6-2013

An Algebraic Approach to Number Theory using Unique Factorization

Mark Sullivan
Union College - Schenectady, NY

An Algebraic Approach to Number Theory using Unique Factorization

By

Mark Sullivan

****

Submitted in partial fulfillment
of the requirements for
Honors in the Department of Mathematics

UNION COLLEGE
June, 2013

ABSTRACT

SULLIVAN, MARK An Algebraic Approach to Number Theory using Unique Factorization. Department of Mathematics, June 2013.

ADVISOR: Karl Zimmermann

Though it may seem non-intuitive, abstract algebra is often useful in the study of number theory. In this thesis, we explore some uses of abstract algebra to prove number theoretic statements. We begin by examining the structure of unique factorization domains in general. Then we introduce number fields and their rings of algebraic integers, whose structures have characteristics that are analogous to some of those of the rational numbers and the rational integers. Next we discuss quadratic fields, a special case of number fields that have important applications to number theoretic problems. We will use the structures that we introduce throughout the thesis to prove several number theoretic statements, including the Fundamental Theorem of Arithmetic, Fermat's Theorem on Sums of Squares, and the Ramanujan-Nagell Theorem, as well as to generate a myriad of other interesting tangentially related results.

Contents

Preface

Unlike analysis and topology, number theory deals mainly with denumerable quantities, rather than the sort of continuity that is present in, for example, the real numbers. This quality allows number theory to act as a suitable framework for working in Abstract Algebra. Lagrange's Theorem and its proof are examples of algebraic concepts that draw upon the language of number theory. Concurrently, because the structure of groups and rings follows that of the integers so closely, there are also many beautiful ways that abstract algebra can also be used to help solve problems in number theory. Thus, many number theoretic results can be proven more easily using results from more general algebraic structures. In this thesis, we will observe this counter-intuitive principle as we apply theorems of unique factorization in rings to number theoretic problems.

The reader of this thesis should be familiar with the definitions that are introduced in a first course in abstract algebra, including, for example, the definitions of "subring," "maximal ideal," and "ring of polynomials." If the reader is not acquainted with these concepts, we recommend reference [5] as a resource.

Throughout this thesis, we will use the notation $\mathbb{Z}$ to denote the set or ring of integers, $\mathbb{Q}$ to denote the set or field of rational numbers, $\mathbb{R}$ to denote the set or field of real numbers, and $\mathbb{C}$ to denote the set or field of complex numbers. A superscript $+$ shall denote the set of positive members of that set. For example, $\mathbb{Z}^+$ shall denote the set $\{x \in \mathbb{Z} | x > 0\}$. If we are dealing with a context that involves an equivalence relation on a set X, and a is an element of X, then we will let $[a]$ denote "the equivalence class of elements of X that are equivalent to a." Further, concerning set relations, we will write $a \in X$ to mean "a is an element of the set X," $X \subseteq Y$ to

mean "the set X is a subset of the set Y," and $X \subsetneq Y$ to mean "the set X is a subset of Y and is not equal to Y." Occasionally we will reverse the direction of some of these symbols, in which case, the placement of the two objects in the sentence that the statement corresponds to shall be reversed. For example, $X \supseteq Y$ shall mean "the set Y is a subset of the set X."

Concerning notational conventions of abstract algebraic concepts, we shall always denote the additive identity of a ring as 0 and the multiplicative identity of any ring that has a multiplicative identity as 1. We will write G/H to denote either "the quotient group of the group G mod its normal subgroup H" or "the quotient ring of the ring G mod its ideal H." For an element a of some ring, we will let (a) denote "the ideal generated by a." If f is a polynomial, we will let ∂f denote "the degree of the polynomial f." We define the degree of the zero polynomial to be less than the degree of any other polynomial. Whenever we are dealing with a nonzero polynomial f, and we write out its terms like $f(x) = c_n x^n + ... + c_1 x + c_0$, then we will always assume that $c_n \neq 0$. Finally, with the exception of $\mathbb{Z}/p\mathbb{Z}$ for a prime number $p \in \mathbb{Z}$, all fields mentioned in this thesis are assumed to have characteristic 0.

1 Unique Factorization Domains

We begin with a discussion of the Euclidean Algorithm. Recall that, in number theory, the Euclidean Algorithm is a series of consecutive applications of the division algorithm that determines the greatest common divisor of two integers. In Ring Theory, a generalization of this algorithm exists that can be used to find an analog of the notion of a greatest common divisor that exists in the integers. This generalization, however, can only be applied to certain rings, and more specifically, certain integral domains. Such an integral domain is called a "Euclidean Domain," as defined below.

Definition 1.1 *An integral domain with unity, D, is a Euclidean Domain provided that there exists a function, known as a field norm, $N : D \longrightarrow \mathbb{Z}^+ \cup \{0\}$, such that $N(0) = 0$ and $\forall\ a, b \in D$ such that $b \neq 0$, $\exists\ q, r \in D$ such that $a = qb + r$, with $r = 0$ or $N(r) < N(b)$. In this case, q is called the quotient, and r the remainder.*

In these Euclidean Domains we will see that there are natural generalizations of the various divisibility properties of the integers. As such, in order to be studied further, the definition of the term "divides" that is commonly associated with the integers should be, in some way, generalized to apply to various other integral domains. For this reason, we will now introduce a definition with a familiar name.

Definition 1.2 *Let C be a commutative ring with $a, b \in C$ and $b \neq 0$. Then a is a multiple of b in C provided that $\exists\ q \in C$ such that $a = bq$.*

We will also say that a divides b in C, denoted $a|b$, when b is a multiple of a in C.

In the context of rings, this definition has some important implications. First, there is an algebraic analog of divisibility for principal ideals.

Proposition 1.3 *Let C be a commutative ring with unity and $a, b \in C$. Then $b|a$ if and only if $(a) \subseteq (b)$.*

Proof ($\Rightarrow$) Let $c_1, c_2 \in C$ with $c_2|c_1$. Then $c_1 = qc_2$ for some $q \in C$. Let $a \in (c_1)$. This implies that $a = c_1q_1$ for some $q_1 \in C$. Then $a = qc_2q_1 = c_2qq_1$, and so $a = c_2q_2$ for some $q_2 \in C$, namely $q_2 = qq_1$. Thus, $a \in (c_2)$.

($\Leftarrow$) Let $(c_1) \subseteq (c_2)$. We know that $c_1 \in (c_1)$, since $c_1 = 1c_1$. Then $c_1 \in (c_2)$, so $c_1 = qc_2$ for some $q \in C$. Thus, $c_2|c_1$. $\square$

Redefining "divides" leads us to redefine "greatest common divisor".

Definition 1.4 *Let $a, b, d \in C$, a commutative ring. Then d is called a greatest common divisor of a and b in C, provided that $d|a$ and $d|b$, and that any other element of C dividing both a and b also divides d.*

It is important to note before moving on that this definition of greatest common divisor provides the possibility that there may be more than one greatest common divisor for any particular pair of elements of a ring. (There is also no guarantee that a greatest common divisor exists for any particular pair of elements; this will be discussed later.) We will soon prove that this ambiguity exists using a proposition. First, we will define the term "unit", in order to make the language that we will use for the upcoming proposition more concise.

Definition 1.5 *Let R be a ring with unity. Then $u \in R$ is a unit in R provided that $\exists\, v \in R$ such that $uv = vu = 1$.*

It turns out that the units of a commutative ring C form an abelian group under the multiplicative operation of C. The proof of this fact is straightforward, and we will

leave it to the reader. We will denote this group, called "the group of units of C", by $U(C)$.

Now, for the sake of a stronger analogy, we will also define "relatively prime."

Definition 1.6 *Let $a, b \in C$, a commutative ring with unity. Then a and b are said to be relatively prime in C provided that the following is true for any $d \in C$: if $d|a$ and $d|b$, then $d \in U(C)$.*

When only one ring is specified, so that there is no possibility for confusion, we will often shorten the phrase "u is a unit in R" to "u is a unit". We will often do the same with the terms "greatest common divisor" and "relatively prime".

Before moving on, we will introduce another definition that will make the terminology easier.

Definition 1.7 *Two elements of an integral domain with unity are said to be associates if some one of them is a product of the other and a unit of that integral domain. In other words, let D be an integral domain and $a \in D$. Then $b \in D$ is an associate of a provided that $b = ua$ for some unit $u \in U(D)$.*

When b is an associate of a, we say that a and b "associate" in the integral domain. It is trivial to show that associating is an equivalence relation, and so we will often say in this case that a and b are "associates". We will now discuss the implications of these definitions for the greatest common divisor.

Proposition 1.8 *Let C be a commutative ring, $u \in U(C)$, $a, b \in C$, and d be a greatest common divisor of a and b in C. Then ud is also a greatest common divisor of a and b.*

Proof Let $u \in U(C)$ and let d be a greatest common divisor of $a, b \in C$. We will show that ud is a common divisor of a and b. Since d is a common divisor of a and b, $\exists\ q_1, q_2 \in C$ such that $a = q_1 d$ and $b = q_2 d$. Since u is a unit, $\exists\ v \in C$ such that $uv = 1$. Then for the aforementioned q_1 and q_2, $a = q_1 vud$ and $b = q_2 vud$. Thus, $ud|a$ and $ud|b$.

We will show that $\forall\ d_1 \in C$ such that $d_1|a$ and $d_1|b$, $d_1|ud$. Let $d_1|a$ and $d_1|b$ for some $d_1 \in C$. Then since d is a greatest common divisor of a and b, $d_1|d$. But $d|ud$, so if $d = qd_1$ for some $q \in C$, then $ud = uqd_1$, and so $d_1|ud$. $\square$

As we said before, this last proposition introduces the somewhat disturbing concept that there may be more than one so-called greatest common divisor. In the set of integers, this ambiguity is not much of an issue, since the only units in the ring of integers are 1 and -1. However, some rings have many units, and some even have infinitely many. Even so, the algebraic analog of divisibility grants us a language that allows us to gracefully express many of the properties of greatest common divisors by using a slightly stronger statement. This removes the need for constantly making awkward statements involving juxtapositions of the words "a" and "greatest." Instead of referring to elements dividing other elements, one can simply discuss the ideals generated by various elements, as in the following proposition.

Proposition 1.9 *Let C be a commutative ring. For $a, b \in C$, neither of which are equal to 0, if $(a, b) = (d)$, then d is a greatest common divisor of a and b.*

Proof Let $a, b \in C$, neither of which are equal to 0, and let $(a, b) = (c)$. We will show first that $c|a$ and $c|b$. We know that $\forall\ x, y \in C$, $ax + by \in (a, b) = (c)$. Let

$y = 0$. Then $\forall\ x \in C, ax \in (c)$, so by definition,

$$(a) = \{ax | x \in C\} \subseteq (c), \tag{1}$$

implying that $c|a$ by Proposition 1.3. By setting $x = 0$, we likewise find that $c|b$. So c is a common divisor of a and b.

Next we will show that $d_1|a$ and $d_1|b$ implies that $d_1|c$. Let $d_1|a$ and $d_1|b$. Then $(a) \subseteq (d_1)$ and $(b) \subseteq (d_1)$ by Proposition 1.3. Thus, $\forall\ x, y \in C, ax, by \in (d_1)$. Since (d_1) is closed under addition, $\forall\ x, y \in C, ax + by \in (d_1)$, so by definition,

$$(c) = (a, b) = \{ax + by | x, y \in C\} \subseteq (d_1). \tag{2}$$

Yet this implies that $d_1|c$ by Proposition 1.3, hence c is a greatest common divisor of a and b. $\square$

We will take this language and apply it in order to create a generalization of the Euclidean Algorithm of number theory, which finds the greatest common divisor of two integers through successive applications of the division algorithm. In this context, the Euclidean Algorithm is quite the same as its version in the integers, except that the steps will involve elements of some integral domain that may or may not be the ring of integers. To demonstrate, consider $a, b \in D$, where D is a Euclidean Domain. Then there exist $q_0, r_0 \in D$ such that $a = q_0 b + r_0$ and $N(r_0) < N(b)$ or $r_0 = 0$. Likewise, there exist $q_1, r_1 \in D$ such that $b = q_1 r_0 + r_1$ and $N(r_1) < N(r_0)$ or $r_1 = 0$, and $q_2, r_2 \in D$ such that $r_0 = q_2 r_1 + r_2$ and $N(r_2) < N(r_1)$ or $r_2 = 0$, and so on, creating a system of equations:

$$a = q_0 b + r_0$$
$$b = q_1 r_0 + r_1$$
$$r_0 = q_2 r_1 + r_2$$
$$r_1 = q_3 r_2 + r_3$$
$$\vdots$$
$$r_{n-2} = q_n r_{n-1} + r_n$$
$$r_{n-1} = q_{n+1} r_n + 0.$$

The process of creating this system of equations is called the Euclidean Algorithm.

The reader may immediately have concern over whether this system of equations is finite, that is, whether there really is a final nonzero remainder r_n as above. Consider the field norm. By the definition of Euclidean Domain, when the Euclidean Algorithm is enacted as above, it is required that the remainders always have field norms less than the corresponding divisors. That is,

$$N(b) > N(r_0) > N(r_1) > N(r_2) > N(r_3) > ... > N(r_n) > 0. \tag{3}$$

The field norm is defined to have nonnegative values only. Thus, the field norms of the remainders in the Euclidean Algorithm form a strictly decreasing sequence of nonegative integers, which, by the Well-Ordering Principle, must end eventually. Thus, the Euclidean Algorithm is guaranteed to end after a finite number of steps. The following theorem demonstrates the importance of this effect.

Theorem 1.10 *Let $0 \neq a, b \in D$, a Euclidean Domain. If r_n is the last nonzero remainder in the Euclidean Algorithm starting with $a = q_0 b + r_0$, then $(r_n) = (a, b)$.*

Proof Let $a, b \in D$, a Euclidean Domain. We will show that $(r_n) = (a, b)$. We claim that $(a, b) = (b, r_0)$.

($\subseteq$) Let $e_1 \in (a, b)$. Then $e_1 = ax + by$ for some $x, y \in D$. Now, $a = q_0 b + r_0$, so

$$e_1 = ax + by = (q_0 b + r_0)x + by = b(q_0 x + y) + r_0 x \in (b, r_0). \tag{4}$$

Thus, $(a, b) \subseteq (b, r_0)$.

($\supseteq$) Let $e_2 \in (b, r_0)$. Then $e_2 = bx + r_0 y$ for some $x, y \in D$. Now, $r_0 = a - q_0 b$, so

$$e_2 = bx + r_0 y = bx + (a - q_0 b)y = ay + b(x - q_0 y) \in (a, b). \tag{5}$$

Thus, $(b, r_0) \subseteq (a, b)$, and so $(a, b) = (b, r_0)$.

The same arguments can be applied to show that $(b, r_0) = (r_0, r_1)$, and that $(r_0, r_1) = (r_1, r_2)$, and so on. Hence, $(r_{i-1}, r_i) = (r_i, r_{i+1})$ for all $0 < i \leq n$, where we define $r_{n+1} = 0$. This implies that

$$(a, b) = (b, r_0) = (r_0, r_1) = ... = (r_n, r_{n+1}) = (r_n, 0) = (r_n). \tag{6}$$

Therefore, $(r_n) = (a, b)$. $\square$

The theorem demonstrates that, because a final nonzero remainder of the Euclidean Algorithm must exist for every pair of elements of a Euclidean Domain, every pair of elements of a Euclidean Domain is guaranteed to have at least one greatest common divisor.

We will now prove that the ring of integers is an example of a Euclidean Domain.

Lemma 1.11 *The ring of integers is a Euclidean Domain.*

Proof First, we know that $\mathbb{Z}$ is an integral domain because no two nonzero integers $a, b \in \mathbb{Z}$ satisfy $ab = 0$. Next, we will show that $\forall\, a, b \in \mathbb{Z}$ with $b \neq 0$, $\exists\, q, r \in \mathbb{Z}$ so that $a = qb + r$ with $r = 0$ or $|r| < |b|$. In other words, we will show that the absolute value function is a suitable field norm for $\mathbb{Z}$. Let $a, b \in \mathbb{Z}$.

Consider the set

$$S = \{a - xb \in \mathbb{Z}^+ \cup \{0\} | x \in \mathbb{Z}\}. \tag{7}$$

We claim that $S \neq \varnothing$. If $b > 0$, then let $x = -|a|$; if $b < 0$, then let $x = |a|$. Then $-xb \geq |a|$, and so $a - xb \geq a + |a| \geq 0$, so $a - xb \in S$, thus $S \neq \varnothing$.

The Well-Ordering Principle implies that S must contain a smallest element, which we will call $s \in S$. Then for some $y \in \mathbb{Z}$, $s = a - yb \geq 0$. We will show that $|s| < |b|$. Suppose for contradiction that $|s| \geq |b|$. If $b > 0$, let $z = 1$; if $b < 0$, let $z = -1$. Therefore, $zb = |b|$. Then

$$a - (y + z)b = a - yb - zb = s - zb = s - |b| = |s| - |b| \geq 0, \tag{8}$$

since by definition of s, $s > 0$. Thus, $a - (y + z)b \in S$. But

$$a - (y + z)b = s - zb = s - |b| < s \tag{9}$$

since $b \neq 0$, so this would imply that s is not actually the least element of S, but rather that $a - (y + z)b$ is, despite the fact that s is defined to be the least element of S. This contradiction implies that $|s| < |b|$. $\square$

We have closely followed the proof on page 17 of reference [3] in writing the previous proof.

The previous lemma will be used to provide an algebraic proof of the Fundamental Theorem of Arithmetic at the end of this chapter. But before that, we will introduce a few more types of rings.

Definition 1.12 *An integral domain with unity D is called a Principal Ideal Domain provided that every ideal in D can be generated by only one element; that is, every ideal D is principal.*

Principal Ideal Domains have an algebraic structure that guarantees that every two members of a Principal Ideal Domain must have a greatest common divisor, since every ideal generated by two members can actually be generated by some one (this is a direct result of Proposition 1.9). The following theorem will indicate that this shared characteristic between Principal Ideal Domains and Euclidean Domains is not a coincidence.

Theorem 1.13 *Every Euclidean Domain is a Principal Ideal Domain.*

Proof Let D be a Euclidean Domain with $I \subseteq D$ an ideal. If $I = \{0\}$, then I is principal, since it is generated by 0. Assume $I \neq \{0\}$, and let $s \in I$ such that $s \neq 0$ and $N(s)$ is minimal among nonzero elements of I. Such an element must exist because of the Well-Ordering Principle, since

$$\{N(a)|a \in I\} \subseteq \mathbb{Z}^+ \cup \{0\}. \tag{10}$$

It is clear that $(s) \subseteq I$, since $s \in I$ and I is an ideal. We will show that $I \subseteq (s)$. Let $a \in I$. Therefore $a = qs + r$ for some $q, r \in D$ with $r = 0$ or $N(r) < N(s)$. Then

$$r = a - qs. \tag{11}$$

But $a, s \in I$, so $r \in I$, and thus $N(r) \geq N(s)$ (because we have defined $N(s)$ to be minimal) or $r = 0$. The first choice leads to a contradiction, since by assumption, if $r \neq 0$, then $N(r) < N(s)$. Thus, $r = 0$, so $a = qs \in (s)$. Therefore, $I = (s)$, and so I is a principal ideal. Hence, D is a Principal Ideal Domain. $\square$

We have closely followed the proof on page 273 of reference [4] in writing the previous proof.

Many important properties of the integers were discovered as results of the study of the intensely interesting numbers known as primes. As such, in order to approach number theoretic problems from an algebraic perspective, it is necessary to have some sort of analog of the prime numbers, and some language in which to express the various theorems of number theory that involve prime numbers. These can be obtained through the study of certain ideals.

Definition 1.14 *Let $P \subsetneq R$ be an ideal in a ring R. Then P is a prime ideal provided that $\forall\ a, b \in R$, if $ab \in P$ then $a \in P$ or $b \in P$. A nonzero element $p \in R$ is called a prime provided that (p) is a prime ideal.*

This definition may take the reader by surprise, as it does not appear, at first, to have any relationship with the fact that prime integers are irreducible, that is, that a prime number has no divisors other than ± 1, itself, and its negative. This is because purely

traditional understandings of the integers that do not take algebraic structures into account tend to conflate the ideas of "prime" and "irreducible", a term that we will define in general now.

Definition 1.15 *Let D be an integral domain. Let $i \in D$ be neither 0 nor a unit. Then i is called irreducible in D provided that $i = ab$ for any $a, b \in R$ implies that either a is a unit or b is a unit. Also, i is called reducible in R provided that it is not irreducible in R.*

The difference between primality and irreducibility is often ignored in introductary courses in number theory because the two coincide in the integers. The set of integers is special among integral domains because an integer that is irreducible is always prime. While this is not true of every integral domain, the converse is, as we shall now prove.

Proposition 1.16 *In an integral domain with unity, all prime elements are irreducible.*

Proof Let D be an integral domain with unity, and $p \in D$ be prime. Then let $p = ab$, for some $a, b \in D$. Then since $p \in (p)$, it is clear that $ab \in (p)$. Since p is prime, (p) is a prime ideal, so $a \in (p)$ or $b \in (p)$. Suppose, with the understanding that the argument will be similar in the other choice, that $a \in (p)$. Then $a = qp$ for some $q \in D$. Then $p = qbp$, so $p(1 - qb) = 0$. Since D is an integral domain, and $p \neq 0$, $1-qb = 0$, so $qb = 1$, and therefore b must be a unit. Thus, p is irreducible. $\square$

While primality implying irreducibility is a property of any and every integral domain with unity, the converse is, as we previously mentioned, not true for every case. For proof of this, consider the following example.

Example 1.17 *Consider the ring* $\mathbb{Z}\left[\sqrt{-5}\right] = \{x + y\sqrt{-5} | x, y \in \mathbb{Z}\}$*, a subring of* $\mathbb{C}$*. This must be an integral domain, since it shares the multiplicative operations of* $\mathbb{C}$*. It is clear that it is a ring with unity because* $1 \in \mathbb{Z}\left[\sqrt{-5}\right]$*. It is a fact that* ± 1 *are the only units in* $\mathbb{Z}\left[\sqrt{-5}\right]$*. The reader may either accept this on faith or wait until we prove this in Chapter 3 using Theorems 3.10 and 3.14. Now, we claim that* 3 *is an irreducible element of* $\mathbb{Z}\left[\sqrt{-5}\right]$*. First, it is a simple matter to show that* $3 \in \mathbb{Z}\left[\sqrt{-5}\right]$*, since* $3 = 3 + 0\sqrt{-5}$*. Next, let* $3 = ab$ *for some* $a, b \in \mathbb{Z}\left[\sqrt{-5}\right]$*. Then* $a = x_1 + x_2\sqrt{-5}$ *and* $b = y_1 + y_2\sqrt{-5}$ *for some* $x_1, x_2, y_1, y_2 \in \mathbb{Z}$*. So* $3 = (x_1 + x_2\sqrt{-5})(y_1 + y_2\sqrt{-5})$*. Thus,*

$$3 = (x_1y_1 - 5x_2y_2) + (x_1y_2 + x_2y_1)\sqrt{-5}. \tag{12}$$

Then since $3 \in \mathbb{Z}$*,*

$$x_1y_2 + x_2y_1 = 0. \tag{13}$$

Likewise,

$$x_1y_1 - 5x_2y_2 = 3 \tag{14}$$

We will now prove by contradiction that $y_2 = 0$*.*

Assume for contradiction that $y_2 \neq 0$*. We see from Equation 14 that*

$$x_1y_1y_2 - 5x_2{y_2}^2 = 3y_2, \tag{15}$$

and Equation 13 implies that

$$x_1 y_2 = -x_2 y_1. \tag{16}$$

Therefore, we find that

$$-x_2 {y_1}^2 - 5x_2 {y_2}^2 = 3y_2. \tag{17}$$

In that case, $-x_2({y_1}^2 + 5{y_2}^2) = 3y_2$*, and by taking the absolute value of both sides of this equation we see that*

$$|x_2|({y_1}^2 + 5{y_2}^2) = 3|y_2|. \tag{18}$$

Distributing the factor on the left side, we see that $|x_2|{y_1}^2 + 5|x_2||y_2|^2 = 3|y_2|$*, and therefore that* $5|x_2||y_2|^2 \leq 3|y_2|$*, which is preposterous for* $y_2 \neq 0$*. So* $y_2 = 0$*.*

With this established, we first deduce that, because $x_1 y_1 - 5x_2 y_2 = 3$*, we must have that* $x_1 y_1 = 3$*. Since* 3 *is irreducible in* $\mathbb{Z}$*, then,* x_1 *or* y_1 *must be* ± 1*, with the other being* ± 3*. So it is clear then, since* $\mathbb{Z}$ *is an integral domain, that* $y_1 \neq 0$*. This means that, since* $x_1 y_2 + x_2 y_1 = 0$ *and* $y_2 = 0$*,* $x_2 = 0$*. Then, referring to the definitions of* a *and* b *that were established in the first paragraph of this example, we see that* $a = x_1$ *and* $b = y_1$*, and one of them is* ± 1*. Thus, one of* a *and* b *must be a unit in* $\mathbb{Z}[\sqrt{-5}]$*, so* 3 *is irreducible in* $\mathbb{Z}[\sqrt{-5}]$*.*

However, despite 3 *being irreducible, we will now show that* (3) *is not a prime ideal. Consider that* $(2 + \sqrt{-5})(2 - \sqrt{-5}) = 9 \in (3)$*. But we will now show that* $2 + \sqrt{-5}, 2 - \sqrt{-5} \notin (3)$ *by contradiction. Assume for contradiction that* $2 \pm \sqrt{-5} = 3q$ *for some* $q \in \mathbb{Z}[\sqrt{-5}]$*. Then* $q = x + y\sqrt{-5}$ *for some* $x, y \in \mathbb{Z}$*.*

Thus,

$$2 \pm \sqrt{-5} = 3x + 3y\sqrt{-5}, \tag{19}$$

and so $(2 - 3x) + (\pm 1 - 3y)\sqrt{-5} = 0$. *In that case,* $2 = 3x$ *and* $\pm 1 = 3y$, *which cannot be true for integers* x *and* y. *So* x *and* y *are simultaneously integral and non-integral, which is a contradiction. Therefore* $2 + \sqrt{-5}, 2 - \sqrt{-5} \notin (3)$, *so it is clear that* (3) *does not satisfy the definition of prime ideal and so* 3 *is simultaneously irreducible and nonprime in* $\mathbb{Z}[\sqrt{-5}]$. □

The previous example all but clearly tells us that those integral domains in which irreducibility implies primality are special. Naturally, then, we seek a characterization of such integral domains. The first observation is that all Principal Ideal Domains, including Euclidean Domains, have this property. Before we prove this statement, though, we will first require a few supporting lemmas and one well-known result from elementary abstract algebra.

