Factorization and the Fundamental Theorem of Arithmetic

Factorization and the Fundamental Theorem of Arithmetic

A *prime number* is a number $p \geq 2$ whose only (positive) divisors are 1 and p. Numbers $m \geq 2$ that are not primes are called *composite numbers*. For example,

prime numbers	$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, \ldots$
composite numbers	$4, 6, 8, 9, 10, 12, 14, 15, 16, 18, 20, \ldots$

Prime numbers are characterized by the numbers by which they are divisible; that is, they are defined by the property that they are only divisible by 1 and by themselves. So it is not immediately clear that primes numbers should have special properties that involve the numbers that they divide. Thus the following fact concerning prime numbers is both nonobvious and important.[1]

Lemma 1. *Let p be a prime number, and suppose that p divides the product ab. Then either p divides a or p divides b (or p divides both a and b).*[2]

Proof. We are given that p divides the product ab. If p divides a, we are done, so we may as well assume that p does not divide a. Now consider what $\gcd(p, a)$ can be. It divides p, so it is either 1 or p. It also divides a, so it isn't p, since we have assumed that p does not divide a. Thus, $\gcd(p, a)$ must equal 1.

[1] A *lemma* is a result that is used as a stepping stone for proving other results.

[2] You may say that this lemma is obvious if we look at the prime factorizations of a and b. However, the fact that a number can be factored into a product of primes in exactly one way is itself a nonobvious fact. We will discuss this further later in this chapter.

Now we use the Linear Equation Theorem with the numbers p and a. The Linear Equation Theorem says that we can find integers x and y that solve the equation

$$px + ay = 1.$$

[Note that we are using the fact that $\gcd(p, a) = 1$.] Now multiply both sides of the equation by b. This gives

$$pbx + aby = b.$$

Certainly pbx is divisible by p, and also aby is divisible by p, since we know that p divides ab. It follows that p divides the sum

$$pbx + aby,$$

so p divides b. This completes the proof of the lemma.[3] □

The lemma says that if a prime divides a product ab, it must divide one of the factors. Notice that this is a special property of prime numbers; it is not true for composite numbers. For example, 6 divides the product $15 \cdot 14$, but 6 divides neither 15 nor 14. It is not hard to extend the lemma to products with more than two factors.

Theorem 2 (Prime Divisibility Property)**.** *Let p be a prime number, and suppose that p divides the product $a_1 a_2 \cdots a_r$. Then p divides at least one of the factors $a_1, a_2, \ldots, a_r$.*

Proof. If p divides a_1, we're done. If not, we apply the lemma to the product

$$a_1(a_2 a_3 \cdots a_r)$$

to conclude that p must divide $a_2 a_3 \cdots a_r$. In other words, we are applying the lemma with $a = a_1$ and $b = a_2 a_3 \cdots a_r$. We know that $p|ab$, so if $p \nmid a$, the lemma says that p must divide b.

So now we know that p divides $a_2 a_3 \cdots a_r$. If p divides a_2, we're done. If not, we apply the lemma to the product $a_2(a_3 \cdots a_r)$ to conclude that p must divide $a_3 \cdots a_r$. Continuing in this fashion, we must eventually find some a_i that is divisible by p. □

[3] When we are proving a statement, we use a little box □ to indicate that we have completed the proof. Some books instead use QED to indicate the end of a proof. The letters QED stand for the Latin phrase *Quod erat demonstrandum*, which roughly means "that which was to be proved." This in turn comes from the Greek phrase $\omega\pi\epsilon\rho\ \epsilon\delta\epsilon\iota\ \delta\epsilon\iota\xi\alpha\iota$, which appears in Euclid's *Elements*.

