UNIQUE FACTORIZATION NOTES

PHIL MAYER

1. Rings

Definition 1.1. Suppose R is a set equipped with two binary operations. Then $(R, +, \cdot)$ is said to be a **ring** if and only if $(R, +)$ is a commutative group and $(R, \cdot)$ is closed.

Since $(R, +)$ is a commutative group, $(R, +)$ is closed, associative, has an identity, and each element in the set has an inverse. We will require our rings to behave nicely on the second binary operation as well: we assume $(R, \cdot)$ is associative, has an identity, and is commutative. While all elements may not necessarily have inverses, these rings will still behave nicely for our purposes. For example, the distributive law $a(b+c) = ab+ac$ applies. Some examples of these rings (called commutative rings with identity) are:

(1) The number systems including $(\mathbb{Z}, +, \cdot)$, $(\mathbb{Q}, +, \cdot)$, $(\mathbb{R}, +, \cdot)$, and $(\mathbb{C}, +, \cdot)$.

(2) $(\mathbb{Z}_n, +, \cdot)$ where $a \in \mathbb{Z}_n$ has an inverse if and only if $\gcd(\{a, n\}) = 1$.

(3) Polynomials with coefficients taken from a commutative ring with identity.

Much like how groups have subgroups, rings also have sub-structures called subrings:

Definition 1.2. $S \subset R$ is a **subring** of R if and only if $\forall a, b \in S, a - b \in S$ and $ab \in S$. Alternatively, S is a subring of R if and only if $S \subset R$, $(S, +)$ is a subgroup of $(R, +)$, and $(S, \cdot)$ is closed.

Definition 1.3. $I \subset R$ is called an **ideal** in R if and only if $\forall a, b \in I, a - b \in I$ and $\forall r \in R,\ ra \in I$.

So an ideal is a subring with an extra condition. One example of an ideal is $n\mathbb{Z}$ in $\mathbb{Z}$. Let's prove that this is true.

Proof. Let $s, t \in n\mathbb{Z}$ and let $k \in \mathbb{Z}$.
Then $s = na$ and $t = nb$ for some $a, b \in \mathbb{Z}$.
But $na - nb = n(a - b) \in n\mathbb{Z}$.
Additionally, notice that $ks = k(na) = n(ka) \in n\mathbb{Z}$.
So $n\mathbb{Z}$ is an ideal in $\mathbb{Z}$. □

An example of a subring that is not an ideal is $\mathbb{Z} \subset \mathbb{Q}$, since $\forall m, n \in \mathbb{Z}, m - n \in \mathbb{Z}$ and $mn \in \mathbb{Z}$, yet choosing $r = \frac{1}{2}$ and $a = 3$, $ra = \frac{3}{2} \notin \mathbb{Z}$.

The final property common in rings that will be helpful while studying unique factorization is the unit.

Definition 1.4. We say a is a **unit** if and only if a has an inverse, denoted a^{-1}.

Interestingly, the set of units of a ring form a group under multiplication. This can be proven easily, mostly from the definition of rings. Since we limit our discussion to commutative rings with identity (that are associative), inverses are the only property left to show. Since each unit has an inverse, the set is clearly a group. It can also be shown that in $\mathbb{Z}_n$, a is a unit if and only if $\gcd(\{a, n\}) = 1$.

2. Integral Domains

The majority of the rings we will work with while studying unique factorization are integral domains. We will expand on integral domains, adding more properties at each level of abstraction in order to define unique factorization domains, principal ideal domains, Euclidean domains, and fields. In order to define an integral domain, we first define zero-divisors.

Definition 2.1. We say that $a, b \in R$ are **zero-divisors** if and only if $a, b \neq 0$ and $ab = 0$. A single element $a \in R$ is a zero-divisor if and only if $\exists b \in R$ such that $ab = 0$.

Note that if a, b are not zero-divisors then $ab = 0 \implies a = 0$ or $b = 0$. For $a \in \mathbb{Z}_n$, a is a zero-divisor if and only if a is not relatively prime to n. For example in $\mathbb{Z}_{12}$, $4 \cdot 3 = 12 \mod 12 = 0$.

Definition 2.2. A ring D is an **integral domain** if and only if D is a commutative ring with identity and no zero-divisors.

As we go forward, we will assume R is just an integral domain (denoted D) unless otherwise stated. The standard examples of infinite and finite integral domains are $\mathbb{Z}$ and $\mathbb{Z}_p$ (where $p \in \mathbb{P}$), respectively. Let's prove the following facts about integral domains.

Remark 2.3. If a is a non-zero-divisor, then $a \neq 0$ can be cancelled. So if $ab = ac$ then $b = c$.

