GAUSS'S LEMMA FOR NUMBER FIELDS

ARTURO MAGIDIN AND DAVID MCKINNON

1. Introduction.

This note arose when the following question was asked on the newsgroup `sci.math`:

Question 1.1. *Can every polynomial with integer coefficients be factored into (not necessarily monic) linear terms, each with algebraic integer coefficients?*

The answer is yes, and follows from a version of Gauss's lemma applied to number fields. Gauss's lemma plays an important role in the study of unique factorization, and it was a failure of unique factorization that led to the development of the theory of algebraic integers. These developments were the basis of algebraic number theory, and also of much of ring and module theory. We take the opportunity afforded by this problem to discuss some of these historical developments, on the (at times flimsy) excuse of introducing the necessary notions for the proofs. It will take us some time to get to the answer to Question 1.1, so we ask for the reader's patience.

To make for easier reading, we give names to many of the results. While some of these are standard, others are the inventions of the authors.

The paper is organized as follows: First we discuss some of the history of unique factorization. In sections 3 and 4 we discuss how Euler and Lamé ran afoul of unique factorization when dealing with Fermat's Last Theorem. Section 5 relates Kummer's own struggle with the failure of unique factorization, and his solution for cyclotomic fields. Section 6 deals with Kronecker's and Dedekind's extension of Kummer's work, setting up the stage for our treatment of Question 1.1. In section 7 we recall the basic notions associated with number fields that we need, and proceed in section 8 to prove Gauss's lemma and a version of its corollaries for number fields, providing an answer to Question 1.1. In section 9 we discuss some related ring theoretic notions and provide an alternative approach to answering our question. Finally, in section 10 we consider the question in the setting of function fields (which

are closely related to number fields), and give an example to show that Question 1.1 has a negative answer there.

2. Unique Factorization and Gauss's Lemma.

Gauss was the first to give a proof of the following fact [9, art. 16]:

Theorem 2.1 (**Fundamental Theorem of Arithmetic**). *Every positive integer can be factored uniquely into a product of prime numbers.*

The proof proceeds in two steps:

Step 1. First, we show that every positive integer can be written as a product of primes in at least one way. This argument relies on a kind of "finiteness" that the positive integers exhibit, namely, that there can be no infinite strictly decreasing sequence of positive integers. If $n = 1$, it equals, by definition, the empty product. Otherwise, it is either a prime, in which case we can write $n = n$, or it is not a prime, in which case we can write it as a product of two strictly smaller positive integers. We can proceed, by applying the same argument to each of those factors, or alternatively, we can appeal to an induction hypothesis to assert that each must be a finite product of primes. We conclude that n itself is a product of primes. (In fact, Gauss skips this step in his proof, saying merely that "it is clear from elementary considerations that any composite number can be [factored] into prime factors.")

Step 2. Once we know that the number can be written as a product of primes in at least one way, we prove the uniqueness by using the "prime divisor property"[8, Prop. 30]:

Proposition 2.2 (**Prime Divisor Property**). *If p is a prime number and p divides the product ab of two integers a and b, then p divides a or p divides b.*

To prove the uniqueness of the factorization, we suppose that

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

are two factorizations of n into primes. From the prime divisor property, it follows that, since p_1 divides the product $q_1 \cdots q_s$, it divides at least one of the q_i. Because q_i is prime, we must have $p_1 = q_i$, at which point we may cancel them and apply induction.

Although the prime divisor property goes back at least to the time of Euclid, unique factorization had not been explicitly formulated (nor proved) until Gauss did so. Gauss points out in the first paragraph of his proof that "it is often wrongly taken for granted" that factorization is unique. In fact, Euclid himself implicitly invokes this fact:

in his discussion of Pythagorean triples a key step uses the fact that, if the product of two relatively prime integers is a square, then each of the integers must itself be a square, a result which requires unique factorization.

An argument following the same broad outline as the proof of the fundamental theorem of arithmetic is used to prove that the ring of polynomials $\mathbb{Q}[x]$ is also a unique factorization domain (UFD), and in fact that $F[x]$ is a UFD whenever F is a field. First, we replace the notion of "prime" with that of "irreducible":

Definition 2.3. Let R be a commutative ring with identity. An element x of R is *irreducible* if (1) x is not a unit (i.e., does not have a multiplicative inverse) and (2) whenever x factors in R as $x = yz$, either y is a unit or z is a unit (i.e., either y or z have a multiplicative inverse).

The finiteness condition necessary to guarantee that every polynomial $f(x)$ in $F[x]$ is a product of irreducibles follows by looking at the degree function, while the fact that every irreducible satisfies the prime divisor property is a corollary of the division algorithm:

Theorem 2.4 (**Division Algorithm for Polynomials over a Field**)**.** *Let F be a field, and let $a(x)$ and $b(x) \neq 0$ be polynomials with coefficients in F. Then there exist unique polynomials $q(x)$ and $r(x)$ in $F[x]$ (the "quotient" and the "remainder", respectively) such that*

$$a(x) = b(x)q(x) + r(x),$$

where either $r = 0$ or $\deg(r) < \deg(b)$.

(This, of course, establishes that $F[x]$ is a Euclidean domain, a condition that is well known to imply unique factorization.)

In [9] Gauss is mostly concerned with integers, however, and if we restrict ourselves to polynomials with integer coefficients, the division algorithm no longer holds. For example, if $a(x) = 3x + 2$ and $b(x) = 2x + 3$, we cannot find $q(x)$ and $r(x) \in \mathbb{Z}[x]$ satisfying the conditions stipulated by the division algorithm. So unique factorization in $\mathbb{Z}[x]$ is harder to establish.

The degree function, together with the fundamental theorem of arithmetic, shows that the first step in the proof of unique factorization can be achieved: every polynomial in $\mathbb{Z}[x]$ can be written as a product of irreducibles, where an irreducible is either a prime integer or a nonconstant irreducible polynomial $f(x)$ in $\mathbb{Z}[x]$ (i.e., $f(x)$ cannot be written as the product of two nonconstant polynomials) whose coefficients have no common factor other than 1 and -1 (this condition is known

as *primitivity*, and we will have much more to say about it). It is the prime divisor property that seems a bit more difficult. That is where Gauss's lemma comes in.

What we would like to do is make use of the already established facts that both $\mathbb{Z}$ and $\mathbb{Q}[x]$ are unique factorization domains. First, we want to show that the primes from $\mathbb{Z}$ remain irreducible in $\mathbb{Z}[x]$, which is of course trivial. It is also clear that if $f(x)$ in $\mathbb{Z}[x]$ is irreducible when we consider it in $\mathbb{Q}[x]$, then it will be the product of an integer times a polynomial that is irreducible in $\mathbb{Z}[x]$ (since all nonzero constants have multiplicative inverses in $\mathbb{Q}[x]$, they do not affect irreducibility; this is not the case in $\mathbb{Z}[x]$, where the only constants with multiplicative inverses are 1 and -1). If we separate out the integer and factor it into primes, the result will be a factorization into irreducibles in $\mathbb{Z}[x]$ for $f(x)$.

The immediate difficulty, however, lies in the converse: is it possible that there is a polynomial $f(x)$ in $\mathbb{Z}[x]$ that can be factored in $\mathbb{Q}[x]$, but not in $\mathbb{Z}[x]$? And the second issue, which we can perhaps already see approaching us, is whether the integer primes are actually primes in $\mathbb{Z}[x]$: if $p|g(x)h(x)$, does it follow that $p|g(x)$ or $p|h(x)$ in $\mathbb{Z}[x]$?

As it happens, both difficulties can be dealt with at the same time. We deal with the latter first by establishing the contrapositive, which is nothing other than Gauss's lemma [9, art. 42]. We then prove unique factorization by first factoring out any constants that we can, using $\mathbb{Q}[x]$ as a "stepping stone" to factor the polynomial part, and then "lifting" this factorization back to $\mathbb{Z}[x]$.

Definition 2.5. Let $f(x) = a_n x^n + \cdots + a_1 x + a_0$ be a nonzero member of $\mathbb{Z}[x]$. The *content* $\text{cont}(f)$ of f is the greatest common divisor of $a_0, \ldots, a_n$.

Definition 2.6. A polynomial $f(x)$ in $\mathbb{Z}[x]$ is *primitive* if $\text{cont}(f) = 1$.

Theorem 2.7 (**Gauss's Lemma**)**.** *The product of primitive polynomials is itself primitive.*

Proof. Let $f(x)$ and $g(x)$ be primitive polynomials in $\mathbb{Z}[x]$, and let $h(x) = f(x)g(x)$. Write

$$\begin{aligned} f(x) &= a_n x^n + \cdots + a_1 x + a_0, \\ g(x) &= b_m x^m + \cdots + b_1 x + b_0, \\ h(x) &= c_{n+m} x^{n+m} + \cdots + c_1 x + c_0. \end{aligned}$$

To prove that $h(x)$ is primitive, consider an arbitrary prime p. We must show p does not divide all the c_i.

Let i_0 be the smallest index such that $p \not| a_{i_0}$; such an index exists since $f(x)$ is primitive. Likewise, let j_0 be the smallest index such that $p \not| b_{j_0}$. We consider $c_{i_0+j_0}$. We have:

$$\begin{aligned} c_{i_0+j_0} &= a_{i_0+j_0}b_0 + a_{i_0+j_0-1}b_1 + \cdots + a_{i_0+1}b_{j_0-1} \\ &\quad + a_{i_0}b_{j_0} \\ &\quad + a_{i_0-1}b_{j_0+1} + \cdots + a_1 b_{i_0+j_0-1} + a_0 b_{i_0+j_0}. \end{aligned}$$

Each summand in the first expression is a multiple of p by the choice of j_0, while each summand in the third expression is a multiple of p by the choice of i_0. Furthermore, $a_{i_0}b_{j_0}$ is not a multiple of p, ensuring that p cannot divide $c_{i_0+j_0}$. □

The argument given also shows that if $p \in \mathbb{Z}$ is a prime and divides the product, then it must divide one of the two factors, showing that integer primes are still primes in $\mathbb{Z}[x]$. Two more auxiliary results demonstrate that $\mathbb{Z}[x]$ is a UFD. They establish the connection between $\mathbb{Q}[x]$ and $\mathbb{Z}[x]$ that allows us to "lift" factorizations from the former to the latter. The first uses, implicitly, unique factorization in $\mathbb{Z}$ in order to express an element of $\mathbb{Q}$ as the quotient of two relatively prime integers. The second uses Gauss's lemma explicitly.

Lemma 2.8 (**Factoring Out the Content**). *If $f(x)$ is a nonzero polynomial in $\mathbb{Q}[x]$, then there exist c_f in $\mathbb{Q}$ and a primitive polynomial $f^*(x)$ in $\mathbb{Z}[x]$ such that $f(x) = c_f f^*(x)$. Moreover, c_f and $f^*(x)$ are unique up to sign.*

Proof. Write

$$f(x) = \left(\frac{a_n}{b_n}\right) x^n + \cdots + \left(\frac{a_0}{b_0}\right),$$

where a_i and b_i are relatively prime integers for $i = 0, 1, \ldots, n$. Multiplying by $b_0 \cdots b_n$ to clear denominators, we obtain:

$$(b_0 \cdots b_n) f(x) = a_n x^n + \cdots + a_0.$$

Let $g(x) = a_n x^n + \cdots + a_0$, and let $c = \text{cont}(g)$. Then $g(x)$ can be written as

$$g(x) = c g^*(x),$$

where $g^*(x)$ is a primitive polynomial. Therefore,

$$(b_0 \cdots b_n) f(x) = c g^*(x).$$

Letting $c_f = c/(b_0 \cdots b_n)$ and $f^*(x) = g^*(x)$ proves the existence assertion.

Uniqueness will follow provided we can prove the following: if $g^*(x) = c \cdot h^*(x)$, where $g^*(x)$ and $h^*(x)$ in $\mathbb{Z}[x]$ are both primitive and c is rational, then $c = \pm 1$. Write $c = u/v$, with relatively prime integers u and v. Then $vg^*(x) = uh^*(x)$. Each coefficient of $uh^*(x)$ must be a multiple of v. As $\gcd(u, v) = 1$ and $h^*(x)$ is primitive, this forces $v = \pm 1$. Similarly, $u = \pm 1$, so $c = \pm 1$, as desired. □

Theorem 2.9 (**Lifting the Factorization**). *Let $f(x)$ belong to $\mathbb{Z}[x]$, and suppose that $f(x)$ admits the factorization $f(x) = G(x)H(x)$ in $\mathbb{Q}[x]$. Then there exist polynomials $g(x)$ and $h(x)$ in $\mathbb{Z}[x]$ such that $g(x)$ is a rational multiple of $G(x)$, $h(x)$ is a rational multiple of $H(x)$, and $f(x) = g(x)h(x)$.*

Proof. We can factor out the content to write:

$$f(x) = c_f f^*(x),\ G(x) = c_g g^*(x),\ H(x) = c_h h^*(x).$$

Then we have $c_f f^*(x) = (c_g c_h) g^*(x) h^*(x)$. By Gauss's lemma, the product $g^*(x)h^*(x)$ is primitive. The uniqueness clause in Lemma 2.8 thus gives $c_f = c_g c_h$ or $c_f = -c_g c_h$. In either case, $c_g c_h$ is an integer, so we let $g(x) = (c_g c_h) g^*(x)$ and $h(x) = h^*(x)$ to complete the proof. □

Corollary 2.10. *The ring $\mathbb{Z}[x]$ is a UFD. The irreducibles in $\mathbb{Z}[x]$ are the integer primes and the primitive polynomials that are irreducible when considered as polynomials in $\mathbb{Q}[x]$.*

It is not hard to see that the same sequence of results holds if we replace $\mathbb{Z}$ with an arbitrary UFD R and $\mathbb{Q}$ with the field of fractions of R. We must, however, modify the lemma on factoring out the content (Lemma 2.8) so that uniqueness means "up to units of R" rather than "up to sign." Since any witness to the fact that a ring R is not a UFD will confirm that $R[x]$ is not either, we obtain:

Theorem 2.11. *Let R be a ring. Then R is a UFD if and only if $R[x]$ is a UFD.*

Gauss employs his lemma in the study of cyclotomic polynomials (see, for example, [9, art. 341]), among other places. The study of cyclotomic polynomials is used in turn to prove Gauss's celebrated result that the construction of a regular N-gon using only compass and straightedge can be achieved if and only if the odd prime factors of N are all distinct *Fermat primes*, meaning primes of the form $2^{2^m} + 1$.

