# FACTORIZATION IN INTEGRAL DOMAINS

PETE L. CLARK

Contents



Thanks to Roy Smith, Timothy Y. Chow, Waldek Hebisch, Arturo Magidin, Gjergji Zaimi, Steve Dalton, Georges Elencwajg, Keith Conrad, Dino Lorenzini, Franz Lemmermeyer, Daniel D. Anderson and Keeko Villaveces for useful comments.

I wish to describe the foundations and some basic aspects of the theory of factorization in integral domains. The issue of uniqueness of factorization is *the* beginning of a systematic study of number theory, and it also plays a key role in the study of hypersurfaces and divisors in algebraic geometry. Moreover, the subject has a richness which makes its study inherently rewarding.

Nevertheless I know of no satisfactory treatment of factorization in any text written for a general mathematical audience. While teaching an undergraduate/basic graduate number theory course, I wrote up some notes on factorization. The temptation to do the subject justice caused the notes to expand to their present form.

My goals are as follows: first to present a more comprehensive (and thus overall more elementary) discussion than has previously appeared in the literature. Second, I wish to highlight connections to number theory and (more briefly) algebraic geometry. Third, I wish to emphasize the recurrent role of two (related) phenomena in the theory: the consideration of norm functions of various sorts and the use of order-theoretic concepts, especially well-ordered sets.

## Classical Roots – The Fundamental theorem of Arithmetic

The study of factorization in rings has its roots in elementary number theory and is thus impressively ancient, going back at least to Euclid of Alexandria (circa 300 BCE). The inspiration for the entire theory is the **Fundamental Theorem of Arithmetic**, which can be stated in two parts as follows:

(FTA1) For all integers $n > 1$, $n$ may be written as a product of prime numbers: $n = p_1 \cdots p_r$ (we say "$n$ admits a prime factorization").
(FTA2) For all integers $n > 1$, the factorization of $n$ into primes is essentially unique: that is, if

$$n = p_1 \cdots p_r = q_1 \cdots q_s,$$

then $r = s$ and after reordering the terms of the product we have $p_i = q_i$ for all $i$.

(FTA1) is quite easy to prove, provided we have in our arsenal some form of mathematical induction. One gets an especially clean proof by using the well-ordering principle. Let $S$ be the set of integers $n > 1$ which *do not* have at least one prime factorization. We wish to show that $S$ is empty so, seeking a contradiction, suppose not. Then by well-ordering $S$ has a least element, say $N$. If $N$ is prime, then we have found a prime factorization, so suppose it is not prime: that is, we may write $N = N_1 N_2$ with $1 < N_1, N_2 < N$. Thus $N_1$ and $N_2$ are too small to lie in $S$ so each have prime factorizations, say $N_1 = p_1 \cdots p_r$, $N_2 = q_1 \cdots q_s$, and then $N = p_1 \cdots p_r q_1 \cdots q_s$ gives a prime factorization of $N$, contradiction!

We generally attribute the Fundamental Theorem of Arithmetic to Euclid, although

neither statement nor proof appears in his *Elements*. Why? Because his Elements contains the following result, now called **Euclid's Lemma**.

**Theorem.** *(*[Euc, Prop. VII.30]*) Let $p$ be a prime number and let $a$ and $b$ be positive integers. If $p \mid ab$, then $p \mid a$ or $p \mid b$.*

Let us recall briefly how Euclid proved Euclid's Lemma. Given two positive integers $a$ and $b$, he gives an (efficient!) algorithm to compute the greatest common divisor of $a$ and $b$: repeated division with remainder. The steps of this **Euclidean Algorithm** can easily be reversed to show that the greatest common divisor $\gcd(a, b)$ of $a$ and $b$ may be expressed as an integral linear combination of them: that is, it gives an *effective proof* of the following result.

**Proposition.** *Let $a$ and $b$ be positive integers. Then there exist $x, y \in \mathbb{Z}$ such that*

$$\gcd(a, b) = xa + yb.$$

Now suppose that a prime $p$ divides $ab$ and that $p \nmid a$. Then $1 = \gcd(a, p)$, so applying the Proposition with $p = b$ gives integers $x$ and $y$ such that $1 = xa + yp$. Multiplying through by $b$ we get $b = xab + ypb$. Since $p \mid ab$ it divides both $xab$ and $ypb$ and thus divides $b$. This completes Euclid's proof of Euclid's Lemma.

Of course an immediate induction argument gives a generalized form of Euclid's Lemma: if a prime $p$ divides a product $a_1 \cdots a_n$, then $p \mid a_i$ for at least one $i$.

The point is that, now that we have disposed of (FTA1), (FTA) is *formally equivalent* to Euclid's Lemma. More precisely:

Assume (FTA), and let $p, a, b$ be positive integers with $p$ prime, such that $p \mid ab$. Thus we may write $pc = ab$. Write out prime factorizations $a = p_1 \cdots p_r$, $b = q_1 \cdots q_s$, $c = v_1 \cdots v_t$ and substitute:

$$pv_1 \cdots v_t = p_1 \cdots p_r q_1 \cdots q_s.$$

Now we have two prime factorizations for the same integer, so we must have that $q_j = p$ for some $j$ or $v_k = p$ for some $k$, and thus $p \mid a$ or $p \mid b$.

Conversely, assume (FTA1) and Euclid's Lemma. We wish to show (FTA2). Again, this can be done very cleanly using well-ordering: assume that the set $S$ of integers $n \geq 2$ which admit more than one factorization into primes is nonempty, and let $N$ be the least element:

$$N = p_1 \cdots p_r = q_1 \cdots q_s.$$

Now $p_1 \mid q_1 \cdots q_s$, so by Euclid's Lemma $p_1 \mid q_j$ for at least one $j$. Since $p_1$ and $q_j$ are both primes, this implies $p_1 = q_j$. Dividing through, we get two different expressions of $\frac{N}{p_1}$ as a product of primes. By the minimality of $N$, these two factorizations must be the same, and therefore the prime factorizations of $N$, which are obtained just by multiplying both sides by $p_1 = q_j$, must have been the same, contradiction.

Since the ancient Greeks did not have mathematical induction in any form, (FTA1) would have been difficult for them to prove. But they probably regarded it as obvious. Thus it seems fair to say that Euclid proved the "hard part" of (FTA). That

(FTA2) lies much deeper than (FTA1) will become increasingly clear as our general study of factorization in integral domains gets properly underway.

I expect that the preceding arguments are familiar to you. In fact, I'm counting on it. In the body of this article, we will study implications among various factorization properties of abstract integral domains. To a remarkable degree, the material of this section provides a blueprint for the general case, in particular for the following implications (all terms will be defined in due course):

Euclidean domain $\implies$ Factorization Domain.

Euclidean domain $\implies$ Principal Ideal Domain $\implies$ Bézout domain $\implies$ GCD-domain $\implies$ Euclid's Lemma Domain (EL Domain).

Factorization Domain + EL Domain $\iff$ Unique Factorization Domain.

It is worth asking: are there other ways to prove FTA? The answer is a resounding yes. Indeed, in the early 20th century, direct proofs of FTA were found by Lindemann [Li33] and Zermelo [Z34]. Their proofs are rather similar, so we speak of the **Lindemann-Zermelo argument**.[1] Here it is:

We claim that the factorization of a positive integer is unique. Assume not; then the set of positive integers which have at least two different standard form factorizations is nonempty, so has a least element, say $N$, where:

$$N = p_1 \cdots p_r = q_1 \cdots q_s. \tag{1}$$

Here the $p_i$'s and $q_j$'s are prime numbers, not necessarily distinct from each other. However, we must have $p_1 \neq q_j$ for any $j$. Indeed, if we had such an equality, then we could cancel and, by an inductive argument we have already rehearsed, reduce to a situation in which the factorization must be unique. In particular $p_1 \neq q_1$. Without loss of generality, assume $p_1 < q_1$. Then, if we subtract $p_1 q_2 \cdots q_s$ from both sides of (1), we get

$$M := N - p_1 q_2 \cdots q_s = p_1(p_2 \cdots p_r - q_2 \cdots q_s) = (q_1 - p_1)(q_2 \cdots q_s). \tag{2}$$

By the assumed minimality of $N$, the prime factorization of $M$ must be unique. However, (2) gives two different factorizations of $M$, and we can use these to get a contradiction. Specifically, $M = p_1(p_2 \cdots p_r - q_2 \cdots q_s)$ shows that $p_1 \mid M$. Therefore, when we factor $M = (q_1 - p_1)(q_2 \cdots q_s)$ into primes, at least one of the prime factors must be $p_1$. But $q_2, \ldots, q_j$ are already primes which are different from $p_1$, so the only way we could get a $p_1$ factor is if $p_1 \mid (q_1 - p_1)$. But this implies $p_1 \mid q_1$, and since $q_1$ is also prime this implies $p_1 = q_1$. Contradiction!

One may also give a direct inductive proof of Euclid's Lemma: see e.g. [Ro63].

Here is yet another proof of (FTA2). Let $n = p_1 \cdots p_r$ be a prime factorization

[1] In an earlier draft, I attributed this result to Hasse, Lindemann and Zermelo because of [Has28], which predates [Li33] and [Z34]. However, [Z10, p. 575] contains a valuable discussion of the history of this result, explaining that it was known to Zermelo in 1912 and passed from him to Hensel to Hasse. Lindemann's work seems to be independent of Zermelo's.

of $n$. Using this factorization we can build a composition series for $\mathbb{Z}/n\mathbb{Z}$ whose successive quotients are $\mathbb{Z}/p_i\mathbb{Z}$. Therefore, if we have two different factorizations

$$n = p_1 \cdots p_r = q_1 \cdots q_s$$

we may apply the Jordan-Hölder theorem [Lan02, Thm. I.3.5] to conclude that the multisets of composition factors agree, which means precisely that the factorization is unique up to the ordering of the factors.

## Basic Terminology

Let $R$ be an integral domain. An element $x \in R$ is a **unit** if there exists $y \in R$ such that $xy = 1$. The set of units of $R$ forms a group under multiplication, denoted $R^\times$.

We say a nonzero, nonunit element $x$ of $R$ is **irreducible** if $x$ has only trivial factorizations: that is, if $x = yz$, then one of $y$ or $z$ is a unit.[2] (Note that it cannot be the case that *both* $y$ and $z$ are units, for then $x$ would itself be a unit.)

Example 0.1: The irreducible elements of $\mathbb{Z}$ are $\pm p$, where $p$ is a prime number.[3]

Example 0.2: In $\mathbb{Q}$, or in any field, there are no irreducible elements, because every nonzero element is a unit.

Let $a$ be any nonzero nonunit in an integral domain $R$. An **irreducible factorization** (or just a **factorization**) of $a$ is an expression

$$a = x_1 \cdots x_n,$$

where each $x_i$ is irreducible. In other words, a factorization is an expression of a nonzero nonunit as a product of irreducible elements.

## 1. Norm functions

An interesting link between number theory and algebra is afforded by the study of "norm functions" on rings, namely on functions $N : R \to \mathbb{N}$. Many rings of number-theoretic interest – e.g, the ring $\mathbb{Z}_K$ of integers in any number field $K$ – come endowed with natural norm functions. On the other hand, many abstract algebraic properties of commutative rings turn out to be equivalent to the existence of a norm function with various properties.

### 1.1. Weak multiplicative norms and multiplicative norms.

We say a function $N : R \to \mathbb{N}$ is a **weak multiplicative norm** if it satisfies:

(MN1) $N(0) = 0$, $N(R \setminus \{0\}) \subset \mathbb{Z}^+$; and
(MN2) For all $x, y \in R$, $N(xy) = N(x)N(y)$.

**Proposition 1.** *Let $N : R \to N$ be a weak multiplicative norm on the ring $R$. Then for any unit $a \in R$, $N(a) = 1$.*

[2] The term **atom** is preferred by most contemporary factorization theorists.

[3] The reader may be wondering why we don't simply call irreducible elements "primes". The important but subtle answer is given in §3.3.

*Proof.* We have $N(1) = N(1 \cdot 1) = N(1) \cdot N(1)$, and since $N(1) \neq 0$,[4] we must have $N(1) = 1$. Similarly, if $a$ is a unit, there exists $b \in R$ such that $ab = 1$ and then $1 = N(1) = N(ab) = N(a)N(b)$, which implies $N(a) = N(b) = 1$. □

For any ring $R$, define $N_0 : R \to \mathbb{N}$ by $N_0(0) = 0$, $N_0(R \setminus \{0\}) = 1$.

We say that a weak multiplicative norm $N : R \to \mathbb{N}$ is a **multiplicative norm** if it satisfies the converse of Proposition 1, i.e.,

(MN3) $x \in R$, $N(x) = 1 \implies x \in R^\times$.

**Proposition 2.** *Let $R$ be a commutative ring.*
*a) $R$ is an integral domain iff $N_0$ is a weak multiplicative norm.*
*b) If $R$ admits any weak multiplicative norm, it is an integral domain.*
*c) The map $N_0$ is a multiplicative norm on $R$ iff $R$ is a field, in which case it is the unique weak multiplicative norm on $R$.*

The proof is straightforward and we leave it to the reader.

### 1.2. **Abstract number rings.**

One of the themes of this article is the phenomenon that factorization properties of commutative rings are implied by the existence of a multiplicative norm satisfying certain additional properties. Along with this we consider the inverse problem: do certain structural properties of rings imply the existence of certain kinds of norms? If so, can one use these properties to construct a *canonical* norm? In this section we consider an important class of rings in which the answer is yes.

Consider the following condition ("finite norms") on a commutative ring $R$:

(FN) For all nonzero ideals $I$ of $R$, $\#R/I < \infty$.

For a ring $R$ satisfying (FN), we can define an **ideal norm function**: $||(0)|| := 0$ and for any $I \neq (0)$, $||I|| = \#R/I$. This gives rise to a norm function on elements in the above sense simply by defining $N(a) = N((a))$, i.e., as the norm of the principal ideal $(a) = \{ra \mid r \in R\}$.

**Lemma 3.** *Let $0 \neq I \subset J$ be ideals in a ring $R$ which satisfies (FN). If $||I|| = ||J||$, then $I = J$.*

The proof is immediate.

**Lemma 4.** *Let $R$ be a ring satisfying (FN), and let $N \in \mathbb{Z}^+$. There are only finitely many nonzero ideals of $R$ of norm equal to $N$.*

*Proof.* Choose any $N + 1$ distinct elements $x_1, \ldots, x_{N+1}$ of $R$, and let

$$S = \{x_i - x_j \mid i \neq j\}.$$

Then $S$ is a finite set consisting of nonzero elements of $R$, say $y_1, \ldots, y_M$. For each $1 \leq i \leq M$ by Lemma 3 there are only finitely many ideals $I$ of $R$ containing $y_i$, so overall the set of all ideals containing some element of $S$ is finite. But if $I$ is any ideal of norm $N$, then $I \cap S$ is nonempty by the pigeonhole principle. □

[4]Here we use that $1 \neq 0$ in $R$.

**Proposition 5.** *For a ring $R$ satisfying (FN), exactly one of the following holds:*
*(i) $N(R) = \{0,1\}$. Then $R$ is a field, and $N$ is a multiplicative norm.*
*(ii) $\{0,1\} \subsetneq N(R)$ and $N(R)$ is finite. Then $R$ is a finite ring which is not a domain, and $N$ is not a weak multiplicative norm.*
*(iii) $\{0,1\} \subsetneq N(R)$ and $N(R)$ is infinite. Then $R$ is an infinite integral domain which is not a field, and $N$ is a multiplicative norm.*

*Proof.* Step 0: Since $0 \neq 1$, we always have $\{0,1\} \subset N(R)$. Moreover, $N(x) = 1 \iff xR = R \iff x \in R^\times$, so $N$ is a weak multiplicative norm on $R$ iff it is a multiplicative norm. If $R$ is a field, then by Proposition we have $N(R) = \{0,1\}$. Henceforth we assume that $R$ is not a field.
Step 1: We claim that if $R$ is not an integral domain, then $R$ is a finite ring. Indeed, let $0 \neq a$ be a zero divisor, so $I = \{r \in R \mid ra = 0\}$ is a nonzero ideal of $R$. Consider the map $a\bullet : R \to R,\ r \mapsto ar$; this is an endomorphism of the underlying additive group $(R,+)$. The image of $a\bullet$ is the principal ideal $aR$ and its kernel is $I$, so

$$R/I \cong aR.$$

By hypothesis, $R/I$ is finite, so $aR$ is finite. Moreover, since $aR$ is a nonzero ideal of $R$, $R/aR$ is finite. But we have a short exact sequence

$$0 \to aR \to R \to R/aR,$$

which shows that $R$ itself is finite.
Step 2: If $R$ is finite, then $N(R)$ is finite, and there exist $x, y \in R \setminus \{0\}$ such that $xy = 0$. In particular $0 = N(xy) \neq N(x)N(y)$, so $N$ is not weakly multiplicative.
Step 3: We claim that if $R$ is an infinite domain which is not a field, then $N(R)$ is infinite. Indeed, in such a ring $R$, there exists $a \in R$ which is neither zero nor a unit. Then for all $n \in \mathbb{Z}^+$, $(a^{n+1}) \subsetneq (a^n)$ – otherwise $a^{n+1} \mid a^n$, so that there exists $x \in R$ with $a^{n+1}x = a^n$, or $a^n(ax - 1) = 0$ so $a = 0$ or $a \in R^\times$. By Lemma 3, we have $N(a) < N(a^2) < \ldots$.
Step 4: We claim that if $R$ is a domain then $N$ is a multiplicative norm. For this it is enough to verify (MN2) for $x, y \neq 0$. Consider the quotient homomorphism $R/(xy) \to R/(x)$. This map is surjective, and its kernel is $(x)/(xy)$. Moreover, since $y$ is not a zero divisor, multiplication by $y$ gives an isomorphism of $R$-modules $\varphi_y : R \to yR$. Since $\varphi_y(xR) = xyR$, passing to the quotient gives $R/x \cong yR/xyR$, and this shows $N(xy) = N(x)N(y)$. □

An **abstract number ring** is an infinite ring satisfying (FN) which is not a field.

**Proposition 6.** *Let $R$ be a domain with additive group isomorphic to $\mathbb{Z}^d$ for some $d \in \mathbb{Z}^+$. Then $R$ is an abstract number ring.*

*Proof.* Since $(R,+) \cong \mathbb{Z}^d$, for every $n \in \mathbb{Z}^+$, the multiplication by $n$ map is injective on $R$. In particular $n \cdot 1 \neq 0$, so $R$ has chracteristic 0. If $R$ were a field, $(R,+)$ would contain the additive group of the rational field $\mathbb{Q}$, which would therefore have to be finitely generated as an abelian group. It isn't, so $R$ is not a field.

It remains to show that $R$ has the finite norm property. Since every nonzero ideal contains a nonzero principal ideal, it is enough to verify that for all $0 \neq \alpha \in R$, $R/(\alpha)$ is finite. Now the elements $1, \alpha, \alpha^2, \ldots, \alpha^n, \ldots$ cannot all be linearly independent over $\mathbb{Z}$, so choose the least positive integer $n$ such that there exist integers $a_0, \ldots, a_n$, not all 0, with $a_n\alpha^n + \ldots + a_1\alpha + a_0 = 0$. If $a_0 = 0$, then

since $R$ is a domain and $\alpha \neq 0$, we could divide through to get a linear dependence relation of smaller degree. So $a_0 \neq 0$. Rewriting the equation as

$$-a_0 = \alpha(a_n\alpha^{n-1} + \ldots + a_1), \tag{3}$$

we see that the ideal $\alpha R$ contains the nonzero integer $a_0$. We have a quotient map $R/a_0R \to R/\alpha R$. As an abelian group, $R/a_0R \cong \mathbb{Z}^d/a_0\mathbb{Z}^d \cong (\mathbb{Z}/a_0\mathbb{Z})^d$. In particular it is finite, hence so is its homomorphic image $R/\alpha R$. □

Let $R$ be a domain which is additively isomorphic to $\mathbb{Z}^d$. We claim that its fraction field $K$ is a degree $d$ field extension of $\mathbb{Q}$ (so in particular, $K$ is a number field). Indeed, for $\alpha \in R$, rewriting (3) as

$$\frac{1}{\alpha} = \frac{-(a_n\alpha^{n-1} + \ldots + a_1)}{a_0}$$

shows that to form the fraction field of $R$ we need only adjoin inverses of the nonzero integers. In other words, the natural map $R \otimes_{\mathbb{Z}} \mathbb{Q} \to K$ is an isomorphism.