Proof. Suppose $ab = ac$.
Then $ab - ac = 0$ so $a(b - c) = 0$.
$\therefore\ b = c$ □

This fact isn't hard to see in action. For example, in $\mathbb{Z}_{12}$, we might ask if $3 \cdot 4 = 3 \cdot 8$. After cancelling the common factor of 3, we see $4 \neq 8$.

Remark 2.4. A unit is never a zero-divisor.

Proof. Suppose $a, b \in R$, where a is a unit and $ab = 0$.
So $a^{-1}(ab) = a^{-1} \cdot 0 \implies b = 0$. □

3. Unique Factorization Domains

Next, we consider integral domains in which elements can be factored uniquely into a product of irreducible elements. Integral domains equipped with this property are called unique factorization domains. In order to define them, we first need to understand what it means to factor an element into irreducibles.

Definition 3.1. We say that a is an **associate** of b if and only if $a = b \cdot u$ where u is a unit.

Definition 3.2. An element $c \in R$ is **irreducible** if and only if when $c = ab$, then a or b is a unit.

In $\mathbb{Z}$, a and $-a$ are associates and the prime numbers are irreducible. For example, $3 = -(-3)$ so 3 and -3 are associates and 3 is irreducible. It will help (for our purposes) to distinguish between irreducible and prime. While in settings like $\mathbb{Z}$ the definitions of prime and irreducible and equivalent, we need to formalize these ideas in order to discuss them in a more abstract setting.

Definition 3.3. An element $c \in R$ is **prime** if and only if when $a \mid bc$, then $a \mid b$ or $a \mid c$.

For example, consider 60 in $2\mathbb{Z}$. Clearly $60 = 2 \cdot 30 = 6 \cdot 10$, and technically 2, 30, 6, and 10 are all irreducible in $2\mathbb{Z}$. But $2 \mid 6 \cdot 10$, yet $2 \nmid 6$ and $2 \nmid 10$ in $2\mathbb{Z}$.

We now arrive at the definition of a unique factorization domain, which encapsulates the theorem that follows it.

Definition 3.4. An integral domain D is called a **unique factorization domain** if and only if every element $a \in D$ that is neither 0 nor a unit can be factored into a product of a finite number of irreducibles. Furthermore, if $a = (p_1 \cdot \dots \cdot p_r) = (q_1 \cdot \dots \cdot q_s)$ where p_i, q_i are irreducibles and $r = s$, then p_i is an associate of q_i.

Theorem 3.5. *(Unique Factorization Theorem) In a unique factorization domain D, any $a \in D$ can be factored uniquely into a product of irreducibles, up to units.*

So consider $R = \mathbb{Z}$. Notice that since the Unique Factorization Theorem is an extension of the Fundamental Theorem of Arithmetic, the Unique Factorization Theorem applies in R. For $R = 2\mathbb{Z}$, the even integers, R is not a unique factorization domain. Firstly, R is not an integral domain (because it does not contain an identity). Additionally, while we can factor elements in R, for example $4 = 2 \cdot 2$, elements like 6 are irreducible since $3 \notin 2\mathbb{Z}$. While in $a \in R = 2\mathbb{Z}$ can be factored if and only if 4 divides a, it is not a unique factorization domain.

4. Principal Ideal Domains

The next class of ring we will study is called a principal ideal domain. In order to investigate them, we need to define the concept of a principal ideal, building on the notion of an ideal in an ordinary ring.

Definition 4.1. $I \subset R$ is a **principal ideal** generated by $a \in I$ if and only if $I = \langle a \rangle = \{ra \mid r \in R, a \in I\}$. If R is not commutative, then the principal ideal is generated by $ra + ar$: $I = \langle a \rangle = \{ra + ar \mid r \in R, a \in I\}$.

Definition 4.2. D is called a **principal ideal domain** if and only if any ideal in D is principal.

The standard example of a principal ideal domain is $\mathbb{Z}$. So given $I \subset \mathbb{Z}$, we might ask what its generators are. Since the ideals in $\mathbb{Z}$ look like $n\mathbb{Z}$, the generator for each I is the least positive integer in I. So the ideal $\{0, \pm 3, \pm 6, \pm 9, \cdots\}$ is generated by 3, denoted $\langle 3 \rangle$. Now observe that $\langle 6 \rangle = \{0, \pm 6, \pm 12, \pm 18, \cdots\}$ and $3 \mid 6$. In general, you can show that for $a, b \in R$, $\langle b \rangle \subset \langle a \rangle$ if and only if $a \mid b$. It can also be shown that for $a, b \in R$, $\langle a \rangle = \langle b \rangle$ if and only if a and b are associates.

