Utrecht University

Faculty of Science

Department of Mathematics

Rough numbers and stick-breaking

Master Thesis

Daily supervisor:
dr. L.P. Grimmelt

Author:
Bart van Hoek

First reader:
dr. D. Schindler

Second reader:
Prof. dr. G.L.M. Cornelissen

July 9, 2019

Abstract

We construct the Buchstab function using a "stick-breaking process". This gives us a new way of expressing the function and thereby some new interpretations. To do this we first look at the basic results concerning rough numbers like the prime number theorem. We analyse the strongest known result for counting rough numbers by Gérald Tenenbaum [2].

Contents

1 Introduction **3**

2 The prime number theorem **5**
2.1 Estimating sums over prime numbers . . . 5
2.2 Abel's summation formula . . . 10
2.3 Mertens' theorem . . . 11
2.4 Perron's formula . . . 12
2.5 The prime number theorem . . . 18

3 Rough numbers **24**
3.1 Smooth numbers . . . 24
3.2 Rough numbers for small y . . . 25
3.3 The Buchstab function . . . 27
3.4 Buchstab with better main term . . . 30
3.5 Strongest known result . . . 32

4 Stick-breaking **35**
4.1 The stick-breaking process . . . 35
4.2 The Dickman function . . . 36
4.3 A naive guess . . . 37
4.4 Stick-breaking for the Buchstab function . . . 38
4.5 Interpretation of the ω_k . . . 41

1 Introduction

It was in 1896 that the prime number theorem was first proven. It was shown, independently, by both Hadamard [10] and de la Vallée Poussin [11] that

$$\pi(x) \approx \int_2^x \frac{1}{\log t} dt. \tag{1}$$

The error term of this approximation has been improved multiple times over the last century and is still being looked at today.

In this master thesis no further research will be done on this topic and only the known literature will be treated. We do this by first studying different sums over prime numbers which ultimately lead to Mertens' second theorem, [13]. One of the most important tools we use in proving Mertens' theorem and thereby in connecting all our different sums is Abel's summation formula. This formula will be essential in each of the three chapters and thus not only the result but also the understanding of how to use it will be important.

After this we prove Perron's formula, which is what we use to derive the prime number theorem. To do this we discuss the Riemann zeta function and the effect its zeros have on the prime number theorem. We will see that the absence of zeros on the line $\mathrm{Re}(s) = 1$ is what in the end leads to the prime number theorem and can be seen as an equivalent statement.

In the second chapter we look at the following.

Definition 1.1. We define *y-rough numbers* as all natural numbers having only primes greater or equal than y.

Despite only being named very recently in 2001 in one by Finch, see [7], rough numbers have been studied long before that. Similar to the first chapter we will only study the y-rough numbers which are bounded by x and denote these by $\phi(x, y)$. We have the following result for rough numbers which is analogous to the prime number theorem for primes. Alexander Buchstab proved in 1937, see [12], that

Theorem 1.2 (Buchstab)**.** *Taking $x \geq y \geq 2$. Then uniformly for $1 \leq u \leq U$ we have*

$$\phi(x, y) = \frac{x\omega(u) - y}{\log y} + O_U\left(\frac{x}{(\log x)^2}\right) \tag{2}$$

It is defined in terms of the Buchstab function $\omega(u)$. The definition and how this Buchstab function arises will be discussed in chapter 3.3. Similar to the prime number theorem the result in equation 2 has been improved multiple times. In the last paragraph we look at the currently strongest result for $\phi(x, y)$ given by Tenenbaum in [2]. We discuss his results and relate them to the things we derived ourselves.

In the final chapter the research part of the thesis is done. To do this we define the following process.

Definition 1.3. We define an infinite number of random variables $(x_1, x_2, ..)$ as a *stick-breaking process* when for every n, x_n is uniformly distributed on the interval $1 - \sum_{k=1}^{n-1} x_k$ such that $\sum_{i=1}^{\infty} x_i = 1$.

In a paper by Brady from 2017, they define a probability on such a stick-breaking process, see chapter 4.3 of [5]. It is shown that this probability is an alternative way of defining the Dickman function, which is the equivalent of the Buchstab function for smooth numbers.

We then try to find another probability measure which results in new way of defining the Buchstab function. This leads us to a new expression for the Buchstab function which was also derived in a completely different manner by Pinsky in [4].

We believe the most interesting part follows after that. When constructing the Buchstab function via stick-breaking, we derive some new functions ω_k. These are not only relevant since they lead to the Buchstab function. We prove these newly derived functions help directly counting the number of rough numbers in a specific range with a fixed amount of prime divisors.

2 The prime number theorem

Prime numbers play an important part in all of mathematics and especially in number theory. It is commonly known there are infinitely many prime numbers. The more interesting question is how "many" prime numbers there are and how do they behave. It seems logical that that the higher we get the smaller the density of primes becomes since there are more possible divisors. In the first chapter we will look more formally into this by studying the famous prime number theorem. The prime number theorem describes the distribution of primes and gives an asymptotic for the prime counting function. It was first proven by both Hadamard and de la Vallée Poussin separately in 1896.

In this chapter we will give a sketch of the proof of the prime number theorem with some important parts given in more detail. More importantly we will discuss some key results proven before the prime number theorem including Mertens' theorem and Perron's formula . Lastly we will try to relate some results using Abel's summation formula. All this will be done by following the approach from chapter 1 of the book *Einführung in die analytische Zahlentheorie* by Jörg Brüdern, [1].

2.1 Estimating sums over prime numbers

A very basic result in number theory is the fact that there are infinitely many prime numbers. Namely assume we have a finite list of prime numbers $p_1, ..., p_l$. The product of all these number added one, $Q = 1 + p_1...p_l$, is by assumption not a prime number and thus has a prime factor p_i. Hence we derive that the difference, namely 1, must also be divisible by p_i. A contradiction, meaning there are infinitely many primes.

The next more challenging step is to look at the density of those primes. How many primes will there be if we are looking at a range of numbers with more than ten digits. A lot less relatively when compared to the amount of primes below 100. To get a more specific idea of this prime density, we will look at the prime counting function which counts the number of primes up until an upper bound.

$$\pi(x) := \#\{p : p \leq x\}$$

Before we look more specifically to $\pi(x)$ we want to find some sort lower bound for the number of primes by taking a close look at the sum over $1/p$. We first look at the following identity which will be essential for creating a bound later.

For $s > 0$ a real number and $x \geq 2$, we have

$$\prod_{p\leq x}(1 - \frac{1}{p^s})^{-1} = \prod_{p\leq x}\sum_{k=0}^{\infty}(\frac{1}{p^s})^k = \prod_{p\leq x}\sum_{k=0}^{\infty}(\frac{1}{p^k})^s.$$

Where we can introduce the sum by using the geometric series. Now taking a close look at the last sum and using the fundamental theorem of arithmetic, we see that in the denominator every number with prime divisors smaller than x appears exactly once. Defining $A(x) := \{n : p \mid n \implies p \leq x\}$, we can rewrite the sum to

$$= \sum_{n\in A(x)} \frac{1}{n^s} = \sum_{n\leq x} \frac{1}{n^s} + \sum_{\substack{n>x,\\ n\in A(x)}} \frac{1}{n^s}.$$

Next we can take $s = 1$ and get the clear estimate

$$\prod_{p \le x} (1 - \frac{1}{p})^{-1} > \sum_{n \le x} \frac{1}{n} > \int_1^x \frac{dt}{t} = \log x.$$

Since the logarithm is continuous and monotonically increasing function for $x \ge 2$, taking it on both sides gives us the inequality,

$$\sum_{p \le x} \log(1 - \frac{1}{p})^{-1} > \log \log x. \tag{3}$$

We now want to extract the sum over $1/p$ from this sum and use the bound we just found. To do this we use the following series expansion for the logarithm. For $0 \le x < 1$ we have the following expansion.

$$\log \frac{1}{1 - x} = \sum_{k=1}^{\infty} \frac{x^k}{k}.$$

This gives us a very nice way of extracting $1/p$, namely,

$$\sum_{p \le x} \log(1 - \frac{1}{p})^{-1} = \sum_{p \le x} \sum_{k=1}^{\infty} \frac{1}{kp^k} = \sum_{p \le x} \frac{1}{p} + \sum_{p \le x} \sum_{k=2}^{\infty} \frac{1}{kp^k}.$$

This leads to the following theorem.

Theorem 2.1. *For $x \le 2$ there is a non-trivial lower bound given by,*

$$\sum_{p \le x} \frac{1}{p} > \log \log x - \frac{1}{2}. \tag{4}$$

Proof. We can write

$$\sum_{p \le x} \frac{1}{p} = \sum_{p \le x} \log(1 - \frac{1}{p})^{-1} - \sum_{p \le x} \sum_{k=2}^{\infty} \frac{1}{kp^k}.$$

So clearly in order to proof the theorem we only need to prove the second term is less than $1/2$, since the first part was already proven by equation 3. We start by rewriting the double sum as

$$\begin{aligned} \sum_{p \le x} \sum_{k=2}^{\infty} \frac{1}{kp^k} \le \frac{1}{2} \sum_{p \le x} \sum_{k=2}^{\infty} \frac{1}{p^k} &= \frac{1}{2} \sum_{p \le x} \Big(\frac{1}{1 - 1/p} - 1/p - 1 \Big) \\ &= \frac{1}{2} \sum_{p \le x} \frac{1}{p(p-1)} < \frac{1}{2} \sum_{n=2}^{\infty} \frac{1}{n(n-1)} = \frac{1}{2}. \end{aligned}$$

For the last equality notice that

$$\frac{1}{n(n-1)} = \frac{1}{n-1} - \frac{1}{n},$$

from which it is clear the sum converges to 1 when writing out the terms. Hence this proves the theorem. □

To get a upper and lower bound for $\pi(x)$ we will first look at the function

$$\theta(x) = \sum_{p \leq x} \log p$$

and try to get bounds for this function. After which it will be a relatively small step to estimating the $\pi(x)$ function in a similar way.

In order to do this we need two lemmas. First of all we have a basic version of Stirling's formula.

Lemma 1.

$$\log n! = n \log n - n + O(\log n) \tag{5}$$

Proof. The proof is relatively simple and only contains one line. We write

$$\log n! = \sum_{k=1}^{n} \log k = \int_1^n \log t \, dt + O(\log n) = n \log n - n + O(\log n)$$

□

Lemma 2. *For every n we can define*

$$e(p) = \sum_{k \geq 1} \left\lfloor \frac{n}{p^k} \right\rfloor,$$

such we can write

$$n! = \prod_{p \leq n} p^{e(p)}.$$

Proof. The proof of this lemma is based on the fundamental theorem of arithmetic. Clearly of all the numbers up until n there are exactly $\left\lfloor \frac{n}{m} \right\rfloor$ divisible by m. In our case we have $\left\lfloor \frac{n}{p^k} \right\rfloor$ numbers smaller or equal to n divisible by p^k. Since this is the case for every k it may be clear the prime p appears in the product exactly $e(p)$ times, resulting in the given formula. □

Now taking the logarithm of what we found in lemma 2 we derive

$$\log n! = \sum_{p \leq n} \log p^{e(p)} = \sum_{p \leq n} \sum_{k \geq 1} \left\lfloor \frac{n}{p^k} \right\rfloor \log p.$$

From this we want to extract the term $\sum_{p \leq n} \left\lfloor \frac{n}{p} \right\rfloor \log p$ and estimate the other (smaller) terms. We saw earlier that $\sum_{k \geq 2} \frac{1}{p^k} = \frac{1}{p(p-1)}$. Using this we can give the estimate,

$$\sum_{p \leq n} \sum_{k \geq 2} \left\lfloor \frac{n}{p^k} \right\rfloor \log p \leq \sum_{p \leq n} n \log p \sum_{k \geq 2} \frac{1}{p^k} = n \sum_{p \leq n} \frac{\log p}{p(p-1)} = O(n).$$

Where the last equality follows since the sum clearly converges.

Including lemma 1 we then derive the equality

$$\log n! = \sum_{p \leq n} \left\lfloor \frac{n}{p} \right\rfloor \log p + O(n) = n \log n - n + O(\log(n)).$$

This leads to the result

$$\sum_{p\le n}\left\lfloor\frac{n}{p}\right\rfloor\log p = n\log n + O(n). \tag{6}$$

Using the fact that for $p > n$, $\left\lfloor\frac{n}{p}\right\rfloor$ vanishes, we derive the following

$$\begin{aligned}\sum_{p\le 2n}\Big(\left\lfloor\frac{2n}{p}\right\rfloor - 2\left\lfloor\frac{n}{p}\right\rfloor\Big)\log p &= \sum_{p\le 2n}\left\lfloor\frac{2n}{p}\right\rfloor\log p - \sum_{p\le n}2\left\lfloor\frac{n}{p}\right\rfloor\log p\\ &= 2n\log 2n + O(2n) - 2n\log n - O(n) = O(n).\end{aligned}$$

We want to relate this to the function $\theta(x)$. Clearly $\lfloor 2x\rfloor \ge 2\lfloor x\rfloor$. Hence it is clear that

$$\sum_{n<p\le 2n}\Big(\left\lfloor\frac{2n}{p}\right\rfloor - 2\left\lfloor\frac{n}{p}\right\rfloor\Big)\log p \le \sum_{p\le 2n}\Big(\left\lfloor\frac{2n}{p}\right\rfloor - 2\left\lfloor\frac{n}{p}\right\rfloor\Big)\log p = O(n).$$

Now using the fact that for $n < p \le 2n$ the term $\left\lfloor\frac{2n}{p}\right\rfloor - \left\lfloor\frac{n}{p}\right\rfloor = 1$ we can extract that

$$\theta(2n) - \theta(n) = \sum_{n<p\le 2n}\log p = \sum_{n<p\le 2n}\Big(\left\lfloor\frac{2n}{p}\right\rfloor - 2\left\lfloor\frac{n}{p}\right\rfloor\Big)\log p = O(n).$$

We want a similar statement for non integer values since θ is defined for real values x. For this we notice that there may be one prime number in between $\theta(\lfloor 2x\rfloor)$ and $\theta(2\lfloor x\rfloor)$, resulting in an added logarithm. Luckily this error term is absorbed in the bigger error term, $O(x)$, and gives us no problem. Namely,

$$\theta(2x) - \theta(x) = \theta(\lfloor 2x\rfloor) - \theta(\lfloor x\rfloor) = \theta(2\lfloor x\rfloor) - \theta(\lfloor x\rfloor) + O(\log x) = O(\lfloor x\rfloor) + O\log x = O(x).$$

The question is if we can in fact also estimate $\theta(x)$ by $O(x)$. This is in fact true since we can use the following handy trick.

$$\theta(x) = \sum_{k=1}^{\infty}\sum_{\frac{x}{2^k}<p\le\frac{x}{2^{k-1}}}\log p = \sum_{k=1}^{\infty}\Big(\theta(\frac{x}{2^{k-1}}) - \theta(\frac{x}{2^k})\Big) = O(\sum_{i=1}^{\infty}\frac{x}{2^k}) = O(x). \tag{7}$$

This leads us to the following theorem.