One says that $R$ is an **order** in the number field $K$. In this case, the canonical norm $N$ we have constructed on $R$ is nothing else than the absolute value of the field norm from $K$ down to $\mathbb{Q}$, restricted to $R$. We leave the proof of this as an exercise for the reader with number-theoretic interests.

Other examples of abstract number rings include: the coordinate ring of an integral affine curve over a finite field, a DVR with finite residue field, and any ring in between an abstract number ring and its fraction field [LeMo72, Thm. 2.3]. In particular, the integral closure of an abstract number ring in its fraction field is an integrally closed abstract number ring. The integral closure of an order $R$ in a number field $K$ is called the **ring of integers** of $K$ and will be denoted here by $\mathbb{Z}_K$. To a very large extent, classical algebraic number theory is the study of properties of the domains $\mathbb{Z}_K$, especially their factorization properties.

### 1.3. **Dirichlet rings.**

We shall give an application of Proposition 1.2 to show that a ring is a domain.

Let $R$ be a ring. The **Dirichlet ring** $\mathcal{D}_R$ is a ring whose elements are the functions $f : \mathbb{Z}^+ \to R$. We define addition pointwise, i.e.,

$$(f + g)(n) := f(n) + g(n),$$

whereas multiplication is given by the convolution product

$$(f * g)(n) = \sum_{d_1 d_2 = n} f(d_1)g(d_2).$$

When $R = \mathbb{R}$ or $\mathbb{C}$, this is often called the ring of arithmetic functions.

**Theorem 7.** *A ring $R$ is an integral domain iff $\mathcal{D}_R$ is an integral domain.*

*Proof.* The map $R \hookrightarrow \mathcal{D}_R$ which sends $r$ to the function which carries 1 to $r$ and every other positive integer to 0 embeds $R$ as a subring of $\mathcal{D}_R$. So if $\mathcal{D}_R$ is a domain, certainly $R$ is. Conversely, by Proposition 1.2 it suffices to construct a weak multiplicative norm function on $\mathcal{D}_R$. The function $N$ which sends the 0 function to 0 and any other function $f$ to the least $n$ such that $f(n) \neq 0$ is easily checked to be a weak multiplicative norm. □

Remark 1.1: Let $R$ be a domain. Then the weak multiplicative norm $N$ constructed on $\mathcal{D}_R$ above is a multiplicative norm if and only if $R$ is a field.

The Dirichlet ring $\mathcal{D}_{\mathbb{C}}$ is often called the ring of **arithmetic functions**.

## 2. Factorization domains

Let us say that a domain $R$ is a **factorization domain** (for short, **FD**) if every nonzero nonunit element has a factorization into irreducibles.[5]

Example 2.1: A field is trivially a FD: it has no nonzero nonunits.

Example 2.2: Part a) of the fundamental theorem of arithmetic asserts that $\mathbb{Z}$ is a FD. The proof was an easy "minimal counterexample" argument.

In practice, most domains encountered in algebra and number theory are factorization domains. We justify this statement by giving two sufficient conditions for a domain to be a FD, each of which is widely applicable.

The idea behind the first condition is extremely simple: factorization ought to be a process of decomposing more complex objects into simpler ones. If to every nonzero element $a$ of $R$ we can assign a positive integer "complexity" $C(a)$ such that in any nontrivial factorization $a = bc$ – i.e., with $b$ and $c$ nonunits – we have $1 \leq C(b),\ C(c) < C(a)$ – then factorizations lower the complexity so that eventually the process must terminate.

In particular any multiplicative norm on $R$ satisfies this key property, so:

**Proposition 8.** *A ring admitting a multiplicative norm is a factorization domain.*

*Proof.* We need only adapt the inductive proof of (FTA1). Indeed, let $N$ be a multiplicative norm on the ring $R$. Suppose for a contradiction that the set $S$ of nonzero nonunits in $R$ which *do not* admit irreducible factorizations is nonempty. Then there exists $x \in S$ with $N(x)$ minimal. Such an $x$ is not irreducible, so it can be factored as $x = yz$, with both $y, z$ nonunits. Then $N(x) = N(y)N(z) \in \mathbb{Z}^+$ and $N(y), N(z) > 1$, so that $N(y), N(z) < N(x)$. But $y$ and $z$, having smaller norms than $x$, each have irreducible factorizations, say $y = y_1 \cdots y_r$ and $z = z_1 \cdots z_s$. Then $x = y_1 \cdots y_r z_1 \cdots z_s$ is an irreducible factorization of $x$, contradiction. □

Now for the second condition. In a domain $R$, we an element $a$ **properly divides** an element $b$ if $b = xa$ and $x$ is *not* a unit. This condition is equivalent to $a \mid b$ but $b \nmid a$ and also to $(a) \supsetneq (b)$. A domain $R$ satisfies the **ascending chain condition on principal ideals** (henceforth **ACCP**) if there does not exist an infinite sequence of elements $\{a_i\}_{i=1}^{\infty}$ of $R$ such that for all $i$, $a_{i+1}$ properly divides $a_i$.

Example 2.3: The integers satisfy ACCP: indeed if the integer $a$ properly divides the integer $b$, then $|a| < |b|$, so an infinite sequence of proper divisors would, again, contradict the well-ordering of the natural numbers.

[5]The term **atomic domain** is used by specialists in the area, but is not so familiar to a general mathematical audience. Our chosen terminology seems more transparent.

Notice that the second condition is more general than the first, i.e., any ring $R$ which admits a multiplicative norm satisfies ACCP.: if $a$ properly divides $b$, $N(a)$ properly divides $N(b)$ and hence $0 \leq N(a) < N(b)$.

**Proposition 9.** *For a commutative ring $R$, the following are equivalent:*
*(i) There are no ascending sequences $(a_1) \subsetneq (a_2) \subsetneq \ldots$ of principal ideals in $R$.*
*(ii) Any nonempty set $\mathcal{F}$ of principal ideals of $R$ has a maximal element. In other words, there exists a principal ideal $I \in \mathcal{F}$ which is not properly contained in any other principal ideal in $\mathcal{F}$.*
*(iii) There is no sequence $\{a_i\}_{i=1}^{\infty}$ in $R$ with $a_{i+1}$ properly dividing $a_i$ for all $i \geq 1$.*

The argument of (i) $\iff$ (ii) comes up many times in this subject, so for efficiency of future use we isolate it in a more abstract form.[6]

**Lemma 10.** *Let $(S, \leq)$ be a partially ordered set. The following are equivalent:*
*(i) There are no infinite sequences*

$$s_1 < s_2 < \ldots < s_n < \ldots \tag{4}$$

*of elements in $S$.*
*(ii) Any nonempty subset $\mathcal{F}$ of $S$ has a maximal element, i.e., there exists $x \in \mathcal{F}$ such that if $y \in \mathcal{F}$ and $x \leq y$ then $y = x$.*

*Proof.* It is easier (and, of course, sufficient) to prove that (i) fails iff (ii) fails. Indeed, if (i) fails, then there exists an infinite sequence as in (4) above, and then $\mathcal{F} = \{s_i\}_{i=1}^{\infty}$ is a nonempty subset of $S$ without a maximal element. Conversely, if (ii) fails, let $\mathcal{F}$ be a nonempty subset of $S$ without maximal elements. Since it is nonempty, there exists $s_1 \in \mathcal{F}$. Since $s_1$ is not maximal, there exists $s_2 \in \mathcal{F}$ with $s_1 < s_2$. Continuing in this way, we build an infinite sequence as in (4). □

*Proof of Proposition 9:* We see that (i) $\iff$ (ii) by applying Lemma 10 to the partially ordered set of principal ideals of $R$, with $(a) \leq (b)$ iff $(a) \subset (b)$. (i) $\iff$ (iii): an infinite sequence $\{a_i\}_{i=1}^{\infty}$ with $a_{i+1}$ properly dividing $a_i$ yields a strictly ascending sequence of principal ideals $(a_1) \subsetneq (a_2) \subsetneq \ldots$, and conversely. □

**Proposition 11.** *A principal ideal domain satisfies ACCP.*

*Proof.* Let $R$ be a principal ideal domain, and suppose for a contradiction that there exists a sequence $\{a_i\}_{i=1}^{\infty}$ in $R$ such that $(a_1) \subsetneq (a_2) \subsetneq \ldots$. Put $I = \bigcup_{i=1}^{\infty}(a_i)$. By assumption $I$ is principal, say $I = (a)$. One the one hand we have $(a) \supset (a_i)$ for all $i$, but on the other hand, the element $a$ must lie in $(a_N)$ for some $N$ and hence also $a_{N+k}$ for all $k \geq 0$. We conclude that $(a_N) = (a_{N+1}) = \ldots$, contradiction. □

**Proposition 12.** *An integral domain satisfying ACCP is a factorization domain.*

*First Proof*: Let $S'$ be the set of all nonzero nonunit elements of $R$ which cannot be factored into irreducibles. Assume, for a contradiction, that $S'$ is nonempty. Then the corresponding set

$$S = \{(x) \mid x \in S'\}$$

[6]The following argument uses "Dependent Choice", a mild form of the Axiom of Choice (AC). However, AC is *equivalent* to the assertion that every nonzero ring has a maximal ideal, a ubiquitously used fact of commutative algebra. Thus in commutative algebra it is standard to assume AC, so no more comments in that direction here.

of principal ideals generated by elements of $S'$ is also nonempty. By ACCP and Remark 1, there exists a maximal element $(x)$ of $S$. Now just follow your nose: by definition of $x$, it is not irreducible, so can be written as $x = yz$ with $y$ and $z$ nonunits. This means that the principal ideals $(y)$ and $(z)$ each strictly contain the principal ideal $(x)$, so by the assumed maximality of $(x)$, both $y$ and $z$ can be factored into irreducibles: $y = y_1 \cdots y_r,\ z = z_1 \cdots z_s$, so (as usual!) we get $x = y_1 \cdots y_r z_1 \cdots z_s$ so $x$ has an irreducible factorization after all, contradiction. □

*Second Proof*: We take a more direct approach. Let $x$ be a nonzero nonunit element. We claim first that there exists a divisor $y$ of $x$ such that $y$ is irreducible. Certainly this holds if $x$ is irreducible, so assume that $x = y_1 z_1$ with both $y$ and $z_1$ properly dividing $x$. If $y_1$ is irreducible again our claim holds, so assume that $y_1 = y_2 z_2$ with $y_2$ strictly dividing $y_1$, and thus $x = y_2 z_1 z_2$ with $(x) \subsetneq (y_1) \subsetneq (y_2)$. Continuing in this way – i.e., replacing $y_n$ by $y_{n+1} z_{n+1}$ with $y_{n+1}, z_{n+1}$ properly dividing $y_n$ if $y_n$ is irreducible – we would get an infinite strictly ascending chain $(y_1) \subsetneq (y_2) \subsetneq \ldots$ of principal ideals, contrary to our assumption. So this cannot be the case, i.e., for some $n$, $y_n$ is an irreducible divisor of $x$.

We have shown that any nonzero nonunit, reducible element $x$ of $R$ can be "partially factored", i.e., written as $x = a_1 y_1$ with $a_1$ irreducible and $y_1$ a nonzero nonunit. If $y$ is irreducible, we have completely factored $x$; if not, the claim applies to $y$, giving $x = a_1 a_2 y_2$ with $(y_2) \subsetneq (y_1)$. Now we argue as above: if this process never terminated, we would produce a strictly ascending sequence $(y_1) \subsetneq (y_2) \subsetneq \ldots$ contradicting ACCP. So for some $n$ we must have $x = a_1 \cdots a_n y_n$ with $y_n \in R^\times$, and thus $x = a_1 \cdots a_{n-1}(y_n a_n)$ is an irreducible factorization of $x$. □

*Third proof*: We will show a stronger statement. Namely, if we have a partial factorization $x = a_1 \cdots a_n$ – i.e., each $a_i$ is a nonunit but not necessarily irreducible, then we define an **elementary move** to be the selection of an(y) index $i$ such that $a_i$ is reducible and the replacement of $a_i$ in the product by $b_1 b_2$, where $a = b_1 b_2$ and $b_1$ and $b_2$ are nonunits. In the preceding proof, we showed that there always exists a sequence of elementary moves which terminates in an irreducible factorization – indeed, this is equivalent to the existence of an irreducible factorization! But now consider the stronger condition that *any* sequence of elementary factorizations eventually terminates. This clearly implies ACCP. We now show the converse: let $R$ be an ACCP domain. A sequence of elementary moves defines a rooted binary tree, and thus an infinite such sequence yields an infinite binary tree. But by König's Infinity Lemma [Kö36], an infinite connected graph in which each vertex has finite degree must admit an infinite path, which gives rise to an infinite strictly ascending chain of principal ideals, contradiction! □

## 3. A deeper look at factorization domains

### 3.1. **A non-factorization domain.**

The ring $\overline{\mathbb{Z}}$ of all algebraic integers is not a factorization domain. In fact, $\overline{\mathbb{Z}}$ is in sense as far from a factorization domain as possible: it has many nonzero nonunit elements, but no irreducible elements![7] We briefly sketch an argument for this: first, there exist nonzero nonunit elements of the ring, for instance the element 2.

[7]Such rings are known to experts as **antimatter domains**: they have no atoms.

Its multiplicative inverse in the fraction field $\overline{\mathbb{Q}}$ (of all algebraic numbers) is $\frac{1}{2}$, and $\frac{1}{2}$ is not an algebraic integer. Second, we claim that there are no irreducible elements in $\overline{\mathbb{Z}}$. Namely, if $x$ is any nonzero nonunit algebraic integer, then one can check that $\sqrt{x}$ is also a nonzero nonunit algebraic integer and $x = \sqrt{x}\sqrt{x}$.

Remark 3.1: It follows from the material of the previous section that any domain which is not a factorization domain is a ring which admits a weak multiplicative norm (e.g. the trivial norm $N_0$) but no multiplicative norm. Thus $\overline{\mathbb{Z}}$ gives a specific example of such a domain. In fact the above argument gives more: because the only positive integer which is an $n$th power for all $n$ is 1, the only weak multiplicative norm on $\overline{\mathbb{Z}}$ is the trivial norm $N_0$.

More generally, if $R$ is any domain which is not a field and for which there exists $n > 1$ such that the $n$th power map $x \mapsto x^n$ is surjective, then $R$ has nonzero nonunits but no irreducible elements, so is not a factorization domain.

### 3.2. **FD versus ACCP.**

Is every factorization domain an ACCP domain? In 1968, the distinguished algebraist P.M. Cohn claimed an affirmative answer [Coh68, Prop. 1.1], however without giving any proof. In 1974 Anne Grams showed that no such proof was possible.

**Theorem 13.** *There exist factorization domains which do not satisfy ACCP.*

*Proof.* See [Gr74]. □

### 3.3. **ACC versus ACCP.**

Many students of ring theory are less familiar with ACCP than with the following:

**Proposition 14.** *For a ring $R$, the following conditions are equivalent:*
*(i) Every nonempty set $S$ of ideals of $R$ has a maximal element, i.e., an element $I \in S$ such that $I$ is not properly contained in any other ideal $J$ of $S$.*
*(ii) (ACC) In any infinite sequence of ideals*

$$I_1 \subseteq I_2 \subseteq \ldots \subseteq I_n \subseteq \ldots$$

*we have equality from some point onward: there exists $N \in \mathbb{Z}^+$ such that for all $k \geq 0$, $I_{N+k} = I_n$.*
*(iii) Every ideal $I$ of $R$ is finitely generated: there exist finitely many elements $x_1, \ldots, x_n$ in $R$ such that*

$$I = \langle x_1, \ldots, x_n \rangle = \{r_1 x_1 + \ldots + r_n x_n \mid r_i \in R\}.$$

*A ring satisfying these equivalent properties is caled **Noetherian**.*

*Proof.* (i) $\iff$ (ii): For any nonempty family $\mathcal{F}$ of subsets of a given set $R$, the condition that that any infinite sequence $I_1 \subset I_2 \subset \ldots$ of elements of $\mathcal{F}$ is equivalent to the condition that every nonempty subset of $\mathcal{F}$ has a maximal element: if (i) does not hold, then there exists a sequence $I_1 \subsetneq I_2 \subsetneq I_3 \subsetneq \ldots$, and then $\{I_n\}_{n=1}^{\infty}$ has no maximal element. Conversely, if (ii) does not hold, then there exists $I_1 \in \mathcal{F}$; since $I_1$ is not maximal, so there exists $I_2 \in \mathcal{F}$ such that $I_2 \supsetneq I_1$, since $I_2$ is not

maximal, there exists $I_3 \in \mathcal{F}$ such that $I_3 \supsetneq I_2$: continuing in this way, we build an infinite strictly ascending chain.
(ii) $\implies$ (iii): If there exists an ideal $I$ which is not finitely generated, then for any $x_1 \in I$, $I_1 := \langle x_1 \rangle \subsetneq I$. Since $I_1$ is finitely generated and $I$ is not, there exists $x_2 \in I \setminus I_1$. Put $I_2 = \langle x_1, x_2 \rangle$, so $I_2 \subset I$. Again, because $I$ is not finitely generated, there exists $x_3 \in I \setminus I_2$. In this way we construct an infinite strictly ascending chain $I_1 \subsetneq I_2 \subsetneq I_3 \subsetneq \ldots$, contradicting (ii).
(iii) $\implies$ (ii): Let $I_1 \subseteq I_2 \subseteq \ldots$ be an infinite sequence of ideals. Then the union $I := \bigcup_{i=1}^{\infty} I_i$ is again an ideal. By assumption, $I$ is finitely generated, so there exist $x_1, \ldots, x_n \in R$ with $I = \langle x_1, \ldots, x_n \rangle$. But since $I$ is the union of the $I_i$'s, for each $1 \leq i \leq n$, there exists $k_i \in \mathbb{Z}^+$ such that $x_i \in I_{k_i}$. Put $k = \max(k_1, \ldots, k_n)$; then $x_1, \ldots, x_n$ are all in $I_k$, so $I = I_k$, which forces $I_k = I_{k+1} = \ldots = I$. □

**Proposition 15.** *A principal ideal domain is a Noetherian domain.*

*Proof.* This is an immediate consequence of the definitions: a PID is a domain in which each ideal can be generated by a single element, whereas a Noetherian ring is one in which each ideal can be generated by finitely many elements. □

Noetherianity is justly regarded as the single most important condition on a ring. This esteem comes in part from the large class of Noetherian rings:

**Theorem 16.** *Let $R$ be a Noetherian ring.*
*a) If $I$ is any ideal of $R$, then the quotient $R/I$ is Noetherian.*
*b) The polynomial ring $R[t]$ is Noetherian.*
*c) The formal power series ring $R[[t]]$ is Noetherian.*

*Proof.* Part a) follows immediately from the inclusion preserving correspondence between ideals of $R/I$ and ideals of $R$ containing $I$, whereas part b) is the celebrated **Hilbert basis theorem**: see e.g. [AM69, Cor. 7.6] or [Lan02, Thm IV.4.1]. Part c) is a variant of the Hilbert basis theorem; see e.g. [Mat89, Thm. 3.3]. □

Unlike ACC, the condition ACCP does not in general pass to quotient rings (even quotient domains). Indeed, it will follow from Theorem 29 that a polynomial ring $\mathbb{Z}[\mathbf{t}] := \mathbb{Z}[(t_i)_{i \in J}]$ in any set $J$ of indeterminates is an ACCP domain. But every commutative ring is isomorphic to a quotient of some ring $\mathbb{Z}[\mathbf{t}]$.

On the other hand, the analogue of Theorem 16b) for ACCP does hold:

**Theorem 17.** *$R$ be an ACCP domain. Then:*
*a) The polynomial ring $R[t]$ is an ACCP domain.*
*b) The formal power series ring $R[[t]]$ is an ACCP domain.*

*Proof.* a) In an infinite ascending chain $(P_i)$ of principal ideals of $R[t]$, $\deg(P_i)$ is a descending chain of non-negative integers, so eventually stabilizes. Therefore for sufficiently large $n$, we have $P_n = a_n P_{n+1}$, where $a_n \in R$ and $(a_{n+1}) \supset (a_n)$. Since $R$ satisfies (ACCP) we have $(a_n) = (a_{n+1})$ for sufficiently large $n$, whence $(P_n) = (P_{n+1})$ for sufficiently large $n$: $R[t]$ satisfies (ACCP).
b) The proof of part a) goes through so long as we make the following modification: we replace the degree of the polynomial $P_i$ by the *order of vanishing* of the formal power series $P_i$: explicitly, if $P = \sum_{n=N}^{\infty} a_n t^n$ with $a_N \neq 0$, then $\operatorname{ord} P = N$. □

In the important paper [AAZ90], the following question is asked: if $R$ is a factorization domain, must $R[t]$ be a factorization domain as well? A negative answer was given by M. Roitman [Roi93]. Note that this, together with Theorem 17, certainly implies Theorem 13, and gives a hint that the property of being a factorization domain is, by itself, not worth very much.