3. Unique factorization and Euler's proof of Fermat's Last Theorem for $n = 3$.

Unique factorization had already made a covert appearance in the study of quadratic forms by Fermat, though it was hard to recognize

it in that guise. We refer the reader to John Stillwell's introduction in [5].

Gauss was very interested in quadratic forms and dedicates section 5 of the *Disquisitiones Arithmeticae* [9] to their study. In the English edition, it comprises over 260 pages, compared with only 107 for sections 1 through 4. It is likely that it was through the study of quadratic forms and their connection to unique factorization that Gauss recognized the unstated assumption of unique factorization of integers, and so was led to state explicitly the fundamental theorem of arithmetic.

Perhaps the first overt appearance occurs in the work of Euler, who ran afoul of unique factorization in his attempt to prove Fermat's Last Theorem for $n = 3$. The idea behind Euler's attempt was similar to Fermat's own proof for the case $n = 4$: to establish an infinite descent, by proving that the existence of a nontrivial solution to $x^3 + y^3 = z^3$, with x, y, and z positive, pairwise coprime integers, necessarily leads to the existence of another solution $x'^3 + y'^3 = z'^3$ with *smaller* integers (i.e., $z' < z$). Since there can be no infinite descending sequence of positive integers, no solution could exist in the first place. Note that the finiteness of the positive integers we discussed earlier is at play here again.

For $n = 3$, Euler starts with $x^3 + y^3 = z^3$, with $\gcd(x, y, z) = 1$. If both x and y are odd, then $x + y$ and $x - y$ are both even, say $2p$ and $2q$, respectively, so $x = p + q$, $y = p - q$, and

$$x^3 + y^3 = (x + y)(x^2 - xy + y^2) = 2p(p^2 + 3q^2).$$

Since x and y are odd and coprime, p and q must be of opposite parities and coprime. And as $x^3 + y^3 = z^3$, $2p(p^2 + 3q^2)$ must be a cube. A similar argument yields the same conclusion if z is odd and one of x or y is even.

At this point we want to show that both $2p$ and $p^2 + 3q^2$ are cubes. If $3 \not| p$, this follows easily by noting that $2p$ and $p^2 + 3q^2$ are relatively prime and appealing to unique factorization; if $3|p$, then we must write $p = 3s$, and then rewrite

$$2p(p^2 + 3q^2) = 3^2 \cdot 2s(3s^2 + q^2),$$

from which we infer that $3^2 \cdot 2s$ and $3s^2 + q^2$ are relatively prime. Each must therefore be a cube (see [7, sec. 2.2] for the details).

Euler noted that one way in which both $2p$ and $p^2 + 3q^2$ are cubes is for p and q to have the forms

$$p = a(a - 3b)(a + 3b), \qquad q = 3b(a - b)(a + b) \tag{3.1}$$

(a similar expression is found for s and q when p is a multiple of 3). If this is indeed the case, then a and b must be relatively prime, because p and q are relatively prime, and must have opposite parities. From here one shows that $2a$, $a-3b$, and $a+3b$ must be pairwise coprime. Since $2p = 2a(a-3b)(a+3b)$ is a cube, each of $2a$, $a-3b$ and $a+3b$ must be a cube. Then $(a-3b)+(a+3b) = 2a$ gives a new solution to the Fermat equation, one with $2a < z^3$, setting up the infinite descent and thus proving the result. A similar argument is used when $3|p$.

At this point, of course, we need to show that the only way for $2p$ and p^2+3q^2 to be both cubes is for p and q to be expressible as in (3.1). It is here that the argument presented by Euler fails (he had, however, other results on quadratic forms that he could have used to establish this claim about p and q).

Euler factors $p^2+3q^2 = (p+q\sqrt{-3})(p-q\sqrt{-3})$ and proceeds to work in $\mathbb{Z}[\sqrt{-3}]$. Since

$$\begin{aligned}(p+q\sqrt{-3})+(p-q\sqrt{-3}) &= 2p\\ (p+q\sqrt{-3})-(p-q\sqrt{-3}) &= 2q\sqrt{-3},\end{aligned}$$

any common divisor of $(p+q\sqrt{-3})$ and $(p-q\sqrt{-3})$ would be a divisor of $2p$ and of $2q\sqrt{-3}$. One can show that both 2 and $\sqrt{-3}$ are irreducible in $\mathbb{Z}[\sqrt{-3}]$ (see the argument to follow). From the fact that p and q have opposite parities it follows that 2 does not divide $p+q\sqrt{-3}$, and from the fact that $3 \not| p$ one deduces that $\sqrt{-3}$ does not divide $p+q\sqrt{-3}$ either. Accordingly, any common divisor of $p+q\sqrt{-3}$ and $p-q\sqrt{-3}$ must in fact be a common divisor of both p and q, which are relatively prime. Hence $p+q\sqrt{-3}$ and $p-q\sqrt{-3}$ have no common divisors in $\mathbb{Z}[\sqrt{-3}]$ other than 1 and -1. Since their product is a cube, Euler concludes that each must be a cube, so in fact we have:

$$\begin{aligned}p+q\sqrt{-3} &= (a+b\sqrt{-3})^3\\ p-q\sqrt{-3} &= (a-b\sqrt{-3})^3\end{aligned}$$

for some integers a and b. It now follows that

$$\begin{aligned}p+q\sqrt{-3} &= a^3+3a^2b\sqrt{-3}-9ab^2-3b^3\sqrt{-3}\\ &= (a^3-9ab^2)+(3a^2b-3b^3)\sqrt{-3},\end{aligned}$$

from which the desired equations (3.1) follow.

The problem, of course, is that hidden in that argument is an assumption of unique factorization: we know that $p-q\sqrt{-3}$ and $p+q\sqrt{-3}$ have no common divisors in $\mathbb{Z}[\sqrt{-3}]$ (other than 1 and -1) and that their product is a cube. *If* we have unique factorization into irreducibles in $\mathbb{Z}[\sqrt{-3}]$, then we are able to conclude that each factor must

itself be a cube. But in fact we do not have uniqueness:

$$(1+\sqrt{-3})(1-\sqrt{-3}) = 4 = (2)(2),$$

and each of the numbers 2, $1+\sqrt{-3}$, and $1-\sqrt{-3}$ is irreducible in $\mathbb{Z}[\sqrt{-3}]$, as we demonstrate shortly. Thus, Euler's argument breaks down.

Euler must have realized that there was something wrong with the argument [7, sec. 2.3], for in a later portion of his *Algebra*, he noted that his methods would indicate that $2x^2 - 5$ cannot be a cube, even though $2(4)^2 - 5 = 27 = 3^3$. Euler attributed the difficulty to the minus sign in the equation, however, and so apparently did not feel that his argument for the $n = 3$ case of Fermat's Last Theorem was in danger.

This is probably a good place to establish the asserted irreducibility of 2, $1+\sqrt{-3}$, and $1-\sqrt{-3}$ in $\mathbb{Z}[\sqrt{-3}]$. A central role in the proof is played by a "norm" function defined from $\mathbb{Z}[\sqrt{-3}]$ to $\mathbb{Z}$. This function allows us to translate certain divisibility questions concerning $\mathbb{Z}[\sqrt{-3}]$ to questions in the more familiar setting of $\mathbb{Z}$.

Each element α of $\mathbb{Z}[\sqrt{-3}]$ can be written uniquely as $\alpha = a + b\sqrt{-3}$ with a and b in $\mathbb{Z}$. We define a *norm* $N\colon \mathbb{Z}[\sqrt{-3}] \to \mathbb{Z}$ as follows:

$$N(a+b\sqrt{-3}) = (a+b\sqrt{-3})(a-b\sqrt{-3}) = a^2 + 3b^2,$$

that is, $N(a+b\sqrt{-3})$ is the product of all images of $a+b\sqrt{-3}$ under the distinct embeddings of $\mathbb{Q}(\sqrt{-3})$ (the field of fractions of $\mathbb{Z}[\sqrt{-3}]$) into $\mathbb{C}$.

Lemma 3.1 (**Properties of the Norm**). *The norm N satisfies:*

(1) *$N(\alpha\beta) = N(\alpha)N(\beta)$ for all α and β in $\mathbb{Z}[\sqrt{-3}]$.*
(2) *If $\alpha|\beta$ in $\mathbb{Z}[\sqrt{-3}]$, then $N(\alpha)|N(\beta)$ in $\mathbb{Z}$.*
(3) *The element α of $\mathbb{Z}[\sqrt{-3}]$ is a unit (i.e., has a multiplicative inverse) if and only if $N(\alpha) = 1$.*

Proof. Statements (2) and (3) follow directly from (1), which can be established through direct computation. □

The two most important points to observe right now are these: first, the properties of the norm provide the "finiteness" needed to accomplish at least the first step in a proof of unique factorization; namely, every element of $\mathbb{Z}[\sqrt{-3}]$ can be expressed as a product of irreducible elements. This is true because any proper divisor in $\mathbb{Z}[\sqrt{-3}]$ of an element α of $\mathbb{Z}[\sqrt{-3}]$ will necessarily have a norm that is a proper divisor of $N(\alpha)$ in $\mathbb{Z}$, ensuring that no infinite descent is possible. And, second, that property (2) gives a way to study divisibility in $\mathbb{Z}[\sqrt{-3}]$ by

referring to divisibility in $\mathbb{Z}$. Since the implication is not reversible, it is not a perfect translation, but even so it is extremely useful.

Consider the number $\sqrt{-3}$. Since $N(\sqrt{-3}) = 3$, it follows from the properties of the norm that $\sqrt{-3}$ is irreducible. Moreover, because

$$N(2) = N(1 + \sqrt{-3}) = N(1 - \sqrt{-3}) = 4,$$

each of 2, $1 + \sqrt{-3}$, and $1 - \sqrt{-3}$ must also be irreducible (a proper divisor of any of the three would have norm 2, but $2 = a^2 + 3b^2$ has no solution with a and b integers).

As the only units of $\mathbb{Z}[\sqrt{-3}]$ are 1 and -1, we see that even though $(2)(2) = (1 + \sqrt{-3})(1 - \sqrt{-3})$, the two factorizations are not related by multiplications by units. In other words, unique factorization does indeed fail in $\mathbb{Z}[\sqrt{-3}]$. This is what Euler failed to take into account.

4. Lamé and a general proof of Fermat's Last Theorem.

In 1847, G. Lamé announced to the Paris Academy that he had found a proof of Fermat's Last Theorem; our account is taken from [7, chap. 4]. His brief sketch consisted in factoring the equation $x^n + y^n = z^n$ as

$$z^n = x^n + y^n = (x + y)(x + \zeta y) \cdots (x + \zeta^{n-1} y),$$

where ζ is a primitive n-th root of unity, say

$$\zeta = \cos(2\pi/n) + i \sin(2\pi/n).$$

Lamé's idea was to obtain a contradiction by proving that each of the factors $(x + \zeta^i y)$ would necessarily be an nth power. He would prove this by showing either that no two of the factors had common divisors in $\mathbb{Z}[\zeta]$ other than units or that there was a factor m common to all n factors such that $(x+y)/m$, $(x+\zeta y)/m$, and so forth had the property that no two had common divisors other than units, and then use a similar argument. A sketch of this argument for a prime exponent p such that $p \nmid xyz$ can be found in [16, chap. 1, Exercises 16–28].

Lamé made a point of mentioning that he had come up with the idea after a casual conversation with Liouville a few months earlier. Liouville, for his part, took the floor immediately after Lamé and cast some doubts on the viability of the latter's program. He quickly pointed out the gap in the argument: to conclude that each factor was an nth power from the fact that there were no common divisors (other than units) of any two, he needed a property analogous to unique factorization for the elements of $\mathbb{Z}[\zeta]$, and this was by no means a given. Lamé acknowledged the gap, and in the following months attempted to fill it. It was not thought a hopeless task: two small cases had already been treated. The case $n = 4$ had been studied by Gauss, who

proved in 1831 that $\mathbb{Z}[i]$ is a UFD. Gauss had been led to study this ring through his interest in biquadratic reciprocity. The case $p = 3$ had been studied by Eisenstein during his analysis of cubic reciprocity, and he had demonstrated in 1844 that $\mathbb{Z}[(-1+\sqrt{-3})/2]$ is also a UFD.

However, although neither Liouville nor Lamé were aware, Kummer had three years earlier published a memoir [14] in which he had shown that the domains $\mathbb{Z}[\zeta]$ did not always enjoy unique factorization. Kummer communicated this to Liouville, who passed the news on to the Academy. Lamé then abandoned his attack on Fermat's Last Theorem, defeated by the failure of unique factorization. Kummer's study of the rings $\mathbb{Z}[\zeta]$ had also been fueled by a desire to obtain higher reciprocity laws, though it is often incorrectly attributed to an interest in proving Fermat's Last Theorem (his results could be used to present a variant of Lamé's argument; see the next section).

5. Kummer and Ideal Numbers.

Starting in 1837, Kummer had begun to study the arithmetic of certain cyclotomic fields, extensions of $\mathbb{Q}$ obtained by adjoining a primitive nth root of unity. Kummer studied divisibility in rings $\mathbb{Z}[\zeta_p]$, where ζ_p is a primitive pth root of unity for a prime p. He quickly found that these rings were not in general UFDs. It was only after several years of effort that he discovered a way to circumvent the difficulties, with the introduction of "ideal numbers."