Theorem 2.2. *We have*

$$\sum_{p\le x}\frac{\log p}{p} = \log x + O(1).$$

Furthermore we have upper and lower bound for $\theta(x)$. Namely for all $x > 2$ there exists two constants c_1 and c_2 such that

$$c_2 x < \theta(x) < c_1 x.$$

Proof. We start by proving the first statement. Looking at the very similar equation 6, we want to lose the floor function. Clearly the difference between $\lfloor x\rfloor$ and x is less than 1. From this we have

$$\sum_{p\le n}\left\lfloor\frac{n}{p}\right\rfloor\log p = \sum_{p\le n}\frac{n}{p}\log p + O(\sum_{p\le n}\log p).$$

Here the error term is simply the function $\theta(x)$, which means that we can use equation 7 to get

$$\sum_{p\leq n} \frac{n}{p} \log p = n \log n + O(n).$$

After division by n this gives us the desired statement

$$\sum_{p\leq n} \frac{\log p}{p} = \log n + O(1).$$

It may be clear the statement is then true for any x. Simply by the fact we can look at $\lfloor x \rfloor$ and that this differs less than 1 from x itself. □

For the second part of the theorem we notice that equation 7 already gives the desired upper bound. What's left to show is that there is another constant which gives a similar lower bound.

For this notice that

$$\sum_{cx<p\leq x} \frac{\log p}{p} = \log x - \log cx + O(1) = \log \frac{1}{c} + O(1).$$

The constant term is independent so we can just choose a small enough constant c and large enough x, such always have both $cx > 2$ and $\log \frac{1}{c} + O(1) > 5$. From this we derive the inequality

$$5 < \sum_{cx<p\leq x} \frac{\log p}{p} \leq \frac{1}{cx} \sum_{cx<p\leq x} \log p \leq \frac{\theta(x)}{cx}.$$

This gives the desired lower bound, proving the result. Now we can finally deduce an estimate of $\pi(x)$ by giving a similar lower and upper bound as before. It is essentially a corollary of the last theorem, but will be labelled as a theorem since it is an important result on its own.

Theorem 2.3. *Let $\pi(x)$ be the prime counting function defined previously. Then there exist two constants c_1 and c_2 greater than 0 such that for all $x > 2$ we have*

$$c_2 \frac{x}{\log x} < \pi(x) < c_1 \frac{x}{\log x}. \tag{8}$$

Proof. This is not that hard to prove using the last result. For the lower bound we can write

$$\pi(x) = \sum_{p\leq x} 1 \geq \sum_{p\leq x} \frac{\log p}{\log x} = \frac{\theta(x)}{\log x} > c_2 \frac{x}{\log x}.$$

The upper bound is a bit harder to see. For $k \geq 2$ we deduce the inequality

$$\theta(x) > \sum_{x^{1/k}<p\leq x} \log p > \frac{\log x}{k} \sum_{x^{1/k}<p\leq x} 1 = \frac{(\log x)(\pi(x) - \pi(x^{1/k}))}{k}.$$

Reordering terms then leads us to

$$\pi(x) < \frac{k\theta(x)}{\log x} + \pi(x^{1/k}) < \frac{c'x}{\log x} + \pi(x^{1/k}).$$

Now since $\frac{x}{\log x}$ is a faster increasing function than $x^{1/k}$ for $k \geq 2$ we conclude,

$$\pi(x) = O(\frac{x}{\log x})$$

□

2.2 Abel's summation formula

Theorem 2.4 (Abel's summation formula)**.** *Let f be a continuously differentiable function and $(a_n)_{n\in\mathbb{N}}$ a sequence of real numbers. If we define*

$$A(x) := \sum_{1\leq n\leq x} a_n,$$

then

$$\sum_{1\leq n\leq x} a_n f(n) = A(x)f(x) - \int_1^x A(t)f'(t)dt.$$

Proof. The theorem is easily proven by rewriting the term on the left. If we first define $m = \lfloor x \rfloor$ then,

$$\begin{aligned}\sum_{1\leq n\leq x} a_n f(n) &= \sum_{1\leq n\leq x} (A(n) - A(n-1))f(n),\\ &= \sum_{1\leq n\leq x} A(n)f(n) - \sum_{0\leq n\leq x-1} A(n)f(n+1).\end{aligned}$$

Now since $A(0)$ vanishes we have,

$$\begin{aligned}&= \sum_{1\leq n\leq x} A(n)f(n) - \sum_{1\leq n\leq x-1} A(n)f(n+1),\\ &= \sum_{1\leq n\leq x-1} A(n)(f(n) - f(n+1)) + A(m)f(m),\\ &= -\sum_{1\leq n\leq x-1} A(n)\int_n^{n+1} f'(t)dt + A(x)f(x) - (A(x)f(x) + A(m)f(m)).\end{aligned}$$

We can then use the fact that $A(x) = A(m)$ and that $A(t) = A(n)$ on the interval $(n, n+1)$,

$$\begin{aligned}&= -\sum_{1\leq n\leq x-1} \int_n^{n+1} A(t)f'(t)dt + A(x)f(x) - \int_m^x A(t)f'(t),\\ &= A(x)f(x) - \int_1^x A(t)f'(t)dt.\end{aligned}$$

□

Corollary 2.5.

$$\sum_{y\leq n\leq x} a_n f(n) = A(x)f(x) - A(y)f(y) - \int_y^x A(t)f'(t)dt.$$

We are now ready to state Mertens' theorem.

2.3 Mertens' theorem

Theorem 2.6 (Mertens' theorem, 1874)**.** *There exists a constant C such that,*

$$\sum_{p\leq x}\frac{1}{p} = \log\log x + C + O(\frac{1}{\log x}).$$

Proof. We apply Abel's summation formula and use the result of theorem 2.2. Taking $a_p = \sum_{p\leq x}\frac{\log p}{p}$ and $f(x) = \log x$ we derive,

$$\sum_{p\leq x}\frac{1}{p} = \sum_{p\leq x}\frac{\log p}{p}\frac{1}{\log p} = \frac{A(x)}{\log x} - \int_2^x -\frac{A(t)}{t\log^2(t)}dt.$$

By theorem 2.2 that we proved earlier $A(x) = \log x + E(x)$. Substituting this into the equation leads to

$$\sum_{p\leq x}\frac{1}{p} = 1 + \frac{E(x)}{\log x} + \int_2^x \frac{1}{t\log(t)}dt + \int_2^x \frac{E(t)}{t\log^2(t)}dt.$$

Making the substitution $y = \log t$ simplifies the integrals to give us

$$\begin{aligned}\sum_{p\leq x}\frac{1}{p} &= 1 + \frac{E(x)}{\log x} + \int_{\log 2}^{\log x}\frac{dy}{y} + \int_2^\infty \frac{E(t)}{t\log^2(t)}dt - \int_x^\infty \frac{E(t)}{t\log^2(t)}dt \\ &= \log\log x + 1 - \log\log 2 + + \int_2^\infty \frac{E(t)}{t\log^2(t)}dt - \int_x^\infty \frac{E(t)}{t\log^2(t)}dt + \frac{E(x)}{\log x}.\end{aligned}$$

Now since we know $E(t) = O(1)$ we see that the first integral gives us an extra constant and for the error term we have

$$\int_x^\infty \frac{E(t)}{t\log^2(t)}dt + \frac{E(x)}{\log x} \leq \int_{\log x}^\infty \frac{C'}{y^2}dy + \frac{C'}{\log x} = \frac{2C'}{\log x}.$$

Together with the leading terms this completes the proof. □

Another good example which demonstrates Abel's summation formula and at the same time gives us a nice result about Mertens and the prime number theorem is the following. We start with Mertens' theorem in the form we proved as our sum

$$A(x) = \sum_{p\leq x}\frac{1}{p} = \log\log x + C + E(x).$$

Here we name the error term to make calculations easier but it may be clear that

$$E(x) = O(\frac{1}{\log x}).$$

Now we use Merten's theorem for the sum and thus take $f(x) = x$ as our function. Clearly with these choices derivative of $f(x)$ is trivial so the calculations are not that hard. We have

$$\begin{aligned}
\sum_{p \le x} 1 &= \sum_{p \le x} \frac{1}{p} p \\
&= x \log\log x + xC + xE(x) - A(2)f(2) - (\int_2^x \log\log t dt + \int_2^x C dt + \int_2^x E(x) dt), \\
&= x \log\log x + xC + xE(x) - 1 - (x \log\log x - \int_2^x \frac{1}{\log x} dt + Cx - 2C + \int_2^x E(x) dt) \\
&= \int_2^x \frac{1}{\log x} dt + O(\frac{x}{\log x})
\end{aligned}$$

Here the second to last equality follows from integration by parts and we substitute back the definition of $E(x)$ at the end. Now we clearly see that the form of Mertens' theorem we just proved is not equivalent to the prime number theorem. In fact we have found something strictly weaker version which in itself is a meaningless result with such a large error term.

It must be said however that there are version of Mertens theorem which do result in the prime number theorem. Namely a stronger error term in Mertens gives us a stronger one for the prime counting function.

2.4 Perron's formula

In this paragraph we prove Perron's formula which is essentially what we are going to use to derive the prime number theorem.

Lemma 3. *For the real values $c, y, T > 0$, we define*

$$\delta(y) := \begin{cases} 0, & \text{if } y < 1, \\ 1/2, & \text{if } y = 1, \\ 1, & \text{if } y > 1. \end{cases}$$

Then the following holds.

$$\left| \frac{1}{2\pi i} \int_{c-iT}^{c+iT} \frac{y^s}{s} ds - \delta(y) \right| < \begin{cases} y^c \min(1, \frac{1}{T|\log y|}), & \text{if } y \neq 1, \\ \frac{c}{T}, & \text{if } y = 1. \end{cases}$$

Meaning that for $T \to \infty$ the integral is equal to $\delta(y)$.

Proof. We first show the easier part which is when $y = 1$. We make a simple substitution to derive

$$\left| \frac{1}{2\pi i} \int_{c-iT}^{c+iT} \frac{ds}{s} - \frac{1}{2} \right| = \left| \frac{1}{2\pi} \int_{-T}^{T} \frac{dt}{c+it} - \frac{1}{2} \right|.$$

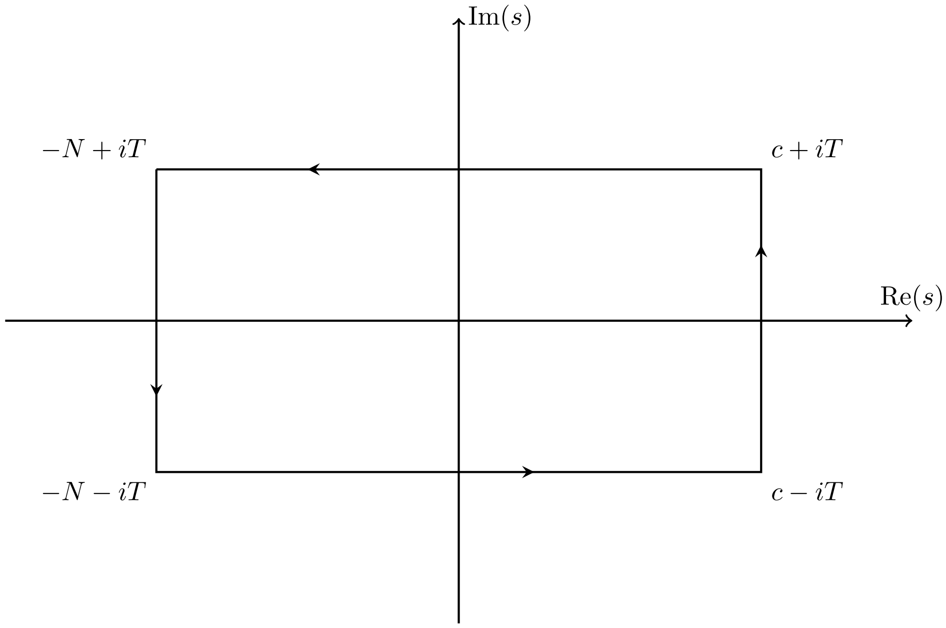
Next we want to split the integral in order to completely get rid of the complex numbers. We notice that

$$\frac{1}{2\pi} \int_{-T}^{T} \frac{dt}{c+it} = \frac{1}{2\pi} \int_0^T \frac{dt}{c+it} + \frac{1}{2\pi} \int_0^T \frac{dt}{c-it} = \frac{1}{\pi} \int_0^T \frac{c}{c^2+t^2} dt.$$

Now again splitting the integral grants us the desired result

$$\begin{aligned}\left|\frac{1}{2\pi i}\int_{c-iT}^{c+iT}\frac{ds}{s}-\frac{1}{2}\right| &= \left|\frac{1}{\pi}\int_{0}^{\infty}\frac{c}{c^2+t^2}dt-\frac{1}{\pi}\int_{T}^{\infty}\frac{c}{c^2+t^2}dt-\frac{1}{2}\right| \\ &= \left|\frac{1}{2}-\frac{1}{\pi}\int_{T}^{\infty}\frac{c}{c^2+t^2}dt-\frac{1}{2}\right| < \left|\frac{1}{\pi}\int_{T}^{\infty}\frac{c}{t^2}dt\right| \\ &< \frac{c}{T}.\end{aligned}$$

For $y \neq 1$ we need some complex analysis and more specifically contour integrals. We won't go into depth into this topic and assume the used results are known. Hence it may be some small details ore omitted. Since both $0 < y < 1$ and $y > 1$ have a very similar proof we will only discuss one of these cases in detail. We take a look at $y > 1$ and choose the contour which has its corners at $c \pm iT$ and $-N \pm iT$ for N a large enough natural number.