## 4. Unique factorization domains

### 4.1. Associates, Prin(R) and G(R).

In order to give a definition of a unique factorization domain, we must specify when two different factorizations of the same nonzero nonunit $x$ are to be regarded as "equivalent." In the case of factorizations of positive integers into prime numbers, we only had to worry about the ordering of the irreducible factors. Of course we still wish to regard two factorizations into irreducibles differing only in the order of the factors as equivalent, but there is more to say. For instance, in $\mathbb{Z}$ we have

$$18 = 2 \cdot 3 \cdot 3 = (2) \cdot (-3) \cdot (-3),$$

and several other choices for the sign besides. The correct generalization of this to an arbitrary domain comes from the following observation: if $x$ is an irreducible element of $R$ and $u$ is a unit of $R$, then $ux$ is also an irreducible element of $R$. Similarly, by multiplying by units we can get many different equivalent-looking factorizations, e.g.

$$a = x_1 \cdots x_r = (ux_1) \cdots (ux_{r-1}) \cdot (u^{1-r} x_r).$$

Thus we need a relation between elements which regards two elements as equivalent iff they differ multiplicatively by a unit. In fact this is itself a well-defined relation: its properties are recorded below.

**Proposition 18.** *Let $R$ be a domain, and let $x, y \in R$. The following are equivalent:*
*(i) $x \mid y$ and $y \mid x$.*
*(ii) There exists a unit $u \in R^\times$ such that $y = ux$.*
*(iii) We have an equality of principal ideals $(x) = (y)$.*
*If $x$ and $y$ satisfy any (hence all) of the conditions above, we say that $x$ and $y$ **are associates** and write $x \sim y$.*

The proof amounts to unwinding the definitions. We leave it to the reader.

In modern mathematics, when one defines an equivalence relation $\sim$ on a structure $X$ it is often worthwhile to explicitly consider the natural map from $X \to X/\sim$ (i.e., from $X$ to the set of $\sim$-equivalence classes of $X$) as being a "quotient map" and to use it to define some structure on $X/\sim$. In particular this is a fruitful perspective to take for the relation of associates on a domain $R$.

It turns out to be best to omit the zero element; namely, define $R^\bullet = R \setminus \{0\}$: this is a commutative monoid, with identity 1. Moreover the relation of associate elements is *compatible* with the monoid structure on $R^\bullet$, in the sense that $x_1 \sim y_1,\ x_2 \sim y_2 \implies x_1 x_2 \sim y_1 y_2$. Thus there is a unique monoid structure on the quotient $R^\bullet/\sim$ which makes the quotient map

$$R^\bullet \to R^\bullet/\sim$$

into a homomorphism of commutative monoids.

The quotient monoid $R^\bullet/\sim$ may be identified with the monoid of principal ideals of $R$ under the usual product: $(a)(b) = (ab)$. We thus denote it by $\mathrm{Prin}(R)$.

Here is another useful way to view $\mathrm{Prin}(R)$ which shows that it has additional structure. Namely, consider divisibility as a relation on $R^\bullet$. It is reflexive and transitive, but it (generally) fails to be anti-symmetric precisely because of the existence of nontrivial associate elements. Thus divisibility is a *quasi-ordering* on $R^\bullet$. Given any quasi-ordered set $(P, \prec)$, there is a canonical way of making a partially ordered set $(S, \leq)$ out of it: namely, we simply introduce the equivalence relation $x \sim y$ if $x \prec y$ and $y \prec x$, and take the quotient $S = P/\sim$. Applying this to $R^\bullet$ gives $\mathrm{Prin}(R)$.

A commutative ring $R$ is a domain iff the monoid $R^\bullet$ is **cancellative**: that is, if $x, y, z \in R^\bullet$ are such that $xz = yz$, then $x = y$. One checks immediately that if $R$ is a domain, this implies that $\mathrm{Prin}(R)$ is cancellative as well. In general, every commutative monoid $M$ has a **group completion**, i.e., a group $G(M)$ and a monoid homomorphism $M \to G(M)$ which is universal for monoid maps from $M$ into a group. A commutative monoid $M$ is cancellative iff the natural map $M \to G(M)$ is injective. In this case, $G(R) := G(\mathrm{Prin}(R))$ is simply $K^\times/R^\times$, where $K$ is the field of fractions of the domain $R$. This gives a third description of $\mathrm{Prin}(R)$: as a submonoid of $G(R)$ it is simply $R^\bullet/R^\times$. The group $G(R)$ is usually called the **group of divisibility** of $R$. The partial ordering on $\mathrm{Prin}(R)$ extends naturally to a partial ordering on $G(R)$: explicitly, for $x, y \in K^\times$, $x \leq y \iff \frac{y}{x} \in R^\bullet$.

Any property of $R$ which is phrased only in terms of divisibility of elements can be faithfully described in the monoid $\mathrm{Prin}(R)$. This includes units, irreducible elements, and the conditions FD and ACCP as well as many to come.

Example: A domain $R$ is a field iff $G(R)$ is trivial.

One implication of this is that one can study factorization at the level of cancellative commutative monoids. This is not a thread that we shall pursue here, although it has certainly been pursued by others in the literature. Our interest is rather as follows: sometimes a property is most cleanly and naturally phrased in terms of the factorization monoid or group.

### 4.2. **Valuation rings.**

We wish to introduce a class of rings for which consideration of the group of divisibility $G(R)$ is especially fruitful.

A domain $R$ is a **valuation ring** if $G(R)$ is a totally ordered abelian group. This is easily seen to be equivalent to the more usual condition that for any $x \in K^\times$, at least one of $x, x^{-1}$ lies in $R^\bullet$. But it is in many respects more graceful. For instance, it follows immediately from Lemma 10 that a valuation ring is an ACCP domain iff $\mathrm{Prin}(R)$ is well-ordered.

**Theorem 19.** *Let $(\Gamma, +)$ be a nontrivial totally ordered commutative group such that $\Gamma^+ := \{x \in \Gamma \mid x \geq 0\}$ is well-ordered. Then $\Gamma$ is isomorphic to $\mathbb{Z}$.*

*Proof.* A totally ordered commutative group is **Archimedean** if for all $x, y > 0$, there exists $n \in \mathbb{Z}^+$ such that $nx > y$. A classical theorem of Hölder [Hö01] asserts that every Archimedean totally ordered group is order isomorphic to a subgroup of $(\mathbb{R}, +)$. Now let $\Gamma \subset \mathbb{R}$ be a nontrivial well-ordered subgroup. If $\Gamma$ is discrete, it is generated by its least positive element, so is isomorphic to $\mathbb{Z}$. If not, 0 is an accumulation point of $\Gamma$, so there exists an infinite strictly decreasing sequence of positive elements of $\Gamma$, so $\Gamma^+$ is not well-ordered.

If $\Gamma$ is non-Archimedean, choose $x, y > 0$ such that for all $n \in \mathbb{Z}^+$, $nx < y$. Then $\{y - nx\}_{n \in \mathbb{Z}^+}$ is an infinite strictly descending chain in $\Gamma^+$. □

Thus any valuation ring $R$ with noncyclic $G(R)$ is not an ACCP domain. But indeed such rings are remarkably plentiful, as the following two results show.

**Theorem 20.** *(Levi* [Le43]*) A commutative group $\Gamma$ admits a total ordering iff it is torsionfree, i.e., has no nontrivial elements of finite order.*

Moreover, let $\Gamma$ be any totally ordered commutative group, and let $k$ be any field. Define $R = k((\Gamma))$ to be the ring of $k$-valued formal series $a = \sum_{g \in \Gamma} a_g t^g$ such that the set $\{g \in \Gamma \mid a_g \neq 0\}$ is well-ordered. An element of $R$ is called a **Malcev-Neumann series** [Mal48] [Ne49]. It is not hard to show that as a partially ordered group, $G(R) \cong \Gamma$.

### 4.3. **Unique factorization domains.**

Finally, we can give the key definition. An integral domain $R$ is a **unique factorization domain** (UFD)[8] if:

(UFD1) = (FD) Every nonzero nonunit admits an irreducible factorization; and

(UFD2) If $a = x_1 \cdots x_r = y_1 \cdots y_s$ are two irreducible factorizations of $a$, then $r = s$, and there exists a permutation $\sigma$ of $\{1, \ldots, r\}$ such that for all $1 \leq i \leq r$, $x_i \sim y_{\sigma(i)}$. That is, after reordering the elements we can pair off each irreducible in the first factorization with an associate irreducible in the second factorization.

Notice that – up to the very minor need to introduce associate elements to discuss unique factorization of possibly negative integers – the conditions (UFD1) and (UFD2) reduce, in the case $R = \mathbb{Z}$, precisely to (FTA1) and (FTA2).

Thus, as in our discussion in the introduction, it is extremely fruitful to introduce a property axiomatizing Euclid's Lemma.

We say an integral domain $R$ is an **EL-domain** if for all irreducible elements $x$ of $R$, if $x \mid ab$, then $x \mid a$ or $x \mid b$. Of course this immediately implies that if $x$ is irreducible and $x$ divides $a_1 \cdots a_n$, then $x \mid a_i$ for some $i$.

**Theorem 21.** *Let $R$ be a factorization domain. Then $R$ is a unique factorization domain iff it is an EL-domain.*

[8]The term **factorial domain** is also commonly used, especially by continental mathematicians.

*Proof.* We showed in the introduction that (FTA) $\iff$ (FTA1) + Euclid's Lemma, which is nothing else than the present result for $R = \mathbb{Z}$. The proof given there extends essentially verbatim to the general case. □

More precisely, we have the following useful characterization of UFDs:

**Theorem 22.** *For an integral domain $R$, the following are equivalent:*
*(i) $R$ is a unique factorization domain.*
*(ii) $R$ is an ACCP domain and is an EL-domain.*
*(iii) $R$ is a factorization domain and an EL-domain.*

*Proof.* (i) $\implies$ (ii): suppose $R$ is a UFD. By Theorem 21, $R$ is an EL-domain. Moreover, suppose $R$ does not satisfy ACCP: $(x_1) \subsetneq (x_2) \subsetneq \ldots$. Then $x_2$ is a nonzero nonunit. Since $(x_3)$ strictly contains $x_2$, there exists a nonzero nonunit $y_1$ such that $x_2 = x_3 y_1$. Since $x_3$ and $y_1$ are both nonzero nonunits, they have unique factorizations into irreducibles, which means that the unique factorization of $x_2$ into irreducibles has at least two irreducible factors. Similarly, there exists a nonzero nonunit $y_2$ such that $x_3 = x_4 y_2$, so $x_2 = x_4 y_2 y_1$, so that we now know that the unique factorization of $x_2$ into irreducibles has at least 3 irreducible factors. Proceeding in this way we can show that the unique factorization of $x_n$ into irreducibles has at least $n$ irreducible factors for any $n \in \mathbb{Z}^+$, which is absurd.
(ii) $\implies$ (iii) by Proposition 12, whereas (iii) $\implies$ (i) by Theorem 21. □

### 4.4. **Prime elements.**

Recall the notion of a **prime ideal** $\mathfrak{p}$ in a ring $R$: this is a proper ideal such that $x, y \in R,\ xy \in \mathfrak{p}$ implies $x \in \mathfrak{p}$ or $y \in \mathfrak{p}$.

Let us define a nonzero element $x$ in a domain $R$ to be a **prime element** if the principal ideal $(x)$ is a prime ideal. Unpacking this, we see that an element $x$ is prime iff $x \mid ab$ implies $x \mid a$ or $x \mid b$.

**Lemma 23.** *a) In any domain $R$, a prime element is irreducible.*
*b) A domain $R$ is an EL-domain exactly when all irreducible elements are prime.*

*Proof.* a) If $x = ab$ with $a$ and $b$ nonunits, then certainly $x \nmid a$ and $x \nmid b$.
b) This is, of course, the definition of an EL-domain. □

In particular, since UFD $\implies$ EL-domain, in any UFD there is no distinction to be made between irreducible elements and prime elements. Conversely, a FD will fail to be a UFD iff there exist irreducible elements which are not prime.

The condition that a domain $R$ be a UFD can also be expressed very elegantly in terms of the group of divisbility $G(R)$ of $\S X.X$. Namely, $R$ is a UFD iff $G(R)$ is isomorphic, as an ordered abelian group, to the direct sum of copies of $(\mathbb{Z}, +)$, one copy for each principal prime ideal.

### 4.5. **Norms on UFDs.**

In this section we give a complete description of all weak multiplicative norms (and also all multiplicative norms) on a UFD.

Let $R$ be a UFD and $N : R \to \mathbb{N}$ be a weak multiplicative norm. As for any domain, if $x$ and $y$ are associate elements of $R$, $y = ux$ for $u \in R^\times$, so

$$N(y) = N(ux) = N(u)N(x) = 1 \cdot N(x) = N(x).$$

Let $\mathcal{P}$ be the set of principal nonzero prime ideals of $R$. For each $\mathfrak{p} \in \mathcal{P}$, choose any generator $\pi_{\mathfrak{p}}$. Put $n_{\mathfrak{p}} := N(\pi_{\mathfrak{p}})$. This data completely determines $N$, since any nonzero element $x$ of $R$ can be written in the form $u \prod_{\mathfrak{p} \in \mathbb{P}} \pi_{\mathfrak{p}}^{x_{\mathfrak{p}}}$ with $x_{\mathfrak{p}} \in \mathbb{N}$ and $x_{\mathfrak{p}} = 0$ for all but finitely many elements of $\mathcal{P}$, and then we must have

$$N(x) = \prod_{\mathfrak{p} \in \mathcal{P}} n_{\mathfrak{p}}^{x_{\mathfrak{p}}} \tag{5}$$

Conversely, by assigning to each $\mathfrak{p} \in \mathcal{P}$ a positive integer $n_{\mathfrak{p}}$, we can define a function $N : R \setminus \{0\} \to \mathbb{Z}^+$ by

$$x = u \prod_{\mathfrak{p} \in \mathbb{P}} \pi_{\mathfrak{p}}^{x_{\mathfrak{p}}} \mapsto N(x) = \prod_{\mathfrak{p} \in \mathcal{P}} n_{\mathfrak{p}}^{x_{\mathfrak{p}}}$$

(and $N(0) := 0$, of course), then $N$ is a weak multiplicative norm.

Again this may be rephrased using the monoid $\mathrm{Prin}(R)$: for any domain $R$, the weak multiplicative norms on $R$ correspond bijectively to monoid homomorphisms $\mathrm{Prin}(R) \to (\mathbb{Z}^+, \cdot)$. If $R$ is a UFD, $\mathrm{Prin}(R) \cong \bigoplus_{\mathbb{P}} (\mathbb{N}, +)$ is the free commutative monoid on the set of principal nonzero prime ideals, so a homomorphism $\mathrm{Prin}(R) \to (\mathbb{Z}^+, \cdot)$ is uniquely specified by the image of each principal nonzero prime ideal.

A multiplicative norm $N$ on a general integral domain may also be expressed in terms of $\mathrm{Prin}(R)$, but not as gracefully: a multiplicative norm determines and is determined by a monoid homomorphism $N : \mathrm{Prin}(R) \to \mathbb{Z}^+$ with the additional property that $N(x) = 1 \iff x = 1$, the identity of $\mathrm{Prin}(R)$. At first glance this looks to be the condition of injectivity, but it is significantly weaker than that: for a homomorphism of monoids $\varphi : M \to N$, that the preimage of $1_N$ is just $1_M$ does not preclude the existence of $x \neq y \in M$ such that $\varphi(x) = \varphi(y) \neq 1$. For instance, if $a, b, d \in \mathbb{Z}^+$ with $a < b$, then in the number ring $\mathbb{Z}[\sqrt{-d}]$, the two elements $a + b\sqrt{-d}$ and $b + a\sqrt{-d}$ each have norm $a^2 + db^2$ but are not associate.

Nevertheless, if $R$ is a UFD things are straightforward: for the weak multiplicative norm $N = (n_{\mathfrak{p}})_{\mathfrak{p} \in \mathcal{P}}$ on $R$ to be a multiplicative norm, it is necessary and sufficient that $n_{\mathfrak{p}} > 1$ for all $\mathfrak{p} \in \mathbb{P}$.

In particular, we record the following simple fact for later use.

**Lemma 24.** *Any UFD admits at least one multiplicative norm.*

It is natural to ask whether there is a characterization of UFDs in terms of the existence of a multiplicative norm function of a certain type. The answer is affirmative and is due to C.S. Queen. We need the following notation: let $R$ be a domain with fraction field $K$, and let $I$ be a nonzero ideal of $R$. We put

$$(R : I) = \{x \in K \mid xI \subset R\}$$

and

$$\overline{I} = (R : (R : I)).$$

It is straightforward to check that $\overline{I}$ is an ideal of $R$ which contains $I$.

**Theorem 25.** *(Queen* [Q96]*) For a ring $R$, the following are equivalent:*
*(i) $R$ is a UFD.*
*(ii) $R$ admits a multiplicative norm $N$ with the following additional property: for all $a, b \in R$ with $a \nmid b$ and $b \nmid a$, there exists $0 \neq c \in \overline{Ra + Rb}$ with $N(c) < \min(N(a), N(b))$.*

We shall content ourselves with the following remarks: first, the ideal $\overline{Ra + Rb}$ is principal if and only if $a$ and $b$ admit a greatest common divisor in the sense of §6. In §6 we will see that in a UFD this is always the case. Using this observation it is easy to prove (i) $\implies$ (ii): we can take the norm with $n_{\mathfrak{p}} = 2$ for all $\mathfrak{p} \in \mathcal{P}$ and then the condition that $a \nmid b$ and $b \nmid a$ implies that their greatest common divisor $c$ properly divides both and therefore has strictly smaller norm.

We define a **Queen norm** to be a multiplicative norm satisfying the additional property (ii) of Theorem 25. Then the theorem may be restated as: a ring is a UFD iff it admits a Queen norm.

## 5. Polynomial and power series rings over UFDs

### 5.1. Polynomial rings.

**Lemma 26.** *Let $x$ be an element of a domain $R$.*
*a) The element $x$ is a unit in $R$ iff it is a unit as an element of $R[t]$.*
*b) The element $x$ is irreducible in $R$ iff it is irreducible as an element of $R[t]$.*
*b) The element $x$ is prime in $R$ iff it is prime as an element of $R[t]$.*

*Proof.* a) Units are mapped to units under any homomorphism of rings, so certainly if $x \in R$ is a unit, $x \in R[t]$ is a unit. Conversely, if there exists $y \in R[t]$ such that $xy = 1$, then taking degrees shows that $\deg y = 0$, i.e., $y \in R$.
b) If $x$ is irreducible in $R$, then by part a) $x$ is not a unit in $R[t]$. Moreover if $x = yz$ is a factorization in $R[t]$, then taking degrees shows that $y, z \in R$ and thus either $y$ or $z$ is a unit in $R$ and hence also in $R[t]$. Conversely, if $x = yz$ is a factorization in $R$ with $y, z$ nonunits, then it is also a factorization in $R[t]$ with $y, z$ nonunits.
c) It is immediate from the definition that if $x$ is prime as an element of $R[t]$ then it is prime as an element of $R$. Conversely, suppose $x$ is a prime element of $R$. Note first that for any polynomial $P = a_d t^d + \ldots + a_1 t + a_0 \in R[t]$, an element $c \in R$ divides $P$ iff $c \mid a_i$ for all $0 \leq i \leq d$. Now let $f = a_m t^m + \ldots + a_1 t + a_0$, $g = b_n t^n + \ldots + b_1 t + b_0$. Seeking a contradiction we shall suppose that $x \mid fg$ but $x \nmid f$ and $x \nmid g$. Let $I$ be the least index $i$ such that $x \nmid a_i$ and let $J$ be the least index $J$ such that $x \nmid b_j$. Then $x$ divides the coefficient of $t^{I+J}$ in $fg$, namely $c_{I+J} = \sum_{i+j=I+J} a_i b_j$. For all $(i, j) \neq (I, J)$ we have $i < I$ or $j < J$ and thus $x \mid a_i$ or $x \mid b_j$: either way $x \mid a_i b_j$. Therefore also $x$ divides the remaining term $a_I b_J$ and since $x$ is prime in $R$, $x \mid a_I$ or $x \mid b_J$, contradiction. □

**Theorem 27.** *(Hensel) If $R$ is a UFD, so is $R[t]$.*

We expect that most readers will have seen Theorem 27 and its proof via Gauss' Lemma on primitive polynomials. For the sake of variety, we will give here a "lemmaless" proof which is modelled on the Lindemann-Zermelo proof of FTA. (Later we will give another very striking proof, due to Nagata.) This argument,

with minor variations, appears several times in the literature. It seems that the first such instance is a paper of S. Borofsky [Bor50].

