Definition: A **Euclidean domain** (or ED) is an integral domain R together with a function $E : R\setminus\{0\} \to \mathbf{N}$, called a **Euclidean norm**, with the property that for all $a, b \in R$ with $a \neq 0$ there exist $q, r \in R$ such that $b = qa + r$ and either $r = 0$ or $E(r) < E(a)$.

1.16 Definition: A **principal ideal domain** (or PID) is an integral domain R such that every ideal I in R is principal.

1.17 Definition: A **unique factorization domain** (or UFD) is an integral domain R with the property that for every nonzero non-unit $a \in R$ we have

(1) $a = a_1 a_2 \cdots a_l$ for some $l \in \mathbf{Z}^+$ and some irreducible elements $a_i \in R$, and
(2) if $a = a_1 a_2 \cdots a_l = b_1 b_2 \cdots b_m$ where $l, m \in \mathbf{Z}^+$ and each a_i and b_j is irreducible, then $m = l$ and for some permutation $\sigma \in S_m$ we have $a_i \sim b_{\sigma(i)}$ for all i.

1.18 Example: Every field is a Euclidean domain, using any function $E : R \setminus \{0\} \to \mathbf{N}$ as a Euclidean norm. Indeed, given $a, b \in R$ with $a \neq 0$ we can choose $q = \frac{b}{a}$ and $r = 0$ to get $b = aq + r$.

1.19 Example: If F is a field then $F[x]$ is a Euclidean domain with norm $E(f) = \deg(f)$.

1.20 Theorem: *Every Euclidean domain is a principal ideal domain.*

Proof: Let R be a Euclidean domain with norm E. Let A be an ideal in R. If $A = \{0\}$ then A is principal with $A = \langle 0 \rangle$. Suppose that $A \neq \{0\}$. Choose a nonzero element $0 \neq a \in A$ of smallest possible Euclidean norm. We claim that $A = \langle a \rangle$. Since $a \in A$ we have $\langle a \rangle \subseteq A$. Let $b \in A$. Choose $q, r \in R$ such that $b = qa + r$ and either $r = 0$ or $E(r) < E(a)$. Note that $r = b - qa \in A$ so we must have $r = 0$ by the choice of a. Thus $b = qa \in \langle a \rangle$.

1.21 Definition: A ring R is called **Noetherian** when it satisfies the following condition, which is called the **ascending chain condition**: for every ascending chain of ideals $A_1 \subseteq A_2, \subseteq A_3 \subseteq \cdots$ in R, there exists $n \in \mathbf{Z}^+$ such that $A_k = A_n$ for all $k \geq n$.

1.22 Theorem: *Every principal ideal domain is Noetherian.*

Proof: Let R be a principal ideal domain. Let $a_1, a_2, a_3, \cdots \in R$ with

$$\langle a_1 \rangle \subseteq \langle a_2 \rangle \subseteq \langle a_3 \rangle \subseteq \cdots .$$

Let $A = \bigcup\limits_{k=1}^{\infty} \langle a_k \rangle$. Note that A is an ideal. Choose $a \in R$ so that $A = \langle a \rangle$. Since $a \in A$, we can choose $n \in \mathbf{Z}^+$ so that $a \in \langle a_n \rangle$. For all $k \geq n$, we have $\langle a_k \rangle \subseteq A = \langle a \rangle \subseteq \langle a_n \rangle \subseteq \langle a_k \rangle$ and so $\langle a_k \rangle = \langle a_n \rangle$.

1.23 Theorem: *Every principal ideal domain is a unique factorization domain.*

Proof: Let R be a principal ideal domain. Let $a \in R$ be a non-zero non-unit. We claim that a has an irreducible factor. If a is irreducible then we are done. Suppose that a is reducible, say $a = a_1 b_1$ where a_1 and b_1 are non-units. Note that $\langle a \rangle \subsetneqq \langle a_1 \rangle$. If a_1 is irreducible then we are done. Suppose that a_1 is reducible, say $a_1 = a_2 b_2$ where a_2 and b_2 are non-units. Then $a = a_1 b_1 = a_2 b_2 b_1$ and $\langle a \rangle \subsetneqq \langle a_1 \rangle \subsetneqq \langle a_2 \rangle$. If a_2 is irreducible then we are done, and otherwise we continue this procedure. Eventually, the procedure must end giving us an irreducible factor a_n of a, otherwise we would obtain an infinite chain of ideals $\langle a \rangle \subsetneqq \langle a_1 \rangle \subsetneqq \langle a_2 \rangle \subsetneqq \cdots$, contradicting the fact that R is Noetherian.

