January 2, 2024

Rings

Table of Contents

---

1. Definitions and constructions

1.1. Definitions, examples, and basic properties

1.1.1. *A ring* is a set R with two binary operations: addition and multiplication, so that

(i) under addition R is an abelian group (which means that the addition is associative, there is a zero 0, and an inverse $-a$ for every $a \in R$),

(ii) the multiplication is associative,

(iii) and the distributive laws hold: for any $a, b, c \in R$, $a(b+c) = ab + ac$ and $(b+c)a = ba + ca$.

If the multiplication in R is commutative, R is said to be *a commutative ring*. If R has a neutral element with respect to mutiplication, this element is called *the identity* of R and is denoted by 1_R or just 1, and R is said to be *a unital ring*.

1.1.2. Examples. (0) A ring consisting of a single element 0 is not forbidden, and is called *the zero ring*. This ring is unital, with $1 = 0$.

(i) $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ are commutative unital rings.

(ii) For any $n \in \mathbb{N}$, $\mathbb{Z}_n$ is a commutative unital ring.

(iii) The ring $2\mathbb{Z}$ of even integers is nonunital.

(iv) The (commutative unital) ring of *Gaussian integers* is the subring $\mathbb{Z}[i] = \{a + bi,\ a, b \in \mathbb{Z}\}$ of $\mathbb{C}$.

(v) The ring of Hamilton's *real quaternions* is $\mathbb{H} = \left\{a + bi + cj + dk,\ a, b, c, d \in \mathbb{R}\right\}$, where the elements i, j, k satisfy the relations of Q_8: $i^2 = j^2 = k^2 = -1$, $ij = -ji = k$, $jk = -kj = i$, $ki = -ik = j$. $\mathbb{H}$ is a noncommutative unital ring.

(vi) If X is a set, the the set of real-valued functions $f\colon X \longrightarrow \mathbb{R}$ is a (commutative, unital) ring under addition $(f + g)(x) = f(x)g(x)$ and multiplication $(fg)(x) = f(x)g(x)$.

(vii) (a subring of the ring of all functions $\mathbb{R} \longrightarrow \mathbb{R}$). The continuous functions vanishing at ∞ (that is, with $f(x) \longrightarrow 0$ as $x \longrightarrow 0$) form a nonunital subring of $C(\mathbb{R})$.

(viii) The real valued polynomials $a_n x^n + \cdots + a_1 x + a_0$, $n \geq 0$, $a_i \in \mathbb{R}$, form a (commutative, unital) ring $\mathbb{R}[x]$ (which is a subring of the ring of functions $\mathbb{R} \longrightarrow \mathbb{R}$). The polynomials with integer coefficients form the ring $\mathbb{Z}[x]$ (a subring of $\mathbb{R}[x]$).

(ix) For any n, $M_{n,n}(\mathbb{R})$ is the ring of $n \times n$ matrices with real entries; this ring is unital and, for $n \geq 2$, noncommutative.

(x) Let X be a set; then the power set $\mathcal{P}(X)$ of X is a (commutative, unital) ring with addition $A+B = A \triangle B$ and multiplication $AB = A \cap B$; it is called *the Boolean ring* of X.

1.1.3. There are only two "general elementary properties" of rings: if R is a ring then

(i) $a0 = 0a = 0$ for all $a \in R$;

(ii) $(-a)b = a(-b) = -ab$ for all $a, b \in R$; in particular, if R is a unital ring, then $(-1)a = -a$ for all $a \in R$.

1.2. The center, zero divisors, units, idempotent, nilpotent, and unipotent elements

1.2.1. Given a ring R, *the center* $Z(R)$ of R is the set $\left\{a \in R : ab = ba \text{ for all } b \in R\right\}$. $Z(R)$ is a commutative subring of R.

1.2.2. A nonzero element a of a ring R is said to be *a left zero divisor* in R if $ab = 0$ for some nonzero $b \in R$, and *a right zero divisor* if $ba = 0$ for some nonzero $b \in R$. In a commutative ring these two notions coincide, and a left=right zero divisor is called *a zero divisor*.

Examples. (i) A nonzero $k \in \mathbb{Z}_n$ is a zero divisor in $\mathbb{Z}_n$ iff $d = \gcd(k, n) \neq 1$ (in which case $k \cdot (n/d) = 0$).

(ii) In the ring of functions $X \longrightarrow \mathbb{R}$, a nonzero function is a zero divisor iff it vanishes at some point of X. In the ring $C(\mathbb{R})$ of continuous functions $\mathbb{R} \longrightarrow \mathbb{R}$, a function $f \neq 0$ is a zero divisor iff $f_{|_I} = 0$ for some interval $I \subset \mathbb{R}$.

1.2.3. If a nonzero element a of a ring R is not a left zero divisor, then it can be "cancelled" from the left: $ab = ac$ implies that $b = c$. Indeed, $ab = ac$ implies that $a(b - c) = 0$, so $b - c = 0$. Similarly, if a is not a right zero divisor, then it can be "cancelled" from the right: $ba = ca$ implies that $b = c$.

1.2.4. A commutative unital ring without zero divisors is called *an integral domain*, or *an ID*. Any integral domain is *cancellative*: if $ab = ac$ with $a \neq 0$, then $b = c$.

Examples. $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, $\mathbb{R}[x]$, $\mathbb{Z}[i]$, and $\mathbb{Z}[x]$ are integral domains. $\mathbb{Z}_n$ is an integral domain iff n is prime. $2\mathbb{Z}$, $\mathbb{H}$, $C(\mathbb{R})$, $M_{n,n}(\mathbb{R})$ for $n \geq 2$, and $\mathcal{P}(X)$ if $|X| > 1$ are not integral domains, for different reasons.

1.2.5. An element u of a unital ring R is said to be *a left unit* in R if $uv = 1$ for some $v \in R$; *a right unit* if $vu = 1$ for some $v \in R$; and *a unit* if it is both a left and a right unit, in which case $uv = vu = 1$ for some $v \in R$, denoted by u^{-1}. Units in a ring R form a group under multiplication, denoted by R^*.

Examples. (i) The only units in $\mathbb{Z}$ are ± 1.

(ii) In $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, $\mathbb{H}$, all nonzero elements are units.

(iii) $k \in \mathbb{Z}_n$ is a unit in this ring iff $\gcd(k, n) = 1$.

(iv) In the ring $\mathbb{Z}[i]$ of Gaussian integers, the units are $\pm 1, \pm i$.

(v) In the ring of *integer quaternions* $\left\{a + bi + cj + dk,\ a, b, c, d \in \mathbb{Z}\right\} \subset \mathbb{H}$, the units form the group Q_8.

(vi) In the ring of functions, or continuous functions, to $\mathbb{R}$ a function is a unit iff it vanishes at no point.

1.2.6. Let R be a unital ring; then no left unit in R is a right zero divisor in R, and no right unit in R is a left zero divisor. In particular, in a commutative unital ring no unit is a zero divisor.

1.2.7. A (nonzero) ring in which every nonzero element is a unit is called *a division ring*. A commutative division ring is called *a field*.

Examples. (i) $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ are fields.

(ii) If p is a prime integer, $\mathbb{Z}_p$ is a field, denoted by $\mathbb{F}_p$.

(iii) $\mathbb{H}$ is a division ring (in which $(a+bi+cj+dk)^{-1} = \frac{a-bi-cj-dk}{a^2+b^2+c^2+d^2}$).

1.2.8. Lemma. *Any finite ring without zero divisors is a division ring. In particular, any finite integral domain is a field.*

(Actually, by *Wedderburn's theorem*, every finite division ring is a field.)

Proof. In such a ring R, the set $R \setminus \{0\}$ is a finite cancellative semigroup, and so, is a group. ∎

1.2.9. A nonzero element e of a ring R is said to be *idempotent* if $e^2 = e$. In a unital ring, an idempotent element is either equal to 1 or is a zero divisor: $e(1-e) = (1-e)e = 0$.

Examples. (i) In a unital ring R the identity 1 is idempotent, and 1 is the only idempotent if R has no zero divisors.

(ii) In $\mathbb{Z}_6$, 3 and 4 are idempotent (and we will see why in 1.9.6).

(iii) A function $f\colon X \longrightarrow \mathbb{R}$ is idempotent iff it takes values 0 and 1 only.

1.2.10. An element z of a ring R is said to be *nilpotent* if $z^n = 0$ for some $n \in \mathbb{N}$. A nilpotent element is either 0 or is a zero divisor: $zz^{n-1} = z^{n-1}z = 0$ (where n is the minimal positive integer for which $z^n = 0$).

Examples. (i) An element $k \in \mathbb{Z}_n$ is nilpotent iff k is divisible by all prime divisors of n: in $\mathbb{Z}_{120}$, only 0, 30, and 60 are nilpotent.

(ii) In $M_{n,n}(\mathbb{R})$ any strictly upper-triangular matrix, with all zeroes below and on the main diagonal, is nilpotent.

1.2.11. An element a of a unital ring R is said to be *unipotent* if $a-1$ is nilpotent, that is, if $(a-1)^n = 0$ for some $n \in \mathbb{N}$.

A unipotent element is always a unit, with $a^{-1} = (1-(1-a))^{-1} = 1+(1-a)+(1-a)^2+\cdots+(1-a)^{n-1}$ (where n is such that $(a-1)^n = 0$).

1.3. Constructions of rings

Let R be a ring; based on R, we may produce new rings:

1.3.1. Subrings. A subset S of R is *a subring of* R if S is a ring under the operations of R; for this it must be that S is a subgroup of R under addition, $S - S = S$, and is closed under multiplication, $S \cdot S \subseteq S$.

Let S be a subring of R. If R is commutative, then S is commutative; if R has no zero divisors, then S has no zero divisors; if R is unital and S contains 1_R, then S is unital with $1_S = 1_R$, and every unit in S is a unit in R.

If R is a subring of a ring A with $R \subseteq Z(A)$ and P is a subset of A, then the minimal ring containing R and P is denoted by $R[P]$; it consists of finite linear combinations with coefficients from R of products of elements of P.

1.3.2. Quotient rings. If R is a ring and S is a subring of R, then the factor R/S is an abelian group, but the multiplication on R may not induce a well defined operation on R/S.

Example. $\mathbb{Z}$ is a subring of $\mathbb{R}$, and multiplication is not defined on $\mathbb{R}/\mathbb{Z}$: we have $\frac{1}{2} = \frac{3}{2}$ in $\mathbb{R}/\mathbb{Z}$, but $\frac{1}{2}\cdot\frac{1}{2} = \frac{1}{4} \neq \frac{3}{2}\cdot\frac{1}{2} = \frac{3}{4}$.

For multiplication to be well defined on R/S so that R/S would be a ring, S must be *an ideal* in R; see 1.5 below. In this case, R/S is commutative if R is; R/S is unital if R is; if a is a zero divisor in R, then $\overline{a}$ (the image of a in R/Z) is either 0 or a zero divisor in R/S; and if u is a unit in R, then $\overline{u}$ is a unit in R/S.

1.3.3. Direct products of rings. For two rings R_1 and R_2, their *direct product* is the ring $R_1 \times R_2$ with $(a_1, b_1) + (a_2, b_2) = (a_1 + a_2, b_1 + b_2)$ and $(a_1, b_1)(a_2, b_2) = (a_1a_2, b_1b_2)$. The direct product of an infinite collection of rings is defined similarly.

1.3.4. Rings of functions. Given a set X, the ring of functions $X \longrightarrow R$ is the direct product $R^X = \prod_{x\in X} R$.

1.3.5. Boolean rings. Given a set X, the Boolean ring $\mathcal{P}(X)$ (see 1.1.2(x)) is isomorphic to the ring $\mathbb{Z}_2^X$ of functions $X \longrightarrow \mathbb{Z}_2$, where a set $A \in \mathcal{P}(X)$ corresponds to the indicator function $1_A(x) = \begin{cases} 1, & x \in A \\ 0, & x \notin A \end{cases}$ of A. (Indeed, then $A \triangle B \mapsto 1_A + 1_B$ and $A \cap B \mapsto 1_A 1_B$.) Every nonzero element in this ring is idempotent.

1.3.6. Group rings. Let G be a group, or a semigroup; *the group ring* RG is defined as the ring of formal sums $\sum_{i=1}^n a_i g_i$ where $n \geq 0$, $a_i \in R$, and g_i are distinct elements of G. The sum and the product of two elements $u = \sum_{i=1}^n a_i g_i$ and $v = \sum_{j=1}^n b_j g_j$ are defined as $u + v = \sum_{i=1}^n a_i g_i + \sum_{j=1}^n b_j g_j$ and $uv = \sum_{\substack{1 \leq i \leq n \\ 1 \leq j \leq m}} a_i b_j g_i g_j$ with collected similar (that is, corresponding to the same g) terms.

Under the identification of $a \in R$ with $a1_G$, R is a subring of RG.

RG is commutative if both R and G are commutative. If R is unital, then RG is unital.

1.3.7. The polynomial ring. The ring $R[x]$ of *polynomials* $f = \sum_{i=0}^n a_i x^i = a_n x^n + \cdots + a_1 x + a_0$ with coefficients from R is the (semi)group ring RG where G is the semigroup $\{1, x, x^2, \ldots\}$. The ring $R[x_1, \ldots, x_k]$ of polynomials in k variables $x_1, \ldots, x_k$ is RG where G is the free commutative semigroup with 1, generated by $x_1, \ldots, x_k$.

If R is commutative, then then $R[x]$, and $R[x_1, \ldots, x_k]$ for any k, are commutative. If R has no zero divisors, then $R[x]$, and $R[x_1, \ldots, x_k]$ for any k, have no zero divisors. If R is unital and has no zero divisors, then the only units in $R[x]$ are those from R.

1.3.8. The ring of formal power series. The formal power series $\sum_{i=0}^\infty a_i x^i$, $a_i \in R$, with the addition $\sum_{i=0}^\infty a_i x^i + \sum_{i=0}^\infty b_i x^i = \sum_{i=0}^\infty (a_i + b_i) x^i$ and the multiplication $\sum_{i=0}^\infty a_i x^i \sum_{j=0}^\infty b_j x^j = \sum_{k=0}^\infty \big(\sum_{i=0}^k a_i b_{k-i}\big) x^k$ form a ring denoted by $R[[x]]$. If R has no zero divisors then $R[[x]]$ has no zero divisors. If R is unital, then $R[[x]]$ is unital; in this case, a series $\sum_{i=0}^\infty a_i x^i$ is a unit iff a_0 is a unit in R.

