

1. Fermat, Euler, and Nonunique Factorization

The purpose of this chapter is mainly motivational. More precise definitions of the objects we will study will be given later.

1.1 Euler and Quadratic Irrationals

Algebraic number theory was born when Euler used algebraic numbers to solve diophantine equations such as $y^2 = x^3 - 2$: Fermat had claimed that $(x, y) = (3, 5)$ is the only solution in natural numbers, and Euler gave a “proof” by writing

$$x^3 = y^2 + 2 = (y - \sqrt{-2})(y + \sqrt{-2}) \quad (1.1)$$

and working with the ring $\mathbb{Z}[\sqrt{-2}] = \{a + b\sqrt{-2} : a, b \in \mathbb{Z}\}$.

The problem with Euler’s idea was that he did not justify all of his claims. Arguing that the two factors on the right hand side of (1.1) were coprime¹, he concluded that each factor had to be a perfect cube,² i.e. that $y + \sqrt{-2} = (a + b\sqrt{-2})^3$ for certain $a, b \in \mathbb{Z}$. Comparing real and imaginary parts yields $y = a^3 - 6ab^2 = a(a^2 - 6b^2)$ and $1 = 3a^2b - 2b^3 = b(3a^2 - 2b^2)$. The last equation tells us that $b \mid 1$, hence $b = \pm 1$. Moreover, $3a^2 - 2b^2 = 1$, hence $a = \pm 1$. Plugging these solutions into $y = a(a^2 - 6b^2)$ shows that $y = \pm 5$ and thus $x = 3$, proving Fermat’s claim.

In order to understand why Euler’s argument is not sufficient, let us consider the diophantine equation $y^2 = x^2 - 5$. Imitating Euler’s proof, we find

$$x^2 = y^2 + 5 = (y - \sqrt{-5})(y + \sqrt{-5}). \quad (1.2)$$

Since the two factors are “coprime”, both of them must be squares; but from $y + \sqrt{-5} = (a + b\sqrt{-5})^2$ we get the equation $1 = 2ab$, which does not have any solutions in integers. This seems to suggest that $y^2 = x^2 - 5$ does not have any integral solutions; but actually $(x, y) = (3, 2)$ is one.³

¹ Here is the first problem: he does not really define what this means.

² This is the second problem: Euler knows that this argument works inside the natural numbers; in fact a proof can be found in Euclid’s elements. But Euler does not explain why this should work in $\mathbb{Z}[\sqrt{-2}]$.

³ Of course there is no need to invoke algebraic numbers for solving $y^2 = x^2 - 5$, because we can write the equation in the form $5 = x^2 - y^2 = (x - y)(x + y)$.

1.2 Fermat and the Harbingers of Nonunique Factorization

The examples above suggest the following question: why does an argument that works well for numbers of the form $a + b\sqrt{-2}$ go wrong for numbers of the form $a + b\sqrt{-5}$? The reason for this strange behavior would not be uncovered until the mid-19th century, although traces of it can be tracked back to the work of Fermat. One of his more famous theorems claims that every prime of the form $4n + 1$ can be written as the sum of two squares. The heart of Fermat's proof was the fact that a divisor of a number of the form $x^2 + y^2$ with $\gcd(x, y) = 1$ also can be represented in the form $x^2 + y^2$; thus from $5 \cdot 13 = 8^2 + 1^2$ we may conclude that 5 and 13 are sums of two squares.⁴

Fermat also found by induction that the same claim holds for the quadratic form $x^2 + 2y^2$; on the other hand he knew that it failed for $x^2 + 5y^2$ because

$$21 = 1^2 + 5 \cdot 2^2 = 4^2 + 5 \cdot 1^2, \quad (1.3)$$

yet 3 and 7 cannot be represented in this form.

The connection with Euler's use of quadratic irrationals becomes apparent when we write (1.3) in the form

$$3 \cdot 7 = (1 + 2\sqrt{-5})(1 - 2\sqrt{-5}) = (4 + \sqrt{-5})(4 + \sqrt{-5}). \quad (1.4)$$

We claim that the factors in these factorizations are all irreducible in the ring $R = \mathbb{Z}[\sqrt{-5}]$, and do not differ just by units.

Before we prove this, let us recall the relevant notations. In a domain R (a commutative ring with 1 and without zero divisors), we say that $b \mid a$ (b divides a) if there exists a $c \in R$ with $a = bc$. The divisors of 1 are called units and form a group R^\times .

A nonunit $p \in R \setminus R^\times$ is called

- irreducible if it only has trivial factorizations: $p = ab$ for $a, b \in R$;
- prime if $p \mid ab$ for any $a, b \in R$ implies that $p \mid a$ or $p \mid b$.