Next we claim that $a = a_1 a_2 \cdots a_l$ for some $l \in \mathbf{Z}^+$ and some irreducible $a_i \in R$. If a is irreducible then we are done. Suppose that a is reducible. Let a_1 be an irreducible factor of a, and say $a = a_1 b_1$. Note that b_1 is not a unit since, if it was then we would have $a \sim a_1$, but a is reducible and a_1 is not. If b_1 is irreducible then we are done. Suppose b_1 is reducible. Let a_2 be an irreducible factor of b_1 and say $b_1 = a_2 b_2$. As above, note that b_2 is not a unit. If b_2 is irreducible then we are done, and otherwise we continue the procedure. Eventually, the procedure must end giving us $a = a_1 a_2 \cdots a_n b_n$ with each a_i and b_n irreducible, otherwise we would obtain an infinite chain $\langle a \rangle \subsetneqq \langle b_1 \rangle \subsetneqq \langle b_2 \rangle \subsetneqq \cdots$.

Finally, we claim that if $a = a_1 a_2 \cdots a_l = b_1 b_2 \cdots b_l$ with $l, m \in \mathbf{Z}^+$ and each a_i and b_j irreducible, then $m = l$ and for some permutation $\sigma \in S_m$ we have $a_i \sim b_{\sigma(i)}$ for all i. Suppose that $a = a_1 a_2 \cdots a_l = b_1 b_2 \cdots b_m$ where $l, m \in \mathbf{Z}^+$ and the a_i and b_j are irreducible. Since $a_1 \big| a_1 a_2 \cdots a_l$, we have $a_1 \big| b_1 b_2 \cdots b_m$. Since a_1 is irreducible and R is a principal ideal domain, it follows from Part (2) of Theorem 1.10 that $\langle a_1 \rangle$ is a non-zero maximal ideal, hence by Theorem 1.13, $\langle a_1 \rangle$ is a non-zero prime ideal, and hence by Part (1) of Theorem 1.10, a_1 is a prime element. Since a_1 is prime and $a_1 \big| b_1 b_2 \cdots b_m$, it follows that $a_1 \big| b_k$ for some k. After permuting the elements b_i we can assume $a_1 \big| b_1$. Since b_1 is irreducible, its divisors are units and associates. and since a_1 is not a unit, we have $a_1 \sim b_1$. Since $a_1 \sim b_1$ we have $b_1 = a_1 u$ for some unit u. Thus we have $a_1 a_2 \cdots a_l = b_1 b_2 \cdots b_m = a_1 u b_2 b_3 \cdots b_m$, and by cancellation, $a_2 a_3 \cdots a_l = u b_2 b_3 \cdots b_m$. A suitable induction hypothesis gives $l = m$ and $a_i \sim b_i$ for all i, after suitably permuting the elements $b_2, \cdots, b_m$.

1.24 Exercise: Show that for each $d \in \{-2, -1, 2, 3\}$ the ring $\mathbf{Z}\big[\sqrt{d}\,\big]$ is a Euclidean domain with Euclidean norm given by $N(a + b\sqrt{d}) = \big|a^2 - db^2\big|$.

1.25 Exercise: Show that the rings $\mathbf{Z}\big[\sqrt{3}\,i\big]$ and $\mathbf{Z}\big[\sqrt{5}\,\big]$ are not unique factorization domains.

1.26 Exercise: Show that $\mathbf{Z}\big[\frac{1+\sqrt{19}\,i}{2}\big]$ is a PID, but not a ED (under any Euclidean norm).

1.27 Theorem: *(Dirichlet's Approximation Theorem) Let* $x \in \mathbf{R} \setminus \mathbf{Q}$.
(1) For every $n \in \mathbf{Z}^+$ *there exist* $p, q \in \mathbf{Z}$ *with* $1 \leq q \leq n$ *such that* $|qx - p| < \frac{1}{n}$.
(2) There exist infinitely many pairs (p, q) *with* $p \in \mathbf{Z}$ *and* $q \in \mathbf{Z}^+$ *such that* $\left|x - \frac{p}{q}\right| < \frac{1}{q^2}$.

Proof: I may include a proof later.

1.28 Theorem: *(Units in Quadratic Integer Rings) Let* $d \in \mathbf{Z}^+$ *be a non-square. Then there exists a unique smallest unit* $u \in \mathbf{Z}[\sqrt{d}\,]$ *with* $u > 1$, *and the set of all units in* $\mathbf{Z}[\sqrt{d}\,]$ *is* $\mathbf{Z}[\sqrt{d}\,]^* = \left\{\pm u^k \middle| k \in \mathbf{Z}\right\}$.

Proof: I may include a proof later.

1.29 Corollary: *(Pell's Equation) Let* $d \in \mathbf{Z}^+$ *be a non-square. Let* u *be the smallest unit in* $\mathbf{Z}[\sqrt{d}\,]$ *with* $u > 1$. *For* $k \geq 0$, *let* $u^k = p_k + q_k\sqrt{d}$ *with* $p_k, q_k \in \mathbf{Z}$. *Then the solutions* $(x, y) \in \mathbf{Z}^2$ *to Pell's Equation* $x^2 - dy^2 = \pm 1$ *are given by* $(x, y) = (\pm p_k, \pm q_k)$ *wher* $0 \leq k \in \mathbf{Z}$.

