

Introduction

We shall here be concerned with the circle of ideas that surrounds the *Fundamental Theorem of Arithmetic*.

First we recall the usual definition of a prime: a *prime number* is a positive integer, other than 1, that has no divisors except itself and 1. For example 2 and 3 are primes, but $6 = 2 \cdot 3$ and $10 = 2 \cdot 5$ are not.

Then the Fundamental Theorem of Arithmetic states that *every positive integer can be factored into primes in an essentially unique way*. For example,

$$\begin{aligned}1 &= 1, \\2 &= 2, \\6 &= 2 \cdot 3, \\10 &= 2 \cdot 5, \\15 &= 3 \cdot 5, \\2499 &= 3 \cdot 7^2 \cdot 17.\end{aligned}$$

By “essentially unique,” we mean unique up to the order of the factors, so that we consider $6 = 2 \cdot 3 = 3 \cdot 2$ to be the same factorization. (Note that 1 is a special case. We think of it as having an “empty” factorization, as it is not divisible by any prime.)

As its name implies, unique factorization is a fundamental property of the positive integers, a property that was known to the ancient Greeks. We will prove this property, and indeed our proof will follow that of Euclid. But we will be interested in examining this proof and seeing what makes it really “work,” with an idea of seeing when we can extend it to more general situations.

For example, let us consider numbers of the form $a + b\sqrt{-1}$ with a and b integers. It turns out, and we shall prove, that numbers of this form also have unique factorization. For example, we have the following factorization

into primes for numbers of this form:

$$\begin{aligned} 3 &= 3, \\ 5 &= (2 + \sqrt{-1})(2 - \sqrt{-1}), \\ 7 &= 7, \\ 11 &= 11, \\ 13 &= (3 + 2\sqrt{-1})(3 - 2\sqrt{-1}), \\ 17 &= (4 + \sqrt{-1})(4 - \sqrt{-1}). \end{aligned}$$

On the other hand, let us consider numbers of the form $a + b\sqrt{-5}$ with a and b integers. Numbers of this form do *not* have unique factorization. For example, we have the following two factorizations of 6 into irreducibles:

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

We can also consider numbers of the form $a + b\sqrt{10}$ with a and b integers. Numbers of this form also do *not* have unique factorization. For example, we have the following two factorizations of 10 into irreducibles:

$$14 = (2)(5) = (\sqrt{10})^2.$$

We have used the word “irreducible” rather than “prime” here as that turns out to be the correct mathematical language.

In fact, we will prove the Fundamental Theorem of Arithmetic in a way that enables us to establish it in many cases, including the two we have mentioned—the ordinary integers, and numbers of the form $a + b\sqrt{-1}$ with a and b integers—simultaneously.

On the other hand, we will also be able to systematically show that in many cases, including the two we have mentioned—numbers of the form $a + b\sqrt{-5}$ with a and b integers, and numbers of the form $a + b\sqrt{10}$ with a and b integers—unique factorization does not hold.

As we will see, instead of unique factorization being the norm and non-unique factorization the exception, the situation is reversed! It is really a very special property, though a crucially important one, of the ordinary integers that the Fundamental Theorem of Arithmetic holds for them.

Chapters 1 and 2 of this book are basically devoted to proving unique and nonunique factorization for ordinary integers and for numbers of the form $a + b\sqrt{D}$. (Here a and b are not always integers, but this is a technical point we will defer until later.)

In Chapter 3, we investigate numbers of the form $a + b\sqrt{-1}$ with a and b integers. Numbers of this form are called the *Gaussian integers*. As we

have remarked, in the Gaussian integers we do have unique factorization into primes, but we would like to know what the primes are. Here we will show that the following is always true (compare the factorizations above): every ordinary prime that leaves a remainder of 3 when divided by 4 remains a prime in the Gaussian integers, but every prime that leaves a remainder of 1 when divided by 4 factors into a product of two “conjugate” primes in the Gaussian integers. In fact, this is closely related to a famous theorem of Fermat: *every prime that leaves a remainder of 1 when divided by 4 can be written as a sum of two squares*. (For example, $5 = 2^2 + 1^2$, $13 = 3^2 + 2^2$, $17 = 4^2 + 1^2$, but also $97 = 8^2 + 5^2$, $101 = 10^2 + 1^2$, and $99989 = 230^2 + 217^2$.)

Actually, we will give several proofs of this theorem. One, due to Euler, is believed to be along the lines of Fermat’s original proof (which he never wrote down for posterity). It uses a technique known as *composition*. Another one uses unique factorization to prove Fermat’s theorem quickly and easily. (It is a bit surprising that this abstract result gives such a concrete fact, but mathematics is full of beautiful surprises.)