Since $\frac{y^s}{s}$ has a simple pole at $s = 0$, by Cauchy's residuum theory we have for the integral over this contour,

$$\oint \frac{y^s}{s}ds = 2\pi i.$$

This contour integral consists of adding 4 normal line integrals which using the triangle inequality gives us

$$\left|\frac{1}{2\pi i}\int_{c-iT}^{c+iT}\frac{y^s}{s}ds-\delta(y)\right| \leq \frac{1}{2\pi}\left|\int_{-N+iT}^{c+iT}\frac{y^s}{s}ds\right|+\frac{1}{2\pi}\left|\int_{-N-iT}^{-N+iT}\frac{y^s}{s}ds\right|+\frac{1}{2\pi}\left|\int_{-N-iT}^{c-iT}\frac{y^s}{s}ds\right|.$$

Here the second integral is easy to estimate by

$$\left|\int_{-N-iT}^{-N+iT}\frac{y^s}{s}ds\right| \leq \left|i\int_{-T}^{T}\frac{y^{-N+it}}{-N+it}dt\right| < \frac{1}{Ny^N}\int_{-T}^{T}1dt = \frac{2T}{Ny^N}.$$

The two other integrals are almost identical so we can estimate them in exactly the same way.

$$\left|\int_{-N\pm iT}^{c\pm iT} \frac{y^s}{s} ds\right| = \left|\int_{-N}^{c} \frac{y^{\omega\pm iT}}{\omega \pm iT} d\omega\right| \leq \frac{1}{T}\int_{-N}^{c} y^\omega d\omega.$$

We created a bound for all three contour integrals. This holds for any N. Hence taking $N \to \infty$ we see that

$$\frac{2T}{Ny^N} \to 0,$$

and that

$$\frac{1}{T}\int_{-\infty}^{c} y^\omega d\omega = \frac{y^c}{T\log y},$$

since $y > 1$.

Using all these results we find

$$\left|\frac{1}{2\pi i}\int_{c-iT}^{c+iT} \frac{y^s}{s} ds - \delta(y)\right| \leq \frac{y^c}{\pi T|\log y|}.$$

For $0 < y < 1$ we can apply the same argument but with a slightly different contour. Choosing N instead of $-N$ means there is no pole inside of our contour, hence by Cauchy we have the integral is equal 0 which coincides with the fact that $\delta(y) = 0$ for $0 < y < 1$. All other steps use very similar estimates as before and we derive the same bound. It looks like we completed the prove and this is indeed the case if $T|\log y| \geq 1$. The only thing left to show is that in the case this is not true we can still say

$$\left|\frac{1}{2\pi i}\int_{c-iT}^{c+iT} \frac{y^s}{s} ds - \delta(y)\right| \leq y^c.$$

For $0 < y < 1$ this is clear when we choose the following smart contour. We draw a line from $c - iT$ to $c + iT$, such that everywhere on that line the real part is greater than c and such that it forms part of a circle around 0 with radius $r^2 = c^2 + t^2$. Then since $y < 1$, is is clear for s on this arch that,

$$\left|\frac{y^s}{s}\right| < \frac{y^c}{r}.$$

Now since there are no poles in our contour, the integral over our contour vanishes. This means that the integral we want to estimate is equal to the integral over the arch K. We derive

$$\left|\frac{1}{2\pi i}\int_K \frac{y^s}{s} ds\right| \leq \frac{1}{2\pi}\frac{y^c}{r}\text{length}(K) < y^c,$$

since the length of the arc is less than πr.

For $y > 1$ we take the opposite contour. Namely the K' the part of the circle you get when leaving out K. So now for all points s on K' we have that $\text{Re}(s) < \text{Re}(c)$. Clearly our pole at $s = 0$ is now within the contour but this only gives an added 1 which is cancelled by the δ-function for $y > 1$. Again the bound

$$\left|\frac{y^s}{s}\right| < \frac{y^c}{r},$$

is viable because we have $y > 1$. We can always estimate the length of K' by $2\pi r$ which as a result gives us the bound of y^c in exactly the same way. This was the last bound needed and hence we proved the lemma. $\square$

Theorem 2.7 (Perron's formula)**.** *If we assume the sum $\sum_{n=1}^{\infty} a(n)n^{-s}$ to be absolutely convergent for $\mathrm{Re}(s) = c$, then Perron's formula says*

$$\sum_{n\leq x}{}' a(n) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}\left(\sum_{n=1}^{\infty} a(n)n^{-s}\right)\frac{x^s}{s}ds.$$

Here the sum is to be interpreted such that when $x \in \mathbb{Z}$, the last term is $\frac{a(x)}{2}$.

Proof.

$$\begin{aligned}\sum_{n\leq x}{}' a(n) &= \sum_{n=1}^{\infty} a(n)\delta(\frac{x}{n}) = \frac{1}{2\pi i}\sum_{n=1}^{\infty} a(n)\int_{c-iT}^{c+iT}\frac{x^s}{sn^s}ds + E \\ &= \frac{1}{2\pi i}\int_{c-iT}^{c+iT}\sum_{n=1}^{\infty}\frac{a(n)}{n^s}\frac{x^s}{s}ds + E\end{aligned}$$

Here the error term follows directly from the lemma we just proved.

$$E = O\left(\sum_{n=1}^{\infty}|a(n)|\left(\frac{x}{n}\right)^c \min(1, \frac{1}{T|\log\frac{x}{n}|})\right).$$

So clearly for $T \to \infty$, we have that $E \to 0$ hence,

$$\sum_{n\leq x}{}' a(n) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}\left(\sum_{n=1}^{\infty} a(n)n^{-s}\right)\frac{x^s}{s}ds,$$

as desired. $\square$

If we want to work with Perron's formula with error term instead of $T \to \infty$ it is needed to bound the error term in a nicer way. As the E given in the proof is very hard to work with. The result can be stated as a corollary.

Corollary 2.8. *Take $c > 0$ such that $\sum_{n=1}^{\infty} a(n)n^{-s}$ is absolutely convergent for $\mathrm{Re}(s) = c$. Now let $T, x > 2$ and define*

$$A_{\max} := \max_{\frac{3}{4}x\leq n\leq\frac{5}{4}x}|a(n)|,$$

then

$$\sum_{n\leq x} a(n) = \frac{1}{2\pi i}\int_{c-iT}^{c+iT}\left(\sum_{n=1}^{\infty}\frac{a(n)}{n^s}\right)\frac{x^s}{s}ds + O\left(\frac{x^c}{T}\sum_{n=1}^{\infty}\frac{|a(n)|}{n^c} + A_{\max} + A_{\max}\frac{x\log x}{T}\right).$$

Proof. We start with

$$E = O\left(\sum_{n=1}^{\infty}|a(n)|\left(\frac{x}{n}\right)^c \min(1, \frac{1}{T|\log\frac{x}{n}|})\right).$$

To try to create a nicer bound for the error term we will approach it in three different regions. Namely

$$|n-x| \le 2, \qquad 2 < |n-x| \le \frac{1}{4}x, \qquad |n-x| > \frac{1}{4}x.$$

In the first region the error term clearly becomes $O(A_{\max})$, since $\frac{x}{n} \approx 1$ and we sum over 4 numbers at most.

Now the second region is the trickiest. First since n is farily close to x and c is a constant greater than 0, we have for some $K > 0$,

$$\left(\frac{x}{n}\right)^c < K\frac{x}{n}.$$

The corresponding sum is then much eachter to bound by an integral. We derive

$$\begin{aligned}
\sum_{|n-x|\le\frac{1}{4}x} |a(n)| \left(\frac{x}{n}\right)^c \min(1, \frac{1}{T|\log\frac{x}{n}|}) &< \frac{K'A_{\max}}{T} \sum_{|n-x|\le\frac{1}{4}x} \frac{x}{n|\log\frac{x}{n}|}, \\
&< \frac{K'A_{\max}}{T}\int_{\frac{3}{4}x}^{x-2} \frac{x}{t\log\frac{x}{t}}dt + \frac{K'A_{\max}}{T}\int_{x+2}^{\frac{5}{4}x} -\frac{x}{t\log\frac{x}{t}}dt, \\
&= \frac{K'A_{\max}}{T}\left(\left[-x\log\log\frac{x}{t}\right]_{\frac{3}{4}x}^{x-2} + \left[x\log\log\frac{x}{t}\right]_{x+2}^{\frac{5}{4}x}\right), \\
&< \frac{K''A_{\max}}{T}\, x\log x.
\end{aligned}$$

We conclude the error term is

$$O\left(A_{\max}\frac{x\log x}{T}\right), \qquad \text{for } |n-x| \le \frac{1}{4}x.$$

In third region we can clearly have the upper bound $\frac{1}{T|\log\frac{x}{n}|} < \frac{C}{T}$, because $\frac{x}{n}$ is sufficiently far from 1. Hence we derive

$$O\left(\frac{x^c}{T}\sum_{n=1}^{\infty}\frac{|a(n)|}{n^c}\right), \qquad \text{for } |n-x| > \frac{1}{4}x.$$

All together this completes the proof. $\square$

Perron's formula gives us a new way to evaluate sums. The question is how we can use this to get to the prime number theorem. A logical first step would be to just look at the prime counting function. Perron's formula then tells us we will need to know the following sum.

$$\sum_{n=1}^{\infty} a(n)n^{-s} = \sum_p p^{-s}.$$

This sum is closely related to the Riemann zeta function as we saw in the proof of theorem 2.1. Using the fact that for $\mathrm{Re}(s) > 1$ we have

$$\zeta(s) = \prod_p \frac{1}{1-p^{-s}},$$

a derivation similar to the proof of theorem 2.1 gives us

$$\log \zeta(s) = \sum_p \sum_{k=1}^{\infty} \frac{1}{kp^{ks}} = \sum_p \frac{1}{p^s} + \sum_p \sum_{k=2}^{\infty} \frac{1}{kp^{ks}}. \tag{9}$$

So we found the sum we we were aiming for. The next step will be, like in the $s = 1$ case, to create a bound for the other sum and showing it converges. We have

$$\sum_{k=2}^{\infty} \frac{1}{kp^{ks}} \leq \sum_{k=2}^{\infty} \frac{1}{p^{ks}} = \frac{1}{p^{2s}} \sum_{k=0}^{\infty} \frac{1}{p^{sk}} = \frac{1}{p^{2s}} \cdot \frac{1}{1 - 1/p^s}.$$

Now since $s > 1$ it may be clear that

$$\frac{1}{p^{2s}} < \frac{1}{p^2} \qquad \text{and} \qquad \frac{1}{1 - 1/p^s} < \frac{1}{1 - 1/p} < 2.$$

We conclude that

$$\sum_p \sum_{k=2}^{\infty} \frac{1}{kp^{ks}} < C.$$

for some real constant. Using the error notation this means that for $\mathrm{Re}(s) > 1$,

$$\log \zeta(s) = \sum_p \frac{1}{p^s} + O(1).$$

It appears that according to Perron's formula we should work with $\log \zeta(s)$, when approximating the prime counting function. The problem here is that while for $\mathrm{Re}(s) > 1$ everything is fine, this will not be the case for $\mathrm{Re}(s) < 1$. Here $\zeta(s)$ has more zeros which lead to singularities for the logarithm. The trick is to consider

$$\frac{\zeta'(s)}{\zeta(s)} = \frac{d}{ds} \log \zeta(s).$$

Here the worst that can happen is for the function to have simple poles and we know this won't be a problem since we can still do contour integrals and such.

Now the question is, apart from having simple poles at most, does $\frac{\zeta'(s)}{\zeta(s)}$ resemble any relevant information. We see that in fact in does, namely

$$\frac{d}{ds} \log \zeta(s) = \sum_p \sum_{k=1}^{\infty} \frac{d}{ds} \frac{1}{kp^{ks}} = \sum_p \sum_{k=1}^{\infty} -\frac{\log p}{p^{ks}}.$$

If we write out the second sum we see

$$-\frac{\zeta'(s)}{\zeta(s)} = \sum_p \log(p)(p^{-s} + p^{-2s} + p^{-3s} + p^{-4s} + ...) = \sum_{n=1}^{\infty} \Lambda(n) n^{-s}.$$

Here

$$\Lambda(n) = \begin{cases} \log p, & \text{if } n = p^k \text{ for } p \text{ a prime,} \\ 0, & \text{otherwise.} \end{cases}$$

has exactly the desired properties and is called the von Mangoldt function. Hence the easiest way towards the prime number theorem will be by calculating the sum over the von Mangoldt function using Perron's formula. From this we will try to alter the sum to get the corresponding result for the prime counting fuction.