1.30 Theorem: *(Prime Elements in the Ring of Gaussian Integers) Every prime element in the ring* $\mathbf{Z}[i]$ *is an associate of exactly one of the following elements.*
(1) $1 + i$,
(2) p, *where* p *is a prime number in* $\mathbf{Z}^+$ *with* $p = 3 \bmod 4$,
(3) $x \pm iy$, *where* $x, y \in \mathbf{Z}$ *with* $0 < y \leq x$ *and* $x^2 + y^2 = p$ *for some prime number* $p \in \mathbf{Z}^+$ *with* $p = 1 \bmod 4$.

Proof: I may include a proof later.

1.31 Corollary: *(Sums of Two Squares) Let* $n \in \mathbf{Z}^+$ *factor as* $n = 2^m \cdot \prod\limits_{\alpha} {p_\alpha}^{k_\alpha} \cdot \prod\limits_{\beta} {q_\beta}^{\ell_\beta}$ *where* $m \in \mathbf{N}$, $k_\alpha, \ell_\beta \in \mathbf{Z}^+$, *the* p_α *are distinct primes with* $p_\alpha = 1 \bmod 4$, *and the* q_β *are distinct primes with* $q_\beta = 3 \bmod 4$. *Then there exists a solution* $(x, y) \in \mathbf{Z}^2$ *to the Sum of Two Squares Equation* $x^2 + y^2 = n$ *if and only if each exponent* ℓ_β *is even, and in this case, the number of solutions* $(x, y) \in \mathbf{Z}^2$ *is equal to* $4 \cdot \prod\limits_{\alpha}(k_\alpha + 1)$.

Proof: I may include a proof later.

1.32 Note: Here are a few remarks about polynomials. Recall that $R[x]$ denotes the ring of polynomials with coefficients in the ring R, and R^R denotes the ring of all functions $f: R \to R$.

(1) A polynomial $f \in R[x]$ determines a function $f \in R^R$. Given $f(x) = \sum\limits_{i=0}^{n} a_i x^i \in R[x]$ we obtain the function $f: R \to R$ given by $f(x) = \sum\limits_{i=0}^{n} a_i x^i$.

(2) Although we do not usually distinguish notationally between the polynomial $f \in R[x]$ and its corresponding function $f \in R^R$, they are not always identical. If the ring R is not commutative then multiplication of polynomials does not agree with multiplication of functions. For $f, g \in R[x]$ given by $f(x) = a + bx$ and $g(x) = c + dx$, in the ring $R[x]$ we have $(fg)(x) = (a + bx)(c + dx) = (ac) + (ad + bc)x + (bd)x^2$, but in the ring R^R we have $(fg)(x) = (a + bx)(c + dx) = ac + adx + bxc + bxdx$.

(3) Equality of polynomials is not the same as equality of functions. For $f, g \in R[x]$ given by $f(x) = \sum\limits_{i=0}^{n} a_i x^i$ and $g(x) = \sum\limits_{i=0}^{m} b_i x^i$ we have $f = g \in R[x]$ if and only if $a_i = b_i$ for all i (and if say $n < m$ then $b_i = a_i = 0$ for $i > n$), but $f = g \in R^R$ if and only if $f(x) = g(x)$ for all $x \in R$. These two notions of equality do not always agree. For example if R is finite then the ring $R[x]$ is infinite but the ring R^R is finite. Indeed if $|R| = n$ then $R[x]$ is countably infinite but $\left|R^R\right| = n^n$. For a more specific example, if $f(x) = x^p - x$ then we have $f \neq 0 \in \mathbf{Z}_p[x]$ (because its coefficients are not equal to zero) but $f = 0 \in {\mathbf{Z}_p}^{\mathbf{Z}_p}$ because, by Fermat's Little Theorem, we have $f(x) = 0$ for all $x \in \mathbf{Z}_p$.

(4) Recall that for $f(x) = \sum\limits_{i=0}^{n} a_i x^i$ with each $a_i \in R$ and $a_n \neq 0$, the element $a_n \in R$ is called the leading coefficient of f, and the positive integer n is called the degree of $f(x)$, and we write $\deg(f) = n$. For convenience, we also define $\deg(0) = -1$. When R is an integral domain, it is easy to see that for $0 \neq f, g \in R[x]$ we have $\deg(fg) = \deg(f) + \deg(g)$. When R is not an integral domain, however, we only have $\deg(fg) \leq \deg(f) + \deg(g)$ because the product of the two leading coefficients can be equal to zero.

(5) When R is an integral domain, because we have $\deg(fg) = \deg(f) + \deg(g)$ for all $0 \neq f, g \in R[x]$, it is easy to see that the units in $R[x]$ are the constant polynomials $f(x) = c$ where c is a unit in R. In particular, when F is a field, the units in $F[x]$ are the elements $f \in F[x]$ with $\deg(f) = 0$.