Later in this chapter we are going to use the Prime Divisibility Property to *prove* that every positive integer can be factored as a product of prime numbers in essentially one way. Unfortunately, this important fact is so familiar to most readers that they will question why it requires a proof. So before giving the proof, I want to try to convince you that unique factorization into primes is far from being obvious. For this purpose, I invite you to leave the familiar behind and enter the[4]

Even Number World

(popularly known as the "$\mathbb{E}$-Zone")

Imagine yourself in a world where the only numbers that are known are the even numbers. So, in this world, the only numbers that exist are

$$\mathbb{E} = \{\ldots, -8, -6, -4, -2, 0, 2, 4, 6, 8, 10, \ldots\}.$$

Notice that in the $\mathbb{E}$-Zone we can add, subtract, and multiply numbers just as usual, since the sum, difference, and product of even numbers are again even numbers. We can also talk about divisibility. We say that a number m $\mathbb{E}$-*divides* a number n if there is a number k with $n = mk$. But remember that we're now in the $\mathbb{E}$-Zone, so the word "number" means an even number. For example, 6 $\mathbb{E}$-divides 12, since $12 = 6 \cdot 2$; but 6 does not $\mathbb{E}$-divide 18, since there is no (even) number k satisfying $18 = 6k$.

We can also talk about primes. We say that an (even) number p is an $\mathbb{E}$-prime if it is not divisible by any (even) numbers. (In the $\mathbb{E}$-Zone, a number is not divisible by itself!) For example, here are some $\mathbb{E}$-primes:

$$2,\ 6,\ 10,\ 14,\ 18,\ 22,\ 26,\ 30.$$

Recall the lemma we proved above for ordinary numbers. We showed that if a prime p divides a product ab then either p divides a or p divides b. Now move to the $\mathbb{E}$-Zone and consider the $\mathbb{E}$-prime 6 and the numbers $a = 10$ and $b = 18$. The number 6 $\mathbb{E}$-divides $ab = 180$, since $180 = 6 \cdot 30$; but 6 $\mathbb{E}$-divides neither 10 nor 18. So our "obvious" lemma is not true here in the $\mathbb{E}$-Zone!

There are other "self-evident facts" that are untrue in the $\mathbb{E}$-Zone. For example, consider the fact that every number can be factored as a product of primes in exactly one way. (Of course, rearranging the order of the factors is not considered a different factorization.) It's not hard to show, even in the $\mathbb{E}$-Zone, that every (even) number can be written as a product of $\mathbb{E}$-primes. But consider the following factorizations:

$$180 = 6 \cdot 30 = 10 \cdot 18.$$

[4] Since this book is not a multimedia product, you'll have to use your imagination to supply the appropriate Twilight Zone music.

Notice that all of the numbers 6, 30, 10, and 18 are $\mathbb{E}$-primes. This means that 180 can be written as a product of $\mathbb{E}$-primes in two fundamentally different ways! In fact, there is even a third way to write it as a product of $\mathbb{E}$-primes,

$$180 = 2 \cdot 90.$$

We are going to leave the $\mathbb{E}$-Zone now and return to the familiar world where odd and even numbers live together in peace and harmony. But we hope that our excursion into the $\mathbb{E}$-Zone has convinced you that facts that seem obvious require a healthy dose of skepticism. Especially, any "fact" that "must be true" because it is very familiar or because it is frequently proclaimed to be true is a fact that needs the most careful scrutiny.[5]

$\mathbb{E}$-Zone Border Crossing—Welcome Back Home

Everyone "knows" that a positive integer can be factored into a product of primes in exactly one way. But our visit to the $\mathbb{E}$-Zone provides convincing evidence that this obvious assertion requires a careful proof.

Theorem 3 (The Fundamental Theorem of Arithmetic)**.** *Every integer $n \geq 2$ can be factored into a product of primes*

$$n = p_1 p_2 \cdots p_r$$

in exactly one way.

Before we commence the proof of the Fundamental Theorem of Arithmetic, a few comments are in order. First, if n itself is prime, then we just write $n = n$ and consider this to be a product consisting of a single number. Second, when we write $n = p_1 p_2 \cdots p_r$, we do not mean that $p_1, p_2, \ldots, p_r$ have to be different primes. For example, we would write $300 = 2 \cdot 2 \cdot 3 \cdot 5 \cdot 5$. Third, when we say that n can be written as a product in exactly one way, we do not consider rearrangement of the factors to be a new factorization. For example, $12 = 2 \cdot 2 \cdot 3$ and $12 = 2 \cdot 3 \cdot 2$ and $12 = 3 \cdot 2 \cdot 2$, but all these are treated as the same factorization.

Proof. The Fundamental Theorem of Arithmetic really contains two assertions.