Now suppose we have a collection of sets A_i. Then $\cup_i A_i = A$ means that an element $a \in A$ lies in at least one A_i. We will use the following theorems to see that every principal ideal domain is a unique factorization domain.

Theorem 4.3. *Suppose $N_1 \subset N_2 \subset \cdots$ where each N_i is an ideal in R, an ascending chain of ideals. Then $N = \cup_i N_i$ is an ideal as well.*

Proof. Let $a, b \in N$.
Then $a \in N_j$ and $b \in N_k$ for some $j, k \in \mathbb{N}$. Assume without loss of generality that $j \leq k$.
But then $N_j \subset N_k$, so $a \in N_k$.
So $a - b \in N_k$ and $\forall r \in R, ra \in N_k$ because N_k is an ideal. □

Theorem 4.4. *Let D be a principal ideal domain. If $N_1 \subset N_2 \subset N_3 \subset \cdots$ is a sequence of ideals $N_i \in D$, then $\exists k \in \mathbb{N}$ such that $N_l = N_k \; \forall l \geq k$. Equivalently, every strictly ascending chain of ideals in a principal ideal domain is of finite length.*

Proof. Let D be a principal ideal domain.
Let $N = \cup_i N_i$ be a sequence of ideals in D.
Then $N = \langle c \rangle$ for some $c \in D$.
So then $c \in N_k$ for some $k \in \mathbb{N}$.
So $N \subset N_k$ since any multiple of c is in N, but $N \supset N_k$ so $N = N_k$. □

Bringing all of these results together, we can then prove the following theorem, which tells us that we can factor non-zero, non-unit elements in a principal ideal domain. The theorem that follows it, that every principal ideal domain is a unique factorization domain, will need a proof for uniqueness.

Theorem 4.5. *In a principal ideal domain, any non-zero, non-unit element can be written as a product of irreducible elements.*

Proof. Suppose D is a principal ideal domain and that $a \in D$ is reducible. If a is irreducible, then we are done.
Then $a = bc$ for some non-unit elements $b, c \in D$.
Since $a \mid b$, $\langle a \rangle \subset \langle b \rangle$.
If both b and c are irreducible, then we are done.
Otherwise, suppose $b = st$ for some non-unit elements $s, t \in D$.
Then $\langle b \rangle \subset \langle s \rangle$.
Continuing the process, we will end up with $\langle a \rangle \subset \langle a_1 \rangle \cdots$, so we have an ascending chain of ideals.
Since D is a principal ideal domain, the chain stops: $\langle a \rangle \subset \langle a_1 \rangle \subset \cdots \langle a_n \rangle$.
Then $a = a_n \cdot w$ for some $w \in D$ and we have a product of irreducibles. □

Theorem 4.6. *Every principal ideal domain is a unique factorization domain*

Proof. The previous theorem shows that if D is a principal ideal domain, then for some $a \in D$ where a is neither zero nor a unit, a has a factorization $a = p_1 p_2 \cdots p_r$ into irreducibles.
So let $a \in D$ have that same factorization, as well as another: $a = q_1 q_2 \cdots q_s$.
Then by a corollary (omitted) for principal ideal domains, we have $p_1 \mid (q_1 q_2 \cdots q_s) \implies p_1 \mid q_j$ for some j.
By changing the order of q_j in the second factorization, we can assume $j = 1$, so $p_1 \mid q_1$.
Then $q_1 = p_1 u_1$, and since p_1 is an irreducible, u_1 is a unit. So p_1 and q_1 are associates.
We then have $p_1 p_2 \cdots p_r = p_1 u_1 q_2 \cdots q_s$, and by the cancellation law, $p_2 \cdots p_r = u_1 q_2 \cdots q_s$.
Continuing this process, we finally arrive at $1 = u_1 u_2 \cdots u_r q_{r+1} \cdots q_s$.
But since the remaining q_i are irreducibles, we must have $r = s$. □

5. Euclidean Domains

The penultimate class of ring that we will use to study unique factorization is the Euclidean domain. As a special type of principal ideal domain, Euclidean domains are also unique factorization domains and integral domains. One approach to defining a Euclidean domain is the following:

Definition 5.1. Let D be an integral domain equipped with some norm function $\nu : D \setminus \{0\} \to \mathbb{Z}^+$. Then D is a **Euclidean domain** if and only if ν satisfies the following:

(1) Given $a, b \neq 0$, $\exists q, r$ with $a = bq + r$ where $r = 0$ or $\nu(r) < \nu(b)$.

(2) $\nu(a) \leq \nu(ab) \, \forall a, b \in D$.