We will borrow from Dedekind's exposition of ideal numbers in [5]. Dedekind explains the situation skillfully, but rather than consider the case of a cyclotomic field, he considers the similar situation in the much simpler ring $\mathbb{Z}[\sqrt{-5}]$. This ring has a norm analogous to the one we introduced earlier for $\mathbb{Z}[\sqrt{-3}]$. Here we define N by

$$N(a+b\sqrt{-5}) = (a+b\sqrt{-5})(a-b\sqrt{-5}) = a^2+5b^2,$$

the product of all images of $(a+b\sqrt{-5})$ under the different embeddings of $\mathbb{Q}(\sqrt{-5})$ (the field of fractions of $\mathbb{Z}[\sqrt{-5}]$) into $\mathbb{C}$. It is now easy to verify that this map also satisfies the properties in Lemma 3.1, so it can be used to establish the fact that every element in $\mathbb{Z}[\sqrt{-5}]$ can be written as a product of irreducibles.

When trying to proceed to uniqueness, however, we again run into a problem: not every irreducible has the prime divisor property. Consider, for example, the two factorizations of 6:

$$6 = 2 \times 3 = (1+\sqrt{-5}) \times (1-\sqrt{-5}).$$

Since $N(2) = 4$, any proper divisor of 2 would necessarily have norm equal to 2, but $a^2+5b^2 = 2$ has no solution for a and b integers, hence

no element can have norm 2. Thus 2 is irreducible, as are $1+\sqrt{-5}$ and $1-\sqrt{-5}$, both of norm 6. Since $a^2+5b^2=3$ also has no integer solution and $N(3)=9$, we also see that 3 is irreducible in $\mathbb{Z}[\sqrt{-5}]$. Furthermore, the two factorizations are patently distinct, and we cannot pass from one to the other by multiplication by suitable units (the only units of $\mathbb{Z}[\sqrt{-5}]$ again being 1 and -1). Thus, $\mathbb{Z}[\sqrt{-5}]$ is not a UFD.

However, Kummer realized that one can tell a lot about the prime factorization of an integer without actually having to factor it into primes: how it behaves with respect to divisibility can help tell the complete story. For example, we already know that we can tell that an integer p greater than 1 is a prime by noting that it has the prime divisor property. For a slightly more complex result, consider the following two conditions:

(i) If n divides a^2b^2, then either n divides a^2 or n divides b^2.
(ii) There exists an integer m such that n does not divide m, but n divides m^2.

An integer n larger than 1 that satisfies condition (i) must be either a prime, or the square of a prime. An integer that satisfies condition (ii) must be divisible by the square of a prime. If we could show that an integer n satisfies *both* (i) and (ii), then it would follow immediately that n is the square of a prime.

Consider the number 2 in $\mathbb{Z}[\sqrt{-5}]$. Clearly, 2 divides a number $a+b\sqrt{-5}$ if and only if both a and b are even integers. The square of $a+b\sqrt{-5}$ is given by

$$(a+b\sqrt{-5})^2=(a^2-5b^2)+(2ab)\sqrt{-5}.$$

Thus, 2 divides $(a+b\sqrt{-5})^2$ if and only if a^2-5b^2 is even. This occurs exactly when a and b have the same parity. If a and b are both odd, then 2 divides $(a+b\sqrt{-5})^2$, despite the fact that it does not divide $a+b\sqrt{-5}$. For example, 2 divides

$$(1+\sqrt{-5})^2=-4+2\sqrt{-5}=2(-2+\sqrt{-5})$$

yet it does not divide $1+\sqrt{-5}$. So 2 satisfies condition (ii).

Note next that if 2 divides the product of two squares, say $(a+b\sqrt{-5})^2$ and $(c+d\sqrt{-5})^2$, then it must divide

$$\Big((a^2-5b^2)(c^2-5d^2)-20abcd\Big)+2\Big(ab(c^2-5d^2)+cd(a^2-5b^2)\Big)\sqrt{-5},$$

whence $(a^2-5b^2)(c^2-5d^2)$ must be even. This implies that at least one of the factors is even, so 2 divides at least one of $(a+b\sqrt{-5})^2$ and $(c+d\sqrt{-5})^2$. Hence 2 satisfies conditions (i) and (ii), and by all rights should be called the "square of a prime."

In fact, one can show that with regards to all divisibility properties in $\mathbb{Z}[\sqrt{-5}]$, the number 2 behaves as if it were the square of a prime. But of course, we know there is no such prime in $\mathbb{Z}[\sqrt{-5}]$. So we introduce an "ideal prime number" α with the property that $\alpha^2 = 2$. Here we are using *ideal* in a sense similar to that found in the dictionary: *"existing as a mental image or in fancy or imagination only"* (*Webster's Ninth New College Dictionary*). We say that a number $a + b\sqrt{-5}$ is divisible by the ideal prime α if and only if its square is divisible by the number 2. More generally, α^k is the highest power of α that divides $a + b\sqrt{-5}$ if and only if 2^k is the highest power of 2 that divides $(a + b\sqrt{-5})^2$.

A similar process leads to the conclusion that the irreducible 3 behaves in all respects as the product of two distinct ideal prime numbers, β and γ, and that the same is true of $1 + \sqrt{-5}$ and $1 - \sqrt{-5}$. One way to establish this is to show that each of them satisfies the following three conditions:

(iii) There exist x and y such that n divides neither x nor y, but divides their product.
(iv) If n divides x^2, then n divides x.
(v) If n divides the product xyz, then n divides at least one of xy, xz, or yz.

The first two conditions are easy to establish. The last is a bit more difficult and lengthy, so we will not prove them here.

We now know that in $\mathbb{Z}[\sqrt{-5}]$, 2 behaves like the square of an ideal prime α, and that each of 3, $1 + \sqrt{-5}$, and $1 - \sqrt{-5}$ behave like the product of two distinct ideal primes. If we are to have some kind of unique factorization in $\mathbb{Z}[\sqrt{-5}]$, then in light of the fact that

$$2 \times 3 = (1 + \sqrt{-5}) \times (1 - \sqrt{-5}),$$

it must be the case that α is one of the prime factors of each of $1+\sqrt{-5}$ and $1-\sqrt{-5}$ and that the other factors are the prime factors of 3. That is, we must have $2 = \alpha^2$, $3 = \beta\gamma$, $1 + \sqrt{-5} = \alpha\beta$, and $1 - \sqrt{-5} = \alpha\gamma$ for distinct ideal primes α, β, and γ. Luckily, as far as divisibility in $\mathbb{Z}[\sqrt{-5}]$ is concerned, these identifications do work out perfectly.

Of course, the "ideal prime numbers" α, β, and γ do not actually exist in $\mathbb{Z}[\sqrt{-5}]$, which is why we call them *ideal* primes, after all. But, in studying divisibility, we can in fact proceed as if they did exist, as if we had unique factorization into primes, whether *actual* — numbers in $\mathbb{Z}[\sqrt{-5}]$ that have the prime divisor property (for example 11 or 13) — or *ideal* (such as α).

We can think of "ideal prime numbers" as analogous to quarks in the study of matter. Elementary particles (the irreducible numbers) are

made up of quarks that combine in different ways. We never observe isolated quarks, we only observe them in combinations making up certain particles. Likewise, we never observe these ideal prime numbers as occurring independently, we only "see" them when they combine with one another to form actual numbers in the domain.

Kummer does not in fact define what "ideal prime numbers" are but instead always speaks of them indirectly, in terms of the divisibility properties of actual numbers. It takes a fair amount of work to make sure that everything does work out properly (and a certain amount of luck: if we attempted to proceed analogously in $\mathbb{Z}[\sqrt{-3}]$, for instance, we would soon encounter unsurmountable difficulties and fail completely).

For the cyclotomic rings $\mathbb{Z}[\zeta_p]$ where ζ_p is a primitive pth root of unity, Kummer proved ([15]):

Theorem 5.1 (**Unique Factorization into Ideal Primes**). *Every element of $\mathbb{Z}[\zeta_p]$ factors uniquely as a product of (ideal and actual) primes.*

The finiteness condition can be established once again through a norm, which maps each element of $\mathbb{Z}[\zeta_p]$ to the product of all its images under the different embeddings of $\mathbb{Q}(\zeta_p)$ (the field of fractions of $\mathbb{Z}[\zeta_p]$) into $\mathbb{C}$. Enough of unique factorization is then recaptured to proceed with the arithmetic almost as usual.

Among other things, Kummer used his ideal prime numbers to establish a variant of Lamé's argument for Fermat's Last Theorem. The argument does run, however, into certain subtle technical difficulties and does not work for an arbitrary prime. Kummer listed a set of two conditions that p would have to satisfy to be able to deduce the $n = p$ case of Fermat's Last Theorem using this approach. He even informed the Berlin Academy that he "had reason to believe" that $p = 37$ did not satisfy them (it does not; see [7]). On the other hand, Kummer considered his proof of Fermat's Last Theorem for the so-called regular primes a by-product of his research into higher reciprocity laws, not the main interest of his development.

6. Kronecker, Dedekind, and algebraic integers.

Kummer's approach had a number of drawbacks, not the least of which was the imprecise nature of "ideal numbers." It was also hard to see how to extend the notion to extensions of $\mathbb{Q}$ other than cyclotomic fields. Although the ideas worked very well in $\mathbb{Z}[\sqrt{-5}]$ and some rings of the form $\mathbb{Z}[\theta]$, they failed completely when applied to others, like $\mathbb{Z}[\sqrt{-3}]$.

The first difficulty lies in finding the correct notion of *number* or *integer* with which to work. In general simply taking $\mathbb{Z}[\theta]$ for the field $\mathbb{Q}(\theta)$ will not do. In some cases, a small adjustment is all that is needed (if one works with $\mathbb{Z}[(1+\sqrt{-3})/2]$ instead of $\mathbb{Z}[\sqrt{-3}]$, all difficulties disappear), but in others there is no apparent way of avoiding problems.

The right definition was eventually given by Dedekind (it had also been discovered independently by Kronecker): the correct generalization of integer is that of "algebraic integer." The concept had already appeared in the work of Eisenstein and others, although it had received no special attention.

An *algebraic number* is a complex number α that satisfies some monic polynomial with rational coefficients. In contrast, an *algebraic integer* is a complex number α which satisfies a monic polynomial with *integer* coefficients.

Every integer is an algebraic integer, and in fact the only rational numbers that are algebraic integers are the integers themselves. This follows, for example, using the rational root test from precalculus. It is a bit harder, but not much, to show that the product and sum of any two algebraic integers is again an algebraic integer, and that the roots of any monic polynomial with algebraic integer coefficients are again algebraic integers.

In general, given a domain D and a subring R of D, we say that an element of D is *integral over* R if it satisfies a monic polynomial with coefficients in D. The algebraic integers are the elements of $\mathbb{C}$ that are integral over $\mathbb{Z}$, and if a complex number is integral over the ring of algebraic integers, then it is an algebraic integer as well.

At this point, it would seem that algebraic integers are too much trouble to be of use. If $\mathcal{A}$ signifies the collection of all algebraic integers, then we do not even have factorization into irreducibles. For if α belongs to $\mathcal{A}$ and α is not a unit there, then $\sqrt{\alpha}$ is also an element of $\mathcal{A}$ and not a unit, so $\alpha = \sqrt{\alpha}\sqrt{\alpha}$ is not irreducible. In particular, no element of $\mathcal{A}$ is irreducible! What is more, we can find many more factorizations of α: for instance, $\alpha = \rho_1\rho_2$, where ρ_1 and ρ_2 are the roots of $x^2 - x + \alpha$. The very first step towards a proof of unique factorization, which we managed for other rings, breaks down in $\mathcal{A}$.

In order to surmount this difficulty, we restrict ourselves to subrings of $\mathcal{A}$ where we do have a natural finiteness condition. This is accomplished by first specifying a *finite* extension K of $\mathbb{Q}$ (called a *number field*) and then considering its *ring of integers* $\mathcal{O}_K = K \cap \mathcal{A}$, the collection of all algebraic integers that are in K. This has the virtue of also providing the correct subring of an arbitrary extension $\mathbb{Q}(\theta)$ in which

to continue the pursuit of unique factorization. The reason $\mathbb{Z}[\sqrt{-3}]$ gives so much trouble is that the collection of all algebraic integers in $\mathbb{Q}(\sqrt{-3})$ is not $\mathbb{Z}[\sqrt{-3}]$, but $\mathbb{Z}[(1+\sqrt{-3})/2]$: $(1+\sqrt{-3})/2$ satisfies the monic polynomial $x^2 - x + 1$.

Another difficulty lies in trying to establish unique factorization into ideal numbers. Kummer's method relied very strongly on properties of $\mathbb{Q}(\zeta_p)$ that are not shared by arbitrary extensions $\mathbb{Q}(\theta)$. In particular, the ring of integers of $\mathbb{Q}(\zeta_p)$ admits a basis over $\mathbb{Z}$ consisting of powers of the same element (namely, ζ_p), so it is of the form $\mathbb{Z}[\zeta_p]$. But some number fields K do not admit such bases: for example, if $K = \mathbb{Q}(\sqrt{7}+\sqrt{10})$, then $\mathcal{O}_K \neq Z[\theta]$ for any θ in $\mathcal{O}_K$ (see [16, chap. 2, Exercise 30]).

Several attempts had been made to generalize Kummer's arguments; in 1865 Selling, a student of Dedekind, produced an argument that in fact ended up in nonsense (it can be made rigorous by using q-adic numbers, but these numbers would not be introduced by Hensel until 1897). Dedekind had attempted a different generalization in 1857 (later redeveloped independently by Zolotareff in 1880), but both Dedekind and Kronecker were stopped by the difficulties presented by any field whose ring of integers was not of the form $\mathbb{Z}[\theta]$. (For a more detailed explanation of the difficulties, see [3, chap. 7].)