Assertion 1. The number n can be factored into a product of primes in some way.

Assertion 2. There is only one such factorization (aside from rearranging the factors).

[5]The principle that well-known and frequently asserted "facts" should be carefully scrutinized also applies to endeavors far removed from mathematics. Politics and journalism come to mind, and the reader will undoubtedly be able to add many others to the list.

We begin with Assertion 1. We are going to give a proof by induction. Don't let this scare you, it just means that first we'll verify the assertion for $n = 2$, and then for $n = 3$, and then for $n = 4$, and so on. We begin by observing that $2 = 2$ and $3 = 3$ and $4 = 2^2$, so each of these numbers can be written as a product of primes. This verifies Assertion 1 for $n = 2, 3, 4$. Now suppose that we've verified Assertion 1 for every n up to some number, call it N. This means we know that every number $n \leq N$ can be factored into a product of primes. Now we'll check that the same is true of $N + 1$.

There are two possibilities. First, $N + 1$ may already be prime, in which case it is its own factorization into primes. Second, $N + 1$ may be composite, which means that it can be factored as $N + 1 = n_1 n_2$ with $2 \leq n_1, n_2 \leq N$. But we know Assertion 1 is true for n_1 and n_2, since they are both less than or equal to N. This means that both n_1 and n_2 can be written as a product of primes, say

$$n_1 = p_1 p_2 \cdots p_r \qquad \text{and} \qquad n_2 = q_1 q_2 \cdots q_s.$$

Multiplying these two products together gives

$$N + 1 = n_1 n_2 = p_1 p_2 \cdots p_r q_1 q_2 \cdots q_s,$$

so $N + 1$ can be factored into a product of primes. This means that Assertion 1 is true for $N + 1$.

To recapitulate, we have shown that if Assertion 1 is true for all numbers less than or equal to N, then it is also true for $N + 1$. But we have checked it is true for 2, 3, and 4, so taking $N = 4$, we see that it is also true for 5. But then we can take $N = 5$ to conclude that it is true for 6. Taking $N = 6$, we see that it is true for $N = 7$, and so on. Since we can continue this process indefinitely, it follows that Assertion 1 is true for every integer.

Next we tackle Assertion 2. It is possible to give an induction proof for this assertion, too, but we will proceed more directly. Suppose that we are able to factor n as a product of primes in two ways, say

$$n = p_1 p_2 p_3 p_4 \cdots p_r = q_1 q_2 q_3 q_4 \cdots q_s.$$

We need to check that the factorizations are the same, possibly after rearranging the order of the factors. We first observe that $p_1 | n$, so $p_1 | q_1 q_2 \cdots q_s$. The Prime Divisibility Property proved earlier in this chapter tells us that p_1 must divide (at least) one of the q_i's, so if we rearrange the q_i's, we can arrange matters so that $p_1 | q_1$. But q_1 is also a prime number, so its only divisors are 1 and q_1. Therefore, we must have $p_1 = q_1$.

Now we cancel p_1 (which is the same as q_1) from both sides of the equation. This gives the equation

$$p_2 p_3 p_4 \cdots p_r = q_2 q_3 q_4 \cdots q_s.$$

Briefly repeating the same argument, we note that p_2 divides the left-hand side of this equation, so p_2 divides the right-hand side, and hence by the Prime Divisibility Property, p_2 divides one of the q_i's. After rearranging the factors, we get $p_2|q_2$, and then the fact that q_2 is prime means that $p_2 = q_2$. This allows us to cancel p_2 (which equals q_2) to obtain the new equation

$$p_3 p_4 \cdots p_r = q_3 q_4 \cdots q_s.$$

We can continue in this fashion until either all the p_i's or all the q_i's are gone. But if all the p_i's are gone, then the left-hand side of the equation equals 1, so there cannot be any q_i's left, either. Similarly, if the q_i's are all gone, then the p_i's must all be gone. In other words, the number of p_i's must be the same as the number of q_i's. To recapitulate, we have shown that if

$$n = p_1 p_2 p_3 p_4 \cdots p_r = q_1 q_2 q_3 q_4 \cdots q_s,$$

where all the p_i's and q_i's are primes, then $r = s$, and we can rearrange the q_i's so that

$$p_1 = q_1 \quad \text{and} \quad p_2 = q_2 \quad \text{and} \quad p_3 = q_3 \quad \text{and} \quad \ldots \quad \text{and} \quad p_r = q_s.$$