1.3.9. Rings of endomorphisms. Let A be an abelian group written additively. *The endomorphisms* of A, that is, homomorphisms $A \longrightarrow A$, under the operations $(\varphi + \psi)(g) = \varphi(g) + \psi(g)$ and $(\varphi\psi)(g) = \varphi(\psi(g))$ form a ring denoted by $\mathrm{End}(A)$.

$\mathrm{End}(A)$ is, usually, noncommutative. $\mathrm{End}(A)$ is unital, with $1 = \mathrm{Id}_A$. The units in $\mathrm{End}(A)$ form the group $\mathrm{Aut}(A)$.

1.3.10. The matrix rings. For any $n \in \mathbb{N}$, *the matrix ring* $M_{n,n}(R)$ is the ring of $n \times n$ matrices with entries from R, with the conventional addition and multiplication. For $n \geq 2$, $M_{n,n}$ is noncommutative even if R is commutative. If R is unital, then $M_{n,n}(R)$ is unital; a matrix $A \in M_{n,n}(R)$ is a unit iff $\det A$ is a unit in R.

When $R = \mathbb{Z}$, $M_{n,n}(\mathbb{Z})$ is isomorphic to the ring $\mathrm{End}(\mathbb{Z}^n)$. For a general ring R, a matrix from $M_{n,n}(R)$ defines an endomorphism φ of the additive group of R^n by $\varphi(u) = Au$, which induces a homomorphism $M_{n,n}(R) \longrightarrow \mathrm{End}(R^n)$; however, this homomorphism is not, generally speaking, surjective. (For example, in the case of commutative R, φ defined this way has the additional redundnat property that $\varphi(au) = a\varphi(u)$ for all $a \in R$.)

1.4. Fields and rings of fractions

1.4.1. Let R be an integral domain. On the set of pairs (a, b) with $a, b \in R$, $b \neq 0$, introduce an equivalence relation $(a_1, b_1) \sim (a_2, b_2)$ is $a_1 b_2 = a_2 b_1$ (which can be checked to be an equivalence relation indeed), and denote the equivalence class of (a, b) by $\frac{a}{b}$. Define $\frac{a_1}{b_1} + \frac{a_2}{b_2} = \frac{a_1 b_2 + a_2 b_1}{b_1 b_2}$ and $\frac{a_1}{b_1} \cdot \frac{a_2}{b_2} = \frac{a_1 a_2}{b_1 b_2}$ (and it can be checked that these addition and multiplication are well defined). Under these operations, the set $F = \left\{ \frac{a}{b},\ a, b \in R,\ b \neq 0 \right\}$ is a field, called *the field of fractions* of R.

The mapping $a \mapsto \frac{a}{1}$ identifies R with a subring of F; the inverse a^{-1} of a nonzero $a \in R$ is $\frac{1}{a} \in F$.

1.4.2. As a corollary we get that any integral domain is a subring of a field.

1.4.3. Examples. (i) The field of fractions of $\mathbb{Z}$ is $\mathbb{Q}$.

(ii) The field of fractions of the polynomial ring $F[x]$, where F is a field, is the field $F(x)$ of rational functions over F.

(iii) The field of fractions of $\mathbb{Z}[x]$ is $\mathbb{Q}[x]$.

(iv) The field of fractions of the ring $F[[x]]$ of power series over a field F is $\left\{ \sum_{i=-n}^\infty a_i x^i,\ n \in \mathbb{N},\ a_i \in F \right\}$.

1.4.4. Here is a more general construction. Let R be a commutative ring, and let D be a nonempty *multiplicative* (that is, with $D \cdot D \subseteq D$) subset of R that does not contain zero or zero divisors. *The ring of fractions* $D^{-1}R$ is the ring of equivalence classes $\frac{a}{d}$ (under the equivalence relation introduced in 1.4.1) of pairs (a, d) with $a \in R$, $d \in D$, with the addition and multiplication as in 1.4.1. $D^{-1}R$ is always a unital ring, with $1 = \frac{d}{d}$ for any $d \in D$. (Notice that $(d, d) \sim (d', d')$ for any other $d' \in D$.) R is a subring of $D^{-1}R$, with $a \in R$ identified with $\frac{ad}{d}$ for any $d \in D$. In $D^{-1}R$, the elements of D are units, with $d^{-1} = \frac{d}{d^2}$.

1.4.5. Even more generally, the condition in 1.4.4 that D contains no zero divisors can be dropped, if the equivalence relation from 1.4.1 is replaced by $(a_1, d_1) \sim (a_2, d_2)$ if $(a_1d_2 - a_2d_1)d = 0$ for some $d \in D$.

1.4.6. The field of fractions F of an integral domain R is"the minimal" field containing R, in the sense that if $\varphi\colon R \longrightarrow K$ is an injective homomorphism to a field K, then φ extends to a homomorphism $F \longrightarrow K$ (by defining $\varphi\big(\frac{a}{b}\big) = \varphi(a)\varphi(b)^{-1}$).

Similarly, if D is a multiplicative set in a commutative ring R, and $\varphi\colon R \longrightarrow S$ is a homomorphism to a unital ring S in which all elements $\varphi(d)$ with $d \in D$ are units, then φ extends to a homomorphism $D^{-1}R \longrightarrow S$.

1.5. Ideals and factorization of rings

1.5.1. Let R be a ring. A subring I of R is called *a left ideal* if $RI \subseteq I$ (that is, for any $a \in I$ and $b \in R$ one has $ba \in I$). A subring I is called *a right ideal* if $IR \subseteq I$, and *a two-sided ideal*, or just *ideal*, if both $RI, IR \subseteq I$. (In a commutative ring all three notions coincide.)

1.5.2. Examples. (0) In any ring R, $0 = \{0\}$ and R itself are ideals. (An ideal I is said to be *nontrivial* if $I \neq 0, R$, and *proper* if $I \neq R$.)

(i) In $\mathbb{Z}$, ideals=subrings=subgroups are the subsets $n\mathbb{Z}$ with $n \in \mathbb{Z}$.

(ii) $\mathbb{Z}$ is a subring but not an ideal in $\mathbb{R}$.

(iii) The subrings $\mathbb{Z}$ and $\mathbb{Z}[x^2]$ of the polynomial ring $\mathbb{Z}[x]$ are not ideals in this ring. The subring $x\mathbb{Z}[x] = \{a_nx^n + \cdots + a_1x,\ a_i \in \mathbb{Z}\}$ is an ideal.

(iv) The ring $C(\mathbb{R})$ of continuous functions on $\mathbb{R}$ is not an ideal in the ring $\mathbb{R}^{\mathbb{R}}$ of functions $\mathbb{R} \longrightarrow \mathbb{R}$. The set of functions vanishing at, say, 0 is an ideal in $\mathbb{R}^{\mathbb{R}}$.

(v) In the matrix ring $M_{n,n}(\mathbb{Z})$, the matrices with zero first column form a left but not right ideal, the matrices with zero first row form a right but not left ideal.

1.5.3. If I is a (two-sided) ideal in a ring R, then multiplication is well defined on the set of cosets of I in R, making R/I a ring, called *the quotient ring* of R.

1.5.4. Let R be a ring. The intersection of any collection of (left, right, two-sided) ideals of R is a (left, right, two-sided) ideal.

The sum $I + J = \big\{a + b,\ a \in I,\ b \in J\big\}$ of two (or any collection of) (left, right, two-sided) ideals is a (left, right, two-sided) ideal.

1.5.5. *The product* IJ of two ideals of a ring R is defined as the set of finite sums $\sum_{i=1}^{n} a_ib_i$ with $a_i \in I$, $b_i \in J$. IJ is a left ideal if I is a left ideal, and a right ideal if J is a right ideal. If I is a right and J is a left ideal, then $IJ \subseteq I \cap J$.

1.5.6. Let R be a unital ring. For a subset P of R, the minimal left ideal containing P is called *the left ideal generated by* P; it is the sum $\sum_{a\in P} Ra$, that is, the set of finite sums $\sum_{i=1}^{n} c_ia_i$ with $c_i \in R$ and $a_i \in S$ for all i. In particular, the left ideal generated by a single element a is the set $Ra = \big\{ca,\ c \in R\big\}$.

Similarly, *the right ideal generated by* P is $\sum_{a\in P} aR$, and *the two-sided ideal generated by* P is $\sum_{a\in P} RaR$.

If R is commutative, then the ideal generated by a set $P \subseteq R$ is denoted by (P). The ideal Ra generated by a single element $a \in R$ is called *principal* and is denoted by (a); it consists of all multiples of a.

1.5.7. For a subset S of a ring R, *the annihilator* of S is the set $\mathrm{Ann}(S) = \big\{a \in R : aS = 0\big\}$. $\mathrm{Ann}(S)$ is a left ideal in R; if S is a left ideal, then $\mathrm{Ann}(S)$ is a two-sided ideal.

1.5.8. In any commutative ring R, nilpotent elements form an ideal, called *the nilradical* of R and denoted by $\mathrm{Nil}(R)$. The ring $R/\mathrm{Nil}(R)$ has no nilpotent elements.

1.6. Homomorphisms of rings

1.6.1. A mapping $\varphi\colon R \longrightarrow S$ from a ring R to a ring S is said to be *a homomorphism* if $\varphi(a+b) = \varphi(a)+\varphi(b)$ and $\varphi(ab) = \varphi(a)\varphi(b)$ for all $a, b \in R$ (that is, if φ is a homomorphism between the additive groups and the multiplicative semigroups of R and S).

1.6.2. An invertible (that is, bijective) homomorphism is called *an isomorphism*; the inverse of an isomorphism is an isomorphism too. Two rings R and S are said to be *isomorphic* if there is an isomorphism $\varphi\colon R \longrightarrow S$; we write $R \cong S$ in this case.

1.6.3. Examples. (0) For any rings R and S, the mapping $\varphi(a) = 0 \in S$ for all $a \in R$ is *the zero homomorphism* $R \longrightarrow S$.

(i) If S is a subring of a ring R, then the embedding $S \longrightarrow R$ is a homomorphism. If S is a quotient ring of R, then the factorization mapping $R \longrightarrow S$ is a homomorphism.

(ii) The mapping $\mathbb{Z} \longrightarrow \mathbb{Z}$, $n \mapsto 2n$, is not a ring homomorphism (though is a homomorphism of additive groups).

(iii) For any unital ring R, there is a homomorphism $\mathbb{Z} \longrightarrow R$ defined by $n \mapsto n1_R$.

(iv) For the ring R^X of functions $X \longrightarrow R$ from a set X to a ring R and a point $x_0 \in X$, the mapping $f \mapsto f(x_0)$, $f \in R^X$, is a ring homomorphism $R^X \longrightarrow R$, called *the evaluation homomorphism* at x_0.

(vi) Also, for any ring R and an element $c \in Z(R)$, the evaluation homomorphism $R[x] \longrightarrow R$ at c is defined by $a_n x^n + \cdots + a_1 x + a_0 \mapsto a_n c^n + \cdots + a_1 c + a_0$.

(vii) Any ring homomorphism $\varphi\colon R \longrightarrow S$ induces homomorphisms $R[x] \longrightarrow S[x]$, $R[[x]] \longrightarrow S[[x]]$, and $M_{n,n}(R) \longrightarrow M_{n,n}(S)$ for all $n \in \mathbb{N}$.

1.6.4. If R is a unital ring and $\varphi\colon R \longrightarrow S$ is a homomorphism, then $\varphi(1_R)$ is either equal to 0 or is an idempotent element of S.

1.6.5. If $\varphi\colon R \longrightarrow S$ is a ring homomorphism, then for any subring P of R, $\varphi(P)$ is a subring of S; and for any subring Q of S, $\varphi^{-1}(Q)$ is a subring of R.

1.6.6. If $\varphi\colon R \longrightarrow S$ is a ring homomorphism, then for any (left, right, two-sided) ideal J of S, $\varphi^{-1}(J)$ is a (left, right, two-sided) ideal in R.

For a (left, right, two-sided) ideal I of R, $\varphi(I)$ may not be an ideal in S; however, $\varphi(I)$ is a (left, right, two-sided) ideal in S if φ is surjective.

1.6.7. For a homomorphism $\varphi\colon R \longrightarrow S$, *the kernel* $\ker\varphi$ of φ is the ideal $\varphi^{-1}(0) = \{a \in R : \varphi(a) = 0\}$.

A ring homomorphism φ is injective iff $\ker\varphi = 0$.

1.7. Isomorphism theorems for rings

1.7.1. The 1st isomorphism theorem. *Let $\varphi\colon R \longrightarrow S$ be a ring homomorphism. Then $R/\ker\varphi \cong \varphi(R)$, under the isomorphism defined by $a \bmod \ker\varphi \mapsto \varphi(a)$.*

1.7.2. The 2nd isomorphism theorem. *Let R be a ring, S be a subring of R, and I be an ideal in R. Then $S+I$ is a subring of R, $S \cap I$ is an ideal in S, and $(S+I)/I \cong S/(S \cap I)$, under the isomorphism that maps $s \bmod I \mapsto s \bmod (S \cap I)$, $s \in S$.*

1.7.3. The 3rd isomorphism theorem. *Let R be a ring and I, J be ideals of R with $I \subseteq J$. Then J/I is an ideal in R/I, and $(R/I)/(J/I) \cong R/J$ under the isomorphism that maps $(a \bmod I) \bmod (J/I) \mapsto a \bmod J$.*

1.7.4. The 4th (lattice) isomorphism theorem. *Let R be a ring and I be an ideal of R. Then the subrings of the quotient ring R/I are in one-to-one correspondence with the subrings of R that contain I: a subring S of R corresponds to the subring S/I or R/I. Under this correspondence intersections of subrings correspond to intersections, sums to sums, and ideals to ideals.*

1.8. Ideals in the ring of fractions

Let R be a commutative ring and let $D \subset R$ be a multiplicative subset of R that doesn't contain 0 or zero divisors. Then the proper ideals in $D^{-1}R$ are in one-to-one correspondence with ideals in D disjoint from D: each proper ideal in $D^{-1}R$ has form $J = D^{-1}I$ for some ideal I in R with $I \cap D = \emptyset$, and such I is unique, defined by $I = J \cap R$.