1.33 Example: In the ring $\mathbf{Z}_4[x]$ we have $(1 + 2x)^2 = 1 + 4x + 4x^2 = 0$, so $f(x) = (1 + 2x)$ is a unit in $\mathbf{Z}_4[x]$.

1.34 Theorem: *(Division Algorithm) Let R be a ring. Let $f, g \in R[x]$ and suppose that the leading coefficient of g is a unit in R. Then there exist unique polynomials $q, r \in R$ such that $f = qg + r$ and $\deg(r) < \deg(g)$.*

Proof: First we prove existence. If $\deg(f) < \deg(g)$ then we can take $q = 0$ and $r = f$. Suppose that $\deg(f \geq \deg(g)$, Say $f(x) = \sum\limits_{i=0}^{n} a_i x^i$ with $a_i \in R$ and $a_n \neq 0$ and $g(x) = \sum\limits_{i=0}^{m} b_i x^i$ with $b_i \in R$ and b_m is a unit. Note that the polynomial $a_n {b_m}^{-1} x^{n-m} g(x)$ has degree n and leading coefficient a_n. It follows that the polynomial $f(x) - a_n {b_m}^{-1} x^{n-m} g(x)$ has degree smaller than n (because the leading coefficients cancel). We can suppose, inductively, that there exist polynomials $p, r \in R[x]$ such that $f(x) - a_n {b_m}^{-1} x^{n-m} g(x) = p(x) g(x) + r(x)$ and $\deg(r) < \deg(g)$. Then we have $f = qg + r$ by taking $q(x) = a_n {b_m}^{-1} x^{n-m} - p(x)$.

Next we prove uniqueness. Suppose that $f = qg + r = pg + s$ where $q, p, r, s \in R[x]$ with $\deg(r) < \deg(g)$ and $\deg(s) < \deg(g)$. Then we have $(q - p)g = s - r$ and so $\deg\big((q-p)g\big) = \deg(s - r)$. Since the leading coefficient of g is a unit (hence not a zero divisor), it follows that $\deg\big((q-p)g\big) = \deg(q - p) + \deg(g)$. If we had $q - p \neq 0$ then we would have $\deg\big((q-p)g\big) \geq \deg(g)$ but $\deg(s - r) < \deg(g)$, giving a contradiction. Thus we must have $q - p = 0$. Since $q - p = 0$ we have $s - r = (q - p)g = 0$. Since $q - p = 0$ and $s - r = 0$ we have $q = p$ and $r = s$, proving uniqueness.

1.35 Corollary: *(The Remainder Theorem) Let R be a ring, let $f \in R[x]$, and let $a \in R$. When we divide $f(x)$ by $(x - a)$ to obtain the quotient $q(x)$ and remainder $r(x)$, the remainder is the constant polynomial $r(x) = f(a)$.*

Proof: Use the division algorithm to obtain $q, r \in R[x]$ such that $f = q(x)(x - a) + r(x)$ and $\deg(r) < \deg(x - a)$. Since $\deg(x - a) = 1$ we have $\deg(r) \in \{-1, 0\}$, and so r is a constant polynomial, say $r(x) = c$ with $c \in R$. Then we have $f(x) = q(x)(x - a) + c$. Put in $x = a$ to get $f(a) = q(a)(a - a) + c = q(a) \cdot 0 + c = c$.

1.36 Corollary: *(The Factor Theorem) Let R be a commutative ring, let $f \in R[x]$ and let $a \in R$. Then $f(a) = 0$ if and only if $(x - a)\big| f(x)$.*

Proof: Suppose that $f(a) = 0$. Choose $q, r \in R[x]$ such that $f(x) = q(x)(x - a) + r(x)$ and $\deg(r) < \deg(x - a)$. Then $r(x)$ is the constant polynomial $r(x) = f(a) = 0$ and so we have $f(x) = q(x)(x - a)$. Since $f(x) = (x - a)q(x)$ we have $(x - a)\big| f(x)$. Conversely, suppose that $(x - a)\big| f(a)$ and choose $p \in R[x]$ so that $f(x) = (x - a)p(x)$. Then $f(a) = (a - a)p(a) = 0 \cdot p(a) = 0$.

1.37 Definition: Let R be a commutative ring, let $f \in R[x]$, and let $a \in R$. We say that a is a **root** of f when $f(a) = 0$. When $f \neq 0$, we define the **multiplicity** of a as a root of f to be the largest $m = m(f, a) \in \mathbf{N}$ such that $(x - a)^m \big| f(x)$ (where we use the convention that $(x - a)^0 = 1$). Note that a is a root of f if and only if $m(f, a) \geq 1$.

1.38 Example: Let $f(x) = x^3 - 3x - 2 \in \mathbf{Q}[x]$. Since $f(x) = (x + 1)^2 (x - 2) \in \mathbf{Q}[x]$, we have $m(f, 2) = 1$ and $m(f, -1) = 2$.

