On Methods for Computing $\pi(x)$

R.P. Leopold
Supervisor : prof. dr. Jaap Top
Mathematics Department
RuG

Contents

1. Introduction

The prime numbers are very important in Mathematics. Their distribution has been studied by mathematicians at least since the ancient Greeks. *Euclid* already proved that there can not be a finite number of them. The prime counting function $\pi(x)$, which counts the number of primes less than or equal to a given positive real number x, is the basis for the study of the distribution of the primes. In 1796 *A.M.Legendre* conjectured that a good approximation to $\pi(x)$ is

$$\frac{x}{log(x) - B}$$

where $B \approx 1.08$. In the same year *C.F. Gauss* conjectured that a good approximation to $\pi(x)$ would be the so-called logarithmic integral

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

Both conjectured results are implied by

$$\lim_{x \to \infty} \frac{\pi(x) log(x)}{x} = 1. \tag{2}$$

This formula implies that the relative error one makes by approximating $\pi(x)$ with $\frac{x}{log(x)}$ goes to zero for x to infinity. The absolute error of this approximation need not be bounded as function of x and indeed it is not.

Around 1849 the russian mathematician *P.N. Chebyshev* tried to prove (2) but did not quite succeed. He did however proof, among other things, that if the limit in (2) exists it must be equal to one. Also he seems to have been the first who used the Zeta function $\zeta(s)$ with s being a complex number instead of a real number.

In 1859 *B. Riemann* published his groundbreaking paper "Ueber die Anzahl der Primzahlen unter einer gegebenen Grösse", in which he uses deep ideas to link the values of $\pi(x)$ to the nontrivial zero's of the ζ-function for which he had found an analytic continuation to the whole complex plane. Using these and extended ideas *Hadamard* and *de la Vallee Poussin* independently proved the so-called *Prime Number Theorem* (2) in 1896.

It seems as if the prime numbers exhibit order globally or asymptotically. Nevertheless their local irregular distribution would seem to imply that for computing $\pi(x)$ for specific values of x one can only use the so-called elementary methods such as for example the sieve of *Eratosthenes*.

The algorithms implemented and used to calculate values of $\pi(x)$ up to a current record of $\pi(10^{23})$ indeed all use variants of the so-called *Extended Meissel-Lehmer* methods. One such method is the method of *J. Lagarias*, *V. Miller* and *A. Odlyzko* [**3**]. These methods use sieving methods combined with smart elementary techniques of finding the needed values of the so-called *Legendre* sum (5) on page 9. They are all based on work by *E. Meissel* in the nineteenth century and *D. H. Lehmer* in the twentieth century. The methods lend themselves to parallel implementation and today the internet is used to compute new values of $\pi(x)$ using distributed computing networks.

One could ask if there can be any more breakthroughs in the computation of $\pi(x)$ or if new values will only be the result of increased computing power. In 1987

J. Lagarias and *A. Odlyzko* published the paper "Computing $\pi(x)$: An Analytic Method" [**4**] in which they present an algorithm for computing values of $\pi(x)$ which uses parametrised families of Mellin transforms. They base their work on the work done by *B. Riemann*. *Lagarias* and *Odlyzko* show that this algorithm can compute values of $\pi(x)$ to within ± 0.1 using time- and space complexities that are superior to the current elementary methods.
In chapter one we will take a look at the sieve of *Eratosthenes* and *Legendre*'s formula on which further work of *Meissel* and *Lehmer* is based. The methods of *Meissel* and *Lehmer* and *D. Mapes* will be presented.
In chapter two we will take a more detailed look at the *Legendre* sum (5), which plays a crucial role in the elementary methods. We will see how, with efficient calculation of this sum, the methods of *Lagarias*, *Miller* and *Odlyzko* and of *M. Deglise* and *J. Rivat* compute $\pi(x)$.
In chapter three we will present parts of the work from *Riemann*'s paper that will be needed to motivate the analytic algorithm of *Lagarias* and *Odlyzko*. In particular a complex contour integral representation for $J(x)$ (17) on page 34, the so-called prime power counting function, will be deduced.
In chapter four the analytic algorithm will be presented and using *Mathematica* computations and *Maple* plots the difficulty in implementation of this algorithm will be hinted at.
P.s. the $log(x)$ function used in this thesis is the natural logarithm unless specified otherwise e.g. $log_2(x)$.

CHAPTER 1

Elementary Methods

1. The Sieve of *Eratosthenes*

Eratosthenes lived from approximately 275 BC to 195 BC, and was an esteemed Greek who was also director of the famous library of Alexandria. He invented or at least published the method, named after him, for finding all primes less than or equal to some bound x, by what is called sieving. It works as follows, make a list of all numbers from 1 up to and including the bound x and start by elliminating all multiples of 2, followed by all multiples of 3, and continue this for all prime numbers upto and possibly including square root of x. When finished with this process you will have a list consisting of all primes smaller than x and greater than square root of x. The reason you only have to use all primes less than or equal to square root of x is that whenever you have a number less than x which is the product of two or more numbers, at least one of those factor numbers must be less than square root of x for obvious reasons. And so a number less than or equal to x will have in its prime decomposition at least one prime number less than square root of x and therefore a composite number below our bound x will be sieved out by this method.
So the primes less than or equal to $\sqrt{x}$ determine firstly the composites below $\sqrt{x}$, secondly they determine the composites in the interval $[\sqrt{x}, x]$ and thirdly, by exclusion this determines the primes in the interval $[\sqrt{x}, x]$. The method is conceptually simple and straightforward to implement but it is not the most efficient method time wise and it also needs a lot of storage capacity. It is also a method that gives you all the primes and with that the number of primes less than or equal to x but when calculating $\pi(x)$ itself we do not necessarilly need to know all primes themselves. The time complexity of the sieve of *Eratosthenes* is $O(xloglog(x))$ and the space complexity is $O(\sqrt{x})$. The space complexity is such because one can sieve in blocks of minimal length $\sqrt{x}$.

2. *Legendre*'s formula

$Adrien - Marie\ Legendre$ lived from 1752 to 1833 and he was a great mathmetician and he devised the first explicit formula for calculating $\pi(x)$. His formula is as follows,

(3)
$$\pi(x) = \lfloor x \rfloor + \pi(\sqrt{x}) - 1 - \sum_{p_i \leq \sqrt{x}} \lfloor \frac{x}{p_i} \rfloor + \sum_{p_i < p_j \leq \sqrt{x}} \lfloor \frac{x}{p_i p_j} \rfloor - \sum_{p_i < p_j < p_k \leq \sqrt{x}} \lfloor \frac{x}{p_i p_j p_k} \rfloor + \cdots$$

here $\lfloor x \rfloor$ equals the largest positive integer less than x and x may be any positive real number. The idea is that the number of composites plus the number of primes plus 1 (because 1 is not a prime but also not composite) equals the number of positive integers less than or equal to our bound x. We have that $\lfloor x \rfloor$ equals the number of positive integers less then or equal to x. The term 1 is for the count of 1 itself. Obviously $\pi(x)$ is for the counting of the primes. The sum terms are as follows. The first sum consists of a term $\lfloor \frac{x}{p_i} \rfloor$ for each prime below square root of x, this equals the number of multiples of this prime below x. This is the same as the number of integers less than or equal to x that have this prime in their decomposition. To see this, notice that $\lfloor x/p \rfloor = n$ if and only if np is the largest multiple of p less than or equal to x. The multiples of p are, p, $2p$, $3p$ and so forth so in this case the n is the number of multiples of p under x.
The terms $\pi(\sqrt{x})$ and the sums over multiple primes are to compensate for integers that counted at least twice when adding the sum over terms $\lfloor \frac{x}{p_i} \rfloor$. Because in this sum we also count the primes less than or equal to $\sqrt{x}$ themselves as terms p_i times 1, we have to add $\pi(\sqrt{x})$. Also when we have a number less than or equal to x that is divisible by two different primes less than or equal to $\sqrt{x}$, it will be counted twice in our sum over the $\lfloor \frac{x}{p_i} \rfloor$. To compensate for that we add the sum over terms $\lfloor \frac{x}{p_i p_j} \rfloor$. Now we have a similar problem in that if we have a integer in our range that is divisible by three different primes, then in the first sum we counted this number three times, the second sum counts this number $\binom{3}{2} = 3$ times and so to compensate we have the third sum over the positive integers divisible by three different primes etcetera.
Now two important things about *Legendre*'s formula (3) at page 7 ,firstly that it is almost nothing more than a precise mathematical description of what it is we do when we perform the sieve of *Eratosthenes*. In the sense that the sum over the terms $\lfloor \frac{x}{p_i} \rfloor$ is the process of counting, for all primes below square root of x, all composite numbers with those primes in their decompositions. As we already saw this captures all composites below x. Although we only need the values of the terms $\lfloor \frac{x}{n} \rfloor$ with $n \leq x$ a product of different primes less than or equal to $\sqrt{x}$, we do not have to actually locate the primes greater than $\sqrt{x}$ as we did with *Eratosthenes*. By exclusion we then know the primes given x. The other terms on the right hand side of the equation are all to compensate for the fact that in the first sum we do not keep track of composites we've allready come across. Secondly if we look at the general term occuring in the sums, $\lfloor \frac{x}{p_1 p_2 \cdots p_{\pi(\sqrt{x})}} \rfloor$, then we can easily see that by counting if a particular prime appears in the denominator, that formally there are $2^{\pi(\sqrt{x})}$ of these terms, which would be too much for any practical computations. In practice, many of these terms will of course not be present because of the fact that the products in the denominators must be less than or equal to x. Also we can then see that the number of sums possible in Legendre's formula must be less then or equal to $log_2(x)$. It can be shown that the number of nonzero terms is aproximately equal to $\frac{6}{\pi^2}(1 - log(2))x$, and so is linear in x. The time complexity of *Legendre*'s formula is $O(x)$ and its space complexity is $O(\sqrt{x})$ [**5**], so only a slight improvement over the sieve of *Eratosthenes*.

3. *Meissel*'s idea

Meissel (1826 – 1895) was a German astronomer who was the first to come up with an efficient formula for calculating $\pi(x)$. *Meissel* worked for quite some years to compute $\pi(x^9)$ by hand and his value was found to be too small by 56, see *Lagarias* and *Odlyzko* [**4**]. As we saw above the primes less than or equal to $\sqrt{x}$ determine the whole lot in the interval $[1, x]$. *Meissel* started with *Legendre*'s formula (3) on page 7 and numbered the primes below $\sqrt{x}$: $p_1, p_2, \cdots, p_a$. With *Eratosthenes* and *Legendre* we obviously have $a = \pi(\sqrt{x})$. Now *Meissel* lowered the value of a. What happens when we do that? This means that in our interval $[1, \sqrt{x}]$ we have at least one prime number greater than p_a that is not counted in $\pi(p_a)$ and which is not represented in our sums over terms $\lfloor \frac{x}{p_1 \cdots p_a} \rfloor$. Hence using *Legendre*'s formula we are not counting primes in $(p_a, \sqrt{x}]$ nor composites divisible by only these primes. So *Meissel* defined,

$$\phi(x, a) = |\{n \leq x : p|n \Rightarrow p > p_a\}|, \tag{4}$$

which is called the partial sieve function and it counts the positive integers less than or equal to x with no prime factor less than or equal to p_a. And we have,

$$\lfloor x \rfloor = \sum_{i \leq a} \lfloor \frac{x}{p_i} \rfloor - \sum_{i<j \leq a} \lfloor \frac{x}{p_i p_j} \rfloor + \sum_{i<j<k \leq a} \lfloor \frac{x}{p_i p_j p_k} \rfloor \cdots + \phi(x, a).$$

With use of the well known möbius function from number theory, we have,

$$\phi(x, a) = \sum \mu(n) \lfloor \frac{x}{n} \rfloor, \tag{5}$$

where we sum over all n having all prime factors less than or equal to p_a.
Recall that the möbius function is defined as :

$$\mu(n) = \begin{cases} 1 & \text{if n = 1,} \\ 0 & \text{if } \mathrm{p}^2 | \text{n for any prime p,} \\ (-1)^t & \text{if } \mathrm{n} = \mathrm{p_1 p_2 \cdots p_t} \text{ for distinct primes p.} \end{cases}$$