We can then show that a Euclidean domain is a type of principal ideal domain by proving the following theorem. The proof that every Euclidean domain is a unique factorization domain follows from our previous proof that every principal ideal domain is a unique factorization domain.

Theorem 5.2. *Every Euclidean domain is a principal ideal domain.*

Proof. Suppose I is an ideal in D such that $I \neq \{0\}$.
We want to show that I is principal (i.e. that it has a single generator).
Pick $d \in I$ with $d \neq 0$ and $\nu(d)$ minimum, possible by the Well-Ordering Principle.
Let $h \in I$.
Then $h = dq + r$ for some $q, r \in D$ where $r = 0$ or $0 \leq \nu(r) < \nu(d)$.
So $r = h - dq$ thus $r \in I$ (since $h, d \in I$ and I is closed).
If $\nu(r) < \nu(d)$, then this contradicts d being the minimum non-zero element of I.
So $r = 0$ in either case, so $h = dq$, meaning every element of I is a multiple of d.
Then I is the principal ideal generated by d.
$\therefore$ This holds for all ideals of D, and so any Euclidean domain is a principal ideal domain. □

Theorem 5.3. *Every Euclidean domain is a unique factorization domain.*

6. Fields

The final class of rings of interest to us is fields, the most well-behaved version of a ring.

Definition 6.1. The set $(F, +, \cdot)$ is a **field** if and only if $(F, +)$ is a group and $(F, \cdot)$ is a commutative group.

Some common examples include $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, and $\mathbb{Z}_p$. Since fields are Euclidean domains, they are also principal ideal domains, unique factorization domains, and integral domains. Thus, they have no-zero divisors for example. We will investigate fields further in our discussion of polynomials.

7. Polynomials

We will call $D[x]$ the set of polynomials with coefficients in D, where D is an integral domain. Typically we will consider $\mathbb{Z}$, $\mathbb{Q}$, or $\mathbb{R}$, though the latter two are fields specifically. We define a polynomial to be a sum, like the following:

$$a_0 + a_1 x + a_2 x^2 \cdots + a_n x^n$$

with $a_i = 0\, \forall i > n$. So polynomials can be thought of as a finite sum of a finite number of non-zero coefficients. We can write polynomials in summation notion, like $p(x) = \sum_{i=0}^{k} a_i x^i$. While the sum of polynomials is easy to express in symbolic form, the product of two polynomials (since we will work in a ring) requires more thought. So given $p(x) = \sum_{i=0}^{k} a_i x^i$ and $q(x) = \sum_{j=0}^{l} a_j x^j$:

$$p(x) \cdot q(x) = \sum_{n=0}^{k+l} c_n x^n$$

with $c_n = a_0 b_n + a_1 b_{n-1} + \cdots + a_n b_0$. While clunky, treating polynomials as objects in a ring makes sense for reasons we will see shortly.

Theorem 7.1. *The set $R[x]$ of all polynomials in an indeterminate x with coefficients in a ring R is a ring under polynomial addition and multiplication. If R is commutative, then so is $R[x]$, and if R has unity $1 \neq 0$, then 1 is also unity for $R[x]$.*

The notion of a polynomial's degree also comes up in this more formal setting.

Definition 7.2. The **degree** of a polynomial $p(x)$, $\deg(p) = n$ if and only if $a_n \neq 0$ and $a_i = 0\, \forall i > n$.

The degree of the sum of two polynomials p and q is found as $\deg(p+q) = \max(\{\deg(p), \deg(q)\})$. The degree of the product of two polynomials is more complicated, however. While it might seem that for all polynomials $\deg(p \cdot q) = \deg(p) + \deg(q)$, this result does not necessarily hold for polynomials with coefficients taken from any ring. For example, in $\mathbb{Z}_4$, $(2x^3 + 3)(2x^2 + x) = (4 \mod 4)x^5 + 2x^4 + 2x^2 + 3x = 2x^4 + 2x^2 + 3x$. So the formula does not hold. Instead, one can prove the following fact.

Remark 7.3. For polynomials $p(x), q(x) \in D[x]$, $\deg(p \cdot q) \leq \deg(p) + \deg(q)$.

One of the most convenient facts about taking this more formalized, abstract approach to studying polynomials is that by studying the ring of that a polynomial's coefficients are taken from, we can often draw conclusions about the set of polynomials itself. For example, it turns out that:

Theorem 7.4. *If D is an integral domain, then $D[x]$ is an integral domain as well.*

Theorem 7.5. *If D is a unique factorization domain, then $D[x]$ is a unique factorization domain as well.*

Additionally, we can think of $D \subset D[x]$ because D can be seen as the set of constant polynomials. This agrees with what we know about integral domains, since the additive and multiplicative identities for polynomials are $0 \cdot x^0 = 0$ and $1 \cdot x^0 = 1$, respectively.