1.39 Exercise: Let p be an odd prime and let $f(x) = x^p - a \in \mathbf{Z}_p[x]$. Find $m(f, a)$.

1.40 Theorem: *(The Roots Theorem) Let R be an integral domain, let $0 \neq f \in R[x]$ and let $n = \deg(f)$. Then*

(1) f has at most n distinct roots in R, and

(2) if $a_1, a_2, \cdots, a_\ell$ are all of the distinct roots of f in R and $m_i = m(f, a_i)$ for $1 \leq i \leq \ell$, then $(x-a_1)^{m_1}(x-a_2)^{m_2}\cdots(x-a_\ell)^{m_\ell}\big|f(x)$ and so $\sum\limits_{i=1}^{\ell} m(f,a) \leq n$.

Proof: We prove Part (1) and leave the proof of Part (2) as an exercise. If $\deg(f) = 0$, then $f(x) = c$ for some $0 \in c \in R$, and so $f(x)$ has no roots. Let f be a polynomial with $\deg(f) = n \geq 1$ and suppose, inductively, that every polynomial $g \in R[x]$ with $\deg(g) = n-1$ has at most $n-1$ distinct roots. Suppose that a is a root of f in R. By the Factor Theorem, $(x-a)\big|f(x)$ so we can choose a polynomial $g \in R[x]$ so that $f(x) = (x-a)g(x)$. Note that $\deg(g) = n-1$ so, by the induction hypothesis, g has at most $n-1$ distinct roots. Let $b \in R$ be any root of f with $b \neq a$. Since $f(x) = (x-a)g(x)$ and $f(b) = 0$ we have $0 = f(b) = (b-a)g(b)$. Since $(b-a)g(b) = 0$ and $(b-a) \neq 0$ and R has no zero divisors, it follows that $g(b) = 0$. Thus b must be one of the roots of g. Since every root b of f with $b \neq a$ is equal to one of the roots of g, and since g has at most $n-1$ distinct roots, it follows that f has at most n distinct roots, as required.

1.41 Note: Here are a few remarks about irreducible polynomials.

(1) When F is a field, we know that $F[x]$ is a unique factorization domain. For $f \in F[x]$ we know that $f = 0$ if and only if $\deg(f) = -1$, and f is a unit if and only if $\deg(f) = 0$, and for $0 \neq f, g \in F[x]$ we know that $\deg(fg) = \deg(f) + \deg(g)$. It follows that for $f \in F[x]$, if $\deg(f) = 1$ then f is irreducible. It also follows that for $f \in F[x]$, if $\deg(f) = 2$ or 3 then f is reducible in $F[x]$ if and only if f has a f has a root in F.

(2) When p is a fairly small prime number and n is a fairly small positive integer, it is easy to list all reducible and irreducible polynomials $f \in \mathbf{Z}_p[x]$ with $\deg(f) \leq n$. Note that it suffices to list monic polynomials (since for $f \in \mathbf{Z}_p[x]$ and $0 \neq c \in \mathbf{Z}_p[x]$ we have $f \sim cf$). We start by listing all monic polynomials of degree 1, that is all polynomials of the form $f(x) = x + a$ with $a \in \mathbf{Z}_p$, and noting that they are all irreducible. Having constructed all reducible and irreducible monic polynomials of all degrees less than n, we can construct all of the reducible monic polynomials of degree n by forming products of the reducible monic polynomials of smaller degree in all possible ways, and then all the remaining monic polynomials of degree n must be irreducible.

(3) For $f \in \mathbf{C}[x]$, we know that f is irreducible if and only if $\deg(f) = 1$. For $f \in \mathbf{R}[x]$, we know that f is irreducible polynomial if and only if either $\deg(f) = 1$ or $f(x) = ax^2 + bx + c$ for some $a, b, c \in \mathbf{R}$ with $a \neq 0$ and $b^2 - 4ac < 0$. For $R = \mathbf{Z}$ or $\mathbf{Q}$, it is a more challenging problem to determine which polynomials are irreducible in $R[x]$. The next few theorems are related to this problem.

1.42 Exercise: List all monic reducible and irreducible polynomials in $\mathbf{Z}_2[x]$ of degree less than 4, then determine the number of irreducible polynomials in $\mathbf{Z}_2[x]$ of degree 4.

1.43 Definition: Let $f \in \mathbf{Z}[x]$. The **content** of f, denoted by $c(f)$, is the greatest common divisor of the coefficients of f. We say that f is **primitive** when $c(f) = 1$.

1.44 Note: Let $a_0, a_1, \cdots, a_n \in \mathbf{Z}$, let $r \in \mathbf{Z}$ and let $d = \gcd(a_0, a_1, \cdots, a_n)$. Then $\gcd\big(ra_0, ra_1, \cdots, ra_n\big) = |r| \gcd(a_0, a_1, \cdots, a_n)$ and $\gcd\big(\frac{a_0}{d}, \frac{a_1}{d}, \cdots, \frac{a_n}{d}\big) = 1$. It follows that for $f(x) = \sum\limits_{i=0}^{n} a_i x^i$ we have $c(rf) = |r| c(f)$ and that if we let $g(x) = \frac{1}{c(f)} f(x)$ then we have $g(x) \in \mathbf{Z}[x]$ and $c(g) = 1$.