Also we can write,

$$\phi(x, a) = 1 + P_1(x, a) + P_2(x, a) + P_3(x, a) + \cdots$$

because 1 has no prime divisors less than or equal to p_a and where

$$P_k(x, a) = |\{n \leq x : n = \prod_{j=1}^{k} p_{m_j}, m_j > a, 1 \leq j \leq k\}|$$

denotes the k-th partial sieve function, which counts the positive integers with exactly k prime factors all greater than p_a. Now we can even define $P_0(x, a) = 1$ and noticing that

$$P_1(x, a) = \pi(x) - a$$

we can write,

$$\pi(x) = a + \phi(x,a) - P_0(x,a) - P_2(x,a) - P_3(x,a) - \cdots .$$

We have that $P_{k-1}(x,a) \neq 0$ and $P_k = 0$ if and only if $x^{1/k} < p_{a+1} \leq x^{1/(k-1)}$. So of course the magnitude of k directly depends on the magnitude of a. Also it might be natural to choose a such that p_{a+1} is the smallest prime greater than $x^{1/k}$ for a specific k for then we have collected as many terms as possible in the $P_{k-1}(x,a)$ before we would need terms residing in $P_k(x,a)$.
In fact $a = \pi(x^{1/3})$ satisfies this property and this value of a implies that $P_{k\geq 3}(x,a) = 0$.
Meissel noticed that we can write

$$\begin{aligned} P_2(x,a) &= |\{n \leq x : n = p_{a+1}p_j, \quad a+1 \leq j\}| \\ &\quad +|\{n \leq x : n = p_{a+2}p_j, \quad a+2 \leq j\}| \\ &\quad +\cdots \\ &= \pi\left(\frac{x}{p_{a+1}}\right) - a + \pi\left(\frac{x}{p_{a+2}}\right) - (a+1) + \cdots \\ &= \sum_{i=a+1}^{\sqrt{x}} \left(\pi\left(\frac{x}{p_i}\right) - (i-1)\right), \quad p_a < p_i \leq \sqrt{x}. \end{aligned}$$

Using the simple formula for the sum of an arithmetic series, $\sum_{i=1}^{k} i = \frac{1}{2}k(k+1)$, we have,

$$P_2(x,a) = -\frac{(\pi(\sqrt{x})-a)(\pi(\sqrt{x})+a-1)}{2} + \sum_{i=a+1}^{\sqrt{x}} \pi\left(\frac{x}{p_i}\right)$$

and *Meissel*'s formula for $\pi(x)$ becomes

$$\begin{aligned} \pi(x) = \lfloor x \rfloor - \sum_{i=1}^{\pi(x^{1/3})} \lfloor \frac{x}{p_i} \rfloor + \sum_{1\leq i<j\leq \pi(x^{1/3})} \lfloor \frac{x}{p_ip_j} \rfloor - \cdots \\ + \left(\frac{(\pi(\sqrt{x})+a-2)(-\pi(\sqrt{x})-a+1)}{2}\right) \\ - \sum_{\pi(x^{1/3})<i\leq \pi(\sqrt{x})} \pi\left(\frac{x}{p_i}\right). \end{aligned} \tag{6}$$

The time complexity of *Meissel*'s algorithm equals $O(\frac{x}{(log(x))^3})$ and it's space complexity is $O(\frac{\sqrt{x}}{log(x)})$ [**5**].

4. Computation of $\pi(1000)$ with *Meissel*'s formula

We are going to compute $\pi(1000)$ with *Meissel*'s formula (6).
We have $x = 1000$, $a = \pi(1000^{1/3} = 10) = 4$ and $\pi(1000^{1/2} = 31.6) = 11$. The first eleven prime numbers are :
$p_1 = 2$, $p_2 = 3$, $p_3 = 5$, $p_4 = 7$, $p_5 = 11$, $p_6 = 13$, $p_7 = 17$, $p_8 = 19$, $p_9 = 23$, $p_{10} = 29$ and $p_{11} = 31$.
We start with the computation of the *Legendre* sum :

$$\begin{gathered}\phi(1000,4) = \lfloor 1000 \rfloor - \\ (\lfloor \frac{1000}{7} \rfloor + \lfloor \frac{1000}{5} \rfloor + \lfloor \frac{1000}{3} \rfloor + \lfloor \frac{1000}{2} \rfloor) + \\ (\lfloor \frac{1000}{7\cdot 5} \rfloor + \lfloor \frac{1000}{7\cdot 3} \rfloor + \lfloor \frac{1000}{7\cdot 2} \rfloor + \lfloor \frac{1000}{5\cdot 3} \rfloor + \lfloor \frac{1000}{5\cdot 2} \rfloor + \lfloor \frac{1000}{3\cdot 2} \rfloor) - \\ (\lfloor \frac{1000}{7\cdot 5\cdot 3} \rfloor + \lfloor \frac{1000}{7\cdot 5\cdot 2} \rfloor + \lfloor \frac{1000}{7\cdot 3\cdot 2} \rfloor + \lfloor \frac{1000}{5\cdot 3\cdot 2} \rfloor) + \\ \lfloor \frac{1000}{7\cdot 5\cdot 3\cdot 2} \rfloor = \\ 1000-(142+200+333+500)+(28+47+71+66+100+166)-(9+23+14+33)+4 = \\ 1000 - 1175 + 478 - 79 + 4 = 228\end{gathered}$$

Secondly we compute the term with $a = 4$ and $\pi(\sqrt{1000}) = 11$:

$$\frac{(11+4-2)(11-4+1)}{2} = 52$$

Thirdly we compute the sum with terms $\pi(\frac{x}{p_i})$ for $i = 5$ up to and including $i = 11$:

$$\begin{gathered}-(\pi(\frac{1000}{11}) + \pi(\frac{1000}{13}) + \pi(\frac{1000}{17}) + \pi(\frac{1000}{19}) + \pi(\frac{1000}{23}) + \pi(\frac{1000}{29}) + \pi(\frac{1000}{31})) = \\ -(24+21+16+15+14+11+11) = -112\end{gathered}$$

So the total sum equals

$$228 + 52 - 112 = 168$$

5. *Lehmer*'s formula

D.H.Lehmer (1905-1991) was an American number theorist who used his adapted version of *Meissel*'s Method to calculate $\pi(10^{10})$ by computer and his value turned out to be too high by one, see *Lagarias* and *Odlyzko* [**4**]. Lehmer put $a = \pi(x^{1/4})$ so that $P_2(x,a) \neq 0$ and $P_3(x,a) \neq 0$ but $P_{k\geq 4}(x,a) = 0$. We have,

$$\begin{aligned}P_3(x,a) &= |\{n \leq x : n = p_{a+1}p_jp_k, \quad a+1 \leq j \leq k\}| \\ &+|\{n \leq x : n = p_{a+2}p_jp_k, \quad a+2 \leq j \leq k\}| \\ &+\cdots \\ &= P_2\left(\frac{x}{p_{a+1}}, a\right) + P_2\left(\frac{x}{p_{a+2}}, a\right) + \cdots\end{aligned}$$

$$= \sum_{i>a} P_2\left(\frac{x}{p_i}, a\right)$$

so that

$$P_3(x,a) = \sum_{i=a+1}^{\pi(x^{1/3})} \sum_{j=i}^{\pi(x\sqrt{x/p_i})} \left(\pi\left(\frac{x}{p_i p_j}\right) - (j-1)\right).$$

And we get *Lehmer*'s formula for $\pi(x)$:

$$\pi(x) = \lfloor x \rfloor - \sum_{i=1}^{\pi(x^{1/4})} \lfloor \frac{x}{p_i} \rfloor + \sum_{1 \le i < j \le \pi(x^{1/4})} \lfloor \frac{x}{p_i p_j} \rfloor - \cdots + \frac{(\pi(\sqrt{x}) + a - 2)(\pi(\sqrt{x}) - a + 1)}{2}$$

$$- \sum_{\pi(x^{1/4}) < i \le \pi(x^{1/3})} \pi\left(\frac{x}{p_i}\right) - \sum_{i=a+1}^{\pi(x^{1/3})} \sum_{j=i}^{\pi(\sqrt{x/p_i})} \left(\pi\left(\frac{x}{p_i p_j}\right) - (j-1)\right).$$

The time complexity of *Lehmer*'s algorithm is $O(\frac{x}{(log(x))^4)}$ and it's space complexity is $O(\frac{x^{1/3}}{log(x)})$ [**5**].

6. The recursion formula for $\phi(x,a)$

In both *Meissel*'s and *Lehmer*'s adaptations of *Legendre*'s formula (3) the so-called *Legendre* sum (5) on page 9 is still what takes the most work in calculating $\pi(x)$. There are two tricks which help in the calculation of this sum. Remembering the definition of $\phi(x,a)$ (4) on page 9, we see that the following must hold :

$$\phi(x,a) = \phi(x,a-1) - \phi\left(\frac{x}{p_a}, a-1\right).$$

This is because the integers in the interval $[1,x]$ not divisible by any of the p_1, p_2, $\cdots$, p_a are the same as those not divisible by p_1, p_2, $\cdots$, p_{a-1} with the exclusion of those that are not divisible by p_a. Those integers not divisible by p_a are less than or equal to $\frac{x}{p_a}$. Using this formula we can recursively break down the computation of $\phi(x,a)$, where we can picture the computation as a binary tree with root $\phi(x,a)$ and which looks like :

(7)

$$\phi(x,a)$$

$$\phi(x,a-1) \qquad -\phi(\frac{x}{p_a}, a-1)$$

$$\phi(x,a-2) \quad -\phi(\frac{x}{p_{a-1}}, a-2) \quad -\phi(\frac{x}{p_a}, a-2) \quad \phi(\frac{x}{p_a p_{a-1}}, a-2)$$

$$\cdots$$

where $\phi(x, a)$ is always the sum of the leaves of this tree. Every node of this tree is uniquely determined by pairs (n, b) with

$$n = p_{a_1} \cdots p_{a_r} \qquad a \geq a_1 > \cdots > a_r \geq b + 1$$

and the node has value $(-1)^r \phi(\frac{x}{n}, b) = \mu(n)\phi(\frac{x}{n}, b)$.
Here n is the product of different primes all less than or equal to p_a, $\mu(n)$ is the möbius function and it correctly captures the sign of the term in that based on the recursion formula above, we have that the sign of a term equals $(-1)^t$ with t the number of primes in n.

7. The first Truncation rule

Another trick that was used by *Meissel* and *Lehmer* is the following.
It consists of the combination of four facts, the first of which is, that when we take $m_k = p_1 p_2 \cdots p_k$ the product of the first k primes with $k \leq a$, then we have

$$\begin{aligned}
\phi(m_k, k) &= \lfloor m_k \rfloor - \sum \lfloor \frac{m_k}{p_i} \rfloor + \sum \lfloor \frac{m_k}{p_i p_j} \rfloor - \cdots \\
&= m_k - \sum \frac{m_k}{p_i} + \sum \frac{m_k}{p_i p_j} - \cdots \\
&= m_k \left(1 - \frac{1}{p_1}\right)\left(1 - \frac{1}{p_2}\right) \cdots \left(1 - \frac{1}{p_k}\right) \\
&= \prod_{i=1}^{k} (p_i - 1) = \varphi(m_k)
\end{aligned}$$

where $\varphi(n)$ is the well known *Euler* φ-function that equals the number of integers q less than or equal to n which have $gcd(n, q) = 1$.
The second fact is that for $x = sm_k + t$ with $0 \leq t < m_k$, we have,

$$\phi(sm_k + t, k) = s\varphi(m_k) + \phi(t, k).$$

To see this we need to see firstly that the function ϕ can be calculated by calculating its value on the disjoint intervals, $[1, m_k)$, $[m_k, 2m_k)$, $\cdots$, $[sm_k, t]$. Secondly we have then to only show that on the intervals of lenght m_k, ϕ takes the same values, or put another way, the function $\phi(x, k)$ is periodic with period m_k. To show this, we first remember that if a number p divides two numbers n and m, it divides their sum $n + m$ and their difference $m - n$, and next we look at a random number n_1 in one of the intervals of length m_k above. If it is not sieved out by the first k primes then we know that $n + m_k$ will not have been sieved out in the next interval. And reversely, if we have a number n_2 in this next interval that's not sieved out then $n_2 - m_k$ is not sieved out in the previous interval. So there is a bijection between the numbers not sieved out in the different intervals.
The third fact is that, when $t > m_k/2$,

$$\phi(t, k) = \varphi(m_k) - \phi(m_k - t, k)$$

this is because the numbers not sieved out in an interval of the form $[cm_k,(c+1)m_k]$ with $c+1 \leq s$ will lie symmetrically around the center $cm_k + \frac{m_k}{2}$ of the interval. Because if a number t in this interval and smaller than this center is not sieved out, which is equivalent with the fact that it is not divisible by any of the primes $p_i < p_k$, then $(c+1)m_k - t$ will also not be divisible by those primes.
The fourth fact is that,

$$\phi(m_k, k) = \phi(m_k - 1, k)$$

simply for the fact that m_k is divisible by the primes less than or equal to p_k and so m_k does not add to the value of $\phi(m_k, k)$.
In practice *Meissel* and *Lehmer* used $a = 4$ corresponding to $m_k = 2\cdot3\cdot5\cdot7 = 210$ to truncate the growth of the tree.

