FACTORIZATION IN INTEGRAL DOMAINS

PETE L. CLARK

Contents

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

Thanks to Roy Smith, Timothy Y. Chow, Waldek Hebisch and Arturo Magidin for useful comments.

of hypersurfaces and divisors in algebraic geometry. Moreover, the subject has a richness which makes its study inherently rewarding.

Nevertheless it is rare to find a satisfactory treatment of factorization in a text at the advanced undergraduate or basic graduate level. So, 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.

However unflattering it may be, I must compare this article to two classic surveys written by two masters of the subject: a 1968 article of P. Samuel [S68] and a 1973 article of P.M. Cohn [C73], both published in the *American Mathematical Monthly.* Samuel's article contains a beautiful discussion of the algebraic-geometric meaning of unique factorization whereas much of Cohn's article discusses the case of factorization in non-commutative rings. My goal here is to present a more comprehensive (and thus overall more elementary) discussion as well as to emphasize connections to number theory and highlight the utility of "norm functions" of various sorts.

By convention, all rings are commutative with a multiplicative identity. Let us agree to exclude from consideration the zero ring, which has a single element $1 = 0$. We will almost entirely be concerned with commutative rings without nonzero divisors of zero, which we call "integral domains" or just "domains".

Introduction

Any nonzero integer has both positive and negative divisors. Our definition of factorization must be sufficiently refined so as to not count such factorizations as

$$3 = (-1) \cdot (-3)$$

as nontrivial. Any integer is divisible not just by 1 but also by -1. Moreover, in any integral domain, every element x is divisible by every **unit** u of the ring. Indeed, the definition of a unit is an element which has a multiplicative inverse – say, u', such that $uu' = 1$ – and then $x = u \cdot (u'x)$ for all x. Thus the first key observation is to regard a factorization $x = yz$ as "trivial" if one of y or z is a unit.

We say a nonzero, nonunit element x of R is **irreducible** if has only trivial factorizations: that is, if $x = yz$, then one of y or z is a unit. (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.[1]

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

$$a = x_1 \cdots x_n,$$

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

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 $a \in R$, $N(a) = 1$.*

Proof. We have $N(1) = N(1 \cdot 1) = N(1) \cdot N(1)$, and since $N(1) \neq 0$,[2] 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.**

Consider the following condition 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**: $N((0)) = 0$ and for any $I \neq (0)$, $N(I) = \#R/I$. 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 $(a) = \{ra \mid r \in R\}$.

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

Proposition 3. *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 $a \in R^{\times}$. It follows easily that $N(a) < N(a^2) < \ldots$.
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.

Example 1.1: Any domain R whose additive group $(R,+)$ is isomorphic to $\mathbb{Z}^d$ for some $d \in \mathbb{Z}^+$ satisfies (FN) and therefore is an abstract number ring. To see this, observe that it is enough to verify the finiteness of R/I for the principal ideal I generated by any nonzero element α of R. 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

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

we see that the ideal αR 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 $R/\alpha R$.

Example 1.2: Let $K/\mathbb{Q}$ is any field extension of finite degree d, and $\mathbb{Z}_K$ is the ring of all elements $\alpha \in K$ which satisfy some monic polynomial relation $\alpha^n + a_{n-1}\alpha^{n-1} + \ldots + a_1\alpha + a_0 = 0$. It is then a basic fact (perhaps the first nontrivial result, in fact) of algebraic number theory that $(\mathbb{Z}_K, +) \cong \mathbb{Z}^d$, so $\mathbb{Z}_K$ is an abstract number ring and hence comes equipped with a standard multiplicative norm, $N(\alpha) = \#(\mathbb{Z}_K/\alpha\mathbb{Z}_K)$.

1.3. **Dirichlet rings.**

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. We suggest that the reader who is unfamiliar with this construction take a few minutes to convince herself that it gives a well-defined ring (especially, that the product is associative, and that the multiplicative identity is $e : 1 \mapsto 1,\ n > 1 \mapsto 0$).

