Sieve of Erasthothenes

1 Introduction

Sieve of Erasthothenes is an algorithm for finding all the prime numbers in a segment $[1; n]$ using $O(n \log\log n)$ opertions.

In the algorithm, we first begin from 2 and cross out all the multiples of 2, i.e, $\{4, 6, 8, \ldots\}$ except 2 itself. Now we move on to the next uncrossed number which is 3 in this case and similarly cross-out its multiples, i.e., $\{6, 9, 12, \ldots\}$. We skip those numbers which were already crossed out before.

Similarly, the next unmarked number is 5, then 7, and so on. In the end, the numbers remaining unmarked are the ones which are prime in $[1; n]$.

2 Asymptotic Analysis

Let us prove that its complexity is indeed $O(n \log\log n)$. For any prime $p \leq n$, the algorithm will perform $\dfrac{n}{p}$ operations. So we need to find the complexity of

$$\sum_{p \leq n} \frac{n}{p} = n \cdot \sum_{p \leq n} \frac{1}{p}, \qquad p \text{ is a prime}$$

We can use the following facts-

1. The number of primes less than or equal to n is approx. $\frac{n}{\ln(n)}$
2. The k^{th} prime number approximately equals $k \ln(k)$

$$\therefore \sum_{p \leq n} \frac{1}{p} \approx \frac{1}{2} + \sum_{k=2}^{\frac{n}{\ln(n)}} \frac{1}{k \ln(k)}$$

The summation itself is approximately of the same order as its integral, so

$$\sum_{k=2}^{\frac{n}{\ln(n)}} \frac{1}{k \ln(k)} \approx \int_{k=2}^{\frac{n}{\ln(n)}} \frac{1}{k \ln(k)} = \ln(\ln(n) - \ln(\ln(n))) - \ln(\ln 2) \approx \ln\ln n$$

Thus, putting it in the above equation,

$$\sum_{p \leq n} \frac{n}{p} \approx n \ln\ln n + o(n)$$