Theorem 1.18 *Let* I *be an ideal of the ring* R. *If* $J \subseteq R$ *and* $I \subseteq J$, *then* J *is an ideal of* R *if and only if* J/I *is an ideal of* R/I. *This is called the Lattice Isomorphism Theorem for rings.*

Proof We will omit this proof. The interested reader may see pages 226-227 of reference [1] for a proof of this statement. □

Corollary 1.19 *Let* M *be an ideal of a commutative ring* C. *Then* C/M *and* M/M *are the only ideals of* C/M *if and only if* M *is maximal in* C, *where*

$$M/M = \{[x] | \forall\, a, b \in M, a \equiv b \text{ if and only if } a - b \in M\} = \{[0]\} = (0), \tag{20}$$

such that $\equiv$ is the equivalence relation defining the quotient ring.

Proof ($\Rightarrow$) Let ${}^{C}/_{I}$ and ${}^{I}/_{I}$ be the only ideals of ${}^{C}/_{I}$. We will show that I is maximal. Theorem 1.18 implies that C and I are ideals of C. Let J be an ideal of C satisfying $I \subseteq J \subseteq C$. Theorem 1.18 implies that ${}^{J}/_{I}$ must be an ideal of ${}^{C}/_{I}$. Thus, either ${}^{J}/_{I} = {}^{I}/_{I}$ or ${}^{J}/_{I} = {}^{C}/_{I}$. Suppose, with the understanding that the other choice is similar, that J satisfies ${}^{J}/_{I} = {}^{C}/_{I}$. We will show that $J = C$. First of all, we know that $J \subseteq C$, by definition of J.

We claim that $C \subseteq J$ Let $c \in C$. Then $[c] \in {}^{C}/_{I} = {}^{J}/_{I}$. But this implies that $c \equiv j$ in I for some $j \in J$. Thus, by definition, $c - j \in I \subseteq J$. Since $c - j, j \in J$, we find that $(c - j) + j = c \in J$. Therefore, $C \subseteq J$. Hence, ${}^{J}/_{I} = {}^{C}/_{I}$ implies that $J = C$, and similar arguments show that ${}^{J}/_{I} = {}^{I}/_{I}$ implies that $J = I$. Thus, any ideal J satisfying $I \subseteq J \subseteq C$ must be equal to either I or C. By definition, then, I is maximal.

($\Leftarrow$) Let M be maximal in C. By Theorem 1.18, it is clear that ${}^{C}/_{M}$ and ${}^{M}/_{M}$ are ideals of ${}^{C}/_{M}$. We will show that ${}^{C}/_{M}$ and ${}^{M}/_{M}$ are the only ideals of ${}^{C}/_{M}$. Let ${}^{R}/_{M}$ be an ideal of ${}^{C}/_{M}$. We will show that ${}^{R}/_{M} = {}^{C}/_{M}$ or ${}^{R}/_{M} = {}^{M}/_{M}$.

Consider the function $\varphi : C \longrightarrow {}^{C}/_{M}$ whose behavior on inputs and outputs is described by $\forall\ a \in C,\ \varphi(a) = [a]$. It is straightforward to show that this function serves as a homomorphism on the rings C and ${}^{C}/_{M}$. We claim that $M \subseteq \varphi^{-1}\left({}^{R}/_{M}\right)$. Let $m \in M$. Since ${}^{R}/_{M}$ is defined, we must have that M is an ideal of R. By definition, $m \equiv 0$ in M, so $[m] = [0] \in {}^{R}/_{M}$. Thus, $\varphi(m) = [m] \in {}^{R}/_{M}$. Therefore, $m \in \varphi^{-1}\left({}^{R}/_{M}\right)$, and so $M \subseteq \varphi^{-1}\left({}^{R}/_{M}\right)$.

Now, we know that

$$\varphi^{-1}\left({}^{R}/_{M}\right) = \{c \in C | \varphi(c) \in {}^{R}/_{M}\} = \{c \in C | [c] \in {}^{R}/_{M}\}$$
$$= \{c \in C | M + c = M + a \text{ for some } a \in R\}$$
$$= \{c \in C | c - a \in R \text{ for some } a \in R\} = \{c \in C | c \in R\} = R. \quad (21)$$

Therefore, $M \subseteq \varphi^{-1}\left({}^{R}/_{M}\right) = R$. Since M is maximal, this implies that $R = M$ or $R = C$. Therefore, ${}^{R}/_{M} = {}^{M}/_{M}$ or ${}^{R}/_{M} = {}^{C}/_{M}$, and so the proof is complete. □

Lemma 1.20 *Let C be a commutative ring with unity. Then $a \in C$ is a unit if and only if $(a) = C$.*

Proof ($\Rightarrow$) Let C be a ring with $u \in C$ a unit. Then $\exists\ v \in C$ so that $uv = 1$. Thus, since (u) is closed under multiplication by an element of C, $1 \in (u)$. Now let $a \in C$. Then since $1 \in (u)$, $a \in (u)$. Hence, $C \subseteq (u)$, and $(u) \subseteq C$ is a matter of definition, so $C = (u)$.

($\Leftarrow$) Let $(a) = C$ for an element $a \in C$, a commutative ring with unity. Then $1 \in (a)$, so $ba = ab = 1$ for some $b \in C$. Therefore a is a unit. □

Lemma 1.21 *Let D be an integral domain with unity and $a, b \in D$. In that case, $(a) = (b)$ if and only if a and b associate.*

Proof ($\Rightarrow$) Let $a, b \in D$, where D is an integral domain, and suppose that $(a) = (b)$. Then $a \in (a) \subseteq (b)$, so $a = q_1 b$ for some $q_1 \in D$. Further, $b \in (b) \subseteq (a)$, so $b = q_2 a$ for some $q_2 \in D$. Therefore, $a = q_1 b = q_1 q_2 a$, and thus, $a(1 - q_1 q_2) = 0$.

We consider two cases: either $a = 0$ or $a \neq 0$. If $a = 0$, then $b = q_2 a = q_2 0 = 0$, and so $a = b$, hence $a = 1b$, and so a and b are associates. If $a \neq 0$, then because D is an integral domain, $1 - q_1 q_2 = 0$, and so $q_1 q_2 = 1$. This implies that q_1 is a unit, and so since $a = q_1 b$, a and b are associates.

($\Leftarrow$) Suppose that a is an associate of b. Then $a = ub$ for some $u \in U(D)$, and so $b|a$, which implies that $(a) \subseteq (b)$ by Proposition 1.3. Therefore, for some $v \in U(D)$, $uv = 1$, so $va = vub = b$. In that case, $a|b$, and so $(b) \subseteq (a)$ by Proposition 1.3. Hence, $(a) = (b)$. $\square$

Lemma 1.22 *In a commutative ring C with unity, an ideal I is maximal if and only if ${}^{C}/_{I}$ is a field.*

Proof Let C be a commutative ring with unity, and M a maximal ideal. We will show that ${}^{C}/_{M}$ is a field. We know by definition that M is maximal if and only if the only ideals in C that are supersets of M are C and M. Corollary 1.19 reveals that M is maximal in C if and only if ${}^{C}/_{M}$ and (0) are the only ideals of ${}^{C}/_{M}$. Let $a \in {}^{C}/_{M}$, where $a \neq 0$. Then $(a) = {}^{C}/_{M}$, since ${}^{C}/_{M}$ and (0) are the only ideals in ${}^{C}/_{M}$. Therefore, a is a unit by Lemma 1.20. Hence, M is maximal if and only if every nonzero element of ${}^{C}/_{M}$ is a unit. This can be true if and only if ${}^{C}/_{M}$ is a field. So M is maximal if and only if ${}^{C}/_{M}$ is a field. $\square$

Lemma 1.23 *For a commutative ring C, the ideal I is prime if and only if ${}^{C}/_{I}$ is an integral domain.*

Proof Let C be a commutative ring and $P \subseteq C$ a prime ideal. We will show that ${}^{C}/_{P}$ is an integral domain. By definition of prime ideal, $P \neq C$. Let $ab \in P$. This

can be true if and only if $(a+P)(b+P) = ab+P = P$ in $^{C}/_{P}$. This, in turn, is equivalent to $[a][b] = [ab] = [0]$, since $^{C}/_{P} = \{x+P | x \in C\}$. But since P is prime, $a \in P$ or $b \in P$. Thus, $[a] = [0]$ or $[b] = [0]$. By definition, then, $^{C}/_{P}$ is an integral domain. The proof of the converse is similar, since all of the statements made in this part of the proof are reversible. $\square$

Lemma 1.24 *In a Principal Ideal Domain, a nonzero ideal is prime if and only if it is maximal.*

Proof ($\Rightarrow$) Let D be a Principal Ideal Domain, and let $P \subseteq D$ be an arbitrary nonzero prime ideal. Our assertion is that this P must be a maximal ideal. Since D is a Principal Ideal Domain, $P = (p)$ for some element $p \in D$, which, by definition, must be a prime. Let $I = (a)$ be some arbitrary ideal where $a \in D$ so that

$$(p) = P \subseteq I = (a). \tag{22}$$

(Since p itself is an example of what a could be, such an a must exist.) Our assertion will be proven, then, when we have that $I = P$ or $I = D$, since then it will be shown that there is no ideal that is both proper superset of P and a proper subset of D.

Since $(p) \subseteq (a)$, $p \in (a)$, so $p = q_1 a$ for some $q_1 \in D$. Thus, $q_1 a \in (p)$, hence $q_1 \in (p)$ or $a \in (p)$. If $a \in (p)$, then $ab \in (p)$ for any and all $b \in D$, so

$$I = (a) \subseteq (p) = P, \tag{23}$$

proving our assertion. Instead then, let us say that $q_1 \in (p)$. Then $q_1 = pq_2$ for some $q_2 \in D$. Then, referring back to our definition of q_1, $p = pq_2 a$, hence

$p - pq_2a = p(1 - q_2a) = 0$. Yet D is an integral domain, and $p \neq 0$ because primes are defined as nonzero. Thus, $1 - q_2a = 0$, so q_2 and a are units. Yet Lemma 1.20 implies that this means that $(a) = D$, so our assertion is proven. Thus, in either case, P is a maximal ideal.

($\Leftarrow$) Let M be a maximal ideal in D. Then Lemma 1.22 indicates that D/M is a field. Therefore, D/M is an integral domain, so by Lemma 1.23, M is prime. $\square$

Now that these are established, we may proceed to prove that Principal Ideal Domains are part of the denomination of "special" rings that have indistinguishability between primes and irreducibles.

Proposition 1.25 *An element of a Principal Ideal Domain is prime if and only if it is irreducible.*

Proof ($\Rightarrow$) That all primes are irreducible is a property of every integral domain, and since a Principal Ideal Domain is an integral domain, we have already proven the first statement in Proposition 1.16.

($\Leftarrow$) Let D be a Principal Ideal Domain, $i \in D$ be an irreducible in D, and $I \supseteq (i)$ be an ideal in D. We will show that i is prime. Because D is a Principal Ideal Domain, $I = (a)$ for some $a \in D$. Then $i \in (a)$, so $i = qa$ for some $q \in D$. But i is irreducible, so q is a unit or a is a unit. If q is a unit, i and a associate in D, so $(i) = (a) = I$ by Lemma 1.21. If a is a unit, then $I = (a) = D$, by Lemma 1.20. Either way, (i) is maximal, so by Lemma 1.24, p is prime. $\square$

We have closely followed the proof on page 284 of reference [4] in writing the previous proof.

It is for this reason that primes and irreducibles are indistinguishable in the ring of integers; the ring of integers, being a Euclidean Domain, is a Principal Ideal Domain, as shown above, and so by Proposition 1.25, every prime is an irreducible and vice-versa in the ring $\mathbb{Z}$. In the same vein, the integers have a particularly interesting property concerning the structure of the ring that pertains to the prime numbers. This property is the Fundamental Theorem of Arithmetic, which states that every nonzero integer other than ± 1 can be written as a product of irreducibles, and that this representation is unique, apart from the order and signs of the factors. In fact, many other rings have a similar property, which is called "unique factorization".

Definition 1.26 *Let D be an integral domain with unity. Then D is a Unique Factorization Domain provided that $\forall\ a \in D$, where a is neither 0 nor a unit, there exist irreducibles $i_1, i_2, ..., i_n \in D$ so that $a = \prod_{k=1}^{n} i_k$, and any other list of irreducibles whose product is a consists of associates of $i_1, i_2, ..., i_n$.*

This particular type of ring will be the main subject of the rest of this section. Though it may seem trivial and uninteresting, the concept of unique factorization plays a very large role in number theory, solving many problems (like the Ramanujan-Nagell Theorem, as we will show in Chapter 4) while simultaneously making others more difficult (such as foiling Kummer's otherwise proof of Fermat's Last Theorem, see pages 305-308 of reference [7] for more information about this tragic example). It turns out, as we shall now prove, that Unique Factorization Domains are also types of rings that have indistinguishable primes and irreducibles.

Proposition 1.27 *An element of a Unique Factorization Domain is prime if and only if it is irreducible.*

Proof ($\Rightarrow$) That primes are irreducible in a Unique Factorization Domain follows from the fact that all Unique Factorization Domains are integral domains and Proposition 1.16.

($\Leftarrow$) Let D be a Unique Factorization Domain and $i \in D$ be an irreducible in D. We will show that i is a prime element of D. Suppose $ab \in (i)$ for some $a, b \in D$. To show that i is prime, we must show that $a \in (i)$ or $b \in (i)$. By assumption, $ab = qi$ for some $q \in D$. Suppose that a can be written as a product of irreducibles as

$$a = a_1 a_2 ... a_m, \tag{24}$$

and b can be written as a product of irreducibles as

$$b = b_1 b_2 ... b_n. \tag{25}$$

Then $a_1 ... a_m b_1 ... b_n = ab = qi$. But i is an irreducible, and the factorizations of a and b as irreducibles are unique up to associates, so i must associate to one of the irreducibles occurring in the factorization of a or in that of b. With the understanding that any other choice is similar, let i associate to a_1. Then

$$a = (ui) a_2 ... a_m \tag{26}$$

for some $u \in U(D)$ such that $a_1 = ui$ and some list of irreducibles $a_2, ..., a_m$. So $i|a$, thus $a \in (i)$. This implies that (i) is a prime ideal, thus i is prime. $\square$

We have closely followed the proof on page 286 of reference [4] in writing the previous proof.

We will soon see that the fact that Unique Factorization Domains also obey the prime-irreducible indistinguishability property is not a coincidence. For now though, we will take a slight detour and discuss a few special properties of certain integral domains. First, we must establish a definition.

Definition 1.28 *An integral domain D with unity is Noetherian provided that every ideal I in D is generated by a finite set of elements in D.*

The first instinct of the reader may be to connect this to Principal Ideal Domains. It is obviously true that all Principal Ideal Domains are Noetherian by definition. This will become important later.

We shall now prove that Noetherianism has two very important implications. First, in a Noetherian domain, there do not exist infinite chains of ideals that are all proper subsets of each other. Next, in every nonempty set of ideals in this domain, there is an element, which we will call a *maximal element*, that is not contained in any other ideal in the set.

Proposition 1.29 *The following are equivalent:*

(i) An integral domain D is Noetherian.

(ii) Given an "ascending chain of ideals" in D, $I_1 \subseteq I_2 \subseteq ... \subseteq I_n \subseteq ...$, there exists some n for which $I_m = I_n$ for all $m \geq n$.

(iii) For every set $S \neq \varnothing$ of ideals in D, $\exists\ I \in S$ such that $\forall\ J \in S$ such that $J \neq I$, $I \cap J \neq I$. (In other words, if $I \neq J$, then $I \not\subseteq J$.)

Proof ($i \Rightarrow ii$) Let D be a Noetherian integral domain. Consider a chain of ideals of D,

$$I_1 \subseteq I_2 \subseteq ... \subseteq I_n \subseteq ... . \tag{27}$$

We will show that there exists a finite number n for which $\forall\ m \geq n$, $I_n \subsetneq I_m$ becomes false. Let

$$I = \bigcup_{k=1}^{\infty} I_k. \tag{28}$$

First we will show that I is an ideal. Let $a, b \in I$. Then $a \in I_a$ for some $I_{l(a)}$ and $b \in I_{l(b)}$ for some $I_{l(b)}$, where $I_{l(a)}$ and $I_{l(b)}$ are two ideals of the union that defines I. Since these ideals are both in a chain of subset relations, we may conclude that either $I_{l(a)} \subseteq I_{l(b)}$ or $I_{l(b)} \subseteq I_{l(a)}$. Suppose, with the understanding that the other case is similar, that $I_{l(a)} \subseteq I_{l(b)}$. Then $a, b \in I_{l(b)}$, so $a + b \in I_{l(b)} \subseteq I$. Let $x \in D$. Then $xa \in I_{l(a)} \subseteq I$. Closure under addition and multiplication by an element of D indicates that I is an ideal.

Since D is Noetherian, I must be finitely generated. Therefore let

$$I = (a_1, a_2, ..., a_m). \tag{29}$$

Since I is a union of the ideals I_j, there must exist some j, call it $l(i)$, such that $a_i \in I_{l(i)}$ for each of the a_i. Now define $I_L = I_{l(m)}$, where $l(m)$ is the largest of the $l(i)$. (We make this choice for $l(m)$ with the understanding that the argument will be similar if we choose any other $l(i)$ to be the largest.) Now I_L contains all of the other $I_{l(i)}$, since these ideals are all members of the ascending chain, with ideals of larger subscripts containing ideals with smaller ones. Then I_L is guaranteed to contain $a_1, a_2, ...,$ and a_m as elements, since they are elements of ideals that are contained

in I_L. Since it is an ideal, I_L is closed under addition and under multiplication by members of D. In that case, by definition of generated ideal,

$$I = (a_1, a_2, ..., a_m) \subseteq I_L. \tag{30}$$

Therefore, I_L is the ideal at which the ascending chain becomes stationary, since no ideal in the ascending chain can properly contain I, as a matter of its construction.

$(ii \Rightarrow iii)$ Let (ii) be true for an integral domain D and let S be a nonempty set of ideals of D. We will show that there exists some $I_n \in S$ such that $\forall\ I_m \in S$, it is false that $I_n \subsetneq I_m$. Suppose for contradiction that there is no such I_n. Then consider an ideal $I_1 \in S$. Then $\exists\ I_2 \in S$ such that $I_1 \subsetneq I_2$. Likewise, $\exists\ I_3 \in S$ such that $I_2 \subsetneq I_3$. If S is finite, then the contradiction becomes obvious at this step, since then there will certainly be a finite number of times that this procedure can be enacted, and therefore an ideal that is the superset of the final relation, which is not properly contained in any other ideal in the chain. Suppose S is infinite. This creates an infinite ascending chain of ideals

$$I_1 \subsetneq I_2 \subsetneq I_3 \subsetneq ..., \tag{31}$$

by virtue of the Axiom of Choice. But condition (ii) forbids the eternal continuation of the properness of the subset relations. Thus, every ideal in the chain is strictly contained some other ideal of S, but the ascending chain of ideals in S becomes stationary at some $I_n \in S$. This contradiction leads us to conclude that there is some I_n that is not strictly contained in any other ideal in S.

$(iii \Rightarrow i)$ Let D be an integral domain in which every nonempty set of ideals

contains an element that is not properly contained in any other element of the set. Consider an ideal I in D. We will show that I is finitely generated. Let S be the set of all finitely generated ideals that are subsets of I. Then $(0) \in S$, so $S \neq \varnothing$. Therefore, there exists some element $J \in S$ so that J is not a proper subset of any other element of S. Assume for contradiction that $J \neq I$. Then let $b_1 \in I \setminus J$. In that case, the ideal that is generated by b_1 and the generators of J is finitely generated and a strict superset of J, while simultaneously being a subset of I, despite that J is maximal in S, and therefore not a proper subset of any element of S. This contradiction leads us to conclude that I is finitely generated, so D is Noetherian. $\square$

We can now prove the capstone theorem of this chapter, which is that while all Euclidean Domains are Principal Ideal Domains, all Principal Ideal Domains are Unique Factorization Domains.

Theorem 1.30 *Every Principal Ideal Domain is a Unique Factorization Domain.*

Proof Let D be a Principal Ideal Domain and $a \in D$ be neither 0 nor a unit. We will first show that a can be written as a product of finitely many irreducible elements of D. Suppose for contradiction that a is not a product of irreducibles. Then a is not an irreducible, and so $\exists\ r_1, r_2 \in D$, neither of which are units, such that $a = r_1 r_2$. If both of r_1 and r_2 are products of irreducibles, then a is as well, so suppose, with the understanding that the other choice is similar, that r_1 is not the product of irreducibles. Now, we know that $a \nmid r_1$, because otherwise $r_1 = q_0 a$ for some $q_0 \in D$, and so $a = q_0 a r_2$, implying that $q_0 r_2 = 1$, which would imply that r_2 is a unit, despite the fact that we have defined it so that it is not a unit. Thus, by Proposition 1.3, we have that $(a) \subsetneq (r_1)$.

We repeat the argument of the previous paragraph to show that $\exists\, r_{11} \in D$ such that $(r_1) \subsetneq (r_{11})$. Continuing, we find $(r_{11}) \subsetneq (r_{111})$ for some $r_{111} \in D$. These arguments can be applied iteratively, resulting in an ascending chain of ideals:

$$(a) \subsetneq (r_1) \subsetneq (r_{11}) \subsetneq (r_{111}) \subsetneq ..., \tag{32}$$

which extends eternally. But D is a Principal Ideal Domain, hence Noetherian, so by Proposition 1.29, this leads us to a contradiction. Thus, our original assumption that a is not a product of irreducibles must be false, to the effect that a is, in fact,a product of irreducibles.

We will now show that the factorization of a into irreducibles is unique up to associates. We proceed by induction on the number of irreducible factors of some irreducible factorization of a, call it n. Since by assumption, a is neither 0 nor a unit, it follows that $n \geq 1$. Let

$$a = p_1 p_2 ... p_n = i_1 i_2 ... i_m, \tag{33}$$

for $m \geq n$ and all p_j and i_k being (not necessarily distinct) irreducibles. Then it is clear that $p_1 | i_1 i_2 ... i_m$, which implies $i_1 i_2 ... i_m \in (p_1)$. Now, due to Proposition 1.25, (p_1) is a prime ideal, and therefore some one of the factors of $i_1 i_2 ... i_m$ is also an element of (p_1). (It is simple to show that one of the irreducibles in this factorization is therefore an element of (p_1), and we leave this proof to the reader.) Assume, with the understanding that the argument would be similar for a different choice, that $i_1 \in (p_1)$. Then $p_1 | i_1$, so $i_1 = p_1 q_1$ for some $q_1 \in D$. But since i_1 is irreducible, q_1 must be a unit. Therefore, p_1 and i_1 are associates. Since D is an

integral domain, we may cancel p_1 to find

$$p_2...p_n = i'_2 i_3...i_m, \tag{34}$$

where $i'_2 = q_1 i_2$ is irreducible. We will now show that $n = m$. Suppose for contradiction, with the understanding that the other case is similar, that $n < m$. Then after using similar arguments to cancel all of the terms on the left, we find from Equation 34 that

$$1 = i'_{n+1} i'_{n+2}...i'_m, \tag{35}$$

where the i'_k are associates of the i_k for $k > n$. In that case, i'_{n+1} is a unit, and simultaneously an irreducible. This contradiction indicates that $n = m$. So then, each of the factors remaining on the left in Equation 34 matches bijectively with one of the factors on the right, where the bijection is multiplication by a certain unit. □

We will now compare two methods of proving the Fundamental Theorem of Arithmetic, as promised.

Theorem 1.31 *Every and any member of the set $\mathbb{Z} - \{0, \pm 1\}$ can be written as a product of a finite set of primes. This representation is unique, up to the signs and orders of the factors. This is called the Fundamental Theorem of Arithmetic.*

Proof (algebraic) We know from Lemma 1.11 that the ring $\mathbb{Z}$ is a Euclidean Domain. Then Theorem 1.13 implies that $\mathbb{Z}$ is a Principal Ideal Domain, and Theorem 1.30 implies in turn that $\mathbb{Z}$ is a Unique Factorization Domain. Thus, any nonunit

integer can be written as a product of a finite set of irreducibles that is unique up to associates. Proposition 1.27 implies that these irreducibles are exactly primes. □

Proof (number-theoretic) Let $n \in \mathbb{Z} \setminus \{0, \pm 1\}$. First we will show that n can be written as a product of irreducibles. Consider the case that n is prime. Then the theorem is proven. Consider the case that n is composite. Then $\exists\ d \in \mathbb{Z}$ such that $d_0|n$ and $1 < d_0 < n$, since there exist divisors of n other than ± 1, itself, and $-n$. Consider the set

$$\{d > 1 | d < n \text{ and } d|n\}. \tag{36}$$

The Well-Ordering Principle guarantees that there exists a minimal element of this set. This element must be an irreducible, since otherwise there is a smaller divisor of n, contradicting its minimality. Call this irreducible p_1. Then $n = p_1 n_1$ for some n_1 with $1 < |n_1| < |n|$.

Suppose, with the understanding that the proof is complete in the other case, that n_1 is composite. Then $\exists\ d_1 \in \mathbb{Z}$ such that $d_1|n_1$ and $1 < d_1 < n_1$. Let

$$p_2 = \min(\{x > 1 \mid x < n_1, x|n_1\}). \tag{37}$$

As with p_1, p_2 must be an irreducible. Then $n = p_1 p_2 n_2$ for some n_2 such that $1 < |n_2| < |n_1|$.

Suppose, with the understanding that the proof is complete in the other case, that n_2 is composite. Then similar arguments show that $n = p_1 p_2 p_3 n_3$ for some positive irreducible p_3 and some n_3 such that $1 < |n_3| < |n_2|$. Continuing this

process, we obtain a strictly decreasing sequence of positive integers

$$|n| > |n_1| > |n_2| > |n_3| > ... > 1, \tag{38}$$

which certainly ends with some final integer n_k. So $n = p_1p_2p_3...p_kn_k$. Then n_k is not composite, since that would require that $n_k = p_{k+1}n_{k+1}$ for some irreducible p_{k+1} and some n_{k+1} with $|n_k| = |n_{k+1}|$, which implies that $|p_{k+1}| = 1$ for an irreducible, which is a contradiction. Thus, n_k is irreducible, and therefore n can be written as a product of primes.