*Proof.* By Theorem 22, it suffices to show that $R[t]$ is an ACCP domain and an EL-domain. By Theorem 17, since $R$ is an ACCP domain, so is $R[t]$. Now, seeking a contradiction, we suppose that $R$ is an EL-domain but $R[t]$ is not. Among the set of all elements in $R[t]$ admitting inequivalent irreducible factorizations, let $p$ be one of minimal degree. We may assume

$$p = f_1 \cdots f_r = g_1 \cdots g_s,$$

where for all $i, j$, $(f_i) \neq (g_j)$ and

$$m = \deg f_1 \geq \deg f_2 \geq \ldots \geq \deg f_r > 0,$$

$$n = \deg g_1 \geq \deg g_2 \geq \ldots \geq \deg g_s > 0,$$

with $n \geq m > 0$. Indeed, by Lemma 26b), any irreducible element $x$ of degree 0 of $R[t]$ is irreducible in $R$. Since $R$ is a UFD, $x$ is prime in $R$ and by Lemma 26c) $x$ is prime in $R[t]$. But then $x$ is associate to one of the irreducible elements on the other side of the equation, so we may cancel them. Moreover if $r = s = 1$ then we have equal factorizations, whereas if exactly one of $r, s$ is equal to one then we have factored an irreducible element: thus $r, s > 1$. It follows that $\deg g < \deg p$.

Suppose the leading coefficient of $f_1$ (resp. $g_1$) is $a$ (resp. $b$). Put

$$q = ap - bf_1 t^{n-m} g_2 \cdots g_s = f_1(af_2 \cdots f_r - bt^{n-m} g_2 \cdots g_s) = (ag_1 - bf_1 t^{n-m}) g_2 \cdots g_s.$$

Thus $q = 0$ implies $ag_1 = bf_1 t^{n-m}$. If, however, $q \neq 0$, then

$$\deg(ag_1 - bf_1 t^{n-m}) < \deg g_1,$$

hence $\deg q < \deg p$ and $q$ has a unique factorization into irreducibles, certainly including $g_2, \cdots, g_s$ and $f_1$. But then $f_1$ must be a factor of $ag_1 - bf_1 t^{n-m}$ and thus also of $ag_1$. Either way $ag_1 = f_1 h$ for some $h \in R[t]$. Since $\deg(ag_1) = \deg g_1 < \deg p$, by induction the factorization of $ag_1$ into irreducibles is unique. It follows that $h = ah_2$, so $ag_1 = f_1 a h_2$, or $g_1 = f_1 h_2$, contradiction. □

By induction, we deduce:

**Corollary 28.** *Let $R$ be a UFD and $n \in \mathbb{Z}^+$. Then $R[t_1, \ldots, t_n]$ is a UFD.*

**Theorem 29.** *Let $R$ be a UFD and let $S = R[t_1, t_2, \ldots]$ be a polynomial ring over $R$ in infinitely many indeterminates. Then $R$ is a non-Noetherian UFD.*

*Proof.* We show $S$ is non-Noetherian by exhibiting an infinite chain of ideals:

$$\langle t_1 \rangle \subset \langle t_1, t_2 \rangle \subset \ldots \subset \langle t_1, \ldots, t_n \rangle \subset \ldots$$

Suppose that for any $n$, $t_{n+1}$ were an element of $\langle t_1, \ldots, t_n \rangle$. In other words, there exist polynomials $P_1, \ldots, P_n$ such that

$$t_{n+1} = P_1 t_1 + \ldots + P_n t_n.$$

Setting $t_1 = \ldots = t_0 = 0,\ t_{n+1} = 1$ gives $1 = 0$ in $R$, a contradiction.

By Theorem 22, to show that $R$ is a UFD it suffices to show that it satisfies the ascending chain condition on principal ideals and Euclid's Lemma. The first is almost immediate: any nonzero element is a polynomial in a finite number of variables, say $P(t_1, \ldots, t_n)$. Any divisor $Q$ of $P$ is again a polynomial in only the variables

$t_1, \ldots, t_n$, so that an ascending chain $(P) \subset (P_2) \subset \ldots \subset (P_n) \subset \ldots$ can be viewed as an ascending chain in the UFD $R[t_1, \ldots, t_n]$, so it stabilizes since UFDs satisfy ACCP. Finally, let $P$ be an irreducible element in $S$. The EL-condition is equivalent to the principal ideal $(P)$ being a prime ideal, which is equivalent to the quoteint $S/(P)$ being an integral domain. But as above $P$ is a polynomial in only finitely many variables, say $P(t_1, \ldots, t_n)$ and if $P(t_1, \ldots, t_n) = X(t_1, \ldots, t_n)Y(t_1, \ldots, t_n)$ with neither $X$ nor $Y$ a unit in $R[t_1, \ldots, t_n]$ then the factorization remains valid in the larger domain $S$, and since $S^\times = R[t_1, \ldots, t_n]^\times = R^\times$, it remains a nontrivial factorization (i.e., neither $X$ nor $Y$ is a unit in $S$). So $P(t_1, \ldots, t_n)$ is irreducible in $R_n := R[t_1, \ldots, t_n]$; since $R_n$ is a UFD, the principal ideal $PR_n$ is prime. But

$$S/PS = R_n[t_{n+1}, t_{n+2}, \ldots]/PR_n[t_{n+1}, t_{n+2}, \ldots] \cong (R_n/PR_n)[t_{n+1}, t_{n+2}, \ldots].$$

Since $(R_n/PR_n)[t_{n+1}, t_{n+2}, \ldots]$ is a domain, so is $S/PS$, so $PS$ is a prime ideal. □

The use of a countably infinite set of indeterminates in Theorem 29 was only a notational convenience: a similar argument shows the same result for any infinite set of indeterminates.

As an application, we can see that ACC is *not* a property that depends only on the group of divisibility. Indeed, the group of divisibility of a UFD is the direct sum of copies of $\mathbb{Z}$ indexed by the nonzero principal prime ideals. It is easy to see that for any cardinal $\kappa$, there exists a PID with exactly $\kappa$ nonzero principal prime ideals, hence the group of divisibility of a polynomial ring in infinitely many indeterminates is isomorphic to the group of divisibility of some PID.

### 5.2. **Power series rings.**

Let $R$ be a UFD, and consider the formal power series domain $R[[t]]$. By Theorems 17 and 22, $R[[t]]$ is an ACCP domain. But must $R[[t]]$ be a UFD?

In contrast to Theorem 27, whose proof essentially goes back to Gauss and thus predates the abstract ring concept, whether a formal power series ring over a UFD must be a UFD was a perplexing problem to 20th century algebraists and remained open for many years. Some special cases were known relatively early on.

**Theorem 30.** *(Rückert* [Rü33]*, Krull* [Kr37]*) Let $k$ be a field, and let $n$ be a positive integer. Then $k[[t_1, \ldots, t_n]]$ is a UFD.*

A significant generalization was proved by Buchsbaum and Samuel, independently, in 1961. We define the **height** of a prime ideal $\mathfrak{p}$ in a ring $R$ to be the supremum of all non-negative integers $n$ such that there exists a strictly ascending chain of prime ideals $\mathfrak{p}_0 \subsetneq \mathfrak{p}_1 \subsetneq \ldots \subsetneq \mathfrak{p}_n = \mathfrak{p}$. In a domain $R$, a prime ideal has height 0 iff it is the zero ideal. A Noetherian domain $R$ is **regular** if for every maximal ideal $\mathfrak{m}$ of $R$, the height of $\mathfrak{m}$ is equal to the dimension of $\mathfrak{m}/\mathfrak{m}^2$ as a vector space over the field $R/\mathfrak{m}$.

**Theorem 31.** *(*[Bu61]*,* [Sa61]*) If $R$ is a regular UFD, then so is $R[[t]]$.*

In the same 1961 paper, Samuel also provided the first example of a UFD $R$ for which $R[[t]]$ is *not* a UFD [Sa61, §4].

One may also ask for analogues of Theorem 29, i.e., about formal power series in countably infinitely many indeterminates. Let $k$ be a field. There is more than one

reasonable way to define such a domain. One the one hand, one could simply take the "union" (formally, direct limit) of finite formal power series rings $k[[t_1, \ldots, t_n]]$ under the evident inclusion maps. In any element of this ring, only finitely many indeterminates appear. However, it is useful also to consider a larger ring, in which the elements are infinite formal $k$-linear combinations of monomials $t_{i_1} \cdots t_{i_n}$. Let us call this latter domain $k[[t_1, \ldots, t_n, \ldots]]$.

In fact, we have seen this domain before: it is isomorphic to the Dirichlet ring $\mathcal{D}_k$ of functions $f : \mathbb{Z}^+ \to k$ with pointwise addition and convolution product. To see this, we use unique factorization in $\mathbb{Z}$! Namely, write enumerate the prime numbers $\{p_i\}_{i=1}^{\infty}$ and write $n \in \mathbb{Z}^+$ as $n = \prod_{i=1}^{\infty} p_i^{a_i}$, where $a_i \in \mathbb{Z}^+$ and $a_i = 0$ for all sufficiently large $i$. Then the map which sends $f \in \mathcal{D}_k$ to the formal power series $\sum_{n \in \mathbb{Z}^+} f(n) \prod_{i=1}^{\infty} t_i^{a_i}$ gives an isomorphism from $\mathcal{D}_k$ to $k[[t_1, \ldots, t_n, \ldots]]$. In 1959, E.D. Cashwell and C.J. Everett used Theorem 30 to prove the following result. A key part of their proof was later simplified by C.F. Martin, who pointed out the applicability of König's Infinity Lemma.

**Theorem 32.** *(*[CE59], [Mar71]*) a) For any field $k$, the ring of formal power series $k[[t_1, \ldots, t_n, \ldots]]$ is a UFD.*
*b) In particular, the ring $\mathcal{D}_{\mathbb{C}} = \{f : \mathbb{Z}^+ \to \mathbb{C}\}$ of airthmetic functions is a UFD.*

In almost any first number theory course one studies unique factorization and also arithmetic functions, including the Dirichlet ring structure (which e.g. leads to an immediate proof of the Möbius Inversion Formula). That arithmetic functions are themselves an example of unique factorization is however a very striking result that does not seem to be well-known to most students or practitioners of number theory. I must confess, however, that as a working number theorist I know of no particular application of Theorem 32. I would be interested to learn of one!

## 6. Greatest common divisors

We recall the definition of a greatest common divisor of two elements $a$ and $b$ in an arbitrary domain $R$. It is an element $d$ of $R$ which is a common divisor of $a$ and $b$ (i.e., $d \mid a$ and $d \mid b$) such that for all $e$ in $R$ with $e \mid a$ and $e \mid b$, we have $e \mid d$.

Of course it is not clear that such elements must exist. A **GCD-domain** is a domain in which any two elements admit at least one greatest common divisor.

Remark 6.1: Let $R$ be any integral domain.
a) If $a = 0$ and $b = 0$, then 0 is a greatest common divisor of $a$ and $b$.
b) If $a$ is arbitrary and $b = 0$, then $a$ is a greatest common divisor of $a$ and $b$.
c) If $a$ is a unit and $b$ is arbitrary, then 1 is a greatest common divisor of $a$ and $b$.

The uniqueness of greatest common divisors is easier to sort out:

**Lemma 33.** *Let $R$ be an integral domain, $a, b \in R$, and suppose $d$ is a greatest common divisor of $a$ and $b$. Then an element $x$ of $R$ is a greatest common divisor of $a$ and $b$ iff $x \sim d$, i.e., iff $x = ud$ for some unit $u \in R^\times$.*

*Proof.* Let $d$ and $d'$ be greatest common divisors of $a$ and $b$. Then $d \mid a$ and $d \mid b$, so $d \mid d'$, and similarly $d' \mid d$. It follows that $d \sim d'$. Conversely, since associate elements have exactly the same divisibility relations, it is clear that any associate of a greatest common divisor is again a greatest common divisor. □

Example 6.1: For two nonzero integers $a$ and $b$, there are two greatest common divisors: $d$ and $-d$. In $\mathbb{Z}$ it is conventional to mean by $\gcd(a,b)$ the unique positive greatest common divisor. However, in a general domain it is convenient to abuse notation slightly by writing $\gcd(a,b)$ for any greatest common divisor of $a$ and $b$, i.e., we tolerate ambiguity up to associate elements.

**Proposition 34.** *Let $R$ be a GCD-domain, $a,b,c \in R$; put $d = \gcd(a,b)$. Then:*
*a)* $\gcd(ab,ac) = a\gcd(b,c)$.
*b)* $\gcd(\frac{a}{d},\frac{b}{d}) = 1$.
*c) If* $\gcd(a,b) = \gcd(a,c) = 1$, *then* $\gcd(a,bc) = 1$.

*Proof.* a) Let $x = \gcd(ab,ac)$. Then $a \mid ab$ and $a \mid ac$ so $a \mid x$: say $ay = x$. Since $x \mid ab$ and $x \mid ac$, $y \mid b$ and $y \mid c$, so $y \mid \gcd(b,c)$. If $z \mid b$ and $z \mid c$, then $az \mid ab$ and $az \mid ac$, so $az \mid x = ay$ and $z \mid y$. Therefore $\gcd(b,c) = y = \frac{1}{a}\gcd(ab,ac)$. Part b) follows immediately. As for part c): suppose $\gcd(a,b) = \gcd(a,c) = 1$, and let $t$ divide $a$ and $bc$. Then $t$ divides $ab$ and $bc$ so $t \mid \gcd(ab,bc) = b\gcd(a,c) = b$. So $t$ divides $\gcd(a,b) = 1$. □

**Proposition 35.** *A GCD-domain is integrally closed in its fraction field.*

Proof: Let $R$ be a GCD-domain with fraction field $K$, and let $\alpha$ be an element of $K$ which satisfies a relation of the form $\alpha^n + a_{n-1}\alpha^{n-1} + ldots + a_1\alpha + a_0 = 0$ with $a_0,\dots,a_{n-1} \in R$. We may write $\alpha = \frac{r}{s}$ with $r,s \in R$, and we may also assume – and this is the crux! – that $\gcd(r,s) = 1$. (Take any representation of $\alpha$ as a quotient of two elements of $R$, and divide numerator and denominator by their gcd.) Then we need only substitute in $\alpha = \frac{r}{s}$ and clear denominators to get

$$r^n + sa_{n-1}r^{n-1} + \ldots + s^{n-1}a_1 r + s^n a_0 = 0,$$

or

$$r^n = -s\left(a_{n-1}r^{n-1} + a_{n-2}sr^{n-2} + \ldots + s^{n-1}a_0\right),$$

so $s \mid r^n$. Since $\gcd(r,s) = 1$, Proposition 34c) implies $\gcd(r^n,s) = 1$. Thus $s$ is a unit, so $\alpha = \frac{r}{s} \in R$.

**Proposition 36.** *A unique factorization domain is a GCD-domain.*

*Proof.* This is an immediate generalization of the usual arguments for $R = \mathbb{Z}$. By Remark 6.1, we know that in any domain, the GCD of $a$ and $b$ necessarily exists except possibly when both $a$ and $b$ are nonzero nonunits. Then, let $x_1,\dots,x_r$ be the set of pairwise nonassociate irreducibles such that any irreducible divisor of either $a$ or $b$ is associate to some $x_i$; we may then write

$$a = x_1^{a_1}\cdots x_r^{a_r},\ b = x_1^{b_1}\cdots x_r^{b_r},$$

with $a_i, b_i \in \mathbb{N}$. Then

$$d = x_1^{\min(a_1,b_1)}\cdots x_r^{\min(a_r,b_r)}$$

is a greatest common divisor of $a$ and $b$. □

Propositions 35 and 36 imply that a UFD is integrally closed.

**Proposition 37.** *A GCD-domain is an EL-domain.*

*Proof.* Suppose $x$ is irreducible and $x \mid yz$. Assume, for a contradiction, that $x \nmid y$ and $x \nmid z$. Then $\gcd(x,y) = \gcd(x,z) = 1$, and by Proposition 34c), $\gcd(x,yz) = 1$, which contradicts $x \mid yz$. □

**Corollary 38.** *A factorization domain is a UFD iff it is a GCD-domain.*

*Proof.* Let $R$ be a factorization domain. Assume first that $R$ is a UFD. Then $R$ is a GCD-domain by Proposition 36. Conversely, assume that $R$ is a GCD-domain. Then it is an EL-domain by Proposition 37, and by Theorem 22 a factorization domain which is an EL-domain is a UFD. □

## 7. GCDs versus LCMs

The definition of GCDs in a domain has an evident analogue for least common multiples. Namely, if $a$ and $b$ are elements of a domain $R$, a **least common multiple** of $a$ and $b$ is an element $l$ such that for all $m \in R$ with $a \mid m$ and $b \mid m$ then $l \mid m$.

Many of the properties of GCD's carry over immediately to LCM's. For instance, if $l$ is an LCM of $a$ and $b$, then $l' \in R$ is an LCM of $a$ and $b$ iff $l'$ is associate to $l$.

**Proposition 39.** *Let $a$ and $b$ be elements in a domain $R$. Then $\operatorname{lcm}(a,b)$ exists iff the ideal $(a) \cap (b)$ is principal, in which case the set of all LCM's of $a$ and $b$ is the set of all generators of $(a) \cap (b)$.*

*Proof.* This is straightforward and left to the reader. □

LCM's exist in any UFD: if

$$a = x_1^{a_1} \cdots x_r^{a_r},\ b = x_1^{b_1} \cdots x_r^{b_r},$$

with $a_i, b_i \in \mathbb{N}$. Then

$$l = x_1^{\max(a_1,b_1)} \cdots x_r^{\max(a_r,b_r)}$$

is a greatest common divisor of $a$ and $b$. Now the simple identity

$$\forall a, b \in \mathbb{N}, \min(a,b) + \max(a,b) = a + b$$

implies that for $a, b$ in any UFD $R$ we have

$$\gcd(a,b)\operatorname{lcm}(a,b) \sim ab.$$

This identity further suggests that the existence of either one of $\gcd(a,b)$, $\operatorname{lcm}(a,b)$ implies the existence of the other. However, this turns out only to be half correct!