We know that primes are always irreducible, and that, in unique factorization domains, irreducibles are prime.

Before we can see why 3 is irreducible in $\mathbb{Z}[\sqrt{-5}]$ we have to determine the units of this ring. This is quite easy:

Proposition 1.1. *Let $m < -1$ be an integer; then the units of the ring $R = \mathbb{Z}[\sqrt{m}] = \{a + b\sqrt{-m} : a, b \in \mathbb{Z}\}$ are $R^\times = \{-1, +1\}$.*

In \mathbb{Z} , the prime 5 only has four possible divisors; going through all possibilities easily shows that $(\pm 3, \pm 2)$ are the only integral solutions.

⁴ From the modern point of view his proof essentially is a "translation" of the fact that $\mathbb{Z}[i]$ is Euclidean into a language avoiding algebraic numbers.

Before we prove this result, let us put $N(a + b\sqrt{m}) = (a + b\sqrt{m})(a - b\sqrt{m}) = a^2 - mb^2$; this is called the norm of $a + b\sqrt{m}$. It is clear from the definition that the norm is multiplicative, i.e., that $N(\alpha\beta) = N(\alpha)N(\beta)$. In fact, if we call $\alpha' = a - b\sqrt{m}$ the conjugate of $\alpha = a + b\sqrt{m}$, then a simple calculation shows that $N(\alpha\beta) = (\alpha\beta)(\alpha\beta)' = \alpha\alpha'\beta\beta' = N(\alpha)N(\beta)$.

Lemma 1.2. *An element $\varepsilon = a + b\sqrt{m} \in R = \mathbb{Z}[\sqrt{m}]$ (here m is a nonsquare integer) is a unit if and only if $N\varepsilon = \pm 1$.*

Proof. If ε is a unit, then there is some $\eta \in R$ with $\varepsilon\eta = 1$. Applying the norm gives $N(\varepsilon)N(\eta) = N(1) = 1$. This is an equation in \mathbb{Z} , hence $N(\varepsilon) = N(\eta) = \pm 1$.

Conversely, assume that $\varepsilon = a + b\sqrt{m} \in R$ satisfies $N(\varepsilon) = \pm 1$. Then $\frac{1}{\varepsilon} = \frac{a - b\sqrt{m}}{a^2 - mb^2} = \pm(a - b\sqrt{m}) =: \eta$ satisfies $\varepsilon\eta = 1$, hence $\varepsilon \in R^\times$. \square

Now we are ready to give the

Proof of Prop. 1.1. From Lemma 1.2 we know that $\varepsilon = a + b\sqrt{m} \in R^\times$ is a unit if and only if $N\varepsilon = a^2 - mb^2 = \pm 1$. Since $m < 0$, this is equivalent to $a^2 - mb^2 = 1$, and for $m < -1$ this holds if and only if $b = 0$ and $a = \pm 1$. \square

In particular, all the factors in (1.4) are nonunits. Now assume that 3 is reducible in R , i.e., there are nonunits $\alpha, \beta \in R$ with $3 = \alpha\beta$. Taking norms shows that $9 = N(3) = N(\alpha)N(\beta)$. Since α and β are nonunits, and since $N\alpha > 0$, we conclude that $N\alpha = N\beta = 3$. Writing $\alpha = a + b\sqrt{-5}$ shows that $3 = a^2 + 5b^2$; but this is impossible in integers.

The same line of reasoning shows that all the factors in (1.4) are irreducible. This is not enough to conclude that $\mathbb{Z}[\sqrt{-5}]$ does not have unique factorization. In fact, consider the factorizations

$$\sqrt{2} \cdot \sqrt{2} = (2 + \sqrt{2})(2 - \sqrt{2})$$

in $R = \mathbb{Z}[\sqrt{2}]$. All the factors in there are irreducible since their norm is ± 2 ; yet the two factorizations do not differ substantially because the factors differ by units: in fact, $2 + \sqrt{2} = \sqrt{2} \cdot (1 + \sqrt{2})$, and $\varepsilon = 1 + \sqrt{2}$ is a unit in R .

On the other hand, 3 and, say, $1 + 2\sqrt{-5}$ do not differ by a unit since their quotient $\frac{1}{3} + \frac{2}{3}\sqrt{-5}$ is not an element of R .