We will now show that the array of primes whose product is n is unique up to the signs of its elements. Let $n \in \mathbb{Z}$ and

$$n = p_1p_2...p_r = q_1q_2...q_s \tag{39}$$

for some $r \leq s$ be two factorizations of n into primes. Assume, with the understanding that any other case is similar due to commutativity of multiplication, that

$$p_1 \leq p_2 \leq ... \leq p_r \tag{40}$$

and

$$q_1 \leq q_2 \leq ... \leq q_s. \tag{41}$$

Then $p_1|q_1q_2...q_s$. We will show that $p_1|q_k$ for some $k \leq s$. If $p_1|q_1$, then this is clear; suppose not. Then p_1 and q_1 are relatively prime, since p_1 and q_1 are distinct primes, so by Euclid's Lemma of number theory, $p_1|q_1q_2...q_s$ implies that $p_1|q_2q_3...q_s$. If $p_1|q_2$, then once again the claim is proven; suppose not. Then $p_1|q_3q_4...q_s$ by similar

arguments. Likewise, we may repeat this process s times and receive the result that at least one of the following statements must be true: $p_1|q_1$, $p_1|q_2$, ... , $p_1|q_s$. Thus, the claim is proven; $p_1|q_k$ for some k.

Now, q_k, being prime, has no positive divisors other than 1 and q_k. We know that $p_1 \neq 1$, by definition of irreducible, so $p_1 = q_k$. We also know that $q_k \geq q_1$, so $p_1 \geq q_1$. By similar arguments, it is possible to show that $q_1 = p_l$ for some $l \leq r$. This indicates that $q_1 \geq p_1$. Thus, $q_1 = p_1$. Knowing that this prime number must not be zero, we may cancel the common factor and find

$$p_2 p_3 ... p_r = q_2 q_3 ... q_s. \tag{42}$$

Similar arguments show that $p_2 = q_2$, and so $p_3 p_4 ... p_r = q_3 q_4 ... q_s$. One may continue this process for $r - 2$ more steps, to find that $p_m = q_m$ for all $m \leq r$, and that $1 = q_{r+1} q_{r+2} ... q_s$. If we let $r \neq s$, we reach a contradiction, since the product of an array of integer primes can never be 1. Thus, $r = s$, and so the factorizations are one and the same, thus the factorization is unique. □

We have closely followed the proof on page 41 of reference [3] in writing the number theoretic proof of the previous theorem.

It is easy to see that the algebraic proof is much longer, if one includes the proofs of the lemmas in an evaluation of the length of the proof. However, the algebraic proof gives a much fuller understanding of the set of integers in the context of more general mathematical structures, since it considers the ring of integers as only a special case of Euclidean Domains. In this regard, the number-theoretic proof is much

more *ad-hoc*. Further, proving theorems in a general context provides a plethora of other tangentially related results. For example, a purely number-theoretic proof that the Gaussian integers has unique factorization would require considerably different methods than those used above.

Consider, for instance, that the number-theoretic proof above uses the fact that there must exist a number p_1 that is guaranteed to be a smallest positive nonunit integer dividing n. If instead we were attempting to prove that the ring of Gaussian integers (that is, the ring consisting of complex numbers with integer coefficients) has unique factorization, then we would have to consider that n could be, for example, $2 + i$. In that case, there would be no such smallest positive nonunit integer, as the reader can easily show. Thus, completely different methods must be used to prove that the Gaussian integers have unique factorization, if one refuses to use algebraic methods. On the other hand, if one uses an algebraic approach, then it is necessary only to prove a variation of Lemma 1.11 that adapts to the Gaussian integers. In other words, if one can show that the Gaussian integers form an integral domain, and that there is a suitable field norm for the Gaussian integers, then one may use the exact same proof as above to prove that unique factorization is possible in the ring of Gaussian integers.

The ease of proving additional results of number theory is a tremendous advantage of algebraic approaches. Mathematicians became acutely aware of this, and defined what we will study in the next chapter: a type of complex number that is defined by its algebraic characteristics.

2 Number Fields

In this chapter, we will establish many important theorems of algebraic number theory. We begin with a few definitions. The language that they impart will be used to devise algebraic perspectives that will be used for the remainder of this thesis.

Definition 2.1 *Let F be a field, and let $E \subseteq F$. Then E is a subfield of F provided that E is also a field under the additive and multiplicative operations of F.*

In the case that E is a subfield of F, we say that F *is a field extension of* E, or that F *extends* E, and this fact is denoted $F : E$.

Definition 2.2 *Let $F : E$ be a field extension. Then consider the vector space $\mathcal{F}$, whose vectors are the members of F. Let addition of vectors $v_1, v_2 \in F$ be defined as $v_1 + v_2$ as in F, and let scalar multiplication of $v \in F$ by $s \in E$ be defined as sv as in F. The dimension of this vector space is called the extension degree of F over E.*

We denote the extension degree of F over E as $[F : E]$. If this extension degree is finite, we call $F : E$ a finite extension. Further, if we let $\mathcal{F}$ be defined as above, then we will often refer to the vector space $\mathcal{F}$ as the vector space "F over E". Now we are ready to define the concept that will be the main focus of this chapter.

Definition 2.3 *Let $F : E$, and suppose $\alpha \in F$. Then α is called algebraic over E provided that $\exists\ p \in E[x]$ such that $p \neq 0$ and $p(\alpha) = 0$. Also, α is called transcendental over E provided that it is not algebraic over E.*

This definition may be rather foreign to the reader, as it may appear that we are now working backwards, in a sense; in the previous chapter we defined algebraic concepts and then found which elements obey them. Now we are defining elements to have certain algebraic characteristics. We therefore must ask for patience, since we must first introduce a few more definitions and propositions before we begin to discuss the fabulously interesting results that pertain to these strange objects. For now, we set up a result from the definitions that we have given so far.

Proposition 2.4 *If $D \subseteq E \subseteq F$ are fields and $F : E$ and $E : D$ are finite extensions, then $[F : D] = [F : E][E : D]$.*

Proof Let A be a basis of F over E, and let B be a basis of E over D. Define $A = \{a_1, a_2, ..., a_n\}$ and $B = \{b_1, b_2, ..., b_m\}$, where $n = [F : E]$ and where $m = [E : D]$. Define

$$AB = \{a_i b_j | a_i \in A \text{ and } b_j \in B\} \tag{43}$$

(let this multiplication be that of F). We will show that AB is a basis of F over D.

First, we will show that AB spans the vector space F over D. Let $v \in F$. Then

$$v = e_1 a_1 + e_2 a_2 + ... + e_n a_n \tag{44}$$

for some $e_1, e_2, ..., e_n \in E$, since A is a basis of F over E. But any

$$e_i = d_{i1} b_1 + d_{i2} b_2 + ... + d_{im} b_m = \sum_{j=1}^{m} d_{ij} b_j \tag{45}$$

for some $d_{i1}, d_{i2}, ..., d_{im} \in D$, since B is a basis of E over D. Thus,

$$v = a_1 \sum_{j=1}^{m} d_{1j} b_j + a_2 \sum_{j=1}^{m} d_{2j} b_j + ... + a_n \sum_{j=1}^{m} d_{nj} b_j = \sum_{i=1}^{n} a_i \sum_{j=1}^{m} d_{ij} b_j. \quad (46)$$

It is not hard to see that each term of the sum in Equation 46 contains $a_i b_j$ for some $1 \leq i \leq n$ and some $1 \leq j \leq m$. Therefore, x is a linear combination of the $a_i b_j$, thus, AB spans F over D.

We will show that AB is linearly independent. Suppose that

$$x_{1,1}a_1b_1 + x_{1,2}a_1b_2 + ... + x_{1,m}a_1b_m + x_{2,1}a_2b_1 + x_{2,2}a_2b_2 + ... + x_{n,m}a_nb_m = 0, \quad (47)$$

for some $x_{1,1}, x_{1,2}, ..., x_{n,m} \in D$. Then we can rearrange this as

$$(x_{1,1}b_1 + x_{1,2}b_2 + ... + x_{1,m}b_m)a_1 + (x_{2,1}b_1 + x_{2,2}b_2 + ... + x_{2,m}b_m)a_2 + ... + (x_{n,1}b_1 + ... + x_{n,m}b_m)a_n = 0. \quad (48)$$

Because A is linearly independent, we know that each of the coefficients multiplying the a_i in the previous equation must be equal to 0. Because B is linearly independent, that implies that each of the coefficients multiplying each of the b_j in the previous equation must be equal to 0. Thus, all of the coefficients must be equal to 0; AB is linearly independent as an implication.

Now that we have this, and the fact that AB spans F over D, we have that AB is a basis of F over D. We will show that AB has a cardinality of nm. The counterclaim is that AB was defined as a set containing nm number of not necessarily

distinct elements, and therefore, $|AB| \leq nm$, but they are not necessarily equal. Therefore, we will show that the elements of AB are distinct. Suppose for contradiction, with the understanding that any other choice is similar, that $a_1b_1 = a_2b_2$. In that case, $a_1b_1 - a_2b_2 = 0$, so

$$s_1a_1b_1 + s_2a_2b_2 = 0, \tag{49}$$

while $s_1 = 1$ and $s_2 = -1$, which are not both zero. Thus, $\{a_1b_1, a_2b_2\}$ is linearly dependent. But we know that $\{a_1b_1, a_2b_2\} \subseteq AB$. We have established that AB is linearly independent, hence, all of its subsets are linearly independent. Thus, $\{a_1b_1, a_2b_2\}$ is simultaneously linearly dependent and linearly independent. This contradiction leads us to conclude that all of the elements of AB are distinct, and therefore, that AB has a cardinality of nm. Thus, while AB is a basis of F over D, its cardinality is $[F : D] = nm = |A||B| = [F : E][E : D]$. $\square$

Next, we will introduce our first result about algebraic numbers, which will motivate a definition for obvious reasons.

Theorem 2.5 *Let $F : E$ and α be algebraic over E. Then there exists a unique monic polynomial of minimum degree, called $p \in E[x]$, satisfying $p(\alpha) = 0$.*

Proof Suppose that α is algebraic over the field E. We will show that there is a unique monic polynomial $p \in E[x]$ with a minimum degree such that $p(\alpha) = 0$. First, we will show that such a monic polynomial with a minimum degree exists. We know that $\exists\, p \in E[x]$ such that $p = 0$ and $p(\alpha) = 0$, since α is algebraic over

F. Let $p = e_n x^n + e_{n-1}x^{n-1} + ... + e_1 x + e_0$. Then

$$e_n\alpha^n + e_{n-1}\alpha^{n-1} + ... + e_1\alpha + e_0 = p(\alpha) = 0. \tag{50}$$

Our convention, as stated in the preface, is that $e_n \neq 0$, unless $p = 0$ (in this case, it does not by definition). In that case, we can divide both sides of the equation by e_n, finding

$$\alpha^n + \frac{e_{n-1}}{e_n}\alpha^{n-1} + ... + \frac{e_1}{e_n}\alpha + \frac{e_0}{e_n} = 0. \tag{51}$$

This indicates that the monic polynomial $p_m(x) = x^n + \frac{e_{n-1}}{e_n}x^{n-1} + ... + \frac{e_1}{e_n}x + \frac{e_0}{e_n}$ (which must exist in $E[x]$, since E is a field and $e_n \neq 0$) has a root at $x = \alpha$. Thus, a monic polynomial satisfying $p(\alpha) = 0$ exists in $E[x]$. Now, the degrees of the monic polynomials in $E[x]$ that satisfy this condition are all positive integers, and so, by the Well-Ordering Principle, there must exist a minimum among them, and so there exists a monic polynomial with minimum degree that satisfies $p_m(\alpha) = 0$ in $E[x]$.

We will now show that the monic polynomial in $E[x]$ with roots at $x = \alpha$ that has minimum degree is unique. Suppose that $p(x) = x^n + e_{n-1}x^{n-1} + ... + e_1 x + e_0$ is a polynomial satisfying $p(\alpha) = 0$ and has a degree that is minimal among the monic polynomials in $E[x]$ satisfying $p(\alpha) = 0$. Suppose for contradiction that the polynomial $p_2(x) = x^n + c_{n-1}x^{n-1} + ... + c_1 x + c_0 \in E[x]$ matches the same description, but $p_2 \neq p$. Then since $p(\alpha) = 0$ and $p_2(\alpha) = 0$, $p(\alpha) - p_2(\alpha) = 0$. Thus,

$$0 = (e_{n-1} - c_{n-1})\alpha^{n-1} + ... + (e_1 - c_1)\alpha + (e_0 - c_0), \tag{52}$$

and this polynomial must be nonzero by assumption that $p_2 \neq p$.

Suppose, with the understanding that the proof is similar in any other case, that $e_{n-1} \neq c_{n-1}$. In that case, we may divide both sides of the equation by $e_{n-1} - c_{n-1}$ and find that

$$0 = \alpha^{n-1} + ... + \frac{e_1 - c_1}{e_{n-1} - c_{n-1}}\alpha + \frac{e_0 - c_0}{e_{n-1} - c_{n-1}}. \tag{53}$$

But then this implies the existence of $p_3(x) = x^{n-1} + ... + \frac{e_1 - c_1}{e_{n-1} - c_{n-1}}x + \frac{e_0 - c_0}{e_{n-1} - c_{n-1}}$, a monic polynomial in $E[x]$ with a root at $x = \alpha$ such that $\partial p_3 = n - 1 < n = \partial p$. This is in spite of the fact that p has the minimum degree among monic polynomials with a root at $x = \alpha$ in $E[x]$. This contradiction leads us to conclude that $p_2 \neq p$ is necessarily false, and so the nonzero monic polynomial p satisfying $p(\alpha) = 0$ that has minimum degree is unique. □

With this theorem in place, it is only natural for us to give a name to this special polynomial.

Definition 2.6 *Let $F : E$, and let $\alpha \in F$ be algebraic over E. Then $p \in E[x]$ is the minimum polynomial of α over E provided that it is the monic polynomial of minimum degree satisfying $p(\alpha) = 0$.*

In other words, the monic polynomial with coefficients in a field that has a root at some algebraic number over that field is called the minimum polynomial of that algebraic number over that field, provided that it is of minimal degree. As said before, Theorem 2.5 guarantees the existence of the minimum polynomial and justifies our use of the word "the" in the definition.

Before moving on, we should introduce some definitions and notation.

Definition 2.7 *Let $F : E$ and $\alpha \in F$. Then the simple extension of E by α, denoted $E(\alpha)$, is defined as the smallest subfield of F such that $\alpha \in E(\alpha)$ and $E(\alpha) : E$. In other words, $E(\alpha)$ is the simple extension of E by α provided that for any field D, $D : E$ and $\alpha \in D$ implies that $D : E(\alpha)$.*

Likewise, we let $E[\alpha]$ denote the smallest subring of F such that $\alpha \in E[\alpha]$ and E is a subring of $E[\alpha]$, defined rigorously in a similar way. At first this notation may seem to be confused with that of the ring of polynomials, but in fact the two mesh perfectly, since $E[\alpha]$ is, in fact, the ring of polynomials in α rather than the variable x. We will prove this with a proposition.

Proposition 2.8 *Let $F : E$ with $\alpha \in F$. Then $E[\alpha] = \{f(\alpha) | f \in E[x]\}$.*

Proof Let $\alpha \in F$, where $F : E$, and consider the ring

$$R = \{f(\alpha) | f(x) \in E[x]\}, \tag{54}$$

where multiplication is defined as traditional multiplication of polynomials and addition is defined as traditional addition of polynomials. We will show that R satisfies the definition of the ring $E[\alpha]$. First of all, we know that $1x \in E[x]$, so evaluating at $x = \alpha$, we have that $\alpha \in R$. Now let $e \in E$. We know that $e \in E[x]$, since it is a constant polynomial. Then evaluating at $x = \alpha$, $e \in R$. Thus, $E \subseteq R$. Now E, being a ring in and of itself, is closed under multiplication and addition, and therefore E is a subring of R, which also contains α.

Now suppose that E is a subring of S and $\alpha \in S$. We will show that R is a

subring of S. Let $r = e_n\alpha^n + e_{n-1}\alpha^{n-1} + ... + e_1\alpha + e_0 \in R$. Because of the closure of rings under multiplication, S, being a ring, must contain all powers of α, since it contains α itself. Likewise, since $E \subseteq S$, any element of E that multiplies a power of α must be an element of S. Finally, the sum of any finite number of elements of S must be an element of S. These facts establish that

$$r = e_n\alpha^n + e_{n-1}\alpha^{n-1} + ... + e_1\alpha + e_0 \in S. \tag{55}$$

Thus, $R \subseteq S$, and it follows easily that R is a subring of S. Hence, R is the minimal ring containing E and α, thus $R = E[\alpha]$. $\square$

Next, we will prove a corresponding proposition for the simple extensions.

Proposition 2.9 *Let $F : E$ and suppose that $\alpha \in F$. Then the simple extension $E(\alpha)$ can be described as the set $E(\alpha) = \left\{ \frac{f(\alpha)}{g(\alpha)} \middle| f, g \in E[x], g(\alpha) \neq 0 \right\}$.*

Proof Let $\alpha \in F$, where $F : E$, and consider the field

$$Q = \left\{ \frac{f(\alpha)}{g(\alpha)} \middle| f, g \in E[x], g(\alpha) \neq 0 \right\}. \tag{56}$$

We could define $\frac{f(x)}{g(x)} = \frac{x}{1}$, in which case $\alpha = \frac{\alpha}{1} = \frac{f(\alpha)}{g(\alpha)} \in Q$. Now let $e \in E$. We could define $\frac{f(x)}{g(x)} = \frac{e}{1}$, in which case $e = \frac{f(\alpha)}{g(\alpha)} \in Q$. Thus, $E \subseteq Q$. Now E, being a field in and of itself, is closed under multiplication, addition, and multiplicative inverses, and therefore E is a subfield of Q, which also contains α.

Now suppose that $P : E$ and $\alpha \in P$. We will show that $P : Q$. Let $a \in Q$.

Then $a = \frac{f(\alpha)}{g(\alpha)}$ for some $f, g \in E[x]$ with $g(\alpha) \neq 0$. Suppose that

$$f(x) = e_n x^n + e_{n-1} x^{n-1} + ... + e_1 x + e_0 \tag{57}$$

and

$$g(x) = c_m x^m + c_{m-1} x^{m-1} + ... + c_1 x + c_0. \tag{58}$$

Then $a = \frac{f(\alpha)}{g(\alpha)} = \frac{e_n \alpha^n + e_{n-1} \alpha^{n-1} + ... + e_1 \alpha + e_0}{c_m \alpha^m + c_{m-1} \alpha^{m-1} + ... + c_1 \alpha + c_0}$. Now, by the same arguments that were used in Proposition 2.8, we have that the numerator and the denominator of a are both elements of P. Because P is a field and the denominator is nonzero, it is clear that $a \in P$. Thus, $Q \subseteq P$, and it follows easily that $P : Q$. Hence, Q is the minimal field containing E and α, thus $Q = E(\alpha)$. $\square$

We will also establish the notation of non-simple extensions. Let $F : E$ with $\alpha_1, \alpha_2, ..., \alpha_n \in F$. Then let $E_1 = E(\alpha_1)$. We define $E(\alpha_1, \alpha_2) = E_1(\alpha_2)$. Recursively, we define $E(\alpha_1, \alpha_2, ..., \alpha_n) = E_{n-1}(\alpha_n)$. It is not hard to show that this field is the minimal field containing E and $\alpha_1, \alpha_2, ..., \alpha_n$, since each additional extension adds the minimal amount of elements to the field, by definition of simple extension.

At this point we will proceed to prove that a simple extension is finite if and only if the extending element is algebraic over the subfield. But before this, we will need some supporting lemmas, and a few propositions. We will begin by proving that any minimum polynomial is an irreducible in its polynomial ring.

Proposition 2.10 *Let $F : E$, and assume that $\alpha \in F$ is algebraic over E. Suppose that $p \in E[x]$ is the minimum polynomial of α over E. Then p is irreducible in $E[x]$.*

Proof Let $p \in E[x]$ be the minimum polynomial of $\alpha \in F$ over E. We will show that p is irreducible. Suppose that $p = fg$ for some $f, g \in E[x]$. Assume for contradiction that neither f nor g are units. Then neither are constant polynomials other than 0 because $U(E[x]) = E \setminus \{0\}$. We also know that $f \neq 0 \neq g$, because if either is zero, then their product, the minimum polynomial, is the zero polynomial, which violates the definition of minimum polynomial. So, with the understanding that f and g are both nonzero and nonconstant,

$$p(\alpha) = f(\alpha)g(\alpha) = 0. \tag{59}$$

Since E is an integral domain, we see that $f(\alpha) = 0$ or $g(\alpha) = 0$. Suppose, with the understanding that the other case is similar, that $f(\alpha) = 0$. But we know that $\partial p = \partial f + \partial g$, so

$$\partial f = \partial p - \partial g < \partial p. \tag{60}$$

Thus, f has a root at $x = \alpha$ and has a degree less than the minimum polynomial. Since the minimum polynomial is the unique monic one of smallest degree that has a root at $x = \alpha$, f must not be monic. Therefore,

$$f(x) = e_n x^n + e_{n-1}x^{n-1} + ... + e_1 x + e_0, \tag{61}$$

for some $e_n, e_{n-1}, ..., e_1, e_0 \in E$. Because E is a field and $e_n \neq 0$, we may define

$$f_0(x) = \frac{f(x)}{e_n} = x^n + \frac{e_{n-1}}{e_n}x^{n-1} + ... + \frac{e_1}{e_n}x + \frac{e_0}{e_n} \in E[x]. \tag{62}$$

It is clear that $\partial f_0 = \partial f < \partial p$, and it is also clear that $f(x) = 0$ if and only if $f_0(x) = 0$, so f_0 has a root at $x = \alpha$. Thus, $f_0 \in E[x]$ is a monic polynomial with a root at $x = \alpha$ that has a degree less than the minimum polynomial of α over E. This contradiction leads us to conclude that our original assumption that neither f nor g are units must be false, and therefore, at least one of them is necessarily a unit in $E[x]$. Therefore, p is irreducible over $E[x]$. $\square$

This next lemma, which we will not prove in this thesis, will establish that the ring of polynomials with coefficients in a field is a Euclidean Domain.

Lemma 2.11 *Let F be a field. Then $F[x]$ is a Euclidean Domain with $N(f) = \partial f$ for all $f \in F[x]$ as the field norm.*

Proof We will omit this proof. The interested reader may see pages 85-87 of reference [6] for a proof of this statement. $\square$

Now we will introduce a lemma that provides a polynomial analog of a simple theorem from number theory. For any $a, b \in \mathbb{Z}\ \exists\ x, y \in \mathbb{Z}$ such that $ax + by = d$, where d is any greatest common divisor of a and b. Thus, if a and b are relatively prime, then there exist x and y such that $ax + by = 1$. Now we will prove a similar thing for polynomials.

Lemma 2.12 *Let F be a field. If $f_1, f_2 \in F[x]$ are relatively prime in $F[x]$, then $\exists\ h_1, h_2 \in F[x]$ such that $h_1 f_1 + h_2 f_2 = 1$.*

Proof Let $f_1, f_2 \in F[x]$ be relatively prime in $F[x]$. Consider the set

$$S = \{yf_1 + zf_2 \neq 0 | y, z \in F[x]\}, \tag{63}$$

and also the set

$$\partial S = \{\partial l | l \in S\}. \tag{64}$$

By the Well-Ordering Principle, $\exists\ s \in \partial S$ such that $\forall\ n \in \partial S$, $s \leq n$. Thus, $\exists\ f_0 \in S$ such that $\partial f_0 = s$ and $f_0 = yf_1 + zf_2$ for some $y, z \in F[x]$. Define $f = \frac{f_0}{c}$, $h_1 = \frac{y}{c}$, and $h_2 = \frac{z}{c}$, where c is the leading coefficient of f_0. It is clear that f is monic, as a matter of its construction. Now,

$$f = h_1 f_1 + h_2 f_2. \tag{65}$$

We will now show that $f = 1$. Lemma 2.11 implies that $\exists\ q, r \in F[x]$ such that $f_1 = qf + r$, where $\partial r < \partial f$ or $r = 0$. This means that $r \notin S$, because f was defined to have the same degree as f_0, which is minimal in ∂S. But

$$f_1 = qf + r = q(h_1 f_1 + h_2 f_2) + r = qh_1 f_1 + qh_2 f_2 + r, \tag{66}$$

and so

$$r = f_1 - qh_1 f_1 - qh_2 f_2 = (1 - qh_1)f_1 + (-qh_2)f_2. \tag{67}$$

By definition of ring, $1 - qh_1, -qh_2 \in F[x]$, so $r \in S$ or $r = 0$. But we have established that $r \notin S$, so $r = 0$. Thus, $f_1 = qf$, implying that $f | f_1$. Similar arguments show that $f | f_2$. Thus, f is a common divisor of two relatively prime elements of

$F[x]$, and is therefore a unit in $F[x]$, hence a constant polynomial. In that case, since we have defined f to be monic, it follows that $f = 1$, because 1 is the only monic constant polynomial. Thus, Equation 65 implies that h_1 and h_2 that satisfy $h_1 f_1 + h_2 f_2 = 1$ must exist. $\square$

Now we are ready to establish another proposition about the notation of simple extensions.

Proposition 2.13 *Let $F : E$ and $\alpha \in F$ be algebraic over E. Then $E(\alpha) = E[\alpha]$.*

Proof ($\subseteq$) Let $F : E$ and $\alpha \in F$, where α is algebraic over E. We will show that $E(\alpha) \subseteq E[\alpha]$. Let $p(x)$ be the minimum polynomial of α over E, and suppose that $\frac{f(\alpha)}{g(\alpha)} \in E(\alpha)$, with $g(\alpha) \neq 0$. Now, for any polynomial $f_1 \in F[x]$ such that $p|f_1$ in $F[x]$, we must have that $f_1(\alpha) = 0$, since

$$f_1(\alpha) = p(\alpha)q(\alpha) = 0q(\alpha) = 0 \tag{68}$$

for some $q \in F[x]$.