When discussed in terms of unique factorization, it is critical to know where a polynomial's coefficients are taken from. For example, in $\mathbb{Q}[x]$, $2 = 2x^0$ is a unit because it has an inverse, namely $\frac{1}{2} = \frac{1}{2}x^0$. Yet, in $\mathbb{Z}[x]$, 2 is an irreducible polynomial. In $\mathbb{C}[x]$, $x^2+1 = (x+i)(x-i)$, but in $\mathbb{R}[x]$ it is irreducible. We can also discuss polynomials with coefficients from $\mathbb{Z}_n$. For example in $\mathbb{Z}_2$, $(x+1)^2 = x^2 + (2 \mod 2)x + 1 = x^2 + 1$. In the more general setting $\mathbb{Z}_p$ where $p \in \mathbb{P}$, $(x+1)^p = x^p + 1$.

If we take a polynomial's constants from a field (or a Euclidean domain), we have a division algorithm for polynomials.

Theorem 7.6. *(Division Algorithm) If $f(x), g(x) \in F[x]$ where F is a field, then $\exists q(x), r(x) \in F[x]$ with $f(x) = g(x) \cdot q(x) + r(x)$ and $r(x) = 0$ or* $\deg(r(x)) < deg(q(x))$.

Proof. Assume $p(x)$ and $d(x)$ are monic (the highest-degree term of each has coefficient 1).
Then consider $S = \{p(x) - d(x)f(x)\}$.
Let $r(x)$ be a polynomial of minimal degree in S.
So there is a $q(x)$ with $p(x) - d(x)q(x) = r(x)$, or $p(x) = d(x) \cdot q(x) + r(x)$.
Suppose $\deg(r) \geq \deg(d)$.
Well $(p(x) - d(x)q(x)) - x^s d(x) = r(x) - x^s d(x)$ where $s = \deg(r) - \deg(d)$.
But if we call $r'(x) = (p(x) - d(x)q(x) - x^s d(x)) = p(x) - d(x)(q(x) - x^s)$ then $r'(x) \in S$.
Yet $\deg(r') < \deg(r)$, the element with the smallest degree in S. So contradiction.
$\therefore \deg(r) < \deg(d)$. □

We can use the Division Algorithm to write reducible polynomials in a field $F[x]$ as the product of irreducibles. This comes from the fact that a field is a unique factorization domain. To summarize:

Theorem 7.7. *If F is a field, then every non-constant polynomial $f(x) \in F[x]$ can be factored into a product of irreducible polynomials in $F[x]$, the irreducibles being unique, up to their order in the factorization and unit (nonzero constant) factors in F.*

The list below details some other facts about polynomials that can be useful in the study of unique factorization.

- An element $a \in F$ is a zero of $f(x) \in F[x]$ if and only if $(x-a)$ is a factor of $f(x)$ in $F[x]$. Equivalently, $f(a) = 0$ if and only if $(x-a) \mid f(x)$.
- We need to differentiate between $f(x)$ as a function and as a polynomial. For example, consider $p(x) = x^2 + 1$ and $q(x) = x^3 + 1$. Both are polynomials in $\mathbb{Z}_2[x]$. As polynomials, $p(x) \neq q(x)$, yet as functions they are equal since $p(0) = q(0) = 1$ and $p(1) = q(1) = 0$.

- Let $f(x) \in F[x]$, and let $f(x)$ be of degree 2 or 3. Then $f(x)$ is reducible over F if and only if it has a zero in F, or in other words has a linear factor.

- The previous fact is not true of degree 4. For example, consider $x^4 + 3x^2 + 2 = (x^2 + 2)(x^2 + 1)$. The two factors are then irreducible in $\mathbb{Q}[x]$.

- If $f(x) \in \mathbb{Z}[x]$, then $f(x)$ factors into a product of two polynomials of lower degrees r and s in $\mathbb{Q}[x]$ if and only if it has such a factorization with polynomials of the same degrees (r and s) in $\mathbb{Z}[x]$.

- If $f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0$ is in $\mathbb{Z}[x]$ with $a_0 \neq 0$, and if $f(x)$ has a zero in $\mathbb{Q}$, then it has a zero $m \in \mathbb{Z}$ and m must divide a_0. In other words, if $f(x)$ can be reduced into a product of linear terms in $\mathbb{Z}[x]$, then the product of the constant terms of each linear factor multiplies to $f(x)$'s constant term.