**Theorem 40.** *For $a, b$ in a domain $R$, the following are equivalent:*
*(i) $\operatorname{lcm}(a,b)$ exists.*
*(ii) For all $r \in R \setminus \{0\}$, $\gcd(ra,rb)$ exists.*

*Proof.* Step 1: i) $\implies$ (ii). Suppose that there exists a least common multiple of $a$ and $b$, say $l$. We claim that $d := \frac{ab}{l}$ is a greatest common divisor of $a$ and $b$. (Note that since $ab$ is a common divisor of $a$ and $b$, $l \mid ab$, so indeed $d \in R$.) Indeed, suppose that $e \mid a$ and $e \mid b$. Then since $\frac{ab}{e}$ is a common multiple of $a$ and $b$, we must have $l \mid \frac{ab}{e}$ and this implies $e \mid \frac{ab}{l}$. Thus $d$ is a GCD of $a$ and $b$.
Step 2: Suppose that for $r \in R \setminus \{0\}$ and $a, b \in R$, $\gcd(ra,rb)$ exists. Then we claim that $\gcd(a,b)$ exists and $\gcd(ra,rb) = r\gcd(a,b)$. Put $g := \frac{\gcd(ra,rb)}{r}$, which is clearly an element of $D$. Since $\gcd(ra,rb)$ divides $ra$ and $rb$, $g$ divides $a$ and $b$. Conversely, if $e \mid a$ and $e \mid b$, then $re \mid ra$ and $re \mid rb$ so $er \mid \gcd(ra,rb)$ and $e \mid g$.
Step 3: We claim that if $l := \operatorname{lcm}(a,b)$ exists then so does $\operatorname{lcm}(ra,rb)$ for all $r \in R \setminus \{0\}$. First note that $rl$ is a common multiple of $ra$ and $rb$. Now suppose $m$ is a common multiple of $ra$ and $rb$, say $m = xra = yrb = r(xa - yb)$. Thus $r \mid m$

and $a \mid \frac{m}{r}$, $b \mid \frac{m}{r}$. So $l \mid \frac{m}{r}$ and $rl \mid m$. Thus $\operatorname{lcm}(ra, rb) = r \operatorname{lcm}(a, b)$.
Step 4: (ii) $\implies$ (i). We may assume that $a$ and $b$ are nonzero, since the other cases are trivial. Suppose $\gcd(ra, rb)$ exists for all $r \in R \setminus \{0\}$. We claim that $l := \frac{ab}{\gcd(a,b)}$ is an LCM of $a$ and $b$. Clearly $l$ is a common multiple of $a$ and $b$. Now suppose that $m$ is a common multiple of $a$ and $b$. Then $ab$ divides both $ma$ and $mb$, so $ab \mid \gcd(ma, mb)$. By Step 2, $\gcd(ma, mb) = m \gcd(a, b)$. Thus $\frac{ab}{\gcd(a,b)} \mid m$. □

**Theorem 41.** *(Khurana* [Kh03, Thm. 4]*) Let $d \geq 3$ be an integer such that $d + 1$ is not prime, and write $d + 1 = pk$ for a prime number $p$ and $k \geq 2$. Then in the domain $R = \mathbb{Z}[\sqrt{-d}]$, the elements $p$ and $1 + \sqrt{-d}$ have a GCD but no LCM.*

*Proof.* Step 1: We claim that $p$ is irreducible as an element of $R$. Indeed, if it were reducible, then by the multiplicativity of the norm map $N(a + b\sqrt{-d}) = a^2 + dp^2$ we could write it as $p = \alpha\beta$, with

$$p^2 = N(p) = N(\alpha\beta) = N(\alpha)N(\beta),$$

and, since $\alpha, \beta$ are nonunits, $N(\alpha), N(\beta) > 1$. But then $N(\alpha) = N(\beta) = p$, i.e., there would be $a, b \in \mathbb{Z}$ such that $a^2 + db^2 = p$. But this is not possible: either $ab = 0$, in which the left hand side is a perfect square, or $a^2 + db^2 \geq d + 1 > p$.
Step 2: $\gcd(p, 1 + \sqrt{-d}) = 1$. Indeed, since $\frac{1}{p} + \frac{1}{p}\sqrt{-d} \notin R$, $p \nmid 1 + \sqrt{-d}$.
Step 3: We claim that $kp$ and $k(1+\sqrt{-d})$ do not have a GCD. Indeed, by Step 2 of the proof of Theorem 40, if any GCD exists then $k$ is a GCD. Then, since $1 + \sqrt{-d}$ divides both $(1 - \sqrt{-d})(1 + \sqrt{-d}) = 1 + d = kp$ and $k(1 + \sqrt{-d})$, $1 + \sqrt{-d}$ divides $\gcd(kp, k(1 + \sqrt{-d}) = k$, i.e., there exist $a, b \in \mathbb{Z}$ such that

$$k = (1 + \sqrt{-d})(a + b\sqrt{-d}) = (a - db) + (a + b)\sqrt{-d},$$

i.e., $a = -b$ and $k = a - db = a + da = a(1 + d)$ and $d + 1 \mid k$, contradicting the fact that $1 < k < d + 1$.
Step 4: Finally, it follows from Theorem 40 that $\operatorname{lcm}(p, 1+\sqrt{-d})$ does not exist. □

Khurana produces similar examples even when $d + 1$ is prime, which implies that for no $d \geq 3$ is $R_d = \mathbb{Z}[\sqrt{-d}]$ a GCD-domain. (In fact, since $(R_d, +) \cong \mathbb{Z}^2$, $R_d$ is an abstract number ring and hence Noetherian, so the notions of EL-domain, GCD-domain and UFD are all equivalent.) Let us give an independent proof:

**Theorem 42.** *For no $d \geq 3$ is $R_d = \mathbb{Z}[\sqrt{-d}]$ an EL-domain.*

*Proof.* As in the proof of Theorem 41 above, the easy observation that the equation $a^2 + db^2 = 2$ has no integral solutions implies that the element 2 is irreducible in $R_d$. Now, since (quite trivially) $-d$ is a square modulo 2, there exists $x \in \mathbb{Z}$ such that $2 \mid x^2 + d = (x + \sqrt{-d})(x - \sqrt{-d})$. But now, if $R_d$ were an EL-domain, the irreducible element 2 would be prime and hence Euclid's Lemma would apply to show that $2 \mid x \pm \sqrt{-d}$, i.e., that $\frac{x}{2} + \frac{1}{2}\sqrt{-d} \in R_d$, a clear contradiction ($\frac{1}{2} \notin \mathbb{Z}$!). □

Note that Theorem 40 has the following immediate consequence:

**Corollary 43.** *(Cohn,* [Coh68, Thm. 2.1]*) For an integral domain $R$, TFAE:*
*(i) Any two elements of $R$ have a greatest common divisor.*
*(ii) Any two elements of $R$ have a least common multiple.*

Thus we need not define an "LCM-domain": these are precisely the GCD domains.

Once again these concepts can be pithily reexpressed in terms of $\mathrm{Prin}(R)$ and $G(R)$. For $x, y \in R^\bullet$, $\gcd(x,y)$, if it exists, is a well-defined element in $\mathrm{Prin}(R)$, its **meet** $x \wedge y$. Similarly, in $\mathrm{Prin}(R)$ $\mathrm{lcm}(x,y)$ is the **join** $x \vee y$. Thus Theorem 43 says that a domain is a GCD-domain iff $\mathrm{Prin}(R)$ is a lattice iff it is a meet semi-lattice iff it is a join semi-lattice.

## 8. More on Principal Ideal Domains

### 8.1. PID implies UFD.

**Theorem 44.** *(Bézout's Lemma) Let $a$ and $b$ be elements in a PID $R$. Then $d = \gcd(a,b)$ exists and moreover can be expressed as a linear combination of $a$ and $b$: there exist $x, y \in R$ such that*

$$ax + by = d.$$

*Proof.* The ideal $\langle x, y \rangle = \{xa + yb \mid x, y \in R\}$ is by assumption principal, i.e., equal to $(d)$ for some $d \in R$. As in the case $R = \mathbb{Z}$, we see easily that $d$ is a greatest common divisor of $a$ and $b$: it is a common divisor since $x, y \in \langle x, y \rangle = (d)$, and if $e \mid a, e \mid b$, then $e \mid ax + by$. But $ax + by = d$, so $e \mid d$. □

**Corollary 45.** *Every PID is a UFD.*

*Proof.* Once again contemplation of the proof of FTA carries us straight through. Namely, a PID is Noetherian and hence satisfies ACCP. By Bézout's Lemma, a PID is a GCD-domain, so we can carry out Euclid's proof of Euclid's Lemma in this context. In fact this is not necessary, since we have already seen that Euclid's Lemma holds in any GCD-domain: Proposition 37. Finally, we apply the "classical" Theorem 21 to deduce that every PID is a UFD. □

### 8.2. Bézout Domains.

In the distinguished tradition of giving a name to any property that appears as a conclusion of a theorem, we define a **Bézout domain** to be an integral domain in which the gcd of any two elements exists and may be expressed as a linear combination of them. In particular, a Bézout domain is a GCD-domain.

It is easy to see that a domain is a Bézout domain iff every finitely generated ideal is principal. In particular, a Noetherian Bézout domain is a PID. In fact this only begins to describe the yearning that a Bézout domain has to be a PID: much weaker finiteness conditions suffice.

**Theorem 46.** *For a Bézout domain $R$, the following are equivalent:*
*(i) $R$ is a PID.*
*(ii) $R$ is Noetherian.*
*(iii) $R$ is a UFD.*
*(iv) $R$ is an ACCP domain.*
*(v) $R$ is a factorization domain.*

*Proof.* We have just mentioned that (i) $\iff$ (ii). That (i) $\implies$ (iii) is Corollary 45. As we well know, for any domain $R$, (iii) $\implies$ (iv) $\implies$ (v). (v) $\implies$ (iii): A Bézout domain is a GCD-domain is an EL-domain, so a Bézout factorization domain is a UFD by Theorem 21. Thus (iii) $\iff$ (iv) $\iff$ (v). (iv) $\implies$ (ii): assume that $R$ is *not* Noetherian. Then it admits an ideal $I$ which is not finitely

generated, which we can use to build an infinite strictly ascending chain of finitely generated ideals $I_1 \subsetneq I_2 \subsetneq \ldots \subsetneq I$. But since $R$ is Bézout, each $I_i$ is principal, contradicting ACCP. □

Recall that a commutative ring is **local** if it has a unique maximal ideal. This happens iff the nonunits of $R$ form an ideal iff they form a commutative group.

**Theorem 47.** *For an integral domain $R$, the following are equivalent:*
*(i) $R$ is a valuation ring.*
*(ii) $R$ is a local Bézout domain.*

*Proof.* Let $R$ be a valuation ring. Let $x, y \in R \setminus R^\times$. Clearly $-x, -y$ are nonunits. Moreover, we have either $x \mid y$ or $y \mid x$, so that $x + y$ is divisible by at least one of the nonunits $x$ and $y$ and is thus not a unit. So $R$ is a local ring. Now let $I = \langle x_1, \ldots, x_n \rangle$ be a finitely generated ideal of $R$. Since $\mathrm{Prin}(R)$ is totally ordered, there exists at least one $i$ such that in the divisibility ordering, $x_i \leq x_j$ for all $1 \leq j \leq n$, i.e., $x_i \mid x_j$. Thus $I = \langle x_i \rangle$. So $R$ is a Bézout domain.

Conversely, let $R$ be a local Bézout domain, and let $x, y \in R^\bullet$. Put $d = \gcd(x, y)$, $x' = \frac{x}{d}$ and $y' = \frac{y}{d}$. Then $x'$ and $y'$ are relatively prime elements of a Bézout domain, so there exist $a, b \in R$ such that $x'a + y'b = 1$, i.e., $R = \langle x', y' \rangle$. But in a local ring any collection of nonunits generate a proper ideal, so that at least one of $x'$ and $y'$ must be a unit, i.e., $x \mid y$ or $y \mid x$. □

**Corollary 48.** *Let $R$ be a domain such that the group of divisibility $G(R)$ is isomorphic, as a partially ordered commutative group, to the integers $\mathbb{Z}$ with the standard ordering. Then $R$ is a local PID.*

*Proof.* Since $G(R)$ is totally ordered, $R$ is a valuation ring. Moreover, since $\mathrm{Prin}(R) \cong (\mathbb{N}, \leq)$ is well-ordered, $R$ satisfies ACCP. Applying Theorem 46 and Theorem 47, we conclude that $R$ is a local PID. □

A PID with a unique maximal ideal is called a **discrete valuation ring** (or **DVR**). Excepting only fields, DVR's have the simplest structure of all integral domains. Therefore a recurrent technique in commutative algebra is to, somehow, study more complicated domains by reducing to the case of DVRs. We will see an important instance of this in §11.2.

On the other hand, we saw in §4 that for any torsionfree commutative group $\Gamma$, there exists an ordering on $\Gamma$ and a valuation ring $R$ with $G(R) \cong \Gamma$, and that so long as $\Gamma$ is not cyclic, $R$ is not an ACCP domain. Thus non-Noetherian Bézout domains exist in abundance.

There are also non-Noetherian Bézout domains with infinitely many maximal ideals. Two striking examples of such are the ring $\overline{Z}$ of algebraic integers [Ka70, Thm. 102] and the ring of entire functions on the complex plane [Hel40].

### 8.3. **Dedekind-Hasse norms.**

We wish to give a criterion for an integral domain to be a PID which is due to R. Dedekind and (later, but independently) H. Hasse. In fact, the Dedekind-Hasse criterion is in terms of a multiplicative norm $N$ on $R$ which satisfies one additional property.

First, consider any multiplicative norm $N : R \to \mathbb{N}$ on an integral domain $R$. We assert that because of the multiplicativity, there is a unique extension of $N$ to a function from the fraction field, $K$, of $R$ to the non-negative rational numbers such that $N(xy) = N(x)N(y)$ for all $x, y \in K$. Indeed, since axiom (MN2) implies $N(1) = 1$, we must have $N(\frac{1}{y}) = \frac{1}{N(y)}$ and thus

$$N\left(\frac{x}{y}\right) = \frac{N(x)}{N(y)}.$$

Since a given element of $K$ has many different representations as a quotient of elements of $R$, we must check that the definition of $N$ is independent of this representation, but this is easy: if $\frac{x_1}{y_1} = \frac{x_2}{y_2}$, then $x_1y_2 = x_2y_1$, so

$$N(x_1)N(y_2) = N(x_1y_2) = N(x_2y_1) = N(x_2)N(y_1),$$

and, since $y_1, y_2 \neq 0$ implies $N(y_1), N(y_2) \neq 0$, we may divide in $\mathbb{Q}$ to get

$$\frac{N(x_1)}{N(y_1)} = \frac{N(x_2)}{N(y_2)}.$$

For example, the usual absolute value $z \mapsto |z|$ on $\mathbb{Z}$ extends multiplicatively to the usual absolute value on $\mathbb{Q}$.

From now on, we will assume without comment that a multiplicative norm has its domain extended to the fraction field $F$ of $R$ as above.

A multiplicative norm $N : F \to \mathbb{Q}$ on the fraction field of an integral domain $R$ is a **Dedekind-Hasse norm** (c.f. [Has28]) if it satisifes the following property:

(HN) For all $x \in F \setminus R$, there exist $a, b \in R$ such that $0 < N(ax - b) < 1$.

Example 8.1: The usual absolute value on $\mathbb{Z}$ is a Dedekind-Hasse norm. Indeed, for any rational number $x$ which is not an integer, we can take $a = 1$ and take $b$ to be $\lfloor x \rfloor$, the greatest integer less than or equal to $x$. Then $0 < x - b < 1$.

**Theorem 49.** *(Hasse,* [Has28]*) For an integral domain $R$, TFAE:*
*(i) $R$ admits a Dedekind-Hasse norm.*
*(ii) $R$ is a PID.*

*Proof.* (i) $\implies$ (ii): Let $N$ be a Dedekind-Hasse norm on $R$, and let $I$ be a nonzero ideal of $R$. Then $I$ contains elements of positive norm. Let $d \in I$ be an element whose norm is positive and is minimal among all elements of $I$. We wish to show that $I = (d)$. So let $i$ be any element of $I$ and put $x := \frac{i}{d}$. If $d \mid i$ then $x \in R$, so assume for a contradiction that $x \in F \setminus R$. Then by assumption there exist $a, b \in R$ such that

$$0 < N\left(\frac{ai}{d} - b\right) < 1.$$

Multiplying through by $d$ we get

$$0 < N(ai - bd) < N(d).$$

So $ai - bd \in I$ has norm positive and smaller than $N(d)$, contradiction!
(ii) $\implies$ (i): Suppose $R$ is a PID. By Corollary 45, $R$ is a UFD, and thus by Lemma 24 $R$ admits a multiplicative norm, say $N$. Let $K$ be the fraction field of

the UFD $R$. Then any $x \in K \setminus R$ can be written as $x = \frac{p}{q}$, where $p, q \in R \setminus \{0\}$, $\gcd(p,q) = 1$ and $q$ is a nonunit, so $N(q) > 1$. Now, applying Proposition 45, we can find elements $a$, $b'$ in $R$ such that $ap + b'q = 1$. Taking $b = -b'$, we have $ap - bq = 1$. Dividing through by $q$ we get $ax - b = \frac{1}{q}$, so

$$0 < N(ax - b) = N(\frac{1}{q}) = \frac{1}{N(q)} < 1.$$

□

Here is an application (a familiar one, but recast in slightly different language).

**Proposition 50.** *Let $F$ be any field. Then the polynomial ring $F[t]$ is a PID.*

*Proof.* Every nonzero polynomial $P(t) = a_n t^n + \ldots + a_0$ has a degree $\deg(P)$ , which is the largest $n \in \mathbb{N}$ such that $a_n \neq 0$. Let us agree that the 0 polynomial has degree $-\infty$. It is easy to check that $\deg(PQ) = \deg(P) + \deg(Q)$. Thus the degree is very much like a norm, only instead of being multiplicative, it is multiplicative-to-additive. That can be remedied, however: put $N(P) = 2^{\deg P}$, with the convention that $N(0) = 2^{-\infty} = 0$. One easily checks that $N$ is a Dedekind-Hasse norm. □

8.4. **Euclidean norms.**

A multiplicative norm $N$ on a domain $R$ is **Euclidean** if for any $a, b \in R$ with $b \neq 0$, there exist $q, r \in R$ such that $a = qb + r$ and $N(r) < N(b)$.

**Proposition 51.** *Any Euclidean norm is a Hasse norm.*

*Proof.* Let $x = \frac{a}{b} \in F \setminus R$. Since the norm is Euclidean, there exist $q, r \in R$ with $N(r) < N(b)$, and then $x - q = \frac{a}{b} - (\frac{a}{b} - \frac{r}{b}) = \frac{r}{b}$, so

$$0 < N(x - q) = N\left(\frac{r}{b}\right) < 1.$$

□

We repackage this result in a form which will be most convenient in number-theoretic applications:

**Proposition 52.** *Let $R$ be an integral domain with fraction field $K$. Suppose that $R$ has a multiplicative norm $N$ with the property that for all $x \in K \setminus R$ there exists $y \in R$ with $N(x - y) < 1$. Then $R$ is a PID.*

8.5. **Case Study I: Quadratic Rings.**

Let $d$ be an integer such that $-d$ is not a square, and consider again the quadratic ring $R_d = \mathbb{Z}[t]/(t^2 + d) = \mathbb{Z}[\sqrt{-d}]$ with fraction field $K_d := \mathbb{Q}(\sqrt{-d})$. This is an abstract number ring, so comes equipped with a canonical norm, in this case equal to $N(a + b\sqrt{d}) = |a^2 - db^2|$. It is an easy exercise to show that this canonical norm is a Euclidean norm if $d = -2, -1, 2, 3$. Indeed, write a general element of $K = \mathbb{Q}(\sqrt{-d})$ as $x = \alpha + \sqrt{-d}\beta$, with $\alpha, \beta \in \mathbb{Q}$. Vhoose $a, b \in \mathbb{Z}$ such that $|\alpha - a|, |\beta - b| \leq \frac{1}{2}$ and put $y = a + \sqrt{-d}b$. Then

$$N(x - y) = |(\alpha - a)^2 - d(\beta - b)^2| < 1.$$

Thus each of the rings $\mathbb{Z}[\sqrt{3}], \mathbb{Z}[\sqrt{2}], \mathbb{Z}[\sqrt{-1}], \mathbb{Z}[\sqrt{-2}]$ is a PID, hence also a UFD.

In contrast, if $d \geq 3$, we saw above that $R_d$ is *not* a PID. However, if $d \equiv 3$

(mod 4), there is an excuse: $R_d$ is not even integrally closed in its fraction field $K_d$. Rather, the ring of integers $\mathbb{Z}_{K_d}$ is $S_d := \mathbb{Z}[\frac{1+\sqrt{-d}}{2}]$. Reasoning as above, it is not hard to show that, for a positive integer $d \equiv 3 \pmod 4$, the canonical norm on $S_d$ is Euclidean iff $d = 3, 7, 11$.

