# Additive Number Theory Talk #13: More Brun's Sieve

**Akash Kumar**

March 20, 2024

# Contents

# 1. Introduction

Last week, Johnny introduced the idea of the sifting function $\sigma(n) = \sum_{d|\gcd(n,P_z)} \mu(d) \in \{0,1\}$. Recall that the idea behind this was that

$$\sum_{d|x} \mu(d) = \begin{cases} 1 & \text{if } x = 1 \\ 0 & \text{otherwise} \end{cases}$$

and we're plugging in $x = \gcd(n, P_z)$ so as to ensure weed out any numbers sharing divisors with $P_z$.

We will now generalize this to $\sigma(n) = \sum_{d|\gcd(n,P_z)} \mu(d) \cdot \chi(d)$, where $\chi$ is some characteristic function. The idea of this is by choosing $\chi_1, \chi_2$ carefully, we'll get **both $\overline{\text{upper}}$ and $\underline{\text{lower}}$ bounds** for $S(A; P_z, x)$.

The first step in doing so is the following proposition, which will allow us to think of $S(\mathcal{A}; P_z, x)$ as the sum of a main term and an error term.

**Proposition 2.2.1.** Let $P^z_{(d)} := \prod_{p \in P_z,\ p \nmid d} p$. Then

$$S(\mathcal{A}, P^z, x) = \sum_{d|P_z} \mu(d)\chi(d)|\mathcal{A}_d| - \sum_{1<d|P_z} \sigma(d) S(\mathcal{A}_d; P^z_{(d)}, x).$$

*Remarks.* Before proving the proposition, I'll discuss some things to clear up what exactly we're claiming. Firstly, to avoid confusion, the $\sigma(d)$ above denotes Connor's $\sigma$ with $\chi$, not Johnny's version.

Next, I'll provide some intuition by going through what the statement looks like for $\chi = 1$. (BE BRIEF HERE.) In this case, the first term is equal to $S(\mathcal{A}, P^z, x)$ exactly (it's the principle of inclusion-exclusion idea), so we expect the second term to be equal to zero. And this is indeed true. (WON'T DISCUSS THIS.) Whenever $\sigma(d) = 1$, this means that $d$ is not divisible by any of the primes $\leq z$. This implies that $S(\mathcal{A}_d; P^z_{(d)}, x) = 0$ because we're trying to weed out multiples of $d$ using primes that don't divide $d$. Hence, every summand in the second term is zero.

This result matches our intuition in thinking about the role of $\chi$. Our goal is to choose the function $\chi$ carefully so that the main term still captures most of the true value of $S(\mathcal{A}, P^z, x)$

while keeping the error term small. When $\chi = 1$, we've opted for an all-main-term, zero-error approach. Now let's get into the proof. (ONLY EXPLAIN MOBIUS INVERSION)

*Proof.*

$$\begin{aligned}
\sum_{d|P_z} \mu(d)\chi(d)|\mathcal{A}_d| &= \sum_{d|P_z} |\mathcal{A}_d^x| \sum_{\delta|d} \mu\left(\frac{d}{\delta}\right)\sigma(d) \quad \text{Mobius inversion} \\
&= \sum_{\delta|P_z} \sigma(\delta) \sum_{t|\frac{P_z}{\delta}} \mu(t)|\mathcal{A}_{\delta t}| \\
&= \sum_{t|P_z} \mu(t)|\mathcal{A}_t| + \sum_{1<\delta|P_z} \sigma(\delta) \sum_{t|\frac{P_z}{\delta}} \mu(t)|\mathcal{A}_{\delta t}| \quad \text{split into } \delta = 1 \text{ or } \delta > 1 \\
&= S(\mathcal{A}, P^z, x) + \sum_{1<\delta|P_z} \sigma(\delta) \sum_{t|\frac{P_z}{\delta}} \mu(t)|A_{\delta t}| \\
&= S(\mathcal{A}, P^z, x) + \sum_{1<d|P_z} \sigma(d) S(\mathcal{A}_d; P^z_{(d)}, x)
\end{aligned}$$

Rearranging gives the desired result. □

# 2. Road to the Theorem

## 2.1 Finding Characteristic Functions

Having proved Proposition 2.2.1, we now seek to find a functions $\chi_1$ and $\chi_2$ that give upper and lower bounds for $S(\mathcal{A}, P^z, x)$. What we're looking for is

$$\sum_{d|P_z} \mu(d)\chi_2(d)|\mathcal{A}_d| \leq S(\mathcal{A}; P_z, x) \leq \sum_{d|P_z} \mu(d)\chi_1(d)|\mathcal{A}_d|$$

The text goes through a bunch of algebra to find some properties that $\chi_1$ and $\chi_2$ must satisfy.