1.45 Theorem: *(Gauss' Lemma)*
(1) For all $f, g \in \mathbf{Z}[x]$ we have $c(fg) = c(f)c(g)$.
(2) Let $0 \neq f \in \mathbf{Z}[x]$ and let $g(x) = \frac{1}{c(f)} f(x) \in \mathbf{Z}[x]$. Then f is irreducible in $\mathbf{Q}[x]$ if and only if g is irreducible in $\mathbf{Z}[x]$.

Proof: Let $f, g \in \mathbf{Z}[x]$. If $f = 0$ or $g = 0$ then we have $c(fg) = 0 = c(f)c(g)$. Suppose that $f \neq 0$ and $g \neq 0$. Let $h(x) = \frac{1}{c(f)} f(x)$ and $k(x) = \frac{1}{c(g)} g(x)$. Then we have $h, k \in \mathbf{Z}[x]$ with $c(h) = c(k) = 1$ and $fg = c(f)c(g)hk$ so that $c(fg) = c(f)c(g)c(hk)$. Thus to prove Part (1) it suffices to show that $c(hk) = 1$. Let $h(x) = \sum\limits_{i-0}^{n} a_i x^i$ and $k(x) = \sum\limits_{i=0}^{m} b_i x^i$ with $a_n \neq 0$ and $b_m \neq 0$. Suppose, for a contradiction, that $c(hk) \neq 1$. Let p be a prime factor of $c(hk)$. Then p divides all of the coefficients of $(hk)(x) = (a_0b_0) + (a_1b_0 + a_0b_1)x + \cdots + (a_nb_m)x^{n+m}$. Since $c(h) = 1$, p does not divide all the coefficients of $h(x)$ so we can choose an index $r \geq 0$ so that $p|a_i$ for all $i < r$ and $p\nmid a_r$. Since $c(k) = 1$ we can choose an index $s \geq 0$ so that $p|b_i$ for all $i < s$ and $p\nmid b_s$. Since p divides every coefficient of $(hk)(x)$, it follows that in particular p divides the coefficient

$$c_{r+s} = a_0 b_{r+s} + a_1 b_{r+s-1} + \cdots + a_r b_s + \cdots + a_{r+s-1} b_1 + a_{r+s}.$$

Since $p|c_{r+s}$ and $p|a_i$ for all $i < r$ and $p|b_i$ for all $i < s$ it follows that $p|a_r b_s$. Since p is prime and $p|a_r b_s$ it follows that $p|a_r$ or $p|b_s$. But r and s were chosen so that $p\nmid a_r$ and $p\nmid b_s$ so we have obtained the desired contradiction. This proves Part (1).

To prove Part (2), let $0 \neq f(x) \in \mathbf{Z}[x]$ and let $g(x) = \frac{1}{c(f)} f(x)$, and note that $c(g) = 1$. Suppose that g is reducible in $\mathbf{Z}[x]$, say $g(x) = h(x)k(x)$ where $h(x)$ and $k(x)$ are non-units in $\mathbf{Z}[x]$. Since $c(h)c(k) = c(hk) = c(g) = 1$ it follows that $c(h) = c(k) = 1$. Note that $h(x)$ cannot be a constant polynomial since if we had $h(x) = r$ with $r \in \mathbf{Z}$, then we would have $|r| = c(h) = 1$ so that $h(x) = \pm 1$, but then h would be a unit. Similarly $k(x)$ cannot be a constant polynomial. Since $h(x)$ and $k(x)$ are nonconstant polynomials in $\mathbf{Z}[x]$, they are also nonconstant polynomials in $\mathbf{Q}[x]$. Since $f(x) = c(f)g(x) = c(f)h(x)k(x)$ and since $c(f)h(x)$ and $k(x)$ are both nonconstant polynomials (hence nonunits) in $\mathbf{Q}[x]$, it follows that $f(x)$ is reducible in $\mathbf{Q}[x]$.

Conversely, suppose that $f(x)$ is reducible in $\mathbf{Q}[x]$, say $f(x) = h(x)k(x)$ where h and k are nonzero, nonunits in $\mathbf{Q}[x]$. Since h and k are nonzero nonunits in $\mathbf{Q}[x]$, they are nonconstant polynomials. Let a be the least common multiple of the denominators of the coefficients of $h(x)$ and let b be the least common multiple of the coefficients of $k(x)$, and note that $ah(x) \in \mathbf{Z}[x]$ and $bk(x) \in \mathbf{Z}[x]$. Let $p(x) = \frac{1}{c(ah)} ah(x)$ and let $q(x) = \frac{1}{c(bk)} bk(x)$ and note that $p(x) \in \mathbf{Z}[x]$ and $q(x) \in \mathbf{Z}[x]$ with $c(p) = c(q) = 1$ and that $\deg(p) = \deg(h)$ and $\deg(q) = \deg(k)$. Since $f(x) = ah(x)\, bk(x) = c(ah)c(bk)p(x)q(x)$ we have $c(f) = c(ah)c(bk)c(pq) = c(ah)c(bk)$ and so $g(x) = \frac{1}{c(f)} f(x) = \frac{1}{c(ah)c(bk)} ah(x)\, bk(x) = p(x)q(x)$. Since $g(x) = p(x)q(x)$ where $p(x)$ and $q(x)$ are nonconstant polynomials in $\mathbf{Z}[x]$, we see that $g(x)$ is reducible in $\mathbf{Z}[x]$.