Hence, any polynomial that divides the minimum polynomial of α over E must have a root at α. The contrapositive of this implies that because $g(\alpha) \neq 0$, $p \nmid g$. Further, by Proposition 2.10, we have that any common divisor of p and g must either be an associate of p or a unit in $E[x]$. Yet no associate of p can divide g, since p divides any associate of p, and therefore, an associate of p dividing g would imply that $p|g$, which we have shown to be false. Thus, any common divisor of p and g must be a unit; the two are relatively prime.

Therefore, by Lemma 2.12, $\exists\, h_1, h_2 \in E[x]$ such that $1 = h_1 p + h_2 g$. By letting

$x = \alpha$, we see that

$$1 = h_1(\alpha)p(\alpha) + h_2(\alpha)g(\alpha) = h_2(\alpha)g(\alpha), \tag{69}$$

since $p(\alpha) = 0$. Therefore, $\frac{f(\alpha)}{g(\alpha)} = \frac{h_2(\alpha)f(\alpha)}{h_2(\alpha)g(\alpha)} = h_2(\alpha)f(\alpha)$. Now, since $E[x]$ is a ring, and $f, h_2 \in E[x]$, $fh_2 \in E[x]$, so

$$f(\alpha)h_2(\alpha) \in E[\alpha]. \tag{70}$$

Thus, $E(\alpha) \subseteq E[\alpha]$.

($\supseteq$) Let $f(\alpha) \in E[\alpha]$. Then $f(\alpha) = \frac{f(\alpha)}{1} \in E(\alpha)$. This implies that $E[\alpha] \subseteq E(\alpha)$, and so $E(\alpha) = E[\alpha]$. $\square$

With all this notation established, we move on to another proposition about algebraic numbers.

Proposition 2.14 *Let $F : E$ and $\alpha \in F$. Then α is algebraic over E if and only if $E(\alpha) : E$ is a finite field extension.*

Proof ($\Rightarrow$) Let $\alpha \in F$ be algebraic over E, a subfield of F. We will show that $[E(\alpha) : E]$ is finite. Let $p = x^m + e_{m-1}x^{m-1} + ... + e_1x + e_0$ be the minimum polynomial of α over E. Consider the vector space V that is spanned over E by the vectors $1, \alpha, \alpha^2, ..., \alpha^{m-1}$, where $m = \partial p$, vector addition is defined as addition in $E(\alpha)$, and scalar multiplication is defined as multiplication by elements of F. (The vectors $1, \alpha, \alpha^2, ..., \alpha^{m-1}$ must all be linearly independent, since otherwise there would exist a nonzero polynomial $c_{m-1}x^{m-1} + ... + c_2x^2 + c_1x + c_0$ that has a root

at $x = \alpha$. Dividing this polynomial by the coefficient c_{m-1}, we get a new monic polynomial with a degree less than that of p which has a root at $x = \alpha$, which is a contradiction of the definition of the minimum polynomial.)

We will show by mathematical induction that $\forall\, n \in \mathbb{Z}^+, \alpha^n \in V$. Consider that

$$\begin{aligned}\alpha^m = 0 + \alpha^m &= -p(\alpha) + \alpha^m \\ &= -\alpha^m - e_{m-1}\alpha^{m-1} - ... - e_1\alpha - e_0 + \alpha^m \\ &= -e_{m-1}\alpha^{m-1} - ... - e_1\alpha - e_0. \quad (71)\end{aligned}$$

Thus, $\alpha^m = \alpha^{m-1+1} \in V$. Assume that $\alpha^{m-1+k} \in V$. Then

$$\alpha^{m-1+k} = s_0 + s_1\alpha^1 + s_2\alpha^2 + ... + s_{m-1}\alpha^{m-1}, \quad (72)$$

for some $s_0, s_1, ..., s_{m-1} \in E$. In that case,

$$\begin{aligned}\alpha^{m+k} &= s_0\alpha^1 + s_1\alpha^2 + s_2\alpha^3 + ... + s_{m-1}\alpha^m \\ &= s_0\alpha^1 + s_1\alpha^2 + s_2\alpha^3 + ... + s_{m-1}(-e_{m-1}\alpha^{m-1} - ... - e_1\alpha - e_0) \in V. \quad (73)\end{aligned}$$

Therefore, any power of α is a member of V.

Any element of $E[\alpha]$ can be written as a linear combination of some powers of α multiplied by elements of E. Therefore, Proposition 2.13 implies that any vector in $E(\alpha)$ can be written as a linear combination of some powers of α multiplied by elements of E. Because V is closed under addition and multiplication by elements of E, this implies that any vector in $E(\alpha)$ is an element of V, since any terms with powers of α higher than m will reduce to linear combinations of powers of α less

than m (as we have just inductively proven). Therefore $E(\alpha) \subseteq V$. It is also clear from the definition of V that $V \subseteq E(\alpha)$. Thus, $E(\alpha)$ is a finite dimensional vector space over E, with dimension m, the order of the basis $\{1, \alpha, \alpha^2, ..., \alpha^{m-1}\}$, and degree of the minimum polynomial of α over E.

($\Leftarrow$) Let $[E(\alpha) : E] = n \in \mathbb{Z}^+$. We know that any subset of a vector space that has a cardinality greater than the dimension of said vector space must be linearly dependent, so any subset of $E(\alpha)$ that more than n elements must be linearly dependent. Therefore, the vectors $1, \alpha, \alpha^2, ..., \alpha^n$ are linearly dependent. Since these are linearly dependent, there exist $s_0, s_1, ..., s_n \in E$ which are not all equal to 0, such that $s_0 1 + s_1\alpha + s_2\alpha^2 + ... + s_n\alpha^n = 0$. This equation indicates that α is a root of the nonzero polynomial $s_0 + s_1 x + s_2 x^2 + ... + s_n x^n \in E[x]$. Therefore, α is algebraic. $\square$

Corollary 2.15 *Let $\alpha \in F$ be algebraic over a subfield E of F and let p be the minimum polynomial of α over E. Then $[E(\alpha) : E] = \partial p$.*

Proof The first half of the above proof justifies this claim. $\square$

Next, we will define the algebraic structure that is of greatest interest to us in the context of algebraic numbers.

Definition 2.16 *Let N be a subfield of $\mathbb{C}$. Then N is called a number field provided that $[N : \mathbb{Q}]$ is finite.*

This marks a considerable loss in generality. We now move from working with fields in general to working with fields that are related to $\mathbb{Q}$. For future reference,

we will refer to $\mathbb{A}$ as the set of elements of $\mathbb{C}$ that are algebraic over $\mathbb{Q}$.

Before moving on, we should establish a few lemmas. First of all, it turns out that number fields contain only algebraic numbers.

Lemma 2.17 *If N is a number field, then $\forall\ \alpha \in N$, $\alpha \in \mathbb{A}$.*

Proof Let $\alpha \in N$, where N is a number field. We will show that α is algebraic over $\mathbb{Q}$. Let $[N : \mathbb{Q}] = n$. (This is finite by definition of number field.) Consider the set $S = \{1, \alpha, \alpha^2, ..., \alpha^n\} \subseteq N$. We consider two cases: either the elements of S are distinct, or they are not.

Consider the case that the elements of S are not distinct. In that case, for some distinct $0 \leq i, j \leq n$, $\alpha^i = \alpha^j$. We define the polynomial $f(x) = x^i - x^j \in \mathbb{Q}[x]$. Then we deduce that $f(\alpha) = \alpha^i - \alpha^j = 0$, implying that α satisfies a nonzero polynomial equation with coefficients in the rational numbers. Hence, $\alpha \in \mathbb{A}$.

Consider the case that the elements of S are distinct. Then the set S is linearly independent, since its cardinality is $n + 1$, which is greater than the dimension of the vector space N. This implies that there exist scalars $s_0, s_1, ..., s_n \in \mathbb{Q}$, of which at least one is nonzero, such that

$$s_0 + s_1\alpha + s_2\alpha^2 + ... + s_n\alpha^n = 0. \tag{74}$$

But this linear combination corresponds to

$$g(x) = s_0 + s_1x + s_2x^2 + ... + s_nx^n \in \mathbb{Q}[x], \tag{75}$$

a nonzero polynomial equation where $g(\alpha) = 0$. Thus $\alpha \in \mathbb{A}$. $\square$

Additionally, the set $\mathbb{A}$ of algebraic numbers forms a field.

Lemma 2.18 *The field $\mathbb{A}$ is a subfield of $\mathbb{C}$.*

Proof Let $\alpha, \beta \in \mathbb{A}$. Then Proposition 2.14 implies that $\mathbb{Q}(\alpha) : \mathbb{Q}$ is finite. Now, since β is algebraic over $\mathbb{Q}$, it must certainly be algebraic over $\mathbb{Q}(\alpha)$, because of the fact that $\mathbb{Q}(\alpha) : \mathbb{Q}$, and therefore, any polynomial with coefficients in $\mathbb{Q}$ will have coefficients in $\mathbb{Q}(\alpha)$. Thus, Proposition 2.14 implies that $\mathbb{Q}(\alpha, \beta) : \mathbb{Q}(\alpha)$ is finite. Therefore, Proposition 2.4 indicates that

$$[\mathbb{Q}(\alpha, \beta) : \mathbb{Q}] = [\mathbb{Q}(\alpha, \beta) : \mathbb{Q}(\alpha)][\mathbb{Q}(\alpha) : \mathbb{Q}]. \tag{76}$$

Thus, $\mathbb{Q}(\alpha, \beta) : \mathbb{Q}$ is finite. In that case, $\mathbb{Q}(\alpha, \beta)$ is a number field, so by Lemma 2.17, it follows that every element of $\mathbb{Q}(\alpha, \beta)$ is algebraic. We have defined $\mathbb{Q}(\alpha, \beta)$ as a field. Therefore, since $\alpha, \beta \in \mathbb{Q}(\alpha, \beta)$, $\alpha + \beta, -\alpha, -\beta, \alpha\beta \in \mathbb{Q}(\alpha, \beta) \subseteq \mathbb{A}$, and $\frac{1}{\alpha}, \frac{1}{\beta} \in \mathbb{Q}(\alpha, \beta) \subseteq \mathbb{A}$ if $\alpha \neq 0 \neq \beta$. Thus, $\mathbb{A}$ is a field. $\square$

Corollary 2.19 *If N is a number field, then $\mathbb{A} : N$.*

Proof Let N be a number field. We know from Lemma 2.17 that $N \subseteq \mathbb{A} \subseteq \mathbb{C}$. We also know that N is closed under all of the same operations as $\mathbb{A}$, which is a field according to Lemma 2.18, and thus N is a subfield of $\mathbb{A}$. $\square$

We have closely followed the proof on page 39 of reference [8] in writing the proof of the previous lemma.

The fact is, the field of algebraic numbers $\mathbb{A}$ in $\mathbb{C}$ is not as interesting as some of its subfields, in particular, the number fields. We see that Proposition 2.14 guarantees that any simple extension of the rationals by an algebraic number is a number field. Soon we will prove the immensely non-intuitive result that the converse is also true; any number field is a simple extension of the field of rationals by an algebraic number. However, we must first build up a few definitions and results. We continue to characterize the minimum polynomial using the next lemma.

Lemma 2.20 *Suppose that $F : E$, and $\alpha \in F$ is algebraic over E, with minimum polynomial p over E. Then for any polynomial $f \in E[x]$, $f(\alpha) = 0$ if and only if $p|f$ in $E[x]$.*

Proof ($\Rightarrow$) We will show that the minimum polynomial of $\alpha \in F$ over E divides any other polynomial in $E[x]$ of which α is a root. Let p be the minimum polynomial of α over E and suppose that $f(\alpha) = 0$ for some $f \in E[x]$. Then by Lemma 2.11, we have that

$$f = qp + r, \tag{77}$$

for some $q, r \in E[x]$ such that $\partial r < \partial p$ or $r = 0$. Then

$$0 = f(\alpha) = q(\alpha)p(\alpha) + r(\alpha) = 0 + r(\alpha) = r(\alpha). \tag{78}$$

But since $\partial r < \partial p$, r is a polynomial with a root at α of degree less than the minimum polynomial. Then since the minimum polynomial is unique, we see that r must not be monic. Thus,

$$r(x) = e_n x^n + e_{n-1}x^{n-1}...e_1 x + e_0 \tag{79}$$

for some $e_0, e_1, ..., e_n \in E[x]$ and $n < \partial p$. Suppose for contradiction that not all of the e_i are equal to 0. Then one can divide by the leading coefficient e_n to find a new polynomial:

$$r_1(x) = x^n + \frac{e_{n-1}}{e_n}x^{n-1}...\frac{e_1}{e_n}x + \frac{e_0}{e_n}. \tag{80}$$

We see that $r_1(\alpha) = 0$, because $r_1(\alpha) = \frac{r(\alpha)}{e_n} = \frac{0}{e_n} = 0$. We also see that $\partial r_1 = \partial r < \partial p$. Thus, r_1 is a monic polynomial with a root at α that has a degree less than the minimum polynomial. This contradiction leads us to conclude that our assumption was false, and that $r = 0$. Thus, since $f = qp + r$, we have that $f = qp$, hence $p|f$ in $E[x]$.

($\Leftarrow$) Refer to the first paragraph of the proof of Proposition 2.13. This establishes that, for any $f \in F[x]$, if $p|f$, then $f(\alpha) = 0$. But if $f \in F[x]$, then certainly $f \in E[x]$, since $F : E$. Therefore, for any $f \in E[x]$ such that $p|f$, $f(\alpha) = 0$. $\square$

Proposition 2.21 *Let $F : E$, and $f \in E[x]$ be a nonzero polynomial. Then $a \in F$ is a root of f if and only if $(x - a)|f$ in $F[x]$.*

Proof ($\Rightarrow$) Let $f \in E[x]$ be nonzero, and suppose that $f(a) = 0$ for some $a \in F$. We will show that $(x - a)|f(x)$ in $F[x]$. By Lemma 2.11, we have that

$$f(x) = q(x)(x - a) + r(x), \tag{81}$$

for some $q, r \in F[x]$ such that $\partial r < \partial(x - a)$ or $r = 0$. Now, $\partial(x - a) = 1$, so $\partial r < 1$, hence, r is a constant polynomial. Equation 81 implies that

$$0 = f(a) = q(a)(a - a) + r(a) = r(a). \tag{82}$$

Yet r is a constant polynomial, so the only way for $r(a) = 0$ is if $r = 0$. Therefore,

$$f(x) = q(x)(x - a). \tag{83}$$

Since $q \in F[x]$, this implies that $(x - a)|f(x)$ in $F[x]$.

($\Leftarrow$) Let $(x - a)|f(x)$ in $F[x]$. Then $f(x) = (x - a)g(x)$ for some $g \in F[x]$. Therefore, $f(a) = (a - a)g(a) = 0$. $\square$

Corollary 2.22 *Let $F : E$ and suppose that $f \in E[x]$ is nonzero with $f(a) = 0$ for some $a \in F$. Then $\exists\ n \geq 1$ such that $f(x) = (x - a)^n f_n(x)$, where $f_n \in F[x]$ and $f_n(a) \neq 0$.*

Proof Let $a \in F$ be a root of $f \in E[x]$ with $f \neq 0$. Then by Proposition 2.21, $(x - a)^1|f(x)$ in $F[x]$, and therefore $f(x) = (x - a)^1 f_1(x)$, for some $f_1 \in F[x]$. If $f_1(a) \neq 0$, then the proof is complete. If not, then by applying Proposition 2.21, we see that $f_1(x) = (x - a)f_2(x)$ for some $f_2 \in F[x]$, hence $f(x) = (x - a)^2 f_2(x)$. If $f_2(a) \neq 0$, then the proof is complete. If not, then we may apply Proposition 2.21 again to find $f(x) = (x - a)^3 f_3(x)$ for some $f_3(x) \in F[x]$.

As we continue to iteratively apply Proposition 2.21, we deduce the following chain of equalities:

$$f(x) = (x - c)f_1(x) = (x - c)^2 f_2(x) = (x - c)^3 f_3(x) = \dots . \tag{84}$$

Examining the degrees of these polynomials and noticing that $\partial(x - a) = 1$, we see that these equalities correspond to the following equalities:

$$\partial f = 1 + \partial f_1 = 2 + \partial f_2 = 3 + \partial f_3 = ..., \tag{85}$$

which imply the inequalities

$$\partial f > \partial f_1 > \partial f_2 > \partial f_3 > ... > 0. \tag{86}$$

(It must be true that the degrees of these polynomials are greater than zero, because $x - a$ cannot divide a constant polynomial.) It is clear from Equation 86 that there cannot be more than ∂f iterative applications of Proposition 2.21. This indicates that $\exists\ n \leq \partial f$ such that Proposition 2.21 cannot be applied more than n times. Therefore, $\forall\ f_{n+1} \in F[x]$, $f_n(x) \neq (x-a)f_{n+1}$. In other words, $(x-a) \nmid f_n(x)$ in $F[x]$. By the contrapositive of Proposition 2.21, this implies that $f_n(a) \neq 0$. This completes the proof. $\square$

The corollary allows us to characterize roots of polynomials using the following definition.

Definition 2.23 *Let $E : F$ be, and $f \in E[x]$ be a nonzero polynomial such that $f(a) = 0$ for some $a \in F$. Suppose that $f(x) = (x-a)^m f_m(x)$ for some $m \in \mathbb{Z}^+$ and $f_m \in F[x]$ such that $f_m(a) \neq 0$. Then m is called the multiplicity of a in f over F.*

Corollary 2.22 guarantees that such an m exists for any root a. We say that $a \in F$ is a *multiple root* or *multiple zero of f in F* if the multiplicity of a in f is greater than 1. To emphasize, the multiplicity of any root of a polynomial must be greater than or equal to 1.

We will establish one more definition before moving on to a lemma concerning these multiple roots.

Definition 2.24 *Let F be a field. Given some nonzero polynomial*

$$p(x) = c_n x^n + c_{n-1}x^{n-1} + ... + \delta_2 x^2 + c_1 x + c_0 \in F[x], \tag{87}$$

we define the formal derivative of p as

$$Dp(x) = nc_n x^{n-1} + (n-1)c_{n-1}x^{n-2} + ... + 2\delta_2 x + c_1 + \in F[x]. \tag{88}$$

Like its namesake, the formal derivative is additive, and also obeys an analog of the product rule. These results are computationally straightforward to prove, and so we leave them to the reader.

We will now establish two more lemmas before proving that any irreducible polynomial has no multiple roots.

Lemma 2.25 *Let $F : E$ and $f \in E[x]$ be a nonzero polynomial. Then $a \in F$ is a multiple zero of f in F if and only if $f(a) = Df(a) = 0$.*

Proof ($\Rightarrow$) Let $F : E$ and suppose that $f \in E[x]$ is a nonzero polynomial, with $a \in F$ a multiple root of f in F. We will show that it is simultaneously a root of f and Df. Suppose that

$$f(x) = (x-a)^m f_m(x), \tag{89}$$

where $m \geq 2$ and $f_m \in F[x]$. (The existence of such an m and f_m is guaranteed by Corollary 2.22 and the fact that a is a multiple root of f in F.) Then it is a simple computational matter to show that

$$Df(x) = (x-a)^m Df_m(x) + m(x-a)^{m-1} f_m(x). \tag{90}$$

Since $m > 1$, this implies that $(x-a)|Df(x)$. Therefore, by Proposition 2.21, we have that $Df(a) = 0 = f(a)$.

($\Leftarrow$) Let $f(a) = Df(a) = 0$. We will show that the multiplicity of a in f is greater than 1. Proposition 2.21 implies that $\exists\ q \in F[x]$ such that

$$f(x) = (x-a)q(x) \tag{91}$$

and therefore, it is easy to show that

$$Df(x) = (x-a)Dq(x) + q(x). \tag{92}$$

This implies that

$$0 = Df(a) = (a-a)Dq(a) + q(a) = q(a). \tag{93}$$

Using Proposition 2.21, we find that $(x-a)|q(x)$. Thus, $\exists\ q_1 \in F[x]$ such that $q(x) = (x-a)q_1(x)$. Therefore,

$$f(x) = (x-a)q(x) = (x-a)(x-a)q_1(x) = (x-a)^2 q_1(x). \tag{94}$$

If q_1 has a root at $x = a$, then the multiplicity of a in f is greater than 2. If not, then the multiplicity is equal to 2. Either way, a is a multiple root of f, and so the statement is proven. $\square$

Lemma 2.26 *Let $F : E$ and $f \in E[x]$. If f and Df are relatively prime in $E[x]$, then f has no multiple zeros in F.*

Proof Let $f \in E[x]$ with f and Df being relatively prime in $E[x]$. We will show that f has no multiple roots in F. Suppose for contradiction that $a \in F$ is a multiple root of f in F. Then by Lemma 2.25, we have that $f(a) = Df(a) = 0$. By Lemma 2.12, $\exists\, h_1, h_2 \in E[x]$ such that

$$h_1 f + h_2 Df = 1. \tag{95}$$

Letting $x = a$, we see that

$$1 = h_1(a)f(a) + h_2(a)Df(a) = h_1(a)(0) + h_2(a)(0) = 0 + 0 = 0. \tag{96}$$

This contradiction leads us to conclude that our original assumption that f has a multiple root at some $a \in F$ is false, and therefore, f has no multiple roots in F. $\square$

We are now ready to prove that irreducible polynomials have no multiple zeros. However, we present this as a lemma as well, because its main function in this thesis is to allow us to prove our next result about number fields.

Lemma 2.27 *Let $F : E$, and suppose that $i \in E[x]$ is irreducible in $E[x]$. Then i has no multiple zeros in F.*

Proof Let $F : E$. We will show that any irreducible polynomial over E has no multiple zeros in F. Let $i \in E[x]$ be irreducible in $E[x]$. We know that $U(E[x]) = E \setminus \{0\}$. Thus, $\partial i > 0$, since otherwise $i = 0$ or i is a unit, which are impossible for an irreducible. Further, since i is irreducible in $E[x]$, $f|i$ in $E[x]$ implies that either $f \in U(E[x])$ or f is an associate of i in $E[x]$, because i is irre-

ducible in $E[x]$.

We consider two cases: either $\partial Di = 0$ or $\partial Di > 0$. (We know that $Di \neq 0$, because that would imply that i is a constant polynomial.) Consider the case that $\partial Di = 0$. Now, the only $f \in E[x]$ that satisfy $f|Di$ are constant polynomials, which are units in $E[x]$. It follows that i and Di are relatively prime in $E[x]$. In that case, Lemma 2.26 implies that i has no repeated zeros in F.

Consider the case that $\partial Di > 0$. Since $\partial i > \partial Di$, $i \nmid Di$ in $F[x]$. Likewise, since i divides any associate of i and "divides" is a transitive relation, no associate of i divides Di in $F[x]$. Therefore, if $f|i$ and $f|Di$ in $F[x]$, then f must not be an associate of i. As we said before, this means that $f \in U(F[x])$ because i is irreducible, hence i and Di are relatively prime. Thus, Lemma 2.26 implies that i has no repeated zeros in $F[x]$. □

Lastly, we will state the Fundamental Theorem of Algebra, for use in the next theorem.

Theorem 2.28 $\forall\ f \in \mathbb{C}[x]$ *such that* $\partial f > 0$, $\exists\ a \in \mathbb{C}$ *such that* $f(a) = 0$. *This is called the Fundamental Theorem of Algebra.*

Proof We will omit this proof. The interested reader may see pages 173-174 of reference [2] for a proof of this statement. □

Now that all of this has been established, we have enough of a basis to describe the algebraic structure of any number field using the following theorem.

Theorem 2.29 *Let N be a number field. Then $N = \mathbb{Q}(\alpha)$ for some $\alpha \in \mathbb{C}$ that is algebraic over $\mathbb{Q}$.*

Proof Let N be a number field. Since $N : \mathbb{Q}$ is finite, there is a basis for N over $\mathbb{Q}$ consisting of $\{\alpha_1, \alpha_2, ..., \alpha_n\}$, for some $\alpha_1, \alpha_2, ..., \alpha_n \in \mathbb{C}$. Clearly then, $\mathbb{Q}(\alpha_1, \alpha_2, ..., \alpha_n) = N$. We will show by mathematical induction on n that this implies that N is a simple extension of $\mathbb{Q}$.

Consider the case that $n = 1$. Then $N = \mathbb{Q}(\alpha_n)$, which is a simple extension. Suppose for the induction hypothesis that $N = \mathbb{Q}(\alpha_1, \alpha_2, ..., \alpha_k)$ implies that $N = \mathbb{Q}(\theta_1)$ for some $\theta_1 \in \mathbb{A}$. We will show in the inductive step that $N = \mathbb{Q}(\alpha_1, \alpha_2, ..., \alpha_{k+1})$ implies that $N = \mathbb{Q}(\alpha)$ for some $\alpha \in \mathbb{A}$. Suppose that $N_1 = \mathbb{Q}(\alpha_1, \alpha_2, ..., \alpha_k)$. Then by the induction hypothesis, $N_1 = \mathbb{Q}(\theta_1)$ for some $\theta_1 \in \mathbb{A}$. Therefore $N = N_1(\theta_2) = \mathbb{Q}(\theta_1, \theta_2)$ for some $\theta_2 \in \mathbb{A}$.

Let p_1 be the minimum polynomial of θ_1 over $\mathbb{Q}$ and let p_2 be the minimum polynomial of θ_2 over $\mathbb{Q}$. Corollary 2.22 guarantees that, for any $\beta_k \in \mathbb{C}$ that is a root of p_1 in $\mathbb{C}$, we may write $p_1(x) = (x - \beta_k)^n f_n(x)$ for some $f_n \in \mathbb{C}[x]$ that does not have a root at β_k. Theorem 2.28 guarantees that this f_n must have a root in $\mathbb{C}$ unless it is a constant polynomial. If it is not constant, then we may apply Corollary 2.22 to write f_n as $f_n(x) = (x - \beta_l)^m f_m(x)$ for any $\beta_l \in \mathbb{C}$ that is a root of f_n and some $f_m \in \mathbb{C}[x]$ that does not have a root at β_l. Again, Theorem 2.28 guarantees that this, too, has a root in $\mathbb{C}$, unless it is a constant polynomial. Iteratively applying these arguments creates a descending sequence of non-negative integers:

$$\partial p_1 > \partial f_n > \partial f_m > ... \geq 0. \tag{97}$$

This process must halt eventually, and at that stage, we will be able to write

$$p_1(x) = (x-\beta_i)^n f_n(x) = (x-\beta_i)^n (x-\beta_j)^m f_m(x) = ... \\ = (x-\beta_1)(x-\beta_2)...(x-\beta_r), \quad (98)$$

where each of the β_i are roots of p_1 in $\mathbb{C}$. (We know that there does not exist a nonunit constant that multiplies the right side of this equation, because that would imply that p_1 is not monic.) Likewise, we will write p_2 as

$$p_2 = (x-\gamma_1)(x-\gamma_2)...(x-\gamma_s), \quad (99)$$

where each of the γ_j are roots of p_2 in $\mathbb{C}$. Such a representation is guaranteed by arguments that are similar to those that guarantee a similar representation's existence for p_1.