- Sometimes we can argue using the degree of a polynomial. If R has no zero-divisors and $f(x), g(x) \in R[x]$ then $\deg(f(x)g(x)) = \deg(f) + \deg(g)$. This is true if R is an integral domain, for example. In rings like $\mathbb{Z}_4$, if two polynomials have leading terms $2x^3$ and $2x^2$ respectively, then the product would have leading term $(4 \mod 4)x^5$.

- In a finite field, we can try all elements to find the roots. For example, in $\mathbb{Z}_5$, $x^3 + 2x^2 + 1$ is irreducible because none of its elements satisfy $x^3 + 2x^2 + 1 = 0$.

8. Greatest Common Divisors

The concept of a greatest common divisor for an algebraic object also comes up here. There are two closely-related definitions we can use.

Definition 8.1. d is the **greatest common divisor** of a and b if and only if:

(1) $d \mid a$ and $d \mid b$.

(2) If $c \mid a$ and $c \mid b$, then $c \mid d$. A second definition of the greatest common divisor (for a less general setting with ordering, for example in the integers) has conclusion $c \leq d$ instead.

To find the greatest common divisor in a unique factorization domain, simply factor the elements of a and b into a product of irreducibles and find the greatest product of common factors. For example, given polynomials $a(x) = (x-1)^3(x+2)^5(x-3)^4$ and $b(x) = (x-1)^2(x+2)^5(x-6)^2$:

$$\gcd(\{a(x), b(x)\}) = (x-1)^2(x+2)^5$$

In a Euclidean domain, we can use the **division algorithm** where we divide a by b and obtain some remainder r. If $r = 0$ or $r = 1$, then the greatest common divisor is b or 1, respectively. For all other remainders after the first step, divide b by r. Continue to divide the divisors and remainders from the previous step until a remainder of 0 or 1 is achieved. The previous remainder (the current divisor) is the greatest common divisor. We can also use the **Euclidean algorithm** to write the greatest common divisor as a linear combination of the two multiples.

9. The Algebraic Integers

In order to define the algebraic integers, we first need to understand what makes an integer square-free. From here, we define the algebraic integers, their norm, and the Gaussian integers in the next section.

Definition 9.1. Some $d \in \mathbb{Z}$ is **square-free** if and only if $d = \pm p_1 \cdot p_2 \cdot \cdots \cdot p_r$ where each $p_i \in \mathbb{P}$ is a distinct prime. So d is not divisible by any square.

With some thought, it can be seen that 6 is square-free, but $8 = 4 \cdot 2$ is not. Letting d be some square-free integer, we can then define the set $\mathbb{Q}\sqrt{d} = \{a + b\sqrt{d} \mid a, b \in \mathbb{Q}\}$. Clearly $\mathbb{Q}\sqrt{d} \subset \mathbb{R}$ if $d > 0$ and $\mathbb{Q}\sqrt{d} \subset \mathbb{C}$ if $d < 0$. Let's prove $\mathbb{Q}\sqrt{d}$ is a field.

Proof. Let $s, t \in \mathbb{Q}\sqrt{d}$.
Then $s = a_1 + b_1\sqrt{d}$ and $t = a_2 + b_2\sqrt{d}$ for some $a_1, a_2, b_1, b_2 \in \mathbb{Q}$.
To show $(\mathbb{Q}\sqrt{d}, +)$ is a commutative group, observe that:
$s + t = (a_1 + b_1\sqrt{d}) + (a_2 + b_2\sqrt{d}) = (a_1 + a_2) + (b_1 + b_2)\sqrt{d}$, so $(\mathbb{Q}\sqrt{d}, +)$ is closed.
Since it is well-known that addition in $\mathbb{C}$ is both associative and commutative, $(\mathbb{Q}\sqrt{d}, +)$ is associative and commutative as well.
$\mathbb{Q}\sqrt{d}$ has additive identity $0 + 0\sqrt{d} = 0$, and an additive inverse $s^{-1} = -a_1 - b_1\sqrt{d}$ exists for each s.
To show $(\mathbb{Q}\sqrt{d}, \cdot)$ is a commutative group, observe that:
$st = (a_1 + b_1\sqrt{d})(a_2 + b_2\sqrt{d}) = (a_1a_2 + b_1b_2d) + (a_1b_2 + a_2b_1)\sqrt{d}$, so $(\mathbb{Q}\sqrt{d}, \cdot)$ is closed.
Multiplication in $\mathbb{C}$ is associative and commutative, so $(\mathbb{Q}\sqrt{d}, \cdot)$ is both associative and commutative.
$\mathbb{Q}\sqrt{d}$ has multiplicative identity $1 + 0\sqrt{d}$, and a multiplicative inverse $s^{-1} = \frac{1}{a_1+b_1\sqrt{d}} \cdot \frac{a_1-b_1\sqrt{d}}{a_1-b_1\sqrt{d}}$ exists for each s.
So $\mathbb{Q}\sqrt{d}$ is a field. □

While the larger set $\mathbb{Q}\sqrt{d}$ is a field, a smaller version called the algebraic integers is not.