To achieve this, our characteristic functions $\chi^{(r)}$ will do two things:

1) Restrict the number of primes dividing $d$: $\nu(d) < r$

2) Restrict the interval that the primes dividing $d$ can come from

The second restriction requires us to produce a partition

$$2 = z_r < z_{r-1} < \cdots < z_1 < z_0 = z$$

We also introduce the following notation: $\beta_n = \gcd\left(d, P_{(z_n,z)}\right)$.

We are now ready to present what we'll take $\chi_1, \chi_2$ to be:

$$\chi_i(d) = \begin{cases} 1 & \text{if } \forall m \in \{1, \ldots, r\},\ \nu(\beta_m) \leq 2b - i - 1 + 2m \\ 0 & \text{otherwise} \end{cases}$$

The variable $b$ above is a constant that is introduced in the algebra on finding necessary properties of $\chi_i$. The intuition is as before: we are restricting the number of primes dividing $d$ as well as the interval from which the prime divisors come.

## 2.2 Steps to Deduce Bounds

Having found suitable $\chi_i$ functions, we now ask: what upper/lower bounds do we get?

Again, I will omit most of the algebra and try to outline the main components of the argument.

I will first discuss an assumption we make about $\omega(p)$. Typically, we've assumed $\omega(p) = O(1)$. (It was $\omega(p) = 1$ for Erasthothenes and $\omega(p) = 2$ for twin primes.) We will use a weaker assumption:

$$\sum_{w \leq p < z} \frac{\omega(p)\ln(p)}{p} \leq \kappa \ln\left(\frac{z}{w}\right) + \eta, \quad 2 \leq w \leq z$$

This basically says that while $\omega(p)$ may not be bounded for all inputs, it is "bounded on average". This is because we're taking the sum over many inputs, and requiring that $\omega$'s behavior is controlled across the sum. It could spike, but infrequently so.

(MENTION IF $\omega(p) = 1$, THEN $\kappa = \eta = 1$ WORKS.)

(WON'T DISCUSS) I'll also discuss the selection of the intervals/partition I mentioned above. The overall idea is to select the numbers $z_n$ with an exponential fall-off in the logarithm. The intervals will be given by

$$\ln z_n = e^{-n\Lambda} \ln z, \quad n = 1, \ldots, r-1$$

where $\Lambda$ is some real number and we set $z_r = 2$.

This is all with the goal of bounding $\frac{W(z_n)}{W(z)}$, an important term that pops out when doing algebra on bounding $S(\mathcal{A}; P_z, x)$.

## 2.3 The Theorem

We are now ready to state the theorem. **Theorem 2.2.2.** Assume that

$$1 \leq \frac{1}{1 - \frac{\omega(\rho)}{\rho}} \leq A,$$

$$\sum_{w \leq p < z} \frac{\omega(p)\ln p}{p} \leq \kappa \ln\left(\frac{\ln z}{\ln w}\right) + \frac{\eta}{\ln w},$$

and

$$|R_d| \leq \omega(d).$$

Let $\lambda$ be such that $0 < \lambda e^{1+\lambda} < 1$. Then

$$S(\mathcal{A}; P^z, x) \leq xW(z)\left(1 + 2\frac{\lambda^{2b+1}e^{2\lambda}}{1-(\lambda e^{1+\lambda})^2}\exp\left((2b+3)\frac{c}{\lambda \ln z}\right)\right) + O\left(z^{2b-1+\frac{2\xi}{e^{\frac{2\lambda}{\kappa}}-1}}\right), \quad \text{(U)}$$

and

$$S(\mathcal{A}; P^z, x) \geq xW(z)\left( 1 - 2\frac{\lambda^{2b}e^{2\lambda}}{1-(\lambda e^{1+\lambda})^2}\exp\left((2b+2)\frac{c}{\lambda \ln z}\right)\right) + O\left(z^{2b-1+\frac{2\xi}{e^{\frac{2\lambda}{\kappa}}-1}}\right), \quad \text{(L)}$$

where

$$c = \frac{\eta}{2}\left(1 + A\left(\kappa + \frac{\eta}{\ln 2}\right)\right),$$

and $\xi = 1 + \epsilon$ for $0 < \epsilon < 1$.

*Intuition.* Let's look within the parentheses next to $xW(z)$, and see how we've made progress from previous talks. The $+1$ doesn't really matter; it's just the $xW(z)$. Now let's look at the remaining portion. This can be thought of as error. Before, we had our error to be $O(2^{\pi(z)})$, which is exponential. Though it's not obvious, one can choose the parameters so that it's less than $2^{\pi(z)}$, which should make sense given that $O(2^{\pi(z)})$ was exponential (bad).