We are left with the following integral

$$\sum_{n\leq x}\Lambda(n) = -\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}\frac{\zeta'}{\zeta}(s)\frac{x^s}{s}ds.$$

This should give us a very good approximation of sum of the von Mangoldt function. Before we do this we can already give a much weaker statement. Namely since

$$\begin{aligned}\sum_{n\leq x}\Lambda(n) &= \sum_{p\leq x}\log p + \sum_{\substack{p,\\ p^2\leq x}}\log p + \sum_{\substack{p,\\ p^3\leq x}}\log p + ...,\\ &= \sum_{k\geq 1}\theta(x^{1/k}),\\ &= \theta(x) + \sum_{k\geq 2}\theta(x^{1/k}),\\ &= \theta(x) + O(\sqrt{x}).\end{aligned}$$

Here the last line follows directly from equation 7. We are now as good as ready to state the theorem itself.

2.5 The prime number theorem

In this section we will give a proof sketch of the prime number theorem. The parts which relate closely to what was discussed before and give a good understanding of the topic will be discussed more broadly while some technical details may be omitted.

We first want to discuss some the zero's of the Riemann zeta function. By the Euler product we know that $\zeta(s)$ does not vanish for any s with real part greater than 1. We also know that it has a pole at $s = 1$. The next step is to examine what happens at the rest of the line $s = 1 + it$ for $t \neq 0$. We will do this by approaching it from right side using the Euler product.

Starting with the expression for the logarithm of $\zeta(s)$ from 9 we have

$$\zeta(s) = \exp\sum_{p}\sum_{k=1}^{\infty}\frac{p^{-ks}}{k}.$$

Writing $s = \sigma + it$ we can split the real and imaginary part to get the following expression.

$$\begin{aligned}
|\zeta(\sigma + it)| &= |\exp \sum_p \sum_{k=1}^{\infty} \frac{p^{-k\sigma}}{k} p^{-ikt}| \\
&= \exp \sum_p \sum_{k=1}^{\infty} \frac{p^{-k\sigma}}{k} |\exp(-\log(p)ikt)| \\
&= \exp \sum_p \sum_{k=1}^{\infty} \frac{p^{-k\sigma}}{k} \cos(kt \log p)
\end{aligned}$$

On first glance this is not a very nice expression since it only seems to give us something even harder with the introduction of the cosine. The trick is the following very simple but essential inequality. We can write for all real numbers α,

$$3 + 4\cos\alpha + \cos 2\alpha = 2 + 4\cos\alpha + 2\cos^2\alpha = 2(1 + \cos\alpha)^2 \geq 0 \tag{10}$$

Clearly $\alpha = kt \log p \in \mathbb{R}$. Hence the question remains how do relate this to our expression found for $\zeta(\sigma + it)$. After noticing the fact that

$$|\zeta(\sigma + it)^4| = \exp\left(4 \sum_p \sum_{k=1}^{\infty} \frac{p^{-k\sigma}}{k} \cos(kt \log p)\right) = \exp \sum_p \sum_{k=1}^{\infty} \frac{p^{-k\sigma}}{k} 4\cos\alpha,$$

it is not that hard to find the other two expression needed for us two use our inequality. We derive that

$$|\zeta(\sigma)|^3 |\zeta(\sigma + it)|^4 |\zeta(\sigma + 2it)| = \exp \sum_p \sum_{k=1}^{\infty} \frac{p^{-k\sigma}}{k} (3 + 4\cos\alpha + \cos 2\alpha) \geq 1. \tag{11}$$

While this may seem an fairly unimportant inequality at first, it was actually one of the most important discoveries in the original proof the prime number theorem.

Namely if now assume that ζ has a pole on the line with real part equal to 1. Thus $\zeta(1 + it_0) = 0$ for some t_0. Then $\zeta(1 + it_0)^4$ has a zero of order at least 4. Since $\zeta(s)$ has a simple pole at $s = 1$ this would mean that

$$\lim_{\sigma \downarrow 1} \zeta(\sigma)^3 \zeta(\sigma + it_0)^4 = 0.$$

Now since $\zeta(1 + 2it_0)$ is constant this result directly contradicts what we just derived in equation 11. Hence we can conclude the following which was instrumental in proving the prime number theorem.

Lemma 4. *For all real numbers $t \neq 0$,*

$$\zeta(1 + it) \neq 0.$$

In fact this result is equivalent to the prime number theorem. Intuitively the idea why this is indeed the case comes down to the following. When using Perron's formula, we want to evaluate $\frac{\zeta'}{\zeta}(s)$. Now we know that this expression has simple poles at the poles and zeros of $\zeta(s)$. Thus any zero on the line $\sigma = 1$ will give us a simple pole. And as we will see later, these poles would in appear in the by us chosen contour. Meaning that we would get multiple residues terms which would definitely alter the result.

To approach it in a somewhat clearer and more formal manner we can look at the "explicit formula" for $\psi(x) = \sum_{n \leq x} \Lambda(n)$. The explicit formula is the most used and also the strongest method for proving the prime number theorem. We will not prove or discuss it in this thesis but it is helpful to look at a simple form of it in this context. We have

$$\psi(x) = x - \sum_{\rho} \frac{x^{\rho}}{\rho} - \frac{\zeta'(0)}{\zeta(0)} \tag{12}$$

Here the sum is over all the zeros ρ of the the Riemann-zeta function hence it may be clear that the zeros of the Riemann-zeta function are of large importance. It is also not hard to see that them being on the line $\sigma = 1$ has some significance. Clearly if $\mathrm{Re}(\rho) = 1$ we see that

$$\frac{x^{\rho}}{\rho} = O(x),$$

while at the same time for $\mathrm{Re}(\rho) < 1$ we have

$$\frac{x^{\rho}}{\rho} = o(x).$$

Here the little o-notation means that in this case

$$\lim_{x \to \infty} \frac{x^{\rho}}{\rho} \Big/ x = 0.$$

Intuitively what this means is that any zero on the line $\sigma = 1$ would result in a error term which is just not small enough. Hence it would result in a weaker version of the prime number theorem.

From this lemma follows a more technical one which in fact is the last step towards the prime number theorem. We will state this lemma without prove. The details can be read in [1].

Lemma 5. *There exists an $\delta > 0$ such that for all the $s = \sigma + it \in \mathbb{C}$ with real part*

$$\sigma \geq 1 - \delta \min\left(1, (\log|t|)^{-9}\right),$$

it holds that $\zeta(s) \neq 0$. Thereby for an s that also satisfies $|s - 1| \geq 2$,

$$\frac{\zeta'}{\zeta}(s) = O((\log|t|)^{9}).$$

The first part of this lemma makes sense since we already saw that for $\sigma \geq 1$, the zeta function has no zeros. The second part gives us a nice way to bound the $\frac{\zeta'}{\zeta}(s)$ in different regions of our contour. We know have everything needed for us to state the prime number theorem.

Theorem 2.9 (Prime number theorem, 1896)**.** *Let C be a positive constant and $\pi(x)$ the prime counting function. Then*

$$\pi(x) = \int_2^x \frac{1}{\log t} dt + O\left(x e^{-C(\log x)^{\frac{1}{10}}}\right)$$

Proof. As discussed in the previous section the best way to approach this is via $\psi(x) = \sum_{n \le x} \Lambda(n)$. Starting from corollary 2.8 we have

$$\psi(x) = -\frac{1}{2\pi i} \int_{c-iT}^{c+iT} \frac{\zeta'}{\zeta}(s) \frac{x^s}{s} ds + O\left(\frac{x^c}{T} \left| \frac{\zeta'}{\zeta}(c) \right| + \log x + \frac{x(\log x)^2}{T} \right). \tag{13}$$

To extract the main term we need to evaluate the integral. Like in the proof of Perron's formula we will draw a rectangular contour with corners $c \pm iT, c' \pm iT$. We choose $c = 1+\lambda$ and $c' = 1-\lambda$. In order for us to have a contour which makes calculations relatively easy, we want all zeros of $\zeta(s)$ to be out of it. Hence by the lemma just stated, 5, we set

$$\lambda = \frac{\delta}{(\log T)^9}.$$

The only pole in our domain is at $s = 1$. It is clear that $\frac{\zeta'}{\zeta}(s)$ only has a simple pole here with residue equal to -1. Hence our residue has the integrand has residue $-x$ at the point $s = 1$. By the residue theorem we derive the following

$$\begin{aligned} -x = &\frac{1}{2\pi i} \int_{c-iT}^{c+iT} \frac{\zeta'}{\zeta}(s) \frac{x^s}{s} d + \frac{1}{2\pi i} \int_{c+iT}^{c'+iT} \frac{\zeta'}{\zeta}(s) \frac{x^s}{s} ds, \\ &+ \frac{1}{2\pi i} \int_{c'+iT}^{c'-iT} \frac{\zeta'}{\zeta}(s) \frac{x^s}{s} ds + \frac{1}{2\pi i} \int_{c'-iT}^{c-iT} \frac{\zeta'}{\zeta}(s) \frac{x^s}{s} ds. \end{aligned}$$

We now want to estimate the last three terms and contain them all in one big error term in order to estimate $\psi(x)$. Since we chose our λ in a smart way we see that for all s on the contour boundary which satisfy $|s-1| \ge 2$ we have

$$\frac{\zeta'}{\zeta}(s) = O(\lambda^{-1}).$$

Firstly the second and fourth integral are roughly identical. Since on these two paths $|\frac{1}{s}| \le \frac{1}{T}, |x^s| < x^c$ and $\left|\frac{\zeta'}{\zeta}(s)\right| = O(\lambda^-1)$, we can quite easily bound the integral. Using that the path length is 2λ we have,

$$\left| \frac{1}{2\pi i} \int_{c+iT}^{c'+iT} \frac{\zeta'}{\zeta}(s) \frac{x^s}{s} ds \right| + \left| \frac{1}{2\pi i} \int_{c'-iT}^{c-iT} \frac{\zeta'}{\zeta}(s) \frac{x^s}{s} ds \right| = O\left(\frac{x^{1+\lambda}}{T} \right).$$

For the third path integral we use a slightly different estimate since the path comes very close to our pole at $s = 1$. We start by parametrising the path by

$$s = 1 - \lambda + it, \qquad -T \le t \le T.$$

Now we know that

$$|s| \ge \frac{\sqrt{2}}{2}(1 - \lambda + |t|).$$

This follows from basic geometry. Namely we get equality in the case $1 - \lambda = |t|$, since in this case we have an isosceles triangle with ratio $1 : 1 : \sqrt{2}$. Because this is the case when the sum of the two right-angled sides is relatively longest, the inequality follows. We conclude

$$|s| \geq \frac{\sqrt{2}}{2}(1 - \lambda + |t|) > \frac{1}{2}(1 + |t|).$$

Now for $|t| \geq 2$ we have by lemma 5

$$\left|\frac{\zeta'}{\zeta}(1 - \lambda + it)\right| = O(\lambda^{-1})$$

For $|t| \leq 2$ we get very close to the simple pole at $s = 1$ hence we know that in this range

$$\left|\frac{\zeta'}{\zeta}(s)\right| = O(\frac{1}{s-1}).$$

So for $|t| < 2$ we indeed have the same bound. Hence on the whole path

$$\left|\frac{\zeta'}{\zeta}(1 - \lambda + it)\right| = O(\lambda^{-1}).$$

Now since clearly $|x^s| < |x^{c'}| = |x^{1-\lambda}|$ we can bound the last integral in the following way.

$$\left|\frac{1}{2\pi i}\int_{c'+iT}^{c'-iT} \frac{\zeta'}{\zeta}(s)\frac{x^s}{s}ds\right| < \frac{x^{c'}}{\pi}\int_{-T}^{T}\left|\frac{\zeta'}{\zeta}(c' + it)\right|\frac{1}{1+|t|}dt.$$

Since we have a bound for the zeta part of the integrand we can integrate what is left to get

$$\left|\frac{1}{2\pi i}\int_{c'+iT}^{c'-iT} \frac{\zeta'}{\zeta}(s)\frac{x^s}{s}ds\right| = O(x^{1-\lambda}\lambda^{-1}\log T) = O(x^{1-\lambda}(\log T)^{10}).$$

Here the last equality follows by our smart to choice of λ.