8. *Mapes*' method

D. Mapes noticed that we can use the correlation between the absence or presence of a prime in the denominators of the terms $\frac{x}{p_1\cdots p_a}$ with the digits 0 and 1 respectively. This gives an imediate ordering of the terms in the sense that the binary string that captures the presence of the respective primes in a specific term, will equal the binary representation of our decimal number, which our term will have in this new ordering. So put

$$T_k(x,a) = (-1)^{\beta_0\beta_1\cdots\beta_{a-1}}\lfloor\frac{x}{p_1^{\beta_0}p_2^{\beta_1}\cdots p_a^{\beta_{a-1}}}\rfloor$$

with,

$$k = \beta_{a-1}2^{a-1} + \beta_{a-2}2^{a-2}\cdots + \beta_1 2 + \beta_0 < 2^a$$

and $\beta_i \in \{0,1\}$. So the binary digit indicating the presence of the highest prime is the digit of the highest power of two. We now write,

$$\phi(x,a) = \sum_{k=0}^{2^a-1} T_k(x,a). \tag{8}$$

By defining the following subsum of this sum,

$$\gamma(M,x,a) = \sum_{M}^{M+2^i-1} T_k(x,a)$$

with M having a binary representation such that 2^i is the highest power of two dividing it or equivalently having zero's as it's last i binary digits. Then it follows that

$$\phi(x,a) = T_0(x,a) + \gamma(1,x,a) + \gamma(2,x,a) + \cdots + \gamma(2^{a-1},x,a). \tag{9}$$

The idea now is to notice that if we have a term $T_k(x,a)$ then we can write $k = k' + k''$ with in k'' the first digits of $k's$ binary representation, and in k'

the higher digits of k. These two $k's$ have respective terms $T_{k'}(x,a)$ and $T_{k''}(x,a)$ and we can substitute the k'-term for x in the k''-term, and get,

$$T_{k+k'}(x,a) = T_{k'}(T_k(x,a),a).$$

Or more precicely, if we have k as above, and

$$k' = \beta_{a-1}2^{a-1} + \beta_{a-2}2^{a-2} + \cdots \beta_i 2^{i-1}$$

and

$$k'' = \beta_{i-1}2^{i-2} + \cdots + \beta_0.$$

So that

$$sign(T_{k'}(x,a)) = (-1)^{\beta_i+\beta_{i+1}+\cdots\beta_{a-1}}.$$

Because k' has $\beta_j = 0$ if $j < i$, and therefore also,

$$|T_{k'}(x,a)| = \lfloor \frac{x}{p_{i+1}^{\beta_i} p_{i+2}^{\beta_{i+1}} \cdots p_a^{\beta_{a-1}}} \rfloor.$$

We define $T_k(-x,a)$ to be $-T_k(x,a)$, because by definition plugging in $-x$ makes no sense but in light of our atempt to plug in a term for x in another term, we want to not be bothered by the signs of the terms. Now substituting we get :

$$T_{k''}(T_{k'}(x,a),a) = sign(T_{k''}(x,a))(-1)^{\beta_0+\beta_1+\cdots\beta_{i-1}} \lfloor \frac{|T_{k'}(x,a)|}{p_1^{\beta_0} p_2^{\beta_1} \cdots p_i^{\beta_{i-1}}} \rfloor$$

Now because $\lfloor \frac{\lfloor \frac{a}{b} \rfloor}{c} \rfloor = \lfloor \frac{a}{bc} \rfloor$, we get after working out the term in brackets, that we can multiply the denominators and because the sets of primes in the respective terms are disjoint we indeed get $T_k(x,a) = T_{k''}(T_{k'}(x,a),a)$.
Now we can even plug in i instead of a here. To see that this is true we have to understand that the a present as variable in $T_k(x,a)$ can be lowered to value i without changing the value of the term as long as k has no more than $i+1$ digits in its binary representation. Because that means that only primes less than or equal to the i-th prime occur in the denominator of the term which means that $T_k(x,a) = T_k(x,i)$. So finally we get,

$$T_{k''}(T_{k'}(x,a),i) = T_k(x,a)$$

Now taking again M as a specific value of k with 2^i the highest power of two occuring in the binary representation of k we can write,

$$\phi(T_M(x,a),i) = \sum_{k=0}^{2^i-1} T_k(T_M(x,a),i) = \sum_{k=0}^{2^i-1} T_{k+M}(x,a) = \sum_{k=M}^{M+2^i-1} T_k(x,a) = \gamma(M,x,a)$$

plugging this in equation (9) on page 14 above,

$$\phi(x,a) = T_0(x,a) + \phi(T_1(x,a),0) + \phi(T_2(x,a),1) + \cdots + \phi(T_{2^{a-1}}(x,a),a-1)$$

substituting in this equation $T_M(x,a)$ for x and i for a gives,

$$\begin{aligned}\phi(T_M(x,a),i) &= T_0(T_M(x,a),i) + \phi(T_1(T_M(x,a),i),0) + \cdots \\ &+\phi(T_2(T_M(x,a),i),1) + \cdots + \phi(T_{2^{i-1}}(T_M(x,a),i),i-1) \\ &= T_M(x,a) + \phi(T_{M+1}(x,a),0) + \phi(T_{M+2}(x,a),i),1) + \cdots \\ &+\phi(T_{M+2^r}(x,a),i),r) + \cdots \phi(T_{M+2^{i-1}}(x,a),i),i-1)\end{aligned}$$

Using the facts that $\phi(T_k(x,a),0) = T_k(x,a)$ by equation (8) on page 14 and $\phi(x,a) = \phi(T_0(x,a),a)$ we can using recursion calculate $\phi(x,a)$ with Mapes's method.
The time complexity for $Mapes'$ Method is the same as its space complexity which equals $O(x^{0.7})$ [**5**].

CHAPTER 2

Extended *Meissel*-*Lehmer* Methods

1. The Structure of the Tree

In preparation for the following methods for calculating $\pi(x)$, we take a closer look at the structure of the binary tree (7) on page 12 above. Notice that if b is held fixed then all nodes have different values of n. And vice versa for every occuring value of n we have different values of b. From this we can see that the leaves, for which $b = 0$, have all different values of n. We already knew this because the leaves correspond to the terms in the *Legendre* sum, for we have $\phi(\frac{x}{n}, 0) = \lfloor \frac{x}{n} \rfloor$, and the *Legendre* sum has terms $\lfloor \frac{x}{n} \rfloor$, with $n \leq x$ ranging over the square free products of primes less than or equal to p_a.
If we look at a random node which is not a leaf, two things can happen :
1) the node has both a left and right descendant ,
2) the node only has a left descendant .
Recall that a general node has the value $\mu(n)\phi(\frac{x}{n}, b)$. Notice that every node which is not a leaf has a left descendant. This is because not having a left descendant is equivalent to $\phi(\frac{x}{n}, b-1) = 0$ which by definition means there are no integers less than or equal to $\frac{x}{n}$ that are not divisible by any of the first $b-1$ primes. But then certainly we already would have $\phi(\frac{x}{n}, b) = 0$. Not having a right descendant means $\phi(\frac{x}{np_b}, b-1) = 0$ which is equivalent to $np_b > x$. But $np_b > x$ implies that $\phi(\frac{x}{n}, b) = 1$ because if it were greater than one then there would have to be an integer less than or equal to $\frac{x}{n}$ that is divisible by p_{b+1}. This is equivalent to $np_{b+1} \leq x$ but we have $p_{b+1} > p_b$ so this cannot happen. So we now have that if a node is not a leaf then it has a left descendant and if it has no right descendant we have $\phi(\frac{x}{n}, b) = 1 = \phi(\frac{x}{n}, b-1)$. The last equality is because of $np_b > x$. The implication of all this is that when the recursion comes to a node with no right descendant it will continue the calculation of the left branch which equals the set of descendants of the left descendant and the left descendant itself of our node. This is a waste of time in that we already know that the ϕ-value equals one.
With *Meissel*, *Lehmer* and *Mapes* we stopped the recursion tree when $a = 5$ which we know was because of the presieving done by the first 4 primes on the interval $[1, 210]$. This can also be called a truncation rule because you truncate the growth of the recursion tree. This truncation rule does not take into account the fact that we can have nodes that have ϕ-value one and for which we can already truncate the tree. A new truncation rule that would take this information in to account is based on a closer examination on what values n can take and what this implies for the values $\phi(\frac{x}{n}, b)$. This is what is done in the following methods, the first by *Odlyzko*, *Miller* and *Lagarias*, and in even more depth by the second method by *Deglise* and *Rivat*.

2. The *Lagarias*, *Miller* and *Odlyzko* Method

In 1985 *Lagarias*, *Miller* and *Odlyzko* [**3**] used the algorithm below to compute values upto $\pi(4 \cdot 10^{16})$. They take $a = \pi(x^{1/3})$ which implies that $P_{k \geq 3}(x, a) = 0$. The new corresponding truncation rule is as follows.
Do not split a node if one of the following holds :

1) $b = 0$ and $n \leq x^{1/3}$
2) $n > x^{1/3}$.

The first condition collects all nodes that have $n \leq x^{1/3}$ and that would have been leaves if no truncation rules were to be aplied. We will call nodes that are leaves when no truncation rule is applied, natural leaves. The second condition will collect two types of nodes, firstly those that also have $b = 0$ or the natural leaves that have $n > x^{1/3}$. Secondly, the nodes with values $\mu(n)\phi(\frac{x}{n}, b)$ with $b \neq 0$ and $np_b > x$. Because $p_b \leq p_a \leq x^{1/3}$ we have to have $n \geq x^{2/3}$ for this to happen. The *Odlyzko*, *Miller* and *Lagarias* Method calculates $\pi(x)$ as

$$\pi(x) = \phi(x, \pi(x^{1/3})) + \pi(x^{1/3}) - 1 - P_2(x, \pi(x^{1/3})).$$

To start this calculation we need to sieve the interval $[1, x^{1/3}]$ so that we have all primes less than or equal to $x^{1/3}$ and $\pi(x^{1/3})$. We put these values in an array called F. To calculate $P_2(x, a)$ we need to sieve the interval $[1, x^{2/3}]$ for which the primes $\leq x^{1/3}$ are sufficient. We have,

$$\phi(x, \pi(x^{1/3})) = \sum_{\substack{1 \leq n \leq x \\ n \leq x^{1/3}}} \mu(n) \lfloor \frac{x}{n} \rfloor + \sum_{(n,b)} \mu(n)\phi(\frac{x}{n}, b)$$

Where the sum over the natural leaves with $n \leq x^{1/3}$ equals

$$\sum_{\substack{1 \leq n \leq x \\ n \leq x^{1/3}}} \mu(n) \lfloor \frac{x}{n} \rfloor, \tag{10}$$

and the sum over the new leaves equals

$$\sum_{(n,b)} \mu(n)\phi(\frac{x}{n}, b), \tag{11}$$

with $n = p_{a_1} \cdots p_{a_r}$, $\pi(x^{1/3}) \geq a_1 > \cdots > a_r = b + 1$ and $n > x^{1/3} > \frac{n}{p_{b+1}}$. We have that $\frac{x}{n} \leq x^{2/3}$ because $n > x^{1/3}$ and if we sieve the interval $[1, x^{2/3}]$ by the primes less than or equal to p_a then by exclusion we will come across all values $\phi(\frac{x}{n}, b)$ for all new leaves. The sieving will be done in no more than $N + 2$ blocks of length $N = \lfloor x^{1/3} \rfloor$. Here again we start with our precomputed array F. Also in the process of calculating $P_2(x, a)$ we also sieved the interval $[1, x^{2/3}]$ so basically we can do both the calculation of $P_2(x, a)$ and of the sum of the new leaves in one loop.