This completes the proof that there is only one way to write n as a product of primes. □

The Fundamental Theorem of Arithmetic says that every integer $n \geq 2$ can be written as a product of prime numbers. Suppose we are given a particular integer n. As a practical matter, how can we write it as a product of primes? If n is fairly small (for example, $n = 180$) we can factor it by inspection,

$$180 = 2 \cdot 90 = 2 \cdot 2 \cdot 45 = 2 \cdot 2 \cdot 3 \cdot 15 = 2 \cdot 2 \cdot 3 \cdot 3 \cdot 5.$$

If n is larger (for example, $n = 9105293$) it may be more difficult to find a factorization. One method is to try dividing n by primes $2, 3, 5, 7, 11, \ldots$ until we find a divisor. For $n = 9105293$, we find after some work that the smallest prime dividing n is 37. We factor out the 37,

$$9105293 = 37 \cdot 246089,$$

and continue checking $37, 41, 43, \ldots$ to find a prime that divides 246089. We find that $43|246089$, since $246089 = 43{\cdot}5723$. And so on until we factor $5723 = 59{\cdot}97$, where we recognize that 59 and 97 are both primes. This gives the complete prime factorization

$$9105293 = 37 \cdot 43 \cdot 59 \cdot 97.$$

If n is not itself prime, then there must be a prime $p \leq \sqrt{n}$ that divides n. To see why this is true, we observe that if p is the smallest prime that divides n, then $n = pm$ with $m \geq p$, and hence $n = pm \geq p^2$. Taking the square root of both sides yields $\sqrt{n} \geq p$. This gives the following foolproof method for writing any number n as a product of primes:

> To write n as a product of primes, try dividing it by every number (or just every prime number) $2, 3, \ldots$ that is less than or equal to $\sqrt{n}$. If you find no numbers that divide n, then n itself is prime. Otherwise, the first divisor that you find will be a prime p. Factor $n = pm$ and repeat the process with m.

This procedure, although fairly inefficient, works fine on a computer for numbers that are moderately large, say up to 10 digits. But how about a number like $n = 10^{128} + 1$? If n turns out to be prime, we won't find out until we've checked $\sqrt{n} \approx 10^{64}$ possible divisors. This is completely infeasible. If we could check $1{,}000{,}000{,}000$ (that's one billion) possible divisors each second, it would still take approximately $3 \cdot 10^{48}$ years! This leads to the following two closely related questions:

Question 1. How can we tell if a given number n is prime or composite?

Question 2. If n is composite, how can we factor it into primes?

Although it might seem that these questions are the same, it turns out that Question 1 is much easier to answer than Question 2. We will later see how to write down large numbers that we know are composite, even though we will be unable to write down any of their factors. In a similar fashion, we will be able to find very large prime numbers p and q such that, if we were to send someone the value of the product $n = pq$, they would be unable to factor n to retrieve the numbers p and q. This curious fact, that it is very easy to multiply two numbers but very difficult to factor the product, lies at the heart of a remarkable application of number theory to the creation of very secure codes.

Exercises

1. Suppose that $\gcd(a, b) = 1$, and suppose further that a divides the product bc. Show that a must divide c.

2. Suppose that $\gcd(a, b) = 1$, and suppose further that a divides c and that b divides c. Show that the product ab must divide c.

3. Let s and t be odd integers with $s > t \geq 1$ and $\gcd(s, t) = 1$. Prove that the three numbers

$$st, \qquad \frac{s^2 - t^2}{2}, \qquad \text{and} \qquad \frac{s^2 + t^2}{2}$$

are pairwise relatively prime; that is, each pair of them is relatively prime. This fact was needed to complete the proof of the Pythagorean triples theorem. [*Hint*. Assume that there is a common prime factor and use the fact (Lemma 1) that if a prime divides a product, then it divides one of the factors.]

4. Give a proof by induction of each of the following formulas. [Notice that (a) is the formula that we can prove using a geometric argument and that (c) is the first n terms of the geometric series.]