In order to avoid these difficulties, it was necessary to give some substance to Kummer's ideal numbers, to have something tangible to work with rather than these shadowy constructs that were never explicitly defined. This was accomplished independently by Kronecker and Dedekind using two very different techniques.

Kronecker and forms. Kronecker was a student and colleague of Kummer. His approach generalized Gauss's theory of forms, the aforementioned subject of section 5 of the *Disquisitiones*. A *form* over a number field K is a homogeneous polynomial in arbitrarily many variables whose coefficients are algebraic integers in K (Gauss had considered binary quadratic forms with integer coefficients). Suitably chosen forms play the role of the ideal prime numbers and ideal numbers of Kummer. In modern terms, we adjoin the ideal prime numbers to our field by first adjoining indeterminates and then taking the quotient by the corresponding (ideal of) relations, in order to create a bigger field K'. If we concentrate only on the algebraic integers of K, then each can be written uniquely as a product of prime elements of $\mathcal{O}_K$ and certain elements of $\mathcal{O}_{K'}$. The former correspond to actual primes, while the latter play the role of the ideal primes.

Of course, there are now new numbers in $\mathcal{O}_{K'}$ that may not be expressible uniquely as products of irreducibles of $\mathcal{O}_{K'}$, but we do not

worry about them. As long as we only care about the algebraic integers of K, we can proceed as Kummer does.

Kronecker's method survives as a major tool in algebraic geometry, but it has had a lesser impact in number theory. He was slow to publish, and he did not have Dedekind's gift for exposition. Although Dedekind's work was mostly ignored when it first appeared, his "theory of ideals" would take center stage in Hilbert's landmark *Zahlbericht* [11] and form the basis of algebraic number theory. Even today, over a hundred years after its appearance, Dedekind's exposition in [5] could still be used as a suitable introduction to the subject (in itself a great testament to its impact and clarity). Nonetheless, it should be mentioned that a complete correspondence can be established between Kronecker's approach and Dedekind's, as Hilbert does in the *Zahlbericht*, so the two methods are, in essence, equivalent.

Dedekind and ideals. Dedekind understood the dangers of the somewhat shadowy approach of Kummer. As he writes[5, pp. 57]:

> While this introduction of new numbers is entirely legitimate, it is nevertheless to be feared at first that the language which speaks of ideal numbers being determined by their products, presumably in analogy with the theory of rational numbers, may lead to hasty conclusions and incomplete proofs. And in fact this danger is not always completely avoided. On the other hand, a precise definition covering *all* the ideal numbers that may be introduced in a particular numerical domain $[\mathcal{O}_K]$, and at the same time a general definition of their multiplication, seems all the more necessary since the ideal numbers do not actually exist in the numerical domain $[\mathcal{O}_K]$.

Dedekind preferred, if at all possible, to make this definition solely in terms of objects he already had in hand. Moreover, he wanted a definition that would "create" all these new objects simultaneously, rather than through a recursive process, and that would allow for ease of calculation. His classical construction of the reals as *Dedekind cuts* illustrates this general philosophy: assuming we understand the rationals and all their arithmetic properties, we define the reals as sets of rational numbers satisfying certain properties and define their operations in terms of operations of the rational numbers.

What is more, his construction of the reals also illustrates another very interesting idea, namely, identifying an object with a set that somehow characterizes it. For the real numbers, constructed with a view towards their order, a real number a is uniquely determined by

the collection of all rational numbers less than or equal to a, so it can be identified with such a set.

Dedekind accomplished a similar program for Kummer's ideal numbers, through the introduction of *ideals*. Dedekind defines ideals to be nonempty subsets of the ring of integers $\mathcal{O}_K$, satisfying two conditions:

(1) If a and b are in I, then both $a+b$ and $a-b$ are also in I.
(2) If a is in I and r is in $\mathcal{O}_K$, then ra is in I.

Dedekind named such collections "ideals" because they would play the role of Kummer's ideal numbers, as we will see shortly.

The motivation for this definition is the observation that the collection of all multiples of a given number (including all multiples of a given ideal number in the cyclotomic case) satisfies these two conditions. In the case of the integers, we can identify a number with the collection of all its multiples, and since we are dealing with a principal ideal domain, every collection which satisfies these conditions corresponds to an integer. (To be more precise, the collections correspond to an equivalence class of integers, where $a \sim b$ if a and b differ by a unit, but we are interested in divisibility; multiplication by units becomes irrelevant).

However, in the case of rings of integers that are not UFDs, there are collections satisfying these two conditions that do not correspond to actual numbers. In the case of $\mathbb{Z}[\sqrt{-5}]$, to consider a now familiar situation, the collection of all multiples of the ideal prime α, where $\alpha^2 = 2$, satisfies the conditions. We simply identify each such collection with a number, ideal or actual, since they will correspond to "all multiples" of that number.

One defines an addition and a multiplication of ideals: the sum $I+J$ of two ideals I and J is the ideal

$$I + J = \{i + j \,|\, i \in I, j \in J\},$$

while the product IJ is somewhat more complicated, defined as follows:

$$IJ = \left\{ \sum i_k j_k \,|\, i_k \in I, j_k \in J \right\},$$

that is, all (finite) sums of products of an element of I by an element of J. It would have been nice to define IJ as the set of all products ij, but unfortunately this is not an ideal, so we have to include all finite sums of such products as well.

Given an element $a \in R$, the principal ideal generated by a, which is denoted (a), is defined by

$$(a) = \{ra \,|\, r \in R\},$$

(i.e., it consists of all multiples of a). It is then easy to verify that a divides b in R if and only if (b) is contained in (a) as sets. We

generalize this observation and say that the ideal I *divides* the ideal J if J is contained in I.

In the case of $\mathbb{Z}$, the division algorithm shows that every ideal is principal and that $(a) + (b)$ is none other than the ideal generated by the greatest common divisor of (a) and (b), as can be expected from the fact that the ideal $(a) + (b)$ is the smallest ideal that contains both (a) and (b) (hence, is generated by the largest number dividing both a and b). To mirror all the divisibility properties of $\mathbb{Z}$ in its ideals, we say an ideal P is a *prime ideal* if whenever P contains a product IJ of two ideals, either P contains I or P contains J (this is equivalent to the usual definition: P is a prime ideal if and only if whenever a product xy lies in P, at least one of x and y must lie in P). It is now a nice exercise to verify that an ideal (p) in $\mathbb{Z}$ is a prime ideal if and only if p is a prime number. In short, all the arithmetic theory of $\mathbb{Z}$ may be restated in terms of ideals, ideal multiplication, ideal division, and prime ideals.

We can then use the ideals of $\mathcal{O}_K$ to play the role of the ideal numbers; rather than defining divisibility by α in terms of divisibility by 2, we define divisibility by α in terms of belonging to the ideal we have identified with α. In general, principal ideals correspond to actual numbers (up to units) — namely, their generators — while ideals that are not principal correspond to Kummer's ideal numbers, with no actual counterpart.The main difficulty in this approach, as Dedekind candidly admits, is showing that the notions of divisibility and multiplication of ideals are connected as we hope. That is, it is easy to verify that if I and J are ideals of $\mathcal{O}_K$, then I and J both divide IJ, in the sense that they both contain it. More difficult, however, is showing that if I divides the ideal I' (i.e., if I contains I'), then there exists a unique ideal J such that $IJ = I'$.

This Dedekind succeeded in doing, though only after considerable effort. Once that was done, the next step was to prove that every ideal can be factored uniquely as a product of prime ideals. Thus, the ideal numbers of Kummer are replaced by ideals, and we can rescue for all rings of integers $\mathcal{O}_K$ as much of unique factorization as Kummer had restored to $\mathbb{Z}[\zeta_p]$.

7. Properties of rings of integers.

It has been a while since we broached the topic, so it may be appropriate to remind our readers of where we were headed when we took them on the foregoing detour through history. We are interested in answering the following question:

Question 1.1. Can every polynomial with integer coefficients be factored into (not necessarily monic) linear terms, each with algebraic integer coefficients?

Rings of integers provide the context in which we can analyze this question. It seems reasonable to wonder if, with unique factorization into prime ideals now at our command, we have rescued enough to push through the results used to establish the UFD property for $\mathbb{Z}[x]$. If we could do that, then starting with a polynomial $f(x)$ from $\mathbb{Z}[x]$, we could let K be its splitting field and then attempt to lift the factorization of $f(x)$ into linear terms in $K[x]$ to a factorization in $\mathcal{O}_K[x]$. Unfortunately, we do not have *quite* enough to be able to do this directly, but a solution suggests itself quickly enough.

First, we recall the necessary definitions and properties of algebraic integers and rings of integers. Some of the results are not trivial, but they are classical, so we will simply refer the reader to a standard textbook in algebraic number theory. What we want to highlight is how these important classical results combine to give an answer to our question, in a way that *almost* parallels the development for $\mathbb{Z}[x]$.

Recall that a complex number α is an algebraic number if and only if there exists a monic polynomial $f(x)$ in $\mathbb{Q}[x]$ such that $f(\alpha) = 0$. A complex number a is said to be an algebraic integer if and only if there exists a monic polynomial $g(x)$ in $\mathbb{Z}[x]$ such that $g(a) = 0$. We denote the collection of all algebraic numbers by $\overline{\mathbb{Q}}$ and the collection of all algebraic integers by $\mathcal{A}$. A *number field* is a finite extension of $\mathbb{Q}$. Given a number field K, its *ring of integers* is the collection of all algebraic integers lying in K, and we denote it by $\mathcal{O}_K$. It is not hard to verify that K is the field of fractions of $\mathcal{O}_K$, so every element of a number field can be written as a quotient of two elements of its ring of integers.

The most important result that we need to know, due to Dedekind, reads as follows [5, sec. 25, Prop. 4]:

Theorem 7.1 (**Unique Factorization of Ideals**). *Let K be a number field, and let $\mathcal{O}_K$ be its ring of integers. Each nonzero ideal of $\mathcal{O}_K$ can be factored uniquely as a product of prime ideals, with the trivial ideal $\mathcal{O}_K$ corresponding to the empty product.*

The general philosophy when working with rings of integers is to avoid the use of divisibility statements in terms of numbers, resorting instead to divisibility in terms of ideals. So we work with ideals rather than elements, translating many (but not all) notions and properties of divisibility from $\mathbb{Z}$ to the ring of integers of a number field.

We say that an ideal $\mathfrak{b}$ of $\mathcal{O}_K$ *divides* the ideal $\mathfrak{a}$ if $\mathfrak{b}$ contains $\mathfrak{a}$; this is in fact equivalent to the existence of an ideal $\mathfrak{c}$ such that $\mathfrak{bc} = \mathfrak{a}$. Likewise, we say two ideals $\mathfrak{a}$ and $\mathfrak{b}$ are *coprime* if their factorizations into prime ideals have no prime ideal factor in common or, equivalently, if the ideal $\mathfrak{a} + \mathfrak{b}$, the smallest ideal dividing both, is the trivial ideal $\mathcal{O}_K$. We call elements a and b of $\mathcal{O}_K$ coprime (or *relatively prime*) if the ideals (a) and (b) are coprime or, equivalently, if the ideal $(a, b) = (a) + (b)$ is the trivial ideal $\mathcal{O}_K$. Since rings of integers are not, in general, UFDs, they are also not in general principal ideal domains (the two conditions are in fact equivalent for rings of integers).

Let K and K' be number fields, with K a subfield of K'. Given an ideal $\mathfrak{a}$ of $\mathcal{O}_K$, we can extend it to an ideal of $\mathcal{O}_{K'}$ by taking the ideal it generates in the larger ring, that is, the ideal $\mathfrak{a}\mathcal{O}_{K'}$. Conversely, given an ideal $\mathfrak{b}$ of $\mathcal{O}_{K'}$, we can restrict it to $\mathcal{O}_K$ by taking $\mathcal{O}_K \cap \mathfrak{b}$. If $\mathfrak{a}$ is not the trivial ideal in $\mathcal{O}_K$, then its extension to $\mathcal{O}_{K'}$ is also nontrivial: if an element a of $\mathfrak{a}$ had a multiplicative inverse in $\mathcal{O}_{K'}$, then this inverse would be in K and therefore would already lie in $\mathcal{O}_K$. We say that the ideal $\mathfrak{b}$ of $\mathcal{O}_{K'}$ *lies over* the ideal $\mathfrak{a}$ of $\mathcal{O}_K$ if $\mathfrak{a} = \mathfrak{b} \cap \mathcal{O}_K$ (we also say that $\mathfrak{a}$ lies under $\mathfrak{b}$).

Theorem 7.2 (**Lying Over**). *Let K and K' be number fields, with K a subfield of K'. If $\mathfrak{p}$ is a prime ideal of $\mathcal{O}_K$, then there exists a prime ideal $\mathfrak{q}$ of $\mathcal{O}_{K'}$ lying over $\mathfrak{p}$.*

One way to see this is to look at the ideal of $\mathcal{O}_{K'}$ generated by $\mathfrak{p}$, factor it into prime ideals, and let $\mathfrak{q}$ be any prime factor. A full proof, together with other properties, is found in [16, Theorem 20].

Note that if $\mathfrak{a} = (a)$ is principal, then its extension is also principal, generated by a. However, it is in general false that an ideal lying under or over a principal ideal must be principal.