The blocks,

$$B_k = [(k-1)N, kN]$$

and a final block of the form $[(k-1)N, x^{2/3}]$ with length probably less than N, are processed in increasing order of the index k. We look at the processing of a general block B_k. From the processed previous block we have available the set of values $\{\phi((k-1)N, j) : 1 \leq j \leq a\}$. We are now going to sieve our current block by the primes $p_1 \cdots p_a$. When we are sieving by a general p_b, which means the primes $p_1, \cdots, p_{b-1}$ are already used to sieve the block B_k, an array A is updated, where A is as follows,

$$\{A(i,j) : 0 \leq \lfloor log_2 N \rfloor, 1 \leq j \leq \lfloor \frac{N}{2^i} \rfloor + 1\}. \tag{12}$$

$A(i,j)$ counts the number of still unsieved elements in the subinterval

$$I_{i,j} = [(k-1)N + (j-1)2^i + 1, (k-1)N + j2^i],$$

with the intervals having $j = \lfloor \frac{N}{2^i} \rfloor + 1$ probably shorter. The array is initialised by puting 1 in every entry. When sieving out, for the first time, an integer $(k-1)N+l$ in this interval, we set the corresponding entries, $A(0,l), A(1, \lfloor \frac{l+1}{2} \rfloor), A(2, \lfloor \frac{l+3}{4} \rfloor), \cdots$, whose corresponding intervals all contain $y = (k-1)N + l$ to 0. This implies that we can calculate $\phi(y, b)$ as follows,

$$\phi((k-1)N + l, b) = \phi((k-1)N, b) + \sum_{r=1}^{m} A(e_r, 1 + \sum_{s<r} 2^{e_s - e_r})$$

where $l = \sum_{r=1}^{m} 2^{e_r}$ is the binary expansion of l. At this time we seek to calculate all values of y in this block such that the pair (y, b) correspond to a leaf.
We have that

$$(k-1)N + 1 \leq y = \lfloor \frac{x}{n} \rfloor \leq kN + 1$$

with $n > x^{1/3} > \frac{n}{p_{b+1}}$, $n = mp_{b+1}$ and $m \leq x^{1/3}$. Also,

$$\frac{x}{(kN+1)p_{b+1}} < m \leq \frac{x}{((k-1)N+1)p_{b+1}},$$

so we are looking for all integers m in the interval

$$J_{k,b} = (\frac{x}{(kN+1)p_{b+1}}, \frac{x}{((k-1)N+1)p_{b+1}}] \bigcap [1, N]$$

for which $\mu(m) \neq 0$ and whose smallest prime factor is less than p_{b+1}. If, while computing array F we also add the $\mu(m)$-values, where m is in $[1, x^{1/3}]$, we have what we need to compute these m-values. For all these found m-values we compute firstly,

$$y = \lfloor \frac{x}{mp_{b+1}} \rfloor = (k-1)N + l$$

and secondly,

$$\mu(n)\phi(y,b) = -\mu(m)\phi(y,b)$$

by using our array $A(i,j)$ (12) on page 19 and a trick, for in our precomputed array F we have the values of $\mu(m)$ and $\mu(n) = -\mu(n/p)$ with p prime. Afterwards we sum these contributions and repeat the process with the next prime p_{b+1}. When block B_k is proccesed by all primes less than or equal to p_a, we update our set of ϕ-values from the previous block and have

$$\{\phi(kN,j) : 1 \leq j \leq a\}.$$

All is set for the next block and after processing all blocks we have summed over the contributions of all new leaves, and we can now easily calculate $\pi(x)$. It is important to note that we have the same sieving limit $x^{2/3}$ for the calculation of $P_2(x,\pi(x^{1/3})$ and for $\phi(\frac{x}{n},b)$ where n square free and greater than $x^{1/3}$. Also beacause we have calculated all contributions of the new set of leaves and can calculate the pairs (n,b) that specify which nodes are new leaves directly, we do in theory not need to use the recurion formula ref. at all in the actual calculation.

3. Complexity of the *Lagarias*, *Miller* and *Odlyzko* method

We will state before hand that in the complexities we calculate the number of elementary operations and not the number of bit operations. Assuming that every elementary operation takes $log(x)$ bit operations the conversion is easy.

$P_2(x,\pi(x^{1/3}))$.

The precomputation of the array F has negligible time and space requirements. Using F we know the primes less than or equal to $x^{1/3}$ and $\pi(x^{1/3})$. The term $P_2(x,\pi(x^{1/3}))$ takes $O(x^{2/3}loglog(x))$ time because we have to sieve $[1,x^{2/3}]$ and the sieving is done in blocks of length $x^{1/3}$ resulting in a space complexity of $O(x^{1/3}log(x))$.

$\phi(x,\pi(x^{1/3}))$.

$\Phi(x,\pi(x^{1/3}))$ is the sum of (10) on page 18 and (11) on page 18. (10) takes no more than $O(x^{1/3})$ time and no more than $O(x^{1/3})$ space for it takes no more than $x^{1/3}$ operations and uses F. For the sum (11) we need to do three things. First we need to sieve the interval $[1,x^{2/3}]$ and update array $A(i,j)$ (12) on page 19. Secondly we need to find the pairs (m,p_{b+1}) of the nodes whose values we need. Thirdly we need to compute the values $\mu(m)\phi(\frac{x}{m},b)$.
Because we have to update array $A(i,j)$ we have that for every integer in the interval $[1,x^{2/3}]$ we have to perform an extra $log(x)$ operations. This is because when we update $A(i,j)$ for a specific integer y, then for all $0 \leq i \leq \lfloor log_2(N)\rfloor$ the process has to mark one $1 \leq j \leq \lfloor \frac{N}{2^i}\rfloor + 1$. This new sieving process costs therefore $O(x^{2/3}log(x) + x^{2/3}loglog(x))$ operations and therefore time.

The (m, p_{b+1}) that correspond to the leaves of which ϕ-values are needed uses array F giving it a space complexity of $O(x^{1/3})$. The time complexity can be shown to be $O(x^{2/3})$, [**3**].
The computations of the values $\mu(n)\phi(\frac{x}{n}, b)$ corresponding to the new leaves consist per leaf in the accessing of array $A(i,j)$ taking $O(log(x))$ operations. Also $O(1)$ elementary operations are needed to change from the $\mu(n)$ value to the $\mu(m)$ value and to multiply the μ- and ϕ-values. This gives a time of $O(x^{2/3}log(x))$ for the computation of the values of the new leaves if the number of new leaves is bounded by $x^{2/3}$. This can be seen from the following,

Lemma 1

There are no more than $x^{2/3}$ new leaves corresponding to condition two of the truncation rule above.

Proof :

Firstly for b fixed there are no two nodes with the same value of n and for fixed n there are no two nodes with the same b as we saw above. This implies that with $n > x^{1/3} \geq m$ we have no more than $x^{1/3}$ m's. Every m needs a corresponding p_{b+1} and of these there are no more than $a = \pi(x^{1/3})$. This gives us an upperbound on the number of newleaves, generated by condition 2 of the truncation rule, of $x^{1/3}\pi(x^{1/3})$. This bound is less than $x^{2/3}$.

End of proof.

Because we have to sum no more than $x^{1/3}$ node values per block, the space complexity for this is $O(x^{1/3})$.
This gives the total algorithm a time complexity equal to $O(x^{2/3}log(x)+x^{2/3}loglog(x)) = O(x^{2/3}log(x))$ and space complexity of $O(x^{1/3})$.

4. The *Deglise* and *Rivat* method

In 1996 *Degise* and *Rivat* [**2**] published the method below with which they had computed $\pi(10^{18})$. The method for calculating $\pi(x)$ used by *Deglise* and *Rivat* is based on a slightly different truncation rule for calculating $\phi(x, a)$ than we had with *Odlyzko*, *Miller* and *Lagarias*. The new truncation rule will correspond to y being a variable having its values in $(x^{1/3}, x^{1/2})$. The new truncation is as follows. Do not split a node if one of the following holds :

1) $b = 0$ and $n \leq y$
2) $n > y$.

With this new truncation rule we have $a = \pi(y)$ and

$$\pi(x) = \phi(x, \pi(y)) + \pi(y) - 1 - P_2(x, \pi(y)).$$

So we have a slightly different P_2-term but because $y \geq x^{1/3}$ we still have $P_{k\geq3}(x, a) =$

0. Condition one will collect natural leaves with $n \leq y$, all other leaves are collected by condition two. Because we have $\frac{x}{y} > y$ we see that the nodes with $\phi(\frac{x}{n}, b) = 1$ or equivalently $\frac{x}{n} < p_b$ are in the set of nodes corresponding to condition two. We have

$$\phi(x, \pi(y)) = \sum_{n=1}^{n \leq y} \mu(n) \lfloor \frac{x}{n} \rfloor + \sum_{(n,b)} \mu(n) \phi(\frac{x}{n}, b)$$

with $n = p_{a_1} \cdots p_{a_r}$, $\pi(y) \geq a_1 > \cdots > a_r = b + 1$ and $n > y > \frac{n}{p_{b+1}}$.
Here the sum of the leaves corresponding to condition one now equals,

$$\sum_{n=1}^{n \leq y} \mu(n) \lfloor \frac{x}{n} \rfloor$$

and the sum corresponding to the new leaves is

$$\sum_{(n,b)} \mu(n) \phi(\frac{x}{n}, b).$$

The difficult part is again calculation of the sum of the leaves corresponding to condition two. We will take a closer look at these new leaves and see that we can more efficiently calculate the sum of these leaves by spliting the calculation over different value ranges of the smallest prime divisor p of n. Every leaf will now be of the following form,

$$\mu(mp) \phi(\frac{x}{mp}, \pi(p) - 1)$$

with m square free, every prime divisor larger than p and $m \leq y$.
We can write the sum over the new leaves as,

$$S = \sum_{p \leq y} \sum_{m \leq y < mp} \mu(mp) \phi(\frac{x}{mp}, \pi(p) - 1)$$

with m having the properties above. We write this sum as $S = S_1 + S_2 + S_3$ with,

$$S_1 = \sum_{x^{1/3} < p \leq y} \sum_{m \leq y < mp} \mu(mp) \phi(\frac{x}{mp}, \pi(p) - 1),$$

$$S_2 = \sum_{x^{1/4} < p \leq x^{1/3}} \sum_{m \leq y < mp} \mu(mp) \phi(\frac{x}{mp}, \pi(p) - 1),$$

$$S_3 = \sum_{p \leq x^{1/4}} \sum_{m \leq y < mp} \mu(mp) \phi(\frac{x}{mp}, \pi(p) - 1).$$

For fixed p in $(x^{1/4}, y]$, suppose that a pair (p, m) contributes to the sum S_1 or sum S_2. By assumption m is a square free number in $(\frac{y}{p}, y]$ and all its prime divisors are greater than p. If m is not prime then $m > p^2$, hence $y \geq m > p^2 > \sqrt{x}$ which contradicts our choice $y \in [x^{1/3}, x^{1/2}]$. Hence m is prime. Also we then have that $q \leq y$ and $y \leq \sqrt{x} < p^2 < pq$ so that every prime $q \in (p, y]$ contributes to S_1 or S_2.

So, noting that $\mu(pq) = 1$, we can write

$$S_1 = \sum_{x^{1/3} < p \le y} \sum_{p < q \le y} \phi(\frac{x}{pq}, \pi(p) - 1)$$

and

$$S_2 = \sum_{x^{1/4} < p \le x^{1/3}} \sum_{p < q \le y} \phi(\frac{x}{pq}, \pi(p) - 1).$$

Calculating S_1.

Because $p > x^{1/3}$ we have $pq > x^{2/3}$ and therefore $\frac{x}{pq} < x^{1/3} < p$. This in turn implies that for all pairs (p, q) contributing to sum S_1 we have $\phi(\frac{x}{pq}, \pi(p) - 1) = 1$. So we only have to count these pairs (p, q) with $x^{1/3} < p < q \le y$, giving,

$$S_1 = \frac{(\pi(y) - \pi(x^{1/3}))^2 - (\pi(y) - \pi(x^{1/3}))}{2}$$

Calculating S_2.

