Problem Set VI: The Multiplicative Structure of $\mathbb{Z}$

November 15, 2025

Thus far, we've managed to prove the following:

Theorem (Fundamental Theorem of Arithmetic). *For every positive integer $n \geq 2$, there exists a unique tuple $(p_1, \ldots, p_k)$ such that:*

1. *$p_1 \leq \ldots \leq p_k$; that is, it's a non-decreasing sequence,*
2. *$\prod_{i=1}^{k} p_i = n$; that is, it's a factorization of n,*
3. *and $p_1, \ldots, p_k$ are all prime.*

This means that we can very efficiently encode a positive integer n as follows:

Definition (Canonical representation). Given $n \in \mathbb{Z}^+$, we define its *canonical representation* as an infinite sequence $(\alpha_i)_{i=1}^{\infty}$ such that α_k is the number of times that the kth prime appears in the prime factorization of n (in other words, its exponent). Given that the prime factorization is finite, this sequence is eventually constantly 0.

Then, the Fundamental Theorem of Arithmetic is telling us that this canonical representation exists and is unique; so let's denote it as $\mathbf{rep}(n)$ (though this notation is non-standard; but we'll use it for convenience). Finally, we define $\mathbf{rep}(1) = (0, 0, 0, \ldots)$, so now all positive integers have a canonical representation.

As another notational convenience, given infinite sequences of integers $a = (a_i)_{i=1}^{\infty}$ and $b = (b_i)_{i=1}^{\infty}$, and a binary function $f : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$, let's denote by $f(a, b)$ the sequence $(f(a_i, b_i))_{i=1}^{\infty}$, consisting of the component-wise application of f; and given a binary relation $R \subseteq \mathbb{Z} \times \mathbb{Z}$, let's also say that aRb whenever $a_i R b_i$ holds for all i.

Theorem (Properties of the canonical representation). *Given positive integers $a, b \in \mathbb{Z}^+$, we have the following:*

- $\mathbf{rep}(ab) = \mathbf{rep}(a) + \mathbf{rep}(b)$ *(so it behaves like a logarithm!)*
- *For integers $n \geq 0$, $\mathbf{rep}(a^n) = n\ \mathbf{rep}(a)$.*
- $\mathbf{rep}(\gcd(a, b)) = \min\{\mathbf{rep}(a), \mathbf{rep}(b)\}$.
- *If $a \mid b$, then $\mathbf{rep}\left(\frac{b}{a}\right) = \mathbf{rep}(b) - \mathbf{rep}(a)$.*
- *We have that $a \mid b$ if and only if $\mathbf{rep}(a) \leq \mathbf{rep}(b)$.*

Proof. Left as an optional exercise. □

Isn't that nice! The only problem is that if we don't have the canonical representation of n at hand, computing $\mathbf{rep}(n)$ can be very costly (integer factorization is a famously difficult problem). Otherwise, this is very useful.

Finally, let's introduce two pieces of notation that are actually standard: for a prime p, we define the *p-adic valuation* $\nu_p(n)$ as the exponent of p in the prime factorization of n. Note that this will be an entry of $\mathbf{rep}(n)$, so p-adic valuations inherit all the properties from **rep** (although we'd now say that $a \mid b$ if and only if, for all prime p, $\nu_p(a) \leq \nu_p(b)$).

Another important piece of notation is the Iverson bracket: given a *statement* S, we say that $[S] = 1$ if S is true, and $[S] = 0$ if S is false. This is generally quite useful in double counting arguments. We can make the following observations connecting regarding the Iverson bracket (think about why these must be the case):

Observation 1. *For all $n \in \mathbb{Z}^+$ and p prime, $\nu_p(n) = \sum_{k=1}^{\infty} [p^k \mid n]$.*

Observation 2. For $n, d \in \mathbb{Z}^+$, $n \text{ div } d = \sum_{k=1}^{n} [d \mid k]$.

Now, we can try our luck at computing $\nu_p(n!)$:

$$\begin{aligned} \nu_p(n!) &= \nu_p\left(\prod_{i=1}^{n} i\right) \\ &= \sum_{i=1}^{n} \nu_p(i) \text{ by the properties of } \mathbf{rep} \\ &= \sum_{i=1}^{n} \sum_{j=1}^{\infty} [p^j \mid i] \text{ by Observation 1} \\ &= \sum_{j=1}^{\infty} \sum_{i=1}^{n} [p^j \mid i] \text{ by Observation 2} \\ &= \sum_{j=1}^{\infty} (n \text{ div } p^j) \end{aligned}$$

This identity is known as *Legendre's formula.*

Instructions: No need to solve them fully, just ponder them. We'll probably go over the solutions of a few of them during class. Have fun!

Problem 1. (Canonical representations, optional)
Prove the presented properties of canonical representations. I'd recommend doing so in the presented order, since the property $\mathbf{rep}(ab) = \mathbf{rep}(a) + \mathbf{rep}(b)$ greatly simplifies the proofs of the others.

Problem 2. (Folklore classic)
When writing 100! in base 10, how many zeroes are there at the end?

Problem 3. (Kummer's theorem)
First, prove that for $a, b, d \in \mathbb{Z}^+$,

$$(a+b) \text{ div } d - a \text{ div } d - b \text{ div } d = [(a \bmod d) + (b \bmod d) \geq d].$$

For nonnegative integers n, k such that $0 \leq k \leq n$, we may define the *binomial coefficient* $\binom{n}{k}$ non-recursively as

$$\binom{n}{k} := \frac{n!}{k!(n-k)!}$$

Use this to find an expression (which may involve Iverson brackets) for $\nu_p\left(\binom{a+b}{a}\right)$ for prime p and non-negative integers a, b.

Problem 4. (Exploring central binomial coefficients)
It is well-known that $\binom{2n}{0} + \binom{2n}{1} ... + \binom{2n}{2n} = 2^{2n} = 4^n$; and in particular, the *central binomial coefficient* $\binom{2n}{n}$ will be the largest term in the sum, meaning that $\binom{2n}{n} \geq \frac{4^n}{2n}$ (why does it entail this?)

Prove that for prime p and $n \in \mathbb{Z}^+$, we have that

$$p^{\nu_p\binom{2n}{n}} \leq 2n$$

What does this tell us, if anything, about the number of primes up to $2n$ (that is, $\pi(2n)$)?