The analogy between ideals and elements is unfortunately not complete. In a UFD, for example, any two elements have a greatest common divisor that is an element, which among other things allows one to take a quotient a/b of elements of the domain and rewrite it in lowest terms (i.e., $a/b = c/d$, where c and d are elements of the domain that are relatively prime). With ideals, however, the greatest common divisor is defined as an ideal and may have no actual counterpart in the domain. One key difficulty in simply extending Gauss's lemma and its associated corollaries to arbitrary number fields is precisely the absence of greatest common divisors: in a number field K, if we have $a, b \in \mathcal{O}_K$, it may be impossible to express a/b in lowest terms. For example, in $K = \mathbb{Q}(\sqrt{10})$, we have $\mathcal{O}_K = \mathbb{Z}[\sqrt{10}]$, where $\sqrt{10}/2$ cannot be expressed in lowest terms. On the other hand, if the ideal (a, b) is

principal, say, generated by d, then we may write $a = dx$ and $b = dy$ for x and y in $\mathcal{O}_K$, in which case $a/b = x/y$ and $(x, y) = \mathcal{O}_K$.

Thus, a final property that we need is the following, which follows from the finiteness of the class number [16, Corollary 2, p. 132]:

Theorem 7.3 (**Extending to a Principal Ideal**). *Let K be a number field, and let $\mathfrak{a}$ be an ideal of $\mathcal{O}_K$. Then there exists a positive integer $k > 0$ such that $\mathfrak{a}^k$ is principal. In particular, there exists a finite extension L of K such that $\mathfrak{a}\mathcal{O}_L$ is principal.*

Our main use of this theorem involves its second clause. To see how to obtain L from the first clause of the theorem, suppose that $\mathfrak{a}^k = (a)$ and that $L = K(a^{1/k})$, where $a^{1/k}$ is any fixed kth root of a, and consider the principal ideal of $\mathcal{O}_L$ generated by $a^{1/k}$. Since

$$(a^{1/k})^k = (a) = \mathfrak{a}^k\mathcal{O}_L = (\mathfrak{a}\mathcal{O}_L)^k,$$

unique factorization into prime ideals shows that $\mathfrak{a}\mathcal{O}_L = (a^{1/k})$, as needed. Note that one way to interpret the first clause of Theorem 7.3 is that an ideal number (that is, an ideal) can be transformed into an actual number (principal ideal) by raising it to a sufficiently high power (i.e., every ideal number is the kth root of an element of $\mathcal{O}_K$ for suitable k).

By extending to a principal ideal we conclude that for any given members a and b of $\mathcal{O}_K$ there exists a finite extension L of K and elements x and y of $\mathcal{O}_L$ such that $a/b = x/y$ and $(x, y) = \mathcal{O}_L$. If (a, b) was already principal, we can take $L = K$. In the case of $\sqrt{10}/2$ in $\mathbb{Q}(\sqrt{10})$, for example, it suffices to go to the extension $L = \mathbb{Q}(\sqrt{10}, \sqrt{2}) = \mathbb{Q}(\sqrt{5}, \sqrt{2})$. In $\mathcal{O}_L$, the ideal $(\sqrt{10}, 2)$ is principal, generated by $\sqrt{2}$. Cancelling that factor, we have $\sqrt{10}/2 = \sqrt{5}/\sqrt{2}$, and the latter fraction expresses this number in lowest terms (relative to $\mathcal{O}_L$).

The extension to principal ideals shows that in the ring $\mathcal{A}$ of all algebraic integers every finitely generated ideal is principal. If

$$(a_1, \ldots, a_n) = (a_1) + \cdots + (a_n),$$

we let $K = \mathbb{Q}(a_1, ..., a_n)$ and extend the ideal generated by $a_1, \ldots, a_n$ in $\mathcal{O}_K$ to a principal ideal in $\mathcal{O}_L$ for some extension L of K. The original ideal in $\mathcal{A}$, which is the extension of this ideal of L, must likewise be principal.

This means in particular, that for any algebraic integers a and b, the ideal (a, b) of $\mathcal{A}$ is principal, so a and b have a greatest common divisor d in $\mathcal{A}$. Moreover, d can be written as a linear combination $d = \alpha a + \beta b$ with α and β in $\mathcal{A}$. This greatest common divisor is unique only up to units, however, and we do not have an obvious choice among them,

as we do in $\mathbb{Z}$. As usual, we will not be overly concerned with this, for we are interested in divisibility. Dedekind calls the fact that any two algebraic integers a and b have a greatest common divisor in $\mathcal{A}$ that can be expressed as a linear combination of a and b an "important theorem," but he notes that "it is not at all easy to prove" without first developing finiteness of the class number [5, pp. 106].

Domains in which every finitely generated ideal is principal are sometimes called *Bézout domains*. The usual proof of Gauss's lemma and associated results can be extended to any Bézout domain, or more generally, to any *GCD-domain*, meaning an integral domain in which any pair of elements have a greatest common divisor (which need not be expressible as a linear combination of them) [13, sec. 1.6, Exercise 8 and Theorem 102].

It is worth noting, however, that $\mathcal{A}$ is not a principal ideal domain: there are ideals which are not even finitely generated. For example, it is not hard to verify that

$$(2, 2^{1/2}, 2^{1/3}, \ldots, 2^{1/n}, \ldots);$$

cannot be finitely generated.

Although in this discussion we have glossed over the finiteness of the class number, most number theorists should recognize it as the arithmetical heart of our argument in the next section. We emphasize Theorem 7.3 instead because it entails a slightly weaker condition. The exact analogy is that the finiteness of the class number asserts that a certain group is finite, whereas Theorem 7.3 asserts merely that the group is a torsion group (i.e., each element has finite order).

8. Gauss's Lemma for Number Fields.

This section is patterned after [2, sec. 11.9]. We only sketch some of the proofs, since they mimic very closely the proof that $\mathbb{Z}[x]$ is a UFD.

Suppose that K is a number field. We say that a polynomial $f(x)$ in $\overline{\mathbb{Q}}[x]$ is *defined over* K if the coefficients of $f(x)$ lie in K. Now let $f(x)$ belong to $\mathcal{A}[x]$, let K be a number field such that $f(x)$ is defined over K, and let $\mathcal{O}_K$ denote the ring of integers of K. Then the coefficients of $f(x)$ all lie in $\mathcal{O}_K$. Accordingly, we can make the following definition:

Definition 8.1. For $f(x)$ in $\mathcal{O}_K[x]$, the *content* $\mathrm{cont}_K(f)$ *of* $f(x)$ *in* K is the ideal of $\mathcal{O}_K$ generated by the coefficients of $f(x)$. The polynomial $f(x)$ is *primitive in* K if $\mathrm{cont}_K(f) = \mathcal{O}_K$.

This is in fact very similar to the definition of *content of a form* given by Kronecker (see [11, chap. 6]). The main difference is that

Kronecker considers the norm of this ideal, so the content is a positive integer rather than just in ideal of $\mathcal{O}_K$.

The content of a polynomial clearly depends on K, since it must be an ideal of $\mathcal{O}_K$. However, the property of being primitive does not depend on the specific K:

Lemma 8.2 (**Independence of Primitivity**). *Let $f(x)$ belong to $\mathcal{A}[x]$, and let K and K' be two number fields over which $f(x)$ is defined. Then $f(x)$ is primitive in K if and only if it is primitive in K'.*

Proof. By looking at the compositum of K and K' (i.e., the smallest field containing both K and K'), it suffices to establish the result when K is contained in K'. If $f(x)$ is not primitive in K, then there is a prime ideal $\mathfrak{p}$ of $\mathcal{O}_K$ such that every coefficient of $f(x)$ lies in $\mathfrak{p}$. (To see this, factor the content into primes, $\mathrm{cont}_K(f) = \mathfrak{p}_1 \cdots \mathfrak{p}_n$, and take $\mathfrak{p} = \mathfrak{p}_i$ for any i.) Now let $\mathfrak{q}$ be a prime of $\mathcal{O}_{K'}$ lying over $\mathfrak{p}$. Then every coefficient of $f(x)$, when considered as a polynomial in $\mathcal{O}_{K'}[x]$, lies in $\mathfrak{q}$, implying that the content of $f(x)$ in K' cannot be the trivial ideal $\mathcal{O}_{K'}$, since $\mathfrak{q}$ divides it. Thus, $f(x)$ is not primitive in K' either.

Conversely, if $f(x)$ is not primitive in K', then there is a prime ideal $\mathfrak{q}$ of $\mathcal{O}_{K'}$ such that every coefficient of $f(x)$ lies in $\mathfrak{q}$. Let $\mathfrak{p} = \mathfrak{q} \cap \mathcal{O}_K$. As $f(x)$ is also defined over K, the coefficients of $f(x)$ lie in $\mathfrak{p}$, so $\mathfrak{p}$ divides the content of $f(x)$ in K. This ensures that $f(x)$ is not primitive in K. □

Since primitivity does not depend on K, we simply say $f(x)$ *is primitive* to mean that it has algebraic integer coefficients and is primitive in any number field K over which it is defined. Gauss's lemma itself is now straightforward, provided we remember to use prime ideals instead of prime numbers:

Theorem 8.3 (**Gauss's Lemma for Number Fields**). *The product of two primitive polynomials is primitive.*

Proof. Proceeding as in the case of $\mathbb{Z}[x]$, consider primitive polynomials $f(x)$ and $g(x)$ in $\mathcal{A}[x]$, and let K be any number field over which both $f(x)$ and $g(x)$ are defined. Write

$$f(x) = \sum a_i x^i, \qquad g(x) = \sum b_j x^j, \qquad h(x) = f(x)g(x) = \sum c_k x^k.$$

Pick an arbitrary prime ideal $\mathfrak{p}$ of $\mathcal{O}_K$, find the smallest indices i_0 and j_0 for which neither a_{i_0} nor b_{j_0} lie in $\mathfrak{p}$, and observe that

$$\begin{aligned} c_{i_0+j_0} &= a_0 b_{i_0+j_0} + \cdots + a_{i_0-1} b_{j_0+1} \\ &\quad + a_{i_0} b_{j_0} \\ &\quad + a_{i_0+1} b_{j_0-1} + \cdots + a_{i_0+j_0} b_0 \end{aligned}$$

does not lie in $\mathfrak{p}$. It follows that $h(x)$ is primitive. $\square$

At this point one might wonder if, with unique factorization of ideals and Gauss's lemma in hand, it might be possible to prove that a factorization of a polynomial $f(x)$ in $\mathcal{O}_K[x]$ lifts from $K[x]$ to $\mathcal{O}_K[x]$. Unfortunately, the results at our disposal are not quite strong enough.

Example 8.4. Consider the field $K = \mathbb{Q}(\sqrt{10})$. The ring of integers of K is known to be $\mathcal{O}_K = \mathbb{Z}[\sqrt{10}]$, which is not a UFD. For instance, $6 = 2 \cdot 3 = (2 + \sqrt{10})(-2 + \sqrt{10})$, where all four of 2, 3, $2 + \sqrt{10}$, and $-2 + \sqrt{10}$ are irreducible. To prove they are indeed irreducible, we once again rely on a norm $N \colon \mathbb{Z}[\sqrt{10}] \to \mathbb{Z}$ given by $N(a+b\sqrt{10}) = a^2 - 10b^2$. The norm has the same basic properties as the norms on $\mathbb{Z}[\sqrt{-3}]$ and $\mathbb{Z}[\sqrt{-5}]$ discussed previously (e.g., it is multiplicative), but now we must expand the characterization of units to those elements of $\mathbb{Z}[\sqrt{10}]$ whose norm equals 1 *or* -1. (For elementary properties of the norm map for general number fields, see [16, chap. 2].)

Note that $N(2) = 4$, $N(3) = 9$, and $N(\pm 2 + \sqrt{10}) = -6$. If any of these were factorable into a product xy of nonunits x and y of $\mathcal{O}_K$, then at least one of x or y would have to have norm ± 2 or ± 3. Since the equations $a^2 - 10b^2 = \pm 2$ and $a^2 - 10b^2 = \pm 3$ have no solutions modulo 5, they have no solutions in $\mathbb{Z}$ either. Thus, there are no elements of $\mathcal{O}_K$ of norm ± 2 or ± 3. All four of the elements in the indicated factorizations of 6 must therefore be irreducible, and hence $\mathcal{O}_K$ is not a UFD (the two factorizations cannot be transformed into one another by multiplication by suitable units because the norms do not match).

Now consider the polynomial $p(x) = 2x^2 - 5$. Its splitting field is $\mathbb{Q}(\sqrt{10})$, for $2x^2 - 5 = 2(x - \sqrt{10}/2)(x + \sqrt{10}/2)$. But there is no way to rewrite this factorization so that all the coefficients lie in $\mathcal{O}_K$.

To see this, assume that $p(x) = (ax + b)(cx + d)$ with a, b, c, and d all in $\mathcal{O}_K$. We then have $ac = 2$. Since 2 is irreducible, we may assume without loss of generality that $a = 1$ and $c = 2$ (by possibly switching a and c and adjusting the factors by a unit). This means that b must be the negative of a root of $p(x)$, but neither $\sqrt{10}/2$ nor $-\sqrt{10}/2$ are algebraic integers. Our original assumption must therefore be wrong, whence $p(x) = 2x^2 - 5$ cannot be factored over $\mathcal{O}_K = \mathbb{Z}[\sqrt{10}]$.

To see where things go wrong in Example 8.4, we go back to the proof that $\mathbb{Z}[x]$ is a unique factorization domain in section 2. It is, in fact, in the step corresponding to the next proposition (factoring out the content) that we find unique factorization of ideals insufficient for our purposes, forcing us to modify our result somewhat:

Proposition 8.5 (**Factoring Out the Content in Number Fields**). *Let $f(x)$ be a nonzero polynomial in $\overline{\mathbb{Q}}[x]$, and let K be a number field over which $f(x)$ is defined. Then there exists a finite extension L of K such that*

$$f(x) = c_f f^*(x),$$

where c_f is a constant lying in L, and $f^*(x)$ is a primitive polynomial defined over L. Moreover, c_f and $f^*(x)$ are unique up to multiplication by units of $\mathcal{O}_L$.