Consider now the case of $d = 19$. We claim that the standard norm $N$ on $S_{19} = \mathbb{Z}[\frac{1+\sqrt{-19}}{2}]$, while not Euclidean, *is* a Dedekind-Hasse norm, so that $S_{19}$ is a PID. We follow [MRR88, Ex. 3.4]: for $\alpha, \beta \in S_{19}^\bullet$, we will find $\theta \in S_{19}$ such that *either*

$$N(\frac{\beta}{\alpha} - \theta) < 1$$

or

$$N(\frac{2\beta}{\alpha} - \theta) < 1.$$

This will show that $N$ is a Dedekind-Hasse norm. We may choose $\theta \in S_{19}$ such that

$$\frac{\beta}{\alpha} - \theta = a + b\sqrt{-19}$$

with $a, b \in \mathbb{Q}$, $|a| \leq \frac{1}{2}$, $|b| \leq \frac{1}{4}$.
Case 1: $|b| \leq \frac{3}{16}$. Then $N(\frac{\beta}{\alpha} - \theta) \leq \frac{235}{256} < 1$, okay.
Case 2: $\frac{3}{16} < b \leq \frac{1}{2}$. Then we may choose $\theta' \in S_{19}$ such that

$$\frac{2\beta}{\alpha} - \theta = a' + b'\sqrt{-19}$$

with $|a'| \leq \frac{1}{2}$, $|b'| \leq \frac{1}{8}$, so $N(\frac{2\beta}{\alpha} - \theta') \leq \frac{35}{64} < 1$. In this case though we have to contemplate the possibility that $N(\frac{2\beta}{\alpha} - \theta') = 0$, i.e., that $\alpha \nmid \beta$ but $\alpha \mid 2\beta$. However, the quotient ring $S_{19}/2S_{19}$ is a domain – recall that this is a finite ring of order $N(2) = 4$, so this is easy to check – so that 2 is a prime element of $S_{19}$. So if $\alpha\delta = 2\beta$, so that either $2 \mid \delta$ – in which case $\alpha \mid \beta$ – or $2 \mid \alpha$, in which case we may work instead with $\frac{\alpha}{2}$ and $\frac{\beta}{2}$, so we are done by induction.

By similar means, it can be shown that the canonical norm on $S_d$ is a Dedekind-Hasse norm for $d = 43, 67, 163$. The following celebrated result, conjectured by Gauss, says that we should look no farther among positive values of $d$.

**Theorem 53.** *(Heegner-Baker-Stark* [Hee52] [Ba67] [St67]*) Let $K_d = \mathbb{Q}(\sqrt{-d})$ be an imaginary quadratic field. Then the ring of integers of $K_d$ is a PID iff $d = 1, 2, 3, 7, 11, 19, 43, 67, 163$.*

The possibility remains that the rings $\mathbb{Z}_{K_d}$ for $d = 19, 43, 67, 163$ are Euclidean with respect to some other norm. Note that it is not at all obvious how to show that a domain is not Euclidean with respect to *any* norm, except to show that it is not a PID! The task is to find – as with Dedekind-Hasse and Queen norms – some intrinsic property equivalent to, or at least implied by, the existence of any Euclidean norm. T. Motzkin did just thus in a brilliant 1949 paper [Mo49]. We will not describe his conditions explicitly, but only record the following consequence.

**Theorem 54.** *(Motzkin* [Mo49]*) For $d = 19, 43, 67, 163$, the quadratic ring $S_d = \mathbb{Z}[\frac{1+\sqrt{-d}}{2}]$ is a PID which is not Euclidean with respect to any norm.*

Thus, for every imaginary quadratic field whose ring of integers is a PID, the canonical norm is a Dedekind-Hasse norm. Was this just good fortune? Not at all.

**Theorem 55.** *Let $R$ be an abstract number ring. Then $R$ is a PID iff the canonical norm is a Dedekind-Hasse norm.*

*Proof.* If the canonical norm on $R$ is a Dedekind-Hasse norm, then by Theorem 49 $R$ is a PID. Conversely, suppose that $R$ is a PID. Glancing back at Theorem 49, we see that what we in fact showed is that *any* multiplicative norm on a PID is a Dedekind-Hasse norm. QED! □

The moral seems to be as follows: the question of whether a domain admits a Euclidean norm can be significantly more subtle than whether it is a PID. In fact, even for the case of rings of integers of number fields, the state of affairs is more intricate than the results presented so far would suggest.

First, we must mention that in contrast to imaginary quadratic fields, it is unproved but widely believed that there are infinitely many squarefree $a > 0$ such that the ring of integers of $\mathbb{Q}(\sqrt{a})$ is a PID. Moreover, restricting e.g. to prime numbers $a > 0$, significant computational data (as well as a probabilistic heuristic due to Cohen and Lenstra that we do not wish to discuss here) supports the belief that a positive proportion of these rings are PIDs.

**Theorem 56.**
*a)* [BSD52] *The real quadratic number fields for which the ring of integers is Euclidean for the standard norm are precisely $\mathbb{Q}(\sqrt{a})$ for*

$$a \in \{2, 3, 5, 6, 7, 11, 13, 17, 19, 21, 29, 33, 37, 41, 57, 73\}.$$

*b)* [W72] *Suppose the Generalized Riemann Hypothesis holds, and let $K$ be a number field which is* not *an imaginary quadratic field. Then the ring of integers $\mathbb{Z}_K$ of $K$ is a PID iff it admits a Euclidean norm.*
*c)* [HM04] *The conclusion of part b) holds unconditionally if $K/\mathbb{Q}$ is Galois and $[K : \mathbb{Q}] \geq 9$.*

The first unconditional example of a real quadratic field with a ring of integers which is Euclidean but not Euclidean with respect to the standard norm was given by D. Clark in 1994 [Cl94]. Clark explicitly (and with significant computer-assisted calculations) constructs an "exotic" Euclidean norm on the ring of integers of $\mathbb{Q}(\sqrt{69})$.

## 9. Localization

Throughout this section, $R$ denotes an integral domain with fraction field $K$.

Our main goal in this section is to present a theorem of Nagata (Theorem 62). One application is a second proof of the fact that if $R$ is a UFD so is $R[t]$.

### 9.1. **Localization in domains.**

If $(M, \cdot)$ is a commutative monoid and $S$ is a subset of $M$, define $\langle S \rangle$ to be the submonoid generated by $S$. This can be described either as the intersection of all submonoids of $M$ which contain $S$, or more explicitly as the set of all finite products $x_1 \cdots x_n$ with $x_i \in S$ (including the empty product, so that always $1 \in \langle S \rangle$).

Now let $R$ be our domain with fraction field $K$, and let $M$ be the monoid of

nonzero elements of $R$ under multiplication. A subset $S$ of $M$ is **multiplicatively closed** (or just **multiplicative**) if $S = \langle S \rangle$. Moreover, a subset $T$ of $S$ is a **set of generators** for $S$ if $\langle T \rangle = S$.

For any subset $S$ of $R \setminus \{0\}$, we define $R_S = R[\{\frac{1}{x} \mid x \in S\}]$, i.e., the subring of the fraction field obtained by adjoining to $R$ all the multiplicative inverses of elements of $R$. We say that $R_S$ is the **localization of R at S**. It is easy to see that $R_S = R_{\langle S \rangle}$, so that it is no loss of generality to restrict to localizations of multiplicatively closed sets. Note that we recover $K$ itself as $R_{R \setminus \{0\}}$.

**Theorem 57.** *Let $R$ be a UFD and $S$ a multiplicative subset. Then the localized ring $R_S$ is again a UFD.*

*Proof.* Let $f = \frac{x}{s}$ be a nonzero nonunit of $R_S$, with $x \in R$ and $s \in S$. Then $x$ is a nonzero nonunit in the UFD $R$, so admits a factorization into prime elements

$$x = \prod_{i=1}^{n} \pi_i^{a_i}.$$

We may assume the ordering is such that $\pi_i \in S$ for $1 \leq i \leq m$ and $\pi_i \in R \setminus S$ for $m < i \leq n$. Then for $1 \leq i \leq m$, $\pi_i \in R_S^\times$, whereas by Lemma X.X, for $i > m$, $\pi_i$ remains prime in $R_S$. Therefore

$$\frac{x}{s} = \left( \frac{\prod_{i=1}^{m} \pi_i^{a_i}}{s} \right) \cdot \pi_{m+1}^{a_{m+1}} \cdots \pi_n^{a_n}$$

expresses $\frac{x}{s}$ as a unit times a product of prime elements. □

9.2. **Saturated subsets.**

A multiplicative subset $S$ is **saturated** if for all $x \in S$ and $y \in R$, if $y \mid x$ then $y \in S$. We define the **saturation** $\overline{S}$ of a multiplicatively closed subset $S$ to be the intersection of all saturated multiplicatively closed subsets containing $S$; equivalently, $\overline{S}$ is obtained from $S$ simply by throwing in all nonzero divisors of all elements of $S$. If $x \in S$ and $y \mid x$, then $ay = x$ for some $a \in R$, and then $\frac{1}{y} = \frac{a}{ay} = a \cdot (\frac{1}{x})$. Thus $R_S = R_{\overline{S}}$, so that we may restrict attention to saturated multipicative sets.

Example 9.1: Any saturated multiplicative subset of $R$ contains $R^\times$. In particular, if $R$ is a field the unique saturated multiplicative subset is $R \setminus \{0\}$.

Example 9.2: If $\mathfrak{p} \subset R$ is a prime ideal, $R \setminus \mathfrak{p}$ is a saturated multiplicative set.

**Proposition 58.** *Let $R$ be a domain and $S \subset R$ a multiplicative set. Then the set of units of $R_S$ is precisely the saturation of the multiplicative set $S$.*

*Proof.* This is straightforward and left to the reader. □

**Proposition 59.** *Let $R$ be a domain, $S$ a saturated multiplicative subset, and $f \in R \setminus S$. If $f$ is prime as an element of $R$, it is also prime as an element of $R_S$.*

*Proof.* Since $f \in R \setminus S$, by Proposition 58 $f$ is not a unit in $R_S$. Let $\alpha, \beta \in R_S$ be such that $f \mid \alpha\beta$ in $R_S$. So there exists $\gamma \in R_S$ such that $\gamma f = \alpha\beta$; putting $\alpha = \frac{x_1}{s_1}$, $\beta = \frac{x_2}{s_2}$, $\gamma = \frac{x_3}{s_3}$ and clearing denominators, we get $s_1 s_2 x_3 f = s_3 x_1 x_2$, so $f \mid r_3 x_1 x_2$. If $f \mid s_3$, then since $S$ is saturated, $f \in S$, contradiction. So, being

prime, $f$ divides $x_1$ or $x_2$ in $R$, hence *a fortiori* in $R_S$ and therefore it also divides either $\frac{x_1}{s_1}$ or $\frac{x_2}{s_2}$ in $R_S$, since these are associates to $x_1$ and $x_2$. □

9.3. **Primal subsets.**

We say that a saturated multiplicative subset $S$ of $R$ is **primal** if it is generated by the units and by the prime elements of $S$.

**Lemma 60.** *An irreducible element of a primal subset is prime.*

*Proof.* Suppose $S$ is primal and $f \in S$ is irreducible. By definition, there exists a unit $u$ and prime elements $\pi_1, \ldots, \pi_n$ such that $f = u\pi_1 \cdots \pi_n$. Since $u\pi_1$ is also prime, we may as well assume that $u = 1$. Then, since $f$ is irreducible, we must have $n = 1$ and $f = \pi_1$. □

**Theorem 61.** *For a factorization domain $R$, the following are equivalent:*
*(i) Every saturated multiplicative subset of $R$ is primal.*
*(ii) $R$ is a UFD.*

*Proof.* Since the set $R^\times$ of units is trivially generated by the empty set of prime elements, both conditions hold if $R$ is a field, so let us now assume otherwise.

Assume (i). Then, since $R$ is a factorization domain which is not a field, there exists an irreducible element $f$ of $R$. Let $S$ be the saturated multiplicative subset generated by $S$, which consists of all units of $R$ together with all divisors of positive powers $f^n$ of $f$. Since $S$ is primal and strictly contains $R^\times$, there must exist a prime element $\pi$ which divides $f^n$ for some $n$. In other words, $f^n \in \pi R$, and since $\pi R$ is prime, we must have that $f = x\pi$ for some $x \in R$. Since $f$ is irreducible we must have $x \in R^\times$, i.e., $f \sim \pi$ and is therefore a prime element. So $R$ is an ACCP domain and an EL-domain and hence a factorization domain by Theorem 22.

Assume (ii), let $S$ be a saturated multiplicative subset of $R$, and suppose that $f \in S \setminus R^\times$. Then $f = u\pi_1^{a_1} \cdots \pi_n^{a_n}$ where the $\pi_i$'s are prime elements. Since each $\pi_i \mid f$, $\pi_i \in S$ for all $i$. It follows that indeed $S$ is generated by its prime elements together with the units of $R$. □

Because of Theorem 61, it is no loss of generality to restate Theorem 57 as: the localization of a UFD at a primal subset is again a UFD. The following elegant result of Nagata may be viewed as a converse.

**Theorem 62.** *(Nagata's Criterion* [Na57]*) Let $R$ be a factorization domain and $S \subset R$ a primal subset. If the localized domain $R_S$ is a UFD, then so is $R$.*

*Proof.* By Theorem 22 it suffices to show that if $f \in R$ is irreducible, $f$ is prime. Case 1: $f \notin S$, so $f$ is not a unit in $R_S$. Since $R_S$ is a UFD, it is enough to show that $f$ is irreducible in $R_S$. So assume not: $f = \frac{x_1}{s_1} \cdot \frac{x_2}{s_2}$ with $x_1, x_2 \in R \setminus S$ and $s_1, s_2 \in S$. hen $s_1 s_2 f = x_1 x_2$. By assumption, we may write $s_1 = up_1 \cdots p_m$ and $s_2 = vq_1 \cdots q_n$, where $u, v \in R^\times$ and $p_i, q_j$ are all prime elements of $R$. So $p_1 \mid x_1 x_2$; since $p_1$ is a prime, we must have either $\frac{x_1}{p_1} \in R$ or $\frac{x_2}{q_2} \in R$. Similarly for all the other $p_i$'s and $q_j$'s, so that we can at each stage divide either the first or the second factor on the right hand side by each prime element on the left hand side, without leaving the ring $R$. Therefore we may write $f = (\frac{1}{uv})\frac{x_1}{t_1}\frac{x_2}{t_2}$ where $t_1, t_2$ are each products of the primes $p_i$ and $q_j$, hence elements of $S$, and also such that $t_1 \mid x_1$, $t_2 \mid x_2$, i.e., the factorization takes place in $R$. Moreover, since $x_i \in R \setminus S$ and

$t_i \in S$, $\frac{x_i}{t_i}$ is not even a unit in $R_S$, hence *a fortiori* not a unit in $R$. Therefore we have exhibited a nontrivial factorization of $f$ in $R$, contradiction.
Case 2: $f \in S$. Since $S$ is primal, by Lemma 60 $f$ is prime. □

Remark: If $S$ is the saturation of a finitely generated multiplicative set, the hypothesis that $R$ is a factorization domain can be omitted.

Application: Let $A$ be a UFD and consider the polynomial ring $R = A[t]$. Put $S = A \setminus \{0\}$. As for any multiplicative subset of a UFD, $S$ is generated by prime elements. But moreover, since $A[t]/(\pi A[t]) \cong (A/\pi A)[t]$, every prime element $\pi$ of $A$ remains prime in $A[t]$, so viewing $S$ as the multiplicative subset of $A[t]$ consisting of nonzero constant polynomials, it too is generated by prime elements. If $F$ is the fraction field of $A$, $R_S = (A[t])_S = F[t]$ is a PID (Proposition 50) and hence a UFD. Applying Nagata's Criterion, we deduce once more that $R = A[t]$ is a UFD.

9.4. **Case study II: affine quadric cones.**

Let $k$ be a field of characteristic different from 2, and let $f(x) = f(x_1, \ldots, x_n) \in k[x_1, \ldots, x_n]$ be a **quadratic form**, i.e., a homogeneous polynomial of degree 2 with $k$ coefficients. We assume that $f$ the associated bilinear form $(x, y) \mapsto \frac{1}{2}(f(x+y) - f(x) - f(y))$ is nonsingular. Equivalently, by making an invertible linear change of variables every quadratic form can be diagonalized, and a quadratic form is nonsingular iff it admits a diagonalization

$$f(x) = a_1 x_1^2 + \ldots + a_n x_n^2 \text{with} a_1, \ldots, a_n \in k^\times. \tag{6}$$

We wish to study the **affine quadric cone** associated to $f$, namely $R_f = k[x]/(f)$. Note that if quadratic forms $f$ and $g$ are isometric – i.e., differ by an invertible linear change of variables – then $R_f \cong R_g$, so we assume if we like that $f$ is in diagonal form as in (6) above. If $n \geq 3$ then every nonsingular diagonal quadratic polynomial is irreducible, so $R_f$ is a domain. If $k$ is quadratically closed – i.e., admits no proper quadatic extension – then conversely any binary ($n = 2$) quadratic form is reducible, so $R_f$ is not a domain. (If $f$ is not quadratically closed, there exist irreducible binary quadratic forms, but we will not consider them here.)

**Theorem 63.** *Let $f = f(x_1, \ldots, x_n) \in \mathbb{C}[x_1, \ldots, x_n]$ be a nonsingular quadratic form. Then $R_f = k[x]/(f)$ is a UFD iff $n \geq 5$.*

*Proof.* By the remarks above, $R_f$ is a domain iff $n \geq 3$, so we may certainly restrict to this case. Because $\mathbb{C}$ is algebraically closed, every quadratic form in $n \geq 2$ variables is **isotropic**, i.e., there exists $0 \neq a \in k^n$ such that $f(a) = 0$: indeed, the first $n-1$ coordinates of $a$ may be chosen arbitrarily. By an elementary theorem in the algebraic theory of quadratic forms [Lam05, Thm. I.3.4], we may make a change of variables to bring $f$ into the form:

$$f(x) = x_1 x_2 + g(x_3, \ldots, x_n).$$

Case 1: Suppose $n = 3$, so that

$$f(x) = x_1 x_2 - a x_3^2$$

for some $a \in k^\times$. In this case, to show that $R_f$ is not a UFD it suffices to show that the images $\overline{x_1}, \overline{x_2}, \overline{x_3}$ of $x_1, x_2, x_3$ in $R_f$ are nonassociate irreducibles, for then $\overline{x_1}\overline{x_2} = a\overline{x_3}^2$ exhibits a non-unique factorization! To establish this, regard

$k[x_1, x_2, x_3]$ as a graded $\mathbb{C}$-algebra in the usual way – with $x_1, x_2, x_3$ each of degree 1 – so that the quotient $R_f$ by the homogeneous ideal $(f)$ inherits a grading. Since $\overline{x_1}$ has degree 1, if it were reducible, it would factor as the product of a degree one element $c_1x_1 + c_2x_2 + x_3x_3 + (f)$ and a degree zero element $r + (f)$, and thus

$$(rc_1 - 1)x_1 + rc_2x_2 + rc_3x_3 \in (f).$$

But the left hand side has degree 1, whereas all nonzero elements in $(f)$ have degree 2 or higher, so $r \in \mathbb{C}[x]^\times$ and therefore the factorization is trivial. The irreducibility of $\overline{x_2}$ and $\overline{x_3}$ is proved in the same way. If $\overline{x_1} \sim \overline{x_3}$ in $R_f$, then we may divide both sides of $\overline{x_1x_2} - a\overline{x_3}^2$ by $\overline{x_1}$ and deduce that also $\overline{x_2} \sim \overline{x_3}$. But in the quotient ring $R_f/(\overline{x_3})$, $\overline{x_3}$ maps to 0 and $\overline{x_1}$ and $\overline{x_2}$ do not, contradiction. So $R_f$ is not a UFD.
Case 2: Suppose $n = 4$, so $f(x) = x_1x_2 + g(x_3, x_4)$, where $g(x_3, x_4)$ is a nonsingular binary form. Here for the first time we use the full strength of the quadratic closure of $k$: since $k^\times = k^{\times 2}$, any two nonsingular quadratic forms in the same number of variables are isometric, so we may assume WLOG that

$$f(x) = x_1x_2 - x_3x_4.$$

Now we argue exactly as in Case 1 above: in $R_f$, the images $\overline{x_1}, \overline{x_2}, \overline{x_3}, \overline{x_4}$ are all non-associate irredcuble elements, so $\overline{x_1x_2} = \overline{x_3x_4}$ is a non-unique factorization.
Case 3: $n \geq 5$. Then $n - 2 \geq 3$, so $g$ is irreducible in the UFD $\mathbb{C}[x_3, \ldots, x_n]$, hence also in $\mathbb{C}[x_2, x_3, \ldots, x_n]$. Therefore $R_f/(\overline{x_1}) = \mathbb{C}[x_1, \ldots, x_n]/(x_1, f) = \mathbb{C}[x_2, \ldots, x_n]/(g)$ is a domain, i.e., $\overline{x}_1$ is a prime element. Moreover,

$$R[\overline{x_1}^{-1}] = \mathbb{C}[x_1, \ldots, x_n, x_1^{-1}]/(x_1x_2 - g)$$
$$\cong \mathbb{C}[x_1, \ldots, x_n, x_1^{-1}]/(x_2 - \frac{g}{x_1}) \cong \mathbb{C}[x_1, x_3, \ldots, x_n, x_1^{-1}]$$

is a localization of the UFD $\mathbb{C}[x_1, x_3, \ldots, x_n]$ hence a UFD. By Nagata's Criterion (Theorem 62), $R_f$ itself is a UFD. □

Now let $k$ be an arbitrary field of characteristic not 2 and $f \in k[x_1, \ldots, x_n]$ a nonsingular quadratic form. Without changing the isomorphism class of $R_q$ we may diagonalize $f$; moreover without changing the ideal $(f)$ we may scale by any element of $k^\times$, so without loss of generality we need only consider forms $x_1^2 + a_2x_2^2 + \ldots + a_nx_n^2$.