This shows that $\mathbb{Z}[\sqrt{-5}]$ does not have unique factorization, and that this is a consequence of Fermat's observation on divisors of numbers of the form $x^2 + 5y^2$. This fact is also responsible for the erroneous result that our second "proof" above has produced. If $\mathbb{Z}[\sqrt{-5}]$ had unique factorization, the given proof would actually be correct, as the following lemma shows:

Lemma 1.3. *Assume that R is a unique factorization domain. If $a, b \in R$ are coprime and if $ab = c^n$ for some $c \in R$, then there exists a unit $u \in R^\times$ and elements $r, s \in R$ such that $a = ur^n$ and $b = u^{-1}s^n$.*

Proof. Since R has unique factorization, a is the product of a unit and certain prime powers. Since a and b are coprime, these primes do not divide b ; since ab is an n -th power, the exponent of each prime in the factorization of a must be a multiple of n . This proves the claim. \square

This result does not hold in $\mathbb{Z}[\sqrt{-5}]$: here $(2 + \sqrt{-5})(2 - \sqrt{-5}) = 3^2$ is a square; since the factors $(2 + \sqrt{-5})$ and $(2 - \sqrt{-5})$ are irreducible (!) and do not differ by a unit, they must be coprime. Yet $\pm(2 + \sqrt{-5})$ is not a square (again because $2 + \sqrt{-5}$ is irreducible).

The insight that nonunique factorization is responsible for the failure of Euler's method in certain cases became common knowledge in the middle of the 19th century and is connected with the work of Dirichlet, Jacobi, Eisenstein, Liouville, Kummer, and Dedekind.

1.3 Dedekind's Ideals

How can we save unique factorization in rings like $\mathbb{Z}[\sqrt{-5}]$? In order to motivate the answer, consider Hilbert's example of the set of integers $M = \{1, 5, 9, \dots, 4n + 1, \dots\}$. In this monoid, the factorization $9 \cdot 49 = 21 \cdot 21$ shows that unique factorization does not hold. The different factorizations can, however, be explained by introducing the "ideal numbers" 3 and 7 and observing that $9 \cdot 49 = 21 \cdot 21$ comes from pairing up the factors in the ideal factorization $441 = 3^2 7^2$ in two different ways.

Now let us do the same in $\mathbb{Z}[\sqrt{-5}]$ by introducing the ideals. Recall that an ideal \mathfrak{a} in a ring R is a set closed with respect to addition and multiplication by ring elements:

$$\begin{aligned} a, b \in \mathfrak{a} &\implies a + b \in \mathfrak{a}; \\ a \in \mathfrak{a}, r \in R &\implies ra \in \mathfrak{a}. \end{aligned}$$

Given elements $a_1, \dots, a_n \in R$ we can define an ideal $\mathfrak{a} = (a_1, \dots, a_n) = \{\sum r_i a_i : r_i \in R\}$; ideals of the form $\mathfrak{a} = (a) = aR$ are called principal ideals.

Now ideals can be multiplied: we simply let $\mathfrak{a}\mathfrak{b} = \{\sum a_i b_i : a_i \in \mathfrak{a}, b_i \in \mathfrak{b}\}$ be the set of all finite sums of products of elements of \mathfrak{a} and \mathfrak{b} . In particular, this implies that e.g. $(a)(b) = (ab)$, $(a)(b_1, b_2) = (ab_1, ab_2)$, $(a_1, a_2)(b_1, b_2) = (a_1 b_1, a_1 b_2, a_2 b_1, a_2 b_2)$, etc. Moreover, the ideal $(1) = R$ consisting of all ring elements is a neutral element with respect to this multiplication.

The factorization (1.4) of elements in $R = \mathbb{Z}[\sqrt{-5}]$ immediately implies a corresponding factorization of principal ideals

$$(3) \cdot (7) = (1 + 2\sqrt{-5})(1 - 2\sqrt{-5}) = (4 + \sqrt{-5})(4 + \sqrt{-5}).$$

But whereas the elements in (1.4) were irreducible, the ideals are not. In fact, write $\mathfrak{p} = (3, 1 + \sqrt{-5})$, $\mathfrak{p}' = (3, 1 - \sqrt{-5})$, $\mathfrak{q} = (7, 4 + \sqrt{-5})$, and $\mathfrak{q}' = (7, 1 - \sqrt{-5})$. Then we find

$$\begin{aligned}\mathfrak{pp}' &= (9, 3(1 - \sqrt{-5}), 3(1 + \sqrt{-5}), 6) \\ &= (3)(3, 1 - \sqrt{-5}, 1 + \sqrt{-5}, 2) = (3)(1) = (3),\end{aligned}$$

because any ideal containing 3 and 2 also contains $3 - 2 = 1$ and hence is the unit ideal. Similarly we get $\mathfrak{qq}' = (7)$; this calculation is left to the reader. Thus we have $(3)(7) = (\mathfrak{pp}')(\mathfrak{qq}')$, and we may hope that the other factorizations can be explained similarly. This does work indeed:

$$\begin{aligned}\mathfrak{pq} &= (21, 3(4 + \sqrt{-5}), 7(1 + \sqrt{-5}), (1 + \sqrt{-5})(4 + \sqrt{-5})) \\ &= (4 + \sqrt{-5})(4 - \sqrt{-5}, 3 + \sqrt{-5}, 1 + \sqrt{-5}) \\ &= (4 + \sqrt{-5})\end{aligned}$$

because the ideal $(4 - \sqrt{-5}, 3 + \sqrt{-5}, 1 + \sqrt{-5})$ contains $7 = 4 - \sqrt{-5} + 3 + \sqrt{-5}$ and $2 = (3 + \sqrt{-5}) - (1 + \sqrt{-5})$, hence $1 = 7 - 3 \cdot 2$. Note that $(3 + \sqrt{-5})$ does not denote an ideal here: it must denote a number inside brackets because the left hand side 2 is a number, and because we have not defined the difference of ideals.

Similarly we find $\mathfrak{p}'\mathfrak{q}' = (4 + \sqrt{-5})$, $\mathfrak{pq}' = (1 - 2\sqrt{-5})$, and $\mathfrak{p}'\mathfrak{q} = (1 + 2\sqrt{-5})$. Thus the nonunique factorization of elements in (1.4) turns into the equality

$$(21) = \mathfrak{pp}'\mathfrak{qq}'$$

of ideals, from which the factorizations of principal ideals

$$(21) = (3)(7) = (1 + 2\sqrt{-5})(1 - 2\sqrt{-5}) = (4 + \sqrt{-5})(4 + \sqrt{-5})$$

results from pairing the ideals \mathfrak{p} , \mathfrak{p}' , \mathfrak{q} and \mathfrak{q}' in different ways.

The first goal now is to prove that this is not accidental, and that factorization into prime ideals holds in any ring of integers of an algebraic number field. This can be shown in various degrees of abstraction. In the next chapter, we give a down and dirty way of doing this in quadratic number fields.

Exercises

- 1.1 Find units $\neq \pm 1$ in the rings $\mathbb{Z}[\sqrt{m}]$ for $m = -1, 2, 3, 5, 6, 7$.
- 1.2 Prove that $(a_1, a_2)(b_1, b_2) = (a_1b_1, a_1b_2, a_2b_1, a_2b_2)$ for ideals in some commutative ring. Generalize.
- 1.3 Show that $6 = (1 + \sqrt{-5})(1 - \sqrt{-5}) = 2 \cdot 3$ is another example of nonunique factorization in $\mathbb{Z}[\sqrt{-5}]$.
- 1.4 Show that $6 = 2 \cdot 3 = (2 + \sqrt{-2})(2 - \sqrt{-2})$ is not an example of nonunique factorization in $\mathbb{Z}[\sqrt{-2}]$.
- 1.5 Explain the different factorizations in Exercise 1 using the ideals $\mathfrak{p} = (2, 1 + \sqrt{-5})$, $\mathfrak{q} = (3, 1 + \sqrt{-5})$, and $\mathfrak{q}' = (3, 1 - \sqrt{-5})$. Show that

1. $(2, 1 - \sqrt{-5}) = \mathfrak{p}$;
2. $\mathfrak{p}^2 = (2)$;
3. $\mathfrak{q}\mathfrak{q}' = (3)$;
4. $\mathfrak{q}^2 = (2 + \sqrt{-5})$.

1.6 Discuss the factorizations $6 = 2 \cdot 3 = -\sqrt{-6}^2$ in $\mathbb{Z}[\sqrt{-6}]$ and $6 = 2 \cdot 3 = (2 + \sqrt{10})(-2 + \sqrt{10})$ in $\mathbb{Z}[\sqrt{10}]$.

1.7 Prove that the only unique factorization domains of the form $\mathbb{Z}[\sqrt{m}]$ with $m \leq 1$ are those for $m = 1$ and $m = 2$.

Hints. First consider the case $m \equiv 2 \pmod{4}$. If $m > 2$, it is composite, say $m = ab$. Now consider the factorizations $m = ab = -\sqrt{-m}^2$. If m is odd and $m \neq 1$, then have a look at the factorization of $m + 1$.

1.8 The last exercise showed that unique factorization domains are rare among the rings $\mathbb{Z}[\sqrt{m}]$ with $m \leq 1$. The situation is better for $m > 1$; nevertheless show that $\mathbb{Z}[\sqrt{m}]$ does not have unique factorization if $m = 2n$ with $n \equiv 1 \pmod{4}$. Does the proof also work if $n \equiv 3 \pmod{4}$?