We split the sum $S_2 = U + V$ as follows,

$$U = \sum_{x^{1/4} < p \le x^{1/3}} \sum_{\substack{p < q \le y \\ q > x/p^2}} \phi(\frac{x}{pq}, \pi(p) - 1)$$

and

$$V = \sum_{x^{1/4} < p \le x^{1/3}} \sum_{\substack{p < q \le y \\ q \le x/p^2}} \phi(\frac{x}{pq}, \pi(p) - 1).$$

Calculating U.

Because

$$p \le \sqrt{\frac{x}{y}} \Rightarrow p^2 \le \frac{x}{y} \Rightarrow q < y \le \frac{x}{p^2} \Rightarrow \frac{x}{pq} > p. \tag{13}$$

We have firstly that we can write,

$$U = \sum_{\sqrt{\frac{x}{y}} < p \le x^{1/3}} \sum_{\substack{p < q \le y \\ q > x/p^2}} \phi(\frac{x}{pq}, \pi(p) - 1).$$

and secondly that all these ϕ-values will also equal one and so we only have to count the number of pairs (p,q) that contribute to sum U. We have,

$$U = \sum_{\sqrt{\frac{x}{y}} < p \le x^{1/3}} |\{q : \frac{x}{p^2} < q \le y\}| = \sum_{\sqrt{\frac{x}{y}} < p \le x^{1/3}} (\pi(y) - \pi(\frac{x}{p^2})).$$

Calculating V.

Because of formula (13) above we can write,

$$V = \sum_{x^{1/4} < p \le \sqrt{\frac{x}{y}}} \sum_{p < q \le y} \pi(\frac{x}{pq}, \pi(p) - 1) + \sum_{\sqrt{\frac{x}{y}} < p \le x^{1/3}} \sum_{p < q \le \frac{x}{p^2}} \pi(\frac{x}{pq}, \pi(p) - 1).$$

So in V all terms for which $p > x^{1/4}$ and $\frac{x}{pq} \ge p$ are collected. Furthermore,

$$x^{1/4} < p \Rightarrow x^{1/2} < p^2 \Rightarrow \frac{x}{p^2} < x^{1/2} \Rightarrow \frac{x}{pq} < x^{1/2} < p^2.$$

and together with $\frac{x}{pq} \ge p$ we get $p \le \frac{x}{pq} < p^2$. This implies,

$$\phi(\frac{x}{pq}, \pi(p) - 1) = 1 + \pi(\frac{x}{pq}) - (\pi(p) - 1) = 2 + \pi(\frac{x}{pq}) - \pi(p).$$

So,

$$V = \sum_{x^{1/4} < p \le \sqrt{\frac{x}{y}}} \sum_{p < q \le y} 2 + \pi(\frac{x}{pq}) - \pi(p) + \sum_{\sqrt{\frac{x}{y}} < p \le x^{1/3}} \sum_{p < q \le \frac{x}{p^2}} 2 + \pi(\frac{x}{pq}) - \pi(p).$$

We split $V = V_1 + V_2$ as follows,

$$V_1 = \sum_{x^{1/4} < p \le \sqrt{\frac{x}{y}}} \sum_{p < q \le y} 2 - \pi(p) + \sum_{\sqrt{\frac{x}{y}} < p \le x^{1/3}} \sum_{p < q \le \frac{x}{p^2}} 2 - \pi(p)$$

where V_1 is the sum over terms $2 - \pi(p)$ of V and

$$V_2 = \sum_{x^{1/4} < p \le \sqrt{\frac{x}{y}}} \sum_{p < q \le y} \pi(\frac{x}{pq}) + \sum_{\sqrt{\frac{x}{y}} < p \le x^{1/3}} \sum_{p < q \le \frac{x}{p^2}} \pi(\frac{x}{pq})$$

is the sum over terms $\pi(\frac{x}{pq})$ of V. V_1 can simply be calculated when we know all primes less than or equal to y.

Calculating V_2.

V_2 is split in a sum over five sums $W_1, \cdots, W_5$, as follows,

$$W_1 = \sum_{x^{1/4} < p \le \frac{x}{y^2}} \sum_{p < q \le y} \pi(\frac{x}{pq}),$$

$$W_2 = \sum_{\frac{x}{y^2} < p \le \sqrt{\frac{x}{y}}} \sum_{p < q \le \sqrt{\frac{x}{p}}} \pi(\frac{x}{pq}),$$

$$W_3 = \sum_{\frac{x}{y^2} < p \le \sqrt{\frac{x}{y}}} \sum_{\sqrt{\frac{x}{p}} < q \le y} \pi(\frac{x}{pq}),$$

$$W_4 = \sum_{\sqrt{\frac{x}{y}} < p \le x^{1/3}} \sum_{p < q \le \sqrt{\frac{x}{p}}} \pi(\frac{x}{pq}),$$

$$W_5 = \sum_{\sqrt{\frac{x}{y}} < p \le x^{1/3}} \sum_{\sqrt{\frac{x}{p}} < q \le \frac{x}{p^2}} \pi(\frac{x}{pq}).$$

Notice that all W-sums need values of $\pi(\frac{x}{pq})$ in $[\frac{x}{y^2}, x^{1/2}]$.
For $\frac{x}{y^2}$ to be in $[x^{1/4}, x^{1/3}]$ it must be that $y \le x^{3/8}$. If $y = x^{1/3}$ we have interestingly that $y = x^{1/3} = \frac{x}{y^2} = \sqrt{\frac{x}{y}}$! This implies that for this value of y which equals the value used by *Lagarias*, *Miller* and *Odlyzko*, the sums $W_2 = W_3 = W_4 = W_5 = U = S_1 = 0$. So for this calculation to make sense we have to have $x^{1/3} = x^{16/48} < y < x^{3/8} = x^{18/48}$. For instance $y = x^{17/48}$.

We will calculate the maximal and minimal values that $\frac{x}{pq}$ will have for W_i $1 \le i \le 5$ and then we know what values $\pi(\frac{x}{pq})$ will be needed in the W_i. Also a closer look at the pairs (p, q) that contribute to the sums W_i. These results will then suggest a calculation per specific W_i.

Calculating W_1.

For p fixed we have $\frac{x}{pq} < \frac{x}{p^2}$ and with $p = x^{1/4}$ we get $\frac{x}{pq} < x^{1/2}$. For p fixed we have $\frac{x}{pq} \ge \frac{x}{py}$ and with $p = \frac{x}{y^2}$ we get $\frac{x}{pq} \ge y$. So we need to sieve at least upto $x^{1/2}$ to get these values. This will be accomplished by sieving the interval $[1, x^{1/2}]$ in blocks of length y.

Calculating W_2.

We have for p fixed that $\frac{x}{pq} < \frac{x}{p^2}$ and with $p = \frac{x}{y^2}$ we get $\frac{x}{pq} < \frac{y^4}{x}$. For p fixed we have $\frac{x}{pq} \ge \frac{x}{p\sqrt{\frac{x}{p}}}$ and with $p = \sqrt{\frac{x}{y}}$ we get $\frac{x}{pq} \ge (xy)^{1/4}$. Now $x^{1/3} \le y \le x^{1/2}$ implies $x^{1/3} \le \frac{y^4}{x} \le x$ and $x^{1/3} \le (xy)^{1/4} \le x^{3/8}$. Here again values of $\pi(\frac{x}{pq})$ with $y \le \frac{x}{pq} \le x^{1/2}$ will be needed and in fact the idea is to calculate W_1 and W_2 at the same time.

Calculating W_3.

We have for p fixed that $\frac{x}{pq} < \frac{x}{p\sqrt{\frac{x}{p}}}$ and with $p = \frac{x}{y^2}$ we get $\frac{x}{pq} < y$. For p

fixed we have $\frac{x}{pq} \geq \frac{x}{py}$ and with $p = \sqrt{\frac{x}{y}}$ we get $\frac{x}{pq} \geq \sqrt{\frac{x}{y}}$. So for W_3 we need only values less than or equal to $\pi(y)$. The idea here is that the computation can be sped up by looking at the function $q \rightarrow \pi(\frac{x}{pq})$ for fixed p. Remember we allready have a list of primes less than or equal to z and $\pi(z)$-values for $z \leq y$. With this list we can check for which values of q the function $q \rightarrow \pi(\frac{x}{pq})$ changes and count the number of q in each subinterval of $[\sqrt{\frac{x}{p}}, y]$ on which this function is constant. After which you multiply the number of q in a subinterval by the constant value the function has on this interval. This works only ofcourse if there are multiple q-values in these subintervals. For $p = \frac{x}{y^2}$ we have $y < q \leq y$ and for $p = \sqrt{\frac{x}{y}}$ we have $\sqrt{\frac{x}{y}} < q \leq y$. Plugging in $y = x^{17/48}$ gives for that latest interval, $x^{31/96} < q \leq x^{34/96}$. So clearly not much room for the values of q, for fixed p, and therefore relatively few different values $\pi(\frac{x}{pq})$.

Calculating W_4.

We have for p fixed that $\frac{x}{pq} < \frac{x}{p^2}$ and with $p = \sqrt{\frac{x}{y}}$ we get $\frac{x}{pq} < y$. For p fixed we have $\frac{x}{pq} \geq \frac{x}{p\sqrt{\frac{x}{p}}}$ and with $p = x^{1/3}$ we get $\frac{x}{pq} \geq x^{1/3}$. For $p = \sqrt{\frac{x}{y}}$ we have $\sqrt{\frac{x}{y}} < q \leq (xy)^{1/4}$ and for $p = x^{1/3}$ we have $x^{1/3} < q \leq x^{1/3}$. Plugging in $y = x^{17/48}$ gives for the first interval, $x^{62/96} < q \leq x^{65/96}$. *Deglise* and *Rivat* state that here there is no advantage to proceed as with W_3 because most values of $\pi(\frac{x}{pq})$ will be distinct, for fixed p. We indeed larger values of q then we had with W_3 so $|\frac{x}{pq_i} - \frac{x}{pq_{i+1}}|$ will be greater for the q_i contributing to sum W_4 .

Calculating W_5.

We have for p fixed that $\frac{x}{pq} < \frac{x}{p\sqrt{\frac{x}{p}}}$ and with $p = \sqrt{\frac{x}{y}}$ we get $\frac{x}{pq} < (xy)^{1/4}$. For p fixed we have $\frac{x}{pq} \geq \frac{x}{p\frac{x}{p^2}}$ and with $p = x^{1/3}$ we get $\frac{x}{pq} \geq \sqrt{\frac{x}{y}}$. For $p = \sqrt{\frac{x}{y}}$ we have $(xy)^{1/4} < q \leq y$ and $p = x^{1/3}$ gives $x^{1/3} < q \leq x^{1/3}$. Plugging in $y = x^{17/48}$ we get $x^{65/192} < q \leq x^{68/192}$. *Deglise* and *Rivat* state that it is advantagious to proceed as with W_3.

Calculating S_3.

The sum S_3 is calculated sieving the interval $[1, \frac{x}{y}]$ by primes less than or equal to $x^{1/4}$ using the sieving method from *Lagarias*, *Miller* and *Odlyzko*. Finally all sums are summed and we have $\pi(x)$.

5. Complexity of the *Deglise* and *Rivat* method

The precomputed array F takes $O(yloglog(x))$ and we then trivially have $\pi(y)$ in time $O(1)$ and space $O(y)$. The sieving of $[1, \frac{x}{y}]$ nescessary for the computation of $P_2(x, \pi(y))$ takes $O(\frac{x}{y}loglog(x))$ time and is done in blocks of length y so the space complexity is $O(y)$. The sum over the leaves corresponding to condition one of the truncation rule takes $O(y)$ time and $O(log(x))$ space. We have,

$$S = S_1 + U + V_1 + W_1 + W_2 + W_3 + W_4 + W_5 + S_3$$

S_1.
We have the values of $\pi(y)$ en $\pi(x^{1/3})$ in our array F so S_1 takes $O(1)$ time and $O(log(x))$ space.

U.
This sum has to sum over no more than $x^{1/3}$ $p's$ so the number of elementairy operations is bounded by $\pi(x^{1/3})$. Using the prime number theorem this is about $\frac{x^{1/3}}{log(x^{1/3})}$ so we have a time of $O(\frac{x^{1/3}}{log(x)})$ and space again $O(log(x))$.

V_1.
For V_1 we have that for every p the number of q can be found in constant time $O(1)$, and we have no more than $\pi(x^{1/3})$ terms, giving a time of $O(x^{1/3})$ and space of $O(log(x))$.