**Theorem 64.** *Let $k$ be a field of characteristic different from 2 and $f = x_1^2 + a_2x_2^2 + \ldots + a_nx_n^2$ a nonsingular quadratic form over $k$. Put $R_f = k[x]/(f)$.*
*a) If $n \leq 2$ then $R_f$ is not an integrally closed domain.*
*b) If $n = 3$, $R_f$ is a UFD iff $f$ is anistropic: $\forall a \in k^n,\ f(a) = 0 \implies a = 0$.*
*c) (i) Suppose $f = x_1^2 - ax_2^2 - bx_3^2 - cx_4^2$. If $a$ is a square in $k$, then $R_f$ is a UFD iff $-bc$ is not a square in $k$.*
*(i) If none of $a$, $b$, $c$, $-ab$, $-ac$, $-bc$ is a square in $k$, then $R_f$ is a UFD iff $-abc$ is not a square.*
*d) If $n \geq 5$, $R_f$ is a UFD.*

Remark: If $f(x_1, \ldots, x_4)$ s a diagonal quadratic form, we may permute and/or rescale the coefficients so that either condition (i) or condition (ii) of part c) holds.

*Proof.* a) It is not hard to show that if $n \leq 2$, $R_f$ is never an integrally closed domain. b) The proof of Theorem 63 goes through to show that if $f$ is isotropic (i.e., not anisotropic), $R_f$ is not a UFD. The anisotropic case is due to Samuel [Sa64]. Part c) is due to T. Ogoma [O74]. Part d) goes back at least to van der Waerden [vdW39]. In [Na57], M. Nagata gives a short proof using Theorem 62. □

It is also interesting to consider affine rings of inhomogeneous quadric hypersurfaces. For instance, we state without proof the following result.

**Theorem 65.** *For $n \geq 1$, let $R_n := \mathbb{R}[t_1, \ldots, t_{n+1}]/(t_1^2 + \ldots + t_{n+1}^2 - 1)$ be the ring of polynomial functions on the $n$-sphere $S^n$.*
*a) (Bouvier* [Bou78]*) If $n \geq 2$, then $R_n$ is a UFD.*
*b) (Trotter* [T88]*) $R_1$ is isomorphic to the ring $\mathbb{R}[\cos\theta, \sin\theta]$ of real trigonometric polynomials, in which $(\sin\theta)(\sin\theta) = (1+\cos\theta)(1-\cos\theta))$ is an explicit non-unique factorization into irreducible elements. Hence $R_1$ is not a UFD.*

## 10. Characterizations of UFDs

**Lemma 66.** *(Multiplicative avoidance) Let $R$ be a commutative ring and $S \subset (R \setminus \{0\}, \cdot)$ a multipicatively closed subset containing 1. Let $\mathcal{I}_S$ be the set of ideals of $R$ which are disjoint from $S$. Then:*
*a) $\mathcal{I}_S$ is nonempty.*
*b) Every element of $\mathcal{I}_S$ is contained in a maximal element of $\mathcal{I}_S$.*
*c) Every maximal element of $\mathcal{I}_S$ is prime.*

*Proof.* a) $(0) \in \mathcal{I}_S$. b) Let $I \in \mathcal{I}_S$. Consider the subposet $P_I$ of $\mathcal{I}_S$ consisting of ideals which contain $I$. Since $I \in P_I$, $P_I$ is nonempty; moreover, any chain in $P_I$ has an upper bound, namely the union of all of its elements. Therefore by Zorn's Lemma, $P_I$ has a maximal element, which is clearly also a maximal element of $\mathcal{I}_S$. c) Let $I$ be a maximal element of $\mathcal{I}_S$; suppose that $x, y \in R$ are such that $xy \in I$. If $x$ is not in $I$, then $\langle I, x\rangle \supsetneq I$ and therefore contains an element $s_1$ of $S$, say

$$s_1 = i_1 + ax.$$

Similarly, if $y$ is not in $I$, then we get an element $s_2$ of $S$ of the form

$$s_2 = i_2 + by.$$

But then

$$s_1 s_2 = i_1 i_2 + (by)i_1 + (ax)i_2 + (ab)xy \in I \cap S,$$

a contradiction. □

**Theorem 67.** *(Kaplansky) An integral domain is a UFD iff every nonzero prime ideal in $R$ contains a prime element.*

*Proof.* Suppose $R$ is a UFD and $0 \neq P$ is a prime ideal. Let $x \in P$ be a nonzero nonunit. Write

$$x = p_1 \cdots p_r$$

a product of prime elements. Then $x \in P$ implies $p_i \in P$ for some $i$, so $(p_i) \subset P$.

Conversely, assume that each nonzero prime ideal of $R$ contains a principal prime. Let $S$ be the set of units of $R$ together with all products of prime elements. One checks easily that $S$ is a saturated multiplicative subset. We wish to show that $S = R \setminus \{0\}$. Suppose then for a contradiction that there exists a nonzero nonunit $x \in R \setminus S$. The saturation of $S$ implies that $S \cap (x) = \varnothing$, and then by Lemma 66 there exists a prime ideal $P$ containing $x$ and disjoint from $S$. But by hypothesis, $P$ contains a prime element $p$, contradicting its disjointness from $S$. □

Corollary 45 is an immediate consequence of Kaplansky's Theorem. Moreover we can derive a criterion for a UFD to a PID, as follows. Define the **dimension** of a ring to be the supremum of all heights of prime ideals.

**Theorem 68.** *For a UFD $R$, the following are equivalent:*
*(i) $R$ is a PID.*
*(ii) $R$ has dimension one, i.e., every nonzero prime ideal is a maximal ideal.*

*Proof.* (i) $\implies$ (ii): Any integral domain which is not a field has nonzero prime ideals so therefore dimension at least one. It suffices to show that in a PID every nonzero prime ideal $\mathfrak{p}$ is maximal. But if not, there exists a prime ideal $\mathfrak{q}$ such that $\mathfrak{p} \subsetneq \mathfrak{q}$. But every ideal is principal, so there exist $p, q \in R$ such that $\mathfrak{p} = (p)$ and $\mathfrak{q} = (q)$. Therefore $p \mid q$, but since $q$ is a prime, and thus irreducible, element, $p$ must be associate to $q$, so that $\mathfrak{p} = \mathfrak{q}$, contradiction.

(ii) $\implies$ (i): Suppose $R$ is a UFD in which each nonzero prime ideal is maximal, and let $\mathfrak{p}$ be a nonzero prime ideal of $R$. By Theorem 67 $\mathfrak{p}$ contains a prime element $p$, so that we have a containment of prime ideals $0 \neq (p) \subset (\mathfrak{p})$. By hypothesis $(p)$ must be maximal, so $(p) = \mathfrak{p}$ and $\mathfrak{p}$ is principal. □

**Corollary 69.** *An abstract number ring is a UFD iff it is a PID.*

*Proof.* We claim that any abstract number ring $R$ has dimension one; in view of Theorem 68, this suffices. So let $0 \neq \mathfrak{p}$ be a prime ideal of $R$. By definition, $R/\mathfrak{p}$ is a finite integral domain, and therefore a field, so $\mathfrak{p}$ is in fact maximal. □

What about the case of dimension greater than one? The following is perhaps the most natural and useful characterization of UFDs among Noetherian domains.

**Theorem 70.** *A Noetherian domain is a UFD iff every height one prime is principal.*

*Proof.* This is a consequence of Krull's *Hauptidealsatz*, which lies beyond the scope of this article. See [Ei95, Cor. 10.6] or [Mat89, Thm. 20.1]. □

Theorem 70 is of basic importance in elementary algebraic geometry. Applied to the polynomial ring $R = k[t_1, \ldots, t_n]$ over a field $k$, it states that any irreducible hypersurface in affine $n$-space is the zero locus of a single polynomial equation.

## 11. The Class Group

Let me now draw the reader's attention to a sobering fact. We have analyzed the definition of UFD and found several "characterizations", but nevertheless most of what we have discussed is of relatively little use in determining whether a concretely given integral domain (say of dimension greater than one) is a UFD. To an extent we simply need to accept that being a UFD is a rather delicate and subtle property – we hope our examples of quadratic rings and affine quadric surfaces have attested to that.

Still, there is one more perspective we wish to give, which is useful both theoeretically and computationally. Let $R$ be an integral domain. To fix ideas, let us assume for now that it is Noetherian, so in particular a factorization domain. Are there quantities that we may compute in order to determine whether $R$ is a UFD?

Recall that a necessary condition for $R$ to be a UFD is that it be integrally closed in its fraction field. For a large class of domains – e.g. domains which are finitely generated over $\mathbb{Z}$, over $\mathbb{Q}$, or over $\mathbb{F}_p$ – algorithms have been implemented which check whether $R$ is integrally closed (and more generally, compute the integral closure). So let us restrict our attention to integrally closed domains.

11.1. **The ideal class group of a Dedekind domain.**

As usual, the simplest case is when $R$ has dimension one. Then a Noetherian, integrally closed domain of dimension one is a **Dedekind domain**, one of the best studied classes of rings. These include, for instance, any integrally closed abstract number ring. Dedekind domains have (indeed, are characterized by) many pleasant properties. Among them is a unique factorization property – but at the level of ideals, not elements. Namely, a domain is a Dedekind domain iff for every nonzero, proper ideal $I$ of $R$, there exist nonzero prime ideals $\mathfrak{p}_1, \ldots, \mathfrak{p}_r$ such that

$$I = \mathfrak{p}_1 \cdots \mathfrak{p}_r.$$

(Because the ideals are assumed to be *prime* and not just unfactorable, the uniqueness of the factorization follows exactly as in the classical case.) In other words, a domain is Dedekind iff the monoid $\mathcal{I}(R)$ of nonzero ideals of $R$ is a free commutative monoid on the nonzero prime ideals. (The identity element is the improper ideal $R$.) Now, inside $\mathcal{I}(R)$ we have $\mathrm{Prin}(R)$, the submonoid of principal ideals. Because $R$ is one-dimensional, it is a UFD iff it is a PID iff $\mathrm{Prin}(R) = \mathcal{I}(R)$. This suggests that we can *quantify* the failure of unique factorization in a Dedekind domain by forming the quotient monoid

$$\mathrm{Cl}(R) := \mathcal{I}(R)/\mathrm{Prin}(R).$$

More explicitly, we define a equivalence relation $\sim$ on $\mathcal{I}(R)$ by $I \sim J$ if there exist $(a), (b) \in \mathrm{Prin}(R)$ such that $(a)I = (b)J$. This relation is compatible with the monoid strcture on $\mathcal{I}(R)$, so the quotient is a monoid. For any domain $R$, the quotient $\mathcal{I}(R)/\mathrm{Prin}(R)$ is a monoid which is trivial iff $R$ is a PID. But this works out especially nicely for Dedekind domains: in fact a domain $R$ is a Dedekind domain iff $\mathcal{I}(R)/\mathrm{Prin}(R)$ is a group, or in other words, that for any nonzero ideal $I$ of $R$, there exists a nonzero ideal $J$ of $R$ such that $IJ$ is principal. The commutative group $\mathrm{Cl}(R)$ is the **ideal class group** of the Dedekind domain $R$.

The class group of a Dedekind domain is an invariant of independent interest and usefulness. Especially, in the case in which $R$ is the ring of integers of a number field $K$, number theorists write simply $\mathrm{Cl}(K)$ for $\mathrm{Cl}(\mathbb{Z}_K)$ out of sheer familiarity: we have over over 100 years of experience studying these groups. For a number field $K$, let $\Delta_K$ be the discrminant of the ring of integers, i.e., the determinant of the trace bilinear form $(x, y) \mapsto \mathrm{Tr}_{K/\mathbb{Q}}(xy)$ with respect to any $\mathbb{Z}$-basis of $\mathbb{Z}_K$. Let $2r_2$ be the number of embeddings $\iota : K \hookrightarrow \mathbb{C}$ such that $\iota(K)$ is not contained in $\mathbb{R}$.

**Theorem 71.** *Let $K$ be a number field.*
*a) (Dedekind) The ideal class group $\mathrm{Cl}(K)$ is a finite commutative group.*
*b) (Minkowski) Every class of $\mathrm{Cl}(K)$ may be represented by an ideal $I$ with*

$$||I|| \leq \sqrt{|\Delta_K|} \left(\frac{4}{\pi}\right)^{r_2} \frac{n!}{n^n}.$$

Note that Lemma 4 shows that b) implies a) and also gives an effective algorithm for computing $\mathrm{Cl}(K)$. This brings up the following:

**Question 1.** *Let $R$ be an abstract number ring. Must $\mathcal{I}(R)/\mathrm{Prin}(R)$ be finite?*

Coming back to the case of rings of integers of number fields, Theorem 71 is a 19th century result. Nowadays there are mathematical software packages which implement much faster algorithms to compute class groups of number fields. For instance, I asked MAGMA to compute the class groups of $\mathbb{Q}(\sqrt{-1000000000000000123})$ and $\mathbb{Q}(\sqrt{100055128505716009})$ and keep track of the time:

`time(ClassGroup(QuadraticField(-1000000000000000123)));`
It took 5.810 seconds to compute that the class group is isomorphic to $\mathbb{Z}/6\mathbb{Z} \times \mathbb{Z}/25131282\mathbb{Z}$.

`time(ClassGroup(QuadraticField(100055128505716009)));`
It took 2.530 seconds to compute that the class group is trivial, i.e., that the ring of integers $R$ of $\mathbb{Q}(\sqrt{100055128505716009})$ is a UFD. This software makes the methods of §8 involving explicit verification that the canonical norm is a Dedekind-Hasse norm look quaint indeed.

### 11.2. **The Picard group and the divisor class group.**

Let $R$ be an integral domain which is *not* a Dedekind domain. Then the monoid $\mathcal{I}(R)/\operatorname{Prin}(R)$ is not a group. Now there are (at least!) two ways to make a commutative monoid $M$ into a commutative group. The first is to take the group completion $G(M)$. However, if $M$ is not cancellative, the homomorphism $M \hookrightarrow G(M)$ is not injective and could well be trivial: i.e., much information is lost. For instance, let $R$ be a non-maximal order in a number field $K$. Then it is known that $\mathcal{I}(R)/\operatorname{Prin}(R)$ is finite. But a finite cancellative commutative monoid is necessarily a group! Thus for a nonmaximal order, $\mathcal{I}(R)/\operatorname{Prin}(R)$ is never cancellative. The other option is simply to restrict to the group of units $M^\times$ of $M$, i.e., the invertible elements. In our case this amounts to taking *invertible* ideals $I$ of $R$ modulo principal ideals. By definition, this is the **Picard group** $\operatorname{Pic}(R)$ of $R$.

We note in passing that a nonzero ideal $I$ of a domain $R$ is invertible iff it is, as an $R$-module, finitely generated, projective and of rank one. That is, $I$ is invertible iff the correspondening quasi-coherent sheaf $\tilde{I}$ on the affine scheme $\operatorname{Spec} R$ is a **line bundle**. Thus this is a special case of the Picard group $\operatorname{Pic} X = H^1(X, \mathcal{O}_X^\times)$ of a scheme. Since one can always pullback line bundles, this suggests that the Picard group should have pleasant functorial properties. And indeed, if $R \to S$ is a homomorphism of domains, then for every invertible ideal $I$ of $R$, $IS$ is an invertible ideal of $S$: indeed, if $J$ is such that $IJ = (a)$, then $JS$ is such that $IS \cdot JS = aS$.

The Picard group of a domain $R$ is certainly an important construction in algebra and number theory. In particular, the Picard group of a nonmaximal order of a quadratic field is intimately related to the study of binary quadratic forms over $\mathbb{Q}$: see [Cox89]. Recall however that we are interested in computing an algebraic invariant of a Noetherian integrally closed domain that will tell us whether it is a UFD. We have seen that in dimension one, $\operatorname{Pic} R = \operatorname{Cl} R$ does what we want. However, in dimension greater than one it turns out that a subtly different group is the right one for the job.

**Proposition 72.** *The ring $D := \mathbb{C}[x, y, z]/(xy - z^2)$ is a Noetherian, integrally closed domain with $\operatorname{Pic} D = 0$ which is not a UFD.*

*Proof.* (Sketch) By Theorem 63, $D$ is not a UFD. As for the integrally closed condition: a geometer would argue that $D$ is singular only at the origin, so its singular locus has codimension $2 > 1$ and that $D$ is a hypersurface, hence a local complete intersection and that this implies "normality" (i.e., that $D$ is integrally closed) [Har77, Prop. II.8.3]. A more elementary algebraic argument can be given: in [Sa58] it is proven that if $k$ is a field, $n$ is a positive integer not divisible by the characteristic of $n$, $f = f(x_1, \ldots, x_n) \in k[x_1, \ldots, x_n]$ has no irreducible factor of multiplicity $n$ or greater, then $k[x_1, \ldots, x_n, y]/(f - y^n)$ is integrally closed. (See also [Mat89, Example 4, p. 65] for a variant which also applies in our situation.) A proof that $\operatorname{Pic} D = 0$ can be found in [Har77, Example II.6.11.3]. □

In fact the proof of Proposition 72 is best understood via some ideas that we are about to introduce, so we will return to it later and be able to give a better (though still not complete) explanation.

Let $R$ be an integrally closed Noetherian domain with fraction field $K$, and let $\Sigma$ be the set of height one prime ideals of $R$. We define the **divisor group** $\operatorname{Div}(R) = \bigoplus_{\mathfrak{p} \in \Sigma} \mathbb{Z}$ to be the free commutative group on $\Sigma$. Also put $\operatorname{Div}^+(R) = \bigoplus_{\mathfrak{p} \in \Sigma} \mathbb{N}$, the subgroup of effective divisors. For each $\mathfrak{p} \in \Sigma$, the localization $R_{\mathfrak{p}}$ is an integrally closed one-dimensional local Noetherian domain, hence a discrete valuation ring [Mat89, Thm. 11.2]. We have a canonical monoid homomorphism

$$R^\bullet \hookrightarrow R_{\mathfrak{p}}^\bullet \to G(R_{\mathfrak{p}}^\bullet) \xrightarrow{\sim} \mathbb{Z}$$

into a group $\mathbb{Z}$, so it extends uniquely to a group homomorphism on the group completion, say

$$\operatorname{ord}_{\mathfrak{p}} : K^\times \to \mathbb{Z}.$$

For $f \in R^\bullet$, we define the **principal divisor**

$$\operatorname{div} f := ((v_{\mathfrak{p}})_{\mathfrak{p} \in \Sigma}.$$

Since $\operatorname{Div}(R)$ is defined to be the direct sum and not the direct product, for this definition to make sense, we must have that for $f \in R^\bullet$, there are only finitely many height one primes containing $f$. This is equivalent to the fact that the quotient ring $R/(f)$ has only finitely many minimal prime ideals. But since $R$ is Noetherian, so is $R/(f)$ and indeed every Noetherian ring has only finitely many minimal primes.