1.46 Corollary: *Let $0 \neq f(x) \in \mathbf{Z}[x]$. Then $f(x)$ is reducible in $\mathbf{Q}[x]$ if and only if $f(x)$ can be factored as a product of two nonconstant polynomials in $\mathbf{Z}[x]$.*

Proof: If $f(x)$ can be factored as $f(x) = g(x)h(x)$, where $g(x)$ and $h(x)$ are nonconstant polynomials in $\mathbf{Z}[x]$, then because $g(x)$ and $h(x)$ are also nonconstant polynomials in $\mathbf{Q}[x]$ (hence nonunits in $\mathbf{Q}[x]$), it follows immediately that f is reducible in $\mathbf{Q}[x]$. Suppose, conversely, that $f(x)$ is reducible in $\mathbf{Q}[x]$. Let $p(x) = \frac{1}{c(f)} f(x)$ and note that $p(x) \in \mathbf{Z}[x]$ with $c(p) = 1$. By Gauss' Lemma, $p(x)$ is reducible in $\mathbf{Z}[x]$. Choose nonunits $k, h \in \mathbf{Z}[x]$ such that $p = kh \in \mathbf{Z}[x]$. Since $c(k)c(h) = c(kh) = c(p) = 1$ we have $c(k) = c(h) = 1$. Since k and h are nonunits with $c(k) = c(h) = 1$, it follows that k and h are nonconstant polynomials (indeed, if $k(x)$ was constant with $k(x) = r$ then we would have $|r| = c(k) = 1$ so that $k(x) = r = \pm 1$, but then $k(x)$ would be a unit). Let $g(x) = c(f)k(x)$ and note that since $k(x)$ is nonconstant, so is $g(x)$. Then we have $f(x) = g(x)h(x)$, which is a product of two nonconstant polynomials in $\mathbf{Z}[x]$.

1.47 Example: Let $f(x) = 6x + 30 \in \mathbf{Z}[x]$. Note that $c(f) = 6$. Since $\deg(f) = 1$, it follows that f is irreducible in $\mathbf{Q}[x]$. But since $c(f) = 6$, it follows that f is reducible in $\mathbf{Z}[x]$, indeed in $\mathbf{Z}[x]$ we have $f(x) = 2 \cdot 3 \cdot (x + 5)$.

1.48 Theorem: *(Rational Roots) Let $f(x) = \sum\limits_{i=0}^{n} c_i x^i$ where $n \in \mathbf{Z}^+$, each $c_i \in \mathbf{Z}$ and $c_n \neq 0$. Let $r, s \in \mathbf{Z}$ with $s \neq 0$ and $\gcd(r, s) = 1$. Then if $f\left(\frac{r}{s}\right) = 0$ then $r | c_0$ and $s | c_n$.*

Proof: Suppose that $f\left(\frac{r}{s}\right) = 0$, that is $c_0 + c_1 \frac{r}{s} + c_2 \frac{r^2}{s^2} + \cdots + c_n \frac{r^n}{s^n} = 0$. Multiply by s^n to get

$$0 = c_0 s^n + c_1 s^{n-1} r^1 + \cdots + c_{n-1} s^1 r^{n-1} + c_n r^n.$$

Thus we have

$$\begin{aligned} c_0 s^n &= -r(c_1 s^{n-1} + \cdots + c_{n-1} s^1 r^{n-2} + c_n r^{n-1}) \text{ and} \\ c_n r^n &= -s\big(c_0 s^{n-1} + c_1 s^{n-2} r^1 + \cdots + c_{n-1} r^{n-1}\big) \end{aligned}$$

and it follows that $r \big| c_0 s^n$ and that $s \big| c_n r^n$. Since $\gcd(r, s) = 1$ we also have $\gcd(r, s^n) = 1$, and since $r \big| c_0 s^n$ it follows that $r | c_0$. Since $\gcd(s, r) = 1$ we also have $\gcd(s, r^n) = 1$, and since $s \big| c_n r^n$ it follows that $s | c_n$.