W_1.
For the computation we need to sieve the interval $[1, \sqrt{x}]$ which takes $O(\frac{x}{y}loglog(x))$ time and $O(y)$ space. In the sum there are no more than $\pi(\frac{x}{y^2})\pi(y)$ terms. This equals approximately $\frac{x/y^2}{log(x/y^2)}\frac{y}{log(y)}$ terms. We can replace the log-terms by $(log(x))^2$ by noting that y equals some power of x, so the powers can be delegated to the constant implied by the O-notation. So we have $O(\frac{x}{y(log(x))^2})$ terms which equals the time complexity and the space needed is of course $O(log(x))$.

W_2.
For p fixed we have $\pi(\sqrt{\frac{x}{p}})$ q's. This gives the number of pairs (p, q), which equals the number of terms in the sum, as,

$$\sum_{\frac{x}{y^2} < p \leq \sqrt{\frac{x}{y}}} \pi(\sqrt{\frac{x}{p}}) \approx \pi(\sqrt{\frac{x}{y}})\pi(\sqrt{\frac{x}{\sqrt{x/y}}}) = \pi(\sqrt{\frac{x}{y}})\pi((xy)^{1/4}) \approx$$

$$\frac{\sqrt{\frac{x}{y}}}{log(\sqrt{\frac{x}{y}})}\frac{(xy)^{1/4}}{log((xy)^{1/4})} = C\frac{x^{3/4}}{y^{1/4}log(x^2)}$$

We know y to be some power of x which implies that we can again replace the log-terms with $log(x)$ but we have to put a constant C in front. This gives a time complexity of $O(\frac{x^{3/4}}{y^{1/4}(log(x))^2})$ and we know that we sieve in blocks of length y giving space complexity $O(y)$.

W_3.
For p fixed we have $\frac{x}{pq} \leq \frac{x}{p\sqrt{\frac{x}{p}}} = \sqrt{\frac{x}{p}}$. So the number of values $\pi(\frac{x}{pq})$ can take

is bounded by $\pi(\sqrt{\frac{x}{p}})$. For each such value it takes constant time to determine which q's correspond to this value. This means that with p fixed it takes no more time than the time needed for $\pi(\sqrt{\frac{x}{p}})$ elementairy operations, giving again a time of $O(\sum_{\frac{x}{y^2}<p\leq\sqrt{\frac{x}{y}}}\pi(\sqrt{\frac{x}{p}})) = O(\frac{x^{3/4}}{y^{1/4}(log(x))^2})$ as we saw above with W_2.

W_4.

For p fixed we have no more than $\pi(\sqrt{\frac{x}{p}})$ q's giving a time of,

$$O(\sum_{\sqrt{\frac{x}{y}}<p\leq x^{1/3}}\pi(\sqrt{\frac{x}{p}})) = O(\pi(x^{1/3})\pi(\sqrt{\frac{x}{x^{1/3}}})) = O(\frac{x^{2/3}}{(log(x))^2})$$

W_5.

or p fixed we have $\frac{x}{pq} \leq \frac{x}{p\frac{x}{p^2}} = p$. So the number of values $\pi(\frac{x}{pq})$ can take is bounded by p. For each such value it takes again constant time to determine which q's correpond to each π-value.This means that with p fixed it takes no more time than the time needed for $\pi(p)$ elementairy operations. This gives a time complexity of,

$$O(\sum_{\sqrt{\frac{x}{y}}<p\leq x^{1/3}}\pi(p)) = O(\pi(x^{1/3})\pi(\sqrt{\frac{x}{y}})) = O(\frac{x^{2/3}}{(log(x))^2})$$

because $y \geq x^{1/3}$.

S_3.

We have seen that in computing S_3 we use the sieve construction of [**3**], to sieve the interval $[1, \frac{x}{y}]$ by primes less than or equal to $x^{1/4}$. This sieve with sieving limit $\frac{x}{y}$ and blocklength y takes $O(\frac{x}{y}log(x)loglog(x))$ time and $O(y)$ space, according to *Deglise* and *Rivat*. Note that with the method of [**3**] we had, with $y = x^{1/3}$, $O(x^{1/3}(log(x)+loglog(x))$ time which is slightly different. This is because the time of *Deglise* and *Rivat* assumes that even if we had already sieved out an integer in a sieving block, than we still carry out the operation of updating array $A(i,j)$ (12) on page 19, which in practice need not happen. But the *Deglise* and *Rivat* time is greater in theory if $loglog(x) > 2$, which is true for $x > 10^4$. All leaves in the sum S_3 are of the form $\mu(mp_b)\phi(\frac{x}{mp_b}, b-1)$ with $m \leq y$ and $b < x^{1/4}$. This implies that the number of leaves in this sum is bounded by $O(y\pi(x^{1/4}))$. When computing the leave values one needs to acces the array $A(i,j)$ (12) using $O(log(x))$ operations. Using the prime number theorem once again, this gives $O(y\frac{x^{1/4}}{log(x)}log(x) = O(yx^{1/4})$ time for computing the leaf values needed in S_3. Time is also needed to compute the values of m we have per sieving block, as we saw above. This can be done in $O(\frac{x}{y})$ time, the case $y = x^{1/3}$ is explained in [**3**]. This gives a total time for the computation of S_3, of $O(\frac{x}{y}log(x)loglog(x) + yx^{1/4})$ and a space of $O(y)$.

Complexity of the algorithm.

Because $y \geq x^{1/3}$ the W_i-sums with time complexity depending on y are all dominated by a time of $O(\frac{x^{2/3}}{(log(x))^2})$. The only other significant time complexities are those of S_3 and of the sieve for $P_2(x, \pi(y))$. This results in a total time for the algorithm of *Deglise* and *Rivat* of,

$$O(\frac{x}{y} loglogx) + \frac{x}{y} log(x) loglog(x) + x^{1/4} y + \frac{x^{2/3}}{(log(x))^2}).$$

Now *Deglise* and *Rivat* plug in $y = x^{1/3}(log(x))^3 loglog(x)$ to get a time complexity of,

$$O(\frac{x^{2/3}}{(log(x))^2}).$$

This finally gives a space complexity of $O(x^{1/3}(log(x))^3 loglog(x))$.

CHAPTER 3

Riemann and $\pi(x)$

1. The Prime Power Counting Function $J(x)$

In 1859 *B. Riemann* published his groundbreaking eight page paper “Ueber die Anzahl der Primzahlen unter einer gegebenen Grösse” in which he connected $\pi(x)$ by way of smart use of function theory to the Zeta function $\zeta(s)$ and its zero's [**1**]. In his paper he starts with :

Lemma 2

Euler's product formula

$$\sum_{n=1}^{\infty}\frac{1}{n^s}=\prod_{p}\frac{1}{1-p^{-s}} \tag{14}$$

for $\Re(s)=a>1$ and p prime.

Proof :

For any prime p we have :

$$\frac{1}{1-p^{-s}}=1+\frac{1}{p^s}+\frac{1}{p^{2s}}+\frac{1}{p^{3s}}+...$$

for $\Re(s)=a>1$ so that $|p^{-s}|<1$ and we have absolute convergence of the geometric series. We have that when a product is taken over a finite number of absolutely convergent series that the product is absolutely convergent and we may arrange the terms of this series in any way possible without altering the series itself. So if we take some prime number q we can look at the formula :

$$|\zeta(s)-\prod_{p\leq q}\frac{1}{1-p^{-s}}|=|\sum\frac{1}{n^s}|,$$

where on the right hand side of this equation we have collected all n containing only prime divisors greater than q. Looking at the right hand side we have :

$$|\sum\frac{1}{n^s}|\leq\sum_{n=q+1}^{\infty}|\frac{1}{n^s}|=\sum_{n=q+1}^{\infty}\frac{1}{n^{\sigma}}$$

with $s=a+it$ and $a>1$ where this last series is a simple and absolutely convergent harmonic series which implies that

$$\lim_{q\to\infty}\sum_{n=q+1}^{\infty}\frac{1}{n^a}=0$$

and therefore,

$$\lim_{q\to\infty}|\zeta(s)-\prod_{p\leq q}\frac{1}{1-p^{-s}}|=0.$$

End of proof.

Following this Riemann uses the right hand side of the Euler product formula (14) on page 31 to write $log\zeta(s)$ as a double series as follows :

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

$Log(1+x)$ has a simpel Taylor series that converges for $|x|<1$ which in our case is no problem with $x=p^{-s}$ and $Re(s)>1$ and gives :

$$-log(1-p^{-s})=p^{-s}+\frac{1}{2}p^{-2s}+\frac{1}{3}p^{-s}+\cdots$$

which equals $\sum_{n=1}^{\infty}n^{-1}p^{-ns}$, giving the double series :

$$\sum_p\sum_{n=1}^{\infty}n^{-1}p^{-ns}.$$

As this series converges absolutely we can interchange the summations and write with Riemann :

$$log\zeta(s)=\sum_p p^{-s}+\frac{1}{2}\sum_p p^{-2s}+\frac{1}{3}\sum_p p^{-3s}+\cdots$$

Now using the identity $p^{-ns}=s\int_{p^n}^{\infty}x^{-s-1}dx$ and first looking only at $n=1$ we get :

$$\sum_p p^{-s}=\sum_p s\int_p^{\infty}x^{-s-1}dx=s\int_2^{\infty}x^{-s-1}dx+s\int_3^{\infty}x^{-s-1}dx+s\int_5^{\infty}x^{-s-1}dx\cdots$$

On the right we have an infinite sum of terms where each term is almost the same in the sense that only the lower integration boundary changes to the next prime. We can change the boundary to 2 for all terms if we just compensate bij multiplying the integrand by a function that is zero from 2 to p if we are looking at the p-term and 1 from then onwards or :

$$\int_2^{\infty}x^{-s-1}dx+\int_3^{\infty}x^{-s-1}dx+\int_5^{\infty}x^{-s-1}dx+...=\int_2^{\infty}\sum_p 1_{[p,\infty)}(x)x^{-s-1}dx,$$

where we have put s in front and we forget about it for a moment. We have :

$$\sum_p 1_{[p,\infty)}(x) = \pi(x).$$

This is what we got for $n = 1$ in our series above so now looking at the general n-th term we get an indicator function, now not starting its interval of definition at p but at p^n and it gets multiplied by $1/n$ so,

$$\int_2^\infty x^{-s-1}dx = \int_p^\infty \frac{1}{n}\sum_p 1_{[p^n,\infty)}(x)x^{-s-1}dx.$$

Taking the series over all n and putting s back in front, we have :

$$log\zeta(s) = s\sum_{n=1}^\infty \int_2^\infty \frac{1}{n}\sum_p 1_{[p^n,\infty)}(x)x^{-s-1}dx = s\int_2^\infty \sum_{n=1}^\infty \frac{1}{n}\sum_p 1_{[p^n,\infty)}(x)x^{-s-1}dx.$$

We have that for all $n \geq 1$,

$$\frac{1}{n}\sum_p 1_{[p^n,\infty)}(x) \leq x,$$

because for n fixed it counts prime powers less than or equal to x and multiplies it with a constant less than one, $\frac{1}{n}$. Also,

$$\int_2^\infty |x \cdot x^{-s-1}|dx = \int_2^\infty x^{-a}dx < \infty,$$

because $\Re(s) = a > 1$. So using x^{-a} we can use the dominated convergence theorem to put the series over n under the integral.
Now we take a look at the double series over the indicator functions,

$$\sum_{n=1}^\infty \frac{1}{n}\sum_p 1_{[p^n,\infty)}(x) = \sum_{n=1}^\infty \frac{1}{n}\sum_{p^n \leq x} 1_{[2,\infty)}(x),$$

where in the second sum we sum over primes p such that $p^n \leq x$. If we define :

$$J(x) = \sum_{n=1}^\infty \frac{1}{n}\sum_{p^n \leq x} 1, \tag{15}$$

then we have :

$$\frac{log\zeta(s)}{s} = \int_2^\infty J(x)x^{-s-1}dx.$$

$J(x)$ is the so-called primepower counting function originally named $f(x)$ by Riemann. The argument of this function slides across the positive reals from left to right and whenever it comes across a primepower p^n it adds $\frac{1}{n}$ starting with zero if we define it zero on the interval $[0,2)$. Also with this slight adjustment we have :

$$\frac{log\zeta(s)}{s} = \int_0^\infty J(x)x^{-s-1}dx. \tag{16}$$

Now we try to extract the J-function with Fourier's inversion formula, which says modulo some constraints on the type of functions allowed, that :

$$\phi(x) = \int_{-\infty}^{\infty} \Phi(t)e^{-itx}dt \Leftrightarrow \Phi(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \phi(x)e^{ixt}dx.$$