Now we estimated all three integral in a error term. Combining this with error term we already have for $\psi(x)$, see equation 13, we see that we are fairly close to the result. Of course $\left|\frac{\zeta'}{\zeta}(c)\right| = O(\lambda^{-1})$ in the same way it followed for c'. Hence putting it all together gives

$$\psi(x) = x + O\left(\frac{x^{1+\lambda}}{T\lambda} + \log x + \frac{x(\log x)^2}{T} + x^{1-\lambda}(\log T)^{10}\right)$$

We can now choose

$$T = \exp(\delta^{-1/9}(\log x)^{1/10}).$$

Here the δ is the one form lemma 5 and can always be chosen small enough for $T > 2$. The question is what happens to the error terms when T is chosen in such a way. First of all

$$\lambda = \frac{\delta}{(\log T)^9} = \frac{\delta^2}{(\log x)^{9/10}}.$$

Now after some substitutions we derive for all different terms within the error term.

$$\begin{aligned}
\frac{x^{1+\lambda}}{T\exp(\log\lambda)} &= x\frac{\exp(\lambda\log x)}{T\lambda} = x\exp((\delta^2 - \delta^{-1/9})(\log x)^{1/10} - \log(\delta^2(\log x)^{-9/10})) \\
\frac{x(\log x)^2}{T} &= x\exp(2\log\log x - \delta^{-1/9}(\log x)^{1/10}) \\
x^{1-\lambda}(\log T)^{10} &= x\exp(-\delta^2(\log x)^{1/10} + \log(\delta^{-10/9}\log x))
\end{aligned}$$

Now these terms become a complete mess when working them out. We wont try to bound or combine even further. What is easier is to notice that clearly $\exp(-\delta^2) < \exp(-\delta^3)$. Looking at all three error terms then certainly all are contained in

$$O(x \exp(-\delta^3 (\log x)^{1/10}).$$

Hence we can conclude

$$\psi(x) = x + O\left(x \exp(-\delta^3 (\log x)^{1/10}\right) \tag{14}$$

This is completely equivalent to the prime number theorem but because we made the statement for the prime counting function we want to alter it a little bit. First remember that at the end of the last section we derived

$$\psi(x) = \theta(x) + O(\sqrt{x})$$

where $\theta(x) = \sum_{p \leq x} \log p$. This gives us the following expression, which is also an equivalent form of the prime number theorem.

$$\sum_{p \leq x} \log p = x + O\left(x \exp(-\delta^3 (\log x)^{1/10}\right)$$

Here the error term $\theta(x)$ is simply absorbed in the bigger one. Now we can use Abel's summation formula to finally get the result for the prime counting function.

$$\begin{aligned}
\pi(x) = \sum_{p \leq x} 1 = \sum_{p \leq x} \log p \frac{1}{\log p} \\
&= \frac{\theta(x)}{\log x} - \int_2^x \theta(t) \left(\frac{1}{\log t}\right)' dt \\
&= \frac{x}{\log x} - \int_2^x t \left(\frac{1}{\log t}\right)' dt + O\left(x \exp(-\delta^3 (\log x)^{1/10}\right).
\end{aligned}$$

Here the error term coming from the integral has a $(\log x)^2$ saving extra and is therefore strictly better hence absorbed by the other one. By integration by parts we also get the right main term

$$\pi(x) = \frac{x}{\log x} - \frac{x}{\log x} + \int_2^x \frac{1}{\log t} dt + O\left(x \exp(-\delta^3 (\log x)^{1/10}\right).$$

Hence the theorem follows for a constant $C > 0$. $\square$

It is nice to check if this is indeed stronger than Mertens theorem. By Abel's summation formula

$$\begin{aligned}
\sum_p \frac{1}{p} = \sum_p 1 \frac{1}{p}, \\
&= \frac{\pi(x)}{x} - \int_2^x \pi(t) \left(\frac{1}{t}\right)' dt, \\
&= \frac{\mathrm{Li}(x)}{x} - \int_2^x \mathrm{Li}(t) \left(\frac{1}{t}\right)' dt + \frac{E(x)}{x} + \int_2^x \frac{E(t)}{t^2} dt, \\
&= \frac{\mathrm{Li}(x)}{x} - \frac{\mathrm{Li}(t)}{t}\Big|_2^x + \int_2^x Li'(t) \frac{1}{t} dt + O\left(\exp(-C(\log x)^{1/10}\right), \\
&= \log\log x + M + O\left(\exp(-C(\log x)^{1/10}\right).
\end{aligned}$$

A result that is strictly stronger than the form of Mertens theorem we derived.

3 Rough numbers

We have studied primes and how they are distributed along all numbers. It is now interesting to look at somewhat different numbers. In 1990 Greene and Knuth defined a natural n as unusual if all its prime divisors were greater or equal to $\sqrt{n}$, see [6]. It was only in 2001 that Finch called these unusual numbers $\sqrt{n}$-rough numbers and the term rough numbers was formally introduced, [7], [8]. More generally a $y-$*rough number* was defined as any integer having prime divisors only equal to or larger than y.

Now trivially all integers will be 2-rough. Thereby for a very large y the first multiple rough numbers will all be primes while the first few 5-rough numbers will also consist of some multiple of 5 or 7. In general however we are not really interested in which numbers are rough and which are not. We want to follow up on the first chapter in which we did not care if a specific number was prime. What interested us was how many there were in a certain range and how they were distributed. We want to do the same here. How many y-rough numbers are there in a certain range and how do they behave asymptotically dependent of y. There are things we would like to look at in this chapter. We will follow the approaches and structure of chapter 7 of [3] and chapter 6 of [2], written by Montgomery and Tenenbaum respectively.

3.1 Smooth numbers

Before we look more into rough number it is nice to define its counterpart as well.

Definition 3.1. A *y-smooth number* is an integer whose primes are at most equal to y. With the corresponding counting function

$$\psi(x, y) := \#\{n \leq x : \ p|n \implies p \leq y\}. \tag{15}$$

The smooth numbers are the opposites of the rough numbers. Compared to the rough numbers there has been done a lot more research into smooth numbers. We will not go into depth about these numbers but we will state some important results because they will be referenced later. When we get similar results for rough numbers it will also be nice to see the resemblance.

Before we derive any asymptotic or other result we first need to define a new variable u.

$$u := \frac{\log x}{\log y} \tag{16}$$

We will see the results in this chapter mainly depend on the ratio between $\log x$ and $\log y$ so that is why defining u makes everything much easier. The important functions in this chapter will therefore be defined with u as a variable.

Certainly $\psi(x, y)$ is trivial for $0 < u < 1$ because in this case $y > x$. Hence any integer smaller than x has only prime divisors smaller than y. In general however it is not that simple. To count smooth numbers we need the following famous function by Dickman in 1930, [9].

Definition 3.2. We define the *Dickman function* $\rho(u)$ as the unique continuous function on $[0, \infty)$ which satisfies the delayed differential equation

$$u\rho'(u) = -\rho(u-1) \qquad \text{for } u > 1 \tag{17}$$

with initial condition

$$\rho(u) = 1, \qquad \text{for } 0 \leq u \leq 1.$$

Using this definition we get the following result for smooth numbers.

Theorem 3.3. *For any $U \geq 0$ and all $0 \leq u \leq U$ with $x \geq 2$ we have*

$$\phi(x, x^{1/u}) = \rho(u)x + O_U(\frac{x}{\log x}) \tag{18}$$

Here the subscript under the error term means the constant is dependent of U. Now we see this results holds for $u < 1$ as well since in this case the amount of smooth numbers is simply equal to x.

3.2 Rough numbers for small y

For smooth numbers we have seen there is an unique defined function which counts them. We will want to do the same for rough numbers but before we do that we first want to approach it in a somewhat different we way.

Definition 3.4. We define the *Möbius function* $\mu(n)$ for any positive integer n as

$$\mu(n) = \begin{cases} 0, & \text{if n has a squared prime factor} \\ 1, & \text{if n=1} \\ (-1)^k, & \text{if n is squarefree and has k prime divisors} \end{cases} \tag{19}$$

Clearly the function is multiplicative,

$$\mu(kl) = \mu(k)\mu(l),$$

when k and l are coprime. This follows from an easy check. Thereby in the case they're not coprime certainly $\mu(kl) = 0$ by definition.

The Möbius function has the following important property.

Lemma 6.

$$\sum_{d|n} \mu(d) = \begin{cases} 1 & \textit{if } n = 1, \\ 0 & \textit{otherwise.} \end{cases} \tag{20}$$

Proof. To see this notice we use that for any $n > 1$ we can write

$$n = p_1^{m_1} ... p_l^{m_l}.$$

Clearly any $d|n$ is then of the form $n = p_1^{m'_1} ... p_l^{m'_l}$. where m'_i can be any integer from 0 to m_i. By definition $\mu(d)$ vanishes when d is not squarefree. Thus we see for $n > 1$

$$\sum_{d|n} \mu(d) = \sum_{\substack{(m'_1, ..., m'_l) \\ m'_i = 0 \text{ or } 1}} \mu(p_1^{m'_1} ... p_l^{m'_l})$$

This is just combinatorics because we have exactly 1 d which consists of l primes, $\binom{l}{l-1}$ different d which consist of $l-1$ primes, etc. So we derive

$$\begin{aligned}\sum_{d|n}\mu(d) &= (-1)^l + \binom{l}{l-1}(-1)^{l-1} + ... - \binom{l}{1} + 1 \\ &= (1-1)^l \\ &= 0\end{aligned}$$

Here we used the binomial theorem theorem

$$\sum_{n=0}^{l}\binom{l}{n}x^n = (1+x)^l.$$

□

To count rough numbers we define a function similar to the one in the smooth number case.

Definition 3.5. We define a function which counts the y-rough numbers bounded by x.

$$\phi(x,y) := \#\{n \leq x : \ p|n \implies p \geq y\} \tag{21}$$

To try to calculate $\phi(x,y)$ we will first try a very straightforward approach using the Möbius function. Let us first define the product over all primes up until y.

$$P(y) = \prod_{p\leq y}.$$

Then clearly using lemma 6

$$\begin{aligned}\phi(x,y) &= \sum_{\substack{n\leq x \\ \gcd(n,P(y)=1)}} 1 \\ &= \sum_{n\leq x}\sum_{d|\gcd(n,P(y))}\mu(d) \\ &= \sum_{d|P(y)}\mu(d)\sum_{\substack{n\leq x \\ d|n}} 1 \\ &= \sum_{d|P(y)}\mu(d)\left(\frac{x}{d}+O(1)\right) \\ &= x\sum_{d|P(y)}\frac{\mu(d)}{d} + O\left(\sum_{d|P(y)} 1\right)\end{aligned}$$

Now notice that

$$\prod_{p\leq y}(1-\frac{1}{p}) = 1 - \frac{1}{2} - \frac{1}{3} + \frac{1}{2\times 3} - \frac{1}{5} + \frac{1}{2\times 5} + \frac{1}{3\times 5} - \frac{1}{2\times 3\times 5}...$$

We get a minus for an odd number of primes and a plus for an even number. So we see

$$\begin{aligned}\phi(x,y) &= x\prod_{p\le y}(1-\frac{1}{p}) + O(2^{\pi(y)})\\ &= x\prod_{p\le y}(1-\frac{1}{p}) + O(2^{\frac{y}{\log y}}).\end{aligned}$$

For very small y this result makes very much sense. Take for example $y = 3$ then certainly $\phi(x,y) = \frac{1}{6}x$. Now we see that for increasing y, the error term rapidly catches up with the main term and the result becomes trivial. Hence this result is not very useful in the broader sense but still interesting when y is of the order $\log\log x$.

3.3 The Buchstab function

We now want to deduce a more general expression for $\phi(x,y)$. We start with an important lemma by Buchstab, [12].

Lemma 7 (Buchstab's identity)**.** *For $x \ge z \ge y \ge 2$,*

$$\phi(x,y) = \phi(x,z) + \sum_{y\le p<z}\phi(\frac{x}{p},p) \tag{22}$$

Proof. Now the proof is relatively simple. Namely the amount of rough numbers which have p as their smallest prime is exactly $\phi(\frac{x}{p},p)$. We can then just sum over all these p and the identity follows directly. □

Lets try to use this to calculate $\phi(x,y)$. Well for $y > x^{1/2}$ the result is trivial and we get

$$\phi(x,y) = \pi(x) - \phi(y) + 1.$$

Notice 1 is always counted as a rough number since it is the empty product of primes. Now lets take the next step and look what happens when $x^{1/3} < y \le x^{1/2}$. We can use our identity and see

$$\phi(x,y) = \phi(x,x^{1/2}) + \sum_{y\le p<x^{1/2}}\phi(\frac{x}{p},p)$$

The nice thing is that we just found the trivial result for the first term. For the second term notice that for all p which we sum over, we have $p > \left(\frac{x}{p}\right)^{1/2}$. Hence using the estimate

$$\pi(x) = \frac{x}{\log x} + O\left(\frac{x}{(\log x)^2}\right)$$

we can write

$$\begin{aligned}\phi(x,y) &= \frac{x}{\log x} + \sum_{y\le p<x^{1/2}}\frac{x}{p\log(x/p)} - \frac{y}{\log y} + O\left(\frac{x}{(\log x)^2}\right)\\ &= \frac{x}{\log x}\left(1 + \sum_{y\le p<x^{1/2}}\frac{\log x}{p\log(x/p)}\right) - \frac{y}{\log y} + O\left(\frac{x}{(\log x)^2}\right)\end{aligned}$$

Using Abel's summation formula with

$$f(t) = \frac{\log x}{t\log(x/t)}$$

and the strong from of the prime number theorem we get

$$\begin{aligned}
\phi(x,y) &= \frac{x}{\log x}\left(1 + f(t)\pi(t)\Big|_y^{x^{1/2}} - \int_y^{x^{1/2}} f'(t)\pi(t)dt\right) - \frac{y}{\log y} + O\left(\frac{x}{(\log x)^2}\right) \\
&= \frac{x}{\log x}\left(1 + \int_y^{x^{1/2}} f(t)\mathrm{Li}'(t)dt\right) - \frac{y}{\log y} + O\left(\frac{x}{(\log x)^2}\right) \\
&= \frac{x}{\log x}\left(1 + \int_y^{x^{1/2}} \frac{\log x}{t\log t\log(x/t)}dt\right) - \frac{y}{\log y} + O\left(\frac{x}{(\log x)^2}\right) \\
&= \frac{x}{\log x}\left(1 - \log\log y + \log(\log x - \log y)\right) - \frac{y}{\log y} + O\left(\frac{x}{(\log x)^2}\right) \\
&= \frac{x(1+\log(u-1))}{\log x} - \frac{y}{\log y} + O\left(\frac{x}{(\log x)^2}\right).
\end{aligned}$$

This is a very lengthy calculation but in the end it gives us an indication for what we are looking for. If we now define a function $\omega(u)$ by $u\omega(u) = 1$ for $1 \le u \le 2$ and

$$u\omega(u) = 1 + \log(u-1), \qquad \text{for} \quad 2 < u \le 3.$$

Then for $x^{1/3} \le y \le x$

$$\phi(x,y) = \frac{x\omega(u) - y}{\log y} + O\left(\frac{x}{(\log x)^2}\right).$$

Clearly in both these ranges the formula checks out. It seems the $\omega(u)$ is defined completely different for both ranges and therefore doesn't have much interest if we cannot connect them. Hence notice that for $2 < u \le 3$

$$(u\omega(u))' = \omega(u-1).$$

This leads us to the following definition by Buchstab, [12].