Suppose, with the understanding that any other choice is similar, that $\beta_1 = \theta_1$ and $\gamma_1 = \theta_2$. We know from Proposition 2.10 and Lemma 2.27 that these β_i must be distinct and that these γ_j must be distinct, because $\mathbb{C} : \mathbb{Q}$. Now suppose that $\beta_i + y\gamma_j = \beta_1 + y\gamma_1$ for some $y \in \mathbb{C}$ and some $i, j \in \mathbb{Z}$. This implies that $y(\gamma_j - \gamma_1) = (\beta_1 - \beta_i)$. By Lemma 2.18, it is legal and defined for us to let $y = \frac{\beta_1 - \beta_i}{\gamma_j - \gamma_1}$, provided that $j \neq 1$. Thus, $\forall\ (i,j) \neq (i,1)$, there exists at most one y satisfying $\beta_i + y\gamma_j = \beta_1 + y\gamma_1$. From this fact, we will choose some $0 \neq c \in \mathbb{Q}$ such that $c \neq y$ for any of these solutions (since there are only finitely many of these equations, and each equation has a unique solution, we are guaranteed that

there exists such a c). Then

$$\beta_i + c\gamma_j \neq \beta_1 + c\gamma_1 \tag{100}$$

for any $1 \leq i \leq r$ and $1 < j \leq s$.

Define

$$\alpha = \theta_1 + c\theta_2. \tag{101}$$

We will show that $\mathbb{Q}(\alpha) = \mathbb{Q}(\theta_1, \theta_2)$. We begin by showing that $\mathbb{Q}(\theta_1, \theta_2) : \mathbb{Q}(\alpha)$. We know that $c \in \mathbb{Q}$, and so $\alpha = \theta_1 + c\theta_2 \in \mathbb{Q}(\theta_1, \theta_2)$. But by definition of $\mathbb{Q}(\alpha)$, $\mathbb{Q}(\alpha)$ is the smallest field extension of $\mathbb{Q}$ containing α; any other field extension of $\mathbb{Q}$ containing α must be a field extension of $\mathbb{Q}(\alpha)$. Hence, $\mathbb{Q}(\theta_1, \theta_2) : \mathbb{Q}(\alpha)$.

We will show that $\mathbb{Q}(\alpha) : \mathbb{Q}(\theta_1, \theta_2)$. We know that $\theta_1 = \alpha - c\theta_2$. Therefore, if $\theta_2 \in \mathbb{Q}(\alpha)$, then $\theta_1 \in \mathbb{Q}(\alpha)$. If that is the case, then $\mathbb{Q}(\alpha) : \mathbb{Q}(\theta_1, \theta_2)$, because $\mathbb{Q}(\theta_1, \theta_2)$ is defined to be the smallest field that extends $\mathbb{Q}$ and contains the elements θ_1 and θ_2. So it suffices to show that $\theta_2 \in \mathbb{Q}(\alpha)$.

Define $f(x) = p_1(\alpha - cx)$. Then

$$f(\theta_2) = p_1(\alpha - c\theta_2) = p_1(\theta_1) = 0, \tag{102}$$

by definition of minimum polynomial. We claim that $f(x)$ has only one root in common with $p_2(x)$. Suppose that ξ is this root, so that

$$f(\xi) = 0 = p_2(\xi). \tag{103}$$

Then

$$f(\xi) = p_1(\alpha - c\xi) = (\alpha - c\xi - \beta_1)(\alpha - c\xi - \beta_2)...(\alpha - c\xi - \beta_r) = 0. \quad (104)$$

In that case, $\alpha - c\xi$ equals one of the β_i. Likewise,

$$p_2(\xi) = (\xi - \gamma_1)(\xi - \gamma_2)...(\xi - \gamma_s) = 0, \quad (105)$$

and so, ξ also equals one of the γ_j. Thus, for some i and j, $\alpha - c\gamma_j = \alpha - c\xi = \beta_i$, hence

$$\beta_i + c\gamma_j = \alpha = \theta_1 + c\theta_2 = \beta_1 + c\gamma_1. \quad (106)$$

Yet based on our definition of c, $j \neq 1$ implies that $\beta_1 + c\gamma_1 \neq \beta_i + c\gamma_j$. By the contrapositive, $\beta_1 + c\gamma_1 = \beta_i + c\gamma_j$ implies that $j = 1$. Thus, $\xi = \gamma_1 = \theta_2$, and this must be the unique root of both $p_2(x)$ and $f(x)$.

We know that θ_2 is algebraic over $\mathbb{Q}(\alpha)$ because it is algebraic over one of its subfields, namely $\mathbb{Q}$. Thus, Theorem 2.5 guarantees that θ_2 has a minimum polynomial over $\mathbb{Q}(\alpha)$. Now let $p_{2\alpha}(x)$ be the minimum polynomial of θ_2 over $\mathbb{Q}(\alpha)$. Then, by Lemma 2.20, we see that $p_{2\alpha}|p_2$ and $p_{2\alpha}|f$ in $\mathbb{Q}(\alpha)[x]$. But p_2 and f have only one root in $\mathbb{C}$ in common, namely θ_2. We deduce from this and Proposition 2.21 that $x - \theta_2$ is a common factor of p_2 and f in $\mathbb{C}[x]$. Further, since there are no other common roots of p_2 and f in $\mathbb{C}[x]$, there are no other factors in $\mathbb{C}[x]$ of the form $x - a$ (where a is a root of p_2 and f) that are common to p_2 and f. By Theorem 2.28, and Proposition 2.21, this implies that there are no other polynomials in $\mathbb{C}[x]$ that divide p_2 and f; $x - \theta_2$ is the only common factor of p_2 and f. Thus,

$p_{2\alpha}(x) = x - \theta_2 \in \mathbb{Q}(\alpha)[x]$, and hence, $\theta_2 \in \mathbb{Q}(\alpha)$. Therefore, $\mathbb{Q}(\alpha) : \mathbb{Q}(\theta_1, \theta_2)$, and so $N = \mathbb{Q}(\theta_1, \theta_2) = \mathbb{Q}(\alpha)$. Thus, N is a simple extension of the field of rationals by α. $\square$

After reading the first chapter, the reader may question why this chapter has seemingly no relation to integers. This thesis is, after all, a treatment of algebraic methods of number theory; why then, one may ask, have we taken so much time discussing algebraic structures that have seemingly no relation to $\mathbb{Z}$? The reason for this diversion has been to build up the background necessary to deal with the properties of integers in a more general context.

In a number field, it is often desirable to define integers by equations that they satisfy. This fits in with the "backwards thinking" that we have been doing for this whole chapter: defining, not characterizing, elements by equations that they satisfy. In that case, by introducing the following definition, we can use all of the properties that we have established above to work with integers in a general setting.

Definition 2.30 *Let $\alpha \in \mathbb{A}$. Then α is an algebraic integer provided that for some monic $p \in \mathbb{Z}[x]$, $p(\alpha) = 0$.*

We will begin our characterization of these integers by noting that they form a subring of $\mathbb{A}$. We introduce the notation of ζ to mean the set of algebraic integers. It turns out that ζ is a ring.

Proposition 2.31 *Under the additive and multiplicative operations of $\mathbb{C}$, ζ forms a subring of $\mathbb{A}$.*

Proof We will omit this proof. The interested reader may see page 47 of reference [8] for a proof of this statement. □

Like $\mathbb{A}$, ζ is not terribly interesting. Rather, its subrings will be of greater consequence. Just as for some number field N, $\mathbb{A} \cap N$ was more important than $\mathbb{A}$ itself, so is $\zeta \cap N$ more important than ζ. We will refer to $\zeta \cap N$ by the notation $Z(N)$. It is a simple matter to show that $Z(N)$ is always a ring, because the intersection of two rings is always a ring. These rings of integers will become important in the next chapter, in which we will discuss quadratic fields. We end this chapter with a simple proposition that may seem inevitable to the reader: the minimum polynomial of an algebraic integer has coefficients in $\mathbb{Z}$.

Proposition 2.32 *Let $\alpha \in N$. Then $\alpha \in Z(N)$ if and only if $p \in \mathbb{Z}[x]$, where p is the minimum polynomial of α over $\mathbb{Q}$.*

Proof ($\Rightarrow$) Let $\alpha \in Z(N)$. Then for some monic $f \in \mathbb{Z}[x]$, $f(\alpha) = 0$. Also, since $\alpha \in \mathbb{A}$, there exists the minimum polynomial $p \in \mathbb{Q}[x]$ so that $p(\alpha) = 0$, by Theorem 2.5. Then, by Lemma 2.20, we have that

$$f = pq \tag{107}$$

for some $q \in \mathbb{Q}[x]$. Let $p = \frac{1}{z_p} p_0$ and $q = \frac{1}{z_q} q_0$, where $p_0, q_0 \in \mathbb{Z}[x]$. (We know that such p_0 and q_0 must exist, because each coefficient of a rational polynomial is a fraction of two integers. By multiplying by the product of the denominators of these fractions, for example, one can produce a new polynomial with integer coefficients.) Let $p_0 = d_p p_1$ and $q_0 = d_q q_1$, where $d_p \in \mathbb{Z}$ is a greatest common

divisor of the coefficients of p_0 and $d_q \in \mathbb{Z}$ is a greatest common divisor of the coefficients of q_0. Thus, $p = \frac{d_p}{z_p} p_1$ and $q = \frac{d_q}{z_q} q_1$, where the coefficients of p_1 have no common irreducible divisors in $\mathbb{Z}$ and the coefficients of q_1 have no common irreducible divisors in $\mathbb{Z}$. Then $f = \frac{d}{z} p_1 q_1$, where $d = d_p d_q \in \mathbb{Z}$ and $z = z_p z_q \in \mathbb{Z}$. Clearing the denominator, $zf = dp_1 q_1$. We cancel any irreducible factors that are common to both d and z to get

$$y_1 f = y_2 p_1 q_1, \tag{108}$$

where $y_1, y_2 \in \mathbb{Z}$ have no common irreducible factors.

We claim that $y_1 = \pm 1$. Assume for contradiction that $y_1 \neq \pm 1$. Then either $y_1 = 0$ or y_1 can be factored into irreducibles, as a result of Theorem 1.31. Either way, y_1 is divisible by at least one irreducible if $y_1 \neq \pm 1$. In that case, $i | y_1$, where $i \in \mathbb{Z}$ is irreducible. Then by Equation 108, $i | y_2 p_1 q_1$. Let

$$p_1 q_1(x) = \sum_{j=0}^{nm} c_j x^j, \tag{109}$$

where $n = \partial p$ and $m = \partial q$. Then for each j, $i | y_2 c_j$. Since $i | y_1$ and y_1 and y_2 have no common irreducible divisors, it is clear that $i \nmid y_2$. Using Theorem 1.31 and Proposition 1.27, we see that i is prime, and therefore $i | c_j$ for each j, since $i | y_2 c_j$ and $i \nmid y_2$.

Let

$$p_1(x) = a_n x^n + a_{n-1} x^{n-1} + ... + a_1 x + a_0 \tag{110}$$

and

$$q_1(x) = b_m x^m + b_{m-1}x^{m-1} + ... + b_1 x + b_0. \tag{111}$$

By definition of p_1, we know that no irreducible in $\mathbb{Z}$ divides all of the a_k simultaneously, and by definition of q_1, we know that no irreducible in $\mathbb{Z}$ divides all of the b_l simultaneously. Let k and l be such that $i \nmid a_k$ but $\forall\ u < k$, $i|a_u$ and $i \nmid b_l$ but $\forall\ v < l$, $i|b_v$. But by definition of multiplication in the ring $\mathbb{Z}[x]$, the coefficient of the term with x^{k+l} is

$$c_{k+l} = a_0 b_{k+l} + a_1 b_{k+l-1} + ... + a_k b_l + ... + a_{k+l-1} b_1 + a_{k+l} b_0. \tag{112}$$

We have defined a_k and b_l so that i divides all of the terms containing a_u or b_v such that $u < k$ or $b < l$. Therefore, i divides every term of c_{k+l} except for $a_k b_l$. This implies that $i \nmid c_{k+l}$, despite the fact that we have already established that $i|c_{k+l}$. This contradiction leads us to the conclusion that our assumption that y_1 is divisible by an irreducible is false. We see that y_1 must therefore be a unit in $\mathbb{Z}$. Hence, $y_1 = \pm 1$.

Returning to Equation 108, we find that

$$f = \pm y_2 p_1 q_1. \tag{113}$$

We know that f is monic. We also know that $p_1, q_1 \in \mathbb{Z}[x]$, and so the leading coefficients of p_1 and q_1 must be integers. Further, y_2 is an integer. Thus, Equation 113 indicates that the leading coefficient of p_1 must be a unit in $\mathbb{Z}$. Hence, either p_1 is monic, or else $-p_1$ is monic. Then the fact that $p = \frac{d_p}{z_p} p_1$ indicates that $\frac{d_p}{z_p} = \pm 1$,

since p is monic. Therefore, either $p = p_1 \in \mathbb{Z}[x]$, or else $p = -p_1 \in \mathbb{Z}[x]$. Either way, this direction of the proof is complete.

($\Leftarrow$) Let $\alpha \in \mathbb{A}$, with $p \in \mathbb{Z}[x]$ being the minimum polynomial of α over $\mathbb{Q}$. Then it is clear that $\alpha \in \zeta$, since α satisfies p, which is a monic polynomial equation with coefficients in $\mathbb{Z}$. Then for any number field N such that $\alpha \in N$, $\alpha \in N \cap \zeta = Z(N)$. $\square$

3 Quadratic Fields

All number fields are finite extensions of the field of rationals, but in order to study number fields in more depth, it is often necessary to choose a degree for the field extension. In this chapter, we will examine the structure of number fields when such a choice is made. This particular choice will leads us to many results that can be useful in applications to number theory. We begin with a definition that will set the choice that we have just mentioned.

Definition 3.1 *Let N be a number field. Then N is called a quadratic field provided that $[N : \mathbb{Q}] = 2$.*

Naturally, all of the results about number fields apply to quadratic fields. However, in this context, the language bears a slight variation: the term "algebraic" is often replaced with "quadratic." For example, "quadratic integers" shall be an alternative way of saying "algebraic integers of a quadratic field." In addition, when there is a possibility for confusion between the members of $\mathbb{Z}$ and the members of ζ, we will refer to the former as "rational integers".

There is an alternative definition of this type of number field. We will state this defintion in the form of a proposition and then prove its equivalence to the above. Before this, though, we will define a term for the sake of future linguistic convenience.

Definition 3.2 *Let $d \in \mathbb{Z}$. Then d is called square-free provided that $d \neq 0, 1$ and $\forall\, y \in \mathbb{Z}$ such that $y \neq \pm 1$, $y^2 \nmid d$.*

With this said, we move on to an alternative definition of quadratic field.

Proposition 3.3 *A number field N is a quadratic field if and only if $N = \mathbb{Q}(\sqrt{d})$ for some square-free $d \in \mathbb{Z}$.*

Proof ($\Rightarrow$) Let $[N : \mathbb{Q}] = 2$. Then $N = \mathbb{Q}(\alpha)$ for some $\alpha \in \mathbb{A} \setminus \mathbb{Q}$, as given by Theorem 2.29 and $\alpha \in \mathbb{Q}$ would imply that $[\mathbb{Q}(\alpha) : \mathbb{Q}] = [\mathbb{Q} : \mathbb{Q}] = 1 \neq 2$. Now, $[\mathbb{Q}(\alpha) : \mathbb{Q}] = 2$, so the minimum polynomial p of α over $\mathbb{Q}$ has degree 2, by Corollary 2.15. Therefore,

$$p(\alpha) = \alpha^2 + a\alpha + b = 0 \tag{114}$$

for some $a, b \in \mathbb{Q}$. We complete the square as follows:

$$\alpha^2 + a\alpha + \frac{a}{4} - \frac{a}{4} + b = \left(\alpha + \frac{a}{2}\right)^2 - \frac{a}{4} + b. \tag{115}$$

We redefine $\frac{a}{4} - b = c$, which, as we can see, is rational. Then

$$\left(\alpha + \frac{a}{2}\right)^2 - c = 0, \tag{116}$$

so $c = (\alpha + \frac{a}{2})^2$. Thus, $\sqrt{c} = \pm(\alpha + \frac{a}{2}) \notin \mathbb{Q}$. (For the remainder of the proof, we will assume that $\sqrt{c} = \alpha + \frac{a}{2} \notin \mathbb{Q}$, with the understanding that the other choice is similar.)

We claim that $\mathbb{Q}(\sqrt{c}) = N$. Let $w + x\sqrt{c} \in \mathbb{Q}(\sqrt{c})$. Then

$$w + x\sqrt{c} = w + x\left(\alpha + \frac{a}{2}\right) = w + x\alpha + x\frac{a}{2} = \left(w + x\frac{a}{2}\right) + x\alpha \in \mathbb{Q}(\alpha). \tag{117}$$

From this we see that $\mathbb{Q}(\sqrt{c}) \subseteq \mathbb{Q}(\alpha) = N$. Let $y + z\alpha \in N = \mathbb{Q}(\alpha)$. Then

$$y + z\alpha = y + z\left(\sqrt{c} - \frac{a}{2}\right) = y + z\sqrt{c} - z\frac{a}{2} = \left(y - \frac{a}{2}z\right) + z\sqrt{c} \in \mathbb{Q}(\sqrt{c}). \quad (118)$$

Thus, $N = \mathbb{Q}(\sqrt{\alpha}) \subseteq \mathbb{Q}(\sqrt{c})$. Therefore, $N = \mathbb{Q}(\sqrt{c})$.

We will now show that $N = \mathbb{Q}(\sqrt{d})$, for some square-free $d \in \mathbb{Z}$. If it is true that $c = -1$, then the proof is complete, since then $N = \mathbb{Q}(\sqrt{-1})$, and -1 is square-free. Suppose that $c \neq -1$. We know that $c = \frac{r}{s}$ for some $r, s \in \mathbb{Z}$ with $s \neq 0$, by definition of $\mathbb{Q}$. We also know that $c \neq 0$ and $c \neq 1$, since $\sqrt{c} \notin \mathbb{Q}$. Therefore, we know that rs has a unique factorization into irreducibles, by Theorem 1.31, since if $\frac{r}{s} \neq 1$ and $\frac{r}{s} \neq 0$, then certainly $rs \neq 1$ and $rs \neq 0$ (since $\mathbb{Z}$ is an integral domain, this also implies that $r \neq 0$ and $s \neq 0$). Let

$$rs = {p_1}^{2n_1}{p_2}^{2n_2}...{p_u}^{2n_u}{q_1}^{2m_1+1}{q_2}^{2m_2+1}...{q_v}^{2m_v+1} \quad (119)$$

be a factorization of rs into irreducibles, with the p_i denoting irreducibles that are raised to even powers in the factorization of rs, and the q_j denoting those that are raised to odd powers. We define $e = {p_1}^{n_1}{p_2}^{n_2}...{p_u}^{n_u}{q_1}^{m_1}{q_2}^{m_2}...{q_v}^{m_v}$, and also $d = q_1q_2...q_v$, so that $rs = e^2d$. We note that d is square-free and that $d, e \in \mathbb{Q}$, with neither being equal to 0. We claim that $N = \mathbb{Q}(\sqrt{c}) = \mathbb{Q}(\sqrt{d})$.

($\subseteq$) We will show that $N = \mathbb{Q}(\sqrt{c}) \subseteq \mathbb{Q}(\sqrt{d})$. Consider that

$$\sqrt{c} = \sqrt{\frac{r}{s}} = \sqrt{\frac{rs}{s^2}} = \frac{1}{s}\sqrt{rs} = \frac{1}{s}\sqrt{e^2d} = \frac{e}{s}\sqrt{d} \in \mathbb{Q}(\sqrt{d}), \quad (120)$$

since $e, s \in \mathbb{Q}$, $s \neq 0$, and $\mathbb{Q}$ is a field. We also know that $\mathbb{Q}(\sqrt{d}) : \mathbb{Q}$ by definition. Further, the definition of $\mathbb{Q}(\sqrt{c})$ indicates that any field that extends $\mathbb{Q}$ and contains the element $\sqrt{c}$ must extend $\mathbb{Q}(\sqrt{c})$. Thus, $\mathbb{Q}(\sqrt{d}) : \mathbb{Q}(\sqrt{c})$, hence $N = \mathbb{Q}(\sqrt{c}) \subseteq \mathbb{Q}(\sqrt{d})$.

($\supseteq$) We will show that $\mathbb{Q}(\sqrt{d}) \subseteq \mathbb{Q}(\sqrt{c}) = N$. Equation 120 indicates that $\sqrt{d} = \frac{s}{e}\sqrt{c}$, since $e \neq 0$. Therefore, by the same arguments, since $\mathbb{Q}(\sqrt{c}) : \mathbb{Q}$ and $\sqrt{d} \in \mathbb{Q}(\sqrt{c})$, it follows that $\mathbb{Q}(\sqrt{c}) : \mathbb{Q}(\sqrt{d})$. Thus, $\mathbb{Q}(\sqrt{d}) \subseteq \mathbb{Q}(\sqrt{c}) = N$, and so $N = \mathbb{Q}(\sqrt{c}) = \mathbb{Q}(\sqrt{d})$.

($\Leftarrow$) Let $N = \mathbb{Q}(\sqrt{d})$, with d being a square-free integer. We will show that N is a quadratic field. Consider that each element of $\mathbb{Q}(\sqrt{d})$ can be written uniquely as $a_1 + b_1\sqrt{d}$ for some $a_1, b_1 \in \mathbb{Q}$. Then any element of this field as a vector space over $\mathbb{Q}$ can be written as a linear combination of 1 and $\sqrt{d}$.

We claim that the set $\{1, \sqrt{d}\}$ is linearly independent. Let $r_1 + r_2\sqrt{d} = 0$. Then $r_1 = -r_2\sqrt{d}$. Therefore, ${r_1}^2 = {r_2}^2 d$. Suppose for contradiction that $r_2 \neq 0$. Then $\left(\frac{r_1}{r_2}\right)^2 = d$. If $\frac{r_1}{r_2} \in \mathbb{Z}$, then d is not square-free. If $\frac{r_1}{r_2} \notin \mathbb{Z}$, then d is not an integer. Either way, we reach a contradiction that leads us to the conclusion that $r_2 = 0$. Then $r_1 = 0$ as well. Thus, $\{1, \sqrt{d}\}$ is linearly independent.

This implies that $\{1, \sqrt{d}\}$ is a basis of $\mathbb{Q}(\sqrt{d})$, and so N has dimension 2 by definition, implying that $[N : \mathbb{Q}] = 2$, hence N is a quadratic field. $\square$

It is possible for a quadratic field to be equal to $\mathbb{Q}(\sqrt{d})$ for some d that is not square-free, as long as its square root is not rational. However, because it is easier to consider only certain integers rather than all of the rational numbers, we will, for the remainder of this thesis, assume that the d that generates a quadratic field is

a squarefree integer. Futher, we will give no references to the above proposition's number as a justification for any results, but rather we will think of Proposition 3.3 and Definition 3.1 as interchangeable definitions of quadratic field. Indeed, many books define quadratic fields as we did in Proposition 3.3, without bothering to mention the language of Definition 3.1.

Further, when we have a quadratic field $N = \mathbb{Q}(\sqrt{d})$, then we will say that N is a *real quadratic field* provided that $d > 0$. On the other hand, when $d < 0$, we will call N a *complex quadratic field*.

Before moving on, we introduce two definitions and three lemmas that follow directly from them.

Definition 3.4 *Let $\alpha = a+b\sqrt{d} \in \mathbb{Q}(\sqrt{d})$, where $\mathbb{Q}(\sqrt{d})$ is a quadratic field. Then the conjugate of α, denoted $\overline{\alpha}$, is defined as $\overline{\alpha} = a - b\sqrt{d}$.*

Definition 3.5 *Let $\alpha = a+b\sqrt{d} \in \mathbb{Q}(\sqrt{d})$, where $\mathbb{Q}(\sqrt{d})$ is a quadratic field. Then the norm of α, denoted $N(\alpha)$, is defined as $N(\alpha) = \alpha\overline{\alpha} = a^2 - b^2 d$.*

The reader must be sure to not confuse the norm with the field norm of a Euclidean Domain. For example, it is not uncommon for a quadratic field to fail to be a Euclidean Domain, in which case the norm is certainly not the same as a field norm.

We should introduce a few lemmas before continuing.

Lemma 3.6 *If $\alpha \in \mathbb{Q}(\sqrt{d})$, where $\mathbb{Q}(\sqrt{d})$ is a quadratic field, then $\alpha + \overline{\alpha} \in \mathbb{Q}$.*

Proof Let $\alpha = a + b\sqrt{d} \in \mathbb{Q}(\sqrt{d})$, where $\mathbb{Q}(\sqrt{d})$ is a quadratic field. Then it is clear that $\alpha + \overline{\alpha} = a + b\sqrt{d} + a - b\sqrt{d} = 2a \in \mathbb{Q}$. $\square$

Lemma 3.7 *If $\alpha \in \mathbb{Q}(\sqrt{d})$, where $\mathbb{Q}(\sqrt{d})$ is a quadratic field, then $N(\alpha) \in \mathbb{Q}$.*

Proof Let $\alpha = a + b\sqrt{d} \in \mathbb{Q}(\sqrt{d})$, where $\mathbb{Q}(\sqrt{d})$ is a quadratic field. Then it is clear that $N(\alpha) = \alpha\overline{\alpha} = (a + b\sqrt{d})(a - b\sqrt{d}) = a^2 - b^2 d \in \mathbb{Q}$. $\square$

Lemma 3.8 *If $\mathbb{Q}(\sqrt{d})$ is a quadratic field and $\alpha, \beta \in \mathbb{Q}(\sqrt{d})$, then the norm is multiplicative,* id est, $N(\alpha\beta) = N(\alpha)N(\beta)$.