Theorem 4. *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.

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.[3]

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.

We wish to give sufficient conditions for a domain R to be a FD. In practice, most domains one encounters in algebra and number theory are FDs, so we are

[3] 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.

looking for conditions which are widely applicable. In fact we will give two different sets of conditions. The first is fundamentally number-theoretic in spirit and can be viewed as a generalization of Example 2.2 in that it relies on the well-ordering of the positive integers. The second is quite general and abstract algebraic in nature.

As for the first condition, the idea 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 5. *A ring admitting a multiplicative norm is a factorization domain.*

Proof. Let N be a multiplicative norm on te ring R. Suppose for a contradiction that the set of nonzero nonunits in R which *do not* admit irreducible factorizations is nonempty. Then, among all such elements x there exists one with $N(x)$ minimal. Such an x certainly 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 we must have $N(y), N(z) < N(x)$. 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. □

Now for the second condition. In a domain R, we say that 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 $(a) \supsetneq (b)$. We say that a domain R satisfies condition **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.

Remark 2.1: 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 6. *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.

Lemma 7. *Let $(S, \leq)$ be a partially ordered set. The following are equivalent:*
(i) There are no infinite sequences

$$s_1 < s_2 < \ldots < s_n < \ldots \tag{1}$$

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 (1) 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 (1). □

Proof of Proposition 6: We see that (i) $\iff$ (ii) by applying Lemma 7 to the partially ordered set of principal ideals of R, with $(a) \leq (b)$ iff $(a) \subset (b)$. (i) $\iff$ (iii) is immediate, since an infinite sequence $\{a_i\}_{i=1}^{\infty}$ with a_{i+1} properly dividing a_i yields an infinite strictly ascending sequence of principal ideals $(a_1) \subsetneq (a_2) \subsetneq \ldots$, and conversely. □

Proposition 8. *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 9. *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

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

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 (for Roy Smith): 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.

Thus we have shown that any nonzero nonunit, reducible element x of R can be "partially factored" in the sense that it can be 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 above claim can be applied to y, getting $x = a_1 a_2 y_2$ with $(y_2) \subsetneq (y_1)$. Now we argue as above: if this process never terminated, we would produce an infinite strictly ascending chain $(y_1) \subsetneq (y_2) \subsetneq \ldots$ contradicting ACCP; therefore 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. □

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! We briefly sketch an argument for this: first, there exist nonzero nonunit elements of the ring, for instance the element 2. 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 nth 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 such that for all $n \in \mathbb{Z}^+$, the map the nth power map $x \in R \mapsto x^n$ is surjective, then R has nonzero nonunits but no irreducible elements, so is not a factorization domain.

3.2. FD versus ACCP.

The merit of the second proof of Proposition 9 is that it shows that the ACCP condition on a domain R shows that an irreducible factorization can be arrived at via a finite number of "elementary factorizations" – i.e., replacement of a nonzero, nonunit reducible element x with yz, where $yz = x$ and y and z each properly divide x. In fact it can be shown that the procedure of the second proof is unnecessarily careful: ACCP guarantees that any sequence of elementary factorizations terminates in an irreducible factorization.

In a widely read 1968 paper [C68], the distinguished algebraist P.M. Cohn claimed an affirmative answer to this question, however without giving any proof. Therefore it caused a bit of a stir when, in 1974, Anne Grams proved the following result.

Theorem 10. *There exist factorization domains which do not satisfy ACCP.*

Proof. See [G74]. □

3.3. **ACC versus ACCP.**

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

Proposition 11. *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 12. *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 13. *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.

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], [L02, Thm IV.4.1].

Unlike ACC, the condition ACCP does not in general pass to quotient rings (even quotient domains). Indeed, it will follow from Theorem 22 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 13b) for ACCP does hold:

Theorem 14. *R be an ACCP domain. Then $R[t]$ is also an ACCP domain.*

Proof. 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). □

4. Unique factorization domains

4.1. **Associates.**

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 15. *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.