Proof. We begin just as before: we write $f(x)$ as

$$f(x) = \frac{a_n}{b_n}x^n + \cdots + \frac{a_0}{b_0}$$

with the a_i and b_i in $\mathcal{O}_K$. We multiply by $b_0 \cdots b_n$ to clear denominators, and we have

$$(b_0 \cdots b_n)f(x) = g(x),$$

with $g(x) \in \mathcal{O}_K$. We cannot simply write $g(x)$ as a constant times a primitive polynomial, however, because $\mathrm{cont}_K(g)$ is only an ideal, and may not be principal. Of course, there is a finite extension L of K such that we can extend $\mathrm{cont}_K(g)$ to a principal ideal in $\mathcal{O}_L$. Since $\mathrm{cont}_K(g)\mathcal{O}_L$ is the same as $\mathrm{cont}_L(g)$, we have $\mathrm{cont}_L(g) = (c)$ with c in $\mathcal{O}_L$. We can then write $g(x) = c \cdot g^*(x)$, with $g^*(x)$ primitive and defined over L. We thus take $c_f = c/(b_0 \cdots b_n)$, an element of L, and $f^*(x) = g^*(x)$.

For uniqueness up to a unit of L, it is again enough to consider the case of $f^*(x) = cg^*(x)$, with both $f^*(x)$ and $g^*(x)$ primitive and defined over L. Write $c = u/v$ with u and v in $\mathcal{O}_L$. Going to a finite extension of L if necessary, we may assume that u and v are relatively prime by the remarks following Theorem 7.3. We now proceed as in the case of $\mathbb{Z}[x]$ to conclude that both u and v are units, making u/v a unit in $\mathcal{O}_L$, as claimed. □

It is worth noting that if $\mathcal{O}_K$ is a unique factorization domain (or if $\mathrm{cont}_K(g)$ is principal), then in the proof of this proposition we may use $L = K$, since we can factor out the content in K itself.

Although we cannot simply lift factorizations from $K[x]$ to $\mathcal{O}_K[x]$, as we saw in Example 8.4, by factoring out the content we are able to lift a factorization from $K[x]$ to $\mathcal{O}_L[x]$, where L is a finite extension of K. This is the message of the following result:

Theorem 8.6 (**Lifting a Factorization**). *Let K be a number field, and let $f(x)$ belong to $\mathcal{O}_K[x]$. If $f(x) = g(x)h(x)$ for polynomials $g(x)$ and $h(x)$ in $K[x]$, then there is a finite extension L of K such that*

$f(x) = G(x)H(x)$, where $G(x)$ is an L-multiple of $g(x)$, $H(x)$ an L-multiple of $h(x)$, and both $G(x)$ and $H(x)$ have coefficients in $\mathcal{O}_L$.

Proof. In view of Proposition 8.5, we can go to a finite extension L of K where we can write

$$f(x) = c_f f^*(x),\ g(x) = c_g g^*(x),\ h(x) = c_h h^*(x),$$

with $f^*(x)$, $g^*(x)$, and $h^*(x)$ primitive, c_f in $\mathcal{O}_L$, and c_g and c_h in L. We then have

$$c_f f^*(x) = f(x) = g(x)h(x) = (c_g c_h) g^*(x) h^*(x),$$

and Gauss's lemma for number fields tells us that $g^*(x)h^*(x)$ is primitive. By the uniqueness of the representation, $c_f = u c_g c_h$ for some unit u of $\mathcal{O}_L$. In particular $c_g c_h$ lies in $\mathcal{O}_L$, permitting us to write

$$f(x) = \Big(c_g c_h g^*(x)\Big) h^*(x),$$

completing the proof. □

Remark 8.7. It is very important to note that in Theorem 8.6 it is typically necessary to pass to an extension L of K. Of course, if we can factor out the content of $f(x)$, $g(x)$, and $h(x)$ in $\mathcal{O}_K$ (as happens if $\mathcal{O}_K$ is a UFD), then we may again set $L = K$ and get the usual result.

Corollary 8.8 (Complete Factorization in $\mathcal{A}[x]$). *Every nonconstant polynomial $f(x)$ in $\mathcal{A}[x]$ can be factored into a product of (not necessarily monic) linear factors, each with algebraic integer coefficients.*

Proof. Let $f(x)$ be a nonconstant polynomial in $\mathcal{A}[x]$, and let K be the splitting field of $f(x)$. We can now lift the factorization in $K[x]$ to a factorization in $\mathcal{O}_L[x]$ for a finite extension L of K. Since we are lifting a factorization into linear terms, the factorization in $\mathcal{O}_L[x]$, a subdomain of $\mathcal{A}[x]$, is a factorization into linear factors as well, and we are done. □

Specializing to integer coefficients we finally obtain the affirmative answer to Question 1.1:

Corollary 8.9. *Every polynomial $f(x)$ in $\mathbb{Z}[x]$ can be factored into a product of (not necessarily monic) linear factors with algebraic integer coefficients.*

Example 8.10. Let us go back to Example 8.4, and find a factorization with algebraic integer coefficients. Recall that $p(x) = 2x^2 - 5$. Any

factorization of $p(x)$ over $\overline{\mathbb{Q}}$ must be equivalent to

$$2x^2 - 5 = 2\left(x + \frac{\sqrt{10}}{2}\right)\left(x - \frac{\sqrt{10}}{2}\right)$$

up to multiplication by elements of $\overline{\mathbb{Q}}$. We want to factor $2 = ab$ in $\mathcal{A}$ in such a manner that the two elements $a\sqrt{10}/2$ and $-b\sqrt{10}/2$ also lie in $\mathcal{A}$.

In $\mathbb{Z}[\sqrt{10}]$, the greatest common divisor of $\sqrt{10}$ and 2 is defined as the ideal $(\sqrt{10}, 2)$, which is not principal since 2 is irreducible and does not divide $\sqrt{10}$. On the level of ideals, we have $(2) = (\sqrt{10}, 2)^2$. To see this, note that

$$(\sqrt{10}, 2)^2 = (10, 2\sqrt{10}, 4),$$

and since all generators are multiples of 2, it is clear that $(\sqrt{10}, 2)^2$ is contained in (2). Because it is also true that $2 = 10 - 2(4)$, it follows that 2 lies in $(\sqrt{10}, 2)^2$, so we have equality.

As (2) is the square of the greatest common divisor of 2 and $\sqrt{10}$, it would suffice to choose $a = b$, doing it in such a way that a is a generator of the ideal $(\sqrt{10}, 2)$ in some extension of K. We adjoin $\sqrt{2}$ to K to obtain the number field $L = \mathbb{Q}(\sqrt{10}, \sqrt{2}) = \mathbb{Q}(\sqrt{5}, \sqrt{2})$, and in $\mathcal{O}_L$ we have $(\sqrt{10}, 2) = (\sqrt{2})$. Taking $a = b = \sqrt{2}$, we then have in $\mathcal{O}_L$:

$$\begin{aligned} p(x) &= 2\left(x + \frac{\sqrt{10}}{2}\right)\left(x - \frac{\sqrt{10}}{2}\right) \\ &= \sqrt{2}\left(x + \frac{\sqrt{10}}{2}\right)\sqrt{2}\left(x - \frac{\sqrt{10}}{2}\right) \\ &= \left(\sqrt{2}x + \sqrt{5}\right)\left(\sqrt{2}x - \sqrt{5}\right), \end{aligned}$$

which is a factorization into linear polynomials with algebraic integer coefficients. It is not hard to verify that $\mathcal{O}_L = \mathbb{Z}[(1 + \sqrt{5})/2, \sqrt{2}]$ (see [16, chap. 2, Exercise 42(c)]), although the factorization may in fact be achieved over the smaller subring $Z[\sqrt{5}, \sqrt{2}]$.

9. Some ring theory.

Dedekind's ideas and exposition, as found in [5], are surprisingly modern. Modulo a few edits (what Dedekind calls a "module" we would nowadays call a "$\mathbb{Z}$-module", for example), it would be right at home in a modern algebra text. Unfortunately, his ideas were not immediately recognized or adopted. For one thing, there was great resistance at

the time to dealing with infinite sets, such as Dedekind's ideals, as completed objects that one could manipulate. Moreover, those who opposed such infinite constructions had a powerful spokesperson in Kronecker.

It was probably not until Hilbert's landmark *Zahlbericht* that the theory came fully into its own. Hilbert brought together several strands and approaches, and he presented a unified treatment for the algebraic theory of numbers as it had been developed up to that point. Hilbert passes effortlessly between Dedekind's theory of ideals and Kronecker's theory of forms but finds in the unique factorization of ideals into prime ideals one of the "foundation pillars" of algebraic number theory. Indeed, these notions are the very language in which we described our answer to Question 1.1.

Dedekind's approach to the problem of unique factorization in number fields did much more than extend Kummer's work and provide the basic framework on which algebraic number theory would later be built. The notions of modules and ideals were taken up by Emil Artin and Emmy Noether in the 1920s, and generalized into what we now call ring theory. Dedekind's influence was powerful. Indeed, according to Stillwell [5, pp. 3]:

> But even then, Emmy Noether used to say "*Es steht schon bei Dedekind*" (It's already in Dedekind), and urged her students to read all of Dedekind's work in ideal theory.

When we first encountered Question 1.1, our minds naturally turned to algebraic number theory, which explains the solution we found. One can think of that solution as a sort of "bottoms up" solution, in which we proceed by taking finite extensions of number fields to get the appropriate greatest common divisors, and thus mimic the proof of lifting the factorization. As noted earlier, this is very much in the spirit of Kummer, Kronecker, and Dedekind.

However, much can be said as well for a "top down" approach, which would proceed instead by asking:

Question 9.1. *For which integral domains D does the analogue of Theorem 2.9 (lifting the factorization) hold?*

This question had already been asked and answered from a purely ring-theoretic point of view, though we were unaware of it. We would be remiss if we did not also take the opportunity to discuss here some of the notions involved. Many deserve to be better known, and they again revolve around ideas of factorization, treading close to the origins of algebraic number theory.

We will need a few considerations before coming back to this question. We again beg the reader's indulgence.

Uniqueness of the factorization. We know that a domain D is a UFD if and only if the polynomial ring $D[x]$ is a UFD. The "only if" direction follows because any witness to the fact that D is not a UFD will show that $D[x]$ is not a UFD either. On the other hand, even though $\mathcal{A}[x]$ is not a UFD, it comes very close to being one, as we see in the following two results:

Theorem 9.2. *Let $f(x)$ be a primitive polynomial in $\mathcal{A}[x]$. If*

$$f(x) = g_1(x) \cdots g_n(x) = h_1(x) \cdots h_n(x),$$

where the $g_i(x)$ and $h_j(x)$ are polynomials in $\mathcal{A}[x]$ of degree 1, then up to a reordering of the $h_j(x)$ there exist units $u_1, \ldots, u_n$ of $\mathcal{A}$ such that $\prod u_i = 1$ and $u_i h_i(x) = g_i(x)$. In particular, each u_i is a unit in $\mathcal{O}_K$ for any number field K containing it.

Proof. Note that if a product of polynomials in $\mathcal{A}[x]$ is primitive, then each of the factors must be primitive. Thus, each of the $g_i(x)$ and $h_j(x)$ is primitive.

Let K be a number field over which the $g_i(x)$ and $h_j(x)$ are defined. Consider both factorizations in $K[x]$, which is a UFD. The two factorizations must be equivalent up to multiplication by elements of K. Up to a reordering of the $h_j(x)$, we may assume that each $g_i(x)$ is a K-multiple of $h_i(x)$.

Fix an index i. Write $g_i(x) = ax + b$ and $h_i(x) = cx + d$, with a, b, c, and d in $\mathcal{O}_K$ and $ac \neq 0$. Since $g_i(x)$ and $h_i(x)$ are both primitive, it follows that each of the ideals (a, b) and (c, d) is the trivial ideal $\mathcal{O}_K$. We also have an element u_i of K such that $g_i(x) = u_i h_i(x)$. By passing to an extension of K if necessary, we may write $u_i = v_i/w_i$ with v_i and w_i algebraic integers and (v_i, w_i) the trivial ideal. Therefore, $w_i g_i(x) = v_i h_i(x)$, hence $(w_i)(a, b) = (v_i)(c, d)$ as ideals. Since both (a, b) and (c, d) are trivial, and since (w_i) and (v_i) are relatively prime, by the unique factorization of ideals it follows that both (w_i) and (v_i) reduce to the unit ideal. We infer that w_i and v_i are algebraic integer units, as then is u_i. A substitution and cancellation now establishes that $\prod u_i = 1$. □

Theorem 9.3. *Let $f(x)$ be a nonzero polynomial in $\mathcal{A}[x]$. Then we can write*

$$f(x) = c_f g_1(x) \cdots g_n(x),$$

where c_f belongs to $\mathcal{A}$, and each $g_i(x)$ is a primitive polynomial in $\mathcal{A}[x]$ of degree 1. Moreover, the factorization is unique up to units of $\mathcal{A}$.

Proof. This follows by combining Proposition 8.5 (factoring out the content in number fields) and Theorem 9.2. □

In short, even though $\mathcal{A}[x]$ is not a UFD, we have a certain "uniqueness up to constants" in factorizations. It is "merely" the fact that the constants cannot be factored into a product of irreducibles (uniquely or otherwise) that prevents $\mathcal{A}[x]$ from being a UFD.

We restate this unique factorization up to constants in the following corollary in a way that is more amenable to generalization for other domains. Recall that $\overline{\mathbb{Q}}$ is the field of fractions of $\mathcal{A}$.

Corollary 9.4 (**Unique Factorization up to Constants**). *Let $f(x)$ be a polynomial in $\mathcal{A}[x]$ of positive degree. Then we can factor $f(x)$ into a product of nonconstant polynomials with coefficients in $\mathcal{A}[x]$,*

$$f(x) = g_1(x) \cdots g_n(x),$$

such that no $g_i(x)$ can be factored as a product of two nonconstant polynomials. Moreover, any two such factorizations of $f(x)$ are equivalent, in the sense that if

$$f(x) = g_1(x) \cdots g_n(x) = h_1(x) \cdots h_m(x),$$

then $n = m$ and up to a reordering of the $h_i(x)$ there exist constants $u_1, \ldots, u_n$ in $\overline{\mathbb{Q}}$ such that $g_i(x) = u_i h_i(x)$ and $\prod u_i = 1$.