Indeed, if I is an ideal in R, then $D^{-1}I$ is an ideal in $D^{-1}R$. If $I \cap D \neq \emptyset$, with $d \in I \cap D$, then $1 = \frac{d}{d} \in D^{-1}I$, so $D^{-1}I = R$. If $I \cap D = \emptyset$, then $\frac{a}{d} \neq 1$ for all $d \in D$, $a \in I$, so $D^{-1}I$ is a proper ideal in

$D^{-1}R$.

On the other hand, let J be an deal in $D^{-1}R$; then $I = J \cap R$ is an ideal in R. We clearly have $D^{-1}I \subseteq J$; also, for every $b \in J$, we have $b = \frac{a}{d}$ for some $a \in R$ and $d \in D$, and $a = bd \in J \cap R = I$, so $b \in D^{-1}I$.

1.9. Direct products and sums of rings

1.9.1. *The direct product* of two rings R_1 and R_2 is the ring $R_1 \times R_2$ as a set with the componentwise addition and multiplication: $(a_1, b_1) + (a_2, b_2) = (a_1 + a_2, b_1 + b_2)$ and $(a_1, b_1)(a_2, b_2) = (a_1a_2, b_1b_2)$. The direct product of two (and finitely many) rings is also called *the direct sum* of R_1 and R_2 and is denoted $R_1 \oplus R_2$.

R_1 and R_2 are naturally identified with the subrings $R_1 \times \{0\}$ and $\{0\} \times \mathbb{R}_2$ of $R_1 \times R_2$. Under this identification, R_1 and R_2 are ideals in $R_1 \times R_2$, and we have $(R_1 \times R_2)/R_1 \cong R_2$, $(R_1 \times R_2)/R_2 \cong R_1$.

1.9.2. The ring $R_1 \oplus R_2$ is commutative iff both R_1 and R_2 are commutative, and is unital iff both R_1 and R_2 are unital, with $1 = (1_{R_1}, 1_{R_2})$.

If the rings R_1 and R_2 are unital, the units in $R_1 \times R_2$ are the pairs (u_1, u_2) where u_i are units in R_i, $i = 1, 2$, so that $(R_1 \times R_2)^* = R_1^* \times R_2^*$ as multiplicative groups.

The product $R_1 \times R_2$ has zero divisors even if R_1 and R_2 do not: $ab = 0$ for any $a \in R_1$ and $b \in R_2$.

1.9.3. If R is a ring with subrings R_1, R_2 such that $R \cong R_1 \times R_2$ under an isomorphism identical on both R_1 and R_2, we say that R is *the (internal) direct product*, or *the (internal) direct sum*, of R_1 and R_2, and write $R = R_1 \times R_2$, or $R = R_1 \oplus R_2$. This is so iff $R_1 \cap R_2 = 0$, $R_1 + R_2 = R$, and $R_1R_2 = 0$; the last condition is satisfied iff R_1 and R_2 are ideals in R (assuming that $R_1 \cap R_2 = 0$).

1.9.4. If R is a unital ring and $R = R_1 \oplus R_2$ for subrings R_1, R_2 of R, then $1 = e_1 + e_2$ for some $e_1 \in R_1$ and $e_2 \in R_2$; e_1 and $e_2 = 1 - e_1$ are identities in the corresponding rings and are idempotent elements in R.

If R is commutative and unital then, conversely, for any idempotent $e \in R$ we have $R = Re \oplus R(1 - e)$.

1.9.5. Let R be unital and $R = R_1 \oplus R_2$. Then any (left/right/two-sided) ideal in R has form $I_1 \times I_2$, where I_1 is an ideal in R_1 and I_2 is an ideal in R_2. Indeed, for an ideal I in R put $I_1 = e_1I$ and $I_2 = e_2I$, where $e_1 \in R_1$ and $e_2 \in R_2$ are such that $e_1 + e_2 = 1$. Then I_1 is an ideal in R_1 (and in R), I_2 is an ideal in R_2 (and in R), $I_1 + I_2 = I$ and $I_1 \cap I_2 = 0$.

1.9.6. Example. Since $\mathbb{Z}_6 \cong \mathbb{Z}_2 \oplus \mathbb{Z}_3$ (see the Chinese remainder theorem for rings below), it must have two idempotent elements distinct from 1: the identity of $\mathbb{Z}_2$ and the identity of $\mathbb{Z}_3$. $\mathbb{Z}_2$ is represented in $\mathbb{Z}_6$ as the subgroup $\{0, 3\}$, where 3 plays the role of identity; and indeed, $3^2 = 3$ in $\mathbb{Z}_6$. $\mathbb{Z}_3$ is represented in $\mathbb{Z}_6$ as the subgroup $\{0, 2, 4\}$, with 4 as the identity since $4 = 1 \bmod 3$; and indeed, $4^2 = 4$ in $\mathbb{Z}_6$.

1.9.7. Given a family R_α, $\alpha \in \Lambda$, of rings, the *direct product* of this family is defined as

$$\prod_{\alpha\in\Lambda} R_\alpha = \left\{(a_\alpha)_{\alpha\in\Lambda},\ a_\alpha \in R_\alpha \text{ for all } \alpha \in \Lambda\right\}$$

with componentwise addition and multiplication, and the *direct sum* of this family is the subring

$$\bigoplus_{\alpha\in\Lambda} R_\alpha = \left\{(a_\alpha)_{\alpha\in\Lambda},\ a_\alpha \in R_\alpha \text{ for all } \alpha \in \Lambda \text{ with } a_\alpha \neq 0 \text{ for only finitely many } \alpha\right\}$$

of $\prod_{\alpha\in\Lambda} R_\alpha$.

$\prod_{\alpha\in\Lambda} R_\alpha$ is unital if all R_α are unital, and then the units in this product are the elements $(u_\alpha)_{\alpha\in\Lambda}$ such that u_α is a unit in R_α for all α. If Λ is infinite, $\bigoplus_{\alpha\in\Lambda} R_\alpha$ is not unital.

2. The theory of "divisibility" of ideals

In this section, R is a commutative unital ring.

2.1. Principal ideals

2.1.1. For $a, b \in R$ we say that a *divides* b, a *is a divisor of* b, or b *is a multiple of* a, and write $a \mid b$, if there exists $c \in R$ such that $ac = b$.

2.1.2. For $a \in R$, *the principal ideal generated by* a is the ideal $(a) = Ra = \{ba,\ b \in R\}$, which consists of all multiplies of a.

2.1.3. The ideal (a, b) generated by two elements $a, b \in R$ is $Ra + Rb = (a) + (b)$; the maximal ideal contained in both (a) and (b) is $(a) \cap (b)$; the product $(a)(b)$ is the ideal (ab).

2.1.4. We have $(0) = 0$ and $(1) = R$. For $a \in R$ we have $(a) = 1$ iff a is a unit in R.

2.1.5. Proposition. *A commutative unital ring R is a field iff R has no ideals except* 0 *and* (1).

Proof. If R is a field and I is a nonzero ideal of R, let $a \in I \setminus \{0\}$; then $(a) = R$ and $(a) \subseteq I$, so $I = R = (1)$. If R is not a field, let a be a nonzero nonunit element of R; then (a) is and ideal of R distinct from 0 and (1). ■

Since fields have no nontrivial ideals, it follows that any nonzero homomorphism from a field to a ring is injective, and that fields have no nontrivial factors.

2.1.6. Proposition. *If R is an integral domain, then for elements $a, b \in R$ one has $(a) = (b)$ iff $a = ub$ for some unit $u \in R$.*

Proof. If $a = ub$ where u is a unit in R, then $b \in (a)$ and so $(b) \subseteq (a)$, and since also $b = u^{-1}a$, then also $(a) \subseteq (b)$.

Now assume that R is an ID, and let $(a) = (b) \neq 0$. Then $a \in (b)$, so $a = ub$ for some $u \in R$, and similarly $b = va$ for some $v \in R$. Then $a = uva$, so $a(1 - uv) = 0$, and so $uv = 1$, which means that u is a unit. ■

2.2. Divisibility of ideals

2.2.1. For two elements $a, b \in R$ we have $a \mid b$ iff $(a) \supseteq b$ iff $(a) \supseteq (b)$. Extending this relation to ideals, for two ideals I and J of R we say that I *divides* J and write $I \mid J$ if $I \supseteq J$.

2.2.2. For ideals $I, J \subseteq R$, *the greatest common divisor* $\gcd(I, J)$ of I and J is the minimal ideal that contains both I and J, which is $(I, J) = I + J$. Likewise, *the least common multiple* $\operatorname{lcm}(I, J)$ of I and J is the maximal ideal that is contained in both I and J, that is, is $I \cap J$.

2.2.3. For two elements $a, b \in R$, *the greatest common divisor* $\gcd(a, b)$ of a and b is the element d of R such that $d \mid a, b$ and for any $c \in R$ with $c \mid a, b$ one has $c \mid d$. (If R is an integral domain then $\gcd(a, b)$, if exists, is defined up to multiplication by a unit.) $\gcd(a, b)$ may not exist. (Here is an example: in the ring $\mathbb{Z}[\sqrt{-5}] = \left\{n + m\sqrt{-5},\ n, m \in \mathbb{Z}\right\}$, the elements $6 = 2 \cdot 3 = (1 + \sqrt{-5})(1 - \sqrt{-5})$ and $2(1 + \sqrt{-5})$ have no gcd.) However "in the sense of ideals" $\gcd((a), (b))$ always exists and equals the ideal (a, b). If (a, b) happens to be principal, $(a, b) = (d)$, then $\gcd(a, b) = d$ "in the sense of elements".

2.2.4. Likewise, *the least common multiple* $\operatorname{lcm}(a, b)$ of $a, b \in R$ is the element $l \in R$ such that $a, b \mid l$ and $l \mid c$ for any c such that $a, b \mid c$. (If R is an integral domain then $\operatorname{lcm}(a, b)$, if exists, is defined up to multiplication by a unit.) $\operatorname{lcm}(a, b)$ may not exist in R, but $\operatorname{lcm}((a), (b))$ always exists and equals $(a) \cap (b)$, and if it is principal, $(a) \cap (b) = (l)$ for some $l \in R$, then $\operatorname{lcm}(a, b) = l$.

2.2.5. The fact that for ideals $I, J \subseteq R$ we have $IJ \subseteq I \cap J$ now reads as $\operatorname{lcm}(I, J) \mid IJ$.

2.3. Comaximal ideals and the Chinese remainder theorem for rings

2.3.1. Two ideals $I, J \subseteq R$ are said to be *comaximal* if $\gcd(I, J) = (1)$, that is, if $I + J = R$.

2.3.2. The Chinese remainder theorem for two ideals. *If ideals $I, J \subseteq R$ are comaximal, then $IJ = I \cap J$, and $R/(IJ) \cong R/I \times R/J$ under the isomorphism $a \bmod (IJ) \mapsto (a \bmod I, a \bmod J)$.*

Proof. $IJ \subseteq I \cap J$ always holds. Since $I + J = (1)$, there are $a \in I$ and $b \in J$ such that $a + b = 1$. Now, for any $c \in I \cap J$ we have $c = 1c = ac + bc$ where $ac \in IJ$ and $bc \in JI = IJ$, so $c \in IJ$.

We have a homomorphism $\varphi\colon R \longrightarrow (R/I) \times (R/J)$ defined by $c \mapsto (c \bmod I, c \bmod J)$. We have $\ker \varphi = I \cap J = IJ$, we only need to show that φ is surjective. Let a and b be as above. Let x and y be arbitrary elements of R; define $c = bx + ay$. Then $c = (1 - a)x + ay = x + (y - x)a = x \bmod I$ and $c = bx + (1 - b)y = y + (x - y)b = y \bmod J$, so $\varphi(c) = (x \bmod I, y \bmod J)$. ■

2.3.3. The following fact allows to extend the Chinese remainder theorem, by induction on k, to the case of several ideals:

Lemma. *Let ideals $I_1, \ldots, I_k \subseteq R$ be comaximal with an ideal $J \subseteq R$, $I_i + J = (1)$ for all i. Then the ideals $\prod_{i=1}^{k} I_i$ and J are also comaximal, $\left(\prod_{i=1}^{k} I_i\right) + J = (1)$.*

Proof. For every i, let $a_i \in I_i$ and $b_i \in J$ be such that $a_i + b_i = 1$. Then

$$1 = \prod_{i=1}^{k}(a_i + b_i) = \prod_{i=1}^{k} a_i + \sum_j \prod_{i=1}^{k} c_{j,i}$$

where for each j at least one of $c_{j,i}$ is equal to b_i and so $\prod_{i=1}^{k} c_{j,i} \in J$, whereas $\prod_{i=1}^{k} a_i \in I_1 \cdots I_k$. ■

2.3.4. The Chinese remainder theorem for k ideals. *Let ideals $I_1, \ldots, I_k \subseteq R$ be pairwise comaximal, $I_i + I_j = (1)$ for any $i \neq j$. Then $\prod_{i=1}^{k} I_i = \bigcap_{i=1}^{k} I_i$, and $R/\bigl(\prod_{i=1}^{k} I_i\bigr) \cong R/I_1 \times \cdots \times R/I_k$ under the isomorphism $a \bmod \bigl(\prod_{i=1}^{k} I_i\bigr) \mapsto \bigl(a \bmod I_1, \ldots, a \bmod I_k\bigr)$.*

2.3.5. Applying the Chinese remainder theorem to integers, we obtain that for any coprime $n, m \in \mathbb{N}$, $\mathbb{Z}_{nm} \cong \mathbb{Z}_n \times \mathbb{Z}_m$ as rings. It follows that $\mathbb{Z}_{nm}^* \cong \mathbb{Z}_n^* \times \mathbb{Z}_m^*$ (as multiplicative groups), and that $\varphi(nm) = \varphi(n)\varphi(m)$, where φ is Euler's function, $\varphi(n) = |\mathbb{Z}_n^*|$.