Definition 3.6. We define the *Buchstab function* $\omega(u)$ as the unique continuously solution to the delayed differential equation

$$(u\omega(u))' = \omega(u-1), \qquad u > 2 \tag{23}$$

with the initial condition for $1 \le u \le 2$

$$u\omega(u) = 1.$$

We can also write the Buchstab function slightly different using the integral from. We then have

$$\omega(u) = \frac{1}{u} + \int_1^{u-1} \omega(v)dv, \qquad u > 2. \tag{24}$$

For the amount of rough numbers we then have the following theorem by Buchstab.

Theorem 3.7 (Buchstab)**.** *Taking $x \geq y \geq 2$. Then uniformly for $1 \leq u \leq U$ we have*

$$\phi(x, y) = \frac{x\omega(u) - y}{\log y} + O_U\left(\frac{x}{(\log x)^2}\right) \tag{25}$$

Proof. We will prove this by induction. For this we only need to look at the induction step since we have seen the result for $u < 3$. Firstly the second main term is only relevant when y comes very close to x since otherwise it is absorbed into the error term. Hence it is only in the trivial case $u < 2$ that this term is of any significance.

Now we assume the result holds for $1 \leq u \leq k$ and try to prove it for $u < k + 1$. Here $k \leq U$. We start in exactly the same way as before

$$\phi(x, x^{1/u}) = \phi(x, x^{1/k}) + \sum_{x^{1/u} < p < x^{1/k}} \phi(\frac{x}{p}, p)$$

Now clearly we can use our induction hypothesis for the first term. Luckily we see for the second term that

$$p^k > x^{\frac{k}{k+1}} = x^{1 - \frac{1}{k+1}} > \frac{x}{p}.$$

Hence we can use it here as well and we derive,

$$\phi(x, x^{1/u}) = \frac{k\omega(k)x}{\log x} + \sum_{x^{1/u} < p < x^{1/k}} \frac{x\omega\left(\frac{\log x}{\log p} - 1\right)}{p \log p} + O_k\left(\frac{x}{(\log x)^2}\right).$$

To see this error term is correct notice that we get an added error term from the sum.

$$\sum_{x^{1/u} < p < x^{1/k}} O\left(\frac{x}{p(\log \frac{x}{p})^2}\right).$$

Clearly

$$\frac{x}{(\log \frac{x}{p})^2} < \frac{x}{(\log x - \log x^{1/k})^2} = \frac{1}{(1 - \frac{1}{k})^2} \frac{x}{(\log x)^2}.$$

For the sum over $1/p$ we can approximated it by Mertens theorem and is of the order

$$\log \log x^{1/k} - \log \log x^{\frac{1}{k+1}} = \log(\frac{k+1}{k}).$$

Hence we see that we do in fact get the correct error term with the constant having some k dependence. This happens every inductive step hence we get something added to the error term every single time. Luckily since we only prove the theorem for $k \leq U$ everything is bounded and it works out.

Now using Abel's summation formula with $\pi(x) = \int_2^x \frac{1}{\log t} dt + R(t)$ in its strongest form and

$$f(t) = \frac{x\omega\left(\frac{\log x}{\log t} - 1\right)}{t \log t},$$

we derive for the sum.

$$f(t)\pi(t)\Big|_{x^{1/u}}^{x^{1/k}} - \int_{x^{1/u}}^{x^{1/k}} f'(t)\pi(t)dt \tag{26}$$

Then for the main term we derive by using integration by parts

$$\int_{x^{1/u}}^{x^{1/k}} \frac{x\omega\left(\frac{\log x}{\log t} - 1\right)}{t\log t}\mathrm{Li}'(t)dt = \int_{x^{1/u}}^{x^{1/k}} \frac{x\omega\left(\frac{\log x}{\log t} - 1\right)}{t(\log t)^2}dt.$$

Substituting $v = \frac{\log x}{\log t}$ and adding the first standard term then leads to

$$\begin{aligned}
&= \frac{k\omega(k)x}{\log x} + \frac{x}{\log x}\int_k^u \omega(v-1)dv. \\
&= \frac{x}{\log x}\left(1 + \int_1^{k-1}\omega(v)dv + \int_{k-1}^{u-1}\omega(v)dv\right) \\
&= \frac{x}{\log x}\left(1 + \int_1^{u-1}\omega(v)dv\right) \\
&= \frac{x\omega(u)u}{\log x}.
\end{aligned}$$

So we indeed get back the inductive step for the main term when subtracting $\frac{y}{\log y}$. Now we still need to check if the error terms we get from $\pi(x)$ by using Abel's summation formula still behave well. Looking at equation 26, but now ignoring the main term and only looking at $R(t)$ we get

$$f(t)R(t)\Big|_{x^{1/u}}^{x^{1/U}} - \int_{x^{1/u}}^{x^{1/U}} f'(t)R(t)dt.$$

Now examining them individually we see that $f(t) = O\left(\frac{x}{t\log t}\right)$ and $f'(t) = O\left(\frac{x}{t^2\log t}\right)$. Thereby $R(t)$ is at least something of the form $O\left(\frac{t}{(\log t)^c}\right)$ for some constant. Hence indeed after combining them they certainly don't blow up over $\frac{x}{(\log x)^2}$ and we are still left witht he same error term.

This completes the inductive step and hence the theorem follows. □

3.4 Buchstab with better main term

In the previous chapter we started with $u < 3$ and in this found how $\phi(x,y)$ was supposed to look like. After which we proved it for all $u < U$. Now we started off with a weak form for the prime number theorem. The question is if we can improve Buchstab's result if instead we use a much stronger result in the first step. Namely

$$\pi(x) = \int_2^x \frac{1}{\log t}dt + R(t).$$

In this paragraph we will ignore the error term and only see what the best main term is supposed to look like. We can do this by following the exact same steps as in standard result. For $x^{1/3} < y \le x^{1/2}$ we have

$$\begin{aligned}
\phi(x,y) &= \phi(x,\sqrt{x}) + \sum_{y \le p < \sqrt{x}} \phi(x/p,p) \\
&\approx \int_0^x \frac{1}{\log t}dt + \sum_{y \le p < \sqrt{x}} \int_0^{x/p} \frac{1}{\log t}dt \\
&= \int_0^x \frac{1}{\log t}dt + \sum_{y \le p < \sqrt{x}} \int_0^{x} \frac{1}{p\log(t/p)}dt \\
&= \int_0^x \frac{1}{\log t}dt + \int_0^{x} \sum_{y \le p < \sqrt{x}} \frac{1}{p\log(t/p)}dt.
\end{aligned}$$

We want to evaluate the sum using Abel's summation formula. Again taking $\pi(s)$ and $f(s) = \frac{1}{s\log(t/s)}$ we have

$$\sum_{y \le p < \sqrt{x}} \frac{1}{p\log(t/p)} = \pi(s)f(s)\Big|_y^{x^{1/2}} - \int_y^{x^{1/2}} \pi(s)f'(s)ds.$$

Now we apply a reverse integration by parts to group all terms into one integral. We derive

$$\approx \int_y^{x^{1/2}} \mathrm{Li}'(s)f(s)ds = \int_y^{x^{1/2}} \frac{1}{\log s}\frac{1}{s\log(t/s)}ds.$$

Now substituting $v = \log s$ we can simplify the integral to get

$$\begin{aligned}
&= \int_{\log y}^{\frac{1}{2}\log x} \frac{1}{v}\frac{1}{\log(t)-v}dv \\
&= \frac{1}{\log t}\int_{\log y}^{\frac{1}{2}\log x} \frac{1}{v} + \frac{1}{\log(t)-v}dv \\
&\approx \frac{\log(\log t - \log y) - \log(\log y)}{\log t} \\
&= \frac{\log(\frac{\log t}{\log y} - 1)}{\log t}.
\end{aligned}$$

Thus we derived an expression for the sum. Substituting this result back in and using that by definition $u\omega(u) = 1 + \log(u-1)$ we derive for the main term.

$$\begin{aligned}
\phi(x,y) &= \int_0^x \frac{1+\log(\frac{\log t}{\log y} - 1)}{\log t}dt) \\
&= \int_0^x \frac{\omega(\frac{\log t}{\log y})}{\log y}du.
\end{aligned}$$

This result clearly checks out in the trivial case when $1 \leq u \leq 2$. Now clearly as we saw for the prime number theorem, when the main term is better we can also have a much stronger error term. How strong this error term can be and that indeed this main term checks out for all u will become apparent in the next paragraph when we look at Tenenbaum's result.

3.5 Strongest known result

The currently strongest result is given by Tenenbaum is his book *Introduction to analytic and probabilistic number theory*, [2]. This result gives the best approximation for $\phi(x, y)$ in every range. It even shows that what we found for very small y is correct. Before we state the theorem and relate it to what we already found we need to define some notation he uses.

Proposition 3.8. *For a specific constant γ,*

$$\lim_{u \to \infty} \omega(u) = e^{-\gamma} \tag{27}$$

This can also be found in [2]. The gamma in the proposition comes from another Mertens theorem which we did not discuss. We can however discuss why the result here makes a lot of sense. Namely u becoming very large is analogous to y being relatively small. Hence the result which we derived for very small y in paragraph 3.2 applies. Then by theorem 3.7 we see that for large u

$$x \prod_{p \leq y} \left(1 - \frac{1}{p}\right) \sim \frac{x\omega(u)}{\log y}.$$

Now Mertens' third theorem, see page 17 of [2], states

$$\prod_{p \leq y} \left(1 - \frac{1}{p}\right) \sim \frac{e^{-\gamma}}{\log y}. \tag{28}$$

Combing the two leads to the desired proposition. It is a very nice result by itself to know what the limit of the Buchstab function itself. Thereby Tenenbaum uses it in his theorem which we will see later.

First he defines a range H_ϵ which is

$$x > x_0(\epsilon), \qquad \exp\left((\log \log x)^{5/3+\epsilon}\right) \leq y \leq x$$

and three functions

$$L_\epsilon(y) := \exp\left((\log y)^{3/5-\epsilon}\right),$$
$$Y_\epsilon(y) := \exp\left((\log y)^{3/2-\epsilon}\right),$$
$$H(u) := \exp\left(\frac{u}{(\log(u+2))^2}\right).$$

The first things that strike us directly are the fact that there is a constant $5/3$ which is similar to what appears in the strongest from of the prime number theorem currently know. Thereby he defines a range which gives a lower bound to y meaning he is probably not going to give the exact same result for all y but rather differentiate according to how large y is.

Combining these three he defines the general error term $E(x, y)$ for some constant $c > 0$ by

$$E(x, y) := H(u)^{-c} L_\epsilon(y)^{-1} + Y_\epsilon(y)^{-1}. \tag{29}$$

For the main term he first defines the integral

$$\mu_y(u) := \int_0^\infty \omega(u - v) y^{-v} dv$$

such that the main term becomes

$$W(x, y) := x\mu_y(u) \frac{e^\gamma \log y}{\zeta(1, y)}. \tag{30}$$

Now finally having defined all notation we can state the theorem.

Theorem 3.9 (Tenenbaum)**.** *For $x \geq y \geq 2$ and an $\epsilon > 0$ we have,*

$$\phi(x, y) - W(x, y) = \begin{cases} O(\psi(x, y)E(x, y)) & (\textit{in the domain } H_\epsilon), \\ O(\psi(x, y)) & (\textit{elsewhere}). \end{cases} \tag{31}$$

We directly see that we have two results according to if y is very small or not. For now let us first examine the main term. We know almost all the terms including

$$\frac{1}{\zeta(1, y)} = \prod_{p \leq y} \left(1 - \frac{1}{p}\right).$$

To understand $\mu_y(u)$ we make the following smart substitution

$$v = \frac{-\log t + \log x}{\log y}$$

which gives

$$\frac{dv}{dt} = -\frac{1}{t \log y}.$$

We can then rewrite the integral to

$$\begin{aligned} \mu_y(u) &= \int_0^x \omega\left(\frac{\log t}{\log y}\right) y^{\frac{\log t - \log x}{\log y}} \frac{dt}{t \log y} \\ &= \int_0^x \frac{\omega\left(\frac{\log t}{\log y}\right)}{\log y} \frac{t}{x} \frac{dt}{t} \\ &= \frac{1}{x} \int_0^x \frac{\omega\left(\frac{\log t}{\log y}\right)}{\log y} dt. \end{aligned}$$

Thus we see that $x\mu_y(u)$ becomes a very well known term. Namely it is exactly what we found in the previous paragraph when trying to construct a better main term for Buchstab. By Mertens second theorem, equation 28, we then see that for sufficiently large y, $W(x, y)$ indeed becomes the main term we found last paragraph. Thus we conclude our idea last paragraph was completely right and Tenenbaum's theorem gives a much stronger main term than the standard Buchstab result.

Now secondly for very small y we have that $\omega\left(\frac{\log t}{\log y}\right)$ approaches $e^{-\gamma}$ very fast hence

$$\mu_y(u) \sim \frac{1}{x \log y} \int_0^x e^{-\gamma} dt \tag{32}$$

$$= \frac{e^{-\gamma}}{\log y}. \tag{33}$$

Hence we see

$$W(x, y) \sim \frac{x}{\zeta(1, y)} = x \prod_{p \leq y} \left(1 - \frac{1}{p}\right).$$

This is exactly what we found in paragraph 3.2 so it appears that Tenenbaum's theorem seems to combine everything. Now for the error terms we notice that in both ranges we have a $\psi(x, y)$ which isn't all that surprising since the smooth numbers are closely related to the rough numbers. Outside of H_ϵ the error term only consists of $\psi(x, y)$ but this is certainly not a problem. Namely in this domain y is small hence there are relatively very few smooth numbers given is a very strong error term.

In the other case when we are in the domain H_ϵ the error term coming from $\psi(x, y)$ gets a lot worse as the amount of smooth numbers increasing. Luckily we have an extra term $E(x, y)$ which gives a lot of extra saving. Here the $H(u)$ term gives a lot of extra saving when u becomes large and the other two are more important when y itself increases.