Setting $s = a+it$ with $a > 1$ and using a simple but clever substitution, $\lambda = log(x)$ we can write equation (16) above as :

$$\frac{log\zeta(a+it)}{a+it} = \int_{-\infty}^{\infty} J(e^\lambda)e^{-(a+it)\lambda}d\lambda,$$

because $e^\lambda = x$, $e^{-(a+it)\lambda} = x^{-(a+it)}$ and $d\lambda = dlogx = \frac{dx}{x}$. Notice that $x \in (0,\infty)$ implies that $\lambda \in (-\infty,\infty)$.
If we now write :

$$\Phi(t) = \frac{log\zeta(a+it)}{a+it} = \frac{1}{2\pi}\int_{-\infty}^{\infty} \phi(\lambda)e^{-it\lambda}d\lambda,$$

then by Fourier inversion we get :

$$2\pi e^{-a\lambda}J(e^\lambda) = \int_{-\infty}^{\infty} \frac{log\zeta(a+it)}{a+it}e^{it\lambda}dt$$

and

$$J(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{log\zeta(a+it)}{a+it}x^{a+it}dt.$$

Now $a > 1$ is fixed here and t goes from minus infinity to positive infinity, so we recover a complex integral over the contour from $a - i\infty$ to $a + \infty$ on the vertical complex line $a = Re(s)$ after adding a term $1/i$ which comes from the fact that the derivative of the contour $c(t) = a+it$ equals i. So writing s instead of $a+it$, we get :

$$J^*(x) = \frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty} \frac{x^s}{s}log\zeta(s)ds. \tag{17}$$

It is important to note that for values of x that equal a prime power the integral (17) on the right does not equal the value of $J(x)$. The value of the integral equals the average of the left- and right limit of $J(x)$ at that point. We can bypass this problem by just making sure we calculate $J(x)$ when we need $J(\lfloor x \rfloor)$ with $x > \lfloor x \rfloor$. But to not be confused we call the integral $J^*(x)$.

2. Inversion and a Formula for $\pi(x)$

Following *Riemann* we will use the function $J(x)$ (15) on page 33 to derive a function for $\pi(x)$. If we consider the definition of $J(x)$, we have :

$$J(x) = \sum_{n=1}^{\infty} \frac{1}{n} \sum_{p^n \leq x} 1 = \sum_{n=1}^{\infty} \frac{1}{n} \sum_{p \leq x^{1/n}} 1 = \sum_{n=1}^{\infty} \frac{1}{n} \pi(x^{1/n}).$$

Here we sum over pairs (p, n) such that $p^n \leq x$. We could of course put all the terms on the right hand side except $\pi(x)$ on the lefthand side with a minus sign and we would have a formula for $\pi(x)$ in terms of $J(x)$ and the terms involving the $\pi(x^{\frac{1}{n}})$, $n > 1$. If we have the values of $\pi(x^{1/n})$ with $n \geq 2$ and we could find a fast way to compute the integral, we would have an effective way to compute $\pi(x)$.
We will now write $\pi(x)$ as a linear combination of values of $J(x)$ as follows :

Lemma 3

$$\pi(x) = \sum_{n=1}^{\infty} \frac{\mu(n)}{n} J(x^{\frac{1}{n}}). \tag{18}$$

Proof :

A well known property of the möbius function is :

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

We have,

$$\sum_{n=1}^{\infty} \frac{\mu(n)}{n} J(x^{\frac{1}{n}}) = \sum_{n=1}^{\infty} (\frac{\mu(n)}{n} \sum_{m=1}^{\infty} \frac{\pi(x^{\frac{1}{mn}})}{m}) = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \frac{\mu(\frac{mn}{m})}{mn} \pi(x^{\frac{1}{mn}})$$

$$= \sum_{t=1}^{\infty} \sum_{m|t} \frac{\mu(\frac{t}{m})}{t} \pi(x^{\frac{1}{t}}) = \sum_{t=1}^{\infty} (\frac{\pi(x^{\frac{1}{t}})}{t} \sum_{d|t} \mu(d)) = \pi(x).$$

End of proof.

If we plug in our integral formula $J^*(x)$ (17) on page 34 for $J(x)$ we obtain Riemann's formula for $\pi(x)$:

$$\pi(x) = \frac{1}{2\pi i} \sum_{n=1}^{\infty} \frac{\mu(n)}{n} \int_{a-i\infty}^{a+i\infty} \frac{x^{\frac{s}{n}}}{s} log\zeta(s) ds. \tag{19}$$

This sum is actually a finite sum, as is the case with the equation of $J(x)$ in terms of terms $\frac{1}{n}\pi(x^{\frac{1}{n}})$, because both $J(x)$ and $\pi(x)$ are zero on the interval $[0, 2)$.

3. Difficulty with the Integral for $J(x)$

If we would want to use formula (19) above for actually calculating $\pi(x)$, we have to calculate our complex contour integral $J^*(x)$ (17) on page 34. This integral is not absolutely convergent. Because the absolute value of the integrand is the product of the absolute value of $\frac{x^s}{s}$ and $log\zeta(s)$ we look at both factors :

$$|log\zeta(a \pm iT)| = |\sum_n \sum_p \frac{1}{n} p^{-n(a \pm iT)}| \leq \sum_n \sum_p \frac{1}{n} p^{-na} = log\zeta(a)$$

so the $log\zeta(s)$ term is bounded by a constant if $a > 1$. The other factor is :

$$|\frac{x^{a \pm iT}}{a \pm iT}| = \frac{x^a}{\sqrt{a^2 + T^2}}$$

and for $T \to \infty$ this behaves like $\frac{1}{|T|}$ which when integrated gives $log(T)$ which shows that the integral, over the absolute value of the integrand, diverges. That also means that we have to specify exactly how we integrate because different calculations can give, when we have non-absolutely convergent integrals, different answers. We define very naturally :

$$\frac{1}{2\pi i} \int_{a-i\infty}^{a+\infty} \frac{x^s}{s} log\zeta(s) ds = \lim_{T \to \infty} \frac{1}{2\pi i} \int_{a-iT}^{a+iT} \frac{x^s}{s} log\zeta(s) ds$$

The integral is difficult to compute, there are no known simple primitives of the integrand and the best one might be able to do is to take a T-value and calculate the integral numerically for this value of T. This also seems problematic.
The integral can be rewritten as a integral of a real function as follows :

Lemma 4

$$J^*(x) = \frac{1}{\pi} \int_0^\infty \Re(\frac{x^{a+it}}{a+it} log\zeta(a+it))dt. \qquad (20)$$

Proof :

We have that if we define the left integrand as $h(s)$ that $\overline{h(s)} = h(\overline{s})$. Complex conjugation commutes with the rational operations so $\overline{\frac{1}{s}} = \frac{1}{\overline{s}}$. Next we have :

$$x^s = x^{a+it} = \exp((a+it)logx) = x^a \exp(itlogx),$$

so writing $s = a + it$ it follows that :

$$\overline{x^s} = \overline{x^a \exp(itlogx)} = x^a \exp(-itlogx) = x^{a-it} = x^{\overline{s}}.$$

Also, writing $s = re^{i\theta}$ with $-\pi < \theta \leq \pi$ one has :

$$log(s) = log(r \exp(i\theta) = log(r) + i\theta \Rightarrow$$

$$\overline{log(s)} = log(r) - i\theta = log(r\exp(-i\theta) = log(\overline{s}).$$

Here we have chosen a principle branch for the complex logarithm. Also,

$$\overline{\zeta(s)} = \overline{\sum_{n=1}^{\infty} \frac{1}{n^s}} = \sum_{n=1}^{\infty} (\overline{\frac{1}{n^s}}) = \sum_{n=1}^{\infty} \frac{1}{n^{\overline{s}}} = \zeta(\overline{s})$$

so we have indeed that $\overline{h(s)} = h(\overline{s})$. Using the fact that $\Re(s) = \frac{s+\overline{s}}{2}$ and writing in accordance with the previous notation,

$$h(a+it) = \frac{x^{a+it}}{a+it} log\zeta(a+it)$$

we can write :

$$\frac{1}{\pi}\int_0^{\infty} \Re(h(a+it))dt = \frac{1}{2\pi}\int_0^{\infty}(h(a+it)+\overline{h(a+it)})dt = \frac{1}{2\pi}\int_0^{\infty}(h(a+it)+h(\overline{a+it}))dt.$$

Converting this back to a complex integral over the vertical line $a = \Re(s)$ we get that this equals :

$$\frac{1}{2\pi i}\int_a^{a+i\infty}(h(s)+h(\overline{s}))ds.$$

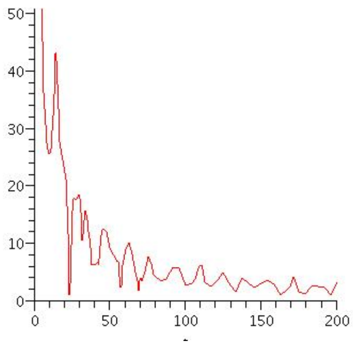
Further,

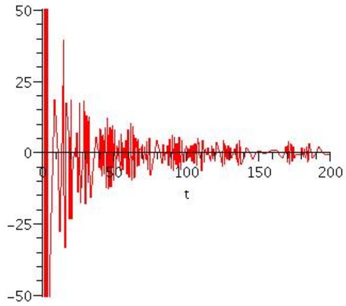
$$\frac{1}{2\pi i}\int_a^{a+i\infty} h(\overline{s})ds = \frac{1}{2\pi i}\int_{a-i\infty}^{a} h(s)ds$$

which is the direct result of the fact that the contour over which we integrate is a vertical line in the complex plane and complex conjugation geometrically equals reflecting in the real line.

End of proof.

Using this real function version of our integral (20) on page 36 we can do some plotting and calculations to see why there are problems if we try to approximate it by taking a finite $T = 200$ boundary and taking $a = 3/2$. For $x = 100$ we get these plots in Maple :





Where on the left we have a figure of $|\frac{x^{3/2+it}}{3/2+it}log\zeta(3/2+it)|$ and on the right we have a figure of $\Re(\frac{x^{3/2+it}}{3/2+it}log\zeta(3/2+it)$. The figures show that if you take an endpoint T on the the horizonal line that when sliding this point to the right or left will give you, upon integrating, results that will differ by relatively significant amounts due to the oscillatory behavior that does not damp out fast enough. If we take a look at these plots we can see that there seem to be some parts where there's relatively more oscillatory activity and parts where there seems to be a lull in this behaviour. Looking at the following table,

T	$J(100)$	T	$J(100)$	T	$J(100)$	T	$J(100)$
5	26.1244	28	30.4073	41	29.5302	25	28.9354
10	30.1454	29	27.6264	42	28.5406	40	28.0949
15	30.8904	30	29.7240	43	28.1926	55	28.8510
20	30.9276	31	28.7211	44	29.8979	70	29.1323
23	28.8445	32	28.0731	45	27.8594	100	28.8046

we have divided the table in blocks of values that have some connection, the first block is where it begins, next we have two blocks that have oscilatory behaviour based on the plot and confirmed by the values in the table and the last block of values represent points of lull in the erratic behaviour. Guessing that these blocks of oscillatory behaviour integrate more or less to zero so that those parts can be thought of as not contributing to the final result of the integral we see that indeed these lull points are rather near the actual value which is $J(100) = 28.5333$. Although there are values of T for which the integral with that boundary would be close enough to the actual value as to be giving the correct value for $\pi(x)$ when after the whole calculation we round to the nearest integer, there is no way of knowing in advance. Here we not only plotted the integrand but we had the actual value to check against! We can see that for other values of x we get different pictures but with the same global behaviour.