By induction on k, or using the version of the Chinese remainder theorem for k ideals, we obtain that if $n = p_1^{r_1} \cdots p_k^{r_k}$ is the prime factorization of $n \in \mathbb{N}$, then $\mathbb{Z}_n \cong \mathbb{Z}_{p_1^{r_1}} \times \cdots \times \mathbb{Z}_{p_k^{r_k}}$ as rings, that $\mathbb{Z}_n^* \cong \mathbb{Z}_{p_1^{r_1}}^* \times \cdots \times \mathbb{Z}_{p_k^{r_k}}^*$ as groups, and that $\varphi(n) = \prod_{i=1}^{k} \varphi(p_i^{r_i}) = \prod_{i=1}^{k} p_i^{r_i - 1}(p_i - 1)$.

2.4. Maximal and prime ideals

2.4.1. An ideal $M \subseteq R$ is said to be *maximal* if $M \neq R$ (that is, is proper) and M is not divisible by any ideal except itself and (1) (that is, $M \subseteq I$ for an ideal I implies that $I = M$ or $I = R$).

M is a maximal ideal in R iff R/M contains no proper nonzero ideal that is, is a field.

2.4.2. A proper ideal $P \subseteq R$ is said to be *prime* if for any ideals $I, J \in R$, if $P \mid IJ$ then $P \mid I$ or $P \mid J$ (that is, $IJ \subseteq P$ implies that $I \subseteq P$ or $J \subseteq P$).

Equivalently, a proper ideal P is prime iff for any $a, b \in R \setminus P$ one has $ab \notin P$, that is, iff R/P is an integral domain. Indeed, if P is prime and $a, b \in R$ are such that $ab \in P$, then $(ab) = (a)(b) \subseteq P$, so $(a) \subseteq P$ or $(b) \subseteq P$, so $a \in P$ or $b \in P$. Conversely, assume that $a, b \notin P$ implies that $ab \notin P$. Let I and J be ideals with $I, J \not\subseteq P$. Let $a \in I \setminus P$ and $b \in J \setminus P$; then $ab \notin P$, so $IJ \not\subseteq P$.

2.4.3. It follows that every maximal ideal is prime, and that the converse is not true.

2.4.4. Examples. (i) The zero ideal of R is prime iff R is an integral domain, and is maximal iff R is a field.

(ii) In $\mathbb{Z}$, an ideal (n), with $n \geq 2$, is prime iff it is maximal iff n is a prime integer.

(iii) If F is a field, then in the ring $F[x, y]$ the ideals (x), (y), $(y - x^2)$ are prime but nor maximal, the ideal (x, y) is maximal, the ideals (x^2) and (xy) are not prime.

(iv) In the ring $\mathbb{Z}[x]$, the ideals (x) and (2) are prime but not maximal, the ideal $(2, x)$ is maximal.

(v) If X is a compact metric (or topological) space, then the maximal ideals in the ring $C(X)$ of continuous functions on X have form $M_c = \bigl\{f \in C(X) : f(c) = 0\bigr\}$, $c \in X$. There are prime nonmaximal ideals in this ring.

2.4.5. If P is a prime ideal in R and S is a subring of R with $S \not\subseteq P$, then $P \cap S$ is a prime ideal in S. (An analogous statement for maximal ideals fails.)

2.4.6. More generally, if $\varphi\colon S \longrightarrow R$ is a homomorphism of (commutative unital) rings and P is a prime ideal in R with $\varphi(S) \not\subseteq P$, then $\varphi^{-1}(P)$ is a prime ideal in S.

If M is a maximal ideal in R, then $\varphi^{-1}(M)$ may not be maximal in S; but it is maximal in the case φ is surjective.

2.4.7. Let $\varphi\colon S \longrightarrow R$ be a surjective homomorphism of (commutative unital) rings. If I is an ideal in S containing $\ker\varphi$, then $\varphi(I)$ is a prime ideal in R iff I is prime, and maximal iff I is maximal.

2.4.8. The following theorem requires Zorn's lemma:

Theorem. *Any nonzero commutative unital ring has a maximal ideal. Moreover, every proper ideal of R is contained in a maximal ideal.*

Proof. Let I be a proper ideal in R (if I is not given, put $I = 0$), and let $\mathcal{I}$ be the set of all proper ideal in R that contain I. If $\mathcal{C}$ is a chain in $\mathcal{I}$, then $J = \bigcup_{L \in \mathcal{C}} L$ is an ideal containing I; since $1 \notin J$, J is proper. Hence, every chain in $\mathcal{I}$ is bounded, and by Zorn's lemma, $\mathcal{I}$ has a maximal element. ■

2.4.9. Proposition. *The intersection of all prime ideals of R is the nilradical* $\mathrm{Nil}(R)$.

Proof. We need to show that an element a of R is contained in all prime ideals of R iff a is nilpotent. If a is nilpotent, $a^n = 0$ for some n, and P is a prime ideal, then $a^n \in P$ implies that $a \in P$ or $a^{n-1} \in P$; but the latter also implies, by induction, that $a \in P$.

On the other hand, assume that $a \in R$ is not nilpotent. Using Zorn's lemma we can find an ideal P that is maximal among ideals containing no power of a. Let's show that P is prime. Assume that there are $b, c \in R \setminus P$ such that $bc \in P$. Then the ideals $P + (b)$ and $P + (c)$ are larger than P and so, $a^n \in P + (b)$, $a^m \in P + (c)$ for some $n, m \in \mathbb{N}$. Then $a^{n+m} \in P^2 + (b)P + (c)P + (bc) \subseteq P$, contradiction. ■

2.4.10. The intersection of all maximal ideals of R is called *the Jacobson radical* of R and is denoted by $\mathrm{Jac}(R)$. The elements of the Jacobson radical can be characterized in the following way:

Proposition. *$a \in \mathrm{Jac}(R)$ iff $1 + ab$ are units in R for all $b \in R$.*

Proof. Let $a \in R$. For any ideal M, $a \in M$ implies that $1 + ab \notin M$ for every $b \in R$; so, if a is contained in all maximal ideals of R, then $1 + ab$ is not contained in any maximal ideal, so is not contained in any proper ideal of R, so is a unit. Conversely, if $a \notin M$ for some maximal ideal M then $ab = 1 \bmod M$ for some b, so $1 - ab \in M$, so $1 + a(-b)$ is not a unit. ■

2.4.11. A ring R is said to be *local* if it has a single maximal ideal M. In this case, M consists of all non-unit elements of R.

2.4.12. Example. If F is a field, then the ring $F[[x]]$ of formal power series with coefficients from F is local with maximal ideal (x).

2.4.13. If P is a prime ideal in R, then the complement $D = R \setminus P$ of P is a multiplicatively closed subset of R. Hence, the ring of fractions $D^{-1}R$ can be constructed. All elements of $D^{-1}R$ outside of the ideal $D^{-1}P$ are invertible, which means that $D^{-1}R$ is a local ring. This procedure of converting a prime ideal into a maximal one is called *the localization of R with respect to P*.

2.5. Radical and primary ideals

2.5.1. For an ideal I of R, *the radical* of I is the ideal $\{a \in R : a^n \in I \text{ for some } n \in \mathbb{N}\}$; it is denoted by $\sqrt{I}$ or $\mathrm{rad}(I)$.

Notice that $\mathrm{rad}(0) = \mathrm{Nil}(R)$ and for any ideal I, $\mathrm{rad}(I)$ is the preimage in R of $\mathrm{Nil}(R/I)$.

2.5.2. An ideal I of R is said to be *radical* if $\mathrm{rad}(I) = I$.

For any ideal I, $\mathrm{rad}(I)$ is a radical ideal. Any prime ideal is radical. The ideals $\mathrm{Nil}(R)$ and $\mathrm{Jac}(R)$ are also examples of radical ideals.

2.5.3. A proper ideal $Q \subseteq R$ is said to be *primary* if R/Q has no zero divisors but nilpotent elements. This means that for any $a, b \in R$ with $a, b \neq Q$ and $ab \in Q$ one has $a^n, b^m \in Q$ for some $n, m \in \mathbb{N}$, or, equivalently, for any $a, b \in R$ with $b \neq Q$ and $ab \in Q$ one has $a^n \in Q$ for some $n \in \mathbb{N}$.

2.5.4. If Q is a primary ideal then its radical $\mathrm{rad}(Q)$ is a prime ideal.

2.5.5. Example. In $\mathbb{Z}$, an ideal (n), with $n \geq 2$, is primary iff $n = p^r$ for some prime $p \in \mathbb{N}$ and $r \in \mathbb{N}$. In this case, $\mathrm{rad}(n) = (p)$.

2.6. The primary decomposition theorem for Noetherian rings

2.6.1. A commutative unital ring is said to be *Noetherian* if every ideal in R is finitely generated.

2.6.2. Examples. $\mathbb{Z}$ is a Noetherian ring. Any field F is a Noetherian ring. The rings $\mathbb{Z}[x]$, $\mathbb{Z}[x, y]$, $\mathbb{Q}[x, y, z]$ are Noetherian (see 2.6.5 and 2.6.6 below). The ring $R[x_1, x_2, \ldots]$ of polynomials in infinitely many variables is not Noetherian. The ring $C([0, 1])$ of continuous functions $[0, 1] \longrightarrow \mathbb{R}$ is not Noetherian.

2.6.3. We say that a ring R *satsfies the ACC (Ascending Chain Condition) for ideals* if any nested sequence $I_1 \subseteq I_2 \subseteq \cdots$ of ideals in R stabilizes: there exists n such that $I_n = I_{n+1} = \cdots$.

Proposition. *A ring R is Noetherian iff it satisfies the ACC for ideals; equivalently, iff any set $\mathcal{I}$ of ideals in R has a maximal element J (in the sense that ideal J is not contained in any other ideal from $\mathcal{I}$).*

Proof. Let R be Noetherian ring and let $I_1 \subseteq I_2 \subseteq \cdots$ be a chain of ideals in R. Let $J = \bigcup_{i=1}^{\infty} I_i$. Then J is a finitely generated ideal in R, $J = (a_1, \ldots, a_k)$. For each j, let $a_j \in I_{i_j}$, and let n be the maximum of $i_1, \ldots, i_k$; then $a_1, \ldots, a_k \in I_n$, so $J = I_n$, and so $I_n = I_{n+1} = \cdots$.

Conversely, if an ideal I in R is not fintely generated, then we can construct a sequence of elements $a_1, a_2, \ldots$ in I such that for every i, $a_{i+1} \notin I_n = (a_1, \ldots, a_n)$. Then the sequence $I_1 \subset I_2 \subset \cdots$ is strictly increasing and never stabilizes.

Now, if every set of ideals of R has a maximal element, then so does every nested sequence $I_1 \subseteq I_2 \subseteq \cdots$ of ideals in R, and thus it must stabilize. And conversely, if R is Noetherian and $\mathcal{I}$ is a set of ideals in R, then any increasing sequence of elements of $\mathcal{I}$ stabilizes, which provides us with a maximal element of $\mathcal{I}$. ■

2.6.4. A subring of a Noetherian ring does not have to be Noetherian. Any quotient ring of a Noetherian ring is Noeitherian. If $R = R_1 \times R_2$, then R is Noetherian iff both R_1 and R_2 are Noetherian.

2.6.5. We also have:

Theorem. *If R is a Noetherian ring, then the polynomial ring $R[x]$ is also Noetherian.*

Proof. Let I be an ideal in $R[x]$. For each i, let J_i be the set of senior coefficients of elements of I of degree $\leq i$; then J_i are ideals in R with $J_1 \subseteq J_2 \subseteq \ldots$. Since R is Noetherian, there is d such that $J_i = J_d$ for all $i \geq d$. For every $i = 1, \ldots, d$ let $\{a_{i,1}, \ldots, a_{i,k_i}\}$ be a set of generators of J_i and let $f_{i,1}, \ldots, f_{i,k_i} \in I$ be such that for every $j = 1, \ldots, k_i$, $a_{i,j}$ is the senior coefficient of $f_{i,j}$. We claim that the set $A = \{f_{i,j},\ i = 1, \ldots, d,\ j = 1, \ldots, k_i\}$ generates I, $I = (A)$. Indeed, let $f \in I$, $\deg f = n$, and let a be the senior coefficient of f. Then $a \in J_i$ where $i = n$ if $n < d$ and $i = d$ if $n \geq d$. Thus there are $c_1, \ldots, c_{k_i} \in R$ such that $a = \sum_{j=1}^{k_i} c_j a_{i,j}$. Now, the polynomial $f - \sum_{j=1}^{k_i} c_j x^{n - \deg f_{i,j}} f_{i,j}$ has degree $< n$ and belongs to I, so by induction on n, it is contained in (A). Hence, $f \in (A)$ as well. ■

2.6.6. Since $R[x, y] = (R[x])[y]$, it follows that if R is Noetherian, then the polynomial ring $R[x, y]$ is also Noetherian, and the polynomial ring $R[x_1, \ldots, x_k]$ is Noetherian for any k. In particular, the rings $\mathbb{Z}[x_1, \ldots, x_k]$ and $F[x_1, \ldots, x_k]$ where F is a field are Noetherian for all k.

2.6.7. If R is a Noetherian ring and a ring S is finitely generated over R, $S = R[\alpha_1, \ldots, \alpha_k]$, then S is isomorphic to a quotient ring of $R[x_1, \ldots, x_k]$: a surjective homomorphism $R[x_1, \ldots, x_k] \longrightarrow S$ is defined by sending $a \mapsto a$ for all $a \in R$ and $x_i \mapsto \alpha_i$ for all i. Hence, such a ring S is also Noetherian.

2.6.8. For Noetherian rings the following generalization of the Fundamental Theorem of Arithmetics holds:

The primary decomposition theorem. *If R is a Noetherian ring, then any proper ideal I of R is representable as a finite intersection $I = \bigcap_{i=1}^{k} Q_i$ of primary ideals.*

In contrast with the unique factorization of integers into a product of powers of distinct primes, the primary decomposition $I = \bigcap_{i=1}^{k} Q_i$ of an ideal in R may not be unique; it can however be shown that the set $\{\mathrm{rad}(Q_i),\ i = 1, \ldots, k\}$ of prime ideals associated with the primary ideals Q_i is defined uniquely.