To describe our next objective, we have to get a bit more technical. The ancient Greeks considered the positive integers, but when we wish to generalize our investigations, we no longer have the idea of positivity. (We cannot make any sense out of saying that $a + b\sqrt{-1}$ is positive.) So we have to consider all integers. But when we do, we see that we have typical factorizations:

$$6 = (2)(3) = (-1)(-2)(3) = (-2)(-3).$$

These are different factorizations, but we do not want to consider these to be essentially different. How do they differ? The answer is that 1 can be factored as $1 = (-1)(-1)$ and we are simply distributing the factors of 1 differently. We give a name to this situation. Factors of 1 are called *units*, and factorizations that differ merely because we have redistributed the units are essentially the same.

We can show that the units in the Gaussian integers are precisely 1, -1 , $\sqrt{-1}$, and $-\sqrt{-1}$. Then we also have the prime factorization in the Gaussian integers:

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

Here the first factor is a unit, so what we see is that, up to a unit factor, 2 is a square in the Gaussian integers.

In factoring numbers, we do not really care about units, but still it is an interesting question—indeed, a very interesting question—to ask what

the units are. We have given the answer for the Gaussian integers, but we can ask the same question in other cases as well. Here we ask the question for numbers of the form $a + b\sqrt{D}$ for D positive and not a perfect square. If $D = 2$ we have units

$$\begin{aligned} 1 &= (3 + 2\sqrt{2})(3 - 2\sqrt{2}), \\ 1 &= (17 + 12\sqrt{2})(17 - 12\sqrt{2}), \\ 1 &= (99 + 70\sqrt{2})(99 - 70\sqrt{2}), \\ 1 &= (577 + 408\sqrt{2})(577 - 408\sqrt{2}). \end{aligned}$$

Note that a factorization $1 = (a + b\sqrt{D})(a - b\sqrt{D})$ gives a solution of the equation $a^2 - b^2D = 1$, and vice versa. Thus the search for units is intimately related to the search for solutions of the equation $a^2 - b^2D = 1$. The units above correspond to solutions for $D = 2$: $1 = 3^2 - 2^2 \cdot 2 = 17^2 - 12^2 \cdot 2 = 99^2 - 70^2 \cdot 2 = 577^2 - 408^2 \cdot 2$. But we can consider this equation for other values of D as well. For example, for $D = 61$ we have the solution

$$1 = (1766319049)^2 - (226153980)^2 \cdot 61$$

and for $D = 109$ we have the solution

$$1 = (158070671986249)^2 - (15140424455100)^2 \cdot 109.$$

In fact, for any such D there are infinitely many solutions (and hence infinitely many units). We shall prove this in Chapter 4 where we investigate the equation $a^2 - b^2D = 1$, known as *Pell's equation*. Our proof is a variant of the *cakravala* method experimentally developed by Indian mathematicians between the ninth and twelfth centuries. This is also a result known to Fermat, and his proof of this result may well have been along the lines of ours, as our proof uses a method of “composition” very closely related to our method in Chapter 3. Also, our proof is constructive, enabling us to find solutions by hand for values of D that are not too large. The above solutions for $D = 61$ and for $D = 109$ were known to have been found by Fermat (by hand, obviously, since computers did not exist in the seventeenth century).

Our investigations in Chapters 1 through 4 can be considerably generalized. To use the appropriate technical language, in these chapters we are considering *quadratic fields*, and we can consider analogous problems for *algebraic number fields*. Indeed, our treatment here parallels the historical development of the subject. Quadratic fields were investigated first, and the phenomena that arose there motivated the development of the general

theory. This subject is known as *algebraic number theory*. In Chapter 5 we survey some of the highlights of this subject. As we have seen, unique factorization of elements holds in the integers \mathbb{Z} , but it does not always hold. While unique factorization of elements is the most straightforward generalization of the situation in the integers, it is not the right generalization. The right generalization is unique factorization of ideals, which does hold. Therefore in Chapter 5 we will focus (though not exclusively) on ideals in general. But we will also provide a wealth of examples, interesting in themselves, that show how quadratic fields fit into the general case. For a precise description of the scope of our investigations in this chapter, we simply refer the reader to the table of contents.

We have three appendices. Appendix A is a careful treatment of mathematical induction, an essential proof technique. Appendix B is a treatment of congruences. Here we begin with the definition, and proceed through linear congruences (including the Chinese Remainder Theorem) and quadratic congruences (including the Law of Quadratic Reciprocity). Appendix C is a technical one, dealing with some of the more complicated cases of results in Chapter 2.