Additionally, we have now introduced lower bounds, which has an interesting application...

# 3. Application to Twin Primes

We are going to show that there are infinitely many $n$ such that $\nu(n(n+2)) \leq 7$. Note that if we could change the 7 to a 2, this would prove the Twin Prime Conjecture, so this is a considerable step in that direction.

For the twin primes problem, we set $\mathcal{A} = \{n(n+2) \mid n(n+2) \leq x\}$. We also have $\omega(2) = 1$ and $\omega(p) = 2$. $\omega(2) = 1$ is so as to not divide by zero in the first assumption of our theorem, and $\omega(p) = 2$ because $n(n+2) \not\equiv 0 \pmod{p}$ rules out two resides.

With this, all the conditions of the theorem hold, and the lower bound given by inequality $(L)$ is positive, so that $\lim\limits_{x\to\infty} S(\mathcal{A}; P^z, x) = \infty$. This shows that infinitely many elements survive the sifting process.

Now we show that these elements that survive satisfy $\nu(n(n+2)) \leq 7$. For this, we set $z = x^{1/8}$. Ideally, we'd have $z = \sqrt{x}$ as any prime divisor of $x$ must be at most $\sqrt{x}$, but this is too ambitious for our sieve. Therefore, we settle for $z = x^{1/8}$. This is just to make the conditions of our theorem work, so if one could develop a stronger sieve, we could perhaps do better than $z = x^{1/8}$.

Ok, so why do we have $\nu(n(n+2)) \leq 7$? Because we set $z = x^{1/8}$, we know that all the prime factors of $n(n+2)$ are greater than $z = x^{1/8}$. So, if we have $n(n+2) = p_1 p_2 \cdots p_r$, then $n(n+2) > \left(x^{1/8}\right)^r = x^{r/8}$. Moreover, by definition of sifting, we have $n(n+2) \leq x$. Therefore, $x^{r/8} < x$, which implies $\frac{r}{8} < 1 \implies r < 8$. Hence, $n(n+2)$ has at most 7 prime factors.

# 4. Application to Coprimality Counting

Let $k, x > 1$ be fixed integers. Our goal is to estimate the number of integers $\leq x$ that are coprime to $k$. In other words, we are interested in the sum

$$\sum_{n \leq x,\ \gcd(n,k)=1} 1$$

Note that if $k = x$, then this is simply Euler's Totient Function $\varphi(x)$. We, however, are interested in when $x$ is much larger than $k$.

It is clear that within intervals modulo $k$, there are $\varphi(k)$ integers coprime to $k$. Yet, if $x$ doesn't land on a multiple of $k$, then the sum depends on how the integers coprime to $k$ are distributed. We will use Brun's Sieve to attack this problem! (MENTION THAT INTUITIVE ANSWER IS $\frac{\varphi(k)}{k}x$)

The set that we will sift is $\mathcal{A} = \{n \mid n \leq x\}$ and the sifting primes are $P = \{p : p \mid k\}$ (we don't want our numbers to share common factors with $k$).

I'll explain why the three assumptions (in order) of Theorem 2.2.2 hold:

- You can check (on Desmos) that $\frac{1}{1-\frac{1}{p}} \leq 2$ for $p \geq 2$, so $A = 2$.
- The inequality holds by previous discussion. (not obvious)
- We have $|A_d| = \frac{x}{d} + R_d$, where $\omega(d) = 1$ and $R_d \leq 1$ (because floor function). So $|R_d| \leq \omega(d)$

Doing similar work to fill in the other parameters of Theorem 2.2.2, we get $\kappa = \eta = 1$, $b = 1$, $\xi = 1.005$, $\lambda = 0.204$. With these parameters, we get

$$S(\mathcal{A}; P, z) \geq xW(z)(1 - o(1)) + O(z^{4.85})$$

Plugging in $z = x^{1/5}$ (as we want the error term to be sublinear to be negligible relative to the main term), we get

$$S(\mathcal{A}; P, z) \geq c \prod_{p|k} \left(1 - \frac{1}{p}\right) x + O(x^{0.97})$$

After all this, we get the following lower and upper bounds, assuming $k$'s prime factors are less than $x^{1/5}$.

$$\boxed{c \cdot \frac{\varphi(k)}{k} x + O(x^{0.97}) \leq \sum_{n \leq x,\ \gcd(n,k)=1} 1 \leq c' \cdot \frac{\varphi(k)}{k} x + O(x^{0.975})}$$

where $c < 1$ and $c' < 4$.

# 5. Bibliography

https://pages.cs.wisc.edu/~cdx/Sieve.pdf