Proof. First, we'll show that any proper non-primary ideal I is representable as an intersection, $I = J \cap K$, of two larger ideals. After factorizing by I, we may assume that $I = 0$, and let $a, b \in R$ be such that $ab = 0$, $b \neq 0$ and $a^n \neq 0$ for all n. Since R is Noetherian, the sequence $\mathrm{Ann}(a) \subseteq \mathrm{Ann}(a^2) \subseteq \cdots$ stabilizes; let n be such that $\mathrm{Ann}(a^{n+1}) = \mathrm{Ann}(a^n)$. Put $J = (a^n)$ and $K = (b)$; let's show that $J \cap K = 0$. Indeed, if $c \in J \cap K$, then $ca = 0$ and $c = da^n$, so $da^{n+1} = ca = 0$, so $d \in \mathrm{Ann}(a^{n+1})$, so $d \in \mathrm{Ann}(a^n)$, so $c = da^n = 0$.

Now, assume that the set of proper ideals of R not representable as an intersection of finitely many primary ideals is nonempty. Since R is Noetherian, this set has a maximal element I. I is not primary, so $I = J \cap K$ for some $J, K \supsetneq I$; by assumption, J and K are representable as intersections of finitely many primary ideals, thus so is I, contradiction. ■

---

3. An introduction to algebraic geometry

Let R be a commutative unital ring. In *algebraic geometry*, a topological space X is constructed for which R is a ring of continuous functions, so that algebraic objects related to R and their properties gain a geometric interpretation. (This space X, called *the spectrum* of R, has strange, but nice properties.)

3.0.1. If R is indeed a ring of continuous functions on a space X, then every point $x \in X$ defines a maximal ideal $M_x = \{a \in R : a(x) = 0\}$ in R. We therefore define X to be the set of prime ideals of R; this set is called *the spectrum* of R and is denoted by $\mathrm{Spec}(R)$. (Dealing with prime ideals rather than with the maximal ones turns out to be more convenient, though the space $X = \mathrm{Spec}(R)$ obtained this way is weird.) Now, for $x = P \in X = \mathrm{Spec}(R)$, "$a \in P$" is interpeted as $a(x) = 0$.

3.0.2. For every $x = P \in X$, we have *the evaluation homomorphism* at x, which maps every "function" $a \in R$ to its value at x, and which is just the factorization homomorphism $R \longrightarrow R/P$. Hence, "the value of the function" $a \in R$ at a point $x = P$ is the image of a in the integral domain R/P.

3.0.3. For "a continuous function" $a \in R$, the set $V_a = \{x \in X : a(x) = 0\}$ of zeroes of a must be a closed subset of X. We have $V_a = \{P \in \mathrm{Spec}(R) : a \in P\} = \{P \in \mathrm{Spec}(R) : (a) \subseteq P\}$. More generally, the set of common zeroes of any family $A \subseteq R$ of continuous functions, $V_A = \{P \in \mathrm{Spec}(R) : F \subseteq P\} = \{P \in \mathrm{Spec}(R) : (F) \subseteq P\}$, must be a closed subset of X.

We use this to define topology on $X = \mathrm{Spec}(R)$: the subsets of X of the form $V_I = \{P \in \mathrm{Spec}(R) : I \subseteq P\}$, where I are ideals in R, are declared to be closed, and their complements in X to be open. The topology on $\mathrm{Spec}(R)$ introduced this way is called *the Zariski topology*.

The Zariski topology is not usually Hausdorff. Moreover, the points of X corresponding to prime non-maximal ideals are not even closed in X, – the closure of a point $P \in \mathrm{Spec}(R)$ is the set V_P.

3.0.4. Examples. (i) $\mathrm{Spec}(\mathbb{Z})$ is countable, it has (a non-closed) point 0 and (closed) points (p) for prime $p \in \mathbb{N}$. The topology on $\mathrm{Spec}(\mathbb{Z})$ is "cofinite" – a set is open iff its complement is a finite set.

(ii) If F is an algebraically closed field (like $\mathbb{C}$), then in the ring $F[x]$ all nonzero prime ideals are maximal and have form $(x - a)$, $a \in F$. $\mathrm{Spec}(F[x])$ is therefore identified, $(a) \leftrightarrow a$, with "the line" F (with one additional "common" point 0 whose closure is the whole line). The Zariski topology on this line is also cofinite.

(iii) $\mathrm{Spec}(\mathbb{R}[x])$, in addition to "real" points $(x - a)$, $a \in \mathbb{R}$, has "complex" points, corresponding to the ideals $(x^2 + ax + b)$ with $a^2 < 4b$. (These points can be seen as pairs $(z, \overline{z})$ of non-real complex-conjugated points in $\mathbb{C}$.)

(iv) If F is algebraically closed, $\mathrm{Spec}(F[x, y])$ is "the plane" F^2 (in which a point (a, b) corresponds to the maximal ideal $(x - a, y - b)$), and, additionally, has the "common" 0 point and many non-closed points that correspond to ideals (f) generated by irreducuble polynomials $f(x, y)$.

For any n, $\mathrm{Spec}(F[x_1, \ldots, x_n])$ is the n-dimensional F-vector space F^n, equipped with the Zariski toplogy: in this topology, a set is closed if it is the zero set of a system of polynomials.

3.0.5. For two "functions" $a, b \in R$, $V_{(a,b)}$ is the set of points where both a and b vanish, that is, $V_{(a,b)} = V_a \cap V_b$. Conversely, on the set $V_a \cup V_b$ any "function" divisible by both a and b vanishes, and we have $V_a \cup V_b = V_{(a)\cap(b)}$.

More generally, for any collection I_α, $\alpha \in \Lambda$, of ideals in R we have $\bigcap_{\alpha\in\Lambda} V_{I_\alpha} = V_{\Sigma_{\alpha\in\Lambda} I_\alpha}$, and for two ideals $I, J \subseteq R$ we have $V_I \cup V_J = V_{I\cap J}$.

3.0.6. Let I be an ideal in R and let $Y = V_I$ be the corresponding closed subset of X. By factorizing R by I we identidy "functions" that agree on Y; this way we obtain a ring of functions on Y. And indeed, $\mathrm{Spec}(R/I) = V_I$.

3.0.7. For any ideal I, $V_{\mathrm{rad}(I)} = V_I$. In particular, $V_{\mathrm{Nil}(R)} = V_0 = X$: the nilpotent elements of R represent "the infenitisimal functions", which vanish everywhere on X.

3.0.8. Let a and b be non-nilpotent zero divisors in R so that $ab = 0$. Then $V_a, V_b \neq X$ and $V_a \cup V_b = V_{(a)\cap(b)} = V_{(ab)} = X$, so that X is representable as a union of two proper closed subsets. Such a topological space is said to be *reducible*, and we see that the existence of zero divisors in R means the reducibility of $\mathrm{Spec}(R)$.

A closed subset $Y = V_I$ of X is therefore irreducible iff R/I has no zero divisors, that is, if I is a prime ideal.

Examples. In the ring $F[x, y]$, where F is a field, for $I = (xy)$ the set V_I is the union of two lines, $\{x = 0\}$ and $\{y = 0\}$. For $J = (y - x^2)$, V_J is the (irreducible) parabola $\{y = x^2\}$.

3.0.9. In the case R is a direct product, $R = I \times J = I \oplus J$, of two its ideals, we have $V_I \cup V_J = V_{I\cap J} = V_0 = X$ and $V_I \cap V_J = V_{I+J} = V_{(1)} = \emptyset$, that is, X is a disjoint union of two its proper closed subsets.

3.0.10. *The dimension* $\dim X = \dim R$ of the space $X = \mathrm{Spec}(R)$ and of the ring R itself, is defined as the maximal length -1 of a nested sequence $Y_0 \subset Y_1 \subset Y_2 \subset \cdots \subset Y_n$ of nonempty irreducible closed subsets of X (a point $\subset$ a curve $\subset$ a surface $\subset \cdots \subset$ an n-dimensional space), or equivalently, as the maximal length -1 of a nested sequence $P_0 \supset P_1 \supset \cdots \supset P_n$ of prime ideals of R.

Examples. (i) For any field F and $n \in \mathbb{N}$, $\dim\big(F[x_1, \ldots, x_n]\big) = n$, due to the sequence $(x_1, \ldots, x_n) \supset (x_2, \ldots, x_n) \supset \cdots \supset (x_n) \supset (0)$ of prime ideals, and the induced sequence $\{0\} \subset F \times \{0\}^{n-1} \subset \cdots \subset F^{n-1} \times \{0\} \subset F^n$ of irreducible closed subsets of the spectrum of this ring.

(ii) $\dim(\mathbb{Z}) = 1$ and $\dim(\mathbb{Z}[x]) = 2$.

3.0.11. If $\eta\colon X \longrightarrow Y$ is a continuous mapping of topological spaces, then a natural "reverse" ring homomorphism φ is defined from the ring of continuous functions on Y to the ring of continuous functions on X, by $\varphi(g) = g \circ \eta$. Given a point $x \in X$, we have $\varphi(g)(x) = 0$ iff $g(\eta(x)) = 0$, so that the preimage $\varphi^{-1}(M_x)$ under φ of the ideal $M_x = \{f : f(x) = 0\}$ of functions on X vanishing at x is the ideal $M_x = \{g : g(\eta(x)) = 0\}$ of functions on Y vanishing at $\eta(x)$.

This allows us, given a ring homomorphism $\varphi\colon S \longrightarrow R$ with $\varphi(1_S) = 1_R$, to define a continuous mapping $\varphi^*\colon \mathrm{Spec}(R) \longrightarrow \mathrm{Spec}(S)$ by $\varphi^*(P) = \varphi^{-1}(P)$, $P \in \mathrm{Spec}(R)$.

If φ is surjective, then φ^* is injective; if φ is injective, then $\varphi^*(\mathrm{Spec}(R))$ is a dense subset of $\mathrm{Spec}(S)$.

Example. The homomorphism $\varphi\colon \mathbb{Z} \longrightarrow R$, $n \mapsto n1_R$, induces a continuous mapping $\varphi^*\colon \mathrm{Spec}(R) \longrightarrow \mathrm{Spec}(\mathbb{Z}) = \{0, (2), (3), (5), \ldots\}$: for a prime ideal $P \in \mathrm{Spec}(R)$, $\varphi^*(P) = \varphi^{-1}(P)$, which is either 0 or (p) for some prime $p \in \mathbb{Z}$. We have $\varphi^*(P) = 0$ if $\varphi(Z) \cap P = 0$, that is, if the homomorphism $\mathbb{Z} \longrightarrow R/P$ is injective and R/P contains a copy of $\mathbb{Z}$; and if $\varphi(p) = p1_R \in P$ for a prime $p \in \mathbb{Z}$, then $\varphi^*(P) = (p)$ and we have an injective homomorphism $\mathbb{Z}_p \longrightarrow R/P$.

---

4. Unique factorization, principal ideal, and Euclidean domains

In this section R is an ID (an integral domain – a commutative unital ring without zero divisors).

4.1. Divisibility of elements, prime and irreducible elements

4.1.1. For $a, b \in R$ we say that *a divides b*, or *a is a divisor of b*, or *b is divisible by a*, or *b is a multiple of a*, and write $a \mid b$ if there is $c \in R$ such that $b = ac$.

4.1.2. For $a, b \in R$ we have $a \mid b$ iff $b \in (a)$ iff $(b) \subseteq (a)$.

4.1.3. Two elements $a, b \in R$ are said to be *associate* if $b = ua$ for some unit $u \in R$. Being associate is an equivalence relation.

Two elements $a, b \in R$ are associate iff $(a) = (b)$.

4.1.4. *The greater common divisor* of $a, b \in R$, or $\gcd(a, b)$, is an element $d \in R$ such that $d \mid a, b$ and if $c \mid a, b$ then $c \mid d$.

$\gcd(a, b)$ may not exist; if it exists, it is defined uniquely up to associates. If the ideal (a, b) is principal, $(a, b) = (d)$ for some $d \in R$, then $\gcd(a, b)$ exists and equals d; in this case $d = xa + yb$ for some $x, y \in R$. It may however happen that $\gcd(a, b)$ exists without (a, b) being principal: in the ring $F[x, y]$, $\gcd(x, y) = 1$ but $1 \notin (x, y)$.

4.1.5. *The least common multiple* of $a, b \in R$, or $\mathrm{lcm}(a, b)$, is an element $l \in R$ such that $a, b \mid l$ and if $a, b \mid c$ then $l \mid c$.

$\mathrm{lcm}(a, b)$ may not exist; if it exists, it is defined uniquely up to associates. If the ideal $(a) \cap (b)$ is principal, $(a) \cap (b) = (l)$ for some $l \in R$, then $\mathrm{lcm}(a, b)$ exists and equals l. Conversely, if $l = \mathrm{lcm}(a, b)$ exists, then $(a) \cap (b) = (l)$. Indeed, since $a, b \mid l$ we have $(l) \subseteq (a) \cap (b)$, and for any $c \in (a) \cap (b)$ we have $a, b \mid c$, so $l \mid c$, so $c \in (l)$.

4.1.6. A nonzero nonunit element $p \in R$ is said to be *prime* if $p \mid (ab)$, $a, b, \in R$, implies that $p \mid a$ or $p \mid b$.

p is a prime element of R iff (p) is a nonzero prime ideal in R.

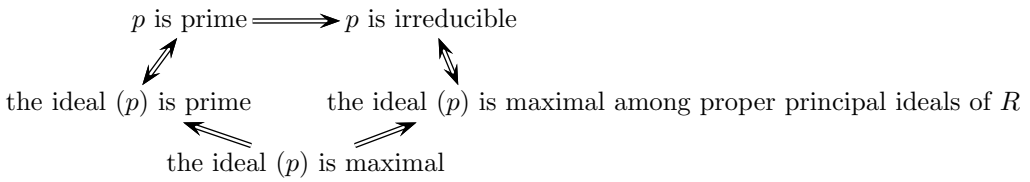
4.1.7. A nonzero nonunit element $p \in R$ is said to be *irreducible* if $a \mid p$, $a \in R$, implies that a is a unit or a is an associate of a. Non-irreducible elements are said to be *reducible*.

p is an irreducible element iff the ideal (p) is maximal in the set of nonzero principal proper ideals of R.