(a) $1 + 2 + 3 + \cdot + n = \dfrac{n(n+1)}{2}$

(b) $1^2 + 2^2 + 3^2 + \cdots + n^2 = \dfrac{n(n+1)(2n+1)}{6}$

(c) $1 + a + a^2 + a^3 + \cdots + a^n = \dfrac{1 - a^{n+1}}{1 - a} \qquad (a \neq 1)$

(d) $\dfrac{1}{1 \cdot 2} + \dfrac{1}{2 \cdot 3} + \dfrac{1}{3 \cdot 4} + \cdots + \dfrac{1}{(n-1)n} = \dfrac{n-1}{n}$

5. This exercise asks you to continue the investigation of the $\mathbb{E}$-Zone. Remember as you work that for the purposes of this exercise, odd numbers do not exist!

(a) Describe all $\mathbb{E}$-primes.

(b) Show that every even number can be factored as a product of $\mathbb{E}$-primes. [*Hint*. Mimic our proof of this fact for ordinary numbers.]

(c) We saw that 180 has three different factorizations as a product of $\mathbb{E}$-primes. Find the smallest number that has two different factorizations as a product of $\mathbb{E}$-primes. Is 180 the smallest number with three factorizations? Find the smallest number with four factorizations.

(d) The number 12 has only one factorization as a product of $\mathbb{E}$-primes: $12 = 2 \cdot 6$. (As usual, we consider $2 \cdot 6$ and $6 \cdot 2$ to be the same factorization.) Describe all even numbers that have only one factorization as a product of $\mathbb{E}$-primes.

6. Welcome to $\mathbb{M}$-World, where the only numbers that exist are positive integers that leave a remainder of 1 when divided by 4. In other words, the only $\mathbb{M}$-numbers that exist are

$$\{1, 5, 9, 13, 17, 21, \ldots\}.$$

(Another description is that these are the numbers of the form $4t + 1$ for $t = 0, 1, 2, \ldots$.) In the $\mathbb{M}$-World, we cannot add numbers, but we can multiply them, since if a and b both leave a remainder of 1 when divided by 4 then so does their product. (Do you see why this is true?)

We say that m $\mathbb{M}$-*divides* n if $n = mk$ for some $\mathbb{M}$-number k. And we say that n *is an* $\mathbb{M}$-*prime* if its only $\mathbb{M}$-divisors are 1 and itself. (Of course, we don't consider 1 itself to be an $\mathbb{M}$-prime.)

(a) Find the first six $\mathbb{M}$-primes.

(b) Find an $\mathbb{M}$-number n that has two *different* factorizations as a product of $\mathbb{M}$-primes.

7.  In this exercise you are asked to write programs to factor a (positive) integer n into a product of primes. (If $n = 0$, be sure to return an error message instead of going into an infinite loop!) A convenient way to represent the factorization of n is as a $2 \times r$ matrix. Thus, if

$$n = p_1^{k_1} p_2^{k_2} \cdots p_r^{k_r},$$

then store the factorization of n as the matrix

$$\begin{pmatrix} p_1 & p_2 & \cdots & p_r \\ k_1 & k_2 & \cdots & k_r \end{pmatrix}.$$

(If your programming language doesn't allow dynamic storage allocation, you'll have to decide ahead of time how many factors to allow.)

(a) Write a program to factor n by trying each possible factor $d = 2, 3, 4, 5, 6, \ldots$. (This is an extremely inefficient method but will serve as a warm-up exercise.)

(b) Modify your program by storing the values of the first 100 (or more) primes and first removing these primes from n before looking for larger prime factors. You can speed up your program when trying larger d's as potential factors if you don't bother checking d's that are even, or divisible by 3, or by 5. You can also increase efficiency by using the fact that a number m is prime if it is not divisible by any number between 2 and $\sqrt{m}$. Use your program to find the complete factorization of all numbers between 1,000,000 and 1,000,030.

(c) Write a subroutine that prints the factorization of n in a nice format. Optimally, the exponents should appear as exponents; but if this is not possible, then print the factorization of (say) $n = 75460 = 2^2 \cdot 5 \cdot 7^3 \cdot 11$ as

```
2^2 * 5 * 7^3 * 11.
```

(To make the output easier to read, don't print exponents that equal 1.)