The map $\operatorname{div} : R^\bullet \to \operatorname{Div}(R)$ is easily seen to be a homomorphism of commutative monoids which factors through $\operatorname{Prin}(R)$. Moreover, since $R$ is an integrally closed Noetherian domain, we have $R = \bigcap_{\mathfrak{p} \in \Sigma} R_{\mathfrak{p}}$ [Ei95, Cor. 11.4], it follows that $\operatorname{div} : \operatorname{Prin}(R) \to \operatorname{Div}(R)$ is *injective*. Once again we get an induced homomorphism on the group completion, $\operatorname{div} : G(R) \to \operatorname{Div}(R)$. Note that $\operatorname{div}^{-1}(\operatorname{Div}^+(R)) = \operatorname{Prin}(R)$. Finally, we define the **divisor class group**

$$\operatorname{Cl}(R) = \operatorname{Div}(R)/\operatorname{div}(G(R)).$$

**Theorem 73.** *Let $R$ be an integrally closed Noetherian domain. TFAE:*
*(i) $R$ is a UFD.*
*(ii)* $\operatorname{Div}^+(R) = \operatorname{Prin}(R)$.
*(iii)* $\operatorname{Cl}(R) = 0$.

*Proof.* (i) $\implies$ (ii): By Theorem 67, every $\mathfrak{p} \in \Sigma$ is principal, say $\mathfrak{p} = (f_{\mathfrak{p}})$. It is easy to see that $\operatorname{div} f_{\mathfrak{p}} = [\mathfrak{p}]$. Therefore the submonoid generated by effective principal divisors is all of $\operatorname{Div}^+(R)$.

(ii) $\implies$ (iii): $\mathrm{Cl}(R) = \mathrm{Div}(R)/G(R)$ is the group completion of the quotient monoid $\mathrm{Div}^+(R)/\mathrm{Prin}(R)$, which we are assuming is trivial.
(iii) $\implies$ (i): Let $\mathfrak{p} \in \Sigma$. Then there exists $x \in K^\times$ such that $\mathfrak{p} = \operatorname{div} x$. Since $\mathfrak{p} \in \mathrm{Div}^+(R)$, $x \in R^\bullet$, i.e., $\mathfrak{p} = (x)$. By Theorem 67, $R$ is a UFD. □

For any integrally closed Noetherian domain $R$, we define a homomorphism $\Phi : \operatorname{Pic} R \to \operatorname{Cl} R$. To do this, let $I$ be an invertible ideal of $R$ and let $\mathfrak{p} \in \Sigma$. Then $IR_{\mathfrak{p}}$ is an ideal in the DVR $R_{\mathfrak{p}}$, hence principal. Therefore we may define the $\mathfrak{p}$-component of $\Phi(I)$ to be $v_{\mathfrak{p}}(IR_{\mathfrak{p}}) \in \mathbb{N}$. As above, since $R$ is Noetherian, we have the $\mathfrak{p}$-component of $I$ is equal to zero except for finitely many primes $\mathfrak{p}$. Thus we get a monoid homomorphism from the monoid of invertible ideals to the divisor group. Moreover, the divisor attached to a principal ideal $(x)$ in this way is indeed $\operatorname{div} x$. Thus the homomorphism factors through a group homomorphism $\Phi : \operatorname{Pic} R \to \operatorname{Cl} R$.

**Theorem 74.** *Let $R$ be an integrally closed Noetherian domain.*
*a) There exists a canonical injective group homomorphism $\Phi : \operatorname{Pic} R \hookrightarrow \operatorname{Cl} R$.*
*b) $\Phi$ is an isomorphism iff $R$ is locally factorial, i.e., the localization of $R$ at each maximal ideal is a UFD.*

*Proof.* See Theorems 11.8 and 11.10 of [Ei95]. We explain only the necessity of the local factoriality condition in part b): $\operatorname{Pic} R$ consists of locally principal ideals, so is certainly trivial when $R$ is a local domain. □

Example ($D = \mathbb{C}[x,y,z]/(xy - z^2)$ revisited): Let $\mathfrak{p} = \langle \overline{y}, \overline{z} \rangle$ be the ideal of $D$ generated by the images of $y$ and $z$. Then $D/\mathfrak{p} = \mathbb{C}[x]$, so $\mathfrak{p}$ is a height one prime of $D$ and thus an element of $\operatorname{Div} D$. In the DVR $D_{\mathfrak{p}}$, $x$ is a unit and $z$ is a uniformizer, so $ord_{\mathfrak{p}}(y) = 2$. The element $y$ does not lie in any other height one prime, so $\operatorname{div} y = 2[\mathfrak{p}]$. On the other hand, let $\mathfrak{m} = \langle \overline{x}, \overline{y}, \overline{z} \rangle$ be the maximal ideal corresponding to the origin. We claim that $\mathfrak{p}D_{\mathfrak{m}}$ is not principal. Indeed, $\mathfrak{m}/\mathfrak{m}^2$ is a 3-dimensional $\mathbb{C}$-vector space generated by the images of $x$, $y$, $z$, whereas $\mathfrak{p} \subset \mathfrak{m}$ and the image of $\mathfrak{p}$ in $\mathfrak{m}/\mathfrak{m}^2$ contains the images of $x$ and $y$, so if $\mathfrak{p}D_{\mathfrak{m}}$ were principal we would have $\dim_{\mathbb{C}} \mathfrak{m}/\mathfrak{m}^2 = 2$. Since $D$ is integrally closed, by Theorem 74a) we may use $\Phi$ to identify $\operatorname{Pic} D$ with a subgroup of $\operatorname{Cl} D$; doing so, what we have shown is that $\mathfrak{p}$ is an order 2 element of $\operatorname{Cl}(D) \setminus \operatorname{Pic}(D)$. A finer analysis – see [Sa64] or [Har77, Example II.6.5.2] – shows that in fact $\operatorname{Cl}(D) = \langle \mathfrak{p} \rangle = \mathbb{Z}/2\mathbb{Z}$, and from this we conclude that $\operatorname{Pic} D = 0$.

By far the most important application of Theorem 74b) is to regular domains. We gave the formal definition in §5.2. Now we supplement it with the remark that the coordinate ring $R = k[V]$ of an affine variety $V$ over a perfect ground field $k$ is regular at a maximal ideal $\mathfrak{m}$ iff $V$ is nonsingular at the closed point corresponding to $\mathfrak{m}$, and $R$ itself is regular iff $V$ is nonsingular at every closed point. The celebrated Auslander-Buchsbaum theorem [AB59] says that a regular local ring is a UFD and thus a regular ring is locally factorial. Thus Theorem 74b) is the analogue for affine integral Noetherian schemes of the equivalence of Cartier and Weil divisors on a nonsingular variety.

It is possible to define the divisor class group of an arbitrary Noetherian domain. Again we let $\operatorname{Div} R$ be the free abelian group on the set $\Sigma$ of height one primes; the integral closure was used only so that for all $\mathfrak{p} \in \Sigma$, $R_{\mathfrak{p}}$ was a DVR and hence there was a canonical monoid homomorphism $v_{\mathfrak{p}} : R_{\mathfrak{p}} \to \mathbb{N}$. In the general case,

if $I$ is an invertible ideal of $R$ and $\mathfrak{p}$ is a height one prime, we may define the component at $\mathfrak{p}$ of $I$ to be the *length* of the $R_{\mathfrak{p}}$-module $R_{\mathfrak{p}}/IR_{\mathfrak{p}}$. In this way, to every invertible ideal we associate an effective divisor. In particular, this works for principal ideals and extends uniquely to a map on $G(R)$, and we may once again define $\operatorname{Cl} R = \operatorname{Div} R/G(R)$. Our assignment of an effective Weil divisor to each effective Cartier divisor induces a homomorphism of groups $\operatorname{Pic} R \to \operatorname{Cl} R$ which is in general neither injective nor surjective.

### 11.3. **Krull domains.**

On the other hand, we know that a UFD is necessarily integrally closed and not necessarily Noetherian, so for our purposes it would be more useful to keep the integrally closed condition and weaken the Noetherian hypothesis. Abstracting the properties that made the construction of $\operatorname{Cl} R$ go through, we are led to the following definition.

An integral domain $R$ is a **Krull domain** if it satisfies the following properties:[9]

(KD1) For each height one prime ideal $\mathfrak{p}$, $R_{\mathfrak{p}}$ is a DVR.
(KD2) $R = \bigcap_{\mathfrak{p}} R_{\mathfrak{p}}$, the intersection extending over the height one ideals.
(KD3) For each $x \in R^\bullet$, the set of height one primes containing $x$ is finite.

These were precisely the properties that we needed to check that an integrally closed Noetherian domain possesses in order to define the divisor class group $\operatorname{Div} R$. Thus the construction of $\operatorname{Div} R$ for a Krull domain is immediate.

**Proposition 75.** *A Krull domain admits a multiplicative norm and is therefore an ACCP domain.*

*Proof.* Let $R$ be a Krull domain. Define the **degree map** $\deg : \operatorname{Div} R \to \mathbb{Z}$ to be the unique homomorphism which sends $[\mathfrak{p}]$ to 1 for each height one prime $\mathfrak{p}$. Then the mapping $0 \mapsto 0$, $f \in R^\bullet \mapsto 2^{\deg(f)}$defines a multiplicative norm on $R$. □

So a domain admits a multiplicative norm if it is an abstract number ring or a Krull domain (in particular, a Noetherian normal domain) and a domain which admits a multiplicative norm is an ACCP domain. I don't know much more than this myself. It seems unlikely that every ACCP domain admits a multiplicative norm, for instance because seemingly weaker properties also imply ACCP. Namely, for any ordinal $\Omega$, define an $\Omega$-norm on a domain $R$ to be a map $N : R^\bullet \to \Omega$ such that for all $x, y \in R$ such that $x$ properly divides $y$, $N(x) < N(y)$. Evidently a domain which admits an $\Omega$-norm is an ACCP domain.

**Question 2.**
*a) Is there a nice characterization of domains which admit a multiplicative norm?*
*b) For which ordinals $\Omega$ does the existence of an $\Omega$-norm imply the existence of a multiplicative norm?*

[9]This is a slight rephrasing of the most standard definition: for the standard definition, which involves a family of discrete valuations, see e.g. [LM71, §8.1]. The equivalence of our definition with the standard one is [LM71, Exercise 8.3].

Coming back to Krull domains: it seem that we have just defined our troubles away, but in fact the class of Krull domains is remarkably natural and robust. We list just a few striking properties.

**Theorem 76.**
*a)* [Na62] *The integral closure of a Noetherian domain is a Krull domain.*
*b)* [LM71, Thm. 8.19] *If $R$ is a Krull domain with fraction field $K$ and $L/K$ is a finite degree field extension, then the integral closure of $R$ in $L$ is a Krull domain.*
*c)* [Gi84] *If $R$ is a Krull domain, so are polynomial and formal power series rings over $R$ in any number of indeterminates.*

The following result is our ultimate characterization of UFDs.

**Theorem 77.** [LM71, Thm. 8.31] *A domain $R$ is a UFD iff it is a Krull domain with $\mathrm{Cl}\, R = 0$.*

We have barely introduced the subject of the divisor class group of a Krull domain. There are whole aspects of the theory that we have not even touched, notably an alternate definition of $\mathrm{Cl}\, R$ based on the important notion of **divisorial ideals**. (For instance, this is closely related to the construction $I \mapsto \overline{I}$ coming up in the definition of a Queen norm.) For a much more systematic development of this theory we recommend [Fo73] and [LM71].

## References


[AAZ90] D.D. Anderson, D.F. Anderson and M. Zafrullah, *Factorization in integral domains*. Journal of Pure and Applied Algebra 69 (1990), 1-19.

[AM69] M.F. Atiyah and I.G. Macdonald, *Introduction to commutative algebra*. Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont. 1969.

[AB59] M. Auslander and D.A. Buchsbaum, *Unique factorization in regular local rings*. Proc. Nat. Acad. Sci. U.S.A. 45 1959 733–734.

[Ba67] A. Baker, *Linear forms in the logarithms of algebraic numbers. I, II, III*. Mathematika 13 (1966), 204-216; ibid. 14 (1967), 102-107; ibid. 14 1967 220–228.

[Bor50] S. Borofsky, *Factorization of polynomials*. Amer. Math. Monthly 57, (1950), 317–320.

[Bou78] A. Bouvier, *Le groupe des classes de l'algbre affine d'une forme quadratique*. Publ. Dép. Math. (Lyon) 15 (1978), no. 3, 53–62.

[BSD52] E.S. Barnes and H.P.F. Swinnerton-Dyer, *The inhomogeneous minima of binary quadratic forms. I.* Acta Math. 87, (1952). 259–323.

[Bu61] D.A. Buchsbaum, *Some remarks on factorization in power series rings*. J. Math. Mech. 10 1961 749–753.

[CE59] E.D. Cashwell and C.J. Everett, *The ring of number-theoretic functions*. Pacific J. Math. 9 (1959) 975–985.

[Cl94] D.A. Clark, *A quadratic field which is Euclidean but not norm-Euclidean*. Manuscripta Math. 83 (1994), no. 3-4, 327–330.

[Coh68] P.M. Cohn, *Bézout rings and their subrings*. Proc. Cambridge Philos. Soc. 64 (1968) 251–264.

[Coh73] P.M. Cohn, *Unique factorization domains*. Amer. Math. Monthly 80 (1973), 1–18.

[Cox89] D.A. Cox, *Primes of the form $x^2 + ny^2$. Fermat, class field theory and complex multiplication*. A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, 1989.

[Ei95] D. Eisenbud *Commutative algebra. With a view toward algebraic geometry*. Graduate Texts in Mathematics, 150. Springer-Verlag, New York, 1995

[Euc] Euclid, *The thirteen books of Euclid's Elements translated from the text of Heiberg. Vol. I: Introduction and Books I, II. Vol. II: Books III–IX. Vol. III: Books X–XIII and Appendix*. Translated with introduction and commentary by Thomas L. Heath. 2nd ed. Dover Publications, Inc., New York, 1956.

[Fo73] R.M. Fossum, *The divisor class group of a Krull domain*. Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 74. Springer-Verlag, New York-Heidelberg, 1973

[Gi84] R. Gilmer, *Commutative semigroup rings*. Chicago Lectures in Mathematics. University of Chicago Press, Chicago, IL, 1984.
[Gr74] A. Grams, *Atomic rings and the ascending chain condition for principal ideals*. Proc. Cambridge Philos. Soc. 75 (1974), 321–329.
[HM04] M. Harper and M.R. Murty, *Euclidean rings of algebraic integers*. Canad. J. Math. 56 (2004), no. 1, 71–76.
[Har77] R. Hartshorne, *Algebraic geometry* Graduate Texts in Mathematics, No. 52. Springer-Verlag, New York-Heidelberg, 1977.
[Has28] H. Hasse, *Über eindeutige Zerlegung in Primelemente oder in Primhauptideale in Integrittsbereichen*. J. reine Angew. Math. 159 (1928), 3-12.
[Hee52] K. Heegner, *Diophantische Analysis und Modulfunktionen*. Math. Z. 56, (1952). 227–253.
[Hel40] O. Helmer, *Divisibility properties of integral functions*, Duke Math. J. 6 (1940), 345-356.
[Hö01] O. Hölder, *Die Axiome der Quantitat und die Lehre vom Mass*. Ber. Verh. Sachs. Ges. Wiss. Leipzig, Math.-Phys. Cl. 53 (1901), 1-64.
[Ka70] I. Kaplansky, *Commutative rings*. Allyn and Bacon, Inc., Boston, Mass. 1970.
[Kh03] D. Khurana, *On GCD and LCM in domains: A Conjecture of Gauss*. Resonance 8 (2003), 72–79.
[Kö36] D. König, *Theorie der Endlichen und Unendlichen Graphen: Kombinatorische Topologie der Streckenkomplexe*. Leipzig: Akad. Verlag, 1936.
[Kr37] W. Krull, *Beiträge zur Arithmetik kommutativer Integritätsbereiche, III, zum Dimensionsbegriff der Idealtheorie*, Mat. Zeit 42 (1937), 745–766.
[Lam05] T.Y. Lam, *Introduction to quadratic forms over fields*. Graduate Studies in Mathematics, 67. American Mathematical Society, Providence, RI, 2005.
[Lan02] S. Lang, *Algebra*. Revised third edition. Graduate Texts in Mathematics, 211. Springer-Verlag, New York, 2002.
[LM71] M.D. Larsen and P.J. McCarthy, *Multiplicative theory of ideals*. Pure and Applied Mathematics, Vol. 43. Academic Press, New York-London, 1971.
[Le43] F.W. Levi, *Contributions to the theory of ordered groups*. Proc. Indian Acad. Sci., Sect. A. 17, (1943). 199–201.
[Li33] F.A Lindemann, *The Unique Factorization of a Positive Integer*. Quart. J. Math. 4 (1933), 319-320.
[LeMo72] K.B. Levitz and J.L. Mott, *Rings with finite norm property*, Can. J. Math., Vol. XXIV, No. 4 (1972), 557-565.
[Mal48] A.I. Malcev, *On the embedding of group algebras in division algebras (Russian)*, Dokl. Akad. Nauk. SSSR 60 (1948), 14991501.
[Mar71] C.F. Martin, *Unique factorization of arithmetic functions*. Aequationes Math. 7 (1971), 211.
[Mat89] H. Matsumura, *Commutative ring theory*. Translated from the Japanese by M. Reid. Second edition. Cambridge Studies in Advanced Mathematics, 8. Cambridge University Press, Cambridge, 1989.
[MRR88] R. Mines, F. Richman and W. Ruitenberg, *A course in constructive algebra*. Universitext. Springer-Verlag, New York, 1988.
[Mo49] T. Motzkin, *The Euclidean algorithm*. Bull. Amer. Math. Soc. 55, (1949). 1142–1146.
[Na57] M. Nagata, *A remark on the unique factorization theorem*. J. Math. Soc. Japan 9 (1957), 143–145.
[Na62] M. Nagata, *Local rings*. Interscience Tracts in Pure and Applied Mathematics, No. 13 Interscience Publishers a division of John Wiley & Sons New York-London 1962
[Ne49] B. H. Neumann, *On ordered division rings*, Trans. Amer. Math. Soc. 66 (1949), 202-252.
[O74] T. Ogoma, *On a problem of Fossum*. Proc. Japan Acad. 50 (1974), 266–267.
[Q96] C.S. Queen, *Factorial domains*. Proc. Amer. Math. Soc. 124 (1996), no. 1, 11–16.
[Ro63] K. Rogers, *Classroom notes: Unique Factorization*. Amer. Math. Monthly 70 (1963), no. 5, 547-548.
[Roi93] M. Roitman, *Polynomial extensions of atomic domains*. Journal of Pure and Applied Algebra 87 (1993), no. 2, 187–199.
[Rü33] W. Rückert, *Zum Eliminationsproblem der Potenzreihenideale*, Math. Ann. 107 (1933), 259–281.

[Sa58] P. Samuel, *Un exemple de variété affine normale*. Soc. Parana. Mat. Anuário (2) 1 1958 4–5.

[Sa61] P. Samuel, *On unique factorization domains*. Illinois J. Math. 5 1961 1–17.

[Sa64] P. Samuel, *Lectures on unique factorization domains*. Notes by M. Pavaman Murthy. Tata Institute of Fundamental Research Lectures on Mathematics, No. 30 Tata Institute of Fundamental Research, Bombay 1964.

[Sa68] P. Samuel, *Unique factorization*. Amer. Math. Monthly 75 (1968), 945–952.

[St67] H.M. Stark, *A complete determination of the complex quadratic fields of class-number one*. Michigan Math. J. 14 1967 1–27.

[T88] H.F. Trotter, *An overlooked example of nonunique factorization*. Amer. Math. Monthly 95 (1988), no. 4, 339–342.

[vdW39] B.L. van der Waerden, *Einftihrung in die algebraische Geometrie*, Berlin, 1939.

[W72] P.J. Weinberger, *On Euclidean rings of algebraic integers*. Analytic number theory (Proc. Sympos. Pure Math., Vol. XXIV, St. Louis Univ., St. Louis, Mo., 1972), pp. 321–332. Amer. Math. Soc., Providence, R. I., 1973.

[Z34] E. Zermelo, *Elementare Betrachtungen zur Theorie der Primzahlen*. Nachr. Gesellsch. Wissensch. Gttingen 1, 43-46, 1934.

[Z10] E. Zermelo, *Collected Works/Gesammelte Werke: Volume I/Band I - Set Theory, Miscellanea/Mengenlehre, Varia*, Springer-Verlag, 2010.

Department of Mathematics, Boyd Graduate Studies Research Center, University of Georgia, Athens, GA 30602-7403, USA

*E-mail address*: pete@math.uga.edu