4.1.8. Lemma. *If p is a prime element of R, then p is irreducible in R.*

Proof. Let $p = ab$ for some $a, b \in R$. Since p is prime, $p \mid a$ or $p \mid b$; w.l.o.g. let $p \mid a$. Since both $a \mid p$ and $p \mid a$, we have that a and p are associate and b is a unit. Hence, p is irreducible. ■

4.1.9. For a nonzero nonunit element $p \in R$ we have the following diagram of implications:



4.1.10. Examples. (i) In $\mathbb{Z}$, all statements in diagram 4.1.9 are equivalent.

(ii) In the ring $\mathbb{Z}[x]$, all irreducible elements are prime. However, the ideals (2) and (x) are prime but not maximal in $\mathbb{Z}[x]$.

(iii) In the ring $\mathbb{Z}[\sqrt{-5}] = \{a + b\sqrt{-5},\ a, b \in \mathbb{Z}\}$, the element 2 is irreducible but not prime: $2 \mid 6 = (1+\sqrt{-5})(1+\sqrt{-5})$ but $2 \nmid (1+\sqrt{-5}), (1+\sqrt{-5})$.

4.2. Principal ideal domains (PIDs)

4.2.1. An integral domain R is said to be *a principal ideal domain*, or *a PID*, if every ideal in R is principal.

4.2.2. In a PID all statements in the diagram 4.1.9 are equivalent, and we get:

Proposition. *Let R be a PID and let $p \in R$ be nonzero and nonunit. Then p is prime iff p is irreducible iff (p) is a prime ideal iff (p) is a maximal ideal.*

4.2.3. Also, if R is a PID, then for any $a, b \in R$ both $\gcd(a, b)$ and $\operatorname{lcm}(a, b)$ exist, and $\gcd(a, b) = xa + yb$ for some $x, y \in R$.

4.2.4. Examples. (i) $\mathbb{Z}$ is a PID.

(ii) If F is a field, the ring $F[x]$ is a PID. (This ring is a Euclidean domain, as we will see later.)

(iii) The ring $\mathbb{Z}[i]$ of Gaussian integers is a PID. (It is also Euclidean.)

(iv) The ring $\mathbb{Z}\left[\frac{1+\sqrt{-19}}{2}\right]$ is a PID (and is not Euclidean).

(v) The rings $\mathbb{Z}[x]$ and $F[x, y]$ are not PIDs.

4.3. Unique factorization domains (UFDs)

4.3.1. An integral domain R is said to be *a unique factorization domain*, or *a UFD*, if every nonzero nonunit element $a \in R$ is representable as a product of irreducible elements: $a = p_1 \cdots p_k$ where $k \geq 1$ and $p_i \in R$ are not necessarily distinct irreducibles, and this representation is unique up to associates and permutations: if $a = q_1 \cdots q_l$ is another representation of a as a product of irreducibles, then $l = k$ and there is a permutation $\sigma \in S_k$ such that for every i, p_i and $q_{\sigma(i)}$ are associate.

4.3.2. In a UFD R, every nonzero nonunit element a is uniquely, up to permutations and associates, representable in the form $a = p_1^{r_1} \cdots p_k^{r_k}$ with $k \geq 1$, $r_i \geq 1$ for all i and p_i being *distinct*, and, moreover, *nonassociate* irreducibles of R.

Sometimes it is convenient to write $a \in R$ in the form $a = up_1^{r_1} \cdots p_m^{r_m}$ where u is a unit and p_i are distinct (i.e. pairwise nonassociate) irreducibles and all $r_i \geq 0$; such a representation is unique up to permutations, associates, and terms of the form p_i^0. In particular, any two nonzero nonunit elements $a, b \in R$ are writable in the form $a = up_1^{r_1} \cdots p_k^{r_k}$ and $b = vp_1^{s_1} \cdots p_k^{s_k}$ where u and v are units, $p_1, \ldots, p_k$ are distinct irreducibles of R, and $r_i, s_i \geq 0$ for all i.

4.3.3. Proposition. *If R is a UFD, then for nonzero nonunit elements $a, b \in R$ with $a = up_1^{r_1} \cdots p_k^{r_k}$ and $b = vp_1^{s_1} \cdots p_k^{s_k}$ where i, v are units, $p_1, \ldots, p_k$ are distinct irreducibles of R and $r_i, s_i \geq 0$ for all i, we have $a \mid b$ iff $s_i \geq r_i$ for all i.*

Proof. Iff $s_i \geq r_i$ for all i, then $b = ac$ for $c = vu^{-1}p_1^{s_1 - r_1} \cdots p_k^{s_k - r_k}$. Conversely, if $b = ac$ with $c = wp_1^{t_1} \cdots p_k^{t_k}$ (where some of t_i can be equal to 0), then $b = uwp_1^{r_1 + t_1} \cdots p_k^{r_k + t_k}$ is the unique factorization of b, so $s_i = r_i + t_i \geq r_i$ for all i. ■

4.3.4. As immediate corollaries we obtain that

Proposition. *If R is a UFD, then $p \in R$ is irreducible iff p is prime in R.*

4.3.5. And

Proposition. *If R is a UFD, then for nonzero nonunit $a, b \in R$, $a = up_1^{r_1} \cdots p_k^{r_k}$ and $b = vp_1^{s_1} \cdots p_k^{s_k}$ where u, v are units, $p_1, \dots, p_k$ are distinct irreducibles of R and $r_i, s_i \geq 0$ for all i, both the $\gcd(a,b)$ and $\operatorname{lcm}(a,b)$ exist in R, and we have $\gcd(a,b) = p_1^{\min(r_1,s_1)} \cdots p_k^{\min(r_k,s_k)}$ and $\operatorname{lcm}(a,b) = p_1^{\max(r_1,s_1)} \cdots p_k^{\max(r_k,s_k)}$.*

4.3.6. As a corollary of Proposition 4.3.5 we get that in a UFD R, $ab = \gcd(a,b)\operatorname{lcm}(a,b)$ (up to associates) for any $a, b \in R$.

4.3.7. Though $d = \gcd(a,b)$ of two elements a, b of a UFD exists, it may be that $d \notin (a,b)$: in the ring $F[x,y]$ where F is a field (which is a UFD, as we will see later), $\gcd(x,y) = 1$, but $1 \notin (x,y)$. For $l = \operatorname{lcm}(a,b)$, however, it is true that $(l) = (a) \cap (b)$.

4.3.8. Theorem. *Every PID is a UFD.*

Proof. Every PID is a Noetherian ring, and we could use the primary decomposition theorem; but the proof is easier for PIDs. Let R be a PID, and assume that the set $\mathcal{A}$ of nonzero nonunit elements of R not representable as products of irreducibles is nonempty. Since R is Noetherian, the set $\{(a) : a \in \mathcal{A}\}$ has a maximal element, an ideal (a). Then a is reducible, $a = bc$, where none of b, c is a unit. Then the ideals (b) and (c) are strictly larger than (a), so $b, c \notin \mathcal{A}$, so b and c are representable as products of irreducibles, so a is also representable as a product of irreducibles, contradiction.

As for uniqueness, assume that $p_1 \cdots p_k = q_1 \cdots q_l$ where p_i and q_j are irreducible, and so prime, elements of R. Then $p_k \mid (q_1 \cdots q_l)$, so p_k divides one of $q_1, \dots, q_l$; w.l.o.g. assume that $p_k \mid q_l$. Since q_l is irreducible, $q_l = up_k$ for some unit u, so q_l is an associate of p_k, and $p_1 \cdots p_{k-1} = \tilde{q}_1 q_2 \cdots q_{l-1}$ where $\tilde{q}_1 = uq_1$ is associate to q_1. By induction on k, $k - 1 = l - 1$ and $\tilde{q}_1, q_2, \dots, q_{l-1}$ are associates, in certain order, of $p_1, p_2, \dots, p_{k-1}$. ■

4.3.9. $\mathbb{Z}$ is a PID. We will see later that the rings $F[x]$ where F is a field, and $\mathbb{Z}[i]$ are PIDs. So, all these rings are UFDs.

We will also learn that if R is a UFD, then the polynomial ring $R[x]$ is a UFD too; this implies that the rings $\mathbb{Z}[x]$, $\mathbb{Z}[x_1, \dots, x_n]$ for any n, and $F[x_1, \dots, x_n]$ for any n are UFDs as well. Notice that these rings are not PIDs.

4.4. Euclidean domains (EDs)

4.4.1. A function $N\colon R \longrightarrow \mathbb{Z}$ is called *a Euclidean norm* on R if $N(0) = 0$, $N(a) \geq 0$ for all $a \in R$, and for any $a, b \in R$ with $b \neq 0$ there are $c, r \in R$ such that $a = bc + r$ and either $r = 0$ (that is, $b \mid a$), or $N(r) < N(b)$.

An integral domain R is said to be *Euclidean* if a Euclidean norm can be defined on it.

4.4.2. Examples. (i) $\mathbb{Z}$ is an ED with the norm $N(a) = |a|$.

(ii) If F is a field, then the ring $F[x]$ is an ED with $N(f) = \deg f$: for any nonzero polynomials $f, g \in F[x]$, $f = gh + r$ for some $h, r \in F[x]$ where either $r = 0$ or $\deg r < \deg g$.

(iii) $\mathbb{Z}[i]$ is an ED with the norm $N(\alpha) = |\alpha|^2$. Indeed, let $\alpha, \beta \in \mathbb{Z}[i]$, $\beta \neq 0$. Let γ be the elemenet of $\mathbb{Z}[i]$ with the minimal distance to α/β, then $|\alpha/\beta - \gamma| \leq \sqrt{2}/2 < 1$. So, $|\alpha - \gamma\beta| < |\beta|$, and for $\delta = \alpha - \gamma\beta$ we have $\alpha = \gamma\beta + \delta$, $N(\delta) = |\delta|^2 < N(\beta) = |\beta|^2$.

4.4.3. Theorem. *Every ED is a PID.*

Proof. Let R be an ED with Euclidean norm N and let I be an ideal in R. Let b be an element of I with the minimal norm. Then for any $a \in I$, if $b \nmid a$, then $a = bc + r$ for some $r = a - bc \in I$ with $N(r) < N(b)$, which is impossible. So, $I = (b)$. ■

4.4.4. We therefore have the following sequence of inclusions: ED $\underset{\neq}{\subset}$ PID $\underset{\neq}{\subset}$ UFD $\underset{\neq}{\subset}$ ID. The following example demonstrate that all these inclusions are, indeed, strict: the ring $\mathbb{Z}\left[\frac{1+\sqrt{-19}}{2}\right]$ is a PID but not an ED; the ring $Z[x]$ is a UFD but not a PID; the ring $\mathbb{Z}[\sqrt{-5}]$ is an ID but not a UFD.

4.4.5. If R is an ED with Euclidean norm N, then for $a, b \in R$ with $b \neq 0$, $\gcd(a, b)$ can be found using *the Euclidean algorithm*: Find $c_0, r_1, c_1, r_2, \ldots, c_k, r_k$ such that

$$a = bc_0 + r_1,\ b = c_1r_1 + r_2,\ r_1 = c_2r_2 + r_3,\ \ldots,\ r_{k-2} = c_{k-1}r_{k-1} + r_k,\ r_{k-1} = c_kr_k$$

with $N(r_1) < N(b)$, $N(r_2) < N(r_3)$, ..., then $r_k \in (a, b)$ and $a, b \in (r_k)$, so $(a, b) = (r_k)$, and $r_k = \gcd(a, b)$.

4.4.6. Let R be an ED with norm N. Then the nonzero elements of R of minimal norm are units in R. Indeed, for such an element u, either $u \mid 1$ or $1 = cu + r$ with $N(r) < N(u)$, which is impossible.

Let v be an element of the minimal norm in $R \setminus \{0, \text{units}\}$. Then v has the property that for any $a \in R$, v divides a or $a - u$ for some unit $u \in R$. (If $v \nmid a$, then $a = cv + u$ with $N(u) < N(v)$, so u is a unit.) Nonzero nonunit elements with this property are called *universal side divisors*, and we see that any Euclidean domain that is not a field must contain a universal side divisor.

4.4.7. Lack of universal side divisors in an integral domain proves that it is non-Euclidean. An example is the quadratic integer ring $\mathcal{O} = \mathbb{Z}\left[\frac{1+\sqrt{-19}}{2}\right]$. The only units in $\mathcal{O}$ are ± 1; if u is a universal side divisor in $\mathcal{O}$, then u must divide 2 or 2 ± 1, and it follows that $u = \pm 2, \pm 3$. But 2 and 3 don't divide $\omega = \frac{1+\sqrt{-19}}{2}$ or $\omega \pm 1$ in $\mathcal{O}$, so $\mathcal{O}$ contains no universal side divisors, and thus is not Euclidean.

4.5. PIDs and the Dedekind-Hasse norm

4.5.1. *The Dedekind-Hasse norm* on an integral domain R is a function $N\colon R \longrightarrow \mathbb{Z}$ such that $N(0) = 0$, $N(a) \geq 0$ for all $a \in R$, and for any $a, b \in R$ with $b \neq 0$ either $(a, b) = (b)$ (that is, $b \mid a$), or there exists a nonzero $r \in (a, b)$ with $N(r) < N(b)$.

4.5.2. Existence of the Dedekind-Hasse norm is a criterion of being a PID:

Theorem. *An intergal domain R has a Dedekind-Hasse norm iff R is a PID. Moreover, in this case there is a Dedekind-Hasse norm N on R that is multiplicative: $N(ab) = N(a)N(b)$ for all $a, b \in R$.*

Proof. Let N be a Dedekind-Hasse norm on R, and let I be an ideal in R. Let b be the element of I of minimal norm $N(b)$ in I. Let a be any element of I; if $b \nmid a$, then there is $r \in (a, b) \subseteq I$ with $N(r) < N(b)$, which is impossible; hence, $b \mid a$, so $a \in (b)$. Hence, $I = (b)$.