4.2. **Unique factorization domains.**

Finally, we can give the key definition. An integral domain R is a **unique factorization domain** (UFD)[4] 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 σ 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.

The condition (UFD2) is logically complex, and one may worry that it is difficult to verify it directly. Note however that this is exactly what was done in the HLZ proof of unique factorization in $\mathbb{Z}$. On the other hand – and also as is the case for $R = \mathbb{Z}$! – Euclid's Lemma is a useful intermediate point.

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

Proof. The argument that one uses in elementary number theory to deduce the fundamental theorem of arithmetic from Euclid's Lemma (and conversely) carries over without essential change to this context. For completeness, we give the details.

Suppose first that R is a UFD, let f be an irreducible element, and suppose that $f \mid xy$, so that $zf = xy$ for some $z \in R$. Now factor both sides of the equation into irreducibles. Since f appears in the left hand factorization, by uniqueness some irreducible f' associate to f must appear in the right hand side, i.e., either $f' \mid x$ or $f' \mid y$. Since f and f' are associates, we conclude $f \mid x$ or $f \mid y$.

Conversely, suppose R is a factorization domain, and $x = f_1 \cdots f_m = g_1 \cdots g_n$ are two irreducible factorizations of the same nonzero nonunit elment x of R. Then $f_1 \mid g_1 \cdots g_n$ so $f_1 \mid g_j$ for some j. It is harmless to reorder the elements, so we may as well assume that $f_1 \mid g_1$. But since g_1 is irreducible, this means f_1 and g_1 are associates, so $g_1 = u_1 f_1$. Therefore we may cancel f_1's to get $f_2 \cdots f_m = u_1 g_2 \cdots g_n$. Proceeding in this way we get units $u_2, \ldots, u_m$ such that $g_i = u_i f_i$ for all i, and thus $1 = u_1 \cdots u_m g_{m+1} \cdots g_n$, so that $g_{m+1} \cdots g_n \in R^\times$. But a product of nonunit elements is never a unit, so that we must have $m = n$. Thus the factorization is unique up to ordering of the factors and associates. □

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

Theorem 17. *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.

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

Proof. (i) $\implies$ (ii): suppose R is a UFD. By Theorem 16, 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 9, whereas (iii) $\implies$ (i) by Theorem 16. □

4.3. 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 18. *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.

4.4. 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 $N : R \setminus \mathbb{N}$ 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

$$N(x) = \prod_{\mathfrak{p} \in \mathcal{P}} n_\mathfrak{p}^{x_\mathfrak{p}} \tag{2}$$

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. This is easy to check directly, but it may be to some readers' taste to have the following conceptual explanation: the set $R\setminus\{0\}$ under multiplication forms a commutative monoid, and the set $R^\times$ of units is a submonoid. Form the quotient monoid $\tilde{R} := (R\setminus\{0\})/R^\times$.[5] Then weak multiplicative norms on R correspond simply to homomorphisms of monoids $N : \tilde{R} \to (\mathbb{Z}^+, \cdot)$. However, if R is a UFD, $\tilde{R}$ is simply the free commutative monoid on the set $\mathcal{P}$ of nonzero principal prime ideals. (Moreover, $\tilde{K}$ is simply the free abelian group on the set $\mathcal{P}$, which, again, may be a more familiar object.) Thus we have simply used the universal property of free commutative monoids (or free commutative groups).

For the weak multiplicative norm $N = (n_\mathfrak{p})_{\mathfrak{p}\in\mathcal{P}}$ to be a multiplicative norm, it is necessary and sufficient that $n_\mathfrak{p} > 1$ for all $\mathfrak{p} \in \mathbb{P}$.

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 19. *(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.

Let us define a **Queen norm** to be a multiplicative norm satisfying the additional property of Theorem 19, so that Theorem 19 can be restated as: a ring is a UFD iff it admits a Queen norm.