Proof Let $\alpha, \beta \in \mathbb{Q}(\sqrt{d})$, where $\mathbb{Q}(\sqrt{d})$ is a quadratic field. We will show that the norm is multiplicative. We know that $\alpha = a_1 + a_2\sqrt{d}$ and $\beta = b_1 + b_2\sqrt{d}$ for some $a_1, a_2, b_1, b_2 \in \mathbb{Q}$. In that case,

$$\begin{aligned} N(\alpha\beta) = N\left((a_1 + a_2\sqrt{d})(b_1 + b_2\sqrt{d})\right) \\ = N((a_1b_1 + a_2b_2d) + (a_1b_2 + a_2b_1)\sqrt{d}) \\ = (a_1b_1 + a_2b_2d)^2 - (a_1b_2 + a_2b_1)^2 d \\ = {a_1}^2{b_1}^2 + 2a_1b_1a_2b_2d + {a_2}^2{b_2}^2d^2 - ({a_1}^2{b_2}^2 + 2a_1b_1a_2b_2 + {a_2}^2{b_1}^2)d \\ = {a_1}^2{b_1}^2 - {a_1}^2{b_2}^2d - {a_2}^2{b_1}^2d + {a_2}^2{b_2}^2d^2 \\ = ({a_1}^2 - {a_2}^2d)({b_1}^2 - {b_2}^2d) = N(\alpha)N(\beta). \quad (121) \end{aligned}$$

Hence, the norm is multiplicative. $\square$

Before continuing, we should mention a slight notational subtlety. So far we have defined $E[\alpha]$ only in a context where E is a field. On the other hand, if R is a ring, we will define $R[\alpha] = a + b\alpha$, where $a, b \in R$. With that said, we continue to another lemma.

Lemma 3.9 *If $\mathbb{Q}(\sqrt{d})$ is a quadratic field, then $\mathbb{Z}[\sqrt{d}] \subsetneq \mathbb{Z}[\frac{1+\sqrt{d}}{2}]$.*

Proof ($\subseteq$) Let $\alpha \in \mathbb{Z}[\sqrt{d}]$. We will show that $\alpha \in \mathbb{Z}[\frac{1+\sqrt{d}}{2}]$. We know that it is true that $\alpha = a + b\sqrt{d}$ for some $a, b \in \mathbb{Z}$. Therefore,

$$\alpha = \frac{2a + 2b\sqrt{d}}{2} = \frac{2a - 2b + 2b + 2b\sqrt{d}}{2}$$
$$= (a - b) + 2b\left(\frac{1+\sqrt{d}}{2}\right) \in \mathbb{Z}\left[\frac{1+\sqrt{d}}{2}\right]. \quad (122)$$

Thus, $\mathbb{Z}[\sqrt{d}] \subseteq \mathbb{Z}[\frac{1+\sqrt{d}}{2}]$.

($\neq$) Consider the element $\alpha = \frac{1+\sqrt{d}}{2} \in \mathbb{Z}[\frac{1+\sqrt{d}}{2}]$. Then $\alpha = \frac{1}{2} + \frac{1}{2}\sqrt{d} \notin \mathbb{Z}[\sqrt{d}]$. Thus, $\mathbb{Z}[\frac{1+\sqrt{d}}{2}] \not\subseteq \mathbb{Z}[\sqrt{d}]$. $\square$

These lemmas may seem like small details at first, but they are vastly important to solving later problems, as we will soon see. Now we will begin to characterize rings of quadratic integers. These integers are split into two categories: those that take the form $a + b\sqrt{d}$ for two integers a and b, and those that take the form $\frac{a+b\sqrt{d}}{2}$, where a and b are either both even integers or both odd integers. The following theorem will establish this distinction.

Theorem 3.10 *Let $\mathbb{Q}(\sqrt{d})$ be a quadratic field. Then $Z(\mathbb{Q}(\sqrt{d})) = \mathbb{Z}[\omega]$, where $\omega = \sqrt{d}$ if and only if $d \not\equiv 1 (\mathrm{mod}\ 4)$, and $\omega = \frac{1+\sqrt{d}}{2}$ if and only if $d \equiv 1 (\mathrm{mod}\ 4)$.*

Proof ($\subseteq$) Let $\mathbb{Q}(\sqrt{d})$ be a quadratic field. We will show that the ring of integers of $\mathbb{Q}(\sqrt{d})$ is a subset of $\mathbb{Z}[\omega]$. Let $\alpha \in Z(\mathbb{Q}(\sqrt{d}))$. Then we know that $\alpha = r + s\sqrt{d}$ for some $r, s \in \mathbb{Q}$, as a virtue of its presence in $\mathbb{Q}(\sqrt{d})$. Let $r = \frac{k}{u}$ and $s = \frac{l}{v}$ for $k, l, u, v \in \mathbb{Z}$, where $u \neq 0 \neq v$ (we know that such integers must exist, by definition of $\mathbb{Q}$). Then $\alpha = \frac{k}{u} + \frac{l}{v}\sqrt{d} = \frac{kv+lu\sqrt{d}}{uv}$. We reduce this fraction into

lowest terms by using the existence of unique factorization into irreducibles in $\mathbb{Z}$: let $p_1, p_2, ..., p_n$ be all of the common irreducible factors of kv and lu, and suppose that $p_1, p_2, ..., p_m$ are factors of uv, where $m \leq n$. Then one can cancel these factors, knowing that none of the primes are 0, and so we are left with a fraction, call it

$$\alpha = \frac{a + b\sqrt{d}}{c}, \tag{123}$$

where no rational prime simultaneously divides all of a, b and c. By Proposition 2.32, we know that $\alpha \in Z(\mathbb{Q}(\sqrt{d}))$ if and only if the coefficients of its minimum polynomial over $\mathbb{Q}$ are rational integers. We consider two cases: either α is rational or it is irrational.

Consider the case that α is rational. Then since it is an algebraic integer by definition, $\exists\ x + w \in \mathbb{Z}[x]$ such that $\alpha + w = 0$. But this is true if and only if $\alpha = -w \in \mathbb{Z} \subseteq \mathbb{Z}[\omega]$, so $\alpha \in \mathbb{Z}[\omega]$, regardless of the value of ω.

Consider the case that α is irrational. Assume for contradiction that the minimum polynomial of α over $\mathbb{Q}$ is of degree 1. Then $\exists\ x + w \in \mathbb{Q}[x]$ such that $\alpha + w = 0$. But then $\alpha = -w \in \mathbb{Q}$, despite the fact that α is irrational. This contradiction leads us to conclude that $\partial p \geq 2$, where p is the minimum polynomial of α over $\mathbb{Q}$. We accept the possibility that $\partial p = 2$. In that case,

$$p(x) = (x - \alpha)(x + y) = x^2 + (y - \alpha)x - y\alpha, \tag{124}$$

for some y such that $y\alpha, y - \alpha \in \mathbb{Q}$. (The minimum polynomial must follow this representation due to Proposition 2.21 and the fact that α is a root of p, which is

monic.) Let $y = -\overline{\alpha}$, so that

$$p(x) = (x - \alpha)(x - \overline{\alpha}) = x^2 - (\overline{\alpha} + \alpha)x + N(\alpha). \tag{125}$$

By Lemma 3.7, this satisfies $y\alpha \in \mathbb{Q}$ and by Lemma 3.6, $-(\alpha - y) \in \mathbb{Q}$. This must be the minimum polynomial, then, since it is a monic polynomial with rational coefficients that is of degree 2 that satisfies $p(\alpha) = 0$, and by Theorem 2.5, any polynomial satisfying these properties must be the unique minimum polynomial of α over $\mathbb{Q}$, because as we have shown, the degree of the minimum polynomial cannot be 1.

As previously stated, $p \in \mathbb{Z}[x]$ by definition of α as an algebraic integer and Proposition 2.32. We deduce from this and the definition of α that

$$-(\overline{\alpha} + \alpha) = -2\frac{a}{c} \in \mathbb{Z} \tag{126}$$

and

$$N(\alpha) = \frac{a^2 - b^2 d}{c^2} = \frac{a^2}{c^2} - \frac{b^2 d}{c^2} \in \mathbb{Z}. \tag{127}$$

We claim that a and c have no prime number $i \in \mathbb{Z}$ as a common factor. Suppose for contradiction that this is not the case. Let $\frac{a^2}{c^2} - \frac{b^2 d}{c^2} = z$ for some $z \in \mathbb{Z}$. Then $a^2 - b^2 d = zc^2$, and so

$$b^2 d = a^2 - zc^2. \tag{128}$$

Since $i^2 | a^2$ and $i^2 | zc^2$ as a consequence of i dividing c, we have that $i^2 | (a^2 - zc^2)$, by the distributive property of multiplication in rings. Hence, $i^2 | b^2 d$. Therefore, $i | b^2$ or $i | d$. If $i | b^2$, then $i | b$, so we reach a contradiction because we have reduced a,

b and c so that they have no common prime factors, and this would imply that $i|a$, $i|b$, and $i|c$. Thus, $i|d$. Let $d = ei$, so that $i^2|b^2d$ implies $i|b^2e$. Then $i|b^2$ or $i|e$. But $i \nmid e$, since that would imply $i^2|d$, while we have assumed d to be squarefree. Thus, $i|b^2$, which, as mentioned, provides a contradiction.

This proves that a and c must have no common prime factors. But $c|2a$, so $2a = cz$ for some $z \in \mathbb{Z}$, which implies that $2|c$ or $2|z$. If $2|c$, then $c = 2t$, and so $2a = 2tz$, which implies that $a = tz$, so t would be a common factor of a and c. This means that $t = 1$ (since otherwise a prime would be a common factor of a and c because a prime would divide any number greater than 1). Thus, $c = 2$ or $2|z$, in which case $a = c\frac{z}{2}$, and so $c|a$, requiring likewise that $c = 1$. Thus, either $c = 1$ or $c = 2$.

Now consider the case that $d \not\equiv 1 (\mathrm{mod}\ 4)$. We will show that $c = 1$. Assume for contradiction that $c = 2$. Then a must be odd in order for a and c to not have 2 as a common prime divisor. If a is odd and c is even, then in order for $\frac{a^2-b^2d}{c^2} \in \mathbb{Z}$, we must have that b^2d is odd, and therefore that b is odd. Now, since $c = 2$, it follows that $\frac{a^2-b^2d}{c^2} = \frac{a^2-b^2d}{4} \in \mathbb{Z}$, so $a^2 - b^2d \equiv 0 (\mathrm{mod}\ 4)$. This implies that $a^2 \equiv b^2d (\mathrm{mod}\ 4)$. But $d \not\equiv 1 (\mathrm{mod}\ 4)$, so either $a^2 \equiv b^2 \equiv 0 (\mathrm{mod}\ 4)$, or else it must be true that $a^2 \not\equiv b^2 (\mathrm{mod}\ 4)$. However, we have already shown that a and b are odd, and therefore that $a^2 \equiv b^2 \equiv 1 (\mathrm{mod}\ 4)$. This contradiction leads us to conclude that $c = 1$, and so $\alpha = \frac{a+b\sqrt{d}}{c} = a + b\sqrt{d} \in \mathbb{Z}[\sqrt{d}] = \mathbb{Z}[\omega]$.

Now consider the case that $d \equiv 1 (\mathrm{mod}\ 4)$. We know that $c = 1$ or $c = 2$. Suppose that $c = 2$. By the same reasoning as above, a and b are both odd. Then

$$\alpha = \frac{a + b\sqrt{d}}{2} = \frac{a - b + b + b\sqrt{d}}{2} = \frac{a-b}{2} + b\left(\frac{1+\sqrt{d}}{2}\right). \tag{129}$$

Since a and b are both odd, we see that $\frac{a-b}{2} \in \mathbb{Z}$. Thus, $\alpha \in \mathbb{Z}[\frac{1+\sqrt{d}}{2}] = \mathbb{Z}[\omega]$. Suppose $c = 1$. Then $\alpha = a + b\sqrt{d} \in \mathbb{Z}[\sqrt{d}]$. By Lemma 3.9, we get that $\alpha \in \mathbb{Z}[\frac{1+\sqrt{d}}{2}] = \mathbb{Z}[\omega]$. Thus, $Z(\mathbb{Q}(\sqrt{d})) \subseteq \mathbb{Z}[\omega]$.

($\supseteq$) Consider the case that $d \equiv 1 (\mathrm{mod}\, 4)$. Let $\alpha \in \mathbb{Z}[\omega]$. By definition of $\mathbb{Z}[\omega]$, $\alpha = a + b(\frac{1+\sqrt{d}}{2}) = \frac{2a+b}{2} + \frac{b\sqrt{d}}{2}$ for some $a, b \in \mathbb{Z}$. Thus, $\alpha \in \mathbb{Q}(\sqrt{d})$. We know from Lemma 2.17 that, since $\mathbb{Q}(\sqrt{d})$ is a number field, α is algebraic over $\mathbb{Q}$, and therefore has a minimum polynomial over $\mathbb{Q}$. By the same arguments as given in the previous direction, we know that the minimum polynomial of α over $\mathbb{Q}$ is $p(x) = x^2 - (\overline{\alpha} + \alpha)x + N(\alpha)$, because it is the unique monic polynomial of minimal degree with rational coefficients that has a root at $x = \alpha$. Then

$$\begin{aligned} p(x) = x^2 - (2a+b)\,x + \left(\left(\frac{2a+b}{2}\right)^2 - \left(\frac{b}{2}\right)^2 d\right) \\ = x^2 - (2a+b)x + \left(\frac{4a^2+4ab+b^2}{4} - \frac{b^2 d}{4}\right) \\ = x^2 - (2a+b)x + \left(a^2 + ab + \frac{b^2(1-d)}{4}\right). \quad (130) \end{aligned}$$

Because we have chosen $d \equiv 1 (\mathrm{mod}\, 4)$, we get that $\frac{1-d}{4} \in \mathbb{Z}$, so the polynomial $p(x)$ has rational integer coefficients and therefore, by Proposition 2.32, it follows that $\alpha \in Z(\mathbb{Q}(\sqrt{d}))$.

Consider the case that $d \not\equiv 1(\mathrm{\ mod\ } 4)$. Let $\alpha \in \mathbb{Z}[\omega]$. Then $\alpha = a + b\sqrt{d}$, where $a, b \in \mathbb{Z}$. Again, the minimum polynomial is $p(x) = x^2 - (\overline{\alpha} + \alpha)x + N(x)$. Then $p(x) = x^2 - 2ax + (a^2 - b^2 d)$. This polynomial has rational integer coefficients, so by Proposition 2.32, we find that $\alpha \in Z(\mathbb{Q}(\sqrt{d}))$. Therefore, $\mathbb{Z}[\omega] \subseteq Z(\mathbb{Q}(\sqrt{d}))$. $\square$

This rather bizarre theorem leads, as one may expect, to many strange and non-intuitive results. We will now begin to characterize the irreducibles in this ring of integers. We begin to characterize these irreducibles with two lemmas, followed by a short but important proposition.

Lemma 3.11 *Let* $\alpha \in Z(\mathbb{Q}(\sqrt{d}))$. *Then* $N(\alpha) \in \mathbb{Z}$.

Proof We know that $\alpha = a + b\omega$ for some $a, b \in \mathbb{Z}$, by Theorem 3.10. Consider the case that $d \not\equiv 1(\mathrm{mod}\ 4)$. Then

$$N(\alpha) = (a + b\sqrt{d})(a - b\sqrt{d}) = a^2 - b^2 d \in \mathbb{Z}. \tag{131}$$

Consider the case that $d \equiv 1(\mathrm{mod}\ 4)$. Then

$$\alpha = a + b\left(\frac{1+\sqrt{d}}{2}\right) = \frac{2a+b}{2} + \frac{b}{2}\sqrt{d}. \tag{132}$$

In that case,

$$\begin{aligned} N(\alpha) &= \left(\frac{2a+b}{2} + \frac{b}{2}\sqrt{d}\right)\left(\frac{2a+b}{2} - \frac{b}{2}\sqrt{d}\right) \\ &= \left(a + \frac{b}{2}\right)^2 - d\frac{b^2}{4} = a^2 + ab - (d-1)\frac{b^2}{4}. \end{aligned} \tag{133}$$

Since $d \equiv 1(\mathrm{mod}\ 4)$, we have that $\frac{d-1}{4} \in \mathbb{Z}$, and therefore that $N(\alpha) \in \mathbb{Z}$. $\square$

Lemma 3.12 *Let* $\alpha \in Z(\mathbb{Q}(\sqrt{d}))$. *Then* α *is a unit in* $Z(\mathbb{Q}(\sqrt{d}))$ *if and only if* $N(\alpha) = \pm 1$.

Proof ($\Rightarrow$) Let α be a unit in $Z(\mathbb{Q}(\sqrt{d}))$. Then $\exists\, u \in Z(\mathbb{Q}(\sqrt{d}))$ such that $\alpha u = 1$. By Lemma 3.8, we have that $N(\alpha)N(u) = 1$. Because $N(\alpha), N(u) \in \mathbb{Z}$ by Lemma

3, we have that $N(\alpha) = \pm 1$.

($\Leftarrow$) Let $N(\alpha) = \pm 1$. Then $\alpha\overline{\alpha} = \pm 1$. In that case, $\alpha(\pm\overline{\alpha}) = 1$, and therefore α is a unit. $\square$

Proposition 3.13 *Let $\alpha \in Z(\mathbb{Q}(\sqrt{d}))$. If $N(\alpha)$ is an irreducible in $\mathbb{Z}$, then α is irreducible in $\alpha \in Z(\mathbb{Q}(\sqrt{d}))$.*

Proof Let $\alpha \in Z(\mathbb{Q}(\sqrt{d}))$. Suppose $N(\alpha)$ is an irreducible in $\mathbb{Z}$, and suppose that $\alpha = \beta\gamma$, where $\beta, \gamma \in Z(\mathbb{Q}(\sqrt{d}))$. We must show that one of β and γ is a unit in $Z(\mathbb{Q}(\sqrt{d}))$. We see that

$$N(\alpha) = N(\beta\gamma) = N(\beta)N(\gamma), \tag{134}$$

by Lemma 3.8. Since $N(\alpha)$ is an irreducible in $\mathbb{Z}$, we must have that either $N(\beta)$ or $N(\gamma)$ is a unit in $\mathbb{Z}$. Suppose, with the understanding that the other choice is similar, that $N(\beta)$ is a unit in $\mathbb{Z}$. Then $N(\beta) = \pm 1$, so by Lemma 3.12, we have that β must be a unit in $Z(\mathbb{Q}(\sqrt{d}))$ as well. $\square$

We continue to characterize the rings of quadratic integers by describing the structure of their groups of units. For our purposes, the following theorem, which describes the structure of complex quadratic rings of integers, will be most important.

Theorem 3.14 *Let $\mathbb{Q}(\sqrt{d})$ be a quadratic field, with $d < 0$ and $d \neq -1, -3$. Then*

$$U(Z(\mathbb{Q}(\sqrt{d}))) = \{\pm 1\}. \tag{135}$$

Also,

$$U(Z(\mathbb{Q}(\sqrt{-1}))) = \{\pm 1, \pm\sqrt{-1}\}, \tag{136}$$

and

$$U(Z(\mathbb{Q}(\sqrt{-3}))) = \left\{\pm 1, \pm\left(\frac{1+\sqrt{-3}}{2}\right), \pm\left(\frac{1-\sqrt{-3}}{2}\right)\right\}. \tag{137}$$

Proof ($\subseteq$) We will show that the group of units of the ring of integers of a complex quadratic field fits the above description. Let $\alpha \in U(Z(\mathbb{Q}(\sqrt{d})))$. Then Lemma 3.12 implies that $N(\alpha) = \pm 1$. We consider two cases: either $d \equiv 1 (\mathrm{mod}\ 4)$, or else $d \not\equiv 1 (\mathrm{mod}\ 4)$.

Consider the case that $d \not\equiv 1 (\mathrm{mod}\ 4)$. Then by Theorem 3.10, $\alpha = a + b\sqrt{d}$, for some $a, b \in \mathbb{Z}$. In that case,

$$\pm 1 = N(\alpha) = N(a + b\sqrt{d}) = (a + b\sqrt{d})(a - b\sqrt{d}) = a^2 - db^2. \tag{138}$$

Yet since this is a complex quadratic field, $d < 0$, hence $-d > 0$, and so we see that $a^2 - db^2 \geq 0 > -1$. Thus,

$$a^2 - db^2 = 1. \tag{139}$$

We consider two cases: either $d = -1$, or $d < -1$.

If $d = -1$, then

$$a^2 + b^2 = 1. \tag{140}$$

The possible solutions to this equation for (a, b) are $(\pm 1, 0)$ and $(0, \pm 1)$. This implies that $\alpha = a + b\sqrt{-1} \in \{\pm 1, \pm\sqrt{-1}\}$. Hence,

$$U(Z(\mathbb{Q}(\sqrt{-1}))) \subseteq \{\pm 1, \pm\sqrt{-1}\}. \tag{141}$$

If $d < -1$, then the second term of Equation 139 will be strictly larger than b^2. This requires that $b = 0$, since otherwise $b^2 \geq 1$, and so the second term of Equation 139 will be greater than 1, implying that a^2 is negative, which is impossible. Therefore, the only solutions for (a, b) are $(\pm 1, 0)$. This implies that $\alpha = \pm 1$, and hence,

$$U(Z(\mathbb{Q}(\sqrt{d}))) \subseteq \{\pm 1\} \tag{142}$$

if $d \not\equiv 1 (\mathrm{mod}\ 4)$ and $d < -1$.

Consider the case that $d \equiv 1 (\mathrm{mod}\ 4)$. In that case, Theorem 3.10 implies that $\alpha = a + b\left(\frac{1+\sqrt{d}}{2}\right)$. Therefore,

$$\begin{aligned} \pm 1 = N(\alpha) &= N\left(a + b\left(\frac{1+\sqrt{d}}{2}\right)\right) = N\left(\frac{2a+b}{2} + \frac{b}{2}\sqrt{d}\right) \\ &= \left(\frac{2a+b}{2} + \frac{b}{2}\sqrt{d}\right)\left(\frac{2a+b}{2} - \frac{b}{2}\sqrt{d}\right) = \frac{(2a+b)^2}{4} - \frac{db^2}{4}. \end{aligned} \tag{143}$$

Yet this is a complex quadratic field, so $d < 0$, hence $-d > 0$, implying that $\frac{(2a+b)^2}{4} - \frac{db^2}{4} \geq 0 > -1$. Thus, $\frac{(2a+b)^2}{4} - \frac{db^2}{4} = 1$. Multiplying both sides of the equation by 4 gives us

$$(2a+b)^2 - db^2 = 4. \tag{144}$$

We consider two cases: either $d = -3$, or $d < -3$. (These are the only cases since we are restricting ourselves to negative d satisfying $d \equiv 1 (\mathrm{mod}\ 4)$.

If $d = -3$, then we define $x = 2a + b$ and $y = b$ to find

$$x^2 + 3y^2 = 4. \tag{145}$$

It is not possible for $|y| > 1$, since then the second term of Equation 145 would be greater than 4, implying that the first term is negative, which is impossible. Thus, $y = 0$ or $y = \pm 1$. If $y = 0$, then $x = \pm 2$, so $2a + b = \pm 2$ while $b = 0$. This implies that $a = \pm 1$, and therefore $\alpha = a + b\left(\frac{1+\sqrt{-3}}{2}\right) = \pm 1$ is a solution. If $y = -1$, then $x = \pm 1$, so $2a + b = \pm 1$ while $b = -1$. Therefore, $a = 1$ or $a = 0$, and hence $\alpha = \left(\frac{1-\sqrt{-3}}{2}\right)$ if $a = 1$ and $\alpha = -\left(\frac{1+\sqrt{-3}}{2}\right)$ if $a = 0$. If $y = 1$, then $x = \pm 1$, so $2a + b = \pm 1$ while $b = 1$. This implies that $a = 0$ or $a = -1$, and therefore $\alpha = \left(\frac{1+\sqrt{-3}}{2}\right)$ is a solution if $a = 0$. or $\alpha = \left(\frac{-1+\sqrt{-3}}{2}\right)$. Compiling all of these possibilities, we find that $\alpha \in \left\{\pm 1, \pm\left(\frac{1+\sqrt{-3}}{2}\right), \pm\left(\frac{1-\sqrt{-3}}{2}\right)\right\}$, and therefore,

$$U(Z(\mathbb{Q}(\sqrt{-3}))) \subseteq \left\{\pm 1, \pm\left(\frac{1+\sqrt{-3}}{2}\right), \pm\left(\frac{1-\sqrt{-3}}{2}\right)\right\}. \tag{146}$$

If $d < -3$, then Equation 144 becomes

$$x^2 - dy^2 = 4, \tag{147}$$

where, as before, $x = 2a + b$ and $y = b$. Now, $d < -3$, but also $d \equiv 1 (\mathrm{mod}\ 4)$, which implies that $d \neq -4$, $d \neq -5$, and $d \neq -6$. Thus, $d \leq -7$. Therefore, unless $y = 0$, we have that the right term of Equation 147 is greater than 4, implying that

the left term is negative, which is impossible. Thus, $y = 0$, and $x = \pm 2$, hence $2a + b = \pm 2$ while $b = 0$. This implies that $\alpha = a + b\left(\frac{1+\sqrt{d}}{2}\right) = \pm 1$. Thus, for $d < -3$, $U(Z(\mathbb{Q}(\sqrt{d}))) \subseteq \{\pm 1\}$. By Equation 142, we also have that $d = -2$ or $d < -3$ with $d \not\equiv 1 (\mathrm{mod}\ 4)$ implies that

$$U(Z(\mathbb{Q}(\sqrt{d}))) \subseteq \{\pm 1\}, \tag{148}$$

so for all $d < 0$ such that $d \neq -1$ and $d \neq -3$, $U(Z(\mathbb{Q}(\sqrt{d}))) \subseteq \{\pm 1\}$.