Compare this with what happens in a ring of integers that is not a UFD:

Example 9.5. Let $D = \mathbb{Z}[\sqrt{10}]$. We saw in Example 8.4 that the polynomial $p(x) = 2x^2 - 5$ cannot be written as the product of two linear polynomials in $D[x]$. The same argument shows that $q(x) = 5x^2 - 2$ cannot be written as a product of two linear polynomials in $D[x]$ either. Consider now the two polynomials in $D[x]$:

$$\begin{aligned} r(x) &= \sqrt{10}x^2 - 7x + \sqrt{10}, \\ s(x) &= \sqrt{10}x^2 + 7x + \sqrt{10}. \end{aligned}$$

We have:

$$\begin{aligned} p(x)q(x) &= (2x^2 - 5)(5x^2 - 2) \\ &= 10x^4 - 29x^2 + 10 \\ &= \left(\sqrt{10}x^2 - 7x + \sqrt{10}\right)\left(\sqrt{10}x^2 + 7x + \sqrt{10}\right) \\ &= r(x)s(x). \end{aligned}$$

.

We claim that neither $r(x)$ nor $s(x)$ can be written as a product of two linear polynomials in $D[x]$. Indeed, if

$$\sqrt{10}x^2 - 7x + \sqrt{10} = (ax + b)(cx + d),$$

with a, b, c, and d in D, then $ab = \sqrt{10}$. By looking at the norms, and remembering that no element has norm 2, we see that up to multiplication by a unit we may assume that $a = 1$ and $c = \sqrt{10}$. But then $-b$ must be a root of $\sqrt{10}x^2 - 7x + \sqrt{10}$, whereas neither of its roots lie in D. We conclude that no such factorization exists. A similar argument shows that the same is true of the other factor.

Example 9.5 shows that in $D[x]$ for $D = \mathbb{Z}[\sqrt{10}]$ we do not have unique factorization up to constants: the factorizations of $10x^4 - 29x^2 + 10$ as $p(x)q(x)$ and as $r(x)s(x)$ are not related by multiplication by constants. When do we get this kind of uniqueness of factorization? Now is the time to get back to Question 9.1: since the ring of polynomials over a field is a UFD, if D is a domain in which we can lift factorizations from $K[x]$ to $D[x]$, then in $D[x]$ we will have unique factorization up to constants. Perhaps somewhat surprisingly, the converse also holds:

Lemma 9.6. *Let D be a domain, and let K be its field of fractions. The following statements are equivalent:*

(a) *Any polynomial $f(x)$ in $D[x]$ of degree at least two that can be factored in $K[x]$ can also be factored as a product of two nonconstant polynomials with coefficients in $D[x]$.*
(b) *The analogue of Theorem 2.9 holds for D.*
(c) *There is unique factorization up to constants in $D[x]$.*

Proof. The fact that (a) and (b) are equivalent is clear. That (c) follows from (a) is the observation that $K[x]$ is a UFD whose units are the nonzero constants. Finally, to see that (c) implies (a), assume that $f(x)$ is a polynomial in $D[x]$ that cannot be written as a product of two nonconstant polynomials in $D[x]$, but is nevertheless not irreducible in $K[x]$, say $f(x) = G(x)H(x)$ with $G(x)$ and $H(x)$ of positive degrees. Multiplying by a suitable constant from D to clear denominators, we have $cf(x) = g(x)h(x)$, where $g(x)$ and $h(x)$ are polynomials of positive degrees with coefficients in D. Therefore, we have $cxf(x) = xg(x)h(x)$. The assumption on $f(x)$ means that the polynomial $cxf(x)$ can be factored in $D[x]$ as $(cx)f(x)$, and each term is of positive degree, and cannot be written as the product of nonconstant polynomials in $D[x]$. On the other hand, we have $x \cdot g(x) \cdot h(x)$ with at least three terms (more if either $g(x)$ or $h(x)$ is further reducible), so $D[x]$ does not have unique factorization up to constants. □

Lifting factorizations and the Schreier property. Most students are familiar with the hierarchy of domains that is presented in most upper division abstract algebra courses:

$$\text{Fields} \subset \text{Euclidean Domains} \subset \text{PIDs} \subset \text{UFDs} \subset \text{Domains},$$

and all of the inclusions are proper. (Note that PID stands for principal ideal domain.) However, there are other important properties of domains that are in turn closely related to factorizations and that are not nearly so well known.

We have already mentioned a few in passing. The Bézout property (that every finitely generated ideal is principal), and the greatest common divisor property (that every pair of elements has a greatest common divisor) are two of these. The latter gives us the ability to lift factorizations. Notions that are closely related to the foregoing properties have been investigated. In [1], there is a list of twelve conditions for domains that are linked with the concept of unique factorization, existence and behavior of greatest common divisors, and polynomial factorizations (including, for example, the "Gauss's Lemma Property," which requires that products of primitive polynomials be primitive). Three of these conditions were considered in detail by P. M. Cohn in his wonderful paper [4]. Once again, they revolve around the notions of divisibility, factorizations, and integral elements. We present these three notions here.

Definition 9.7. A domain D is a *pre-Schreier domain* if the following holds for every a, b, and c in D: if $a|bc$, then there exist β and γ in D such that $a = \beta\gamma$, $\beta|b$, and $\gamma|c$. Equivalently, D is a pre-Schreier domain if and only if any two factorizations of an element have common refinements.

Definition 9.8. A domain D is a *Schreier domain* if it is a pre-Schreier domain that is integrally closed in its field of fractions (i.e., any element of the field of fractions of D that satisfies a monic polynomial with coefficients in D already lies in D).

Definition 9.9. We say that a domain D has the *AP-property* (atoms are primes) if every irreducible element a of D satisfies the prime divisor property (i.e., for any elements b and c of D satisfying $a|bc$, either $a|b$ or $a|c$).

We have the following implications among these properties:

$$\begin{array}{ccccccccc} \text{PID} & \subset & \text{UFD} & & & & & & \\ \cap & & \cap & & & & & & \\ \text{Bézout} & \subset & \text{GCD} & \subset & \text{Schreier} & \subset & \text{pre-Schreier} & \subset & \text{AP}. \end{array}$$

All inclusions are known to be proper (see [1]). One connection between these notions and unique factorization is the following[4, Theorem 2.3]:

Theorem 9.10. *A domain D is a UFD if and only if it is a Schreier domain such that each nonunit of D can be factored into a product of irreducible elements in at least one way.*

The Schreier property is exactly the missing link between being able to factor into irreducibles in at least one way and being able to do so (modulo units) in exactly one way. Since in rings of integers we always have factorization in at least one way, the Schreier property for a given $\mathcal{O}_K$ is equivalent to the UFD property.

The Bézout property, a consequence of the finiteness of the class number, implies that the ring $\mathcal{A}$ of all algebraic integers is a Schreier domain. It is precisely the Schreier property that characterizes the domain in which the analogue of lifting the factorization holds, thus answering Question 9.1. One of the key ingredients in the proof of this result once again brings us close to the very foundations of algebraic number theory. It is a generalization of Dedekind's "Prague Theorem," which Hilbert used in the *Zahlbericht* as the key step in establishing the unique factorization of ideals into prime ideals. We quote it here in its original version for algebraic integers:

Theorem 9.11 (**Dedekind's Prague Theorem**). *Let $f(x)$ and $g(x)$ be polynomials with algebraic integer coefficients,*

$$\begin{aligned} f(x) &= a_r x^r + a_{r-1}x^{r-1} + \cdots + a_0, \\ g(x) &= b_s x^2 + b_{s-1}x^{s-1} + \cdots + b_0. \end{aligned}$$

If every coefficient of the product $f(x)g(x)$ is divisible by the integer m, then each of the numbers $a_0b_0, a_0b_1, \ldots, a_0b_s, a_1b_0, \ldots, a_rb_s$ is also divisible by m.

A proof of this result, together with the classical proof of unique factorization into prime ideals, can be found in [18, Lemma 8.12].

Through suitable definitions, one can interpret the Prague theorem as saying that the content of a product of polynomials is the product of the contents. To make this precise, however, requires a modification of our definition of content to something akin to Kronecker's definition, which makes the content of such a polynomial a positive integer rather than an ideal. We will not go into it here. Viewed in that light, the Prague theorem is itself a generalization of Gauss's lemma, so we are still circling the same notions with which we began.

Using a generalization of the Prague theorem for more general rings, P. M. Cohn proved the following theorem:

Theorem 9.12 (**P. M. Cohn**). *If D is a Schreier domain and x is an indeterminate, then $D[x]$ is a Schreier domain.*

This result gives the top-down answer to our question (see [17]):

Theorem 9.13. *Let D be a domain, and let K be its field of fractions. The analogue of Theorem 2.9 (lifting of factorization) holds for $D[x]$ if and only if D is a Schreier domain.*

Proof. First suppose that D is a Schreier domain. By Theorem 9.12 $D[x]$ is a Schreier domain as well. Let $f(x) = G(x)H(x)$ be a factorization of a polynomial $f(x)$ from $D[x]$ in $K[x]$. Multiplying by suitable constants to clear denominators, we obtain c in D such that $cf(x) = g(x)h(x)$, with $g(x)$ and $h(x)$ in $D[x]$, $g(x)$ a D-multiple of $G(x)$, and $h(x)$ a D-multiple of $H(x)$.

Since $f(x)|g(x)h(x)$, the Schreier property implies the existence of $a(x)$ and $b(x)$ in $D[x]$ such that $f(x) = a(x)b(x)$, $a(x)|g(x)$, and $b(x)|h(x)$. By considering the degrees of $a(x)$, $g(x)$, $G(x)$, $b(x)$, $h(x)$, and $H(x)$, we see that $\deg(a) = \deg(G)$ and $\deg(b) = \deg(H)$. Accordingly, we can lift factorizations from $K[x]$ to $D[x]$.

Conversely, suppose that the analogue of Theorem 2.9 holds for $D[x]$. We must prove that D is integrally closed and satisfies the pre-Schreier property. Consider an element k of K that is integral over D, and let

$$f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0$$

be a monic polynomial in $D[x]$ such that $f(k) = 0$. In $K[x]$, we have $f(x) = (x - k)g(x)$ for some $g(x)$; lifting this factorization, we find that $f(x) = a(x)b(x)$ for polynomials $a(x)$ and $b(x)$ in $D[x]$, and such that $a(x)$ is linear and has k as a root. Multiplying by suitable units if necessary, we may assume that both $a(x)$ and $b(x)$ are monic. Then $a(x) = x - k$, which implies that k belongs to D. Thus D is integrally closed.

To see that D satisfies the pre-Schreier property, let a, b, and c be elements of D such that $a|bc$. If either $b = 0$ or $c = 0$, then to verify the pre-Schreier property we simply factor a as $a = a \cdot 1$ or as $a = 1 \cdot a$, respectively. If $a = 0$, then we must have $b = 0$ or $c = 0$, so we may assume that $abc \neq 0$. Let r in D be such that $ar = bc$. Consider the polynomial

$$f(x) = a\left(x - \frac{b}{a}\right)\left(x - \frac{c}{a}\right) = ax^2 - (b + c)x + r$$

which has coefficients in $D[x]$ and has been factored in $K[x]$. Since we can lift factorizations, we must have $f(x) = g(x)h(x)$, with $g(x) = \alpha x - \beta$, $h(x) = \delta x - \gamma$, and α, β, γ, and δ in D. Exchanging $g(x)$

and $h(x)$ if necessary, we may assume that $\beta/\alpha = b/a$, so $a\beta = b\alpha$; and $\gamma/\delta = c/a$, so $a\gamma = c\delta$. We also have $a = \alpha\delta$ and $r = \beta\gamma$. Thus we see that $bc = ar = a\beta\gamma = c\beta\delta$, implying that $b = \beta\delta$. Similarly, $bc = a\beta\gamma = b\alpha\gamma$, whence $c = \alpha\gamma$. Thus we have factored $a = \alpha\delta$, where $\alpha|c$ and $\delta|b$, confirming that D satisfies the pre-Schreier property. We conclude that if we can lift factorizations, then D is a Schreier domain, as claimed. □

It is not hard to verify that we can lift factorizations of *monic* polynomials in $D[x]$ if and only if D is integrally closed, so the notion of integral closure is also closely tied to factorizations. Theorem 9.13 provides a "top-down" answer to Question 1.1: invoke Theorem 7.3 (extending to a principal ideal) to show that $\mathcal{A}$ is a Schreier domain, note that its field of fractions is algebraically closed, and lift the factorization from $\overline{\mathbb{Q}}[x]$ to $\mathcal{A}[x]$.

We also obtain:

Corollary 9.14. *Let K be a number field, and let $\mathcal{O}_K$ be its ring of integers. The analogue of Theorem 2.9 holds for $\mathcal{O}_K$ if and only if $\mathcal{O}_K$ is a UFD.*

10. Dedekind Domains and Function Fields.

Before finishing, we exhibit a situation that closely parallels the case of number fields, but in which Question 1.1 has a negative answer. Most of the details require, unfortunately, some heavy technical machinery, so we merely assert many of the necessary results and point the reader to suitable references.

The ring of integers in a number field is an example of a Dedekind domain, which is an integral domain in which ideals can be uniquely factored into prime ideals. (There are many equivalent definitions of a Dedekind domain; see [12, sec. 10.2]) The other main example of a Dedekind domain is the coordinate ring of a nonsingular curve over an algebraically closed field. These two cases are referred to by their fields of fractions, as "the number field case" and "the function field case," respectively.