4 Stick-breaking

4.1 The stick-breaking process

In this chapter we we will discuss the so-called stick-breaking process and define a probability measure on it which we will show to be equal to the Buchstab function. This has already been done for the Dickman function by Brady in chapter 4.3 of [5]. We will first follow and discuss the methodology in this paper with respect to stick-breaking and how the probability is defined for the Dickman function. After which we will do something similar for the Buchstab function.

Firstly a stick of length 1 is chosen which we cut at a random place with uniform probability. We throw away the part to the left of the cut and cut what is left of the stick again with uniform probability at a random point. Throwing away the part to the left of the cut again we continue this process an infinite number of times giving us an infinite number of pieces adding up to 1. We define the process in a more formal way as is done in the cited paper.

Definition 4.1. We define an infinite number of random variables $(x_1, x_2, ..)$ as a *stick-breaking process* when for every n, x_n is uniformly distributed on the interval $1 - \sum_{k=1}^{n-1} x_k$ such that $\sum_{i=1}^{\infty} x_i = 1$.

The defined stick-breaking adheres a nice rearrangement property.

Proposition 4.2. *Let $(x_1, x_2, ...)$ be a stick-breaking process. Then naturally we get a sequence of intervals of length x_i, by $I_i = [x_{i-1}, x_{i-1} + x_i)$. Now let m be a uniformly chosen random point on the interval $[0, 1)$. Then the rearrangement principle says that the length of the interval containing m is uniformly distributed between 0 and 1.*

Proof. We start by defining a probability

$$f(u) := \mathbb{P}(I_n \text{ which contains } m \text{ has length at most } u).$$

To calculate this function we use a trick which is often used when doing calculations with stick-breaking since it works in a very nice way. We split $f(u)$ into two different cases. Namely m being in the first interval I_1 or not. Now this first probability is very easy to calculate since the likely-hood of m being in I_1 is simply x_1. Hence we can write,

$$f(u) = \int_0^u x_1 dx_1 + \int_0^1 (1 - x_1) f\left(\frac{u}{1 - x_1}\right) dx_1.$$

Here the second integral is clear since the probability of it being in the other intervals is $(1 - x_1)$ and we just get the back f with a scaled u because of the absence of I_1.

Now notice that when $\frac{u}{1-x_1} \geq 1$, $f\left(\frac{u}{1-x_1}\right) = 1$. Hence we can manipulate the integral in the following way.

$$\begin{aligned}
f(u) &= \frac{u^2}{2} + \int_{1-u}^1 (1 - x_1) dx_1 + \int_0^{1-u} (1 - x_1) f\left(\frac{u}{1 - x_1}\right) dx_1, \\
&= u^2 + \int_0^{1-u} (1 - x_1) f\left(\frac{u}{1 - x_1}\right) dx_1.
\end{aligned}$$

After making the substitution $v = \frac{u}{1-x_1}$ we then derive,

$$f(u) = u^2 + \int_u^1 \frac{u^2}{v^3} f(v)dv.$$

Now clearly this equality is solved by $f(u) = u$ in the interval $[0, 1]$. It also the unique solution which can be seen by evaluating the following integral.

$$\begin{aligned}
\int_0^1 |f(u) - u|du &= \int_0^1 |u^2 + \int_u^1 \frac{u^2}{v^3} f(v)dv - u|du \\
&= \int_0^1 | \int_u^1 \frac{u^2}{v^3}(f(v) - v)dv|du \\
&\leq \int_0^1 \int_u^1 \frac{u^2}{v^3}|f(v) - v|dvdu \\
&= \int_0^1 \int_0^v \frac{u^2}{v^3}|f(v) - v|dudv \\
&= \frac{1}{3}\int_0^1 |f(v) - v|dv.
\end{aligned}$$

We derive the integral must be zero meaning that our solution is unique. □

Now we can use this stick-breaking process to define the Dickman function used in counting smooth numbers.

4.2 The Dickman function

Definition 4.3. We define the function $\rho(u)$ as the probability that all the pieces of the stick-breaking process, I_n, are smaller than $\frac{1}{u}$.

Proposition 4.4. *$\rho(u)$ satisfies the following identities*

$$\rho(u) = \frac{1}{u}\int_{u-1}^u p(v)dv$$

and

$$u\rho'(u) = -\rho(u - 1).$$

Hence $\rho(u)$ is the Dickman function.

Proof. We use a fairly similar trick to before. First let $(x_1, x_2, ...)$ be a stick-breaking process. Now clearly $\frac{1}{1-x_1}(x_2, x_3, ...)$ is also stick-breaking process. Since we still want our x_i to not exceed $\frac{1}{u}$ our new variable must be less than $\frac{1}{(1-x_1)u}$. Using this trick we can write

$$\rho(u) = \int_0^{\frac{1}{u}} \rho((1 - x_1)u)dx_1.$$

If we now substitute $v = (1 - x_1)u$ with $dx_1 = -\frac{1}{u}dv$ we derive,

$$\rho(u) = \frac{1}{u} \int_{u-1}^{u} \rho(v) dv.$$

This is the first identity. The second one follows by multiplying u to the other side and afterwards taking the derivative on both sides.

$$\rho(u) + u\rho'(u) = \rho(u) - \rho(u-1).$$

Since $\rho(u) = 1$ for $u < 1$ is trivially satisfied we conclude that we derived the delayed differential equation which has as its unique solution the Dickman function. Hence we see that indeed $\rho(u)$ is the Dickman function. □

4.3 A naive guess

We have found the Dickman function in a fairly simple way. The question is why this works. While we cannot give an ultimate answer we do have a good intuitive idea why it works. We could see our stick as a number n of which we are cutting primes. In the stick-breaking all pieces have to add up 1 while all primes ultimately multiply to n. So asking the probability that all pieces of the stick are ultimately smaller than some constant is similar to asking all primes to be smaller than a certain number. This makes the Dickman function into a probability function which gives the probability that a number n is smooth.

Now since the Buchstab function and the Dickman function have a lot of resemblance the next logical step is asking if we can maybe define a similar stick-breaking probability which gives us the Buchstab function. Now we have seen that stick-breaking works for the Dickman function so lets start in an almost identical fashion.

We define exactly the same stick-breaking process with each interval corresponding to a x_i. If the Dickman function appears from each piece being smaller than $\frac{1}{u}$, maybe Buchstab appears from the exact opposite. So starting with the most basic possible guess we define

$$\omega'(u) := \mathbb{P}(\text{For all } i : x_i > \frac{1}{u}).$$

Continuing the process the same as before we can eliminate the first coordinate x_1 such that

$$\frac{1}{1 - x_1}(x_2, x_3, ...)$$

is a new stick-breaking process. To uphold the same conditions we scale the variable u. We derive,

$$\omega'(u) = \int_{\frac{1}{u}}^{1} \omega((1 - x_1)u) dx_1$$

which after the variable change $v = (1 - x_1)u$ leads us to,

$$\omega'(u) = \frac{1}{u} \int_{0}^{u-1} \omega(v) dv.$$

To calculate this integral we see that by definition $\omega'(u) = 0$ for $0 \leq u \leq 1$. Hence after evaluating the integral we see that $\omega'(u)$ is also equal to 0 for $0 \leq u \leq 2$. Continuing the process inductively results in

$$\omega'(u) = 0, \qquad \text{for all } u \geq 0.$$

Hence we defined a completely meaningless probability. It may clear that we could have seen this coming. Namely we cannot make infinite cuts and expect them all to be larger than $\frac{1}{u}$ while still adding up to 1. So $\omega'(u)$ being equal to 0 was clear from the start.

Sadly $\omega'(u)$ is not the Buchstab function but it is not like our guess was completely useless. As can be seen we do in fact get the delayed differential equation of the Buchstab function. The problem is that our starting values make very little sense. Thus we want to define something similar which fixes the starting values in the correct way.

Firstly we don't want our integral to start at 0 but at 1 so we need some different bound. Secondly it is clear the we cannot define a probability over an infinite number of pieces if they must have a minimum length. Thus we see that we are looking at a number of pieces bounded by $\lfloor u \rfloor$ if their minimum length is $1/u$. In the end these heuristics lead to the following result.

4.4 Stick-breaking for the Buchstab function

We now define a stick-breaking which leads to the Buchstab function in a similar way to the Dickman function.

Definition 4.5. Let $(x_1, x_2, ...)$ be a stick-breaking process and define for $u > 1$ a stick-breaking probability by

$$\omega_k(u) := \mathbb{P}(\text{For } 1 \leq i \leq k-1 : \ \frac{1}{u} < x_i < 1 - \frac{1}{u} - \sum_{j<i} x_j \text{ and } x_k > 1 - \frac{1}{u} - \sum_{j<k} x_j). \tag{34}$$

We set $\omega_k(u) = 0$ if $u < 1$.

When we define our probability like this we get back the Buchstab function in the following way.

Theorem 4.6. *If we define*

$$\omega(u) := \sum_{k \leq u} \omega_k(u), \tag{35}$$

then $\omega(u)$ satisfies

$$u\omega(u) = 1 + \int_1^{u-1} \omega(v)dv.$$

Therefore $\omega(u)$ is the Buchstab function.

Proof. To see this first notice that ω_1 is very easily derived. Namely

$$\begin{aligned}\omega_1(u) &= \mathbb{P}(x_1 > 1 - \frac{1}{u}) \\ &= \int_{1-\frac{1}{u}}^{1} 1 dx_1 \\ &= \frac{1}{u}.\end{aligned}$$

For $k \geq 2$ it becomes somewhat trickier. We would like to get some sort of integral similar to our failed try. So using the same stick as for the Dickman function we have,

$$(x_1', x_2', ...) = \frac{1}{1 - x_1}(x_2, x_3, ...)$$

is also a stick-breaking process. Now we still want it to satisfy the same conditions. Meaning that if for example $x_2 > 1/u$, then we get the inequality

$$(1 - x_1)x_2' > \frac{1}{u} \qquad \text{or} \qquad x_2' > \frac{1}{(1 - x_1)u}$$

This leads us to a similar guess as before, which is that we must evaluate $\omega_{k-1}((1 - x_1)u)$ for our "new" stick-breaking process. Here the $(k - 1)$ is clear since the new stick-breaking process has one less "cut". Now to check this is indeed what we want, we multiply the $1 - x_1$ out of the fraction. We derive

$$\begin{aligned}\omega_{k-1}((1 - x_1)u) = \mathbb{P}(\text{For } 1 \leq i \leq k - 2 :\ & \frac{1}{u} < (1 - x_1)x_i' < (1 - x_1) - \frac{1}{u} - (1 - x_1)\sum_{j<i} x_j' \\ & \text{and } (1 - x_1)x_{k-1}' > (1 - x_1) - \frac{1}{u} - (1 - x_1)\sum_{j<k-1} x_j').\end{aligned}$$

Now notice that $(1 - x_1)x_i' = x_{i+1}$. Using this we see that

$$\begin{aligned}\omega_{k-1}((1 - x_1)u) = \mathbb{P}(\text{For } 2 \leq i \leq k - 1 :\ & \frac{1}{u} < x_i < (1 - x_1) - \frac{1}{u} - \sum_{j<i} x_j \\ & \text{and } x_k > (1 - x_1) - \frac{1}{u} - \sum_{2\leq j<k} x_j).\end{aligned}$$

But this probability is exactly the last $(k - 1)$ conditions of $\omega_k(u)$. Since this first condition is simple we derive,

$$\omega_k(u) = \int_{\frac{1}{u}}^{1-\frac{1}{u}} \omega_{k-1}((1 - x_1)u)dx_1.$$

Substituting $v = (1 - x_1)u$ and with that $dx_1 = -\frac{dv}{u}$ we deduce

$$\omega_k(u) = \frac{1}{u}\int_1^{u-1} \omega_{k-1}(v)dv. \tag{36}$$

From this the result follows quite directly. Namely

$$\begin{aligned}\omega(u) &= \frac{1}{u} + \sum_{2\leq k\leq u} \omega_k(u) \\ &= \frac{1}{u} + \sum_{2\leq k\leq u} \frac{1}{u}\int_1^{u-1} \omega_{k-1}(v)dv \\ &= \frac{1}{u} + \frac{1}{u}\int_1^{u-1} \sum_{k\leq u-1} \omega_k(v)dv\end{aligned}$$

Now since $v \leq u-1$ we can just replace the sum by $\omega(v)$ and we are left with the Buchstab function,

$$\omega(u) = \frac{1}{u} + \frac{1}{u}\int_1^{u-1} \omega(v)dv.$$

□

So we found a very nice way to write the Buchstab function $\omega(u)$ as a sum. These ω_k can all be expressed in terms of ω_{k-1} as seen in equation 36. Now we can even improve this expression.

Corollary 4.7.

$$\omega_k(u) = \frac{1}{u}\int_{k-1}^{u-1} \omega_{k-1}(v)dv. \tag{37}$$

Proof. If we start by looking at $\omega_2(u)$ using equation 36 then we see it clearly vanishes for all $u \leq 2$. This results in $\omega_3(u)$ vanishing for for $u \leq 3$ because we are then integrating over the zero function. As this process continues inductively we that our result follows and we can give an even better expression of the ω_k

Now if we look directly at the definition of the ω_k via our stick-breaking this result does not come as a surprise. Namely in our stick-breaking probability we want every x_i to cut off a piece of the stick that is larger than $\frac{1}{u}$. Hence in this case the first $k-2$ pieces combine to a length greater than $\frac{k-2}{u}$. This means that for our second to last condition, which concerns x_{k-1}, we demand

$$\frac{1}{u} < x_{k-1} < 1 - \frac{1}{u} - \sum_{j<k-1} x_j < 1 - \frac{1}{u} - \frac{k-2}{u} = 1 - \frac{k-1}{u}.$$

Hence we see that setting $k = u$ would give us the following condition.

$$\frac{1}{u} < x_{k-1} < \frac{1}{u}.$$

The probability of this happening is zero hence as we derived earlier

$$\omega_k(u) = 0 \quad \text{for } u \leq k.$$

□

The Buchstab function is generally defined by its delayed differential equation. It would be interesting to see if we could write it in a more direct way using our new expression. We can for example directly calculate $\omega_2(u)$ for example. This means for $u < 3$ we have

$$\begin{aligned}\omega(u) &= \frac{1}{u} + \frac{1}{u}\int_1^{u-1} \frac{1}{v}dv \\ &= \frac{1}{u} + \frac{\log(u-1)}{u}.\end{aligned}$$

For $u < 4$ we could do something similar by directly integrating the $\frac{\log(u-1)}{u}$ and adding it. As the process continues inductively we derive for $\omega_k(u)$.

$$\omega_k(u) = \frac{1}{u}\int_{k-1}^{u-1} \frac{1}{x_{k-1}} \int_{k-2}^{x_{k-1}-1} \frac{1}{x_{k-2}} ... \frac{1}{x_2} \int_{1}^{x_2-1} \frac{1}{x_1} dx_1 ... dx_{k-2} dx_{k-1}.$$

This leads us to another expression for the Buchstab function.