($\supseteq$) It is clear that ± 1 are units in the ring of integers of any complex quadratic field, since $(1)(1) = 1$ and $(-1)(-1) = 1$. From Theorem 3.10, we know that $\pm\sqrt{-1} \in Z(\mathbb{Q}(\sqrt{-1}))$. It is a simple computational matter to show that each of $\pm\sqrt{-1}$ is the multiplicative inverse of the other. It is clear then, that

$$\{\pm 1, \pm\sqrt{-1}\} \subseteq U(Z(\mathbb{Q}(\sqrt{d}))). \tag{149}$$

Also using Theorem 3.10, $\pm\left(\frac{1+\sqrt{-3}}{2}\right), \pm\left(\frac{1-\sqrt{-3}}{2}\right) \in Z(\mathbb{Q}(\sqrt{-3}))$. It is a simple matter to show that the norm of each of these four is equal to 1. By Lemma 3.12, this implies that each of these are units. Thus,

$$\left\{\pm 1, \pm\left(\frac{1+\sqrt{-3}}{2}\right), \pm\left(\frac{1-\sqrt{-3}}{2}\right)\right\} \subseteq U(Z(\mathbb{Q}(\sqrt{-3}))). \tag{150}$$

This completes the proof. □

We have closely followed the proof on page 39 of reference [8] in writing the previous proof. For more information about groups of units in quadratic integer rings,

including a proof of the fact that if $d > 0$, then $U(Z(\mathbb{Q}(\sqrt{d})))$ is an infinite group, see reference [7], especially near pages 274-275.

It is a fact that factorization into irreducibles is guaranteed for the ring of integers of any number field (we will not show this, yet the interested reader may see pages 87-88 of reference [8] for a proof). However, this factorization is not always unique. Further, the factorization is not always unique for a ring of integers of a quadratic field. Likewise, some quadratic fields are more interesting than others. The next theorem will characterize several of the quadratic integer rings, including $Z(\mathbb{Q}(\sqrt{-1}))$ (the Gaussian integers) and $Z(\mathbb{Q}(\sqrt{-7}))$, which we will use later. Before this, though, we will introduce a lemma.

Lemma 3.15 *Suppose that $\mathbb{Q}(\sqrt{d})$ is a quadratic field. If it is a true statement that $\forall\ \varepsilon \in \mathbb{Q}(\sqrt{d})\ \exists\ \kappa \in Z(\mathbb{Q}(\sqrt{d}))$ such that $N(\varepsilon - \kappa) < 1$, then $Z(\mathbb{Q}(\sqrt{d}))$ is a Euclidean Domain.*

Proof Let $\mathbb{Q}(\sqrt{d})$ be a quadratic field that satisfies the hypothesis of the lemma. Suppose that $\alpha, \beta \in Z(\mathbb{Q}(\sqrt{d}))$ with $\beta \neq 0$ and define $\varepsilon = \frac{\alpha}{\beta}$. We will produce $\gamma, \delta \in Z(\mathbb{Q}(\sqrt{d}))$ such that $\alpha = \beta\gamma + \delta$ with either $\delta = 0$ or $N(\delta) < N(\beta)$.

Consider the case that $\varepsilon \in Z(\mathbb{Q}(\sqrt{d}))$. Then $\gamma = \varepsilon$, $\delta = 0$ satisfies the condition. Consider the case that $\varepsilon \notin Z(\mathbb{Q}(\sqrt{d}))$. We know that $\varepsilon \in \mathbb{Q}(\sqrt{d})$. By the assumption, then, $\exists\ \kappa \in Z(\mathbb{Q}(\sqrt{d}))$ such that $N(\varepsilon - \kappa) < 1$. Then, with this κ, we have

$$1 > N(\varepsilon - \kappa) = N\left(\frac{\alpha}{\beta} - \kappa\right) = N\left(\frac{\alpha}{\beta} - \frac{\kappa\beta}{\beta}\right) = N\left((\alpha - \kappa\beta)\,\frac{1}{\beta}\right)$$
$$= N(\alpha - \kappa\beta)\frac{1}{\beta}\overline{\left(\frac{1}{\beta}\right)} = N(\alpha - \kappa\beta)\frac{1}{\beta}\frac{1}{\overline{\beta}}. \quad (151)$$

(The last equality is easy to show, and we leave its proof to the reader.) Then $N(\alpha - \kappa\beta)\frac{1}{N(\beta)} = N(\varepsilon - \kappa) < 1$, so by clearing the denominator, we find that $N(\alpha - \kappa\beta) < N(\beta)$. Thus, setting $\delta = \alpha - \kappa\beta$ and $\gamma = \kappa$ satisfies the condition, proving that $Z(\mathbb{Q}(\sqrt{d}))$ is a Euclidean Domain. $\square$

Now that we have established that the hypothesis of Lemma 3.15 is a sufficient condition to being a Euclidean Domain, we need only show that it is true for whatever quadratic field we intend in order to show that that quadratic field's ring of integers is a Euclidean Domain.

Theorem 3.16 *If $d \in \{-1, -2, -3, -7, -11\}$, then the norm serves as a Euclidean field norm in the ring of integers $Z(\mathbb{Q}(\sqrt{d}))$.*

Proof Suppose that $\varepsilon \in \mathbb{Q}(\sqrt{d})$, for $d \in \{-1, -2, -3, -7, -11\}$. We will show that $\exists\ \kappa \in Z(\mathbb{Q}(\sqrt{d}))$ such that $N(\varepsilon - \kappa) < 1$. We consider two cases: either $d \not\equiv 1 (\mathrm{mod}\ 4)$ or $d \equiv 1 (\mathrm{mod}\ 4)$.

Consider the case that $d \not\equiv 1 (\mathrm{mod}\ 4)$, so that $d = -1$ or $d = -2$. Then, by Theorem 3.10, we have that $Z(\mathbb{Q}(\sqrt{d})) = \mathbb{Z}[\sqrt{d}]$. Let $\varepsilon = r + s\sqrt{d}$. We must produce a $\kappa = x + y\sqrt{d}$, where $x, y \in \mathbb{Z}$, such that $N(\varepsilon - \kappa) < 1$. Let x be a rational integer such that $|r - x|$ is minimal, and let y be the rational integer such that $|s - y|$ is minimal. Then $|r - x| \leq \frac{1}{2}$ and $|s - y| \leq \frac{1}{2}$. Therefore,

$$N(\varepsilon - \kappa) = N\left((r - x) + (s - y)\sqrt{d}\right) = (r - x)^2 - d(s - y)^2$$
$$\leq \left(\frac{1}{2}\right)^2 + 2\left(\frac{1}{2}\right)^2 = \frac{3}{4} < 1. \quad (152)$$

Thus, the hypothesis of Lemma 3.15 is satisfied, and so $Z(\mathbb{Q}(\sqrt{-1}))$ and $Z(\mathbb{Q}(\sqrt{-2}))$ are Euclidean Domains.

Consider the case that $d \equiv 1 (\mathrm{mod}\ 4)$, so that $d = -3$, $d = -7$, or $d = -11$. Then by Theorem 3.10, we have that $Z(\mathbb{Q}(\sqrt{d})) = \mathbb{Z}[\frac{1+\sqrt{d}}{2}]$. Thus, we must produce a $\kappa = x + y(\frac{1+\sqrt{d}}{2})$, where $x, y \in \mathbb{Z}$, such that $N(\varepsilon - \kappa) < 1$. Let y be a rational integer such that $|2s - y|$ is minimal, and let x be a rational integer such that $|(r - \frac{1}{2}y) - x|$ is minimal. Then $|2s - y| \leq \frac{1}{2}$, and so $|s - \frac{1}{2}y| \leq \frac{1}{4}$ and likewise $|(r - \frac{1}{2}y) - x| \leq \frac{1}{2}$. Then

$$\begin{aligned} N(\varepsilon - \kappa) = N\left(\left(r - x - \frac{1}{2}y\right) - \left(s - \frac{1}{2}y\right)\sqrt{d}\right) \\ = \left(r - x - \frac{1}{2}y\right)^2 - d\left(s - \frac{1}{2}y\right)^2 \\ \leq \left(\frac{1}{2}\right)^2 + 11\left(\frac{1}{4}\right)^2 = \frac{15}{16} < 1. \quad (153) \end{aligned}$$

Thus, the hypothesis of Lemma 3.15 is satisfied, and so $Z(\mathbb{Q}(\sqrt{-3}))$, $Z(\mathbb{Q}(\sqrt{-7}))$ and $Z(\mathbb{Q}(\sqrt{-11}))$ are Euclidean Domains. $\square$

We have closely followed the proof on pages 99-100 of reference [8] in writing the previous proof.

Notice that the argument above would not have worked if d had been less than -11. In fact, though we will not do so in this thesis, one can prove that $Z(\mathbb{Q}(\sqrt{d}))$ is not Euclidean for $d < -11$ (for a proof of this, see page 101 of reference [8]). Likewise, the d that yield $Z(\mathbb{Q}(\sqrt{d}))$ that are Euclidean Domains or Unique Factorization Domains are very non-intuitive, even seemingly random. The above is only

one example of many *ad hoc* proofs to show that certain d have these properties. The full catalog of rings of integers in quadratic fields has yet to be finished and is a topic of current research. A full discussion of this topic could constitute another thesis in and of itself. The topic of this thesis, however, is the use of unique factorization to solve number-theoretic problems. Concurrently, we have proven the above with the main intention of applying it to the proofs of two important number theoretic results. The next chapter will show these two examples of consequences of the above theorem.

4 Applications of Unique Factorization

In this final chapter, we will discuss two results of number theory that can be proven in an efficient and simple way using theorems of Abstract Algebra that we have built up in the first three chapters of this thesis. The first is Fermat's Theorem on Sums of Squares. Before discussing this, however, we will construct an interconnected base of lemmas.

Lemma 4.1 *Let F be a finite field. Then $U(F)$ is cyclic.*

Proof We will omit this proof. The interested reader may see page 185 of reference [6] for a proof of this statement. □

Lemma 4.2 *Let $a, n \in \mathbb{Z}$. If a and n are relatively prime, then $[a] \in U\left(\mathbb{Z}/n\mathbb{Z}\right)$.*

Proof Let a and n be relatively prime. We direct the reader to page 23 of reference [3] for a proof that this implies that $\exists\ x, y \in \mathbb{Z}$ such that $ax + ny = 1$. Thus, $ax - 1 = ny$, so $ax - 1 \in n\mathbb{Z}$. Therefore, under the equivalence relation that defines $\mathbb{Z}/n\mathbb{Z}$, $[ax] = [1]$, so $[a][x] = 1$, hence $[a]$ has a multiplicative inverse in $\mathbb{Z}/n\mathbb{Z}$, and therefore $a \in U\left(\mathbb{Z}/n\mathbb{Z}\right)$. □

Lemma 4.3 *Let G be a cyclic group, and suppose that $\exists\, n \in \mathbb{Z}^+$ such that n divides the order of G. Then $\exists\ a \in G$ such that $o(a) = n$.*

Proof Let $G = \{e, a, a^2, ..., a^{m-1}\}$, where $e = a^m$ is the identity element, and $|G| = m$. Assume that $n|m$ for some $n \in \mathbb{Z}^+$. We will show that there exists an element that generates a cyclic subgroup of order n. Let $l = \frac{m}{n}$, and let $b = a^l \in G$. We claim that the order of b in G is n. First of all, we know that it is true that

$b^n = (a^l)^n = a^{nl} = a^m = e$. Thus, $o(b)|n$. Assume for contradiction that $o(b) < n$. Then $b^k = e$ for some $k < n$. In that case, $e = b^k = (a^l)^k = a^{kl}$, despite that $kl < nl = m$. Therefore, the order of G is not m, despite the fact that we have defined the order of G to be m. This contradiction leads us to conclude that $o(b) = n$. Thus, the cyclic subgroup generated by b has an order of n, as desired. □

Lemma 4.4 *If $p \in \mathbb{Z}$ is a prime number such that $p \equiv 1(\mathrm{mod}\ 4)$, then $\exists\ n \in \mathbb{Z}$ such that $p|n^2 + 1$.*

Proof Let $p \equiv 1(\mathrm{mod}\ 4)$. We will show that $p|n^2 + 1$ for some $n \in \mathbb{Z}$. The elements of $\mathbb{Z}/p\mathbb{Z}$ are equivalence classes $[0], [1], [2], ..., [p-1]$. But all of the integers $1, 2, ..., p-1$ are relatively prime to p, so from Lemma 4.2 it follows that $U\left(\mathbb{Z}/p\mathbb{Z}\right) = \{[1], [2], ..., [p-1]\}$, which has order $p-1$. By assumption, $4|p-1$, so 4 divides the order of this group of units. We also know that $\mathbb{Z}/p\mathbb{Z}$ is a finite field, since each of its elements is a unit except for $[0]$. Then $U\left(\mathbb{Z}/p\mathbb{Z}\right)$ is cyclic, by Lemma 4.1. Then by Lemma 4.3, $\exists\ [n] \in U\left(\mathbb{Z}/p\mathbb{Z}\right)$ such that $o([n]) = 4$. In that case, $[n]^4 = [1]$, so by definition of the equivalence relation that defines $\mathbb{Z}/p\mathbb{Z}$, $p|n^4 - 1$. We know that $n^4 - 1 = (n^2 + 1)(n^2 - 1)$. Since p is a prime in $\mathbb{Z}$, then, we see that $p|n^2 \pm 1$. Therefore, $n^2 \equiv \pm 1(\mathrm{mod}p)$. However, it cannot be true that $n^2 \equiv 1(\mathrm{mod}p)$, since then the order of $[n]$ in $U\left(\mathbb{Z}/p\mathbb{Z}\right)$ would be 2 or less, in spite of our understanding that $o(n) = 4$ in $U\left(\mathbb{Z}/p\mathbb{Z}\right)$. Thus, $n^2 \equiv -1(\mathrm{mod}p)$, so $p|n^2 + 1$ for some $n \in \mathbb{Z}$, as desired. □

Lemma 4.5 *If an integer p is simultaneously irreducible in $\mathbb{Z}$ and reducible in $\mathbb{Z}[i]$, then $p = a^2 + b^2$ for some $a, b \in \mathbb{Z}$. If we let $x = a^2$ and $y = b^2$, then the only representations of p as the sum of two perfect squares are $p = x + y$ and $p = y + x$.*

Proof Let p be irreducible in $\mathbb{Z}$ and reducible in $\mathbb{Z}[i]$. Then $\exists\ \alpha, \beta \in \mathbb{Z}[i]$ that are not units that satisfy $p = \alpha\beta$. Then by Lemma 3.8, $N(p) = N(\alpha\beta) = N(\alpha)N(\beta)$. We also know that $N(p) = p^2$, by the definition of the field norm on the Gaussian integers. Therefore, since the ring of integers is a Unique Factorization Domain by Theorem 1.31, we see that p^2 is a unique factorization of the rational integer $N(\alpha)N(\beta)$ into primes:

$$N(\alpha)N(\beta) = p^2. \tag{154}$$

Since p is a prime, it divides one of $N(\alpha)$ and $N(\beta)$, as a virtue of dividing their product. Suppose, with the understanding that the other choice is similar, that p divides $N(\alpha)$, so that $N(\alpha) = kp$ for some $k \in \mathbb{Z}$. Then $kN(\beta) = p$. Since all primes are irreducible in $\mathbb{Z}$, this means that one of k and $N(\beta)$ is a unit, and the only units of $\mathbb{Z}$ are ± 1. $N(\beta)$ cannot be a unit, since $N(\beta) > 0$ by definition of the Gaussian integers as a complex quadratic field, and $N(\beta) = 1$ would imply that β is a unit because of Lemma 3.12, contrary to our definition. Thus, k must be a unit, and therefore $N(\alpha) = kp = \pm p$. Because $\mathbb{Z}[i]$ is a complex quadratic field,

$$N(\alpha) = p, \tag{155}$$

by virtue of the norm being nonnegative. Note that, by Proposition 3.13, α is an irreducible in $\mathbb{Z}[i]$, as is $\overline{\alpha}$, since their norm $N(\alpha) = N(\overline{\alpha}) = \alpha\overline{\alpha} = p$, which is a rational prime. Now, since $\alpha \in \mathbb{Z}[i]$, we know that $\alpha = a + bi$ for some $a, b \in \mathbb{Z}$. Then

$$N(\alpha) = \alpha\overline{\alpha} = (a + bi)(a - bi) = a^2 + b^2 = p. \tag{156}$$

Thus, p can be represented as a sum of the perfect squares a^2 and b^2. Futher, since $\mathbb{Z}[i]$ is a Unique Factorization Domain, $p = (a + bi)(a - bi)$ is a unique factorization of p into irreducibles excepting multiplication by units or reversal of order. Any representation within these constraints will still yield that $p = a^2 + b^2$ (since addition is commutative, and any series of units that multiply the right side of the equation $p = (a+bi)(a-bi)$ must multiply to 1), and so this representation of p as a sum of two perfect squares is unique up to reversal of order. $\square$

We are now ready to construct the algebraic proof of Fermat's Theorem on sums of squares.

Theorem 4.6 *The prime number $p \in \mathbb{Z}$ can be represented as a sum of perfect squares, $p = a^2 + b^2$, with $a, b \in \mathbb{Z}$, if and only if $p = 2$ or $p \equiv 1 (\mathrm{mod}\ 4)$. If a and b are two integers that satisfy this representation, then the only representations of p as the sum of two perfect squares are $p = x + y$ and $p = y + x$, where $x = (\pm a)^2$ and $y = (\pm b)^2$. This is called Fermat's Theorem on Sums of Squares.*

Proof ($\Rightarrow$) Suppose that a prime p satisfies $p = a^2 + b^2$ for some $a, b \in \mathbb{Z}$. We will show that $p = 2$ or $p \equiv 1 (\mathrm{mod}\ 4)$. If $a = b = 1$, then $p = 2$. Suppose that $p \neq 2$. Then p is odd, and therefore either $p \equiv 1 (\mathrm{mod}\ 4)$ or $p \equiv 3 (\mathrm{mod}\ 4)$. We know that a is either even or odd. If a is even, then $a \equiv 0 (\mathrm{mod}\ 4)$ or $a \equiv 2 (\mathrm{mod}\ 4)$, hence we see that $a^2 \equiv 0 (\mathrm{mod}\ 4)$, and if a is odd, then $a \equiv 1 (\mathrm{mod}\ 4)$ or $a \equiv 3 (\mathrm{mod}\ 4)$, hence $a^2 \equiv 1 (\mathrm{mod}\ 4)$. The same argument applies to b, and so the sum $a^2 + b^2$ is equivalent to 0, 1, or 2 modulo 4. Therefore, the only possible situation is that $p \equiv 1 (\mathrm{mod}\ 4)$.

($\Leftarrow$) Let $p \in \mathbb{Z}$ be a prime such that $p = 2$ or $p \equiv 1 (\mathrm{mod}\ 4)$. We will show

that p can be represented as a sum of perfect squares that are unique up to sign and order. Consider the case that $p = 2$. If so, then $p = 1 + 1 = 1^2 + 1^2$. Let $a, b \in \mathbb{Z}$ be such that $p = 2 = a^2 + b^2$. Then since $a^2, b^2 > 0$, either one of a^2 and b^2 must be 0 and the other must be 2, or else both must be equal to 1. But 2 is not a perfect square, so 1 and 1 are the only perfect squares whose sum is equal to 2. Consider the case that $p \equiv 1 (\mathrm{mod}\ 4)$. If that is the case, then by Lemma 4.4, $p|n^2 + 1$ for some $n \in \mathbb{Z}$. But $n^2 + 1 = (n + i)(n - i)$ in $\mathbb{Z}[i]$, so in the Gaussian Integers,

$$p|(n + i)(n - i). \tag{157}$$

Suppose for contradiction that p is irreducible in $\mathbb{Z}[i]$. We know from Theorem 3.16 that $\mathbb{Z}[i]$ is a Euclidean Domain, since it is the ring of integers of $\mathbb{Q}(\sqrt{-1})$ by Theorem 3.10. Using Theorems 1.13 and 1.30, we find that this implies that $\mathbb{Z}[i]$ is a Unique Factorization Domain. By Proposition 1.27, then, p is a prime in $\mathbb{Z}[i]$. If that is correct, then since $p|(n + i)(n - i)$,

$$p|n \pm i \tag{158}$$

in $\mathbb{Z}[i]$. In that case, for some $r + si \in \mathbb{Z}[i]$,

$$n \pm i = (r + si)p = rp + spi. \tag{159}$$

By the properties that $\mathbb{Z}[i]$ inherits from $\mathbb{C}$, $\pm i = spi$, and therefore $\pm 1 = sp$. Therefore $1 = \pm sp = (\pm s)p$, and so p is a unit. However, p must not be a unit, because irreducibles are, by definition, not units. Thus, we reach a contradiction

that leads us to conclude that p is reducible in $\mathbb{Z}[i]$. If that is true, then by Lemma 4.5, we see that the statement is entirely proven. $\square$

Notice the interplay between algebra and number theory that has taken place in this proof. Lemma 4.2 was essentially an algebraic phrasing of the common number-theoretic issue of modular units. On the other hand, Lemma 4.4 draws heavily on concepts from the theory of cyclic groups, arriving almost startlingly at the implication that $4|p-1$ can only be true if there exists an n satisfying $p|n^4-1$. The almost magical ease of this transition is an example of the ways that algebra can make extremely non-intuitive number-theoretic results appear almost effortlessly. Lemma 4.5 relies heavily on the uniqueness of factorization in order to prove the uniquenes of representation of the prime as the sum of two perfect squares. The theorem's proof when determining that assuming that the prime is irreducible implies that it is prime in the Gaussian integers is also highly dependent on unique factorization. One could imagine how *ad hoc* and awkward these dealings would have been if only a number-theoretic language had been used. Only looking at a broader scope of mathematics and using number theory and algebra as separate parts of the same machine can one attain a mechanism that both succintly proves theorems and also generates more interesting topics.

We conclude this thesis with a discussion of the proof of one more number-theoretic theorem with a heavily algebraic proof, a theorem proposed by the great mathematician Srinivasa Ramanujan and proven by Trygve Nagell. We will begin with a notational convention, and a lemma. Even though congruence modulo n is usually defined as being an equivalence relation in $\mathbb{Z}/n\mathbb{Z}$, for the following proofs,

we will consider congruence in a more general setting. Let $a, b, n \in Z(\mathbb{Q}(\sqrt{d}))$. Then $a \equiv b (\mathrm{mod}\ n)$ will denote the statement $a - b \in (n)$, where (n) is the ideal generated by n in $Z(\mathbb{Q}(\sqrt{d}))$.

Lemma 4.7 *Let $r, s, n \in \mathbb{Z}^+$. Then*

$$\left(1 + r\left(\sqrt{-7}\right)^s\right)^{(7^n)} \equiv 1 + 7^n r\left(\sqrt{-7}\right)^s (\mathrm{mod}\ 7^{n+1}). \tag{160}$$

Proof Let $r, s, n \in \mathbb{Z}^+$. We will show by mathematical induction that the statement is true. Let $n = 1$. We will show that the statement holds; *id est*, we will show that $\left(1 + r\left(\sqrt{-7}\right)^s\right)^7 \equiv 1 + 7r\left(\sqrt{-7}\right)^s (\mathrm{mod}\ 7^2)$. By the binomial theorem, we have that

$$\begin{aligned}\left(1 + r\left(\sqrt{-7}\right)^s\right)^7 = \binom{7}{0} + \binom{7}{1} r\left(\sqrt{-7}\right)^s + \binom{7}{2} r^2\left(\sqrt{-7}\right)^{2s} + \\ \binom{7}{3} r^3\left(\sqrt{-7}\right)^{3s} + ... + \binom{7}{7} r^7\left(\sqrt{-7}\right)^{7s} = \\ 1 + 7r\left(\sqrt{-7}\right)^s + \frac{(7)(6)}{2} r^2(-7)^s + \\ \frac{(7)(6)(5)}{(3)(2)} r^3(-7)^s\left(\sqrt{-7}\right)^s + ... + r^7\left(\sqrt{-7}\right)^{7s}.\end{aligned} \tag{161}$$

Now, 6 is divisible by both 2 and $(3)(2)$ in $Z(\mathbb{Q}(\sqrt{d}))$ (we know this because any rational integer is a quadratic integer, as given by Theorem 3.10). Thus, we find that terms 2 and 3 (counting from 0) are divisible by 7^2. Any terms after term 3 are also divisible by 7^2, because they possess powers of $\sqrt{-7}$ that are greater than or equal to $\sqrt{-7}^4 = 7^2$. Thus,

$$\left(1 + r\left(\sqrt{-7}\right)^s\right)^7 \equiv 1 + 7r\left(\sqrt{-7}\right)^s (\mathrm{mod}\ 7^2). \tag{162}$$

Therefore, the lemma holds for $n = 1$. Let $n = l - 1$ for some arbitrary $l \geq 2$ and suppose that the statement is true for n as the inductive hypothesis. We will show that the statement is true for $n + 1 = l$. By the induction hypothesis,

$$\left(1 + r\left(\sqrt{-7}\right)^s\right)^{\left(7^{l-1}\right)} = 1 + 7^{l-1} r\left(\sqrt{-7}\right)^s + 7^l z, \tag{163}$$

and therefore,

$$\begin{aligned}\left(1 + r\left(\sqrt{-7}\right)^s\right)^{\left(7^l\right)} &= \left(\left(1 + r\left(\sqrt{-7}\right)^s\right)^{\left(7^{l-1}\right)}\right)^7 \\ &= \left(1 + 7^{l-1} r\left(\sqrt{-7}\right)^s + 7^l z\right)^7. \end{aligned} \tag{164}$$

We use the binomial theorem to expand the right side of this equation, considering $1 + 7^{l-1} r\left(\sqrt{-7}\right)^s$ and 7^l as the two arguments:

$$\begin{aligned}\left(1 + r\left(\sqrt{-7}\right)^s\right)^{\left(7^l\right)} &= \binom{7}{0}\left(1 + 7^{l-1} r\left(\sqrt{-7}\right)^s\right)^7 \\ &\quad + \binom{7}{1}\left(1 + 7^{l-1} r\left(\sqrt{-7}\right)^s\right)^6 7^l z \\ &\quad + \binom{7}{2}\left(1 + 7^{l-1} r\left(\sqrt{-7}\right)^s\right)^5 7^{2l} z^2 + \ldots + \binom{7}{7} 7^{7l} z^l \\ &= \left(1 + 7^{l-1} r\left(\sqrt{-7}\right)^s\right)^7 + 7^{l+1}\alpha, \end{aligned} \tag{165}$$

for some $\alpha \in Z(\mathbb{Q}(\sqrt{-7}))$. We apply the binomial theorem again to get

$$\left(1 + r\left(\sqrt{-7}\right)^s\right)^{(7^l)} = \binom{7}{0} + \binom{7}{1}7^{l-1}r\left(\sqrt{-7}\right)^s + \binom{7}{2}\left(7^{l-1}r\left(\sqrt{-7}\right)^s\right)^2 + ...$$
$$+ \binom{7}{7}\left(7^{l-1}r\left(\sqrt{-7}\right)^s\right)^7 + 7^{l+1}\alpha = 1 + 7^l r\left(\sqrt{-7}\right)^s + 7^{l+1}\beta + 7^{l+1}\alpha$$
$$\equiv 1 + 7^l r\left(\sqrt{-7}\right)^s (\text{mod } 7^{l+1}), \quad (166)$$

for some $\beta \in Z(\mathbb{Q}(\sqrt{d}))$. Examining this, we see that this is the statement in the case of l. Thus, the statement is true by mathematical induction. $\square$

We are now able to move on to the heavily algebraic proof of the Ramanujan-Nagell Theorem.