Definition 9.2. The set $\mathbb{Z}\sqrt{d}$ is called the **algebraic integers**, defined as the following where $a, b \in \mathbb{Z}$:

$$\mathbb{Z}\sqrt{d} = \begin{cases} a + b\sqrt{d} & d \equiv 2 \text{ or } 3 \bmod 4 \\ \frac{a}{2} + \frac{b}{2}\sqrt{d} & d \equiv 1 \text{ or } 4 \bmod 4 \end{cases}$$

While $\mathbb{Z}\sqrt{d}$ is not a field, it is still an integral domain. We might ask: is it a unique factorization, principal ideal, or Euclidean domain as well? If not, does $\mathbb{Z}\sqrt{d}$ fall into one of these categories under any conditions? In order to start investigating these questions, we have to equip the algebraic integers with a norm:

Definition 9.3. $\nu : \mathbb{Z}\sqrt{d} \setminus \{0\} \to \mathbb{Z}$ is the **norm function** in $\mathbb{Z}\sqrt{d}$, given by $\nu(a+b\sqrt{d}) = (a+b\sqrt{d})(a-b\sqrt{d}) = a^2 - b^2 d$.

Recall that in order for an integral domain D to be a Euclidean domain, $\nu : D \setminus \{0\} \to \mathbb{Z}^+$ must exist and satisfy the following properties:

(1) Given $a, b \neq 0$, $\exists q, r$ with $a = bq + r$ where $r = 0$ or $\nu(r) < \nu(b)$.

(2) $\nu(a) \leq \nu(ab) \, \forall a, b \in D$.

The standard examples are $\nu(x) = |x|$ on $\mathbb{Z}$ and $\nu(p(x)) = \deg(p(x))$ for $p(x)$ in some field $F[x]$ with coefficients taken from a field F. In $\mathbb{Z}\sqrt{d}$, the norm function does not always satisfy both conditions, and so $\mathbb{Z}\sqrt{d}$ is not a Euclidean domain in general. It turns out that $\mathbb{Z}\sqrt{d}$ is a Euclidean domain for the following values of d:

$$d = -11, -7, -3, -2, -1, 1, 2, 3, 5, 6, 7, 11, 13, 17, 19, 21, 29, 33, 37, 41, 57, 73$$

More generally, $\mathbb{Z}\sqrt{d}$ is a unique factorization domain for:

$$d = -1, -2, -3, -7, -11, -19, -43, -67, -163$$

Let's now consider $D = \mathbb{Z}(\sqrt{-3})$. In D, $4 = 2 \cdot 2 = (1 + 2\sqrt{-3})(1 - 2\sqrt{-3})$. We know that D is a unique factorization domain (given above), and the fact that we have two factorizations does not contradict it. Both feature two irreducible elements, but we need to show that the factors are associates in order to be sure. Let's first prove that one of the factors is irreducible (which can be extended to the other three).

Proof. Assume $(1 + \sqrt{-3})$ is reducible in D.
Then $(1 + \sqrt{-3}) = (a + b\sqrt{-3})(c + d\sqrt{-3})$.
Taking the norm of both sides, we have $\nu(1 + \sqrt{-3}) = \nu((a + b\sqrt{-3})(c + d\sqrt{-3}))$.
But $\nu(1 + \sqrt{-3}) = (1)^2 + 3 \cdot (1)^2 = 4$ and separating the two norms:
$\nu((a + b\sqrt{-3})(c + d\sqrt{-3})) = \nu(a + b\sqrt{-3})\nu(c + d\sqrt{-3})$.
So we need $a^2 + 3 \cdot b^2 = c^2 + 3 \cdot d^2 = \pm 2$.
Then clearly $b = 0$ and $d = 0$ since each term is positive, so $a = \pm\sqrt{2}$ and $c = \pm\sqrt{2}$.
But $a, c \in \mathbb{Z}$ so we have a contradiction.
So $(1 + \sqrt{-3})$ is irreducible in D. □

We now attempt to prove that 2 and $(1 + \sqrt{-3})$ are associates. The proof can be adapted to show that 2 and $(1 - \sqrt{-3})$ are associates as well.

