MSCM Number Theory

Alan Zhou

2024-25

Contents

1 Proving the Fundamental Theorem of Arithmetic

The **fundamental theorem of arithmetic** (FTA) states that every positive integer can be written as a product of primes, and that moreover, there is only one way to do this (aside from rearranging the primes in the product). For example, 2024 and 2025 can be written as

$$2024 = 2 \cdot 2 \cdot 2 \cdot 11 \cdot 23 \quad \text{and} \quad 2025 = 3 \cdot 3 \cdot 3 \cdot 3 \cdot 5 \cdot 5.$$

For a given positive integer, the corresponding product of primes is its *prime factorization*.

This handout discusses how to prove the FTA. Knowledge of the material that is already in the textbook (AoPS Volume 1) or in class examples is assumed.

These notes are also available at https://azhou5849.github.io/teaching/.

1.1 Every positive integer is a product of primes

The fact that there is *a* prime factorization requires no additional theory.

Theorem 1.1. *Every positive integer (greater than 1) is prime or can be written as a product of primes.*

A sketch of the proof is as follows: given a positive integer $n > 1$, if n is prime, we are done. Otherwise, we can write n as a product of two positive integers smaller than n, say $n = ab$ where $1 < a, b < n$. If a is prime, then we do not need to do any further work with it, and otherwise, we can write a as a product of two positive integers smaller than n. Continuing in this way going down all possible branches of the resulting factor tree, we have to stop eventually, and this happens precisely when all of the branches end at primes.

Remark. To take this sketch and turn it into a rigorous proof, we can use mathematical induction or a proof by contradiction where we consider the minimal counterexample.

What remains is to prove that there is only one prime factorization (aside from rearrangement), and the rest of the handout is dedicated to this goal.

1.2 The euclidean algorithm

Given integers a and b, any common divisor d also divides $a + b$ and $a - b$. From this, it follows that we always have $\gcd(a, b) = \gcd(a - b, b)$. For example,

$$\begin{aligned} \gcd(70, 91) &= \gcd(70, 21) \\ &= \gcd(49, 21) = \gcd(28, 21) = \gcd(7, 21) \\ &= \gcd(7, 14) = \gcd(7, 7) = \boxed{7}. \end{aligned}$$

This method for calculating GCD is the *euclidean algorithm*.

Instead of repeatedly subtracting the same number over and over again, as in the second line above, we can jump to the end by dividing and extracting the remainder. The same example would then look like

$$\begin{aligned} \gcd(70, 91) &= \gcd(70, 21) && 91 = 1 \cdot 70 + 21 \\ &= \gcd(7, 21) && 70 = 3 \cdot 21 + 7 \\ &= \gcd(7, 0) && 21 = 3 \cdot 7 + 0 \\ &= \boxed{7}. \end{aligned}$$

Remark. Most GCD calculations in middle school competitions involve numbers small enough that we could use factorization instead. The advantage of the euclidean algorithm emerges when we deal with numbers that are difficult to factor efficiently. For instance,

$$\begin{aligned}
\gcd(582133, 268381) &= \gcd(45371, 268381) & 582133 &= 2 \cdot 268381 + 45371 \\
&= \gcd(45371, 41526) & 268381 &= 5 \cdot 45371 + 41526 \\
&= \gcd(3845, 41526) & 45371 &= 1 \cdot 41526 + 3845 \\
&= \gcd(3845, 3076) & 41526 &= 10 \cdot 3845 + 3076 \\
&= \gcd(769, 3076) & 3845 &= 1 \cdot 3076 + 769 \\
&= \gcd(769, 0) & 3076 &= 4 \cdot 769 + 0 \\
&= \boxed{769}.
\end{aligned}$$

For reference, the prime factorizations are $582133 = 757 \cdot 769$ and $268381 = 349 \cdot 769$.

1.3 Bézout's lemma

The euclidean algorithm lets us write the GCD in terms of the original integers. More specifically,

Lemma 1.2 (Bézout)**.** *Let a and b be positive integers. There exist integers m and n (which are not necessarily positive) such that*

$$ma + nb = \gcd(a, b).$$

Using the example of 70 and 91, we have

$$7 = 70 - 3 \cdot 21,$$

from the last step of the euclidean algorithm that involves a non-zero remainder. The previous step gives us $21 = 91 - 1 \cdot 70$, and we substitute this in to get

$$7 = 70 - 3 \cdot (91 - 1 \cdot 70) = 4 \cdot 70 - 3 \cdot 91.$$

To match the form of the lemma, this says that

$$4 \cdot 70 + (-3) \cdot 91 = \gcd(70, 91).$$

A more involved example is that

$$\begin{aligned}
769 &= 1 \cdot 3845 - 1 \cdot 3076 = 1 \cdot 3845 - (41526 - 10 \cdot 3845) \\
&= 11 \cdot 3845 - 1 \cdot 41526 = 11 \cdot (45371 - 1 \cdot 41526) - 1 \cdot 41526 \\
&= 11 \cdot 45371 - 12 \cdot 41526 = 11 \cdot 45371 - 12 \cdot (268381 - 5 \cdot 45371) \\
&= 71 \cdot 45371 - 12 \cdot 268381 = 71 \cdot (582133 - 2 \cdot 268381) - 12 \cdot 268381 \\
&= 71 \cdot 582133 - 154 \cdot 268381.
\end{aligned}$$

1.4 Divisibility by a prime

Proposition 1.3. *Let p be a prime number and let a and b be integers. If p divides ab, then p divides a or p divides b.*

Proof. It suffices to show that if p divides ab but p does not divide a, then p divides b. Since the only divisors of p are 1 and p, we must have $\gcd(a, p) = 1$ or $\gcd(a, p) = p$. The latter implies p divides a, so it must be the case that $\gcd(a, p) = 1$. Now, Bézout's lemma implies that we can find integers m and n such that

$$ma + np = 1.$$

Multiplying through by b,

$$m(ab) + (nb)p = b.$$

Both terms of the left hand side are divisible by p, so the right hand side b is divisible by p. □

It follows that if a prime p divides a product of integers of any length (not necessarily just two factors), then p divides at least one of the integers in the product.

1.5 Prime factorization is unique

We can now prove that no positive integer can have two genuinely different prime factorizations. Suppose that the positive integer n has prime factorizations

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

where $p_1 \leq p_2 \leq p_3 \leq \cdots \leq p_r$ and $q_1 \leq q_2 \leq q_3 \leq \cdots \leq q_s$ are primes. Suppose $p_1 \leq q_1$. By the previous section's result, the prime number p_1 divides the product $q_1 q_2 q_3 \cdots q_s$, so p_1 must divide one of the factors $q_1, q_2, q_3, \ldots, q_s$. Since those factors are all prime, the only way this can happen is for p_1 to be equal to one of those factors. Since $p_1 \leq q_1$, which is the smallest of the factors, we must have $p_1 = q_1$. If we suppose instead that $q_1 \leq p_1$, we reach the same conclusion with the same argument. Thus we have $p_1 = q_1$ in any case, and we can cancel this common factor to get

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

We then continue this process to show $p_2 = q_2$, then $p_3 = q_3$, and so on. If one side ever runs out of factors before the other, then we end up with one side of the equation being 1 and the other side being a product of one or more primes, which is impossible. Thus the two factorizations have the same length and use the same primes.