Theorem 4.8 *Let* $(x, n) \in \mathbb{Z}^2$. *The only solutions* (x, n) *of* $x^2 = 2^n - 7$ *are* $(\pm 1, 3), (\pm 3, 4), (\pm 5, 5), (\pm 11, 7)$, *and* $(\pm 181, 15)$. *This is called the Ramanujan-Nagell Theorem.*

Proof We will show that the above are the only possibilities for ordered pairs of x and n. If $n \leq 2$, then it is clear that $x \notin \mathbb{R}$, since then $x^2 = 2^n - 7 < 0$. Therefore let $n \geq 3$. We see that, because 2^n is even, x must be odd because $x^2 = 2^n - 7$. We consider two cases: either n is even, or n is odd.

Consider the case that n is even. Then because $2^{\frac{n}{2}} \in \mathbb{Z}$, there is a factorization in $\mathbb{Z}$ as

$$7 = 2^n - x^2 = (2^{\frac{n}{2}} + x)(2^{\frac{n}{2}} - x). \quad (167)$$

Because 7 is irreducible in $\mathbb{Z}$, and by Theorem 1.31, we then see that one of these factors is a unit, so one of $2^{\frac{n}{2}} + x$ and $2^{\frac{n}{2}} - x$ is equal to ± 1, and the other is equal

to ± 7. If x is positive, then $2^{\frac{n}{2}} + x > 2^{\frac{n}{2}} - x$, so $2^{\frac{n}{2}} + x = 7$, since it is the sum of two positive integers. Likewise, $2^{\frac{n}{2}} - x = 1$, by process of elimination. If x is negative, then $2^{\frac{n}{2}} + x < 2^{\frac{n}{2}} - x$, and then $2^{\frac{n}{2}} - x = 2^{\frac{n}{2}} + (-x) = 7$, because it is the sum of two positive integers, and $2^{\frac{n}{2}} + x = 1$, by process of elimination. In either case, we can add these two to find that $2^{1+\frac{n}{2}} = 8$. It follows that $n = 4$ and $x = \pm 3$ are the only possible solutions for an even n.

Consider the case that n is odd. If $n = 3$, then it is clear from brute force calculation that $x = \pm 1$ are the only solutions. Therefore let $n > 3$. By Theorem 3.16, we know that $Z(\mathbb{Q}(\sqrt{-7}))$ is a Euclidean Domain. Using Theorems 1.13 and 1.30, it is clear that $Z(\mathbb{Q}(\sqrt{-7}))$ is a Unique Factorization Domain. Now, in $Z(\mathbb{Q}(\sqrt{-7}))$,

$$2 = \left(\frac{1+\sqrt{-7}}{2}\right)\left(\frac{1-\sqrt{-7}}{2}\right) = N\left(\frac{1\pm\sqrt{-7}}{2}\right). \tag{168}$$

(It is a simple arithmatic task to show that $2, \frac{1+\sqrt{-7}}{2} \in Z(\mathbb{Q}(\sqrt{-7}))$ using Theorem 3.10.) By Proposition 3.13, $\frac{1\pm\sqrt{-7}}{2}$ are irreducibles in $Z(\mathbb{Q}(\sqrt{-7}))$, since their norm is an irreducible in $\mathbb{Z}$. Because $n > 2$, $4|2^n$, and therefore $4|x^2 + 7$ in $Z(\mathbb{Q}(\sqrt{-7}))$. Let $m = n - 2$. Then $\frac{x^2+7}{4} = 2^m$. In that case, using Equation 168,

$$\left(\frac{x+\sqrt{-7}}{2}\right)\left(\frac{x-\sqrt{-7}}{2}\right) = \left(\frac{1+\sqrt{-7}}{2}\right)^m\left(\frac{1-\sqrt{-7}}{2}\right)^m, \tag{169}$$

where the right side is a unique factorization into irreducibles in $Z(\mathbb{Q}(\sqrt{-7}))$.

We will show that $\frac{1+\sqrt{-7}}{2}$ cannot simultaneously divide $\frac{x+\sqrt{-7}}{2}$ and $\frac{x-\sqrt{-7}}{2}$. Suppose for contradiction that $\frac{x+\sqrt{-7}}{2} = \gamma_1(\frac{1+\sqrt{-7}}{2})$ and $\frac{x-\sqrt{-7}}{2} = \gamma_2(\frac{1+\sqrt{-7}}{2})$, where $\gamma_1, \gamma_2 \in Z(\mathbb{Q}(\sqrt{d}))$. Then by the distributive axiom of the ring of quadratic inte-

gers,

$$\sqrt{-7} = \frac{x+\sqrt{-7}}{2} - \frac{x-\sqrt{-7}}{2} = (\gamma_1 - \gamma_2)\left(\frac{1+\sqrt{-7}}{2}\right). \tag{170}$$

It is simple to verify that this implies that $\gamma_1 - \gamma_2 = \frac{3}{2} + \frac{1}{2}\left(\frac{1+\sqrt{-7}}{2}\right) \notin Z(\mathbb{Q}(\sqrt{-7}))$, by Theorem 3.10. Thus, $\gamma_1, -\gamma_2 \in Z(\mathbb{Q}(\sqrt{d}))$, but their sum is not, despite $Z(\mathbb{Q}(\sqrt{d}))$ being a ring. This contradiction leads us to the conclusion that $\frac{1+\sqrt{-7}}{2}$ cannot simultaneously divide $\frac{x-\sqrt{-7}}{2}$ and $\frac{x+\sqrt{-7}}{2}$. Thus, $\left(\frac{1+\sqrt{-7}}{2}\right)^m$ cannot simultaneously divide the two. Yet because $Z(\mathbb{Q}(\sqrt{-7}))$ is a unique factorization domain, we deduce from Proposition 1.27, that $\frac{1+\sqrt{-7}}{2}$ is a prime, by virtue of being irreducible. Thus, $\frac{1+\sqrt{-7}}{2} | \frac{x\pm\sqrt{-7}}{2}$. But since it cannot divide both of them, if $\frac{1+\sqrt{-7}}{2}$ divides one of $\frac{x\pm\sqrt{-7}}{2}$, then $\left(\frac{1+\sqrt{-7}}{2}\right)^m$ divides that same one and not the other.

By the same arguments, if $\frac{1-\sqrt{-7}}{2}$ divides one of $\frac{x\pm\sqrt{-7}}{2}$, then $\left(\frac{1-\sqrt{-7}}{2}\right)^m$ divides that same one and not the other. Now, suppose for contradiction that there is some $\frac{x\pm\sqrt{-7}}{2}$ that is divisible by both of them. Then neither of these divides $\frac{x\mp\sqrt{-7}}{2}$, and so $\frac{x\mp\sqrt{-7}}{2}$ must be a unit, by definition of unique factorization. Then by Theorem 3.14, we have that $\frac{x\mp\sqrt{-7}}{2} = (-1)^{e_1}$ for some $e_1 \in \mathbb{Z}$. But this implies that $x = 2(-1)^{e_1} \pm \sqrt{7}$, which contradicts our definition of x as a rational integer. Thus, each of $(\frac{1\pm\sqrt{-7}}{2})^m$ divides exactly one of $\frac{x\pm\sqrt{-7}}{2}$. Therefore,

$$\frac{x \pm \sqrt{-7}}{2} = \delta_1\left(\frac{1+\sqrt{-7}}{2}\right)^m \tag{171}$$

and

$$\frac{x \mp \sqrt{-7}}{2} = \delta_2\left(\frac{1-\sqrt{-7}}{2}\right)^m, \tag{172}$$

for some $\delta_1, \delta_2 \in Z(\mathbb{Q}(\sqrt{-7}))$.

From this, we deduce that

$$\left(\frac{1+\sqrt{-7}}{2}\right)^m \left(\frac{1-\sqrt{-7}}{2}\right)^m = \left(\frac{x+\sqrt{-7}}{2}\right)\left(\frac{x-\sqrt{-7}}{2}\right) = \delta_1\delta_2\left(\frac{1+\sqrt{-7}}{2}\right)^m \left(\frac{1-\sqrt{-7}}{2}\right)^m. \quad (173)$$

It is clear that $Z(\mathbb{Q}(\sqrt{-7}))$ is an integral domain, since it is a subring of $\mathbb{C}$. Therefore,

$$(1-\delta_1\delta_2)\left(\frac{1+\sqrt{-7}}{2}\right)^m \left(\frac{1-\sqrt{-7}}{2}\right)^m = 0 \quad (174)$$

implies that $1-\delta_1\delta_2 = 0$, which implies that δ_1 and δ_2 are units. By Theorem 3.14, then, we have that $\delta_1, \delta_2 \in \{\pm 1\}$. Further, in order for $\delta_1\delta_2 = 1$ in this case, we must have that $\delta_1 = \delta_2 = (-1)^{e_2}$ for some $e_2 \in \mathbb{Z}^+$. Then it follows that $\frac{x\pm\sqrt{-7}}{2} = (-1)^{e_2}\left(\frac{1+\sqrt{-7}}{2}\right)^m$ and $\frac{x\mp\sqrt{-7}}{2} = (-1)^{e_2}\left(\frac{1-\sqrt{-7}}{2}\right)^m$. From this we derive the fact that

$$\pm\sqrt{-7} = \left(\frac{1+\sqrt{-7}}{2}\right)^m - \left(\frac{1-\sqrt{-7}}{2}\right)^m. \quad (175)$$

We claim that the plus sign cannot occur. Suppose for contradiction that

$$\sqrt{-7} = \left(\frac{1+\sqrt{-7}}{2}\right)^m - \left(\frac{1-\sqrt{-7}}{2}\right)^m. \quad (176)$$

Then setting $\alpha = \frac{1+\sqrt{-7}}{2}$ and $\beta = \frac{1-\sqrt{-7}}{2}$, we get that

$$\alpha^m - \beta^m = \alpha - \beta. \quad (177)$$

We see that $1 + \beta^2 = -\alpha$, and therefore $\beta^2 | \alpha + 1$, hence

$$\alpha \equiv -1 (\mathrm{mod}\ \beta^2). \tag{178}$$

Then $\alpha^2 \equiv 1 (\mathrm{mod} \beta^2)$, so

$$\alpha^m = \alpha(\alpha^2)^{\frac{m-1}{2}} \equiv \alpha(1)^{\frac{m-1}{2}} \equiv \alpha (\mathrm{mod}\ \beta^2). \tag{179}$$

Therefore, since $\alpha^m = \alpha - \beta + \beta^m$ by assumption, and because $n > 3$ implies that $m > 1$, we have that

$$\alpha^m \equiv \alpha - \beta (\mathrm{mod}\ \beta^2) \tag{180}$$

(because if $m > 1$, then $\beta^2 | \beta^m$). But then $\alpha \equiv \alpha - \beta (\mathrm{mod}\, \beta^2)$, and therefore $\beta \equiv 0 (\mathrm{mod}\, \beta^2)$, hence $\beta^2 | \beta$, despite that β is not a unit in $Z(\mathbb{Q}(\sqrt{-7}))$. This contradiction leads us to conclude that our original assumption is invalid. *Id est*, we must have that $\alpha^m - \beta^m \neq \alpha - \beta = \sqrt{-7}$. Thus, $\alpha^m - \beta^m = -\sqrt{-7}$.

Now, by the binomial theorem, we have that

$$\alpha^m = \sum_{j=0}^{m} \binom{m}{j} \left(\frac{1}{2}\right)^{m-j} \left(\frac{\sqrt{-7}}{2}\right)^j \tag{181}$$

and

$$\beta^m = \sum_{j=0}^{m} \binom{m}{j} \left(\frac{1}{2}\right)^{m-j} \left(-\frac{\sqrt{-7}}{2}\right)^j. \tag{182}$$

Therefore,

$$-\sqrt{-7} = \alpha^m - \beta^m = \sum_{j=0}^{m} \binom{m}{j} \left(\frac{1}{2}\right)^{m-j} \left(\left(\frac{\sqrt{-7}}{2}\right)^j - \left(-\frac{\sqrt{-7}}{2}\right)^j \right), \tag{183}$$

and so, because cancellation is possible in an integral domain,

$$-2^m = \sum_{j=1}^{m} \binom{m}{j} (1-(-1)^j)(-7)^{\frac{j-1}{2}} \tag{184}$$

(note that the factor of $(1-(-1)^j)$ has the effect of nullifying terms with an even j). Expanding this by the binomial theorem,

$$\begin{aligned} -2^m = \binom{m}{1}(2) - \binom{m}{3}(2)7 + \binom{m}{5}(2)7^2 + ... + \binom{m}{m}(2)(-7)^{\frac{m-1}{2}} \\ \equiv 2m (\mathrm{mod}\ 7). \end{aligned} \tag{185}$$

Now we have that $-2^{m-1} \equiv m(\ \mathrm{mod}\ \ 7)$ for any m such that $n = m+2$ is a solution for $n > 3$. (We may cancel the 2 from both sides of the congruence because 2 and 7 are relatively prime, as given by a well-known theorem of number theory. For a formal statement and proof of this theorem, see pages 56-57 of reference [7].) Further, any solution for m must satisfy $-2^{m-1} \equiv m (\mathrm{mod}\ 7)$ if $n > 3$, since all of the statements that we have made thus far are reversible.

By Lemma 1.11, we have that $m - 1 = 6q + r$ for some $q, r \in \mathbb{Z}$, where $0 \leq r < 6$. Then

$$-2^{m-1} = -2^{6q+r} = -2^{6q}2^r = -(2^3)^{2q}2^r \equiv -2^r (\mathrm{mod}\ 7). \tag{186}$$

Now, since $m - 1 = 6q + r \equiv r (\mathrm{mod}\ 6)$, and $0 \leq r < 6$, there are six separate possibilities for $m - 1$. Consider the case in which either $m - 1 \equiv 0 (\mathrm{mod}\ 6)$ or $m - 1 \equiv 3 (\mathrm{mod}\ 6)$. Note that $-2^0 \equiv -1 \equiv -2^3 (\mathrm{mod}\ 7)$. Then we deduce that

$-2^r \equiv -1 (\mathrm{mod}\ 7)$. Since, as we have shown, $-2^{m-1} \equiv -2^r (\mathrm{mod}\ 7)$, we have, from the previous paragraph, that $m \equiv -1 (\mathrm{mod}\ 7)$. Suppose for contradiction that $m - 1 \equiv 3 (\mathrm{mod}\ 6)$. Then $m \equiv 4 (\mathrm{mod}\ 6)$, and therefore $m = 4 + 6q_1$ for some $q_1 \in \mathbb{Z}$. But then m would be even, despite that $m = n - 2$, with n being odd. Therefore, $m \equiv 1 (\mathrm{mod}\ 6)$. Then the simultaneity of this congruence as well as the congruence $m \equiv -1 \equiv 6 (\mathrm{mod}\ 7)$ implies that

$$m \equiv 13 (\mathrm{mod}\ 42) \tag{187}$$

is the only possibility modulo 42 if $m - 1 \equiv 0 (\mathrm{mod}\ 6)$ or $m - 1 \equiv 3 (\mathrm{mod}\ 6)$, as can be shown by the Chinese Remainder Theorem. Using very similar arguments, we find that that

$$m \equiv 3 (\mathrm{mod}\ 42) \tag{188}$$

or

$$m \equiv 5 (\mathrm{mod}\ 42) \tag{189}$$

are the only other possibilities for m modulo 42.

We have shown that the only solutions for m are $m \equiv 3, 5$ or $13 (\mathrm{mod}\ 42)$. We now claim that $m = 3$, $m = 5$, or $m = 13$ are the only possible solutions for m given the current constraints that $n > 3$. First of all, $n = m + 2 = 3 + 2 = 5$ satisfies the Ramanujan equation provided that $x = \pm 5$. If $n = m + 2 = 5 + 2 = 7$, n satisfies the equation provided that $x = \pm 11$. If $n = m + 2 = 13 + 2 = 15$, n satisfies the equation provided that $x = \pm 181$. To verify that these values of x are the unique solutions for these choices of n is a straightforward matter of computation. Let $m_1 \equiv m (\mathrm{mod}\ 42)$ for some $m_1 \in \mathbb{Z}$, and suppose that $m_1 + 2$

and $m+2$ both satisfy n in the Ramanujan equation. Assume for contradiction that $m_1 \neq m$, and suppose that $m_1 > m$.

Define $l \in \mathbb{Z}$ so that 7^l is the highest power that divides $m_1 - m$. We know that $l \geq 1$, since $7|42$ and $42|m_1 - m$. Then

$$m_1 - m = 7^l 6k, \tag{190}$$

where $k \in \mathbb{Z}$ and $7 \nmid k$. Now, define $\theta = \left(\frac{1+\sqrt{-7}}{2}\right)^{m_1-m}$. With this definition, it follows easily that

$$2^{m_1-m}\theta = (1+\sqrt{-7})^{m_1-m}. \tag{191}$$

We claim that $\theta \equiv (1+\sqrt{-7})^{m_1-m} (\mathrm{mod}\ 7^{l+1})$. Note that, since the order of $[2]$ in $\mathbb{Z}/7\mathbb{Z}$ is 3, $2^{6k} \equiv 1 (\mathrm{mod}\ 7)$. It follows that $2^{6k} = 1 + 7k_1$, for some $k_1 \in \mathbb{Z}$. Therefore,

$$2^{m_1-m} = (1+7k_1)^{(7^l)}. \tag{192}$$

By Lemma 4.7, it follows that

$$2^{m_1-m} \equiv 1 + 7^l(-k_1)(-7) \equiv 1 + 7^{l+1}k_1 \equiv 1 (\mathrm{mod}\ 7^{l+1}). \tag{193}$$

This leads us to the fact that $\theta \equiv 2^{m_1-m}\theta \equiv (1+\sqrt{-7})^{m_1-m} (\mathrm{mod}\ 7^{l+1})$.

We will now show that $\theta \equiv 1 + (m_1 - m)\sqrt{-7} (\mathrm{mod}\ 7^{l+1})$. From Lemma 4.7, we have that $(1+\sqrt{-7})^{(7^l)} \equiv 1 + 7^l\sqrt{-7} (\mathrm{mod}\ 7^{l+1})$. Then

$$(1+\sqrt{-7})^{m_1-m} \equiv \left((1+\sqrt{-7})^{(7^l)}\right)^{6k} \equiv (1+7^l\sqrt{-7})^{\frac{m_1-m}{7^l}} (\mathrm{mod}\ 7^{l+1}). \tag{194}$$

By the binomial theorem, we now have that

$$(1+\sqrt{-7})^{m_1-m} \equiv 1 + \frac{m_1-m}{7^l}7^l\sqrt{-7} + \binom{6k}{2}(7^l\sqrt{-7})^2 + ... \\ + \binom{6k}{6k}(7^l\sqrt{-7})^{6k} (\text{mod } 7^{l+1}). \quad (195)$$

All of the terms after term 1 (counting from 0) have factors of 7^{l+1} in them that are not part of the binomial coefficients. As for term 1, the factors of 7^l may safely cancel. We are left to conclude that

$$\theta \equiv (1+\sqrt{-7})^{m_1-m} \equiv 1 + (m_1-m)\sqrt{-7} (\text{mod } 7^{l+1}). \quad (196)$$

We will now show that $\alpha^{m_1} \equiv \alpha^m + 4^m(m_1-m)\sqrt{-7} (\text{mod } 7^{l+1})$. First of all, since $2^m\alpha^m = (1+\sqrt{-7})^m$ by definition of α, it is not hard to show, using the binomial theorem, that

$$2^m\alpha^m \equiv 1 + m\sqrt{-7} (\text{mod } 7). \quad (197)$$

Multplying both sides of this equation by 4^m, we see that

$$8^m\alpha^m \equiv \alpha^m \equiv 4^m(1+m\sqrt{-7}) (\text{mod } 7). \quad (198)$$

Thus,

$$\alpha^m = 4^m(1+m\sqrt{-7}) + 7z, \quad (199)$$

for some $z \in \mathbb{Z}$. Now, since $\alpha^{m_1} = \alpha^m x$, it is true that

$$\alpha^{m_1} \equiv \alpha^m(1 + (m_1 - m)\sqrt{-7}) \equiv \alpha^m + \alpha^m(m_1 - m)\sqrt{-7} (\text{mod } 7^{l+1}). \quad (200)$$

In that case, based on what we have just derived about α^m,

$$\begin{aligned}
\alpha^{m_1} &\equiv \alpha^m + (4^m(1 + m\sqrt{-7}) + 7z)(m_1 - m)\sqrt{-7} (\text{mod } 7^{l+1}) \\
&\equiv \alpha^m + 4^m(m_1 - m)\sqrt{-7} + 4^m m\sqrt{-7}(m_1 - m)\sqrt{-7} + 7z(m_1 - m)\sqrt{-7} \\
&\equiv \alpha^m + 4^m(m_1 - m)\sqrt{-7} + 4^m m(-7)7^l 6k + 7z7^l 6k\sqrt{-7} \\
&\equiv \alpha^m + 4^m(m_1 - m)\sqrt{-7} (\text{mod } 7^{l+1}). \quad (201)
\end{aligned}$$

Likewise,

$$\beta^{m_1} \equiv \beta^m - 4^m(m_1 - m)\sqrt{-7} (\text{mod } 7^{l+1}). \quad (202)$$

But in order for m_1 and m to satisfy the Ramanujan equation for $n = m_1 + 2$ and $n = m + 2$, we must have that $\alpha^m - \beta^m = \alpha^{m_1} - \beta^{m_1} = -\sqrt{-7}$. Thus, $\alpha^{m_1} - \beta^{m_1} - (\alpha^m - \beta^m) = 0$. Therefore, since

$$\alpha^{m_1} - \beta^{m_1} \equiv \alpha^m - \beta^m + 2(4^m)(m_1 - m)\sqrt{-7} (\text{mod } 7^{l+1}), \quad (203)$$

we have that

$$2(4^m)(m_1 - m)\sqrt{-7} \equiv \alpha^{m_1} - \beta^{m_1} - (\alpha^m - \beta^m) \equiv 0 (\text{mod } 7^{l+1}). \quad (204)$$

Now, since $m_1 - m = 7^l 6k$, we have that $0 \equiv 2(4^m)7^l 6k\sqrt{-7} (\mathrm{mod}\ 7^{l+1})$, and so

$$12k4^m\sqrt{-7} \equiv 0 (\mathrm{mod}\ 7). \tag{205}$$

We multiply both sides by $3(2^m)$ to get that

$$36(8^m)k\sqrt{-7} \equiv k\sqrt{-7} \equiv 0 (\mathrm{mod}\ 7). \tag{206}$$

Thus, $k\sqrt{-7} = 7\rho$, for some $\rho \in Z(\mathbb{Q}(\sqrt{-7}))$. Because it is an integral domain, we may cancel $\sqrt{-7}$ from both sides, arriving at

$$k = -\rho\sqrt{-7}. \tag{207}$$

By Theorem 3.10, $\rho = x_1 + y_1\left(\frac{1+\sqrt{-7}}{2}\right)$ for some $x_1, y_1 \in \mathbb{Z}$. Then

$$k = -\frac{2x_1 + y_1}{2}\sqrt{-7} + \frac{7y_1}{2} = \frac{7y_1}{2}, \tag{208}$$

as $\frac{-2x_1 - y_1}{2} = 0$ must be true in order for k to be a rational integer. Thus, $2k = 7y_1$. But this equation is true in the rational integers. In the rational integers, 7 is prime, so $7|2k$ implies that $7|2$ or $7|k$. The first is obviously impossible. The second is also impossible, because we have defined k to be indivisible by 7. Thus, we reach a contradiction that leads us to conclude that $m_1 = m$. Therefore, the theorem is proven. □

Note how purely number theoretic the statement of the Ramanujan-Nagell Theorem appears. Much like Fermat's Last Theorem, it would seem at first glance that

the theorem has absolutely nothing to do with operations on sets, and may even appear to be obvious. However, both of these presumptions are quite wrong, both for the Ramanujan-Nagell Theorem and for Fermat's infamous Last Theorem.

After following the maddeningly winding path that the proof of the Ramanujan-Nagell Theorem takes, the reader may even find him- or herself confused as to how unique factorization was primarily used in the proof. First of all, if we had not been aware of the fact that $Z(\mathbb{Q}(\sqrt{d}))$ was a Unique Factorization Domain as a result of Theorem 3.16, then we would never have been able to use Proposition 1.27, and thus, we would not have been able to come to the conclusion that $\pm\sqrt{-7} = \left(\frac{1+\sqrt{-7}}{2}\right)^m - \left(\frac{1-\sqrt{-7}}{2}\right)^m$. This would have made it impossible for us to move on. Further, purely number theoretic methods would have forbidden our extensive use of the properties of $\sqrt{-7}$. Therefore, an algebraic approach to this proof is highly desirable.

References

[1] J. A. Beachy and W. D. Blair, *Abstract Algebra with a Concrete Introduction*, Prentice-Hall, Inc., 1990.

[2] J. W. Brown and R. V. Churchill, *Complex Variables and Applications*, eighth edition, The McGraw-Hill Companies, Inc., 2009.

[3] D. M. Burton, *Elementary Number Theory*, sixth edition, The McGraw-Hill Companies, Inc., 2007.

[4] D. S. Dummit and R. M. Foote, *Abstract Algebra*, third edition, John Wiley and Sons Inc., 2004.

[5] I. N. Herstein, *Abstract Algebra*, third edition, John Wiley and Sons, Inc., 1999.

[6] T. W. Hungerford, *Abstract Algebra, an Introduction*, second edition, Saunders College Publishing, 1997, 1990.

[7] H. M. Stark, *An Introduction to Number Theory*, MIT Press, 1989.

[8] I. N. Stewart and D. O. Tall, *Algebraic Number Theory*, second edition, Cambridge University Press, 1987.