Corollary 4.8.

$$\omega(u) = \frac{1}{u} + \frac{1}{u} \sum_{2\leq k\leq u} \int_{k-1}^{u-1} \int_{k-2}^{x_{k-1}-1} ... \int_{1}^{x_2-1} \prod_{i=1}^{k-1} \frac{dx_i}{x_i} \tag{38}$$

This result is very nice since we can now write the Buchstab function now just as a delayed differential equation. There are some other ways to write Buchstab function directly and not as a delayed differential equation. All of them are however quite complex and and much harder to intuitively understand than our expression.

The expression we found is also not completely new. It was discovered two years ago already in a paper by Ross Pinsky, [4]. Equation 1.25 on page 15 there gives the same expression we just found. He derives it however in a more advanced probabilistic manner which we will not go in to. Nevertheless the result is exactly the same and gives us more confidence that indeed what we found is correct. Like in the paper our new expression of Buchstab automatically gives us a new way of expressing the constant $e^{-\gamma}$.

$$e^{-\gamma} = \lim_{n\to\infty} \frac{1}{n} \sum_{2\leq k\leq n} \int_{k-1}^{n-1} \int_{k-2}^{x_{k-1}-1} ... \int_{1}^{x_2-1} \prod_{i=1}^{k-1} \frac{dx_i}{x_i} \tag{39}$$

In his paper he isn't interested in the individual terms which make up the Buchstab function. We however will try to analyse what these ω_k mean and why we can split up the Buchstab function in such a way that it still all makes sense.

4.5 Interpretation of the ω_k

Writing the Buchstab function as a sum over different functions which are all related is a nice result by itself. The question is if these $\omega_k(u)$ resemble anything themselves. It could be that they are just a handful of terms which nicely combine to make the Buchstab function but are meaningless otherwise.

Ignoring error terms we start with what we know for rough numbers in its simplest form.

$$\phi(x, y) = \frac{x\omega(u)}{\log y} = \frac{x\omega(u)u}{\log x}.$$

We know that $\omega_1(u) = \frac{1}{u}$ and hence

$$\frac{x\omega_1(u)u}{\log x} = \frac{x}{\log x}.$$

This is the approximation for the amount of prime numbers. While this could mean a number of things it seems like the ω_k could maybe be related to the amount of prime factors. This idea gets even more plausible when we remember how $\omega(u)$ was defined. Namely a sum over all integers k less than u. Since by definition $y = x^{1/u}$ we know that any $n \in \phi(x, y)$ has at most $\lfloor u \rfloor$ prime divisors. Certainly any n having more prime divisors than u which are larger than y would be a number exceeding x since $y^u = (x^{1/u})^u$.

now we can make a pretty good guess of what the $\omega_k(u)$ resemble. We have that ω_1 counts rough numbers with one prime divisor (primes). Thereby the sum over all ω_k stops at the $k = \lfloor u \rfloor$ which coincidentally is also the maximum amount of prime divisor a rough number n can have. This leads us too the following educated guess. Namely that every $\omega_k(u)$ counts the rough numbers with exactly k prime divisors.

Definition 4.9. We define

$$\phi_k(x, y) = \#\{n \leq x : \ p|n \implies p \geq y, \ n \text{ has exactly } k \text{ prime divisors}\}.$$

Such that

$$\phi(x, y) = \sum_{k=1}^{\lfloor u \rfloor} \phi_k(x, y). \tag{40}$$

We split up the rough numbers by grouping them according to how many prime divisors they have. In this way we get the same number of terms, $\lfloor u \rfloor$, as we have in our new definition of the Buchstab function. That we indeed have the expected relation between them follows from the following theorem.

Theorem 4.10. *Let $x \geq y \geq 2$ and $k \in \mathbb{N}$. Then uniformly for $1 \leq u \leq U$*

$$\phi_k(x, y) = \frac{x\omega_k(u)u}{\log x} + O_{U,k}\left(\frac{x}{(\log x)^2}\right) \tag{41}$$

Proof. We will prove by induction over u. Assume that for $y \geq x^{1/a}$ the formula holds. We want to show that it also does for

$$x^{\frac{1}{a+1}} \leq y < x^{1/a}.$$

Firstly this is trivial for $a = 1$ hence we continue by proving the induction step. By Buchstab's identity

$$\phi_k(x, y) = \phi_k(x, x^{1/a}) + \sum_{y \leq p < x^{1/a}} \phi_{k-1}(\frac{x}{p}, p) \tag{42}$$

$$= \frac{x\omega_k(a)a}{\log x} + O_{a,k}\left(\frac{x}{(\log x)^2}\right) + \sum_{y \leq p < x^{1/a}} \frac{x\omega_{k-1}\left(\frac{\log x}{\log p} - 1\right)}{p \log p} + O\left(\sum_{y \leq p < x^{1/a}} \frac{x/p}{(\log x)^2}\right) \tag{43}$$

$$= \frac{x\omega_k(a)a}{\log x} + \sum_{y \leq p < x^{1/a}} \frac{x\omega_{k-1}\left(\frac{\log x}{\log p} - 1\right)}{p \log p} + O_{a,k}\left(\frac{x}{(\log x)^2}\right). \tag{44}$$

Here we also used our induction hypothesis within the sum. The fact that this is allowed follows directly from

$$p^a > y^a > x^{\frac{a}{a+1}} > \frac{x}{p}.$$

In the error term we get some other constant depending on U. This follows in identically the same way as in the proof the Buchstab function in the previous chapter.

Now moving on we can extract $\frac{x}{\log x}$ from equation 44 such that

$$\phi_k(x,y) = \frac{x}{\log x}\left(\omega_k(a)a + \log x \sum_{y<p\leq x^{1/a}} \frac{\omega_{k-1}\left(\frac{\log x}{\log p} - 1\right)}{p\log p}\right) + O_{a,k}\left(\frac{x}{(\log x)^2}\right).$$

Now we can continue like usual and evaluate the sum using Abel's summation formula. We use as our sum $\pi(t)$ in it strongest form. Thereby we take

$$f(t) = \frac{\omega_{k-1}\left(\frac{\log x}{\log t} - 1\right)}{t\log t}.$$

This is clearly a continuously differential function in the defined domain. We derive

$$\sum_{y<p\leq x^{1/a}} \frac{\omega_{k-1}\left(\frac{\log x}{\log p} - 1\right)}{p\log p} = \pi(t)f(t)\Big|_y^{x^{1/a}} - \int_y^{x^{1/a}} \pi(t)f'(t)dt$$

Well clearly the way $f(t)$ is defined we would ideally not want to take it derivative. Lucky by integration by parts it is clear this becomes

$$\begin{aligned}\sum_{y<p\leq x^{1/a}} \frac{\omega_{k-1}\left(\frac{\log x}{\log p} - 1\right)}{p\log p} &= \int_y^{x^{1/a}} \mathrm{Li}'(t)f(t)dt + R(t),\\ &= \int_y^{x^{1/a}} \frac{\omega_{k-1}\left(\frac{\log x}{\log t} - 1\right)}{t(\log t)^2}dt + R(t).\end{aligned}$$

Here the error term comes from strongest form of the prime number theorem. As we saw in the proof of the Buchstab function this error term is much stronger and therefore absorbed into the term

$$O_{a,k}\left(\frac{x}{(\log x)^2}\right).$$

Now we want to evaluate this integral using a smart substitution.

Setting

$$v = \frac{\log x}{\log t},$$

which gives us

$$\frac{dv}{dt} = -\frac{\log x}{t(\log t)^2},$$

simplifies the integral a large amount. We deduce using the definition of the ω_k

$$\begin{aligned}\phi_k(x,y) &= \frac{x}{\log x}\left(\omega_k(a)a + \int_a^u \omega_{k-1}(v-1)dv\right)\\ &= \frac{x}{\log x}\left(\int_{k-1}^{a-1} \omega_{k-1}(v)dv + \int_{a-1}^{u-1} \omega_{k-1}(v)dv\right)\\ &= \frac{x}{\log x}\int_{k-1}^{u-1} \omega_{k-1}(v)dv\\ &= \frac{x\omega_k(u)u}{\log x}.\end{aligned}$$

We conclude that indeed the number of rough number with exactly k prime divisors is counted by each ω_k. $\square$

Now it is interesting to examine the ω_k in the limit. We know that

$$\lim_{u\to\infty} \omega(u) = e^{-\gamma}.$$

Since $\omega(u) = \sum_{k=1}^{\lfloor u \rfloor} \omega_k(u)$, maybe ω_k also approaches some sort of constant C_k. These would then all add up to $e^{-\gamma}$ when taking the sum over all integers k.

A short look at ω_k makes it clear this is not the case and these C_k do not exist.

Proposition 4.11. *For all $k \in \mathbb{N}$*

$$\lim_{u\to\infty} \omega_k(u) = 0. \tag{45}$$

Proof. The proof is very straightforward and can be seen in two different ways. First of all by looking directly at our stick-breaking definition of $\omega_k(u)$, equation 34. Namely the last condition is

$$x_k > 1 - \frac{1}{u} - \sum_{j<k} x_j.$$

This means that in the limit we would want

$$x_k > 1 - \sum_{j<k} x_j.$$

which is impossible by definition. Namely in our stick-breaking process x_k is bounded above by $1 - \sum_{j<k} x_j$ which is the size of the stick left for x_k.

This could also be seen when looking at our integral definition of $\omega_k(u)$.

Lets assume

$$\lim_{u\to\infty} \omega_{k-1}(u) = 0.$$

Then for all $\epsilon > 0$ there exists a M such that for $v > M$

$$\omega_{k-1}(v) < \epsilon.$$

Using this we can write

$$
\begin{aligned}
\omega_k(u) &= \frac{1}{u}\int_{k-1}^{u-1} \omega_{k-1}(v)dv \\
&= \frac{1}{u}\int_{k-1}^{M} \omega_{k-1}(v)dv + \frac{1}{u}\int_{M}^{u-1} \omega_{k-1}(v)dv \\
&< \frac{1}{u}\int_{k-1}^{M} 1dv + \frac{1}{u}\int_{M}^{u-1} \epsilon dv \\
&= \frac{M-k+1}{u} + \frac{\epsilon(u-1-M)}{u}
\end{aligned}
$$

So we can conclude that

$$\lim_{u\to\infty} \omega_k(u) = 0.$$

Now since for $\omega_1(u) = \frac{1}{u}$ the result trivially follows, it follows for all k by induction. □

So although $\omega(u)$ converges to a constant each individual term in the sum goes to 0. This is entirely possible if we remember the sum runs up to $\lfloor u \rfloor$. Hence the $\omega_k(u)$ and the sum go to infinity at the same time. So what happens is that the $\omega_k(u)$ don't approach zero "fast enough" so each ω_k still contributes a very small part to ω.

References

[1] Jörg Brüdern, *Einführung in die analytische Zahlentheorie*, Springer-Lehrbuch, Springer-V erlag Berlin Heidelberg GmbH, Berlin, 1994.

[2] Gérald Tenenbaum, *Introduction to analytic and probabilistic number theory*, Cambridge studies in advanced mathematics 46, Cambridge university press, Cambridge, 1995.

[3] Hugh L. Montgomery, Robert C. Vaughan, *Multiplicative number theory I: classical theory*, Cambridge studies in advanced mathematics 97, Cambridge university press, Cambridge, 2006.

[4] Ross G. Pinsky, *A natural probabilistic model on the integers and its relation to Dickman-type distributions and Buchstabs function*, arXiv:1606.02965, 2016.

[5] Zarathustra Elessar Brady, *Sieves and iteration rules*, 2017.

[6] Greene, D. H. and Knuth, D. E. *Mathematics for the Analysis of Algorithms*, 3rd ed. Boston, pp. 95-98, 1990.

[7] Finch, S. R. *RE: Unusual Numbers.* 27 Aug 2001.

[8] Finch, S. R. *Stieltjes Constants.* 2.21 in Mathematical Constants. Cambridge, England: Cambridge University Press, pp. 166-171, 2003.

[9] Dickman, K. *On the frequency of numbers containing prime factors of a certain relative magnitude*, Arkiv fr Matematik, Astronomi och Fysik. 22A (10): 114.

[10] Hadamard, J. *Sur la distribution des zéros de la fonction $\zeta(s)$ et ses conséquences arithmétiques.* Bull. Soc. math. France 24, 199-220, 1896.

[11] de la Valle Poussin, C.-J. *Recherches analytiques la thorie des nombres premiers.* Ann. Soc. scient. Bruxelles 20, 183-256, 1896.

[12] A.A. Buchstab, *Asymptotic estimates of a general number-theoretic function* Mat. Sb. , 44 (1937) pp. 12391246

[13] F. Mertens, *Ein Beitrag zur analytischen Zahlentheorie*, Journal für die reine und angewandte Mathematik, 4662, Math. 78, Berlin, 1874.