Proof. If the two factors are indeed associates, then we can write $2 \cdot u = (1 + \sqrt{-3})$.
Well consider $u = (\frac{1}{2} + \frac{1}{2}\sqrt{-3})$.
Clearly $2 \cdot (\frac{1}{2} + \frac{1}{2}\sqrt{-3}) = (1 + \sqrt{-3})$, but we need to show $(\frac{1}{2} + \frac{1}{2}\sqrt{-3})$ is a unit.
So $\nu(\frac{1}{2} + \frac{1}{2}\sqrt{-3}) = (\frac{1}{2})^2 + 3 \cdot (\frac{1}{2})^2 = 1$. So 2 and $(1 + \sqrt{-3})$ are associates, and this agrees with the Unique Factorization Theorem. □

It is also worth investigating whether or not elements are prime in $\mathbb{Z}\sqrt{d}$. For example, 2 is irreducible but not prime in $\mathbb{Z}(\sqrt{-3})$.

Proof. Suppose $2 \mid (1+\sqrt{-3})(1-\sqrt{-3})$.
But $2 \nmid (1+\sqrt{-3})$ and $2 \nmid (1-\sqrt{-3})$.
Yet suppose it does: then $(1+\sqrt{-3}) = 2 \cdot (a + b\sqrt{-3}) = 2a + 2b\sqrt{-3}$.
But then $a = \frac{1}{2}$, yet $a \in \mathbb{Z}$, so we have a contradiction.
So 2 is not prime in $\mathbb{Z}(\sqrt{-3})$. □

10. The Gaussian Integers

Definition 10.1. The set $\mathbb{Z}[i]$ is called the **Gaussian integers**. It is a special case of the algebraic integers defined as $\mathbb{Z}[i] = \{a + bi \mid a, b \in \mathbb{Z}\}$.

Clearly, Gaussian integers are just algebraic integers with $d = -1$. As previously discussed, $\mathbb{Z}[i]$ is a Euclidean domain, and thus a principal ideal domain and a unique factorization domain. Note that for some Gaussian integer $(a + bi) \in \mathbb{Z}[i]$, $\nu(a + bi) = a^2 + b^2$. Since units in the algebraic integers have norm ± 1, the units in $\mathbb{Z}\sqrt{d}$ are $\pm 1, \pm i$. This changes the reducibility of elements. For example, while 2 is irreducible in $\mathbb{Z}$, it is reducible in $\mathbb{Z}[i]$ because $2 = (1 + i)(1 - i)$. Overall, the irreducibles in $\mathbb{Z}[i]$ have three forms:

(1) $p = i \pm 1$, with norm 2.

(2) $p \equiv 3 \mod 4$, with norm p^2.

(3) $x = a + bi$ where $a^2 + b^2 = p$ and $p \equiv 1 \mod 4$.

Let's prove these facts.

Proof. Assume $1 + i$ is reducible.
Then $1 + i = (a + bi)(c + di)$ for some $a, b, c, d \in \mathbb{Z}$.
But $\nu(1 + i) = 2$, so one of the factors must be a unit. □

Proof. Suppose $p \equiv 3 \mod 4$ and that we can factor it into $p = (a + bi)(c + di)$ for some $a, b, c, d \in \mathbb{Z}$.
Then $\nu(p) = p^2 = (a^2 + b^2)(c^2 + d^2)$.
$\implies a^2 + b^2 = p$ and $c^2 + d^2 = p$.
Mod 4, the only squares are 0 and 1.
So $a^2 + b^2 \equiv 0, 1,$ or $2 \mod 4$, and same for $c^2 + d^2$.
But since $p \equiv 3 \mod 4, p^2 = 1$. So both $a^2 + b^2, c^2 + d^2 = 1$.
So p is irreducible. □

Proving the third fact follows trivially from the following theorem originally proposed by Fermat:

Theorem 10.2. *The equation $a^2 + b^2 = p$ is solvable in the integers if and only if $p \equiv 1 \mod 4$.*

Let's conclude with some factorizations in $\mathbb{Z}[i]$. Suppose we take $2 + 5i$. Its norm is $\nu(2 + 5i) = 2^2 + 5^2 = 4 + 25 = 29$. Since 29 is prime in $\mathbb{Z}$, $29 \equiv 1 \mod 4$ so $2 + 5i$ is irreducible in $\mathbb{Z}[i]$. How about $3 + 4i$? Well $\nu(3 + 4i) = 3^2 + 4^2 = 9 + 16 = 25$. So we need to find $a, b, c, d \in \mathbb{Z}$ such that $25 = 5 \cdot 5 = (a^2 + b^2)(c^2 + d^2)$. Note that neither of the two factors can have norm -5 since $a, b, c, d \in \mathbb{Z}$ and the square of each is non-negative. After a small amount of trial and error, the factorization $3 + 4i = (-1 + 2i)(-1 - 2i)$ can be found.