5. Polynomial rings over UFDs

Theorem 20. *(Gauss) If R is a UFD, so is $R[t]$.*

Remark 5.1: The most common proof of Theorem 20 uses Gauss' Lemma on primitive polynomials. We will give a "lemmaless" proof which is modelled on the Hasse-Lindemann-Zermelo proof of unique factorization in $\mathbb{Z}$ [H28] [L33] [Z34].

[5] If one is uncomfortable with quotients of monoids, one can simply view $\tilde{R}$ as a submonoid of the quotient group $\tilde{K} := (K \setminus \{0\})/R^\times$, where K is the fraction field of R.

This argument, with minor variations, appears several times in the literature; it seems that the first such instance is a paper of S. Borofsky [B50].

Proof. By Theorem 17, it suffices to show that $R[t]$ is a factorization domain and an EL-domain. By Theorem 14, since R is a FD, 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

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

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

with $n \geq m > 0$. Suppose the leading coefficient of f_1 (resp. g_1) is a (resp. b). Put

$$q = ap - bf_1x^{n-m}g_2 \cdots g_s = f_1(af_2 \cdots f_r - bx^{n-m}g_2 \cdots g_s) = (ag_1 - bf_1x^{n-m})g_2 \cdots g_s.$$

Thus $q = 0$ implies $ag_1 = bf_1x^{n-m}$. If, however, $q \neq 0$, then

$$\deg(ag_1 - bf_1x^{n-m}) < \deg g_1,$$

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 $ag_1 - bf_1x^{n-m}$ and thus also of ag_1. Either way $ag_1 = f_1h$ for some $h \in R[t]$. Since a is constant and f_1 is irreducible, this implies $h = ah_2$, so $ag_1 = f_1ah_2$, or $g_1 = f_1h_2$, contradiction. □

By induction, we deduce:

Corollary 21. *Let R be a UFD and $n \in \mathbb{Z}^+$. Then $R[t_1, \ldots, t_n]$ is a UFD.*

Theorem 22. *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

$$t_{n+1} = P_1t_1 + \ldots + P_nt_n.$$

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

By Theorem 17, 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. □

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 whic 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 23. *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 24. *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 25. *A GCD-domain is integrally closed in its fraction field.*

Proof: Let R be a GCD-domain with fraction field K, and let α 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

$a_0, \ldots, a_{n-1} \in R$. 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 α 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 24c) implies $\gcd(r^n, s) = 1$. Thus s is a unit, so $\alpha = \frac{r}{s} \in R$.

Proposition 26. *A unique factorization domain is a GCD-domain.*

Proof. This is an immediate generalization of the usual arguments for $R = \mathbb{Z}$. By Remark 5.1, we know that GCD necessarily exists except possibly when both a and b are nonzero nonunits. Then, let $x_1, \ldots, x_r$ 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 25 and 26 imply that a UFD is integrally closed (c.f. §3.1).

Proposition 27. *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 24c), $\gcd(x, yz) = 1$, which contradicts $x \mid yz$. □

Corollary 28. *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 26. Conversely, assume that R is a GCD-domain. Then it is an EL-domain by Proposition 27, and by Theorem 17 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 29. *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 30. *(Khurana,* [K03, Thm. 2]*) For a, b in a domain R, TFAE:*
(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 31. *(Khurana,* [K03, 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 α, β 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 30, 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 30 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 32. *For no $d \geq 3$ is $R_d = \mathbb{Z}[\sqrt{-d}]$ an EL-domain.*

Proof. As in the proof of Theorem 31 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 $\frac{x}{2} + \frac{1}{2}\sqrt{-d} \in R_d$, which is a clear contradiction ($\frac{1}{2}$ is not an integer!). □

Note that Theorem 30 has the following immediate consequence:

Corollary 33. *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.

In particular, there is no need to formally define an "LCM-domain", because this is equivalent to a GCD-domain.