Now, assume that R is a PID. Then R is a UFD. Introduce the following norm on R: put $N(0) = 0$, $N(u) = 1$ is u is a unit, and for nonunit nonzero $a \in R$ that is a product of k irreducibles, put $N(a) = 2^k$. Then N is multiplicative. Next, for any $a, b \in R$ with $b \neq 0$ and $b \nmid a$, we have $(a, b) = (r)$ for some $r \in R$, and since $b \in (r)$ but $(b) \neq (r)$ we get that $r \mid b$ but $b \nmid r$, so $N(r) < N(b)$. Hence, N is a multiplicative Dedekind-Hasse norm. ■

5. Quadratic integer rings

Let $D \neq 0, 1$ be a square-free integer.

5.1. Quadratic fields and the field norm

5.1.1. *The quadratic field* associated with D is the subfield $\mathbb{Q}[\sqrt{D}] = \left\{a + b\sqrt{D},\ a, b \in \mathbb{Q}\right\}$ of $\mathbb{C}$. If $D > 0$, then $\mathbb{Q}[\sqrt{D}]$ is *real*, that is, $\subseteq \mathbb{R}$; if $D < 0$ it is *non-real*.

5.1.2. The mapping $a + b\sqrt{D} \mapsto a - b\sqrt{D}$ is an automorphism of $\mathbb{Q}[\sqrt{D}]$; the element $\overline{\alpha} = a - b\sqrt{D}$ is called *the conjugate* of the element $\alpha = a + b\sqrt{D}$.

In the case $D < 0$, $\overline{\alpha}$ is the complex conjugate of $\alpha \in \mathbb{Q}[\sqrt{D}]$.

5.1.3. *The field norm*, or just *the norm* on $\mathbb{Q}[\sqrt{D}]$ is the function $N\colon \mathbb{Q}[\sqrt{D}] \longrightarrow \mathbb{Q}$ defined by $N(\alpha) = \alpha\overline{\alpha}$, $\alpha \in \mathbb{Q}[\sqrt{D}]$. For $\alpha = a + b\sqrt{D}$ we have $N(\alpha) = a^2 - b^2D$.

We have $N(\alpha) = 0$ iff $\alpha = 0$. In the case $D < 0$, $N(\alpha) = |\alpha|^2$, where $|\alpha|$ is the absolute value of $\alpha \in \mathbb{C}$, and is positive for all nonzero $\alpha \in \mathbb{Q}[\sqrt{D}]$; if $D > 0$, then N takes negative values as well.

5.1.4. The norm N is multiplicative: for any $\alpha, \beta \in \mathbb{Q}[\sqrt{D}]$,

$$N(\alpha\beta) = (\alpha\beta)\overline{(\alpha\beta)} = \alpha\overline{\alpha}\beta\overline{\beta} = N(\alpha)N(\beta).$$

5.2. The ring of integers in a quadratic field

5.2.1. *The ring of integers* $\mathcal{O} = \mathcal{O}_{\mathbb{Q}[\sqrt{D}]}$ in $\mathbb{Q}[\sqrt{D}]$, or *the quadratic integer ring* (associated with D), is the subring of $\mathbb{Q}[\sqrt{D}]$ of elements α satisfying a monic quadratic equation with integer coefficients: $\alpha^2+c\alpha+d=0$ for some $c,d\in\mathbb{Z}$. In the case $D\neq 1 \bmod 4$, we have $\mathcal{O}=\mathbb{Z}[\sqrt{D}]=\left\{a+b\sqrt{D},\ a,b\in\mathbb{Z}\right\}$; in the case $D=1 \bmod 4$, $\mathcal{O}=\left\{\frac{a+b\sqrt{D}}{2},\ a,b\in\mathbb{Z},\ a=b \bmod 2\right\}$. If we put $\omega=\sqrt{D}$ for $D\neq 1 \bmod 4$ and $\omega=\frac{1+\sqrt{D}}{2}$ for $D=1 \bmod 4$, then, in both cases, $\mathcal{O}=\mathbb{Z}[\omega]=\{a+b\omega,\ a,b\in\mathbb{Z}\}$.

Remark. Nobody forgives us to consider the ring $\mathbb{Z}[\sqrt{D}]$ in the case $D=1 \bmod 4$ though.

5.2.2. *The conjugation* $\alpha=a+b\sqrt{D}\mapsto\overline{\alpha}=a-b\sqrt{D}$ is an automorphism of $\mathcal{O}$.

5.2.3. The field norm N takes only integer values on $\mathcal{O}$, $N(\mathcal{O})\subseteq\mathbb{Z}$. Indeed, for $\alpha=a+b\omega$ with $a,b\in\mathbb{Z}$ we have $N(\alpha)=\alpha\overline{\alpha}=a^2+b^2\omega\overline{\omega}$ and $\omega\overline{\omega}=D\in\mathbb{Z}$ if $\omega=\sqrt{D}$ and $\omega\overline{\omega}=\frac{1-D}{4}\in\mathbb{Z}$ if $D=1 \bmod 4$.

It follows from multiplicativity of N that for $\alpha,\beta\in\mathcal{O}$, if $\alpha\ \big|\ \beta$ then $N(\alpha)\ \big|\ N(\beta)$. (This fact is very useful for determining whether one element of $\mathcal{O}$ is divisible by another.)

5.2.4. An element $u\in\mathcal{O}$ is a unit in $\mathcal{O}$ iff $N(u)=\pm1$. Indeed, if $uv=1$ for some v, then $N(u)N(v)=1$ so $N(u)=\pm1$. Conversely, if $N(u)=\pm1$, then $u\overline{u}=\pm1$, so $\pm\overline{u}=u^{-1}$.

For $D<0$, $D\neq-1,-3$, the ring $\mathcal{O}$ has only two units, ±1. The ring of Gaussian integers $\mathcal{O}_{\mathbb{Q}[\sqrt{-1}]}=\mathbb{Z}[i]$ has four units, $\pm1,\pm i$. The ring $\mathcal{O}_{\mathbb{Q}[\sqrt{-3}]}$ has six units, ±1, $\frac{\pm1\pm\sqrt{-3}}{2}$.

For $D>0$, the group of positive units in the ring $\mathcal{O}$ is infinite cyclic: there is a unit $u>1$ (called *the fundamental unit* of $\mathcal{O}$) such that all positive units have form u^n, $n\in\mathbb{Z}$, and the group of all units in $\mathcal{O}$ is $\{\pm u^n,\ n\in\mathbb{Z}\}$. (For example, the fundamental unit in $\mathcal{O}_{\mathbb{Q}[\sqrt{2}]}$ is $1+\sqrt{2}$.)

5.2.5. As an example, consider the ring $\mathcal{O}=\mathbb{Z}[\sqrt{-5}]$. In this ring, $6=2\cdot3=\alpha\overline{\alpha}$ where $\alpha=1+\sqrt{-5}$. We have $N(2)=4$, $N(3)=9$, $N(\alpha)=N(\overline{\alpha})=6$, and it is easy to check that all these elements are irreducible. However, these elements are not prime: 2, for example, satisfy $2\ \big|\ \alpha\overline{\alpha}$, but $2\not|\ \alpha$ and $2\not|\ \overline{\alpha}$. Thus, 6 has two different factorizations into irreducibles, and $\mathcal{O}$ is not a UFD.

$\gcd(2\alpha,6)$ and $\operatorname{lcm}(2,\alpha)$ do not exist in $\mathcal{O}$. Indeed, 2 and α divide both 2α and 6, so if $\delta=\gcd(2\alpha,6)$ or $\delta=\operatorname{lcm}(2,\alpha)$, then δ must divide both 2α and 6 and be divisible by both 2 and α: But then $N(\delta)$ must divide both $N(2\alpha)=24$ and $N(6)=36$, and be divisible by both $N(2)=4$ and $N(\alpha)=6$, so we must have $N(\delta)=12$. However, there is no $\delta\in\mathcal{O}$ with $N(\delta)=12$.

Also, $\gcd(2,\alpha)$ exists in $\mathcal{O}$, – it is equal to 1; but $1\notin(2,\alpha)$, and so $(2,\alpha)\neq(1)$.

But if we deal with ideals instead of elements, everything must be "ideal" in $\mathcal{O}$: the ideal (6) must have a unique factorization into a product of prime ideals. ($\mathcal{O}$ is a "1-dimensional" ring where nonzero prime ideals are maximal, primary ideals are products of prime ideals, and the intersection of distinct prime ideals coincide with their product.) The problem is that the ideals (2), 3, (α), and $(\overline{\alpha})$ are not prime (and so, not maximal). If we put $I_2=(2,\alpha)$, $I_3=(3,\alpha)$, and $\overline{I}_3=(3,\overline{\alpha})$, then we can check that the ideals $I_2,I_3,\overline{I}_3$ are prime (and maximal), and $(2)=I_2^2$, $(3)=I_3\overline{I}_3$, $(\alpha)=I_2I_3$, $(\overline{\alpha})=I_2\overline{I}_3$. And now, both $(6)=(2)(3)=I_2^2I_3\overline{I}_3$ and $(6)=(\alpha)(\overline{\alpha})=I_2I_3I_2\overline{I}_3$ give the same factorization of (6) into a product of maximal ideals.

5.2.6. Let us now consider $\mathcal{O}=\left\{\frac{a+b\sqrt{-19}}{2},\ a,b\in\mathbb{Z},\ a=b \bmod 1\right\}=\left\{a+b\omega,\ a,b\in\mathbb{Z}\right\}$, $\omega=\frac{1+\sqrt{-19}}{2}$. In this ring, $N(a+b\omega)=|a+b\omega|^2$ (where $|\cdot|$ is the absolute value in $\mathbb{C}$), and we can compute that it is equal to $a^2+ab+5b^2$. The only units in this ring are ±1. Notice also that for $\gamma=a+b\omega$, $N(\gamma)$ is even iff both a and b are even, so that $\gamma/2\in\mathcal{O}$.

$\mathcal{O}$ is the standard example of a PID that is not an ED. Firstly, $\mathcal{O}$ is not a ED since it has no a universal side divisors. Indeed, if $v=a+b\omega$ is a universal side divisor, then $v\ \big|\ 2$ or $v\ \big|\ 2\pm1$. Since v is not a unit, $v\ \big|\ 2$ or 3, so $N(v)\ \big|\ 4$ or 9, but it is clearly impossible if $b\neq0$, so $v=\pm2$ or $v=\pm3$. However, $2,3\not|\ \omega$ and $2,3\not|\ \omega\pm1$.

Secondly, $\mathcal{O}$ is a PID since N is a Dedekind-Hasse norm on $\mathcal{O}$. To prove this, let $\alpha,\beta\in\mathcal{O}$ with $\beta\neq0$, $\beta\not|\ \alpha$; we need to show that for some $x,y\in\mathcal{O}$ we have $x\alpha+y\beta\neq0$ and $N(x\alpha+y\beta)<N(\beta)$. Put $\gamma=\alpha/\beta$ and rewrite this as $x\gamma+y\neq0$ and $|x\gamma+y|<1$. After replacing γ by $\gamma+z$ for some $z\in\mathcal{O}$ and by $-\gamma$ if needed, we may assume that γ lies is in the parallelogram $\left(0,\frac{1}{2},\frac{1+\omega}{2},\frac{\omega}{2}\right)$, so that $0<\operatorname{Im}(\gamma)\leq\frac{\sqrt{19}}{4}$. If $\operatorname{Im}(\gamma)<\frac{\sqrt{3}}{2}$, then $|\gamma-n|<1$ for some $n\in\mathbb{Z}$. If $\frac{\sqrt{3}}{2}\leq\operatorname{Im}(\gamma)\leq\frac{\sqrt{19}}{4}$, then $0\leq\operatorname{Im}(\omega-2\gamma)\leq\frac{\sqrt{19}}{2}-\sqrt{3}<\frac{\sqrt{3}}{2}$, and so, $|\omega-2\gamma-n|<1$ for some $n\in\mathbb{Z}$. In any case, $|x\gamma+y|<1$ for some $x\in\{\pm1,\pm2\}$, $y\in\mathcal{O}$. If, for

these x and y, we have $x\gamma + y = 0$, then in the case $x = \pm 1$ we have $\gamma \in \mathcal{O}$ so $\beta \,\big|\, \alpha$; and in the case $x = \pm 2$ and $\gamma \notin \mathcal{O}$ we have $\gamma = \pm\frac{y}{2}$, so $\overline{y}\gamma = \pm\frac{|y|^2}{2} \in \frac{1}{2}\mathbb{Z} \setminus \mathbb{Z}$, so $|\overline{y}\gamma - n| = \frac{1}{2}$ for some $n \in \mathbb{Z}$.

5.2.7. It is known (but we don't prove) that:
For $D = -1, -2, -3, -7, -11$; 2, 3, 5, 6, 7, 11, 13, 17, 19, 21, 29, 33, 37, 41, 57, 73, and only these D, the ring $\mathcal{O}$ is *norm Euclidean* (that is, Euclidean with respect to its field norm).
For $D = 14$, 69 the ring $\mathcal{O}$ is Euclidean with respect to some other norm, different from its field norm.
For $D = -19, -43, -67, -163$ the ring $\mathcal{O}$ is a PID but not an ED. It is conjectured, but not proved, that $\mathcal{O}$ is a PID for infinitely many positive D.
A quadratic integer ring is a UFD iff it is a PID.

5.3. Prime ideals and elements in quadratic integer rings

Let $D \neq 0, 1$ be a square-free integer and let $\mathcal{O} = \mathcal{O}_{\mathbb{Q}[\sqrt{D}]}$. We will now describe the prime ideals (and so, the prime elements) of $\mathcal{O}$.

5.3.1. If P is a nonzero prime ideal in $\mathcal{O}$, then $P \cap \mathbb{Z}$ is a nonzero prime ideal in $\mathbb{Z}$, and so $P \cap \mathbb{Z} = (p)$ for some prime $p \in \mathbb{Z}$. Thus, the nonzero prime ideals in $\mathcal{O}$ "split" between prime integers.