1.49 Exercise: Show that $\sqrt{1 + \sqrt{2}} \notin \mathbf{Q}$.

1.50 Theorem: *(Modular Reduction) Let $f(x) = \sum\limits_{i=0}^{n} c_i x^i$ with $n \in \mathbf{Z}^+$, $c_i \in \mathbf{Z}$ and $c_n \neq 0$. Let p be a prime number with $p \not| c_n$. Let $\overline{f}(x) = \sum\limits_{i=0}^{n} \overline{c_i}\, x^i \in \mathbf{Z}_p[x]$ where $\overline{c_i} = [c_i] \in \mathbf{Z}_p$. If $\overline{f}$ is irreducible in $\mathbf{Z}_p[x]$ then f is irreducible in $\mathbf{Q}[x]$.*

Proof: Suppose that $f(x)$ is reducible in $\mathbf{Q}[x]$. By the corollary to Gauss' Lemma, we can choose two nonconstant polynomials $g, h \in \mathbf{Z}[x]$ such that $f = gh \in \mathbf{Z}[x]$. Write $g(x) = \sum\limits_{i=0}^{k} a_i x^k \in \mathbf{Z}[x]$ and $h(x) = \sum\limits_{i=0}^{\ell} b_i x^i \in \mathbf{Z}[x]$ with $a_k \neq 0$, $b_\ell \neq 0$ and $k, \ell \geq 1$. Let $\overline{g} = \sum\limits_{i=0}^{k} \overline{a}_i x^i \in \mathbf{Z}_p[x]$ and $\overline{h}(x) = \sum\limits_{i=0}^{\ell} \overline{b}_i x^i \in \mathbf{Z}_p[x]$, and note that $\overline{f} = \overline{g}\,\overline{h} \in \mathbf{Z}_p[x]$. Since $c_n = a_k b_\ell$ and $p \not| c_n$ it follows that $p \not| a_k$ and $p \not| b_\ell$ in $\mathbf{Z}$ so $\overline{a}_k \neq 0$ and $\overline{b}_\ell \neq 0$ in $\mathbf{Z}_p$. Thus $\deg(\overline{g}) = \deg(g) = k$ and $\deg(\overline{h}) = \deg(h) = \ell$ so that $\overline{g}$ and $\overline{h}$ are nonconstant polynomials in $\mathbf{Z}_p[x]$, and so the polynomial $\overline{f} = \overline{g}\overline{h}$ is reducible in $\mathbf{Z}_p[x]$.

1.51 Exercise: Prove that $f(x) = x^5 + 2x + 4$ is irreducible in $\mathbf{Q}[x]$ by working in $\mathbf{Z}_3[x]$.

1.52 Theorem: *(Eisenstein's Criterion) Let* $f(x) = \sum\limits_{i=0}^{n} c_i x^i$ *with* $n \in \mathbf{Z}^+$, $c_i \in \mathbf{Z}$ *and* $c_n \neq 0$. *Let* p *be a prime number such that* $p_i | c_i$ *for* $0 \leq i < n$ *and* $p \nmid c_n$ *and* $p^2 \nmid c_0$. *Then* f *is irreducible in* $\mathbf{Q}[x]$.

Proof: Suppose, for a contradiction, that $f(x)$ is reducible in $\mathbf{Q}[x]$. By the corollary to Gauss' Lemma, we can choose two nonconstant polynomials $g, h \in \mathbf{Z}[x]$ such that $f = gh \in \mathbf{Z}[x]$. Write $g(x) = \sum\limits_{i=0}^{k} a_i x^k \in \mathbf{Z}[x]$ and $h(x) = \sum\limits_{i=0}^{\ell} b_i x^i \in \mathbf{Z}[x]$ with $k, \ell \geq 1$ and $a_k \neq 0$, $b_\ell \neq 0$. Since $c_0 = a_0 b_0$ and $p | c_0$ but $p^2 \nmid c_0$, it follows that p divides exactly one of the two numbers a_0 and b_0. Suppose that p divides a_0 but not b_0 (the case that p divides b_0 but not a_0 is similar). Since $p | c_1$, that is $p \big| (a_0 b_1 + a_1 b_0)$, and $p | a_0$ it follows that $p | a_1 b_0$, and since $p \nmid b_0$ it follows that $p | a_1$. Since $p | c_2$, that is $p \big| (a_0 b_2 + a_1 b_1 + a_2 b_0)$ and $p | a_0$ and $p | a_1$, it follows that $p | a_2 b_0$, and since $p \nmid b_0$ it then follows that $p | a_2$. Repeating this argument we find, inductively, that $p | a_i$ for all $i \geq 0$, and in particular we have $p | a_k$. Since $c_n = a_k b_\ell$ and $p | a_k$ it follows that $p | c_n$, giving the desired contradiction.

1.53 Example: Note that $f(x) = 5x^5 + 3x^4 - 18x^3 + 12x + 6$ is irreducible in $\mathbf{Q}[x]$ by Eisenstein's Criterion using $p = 3$.

1.54 Exercise: Let p be a prime number. Show that $f(x) = 1 + x + x^2 + \cdots + x^{p-1}$ is irreducible in $\mathbf{Q}[x]$,