8. More on Principal Ideal Domains

8.1. **PID implies UFD.**

Proposition 34. *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 35. *Any principal ideal domain is a unique factorization domain.*

Proof. We need only put together previously proved results. We know:
A PID is a Noetherian domain (Proposition 7)
A Noetherian domain is a Factorization Domain (Proposition 8).
A PID is a GCD-domain (Proposition 34).
A GCD-domain which is also a factorization domain is a UFD (Corollary 28). □

8.2. **Hasse norms.**

We wish to give a criterion for an integral domain to be a PID which is due to H. Hasse. In fact, Hasse's 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.

Definition: A multiplicative norm $N : F \to \mathbb{Q}$ on the fraction field of an integral domain R is a **Hasse norm** (c.f. [H28]) 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 $\mathbb{Q}$ is a Hasse norm. Indeed, for any rational number x which is not an integer, we can take $a = 1$ and take b to be $\lfloor x \rangle$, the greatest integer less than or equal to x. Then $0 < x - b < 1$.

Theorem 36. *(Hasse,* [H28]*) For an integral domain R, TFAE:*
(i) R admits a Hasse norm N.
(ii) R is a PID.

Proof. (i) $\implies$ (ii): Let I be a nonzero ideal of R, so 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, and define the norm of $x \in R$ as follows: $N(0) = 0$, for any unit u, $N(u) = 1$; otherwise, if $x = x_1 \cdots x_r$ is a product of r irreducible elements (counting multiplicities), then put $N(x) = 2^r$. It is immediate from the uniqueness of factorization that N is a multiplicative norm, so it remains to be seen that it is a Hasse norm. But any element $x \in F \setminus R$ 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 34, 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 37. *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 $a_n \neq 0$. By convention, we decree 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 $N(0) = 2^{-\infty} = 0$. We claim that N is a Hasse norm on $F[t]$. Let's check:

• $N(P) = 0 \iff 2^{\deg P} = 0 \iff P = 0$.
• $N(P) = 1$ iff $2^{\deg P} = 1 \iff \deg(P) = 0$ iff P is a nonzero constant polynomial. These are indeed precisely the units of $F[t]$.
• $N(PQ) = 2^{\deg(PQ)} = 2^{\deg P + \deg Q} = 2^{\deg P} 2^{\deg Q} = N(P)N(Q)$.
• Let $x(t) = \frac{A(t)}{B(t)}$ be a rational function with $B(t)$ not dividing $A(t)$. Then by the division algorithm for polynomials, we may write $A(t) = Q(t)B(t) + R(t)$, where $\deg R(t) < \deg B(t)$, so

$$x(t) = Q(t) + \frac{R(t)}{B(t)},$$

and thus

$$0 < N(x(t) - Q(t)) = N\left(\frac{R(t)}{B(t)}\right) < 1.$$

□

8.3. **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 38. *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 39. *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.*

In the literature, one generally calls a **Euclidean domain** an integral domain for which there exists a Euclidean norm (although the precise definition of a Euclidean norm can be varied somewhat without changing the class of rings one gets: for instance, the multiplicativity is not essential). The advantage of a Euclidean norm is that, as in Proposition 39, it is especially straightforward to check whether a given norm is Euclidean. The advantage of a general Hasse norm is that, as we saw, every PID is guaranteed to have one, whereas in general a PID need not admit any Euclidean norm, and whether it does turns out to be a much more subtle question than whether it is a PID. Thus it may be that the notion of a Euclidean domain is given undue prominence in most expositions of the theory of factorization.

9. Localization

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

The idea of "localizing" a domain by adjoining the multiplicative inverses of some of its elements is simple and useful, but for some reason it seems to be deferred until a full-fledged graduate course in commutative algebra, where it appears as a special case of a more general construction, in which the possibility of zero divisors causes some additional complications. It seems to me that the construction of the fraction field of an integral domain should be a mainstay of undergraduate algebra – it requires no more and no less than an understanding of the construction of the construction of the rational numbers as equivalence classes of pairs of integers – and from there localization ought to be a relatively easy sell.