There is a strong general theory of Dedekind domains. The close connection between the number field and function field cases was in fact noted by Dedekind, who together with Weber applied these ideas to develop an arithmetic theory of Riemann surfaces [6], which marks the beginning of modern algebraic and arithmetic geometry. The fact that they are analogous was made even clearer when Dedekind developed the notions of ideals and ideal multiplication. His first proofs of

unique factorization were very computational, and relied throughout on the exact nature of the elements in the rings in question.[1] Later on he abstracted the key properties of the ideals, and he proved most of his results in terms of those properties rather than in a computational manner. For any specific setting, the exact nature of the rings in question would be used to show that ideals satisfied those key properties, but after that all further results would follow automatically. Once again, a very modern approach.

Thus, many of our results so far hold for general Dedekind domains: unique factorization of ideals, the lying over theorem for prime ideals, Lemma 8.2, and Gauss's lemma can all be proved in the general setting. Unfortunately, the key property we used to prove the factorization result, Theorem 7.3, fails to hold in general: if $\mathfrak{a}$ is an ideal in the function field analogue of the ring of integers, it is possible that $\mathfrak{a}^k$ is not principal for any positive integer k and that there is no extension of the underlying field in whose ring of integers $\mathfrak{a}$ becomes principal. Because of this failure, the proof of an analogue of the complete factorization theorem for function fields breaks down. We exhibit an explicit counterexample here.

A precise definition of "function field" would take us too far away from the main discussions on this paper. Thus, we discuss only the two specific examples necessary to answer Question 1.1 for general Dedekind domains. The easiest example of a function field is the field $\mathbb{C}(x)$ of rational functions of a single variable x; this is the analogue of the field of rational numbers $\mathbb{Q}$. Inside $\mathbb{C}(x)$ is the ring of polynomials $\mathbb{C}[x]$, which is the function field analogue of $\mathbb{Z}$.

Since $\mathbb{C}[x]$ is a unique factorization domain, it will certainly not furnish the counterexample we seek. This motivates us to find a second example of a function field. We consider the larger ring $A = \mathbb{C}[x, y]/(p(x, y))$, where $p(x, y)$ is a nonconstant polynomial in two variables, satisfying an additional technical hypothesis.[2] We consider the following polynomial:

$$p(x, y) = y^2 - x^3 + 3x - 49$$

[1] In his exposition of ideals published in 1871, for example, Dedekind does not even define multiplication of ideals until *after* he has proven the unique factorization theorem. Though this sounds paradoxical, recall that divisibility was defined in terms of inclusion of ideals, not in terms of multiplication. Dedekind proved that every ideal was the intersection of all prime ideal powers that divide it.

[2] The technical hypothesis is that there should be no solutions (x, y) in complex numbers x and y to the equations $p_x(x, y) = p_y(x, y) = p(x, y) = 0$, where p_x and p_y denote partial derivatives.

For this choice of $p(x, y)$, the ring A is a Dedekind domain. If we denote its field of fractions by $K = \mathbb{C}(x)[y]/(p(x, y))$, then K is a function field, and A is the analogue of the ring of integers $\mathcal{O}_K$ in the number field setting. Just as in the number field case, the ring A is precisely the set of elements α of K which satisfy a monic polynomial with coefficients in $\mathbb{C}[x]$:

$$\alpha^n + q_{n-1}\alpha^{n-1} + \cdots + q_1\alpha + q_0 = 0,$$

where $q_i = q_i(x)$ in $\mathbb{C}[x]$ for each i. In other words, A is the set of $\mathbb{C}[x]$-*integers* in K.

Notice that A is closely associated with the following elliptic curve C:

$$y^2 = x^3 - 3x + 49$$

The nonzero prime ideals of A correspond naturally to points lying on C, via the following one-to-one correspondence:

$$(a, b) \in C \longleftrightarrow \mathfrak{p} = (x - a, y - b)$$

For example, the prime ideal $(x, 7+y)$ corresponds to the point $(0, -7)$. The ideal $(x - a, y - b)$ is precisely the set of polynomials in A which vanish at the point (a, b).[3] If the point (a, b) does not lie on C, then the ideal $(x-a, y-b)$ will be the unit ideal of A, and in particular will not be prime. It is not at all obvious that every nonzero prime ideal of A has the form $(x - a, y - b)$ for some point (a, b) in C, but it is nevertheless true.

This correspondence also sheds some light on prime factorization. For instance, the ideal $(7 + y)$ can be factored in A as:

$$(7 + y) = (x, 7 + y)(x - \sqrt{3}, 7 + y)(x + \sqrt{3}, 7 + y)$$

which reflects the fact that the three points $(0, -7)$, $(\sqrt{3}, -7)$, and $(-\sqrt{3}, -7)$ are precisely the points of the curve C such that $7 + y = 0$.

Another factorization, which will be a key player in our counterexample, is the factorization of the ideal (x): namely,

$$(x) = (x, 7 + y)(x, 7 - y),$$

again corresponding to the fact that $(0, \pm 7)$ are the only two points in C with x-component equal to 0. This is a factorization of (x) into two *distinct* prime ideals of A; if the two ideals were the same, then they would both contain $(7 + y) + (7 - y) = 14$, which is obviously untrue because 14 is a unit of A.

[3]Technically, the elements of A are equivalence classes of polynomials, but if $f(x, y)$ and $g(x, y)$ are congruent modulo $(p(x, y))$, then they have the same value at a point (a, b) of C.

The important feature of A, for our purposes, is the failure of the analogue of Theorem 7.3. In fact, for most prime ideals $\mathfrak{p}$ of the ring A, the ideal $\mathfrak{p}^k$ is not principal for any positive integer k. In particular, we have the following result:

Theorem 10.1. *For every positive integer k, the ideal $(x, 7+y)^k$ is not principal.*

A proof of Theorem 10.1 is well beyond the scope of this paper, unfortunately.[4] Taking it for granted, however, we are now able to describe the promised counterexample to the analogue of Question 1.1 for function fields. Let $f(T)$ be the following polynomial, with coefficients in A:

$$\begin{aligned} f(T) &= xT^2 + 14T + (3 - x^2) \\ &= x\left(T + \frac{7+y}{x}\right)\left(T + \frac{7-y}{x}\right) \end{aligned} \tag{10.1}$$

Any factorization of $f(T)$ over $\overline{K}$ (the algebraic closure of K) will be equivalent to Equation 10.1 up to multiplication by elements of $\overline{K}$. To show that we cannot factor $f(T)$ into linear factors with $\mathbb{C}[x]$-integer coefficients, it suffices to show that we cannot factor x as $x = ab$ with a and b integral over $\mathbb{C}[x]$, and such that the following two elements of $\overline{K}$ are also integral over $\mathbb{C}[x]$:

$$\frac{a(7+y)}{x} \quad \text{and} \quad \frac{b(7-y)}{x}.$$

If $a(7+y)/x$ is an integral element, the ideal (x) must divide the ideal $(a)(7+y)$. From the factorizations of (x) and $(7+y)$ already computed, this means that (a) contains $(x, 7+y)$. Similarly, we must also have that (b) contains $(x, 7-y)$. As $(x) = (a)(b) = (x, 7+y)(x, 7-y)$, this implies that $(a) = (x, 7+y)$ and $(b) = (x, 7-y)$, where the ideals should now be interpreted in the ring of $\mathbb{C}[x]$-integers in some finite extension L of K.

Thus, in order to prove that $f(T)$ cannot factor into linear factors with integral coefficients, it suffices to prove that the ideal $(7+y, x)$ is not principal in the ring of $\mathbb{C}[x]$-integers of any finite extension of K.

[4]The multiplication of ideals in A turns out to be closely related to the group of points of C, described in much more generality in [19] and [20]. In particular, if a prime ideal $\mathfrak{p}$ corresponds to a point P on C, then $\mathfrak{p}^k$ is principal if and only if the order of P divides k in the group of points. Since only countably many of the uncountably many points on C have finite order, $\mathfrak{p}^k$ is almost always not principal. In particular, using techniques of arithmetic geometry, it is a straightforward matter to confirm that the point $(0, -7)$ has infinite order in the group of points on C, and therefore that the ideals $(x, 7+y)^k$ are not principal for any $k > 0$.

Just as in the number field case, it turns out that this is equivalent to showing that $(7+y,x)^k$ is not principal in A, which is precisely the content of Theorem 10.1. Thus, $f(T)$ cannot be factored into linear factors with integral coefficients.

In fact, notice that even though x divides the product $(7+y)(7-y)$ (since $49-y^2 = 3x-x^3$ in A), we cannot factor x into a product of an element that divides $7+y$ and one that divides $7-y$, even in the integral closure of A in $\overline{K}$. In other words, the integral closure of A in $\overline{K}$ is not a pre-Schreier domain, and so we cannot always lift factorizations from $\overline{K}$, as noted in Theorem 9.13.

11. Final remarks.

Question 1.1 goes to the very heart of Dedekind's vision of algebraic number theory: exploring not only the parallels between the rationals and number fields and between the integers and rings of integers, but also the circumstances under which those parallels break down. Our answer has taken us on a tour of the history of some of the fundamental building blocks of ring theory and algebraic number theory, and even when we thought we were leaving number theory behind for the wider field of ring theory, we found ourselves drawn back to the notions of factorizations and to Dedekind's work. It is easy to understand, then, Emmy Noether's dictum: *Es steht schon bei Dedekind.*

Acknowledgements.

The authors thank Derek Holt and George Bergman for helpful comments; we also thank Bill Dubuque for many suggestions, and particularly for bringing P. M. Cohn's paper to our attention. We are grateful to the anonymous referee for glimpsing some potential of the paper based on our original submission and suggesting that we build the paper around the historical investigation of factorizations.

References

[1] D. D. Anderson and R. O. Quintero, Some generalizations of GCD-domains, in *Factorization in Integral Domains (Iowa City, IA, 1996)*, Lecture Notes in Pure and Appl. Math., no. 189, Marcel Dekker, New York, 1997, 189–195.

[2] G. Birkhoff and S. Mac Lane, *A Survey of Modern Algebra*, 3rd ed., Mac Millan, New York, 1970.

[3] N. Bourbaki, *Elements of the History of Mathematics* (trans. J. Meldrum), Springer-Verlag, New York, 1991.

[4] P. M. Cohn, Bezout rings and their subrings. *Proc. Cambridge Philos. Soc.* **65** (1968) 251–264.

[5] R. Dedekind, *Theory of Algebraic Integers* (trans. J. Stillwell), Cambridge University Press, Cambridge, 1996.
[6] R. Dedekind and H. Weber, Theorie der algebraischen Functionen einer Veränderlichen. *J. Reine Angew. Math.* **92** (1882), 181–290.
[7] H. M. Edwards, *Fermat's Last Theorem: A Genetic Introduction to Algebraic Number Theory*, Springer-Verlag, New York, 1977.
[8] Euclid, *Elements, Book VII.*
[9] C. F. Gauss, *Disquisitiones Arithmeticae*, (trans. A. A. Clarke), Springer-Verlag, New York, 1986.
[10] R. Hartshorne, *Algebraic Geometry*, Springer-Verlag, New York, 1977.
[11] D. Hilbert, *The Theory of Algebraic Number Fields,* (trans. I. T. Adamson), Springer-Verlag, New York, 1998.
[12] N. Jacobson, *Basic Algebra II*, 2nd ed., W. H. Freeman, New York, 1989.
[13] I. Kaplansky, *Commutative Rings*, revised ed., University of Chicago Press, Chicago, 1974.
[14] E. E. Kummer, De numeris complexis, qui radicibus unitatis et numeris integris realibus constant. Gratulationschrift der Univ. Breslau zur Jubelfeier der Univ. Königsberg. Reprinted as: Sur les nombres complexes qui sont formés avec les nombres entiers réels et les racines de l'unité. *Journ. de Math.,* (1) v. XII (1847), 185–212.
[15] E. E. Kummer, Ueber die Zerlegung der aus Wurzeln der Einheit gebildeten complexen Zahlen in Primfactoren. *J. de Crelle,* v. XXXV (1847), 327–367.
[16] D. A. Marcus, *Number Fields*, Springer-Verlag, New York, 1977.
[17] S. McAdam and D. E. Rush, Schreier rings, *Bull. London Math. Soc.*, no. 10 (1978), 77–80.
[18] H. Pollard and H. G. Diamond, *The Theory of Algebraic Numbers*, 3rd ed., Dover, Mineola NY, 1998.
[19] J. H. Silverman, *The Arithmetic of Elliptic Curves*, Springer-Verlag, New York, 1986.
[20] J. H. Silverman and J. Tate, *Rational Points on Elliptic Curves*, Springer-Verlag, New York, 1992.

ARTURO MAGIDIN earned his degrees at the Universidad Nacional Autónoma de México (Matemático, 1993), and at UC Berkeley (Ph.D. 1998) as a student of George Bergman. He then spent four years at the Instituto de Matemáticas at the UNAM and has been at the University of Montana for three, only the second time in his life that he has lived north of his coauthor. Although his main research has been at the intersection of general algebra and group theory, he considers himself a "number theory groupie" and is always keen to try his hand at it when time permits.
Dept. of Mathematical Sciences, University of Montana, Missoula MT 59801-0864
magidin@member.ams.org

DAVID MCKINNON earned a perfectly respectable BA/MA degree in geometry from Harvard in 1992, whereafter he was almost immediately seduced by the lure of number theory. This disreputable path led him to a Ph.D. in 1999 from UC Berkeley under the delightful direction of Paul Vojta. Since then, he has held a postdoctoral position at Tufts and a tenure-track position at the University of Waterloo, and he was even lucky enough to become the husband of Jennifer and the father of Heather and Robert. He wonders if he is the only number theorist whose Erdős number depends on symplectic topologists.
Pure Mathematics Department, University of Waterloo, Waterloo, ON N2L 3G1
dmckinnon@math.uwaterloo.ca