Geometrically, we have a projection $\pi\colon \mathrm{Spec}(\mathcal{O}) \longrightarrow \mathrm{Spec}(\mathbb{Z})$, under which the set $\mathrm{Spec}(\mathcal{O})$ of prime ideals of $\mathcal{O}$ splits to "fibers" $V_p = \pi^{-1}(p)$ of prime ideals (p) in $\mathbb{Z}$.

5.3.2. Let p be a positive prime in $\mathbb{Z}$. The fiber $V_p = \big\{P \in \mathrm{Spec}(\mathcal{O}) : p \in P\big\}$ is the set of prime ideals in $\mathcal{O}$ containing (p), and is isomorphic to $\mathrm{Spec}(\mathcal{O}/(p))$, – the set of prime ideals in the ring $\mathcal{O}/(p)$. Since conjugation $\alpha \mapsto \overline{\alpha}$ preserves (p), it is still defined on $\mathcal{O}/(p)$.

We have $\mathcal{O} = \{a + b\omega,\ a, b \in \mathbb{Z}\}$, where $\omega = \sqrt{D}$ if $D \neq 1 \bmod 4$ and $\omega = \frac{1+\sqrt{D}}{2}$ if $D = 1 \bmod 4$, and $(p) = \{ap + bp\omega,\ a, b \in \mathbb{Z}\}$. As a group under addition, $\mathcal{O}/(p) \cong \mathbb{Z}_p^2$, but may have different multiplicative structures:

(i) If $\mathcal{O}/(p)$ is an integral domain, then it is a field, and has no nontrivial ideals. Then (p) is a prime and maximal ideal in $\mathcal{O}$, and V_p is the singleton $\{(p)\}$. In this case the element p is prime in $\mathcal{O}$, and we say that the prime p is *inert*. The field $\mathcal{O}/(p)$ has p^2 elements.

(ii) If $\mathcal{O}/(p)$ has two *non-associate* zero divisors ξ and η with $\xi\eta = 0$, then the ideals (ξ) and (η) are nonzero, maximal (since they are maximal proper subgroups of $\mathcal{O}/(p) \cong \mathbb{Z}_p^2$), and satisfy $(\xi) \cap (\eta) = 0$. So $\mathcal{O}/(p) = (\xi) \times (\eta)$, and $\mathcal{O}/(p)$ has no other nontrivial ideals. We will see in 5.3.4 that the elements ξ and η are conjugate (up to association), so that $(\eta) = \overline{(\xi)}$. The preimages P and $\overline{P}$ in $\mathcal{O}$ of (ξ) and $\overline{(\xi)}$ are prime (and maximal) ideals, with $P\overline{P} = P \cap \overline{P} = (p)$, and $V_p = \{P, \overline{P}\}$. We say that the prime p *splits* in $\mathcal{O}$ in this case. The fields $\mathcal{O}/P$ and $\mathcal{O}/\overline{P}$ have p elements and are isomorphic to $\mathbb{F}_p$.

(iii) The last case is that $\mathcal{O}/(p)$ is not an integral domain, but all zero divisors in $\mathcal{O}$ are nilpotent. (That is, (p) is a primary ideal.) In this case $\mathrm{Nil}(\mathcal{O}/(p))$ is the only nontrivial ideal of $\mathcal{O}/(p)$, its preimage $P \subseteq \mathcal{O}$ is maximal and self-conjugate: $\overline{P} = P$, and we have $V_p = \{P\}$, and $P^2 = (p)$. We then say that (p) *ramifies* in $\mathcal{O}$.

5.3.3. We see that every prime $p \in \mathbb{Z}$ either generates a maximal ideal (p) in $\mathcal{O}$ (is inert), or "splits", $(p) = P\overline{P}$, into a product of two conjugate maximal ideals, or is the square, $(p) = P^2$, of a single maximal ideal (ramifies), and every prime ideal P in $\mathcal{O}$ is associated this way with a prime integer p. (Geometrically, $\pi\colon \mathrm{Spec}(\mathcal{O}) \longrightarrow \mathrm{Spec}(\mathbb{Z})$ is a 2-to-1 covering, with exception to inert and ramification points where it is just 1-to-1, and the conjugation acts as a transposition in the fibers.)

Examples. (i) In the ring $\mathcal{O}_{\mathbb{Q}[\sqrt{-1}]} = \mathbb{Z}[i]$, 2 ramifies: $(2) = (1+i)^2$, 3 is inert, 5 splits: $(5) = (2+i)(2-i)$.
(ii) In the ring $\mathcal{O}_{\mathbb{Q}[\sqrt{-5}]}$, 2 and 5 ramify: $(2) = (2, 1+\sqrt{-5})^2$ and $(5) = \big(\sqrt{-5}\big)\big(\sqrt{-5}\big)$, 3 splits: $(3) = (3, 2+\sqrt{-5})(3, 2-\sqrt{-5})$, and 11 is inert.

5.3.4. The ring $\mathcal{O}$ is isomorphic to $\mathbb{Z}[x]/(f)$, where f is the monic quadratic polynomial satisfied by ω: $f = x^2 - D$ if $D \neq 1 \bmod 4$ and $f = x^2 - x + \frac{1-D}{4}$ if $D = 1 \bmod 4$. Thus

$$\mathcal{O}/(p) \cong (\mathbb{Z}[x]/(f))/(p) \cong \mathbb{Z}[x]/(f, p) \cong (\mathbb{Z}[x]/(p))/(f) \cong \mathbb{F}_p[x]/(f).$$

The conjugation automorphism acts on $\mathbb{F}_p[x]/(f)$ by $x \mapsto -x$ if $D \neq 1 \bmod 4$, and $x \mapsto 1 - x$ if $D = 1 \bmod 4$.

The ring $\mathbb{F}_p[x]$ is a PID. If the polynomial f is irreducible in this ring, then $\mathcal{O}/(p)$ is a field, that is, p is inert. If f splits in $\mathbb{F}_p[x]$ to distinct factors, $f = (x-\alpha)(x-\beta)$ with $\alpha \neq \beta$, then $\mathcal{O}/(p)$ has non-associate

zero divisors, that is, p splits. (In which case conjugation transposes $x-\alpha$ and $x-\beta$.) And if f is a square, $f=(x-\alpha)^2$, in $\mathbb{Z}_p[x]$, then $\mathcal{O}/(p)$ has a nontrival nilradical, that is, p ramifies.

Examples. (i) Let $D=-1$, so that $\mathcal{O}_{\mathbb{Q}[\sqrt{-1}]}=\mathbb{Z}[i]$. The polynomial $f=x^2+1$ is irreducible in $\mathbb{Z}_3[x]$, so 3 is inert (is prime) in $\mathbb{Z}[i]$. In $\mathbb{Z}_5[x]$, f splits to distinct factors: $x^2+1=(x-2)(x-3)$, so 5 splits. And in $\mathbb{Z}_2[x]$, f is a square: $f=(x+1)^2$, so 2 ramifies.

(ii) Let $D=-5$. The polynomial x^2+5 is reducible in $\mathbb{Z}_2[x]$ ($f=(x+1)^2$), in $\mathbb{Z}_3[x]$ ($f=(x-1)(x-2)$), and in $\mathbb{Z}_7[x]$ ($f=(x-3)(x-4)$), so 2 ramifies and 3, 7 split in $\mathcal{O}_{\mathbb{Q}[\sqrt{-5}]}$. In $\mathbb{Z}_{11}[x]$, f is irreducible, so 11 is inert (is prime) in $\mathcal{O}_{\mathbb{Q}[\sqrt{-5}]}$.

5.3.5. The ideals P above can be principal or not; in the former case they are generated by prime elements of $\mathcal{O}$, and every prime element of $\mathcal{O}$ appears this way.

If a prime $p\in\mathbb{Z}$ splits or ramifies, $(p)=P\overline{P}$, and P is principal, $P=(\pi)$, then $p=u\pi\overline{\pi}$ for some unit $u\in\mathcal{O}$. Let $\pi=a+b\sqrt{D}$; then $N(p)=p^2=N(\pi)N(\overline{\pi})$, so $N(\pi)=a^2-b^2D=\pm p$. Conversely, if there are $a,b\in\mathbb{Z}$ or $\in\frac{1}{2}\mathbb{Z}$ (for $D=1 \bmod 4$) such that $\pm p=a^2-b^2D$, then $p=\pm(a+b\sqrt{D})(a-b\sqrt{D})$, so p is reducible. Hence, the prime elements of $\mathcal{O}$ are the inert prime integers p, which are not representable in the form $p=\pm(a^2-b^2D)$, and the elements of the form $\pi=a+b\sqrt{D}$, $\overline{\pi}=a-b\sqrt{D}$ such that $\pi\overline{\pi}=a^2-b^2D=\pm p$ for a non-inert prime integer p. (Notice that, if $\mathcal{O}$ is not a PID, it may happen that a prime $p\in\mathbb{Z}$ is irreducible in $\mathcal{O}$, thus isn't of the form a^2-b^2D, but not prime, as the example of $3\in\mathbb{Z}[\sqrt{-5}]$ demonstrates.)

5.4. Primes in $\mathbb{Z}[i]$ and representation of positive integers as a sum of two squares

5.4.1. If we apply the results of 5.3 to the PID $\mathbb{Z}[i]=\mathcal{O}_{\mathbb{Q}[\sqrt{-1}]}$ of Gaussian integers, we get the following. The primes=irreducibles in $\mathbb{Z}[i]$ are of two sorts: the "'inert" primes $p\in\mathbb{N}$, not represnetable in the form $p=a^2+b^2$ with $a,b\in\mathbb{Z}$, and the non-real primes $\pi=a+bi$, $\overline{\pi}=a-bi$ for which $N(\pi)=a^2+b^2=p$ is a prime in $\mathbb{N}$ that "splits" or "ramifies".

5.4.2. For a prime $p=a^2+b^2$ to ramify we must have $a+bi=u(a-bi)$ where $u=\pm1,\pm i$; this is only possible if $p=2$, for which $2=-i(1+i)^2$.

A prime $p\neq 2$ splits iff the polynomial x^2+1 splits, $x^2+1=(x-k)(x+k)$, in $\mathbb{Z}_p[x]$. This is so iff there is $k\in\mathbb{Z}_p$ such that $k^2=-1$, that is, k has order 4 in $\mathbb{Z}_p^*\cong\mathbb{Z}_{p-1}$; such an element exists iff $4\mid p-1$, that is, iff $p=1 \bmod 4$.

We therefore obtain: in $\mathbb{Z}[i]$, the prime 2 ramifies; a prime $p\geq 3$ is inert iff $p=3 \bmod 4$, and splits iff $p=1 \bmod 4$.

5.4.3. Hence, the primes in $\mathbb{Z}[i]$ are (up to associates): $1+i$; all prime integers $p=3 \bmod 4$; and two conjugate non-real numbers $\pi=a+bi$ and $\overline{\pi}=a-bi$ with $\pi\overline{\pi}=a^2+b^2=p$ for each prime integer $p=1 \bmod 4$:

$$1+i,\ 3,\ 2+i,\ 2-i,\ 7,\ 11,\ 3+2i,\ 3-2i,\ 4+i,\ 4-i,\ 19,\ldots$$

5.4.4. As a corollary we obtain a criterion of the representativity of a positive integer as a sum of two squares. Let $n=a^2+b^2$, $a,b\in\mathbb{N}$; then in $\mathbb{Z}[i]$, $n=(a+bi)(a-bi)$. Factorize $a+bi$ into irreducibles in $\mathbb{Z}[i]$,

$$a+bi=u(1+i)^k p_1^{r_1}\cdots p_l^{r_l}\pi_1^{s_1}\overline{\pi}_1^{t_1}\cdots\pi_d^{s_d}\overline{\pi}_d^{t_d} \tag{5.1}$$

where $u=\pm1,\pm i$, p_i are distinct inert integer primes, π_j are non-real primes corresponding to distinct non-inert integer primes q_j, $r_i\geq 1$, $s_j,t_j\geq 0$, and $s_j+t_j\geq 1$ for all i, j. Then

$$a-bi=\overline{a+bi}=\overline{u}(1-i)^k p_1^{r_1}\cdots p_l^{r_l}\overline{\pi}_1^{s_1}\pi_1^{t_1}\cdots\overline{\pi}_d^{s_d}\pi_d^{t_d}$$

and

$$n=(a+bi)(a-bi)=2^k p_1^{2r_1}\cdots p_l^{2r_l}q_1^{s_1+t_1}\cdots q_d^{s_d+t_d}.$$

We see that for n to be representable in the form $n=a^2+b^2$ it is necessary that in the prime factorization $n=2^k p_1^{x_1}\cdots p_k^{x_k}q_1^{y_1}\cdots q_d^{y_d}$ with $p_1,\ldots,p_l=3 \bmod 4$ and $q_1,\ldots,q_d=1 \bmod 4$ all exponents x_i are even. And this condition is also sufficient: if n factorizes this way, then a and b can be defined by the formula (5.1) where for each i, $r_i=x_i/2$, and for each j, π_j is a non-real prime with $N(\pi_j)=q_j$ and $s_j+t_j=y_j$. It also follows that this n has $(y_1+1)\cdots(y_d+1)$ (not necessarily all distinct) representations in the form $n=a^2+b^2$, corresponding to distinct decompositions of each of y_j as a sum s_j+t_j.

5.4.5. Example. $5850 = 2 \cdot 3^2 \cdot 5^2 \cdot 13$ satisfies the criterion, – the only its inert prime divisor, 3, appears in an even power. For the rest of divisors we have $5 = (2+i)(2-i)$ and $13 = (3+2i)(3-2i)$. So, to write 5850 in the form $\alpha\overline{\alpha}$, we have, up to multiplication by units, the following options for α: $\alpha = (1+i)3(2+i)^2(3+2i) = -51+57i$, $\alpha = (1+i)3(2+i)^2(3-2i) = 33+69i$, and $\alpha = (1+i)3(2+i)(2-i)(3+2i) = 15+75i$. (The other 3 options are obtained by replacing α by $\overline{\alpha}$ and give the same final result.) The corresponding representations of 5850 as sums of squares are, respectively, $51^2 + 57^2$, $33^2 + 69^2$, and $15^2 + 75^2$.