Perhaps localization is absent in most introductory algebra courses because standard treatments of localization make heavy use of ideal-theoretic language. However, as usual in our study of factorization in domains, we can go a long way by considering only principal ideals.

Our main goal in this section is to give a complete presentation of a theorem of Nagata (Theorem 45), which is justly admired by the experts but seems not to be well enough known. An immediate application is a striking 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 40. *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 $i > m$, π_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 41. *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 42. *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 41 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 43. *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 44. *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 π which divides f^n for some n. In other words, $f^n \in \pi R$, and since πR 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 17.

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 π_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 44, it is no loss of generality to restate Theorem 40 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 45. *(Nagata* [N57]*) 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 17 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 43 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 π 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. But if F is the fraction field of A, $R_S = (A[t])_S = F[t]$ which is, by Proposition 37 above, a PID and hence a UFD. Nagata's theorem applied to R and S now tells us (again) that $R = A[t]$ is a UFD.

Nagata himself used Theorem 45 to study affine quadric surfaces.

Theorem 46. *(Nagata,* [N57]*) Let k be a field of characteristic different from* 2*, and let $a_1, \ldots, a_n \in k^\times$. If $n \geq 5$, $k[t_1, \ldots, t_n]/(a_1x_1^2 + \ldots + a_nx_n^2)$ is a UFD.*

The condition $n \geq 5$ is sharp, since the ring $k[t_1, t_2, t_3, t_4]/(t_1t_2 - t_3t_4)$ is not a UFD: the images $\overline{t_1}, \ldots, \overline{t_4}$ of $t_1, \ldots, t_4$ in R can be seen to be nonassociate irreducible elements, so $\overline{t_1t_2} = \overline{t_3t_4}$ gives distinct factorizations into irreducibles.[6]

There are many other interesting results on factoriality and class groups of quadric surfaces, which are unfortunately beyond the scope of our ambitions. But we cannot resist mentioning the following:

Theorem 47. *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) If $n \geq 2$, then R_n is a UFD.
b) 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.

For further results on such rings see [S64], [S68], [O74], [B78], [R86].

9.4. **Localized norms.**

Let R be a UFD, $S \subset R$ a multiplicative subset, and N a multiplicative norm on a UFD R. Earlier we saw that N extends uniquely to a multiplicative function $K^\times \rightarrow \mathbb{Q}^{>0}$, so in particular we have a multiplicative function $N : R_S \rightarrow Q^{>0}$. However, this is not a norm because (as long as S contains at least one nonunit) it is not $\mathbb{Z}$-valued on R_S. However, this can easily be remedied: we may assume

[6] The details can be found in the wikipedia article on UFDs.

without loss of generality that S is saturated, and then by Theorem 44, S is generated as a multiplicative set by the units of R together with the prime elements $\mathcal{P}_S$ of S. Using the classification of multiplicative norms given in §4.4, it follows there is a unique multiplicative norm N_S on R_S such that

- If π is a prime element of $\mathcal{P}_S$, $N_S(\pi) = 1$;
- If π is a prime element of $R \setminus \mathcal{P}_S$, $N_S(\pi) = N(\pi)$.

One can then easily check the following:

Proposition 48.
Let R be a UFD, S a multiplicative subset, and N a multiplicative norm on R.
a) If N is a Queen norm on R, N_S is a Queen norm on R_S.
b) If N is a Hasse norm on R, N_S is a Hasse norm on R_S.
Thus any localization of a UFD is a UFD, and any localization of a PID is a PID.

To be sure, these facts are straightforward to prove directly. We saw this above for a UFD; for a PID it suffices to verify that if I is any ideal of R_S, then any generator of the principal ideal $I \cap R$ is also a generator of I. Nevertheless the fact that they can also be proved using norm considerations seems interesting.

