Prime factorization

We will prove that every integer n which is not prime has a *unique* prime factorization. (Normally, we do not worry about the factors of 1 and n itself).

Of course this makes good sense. If a number is even, with its last digit in $\{0, 2, 4, 6, 8\}$, then it is even and that's the end of it. If we should imagine that there were two *different* factorizations, surely 2 must be included in each of them. But this is true for every factor of n.

This theorem is also called *the fundamental theorem of arithmetic*. As the name says, it's fundamental, so it will be worth some effort for the proof. The classic approach from Euclid is straightforward, but relies on some tricky preliminary work. The approach given here is from the famous English mathematician, G.H. Hardy.

The list of the prime factors of n goes like this:

$$n = p_1 \cdot p_2 \ldots p_k$$

A factor may be repeated. An example is:

$$12 = 2 \cdot 2 \cdot 3$$

To compare two factorizations for whether they are the same or different, we suppose they are sorted (say, from smallest to largest).

More examples for relatively small numbers:

```
  39 = 3.13
 144 = 2.2.2.2.3.3
 210 = 2.3.5.7
2310 = 2.3.5.7.11
```

Sometimes the factors can be hard to find.

Example: let's try 123456789. By using the digit addition trick, we can tell that this number is divisible by 9 (I get $9 + 9 + 9 + 9 + 9$).

At first it seems easy. I found two factors of 3, leaving 13717421.

I was forced to use Python to show that 3607 and 3803 are the two prime factors of that number, which once found, are easily confirmed. Factoring is a hard problem.

preliminary

Before starting the proof, some background. When we say that one integer *evenly divides* another, we write $a|n$ (a is a factor of n or a divides n). This means there exists another integer k such that

$$a \cdot k = n$$

a times k is exactly equal to n with no remainder.

Take care to distinguish $a|b$ (a divides b) from a/b (a divided by b).

If there is a second number m such that a also evenly divides m, we write $a|m$ and naturally there is a j such that

$$a \cdot j = m$$

Addition or subtraction of $m + n$ gives

$$m + n = a \cdot j + a \cdot k = a(j + k)$$

$$m - n = a \cdot j - a \cdot k = a(j - k)$$

- If $a|m$ and $a|n$, a also divides their sum or difference.

In fact, if $x \pm y = z$, and if a divides any two of them, it divides the third.

We rely on this in the following proof.

Also, we can manually find the factors for all numbers up to some reasonable lower limit. Checking for divisibility by 2, 3 and 5 is easy. So for each number, we check for divisibility by 2. If not divisible by 2, continue by testing 3, then 5. Then try 7.

Let's stop with 120, checking all numbers smaller than $121 = 11^2$.

In this way, we will discover a number of new primes. In any event, all composite (non-prime) numbers to this point will be found to have a unique prime factorization. Therefore, if there is a number with two different factorizations, it is larger than 120.

Crucially, if such numbers exist, there must be a smallest such number.

abnormal numbers

Hardy and Wright (*Theory of Numbers*, sect. 2:11) have a proof of prime factorization which is quite elegant.

Proof. By contradiction.

Hardy:

> Let us call numbers which can be factored into primes in more than one way, *abnormal*, and let n be the smallest abnormal number.

Start by supposing that there are two different factorizations of n:

$$n = p_1 \cdot p_2 \dots p_k$$

and

$$n = q_1 \cdot q_2 \ldots q_j$$

where the p's and q's are all primes.

Different factorizations

As a preliminary result, consider the possibility that some factor appears in both factorizations, so that some p is equal to a q.

We will rearrange if necessary so the shared factor is listed first: let $p_1 = q_1$ and then write

$$n = p_1 \cdot p_2 \ldots p_k = p_1 \cdot q_2 \ldots q_j$$

But now, n/p_1 (n divided by p_1) is abnormal, because it has two different prime factorizations.

But this is impossible, because n is the smallest abnormal number.

Therefore, no p is a q and no q is a p. If there exist abnormal numbers with two factorizations, those factorizations must be completely different.

This already says something strong. Imagine what it would mean for 3 to be a factor of n by one reckoning, and not be a factor of n by some other factorization.

inequality

Consider p_1, the least factor in the factorization of $n = p_1 \cdot p_2 \ldots p_k$.

It is certainly possible that $p_1 \cdot p_1 = n$. This happens when n is a perfect square, like $9 = 3 \cdot 3$. Since the p's and q's are distinct, at most, this is true of p_1 or q_1 but not both.

Suppose that $p_1 \cdot p_1 = n$.

Since $q_1 \cdot q_1 \neq n$, there must be some other factor q_j such that $q_1 \cdot q_j = n$. Since q_1 is the least of the q's, we have that $q_1 < q_j$.

$$q_1 \cdot q_1 < q_1 \cdot q_j = n$$

Now, $p_1^2 = n$ and $q_1^2 < n$ so

$$p_1^2 \cdot q_1^2 < n^2$$

$$p_1 \cdot q_1 < n$$

The other possibility is that $p_1 \cdot p_1 < n$ and the same is true for q. The above inequality still holds. In all cases, $p_1 \cdot q_1 < n$.

We also know that neither p_1 nor q_1 is equal to zero, since they are factors of n. So $p_1 \cdot q_1 > 0$ and we can write:

$$0 < p_1 \cdot q_1 < n$$

another number

We will work out the properties of a new number. Let

$$N = n - p_1 \cdot q_1$$

so

$$p_1 \cdot q_1 = n - N$$

Plug this result into the inequality from above.

$$0 < n - N < n$$

Since $0 < n - N$ we have that $N < n$, and since $n - N < n$, $N > 0$. Thus

$$0 < N < n$$

We also know that N is not abnormal, since n is the smallest abnormal number.

contradiction

We're given that $p_1|n$. From the equality

$$N = n - p_1 \cdot q_1$$

and our preliminary result about what factorization means, it must be that $p_1|N$. By the same reasoning, $q_1|N$.

We derive the contradiction as follows.

Both p_1 and q_1 appear in the factorization of N, and that factorization is unique, *since N is not abnormal.* Therefore, p_1 times q_1 *also* divides N.

Consider again

$$N = n - p_1 \cdot q_1$$

Certainly, $p_1 \cdot q_1$ divides itself, and we have concluded that $p_1 \cdot q_1$ divides N, so it follows that $p_1 \cdot q_1$ divides n, using our standard result about factorization.

This is a difficulty. One way to show it is to look at n/p_1. From what we've said, q_1 divides n/p_1.

But n/p_1 is less than n and has the unique prime factorization $p_2 \cdot p_3 \ldots p_k$.

Since q_1 is not a p, this is impossible.

Hence there cannot be any abnormal numbers.

□