10. Kaplansky's Characterization of UFDs

Lemma 49. *(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 50. *(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 49 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. □

Notice that Corollary 35, that every PID is a UFD, is an immediate consequence of Kaplansky's Theorem. Moreover we can derive a criterion for a UFD to a PID, as follows. 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. Moreover we define the **dimension** of a ring to be the supremum of all heights of prime ideals.

Theorem 51. *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 50 $\mathfrak{p}$ 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 52. *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 51, 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. □

11. Some additional topics

Some topics that did not make it into this draft but may still deserve a place:

1. The Cashwell-Everett theorem: if R is a field, the Dirichlet ring $\mathcal{D}_R$ is a UFD.

2. A proof that the ring of integers of $\mathbb{Q}(\sqrt{-19})$ is not Euclidean for any norm.

3. A proof that the standard norm on $\mathbb{Q}(\sqrt{-19})$ is a Hasse norm.

4. Standard-Euclidean quadratic fields versus known and conjectured results about class number one number fields being Euclidean with respect to some norm.

5. Examples to show: the localization of an ACCP domain need not be ACCP; a

polynomial ring over a FD need not be a FD.

6. Integral closure of UFDs; compatibility under localization.

7. A discussion of divisorial ideals.

8. A discussion of Krull domains, including the characterization as the domains which are completely integrally closed and satisfy ACC on divisorial ideals. In particular a Noetherian normal domain is Krull.

9. A characterization of UFDs as Krull domains with trivial class group.

10. The theorem $\mathrm{Cl}(R[t]) = \mathrm{Cl}(R)$.

11. More discussion of the group of divisibility $K^\times/R^\times$ of a domain.

12. Factorization issues recast in the language of commutative monoids.

Finally, two questions that are (to the best of my nonexpert knowledge) open:

Q1: Is there a nice characterization of rings R which admit a multiplicative norm? This class contains: UFDs (and half-factorial domains, which we did not discuss and do not plan to) and abstract number rings and is contained in ACCP. That leaves a lot of room: e.g. does every Dedekind domain admit a multiplicative norm?

Q2: For which of the other class number one imaginary quadratic fields is the standard norm a Hasse norm? What is known about other class number one fields?

References

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

[B50] S. Borofsky, *Factorization of polynomials.* Amer. Math. Monthly 57, (1950), 317–320.

[B78] A. Bouvier, *Le groupe des classes de l'algbre affine d'une forme quadratique.* Publ. Dép. Math. (Lyon) 15 (1978), no. 3, 53–62.

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

[C68] P.M. Cohn, *Bézout rings and their subrings.* Proc. Cambridge Philos. Soc. 64 (1968) 251–264.

[C73] P.M. Cohn, *Unique factorization domains.* Amer. Math. Monthly 80 (1973), 1–18.

[G74] A. Grams, *Atomic rings and the ascending chain condition for principal ideals.* Proc. Cambridge Philos. Soc. 75 (1974), 321–329.

[H28] H. Hasse, *Über eindeutige Zerlegung in Primelemente oder in Primhauptideale in Integrittsbereichen.* J. reine Angew. Math. 159, 3-12, 1928.

[K03] D. Khurana, *On GCD and LCM in domains: A Conjecture of Gauss.* Resonance 8 (2003), 72–79.

[L02] S. Lang, *Algebra.* Revised third edition. Graduate Texts in Mathematics, 211. Springer-Verlag, New York, 2002.

[L33] F.A Lindemann, *The Unique Factorization of a Positive Integer.* Quart. J. Math. 4, 319-320, 1933.

[N57] M. Nagata, *A remark on the unique factorization theorem.* J. Math. Soc. Japan 9 (1957), 143–145.

[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.

[R86] A. Ryckaert, *Factorialité de l'algèbre affine de certaines formes quadratiques.* Bull. Soc. Math. Grèce (N.S.) 27 (1986), 111–126.

[S64] 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.

[S68] P. Samuel, *Unique factorization.* Amer. Math. Monthly 75 (1968), 945–952.

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

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

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