4. *Riemann*'s solution

Riemann himself plugged in another more complex formula for $log\zeta(s)$ to get, after some smart mathematics,

$$J^*(x) = Li(x) - \sum_\rho Li(x^\rho) - log2 + \int_x^\infty \frac{dt}{t(t^2-1)logt}.$$

In which we have that $Li(x)$ is the socalled logarithmic integral defined as :

$$Li(x) = \int_2^x \frac{dt}{logt}.$$

This integral grows with x but slowly and steadily. It equals Gauss's guessed approximation for $\pi(x)$ as we saw (1 on page 4. The sum is over the complex zero's of the Zeta function and it is conditionally convergent, thus needs a specification as to in what order the terms are summed. The correct answer here is that the $\rho's$ get the same order as their absolute values starting with the smallest. This term is responsible for what *Riemann* himself called periodicity and what we have here called oscillatory behaviour. The $log2$ is equal to 0.693147.. and the last term is also small and certainly decreasing in a smooth way towards zero. *Riemann* used the first term as aproximation to $J(x)$ and plugging that in to our Möbius formula (18) on page 35 for $\pi(x)$ gives Riemann's approximating function for $\pi(x)$:

$$R(x) = Li(x) - \frac{1}{2}Li(x^{\frac{1}{2}}) - \frac{1}{3}Li(x^{\frac{1}{3}}) - \frac{1}{5}Li(x^{\frac{1}{5}}) + \frac{1}{6}Li(x^{\frac{1}{6}}) + \cdots$$

or

$$\pi(x) \approx \sum_{n=1}^{\infty} \frac{\mu(n)}{n} Li(x^{\frac{1}{n}}).$$

Based on empirical evidence we have that Riemann's approximation is reasonably better than that of Gauss's at least up to large values of x, but this is not true in general and for very high values of x we can have a reversal of this situation.

CHAPTER 4

An Analytic method for calculating $\pi(x)$

1. The Mellin Transform Pair

In 1987 *Lagarias* and *Odlyzko* [**4**] start with the fomula,

$$J^*(x) = \sum_p \sum_{\substack{n=1 \\ p^n \leq x}}^{\infty\,'} \frac{1}{n} = \frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} \frac{x^s}{s} log\zeta(s) ds$$

where the prime in the sum indicates that in case $x = p^n$ we should add for that specific value of n, not $\frac{1}{n}$ but $\frac{1}{2n}$. This is in accordance with (17) on page 34. They simply note that you can see the truth of this formula by starting with the integral on the right and using the identity :

$$log\zeta(s) = \sum_p \sum_{n=1}^{\infty} \frac{1}{n} p^{-ns}$$

so that :

$$\frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} \frac{x^s}{s} log\zeta(s) ds = \sum_p \sum_{n=1}^{\infty} \frac{1}{n} \left(\frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} \frac{(x/p^n)^s}{s} ds \right).$$

Since we have :

Lemma 5

(21) $$\frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} \frac{(x/n)^s}{s} ds = \begin{cases} 1 & n < x \\ 1/2 & n = x \\ 0 & n > x \end{cases},$$

the formula follows.

Proof of the Lemma :

Consider

$$\lim_{T\to\infty} \frac{1}{2\pi i} \int_{a-iT}^{a+iT} \frac{x^s}{s} ds,$$

with $a > 0$ and $x > 0$.

we distinguish three cases, $0 < x < 1$, $x = 1$ and $x > 1$. Starting with the first case, we have

$$|\frac{1}{2\pi i}\int_{a-iT}^{a+iT}\frac{x^s}{s}ds| \leq \frac{x^a}{\pi T|logx|}.$$

To proof this inequality take $K > a$ and consider the rectangle $a \leq \Re(s) \leq K$, $-T \leq \Im(s) \leq T$. Because $a > 0$, the function x^s/s has no singularity for values in this rectangle. By Cauchy's integral theorem, the integral of x^s/s over the boundary of this rectangle must equal 0. It follows that

$$\frac{1}{2\pi i}\int_{a-iT}^{a+iT}\frac{x^s}{s}ds = -\frac{1}{2\pi i}\int_{a+iT}^{K+iT}\frac{x^s}{s}ds + \frac{1}{2\pi i}\int_{a-iT}^{K-iT}\frac{x^s}{s}ds + \frac{1}{2\pi i}\int_{K-iT}^{K+iT}\frac{x^s}{s}ds.$$

The last integral's modulus is bounded by $\frac{(x^K/K)(2T)}{2\pi}$. For the other two the absolute value is bounded by

$$\frac{1}{2\pi}\int_a^K\frac{x^\sigma}{T}d\sigma = \frac{x^K - x^a}{2\pi T logx}.$$

It follows that

$$|\frac{1}{2\pi i}\int_{a-iT}^{a+iT}\frac{x^s}{s}ds| \leq \frac{|x^K - x^a|}{\pi T|logx|} + \frac{x^K 2T}{2\pi K}.$$

If we first let $K \to \infty$ we get $x^K \to 0$. Letting $T \to \infty$ makes the whole converge to zero.

In case $x = 1$ we find that

$$\lim_{T\to\infty}\frac{1}{2\pi i}\int_{a-iT}^{a+iT}\frac{x^s}{s}ds = \lim_{T\to\infty}\frac{1}{2\pi}\int_{-T}^{T}\frac{dt}{a+it} =$$
$$\lim_{T\to\infty}\left(\frac{1}{2\pi}\int_{-T}^{T}\frac{adt}{a^2+t^2} - i\frac{1}{2\pi}\int_{-T}^{T}\frac{tdt}{a^2+t^2}\right)$$

where we multiplied $\frac{1}{a+it}$ with $\overline{a+it} = a - it$. Also,

$$\lim_{T\to\infty}\int_{-T}^{T}\frac{tdt}{a^2+t^2} = \lim_{T\to\infty}\frac{1}{2}log(a^2+t^2)|_{-T}^{T} = 0.$$

Making the substituation $u = \frac{t}{a}$ we get that the limit of the first integral on the righthand side equals

$$\lim_{T\to\infty}\frac{1}{2\pi}\int_{\frac{-T}{a}}^{\frac{T}{a}}\frac{du}{1+u^2} = \frac{1}{2},$$

as can be seen by using the fact that $\int_{-\infty}^{\infty}(1+u^2)^{-1}du = \pi$.

Lastly for $x > 1$ we consider the integral of $\frac{x^s}{s}$ over a rectangle of the form $-K \leq \Re(s) \leq a$, $-T \leq \Im(s) \leq T$. Since the integrand is analytic in this rectangle and $s = 0$ lies inside the rectangle, Cauchy's integral formula states that the integral around the boundary of the rectangle must equal the residu of the integrand in zero, which equals one, hence,

$$\frac{1}{2\pi i}\int_{a-iT}^{a+iT}\frac{x^s}{s}ds+\frac{1}{2\pi i}\int_{a+iT}^{-K+iT}\frac{x^s}{s}ds+\frac{1}{2\pi i}\int_{-K+iT}^{-K-iT}\frac{x^s}{s}ds+\frac{1}{2\pi i}\int_{-K-iT}^{a-iT}\frac{x^s}{s}ds=1.$$

This implies

$$|\frac{1}{2\pi i}\int_{a-iT}^{a+iT}\frac{x^s}{s}ds-1|\leq\frac{1}{2\pi}\int_{-K}^{a}\frac{x^\sigma}{T}d\sigma+\frac{1}{2\pi}\frac{x^{-K}2T}{K}+\frac{1}{2\pi}\int_{-K}^{a}\frac{x^\sigma}{T}d\sigma$$

$$=\frac{x^a-x^{-K}}{\pi Tlogx}+\frac{x^{-K}T}{\pi K}.$$

Letting $K\to\infty$ eventually gives

$$|\frac{1}{2\pi i}\int_{a-iT}^{a+iT}\frac{x^s}{s}ds-1|\leq\frac{x^a}{\pi Tlogx}$$

which gives that the limit for $T\to\infty$ equals zero. Now because $0<\frac{x}{n}<1\Leftrightarrow n>x$, $\frac{x}{n}=1/2\Leftrightarrow n=x$ and $\frac{x}{n}>1\Leftrightarrow n<x$, the lemma follows.

End of proof.

When we look once more at this integral form of the indicator function (21) on page 41, we actually have

$$c(u)=\frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty}\frac{(x/u)^s}{s}ds=\frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty}\frac{x^s}{s}u^{-s}ds$$

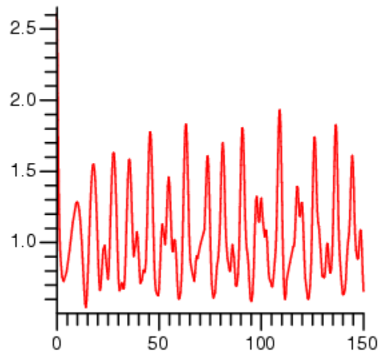
where we have plugged in the real variable u and $c(u)$ is one part of a so-called Mellin transform pair,

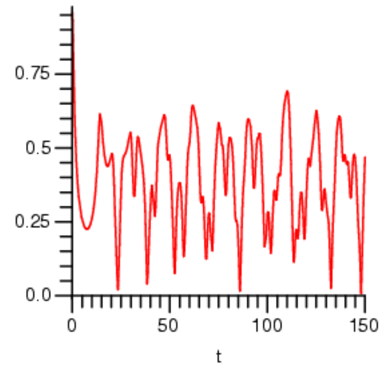
$$c(u)=\frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty}F(s)u^{-s}ds$$

and,

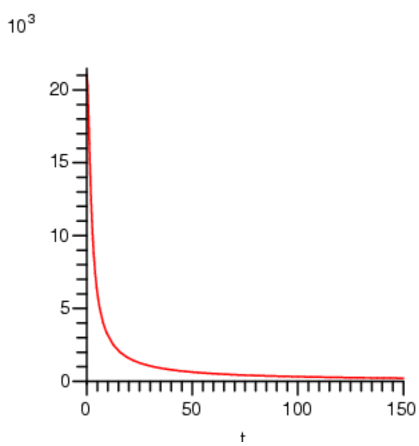
$$F(s)=\int_{0}^{\infty}c(u)u^{s-1}du$$

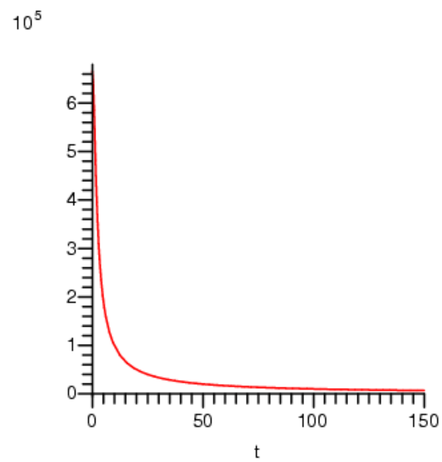
so in our case, $F(s)=x^s/s$ and $c(u)$ is the indicator function (21) above. Now as we've seen in our integral formula for $J(x)$ we have an integrand which is the product of this $F(s)$ and $log\zeta(s)$. The erratic oscillatory behaviour is because of the influence of the Zeta function and the $F(s)$ is decreasing smoothly but not fast enough to damp out this Zeta influence from some x onwards. Taking $a=3/2$, consider some plots to see this more clearly,



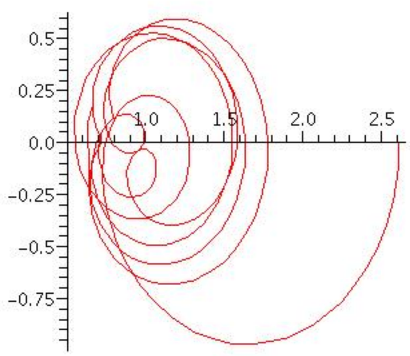


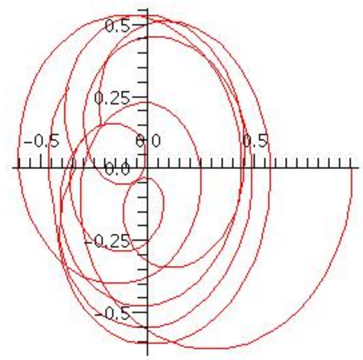
On the left we have a figure of the absolute value of the Zeta function on the vertical line $a = 3/2$. On the right we have a figure of the absolute value of the logarithm of the Zeta function on this line. The important thing to notice is that based on these figures we see the logarithm having the same oscillatory behaviour as the Zeta function and that the absolute value of both functions is small on this range of t-values. Now looking at the absolute value of $\frac{x^{3/2+it}}{3/2+it}$ for values $x = 1000$ and $x = 10000$,



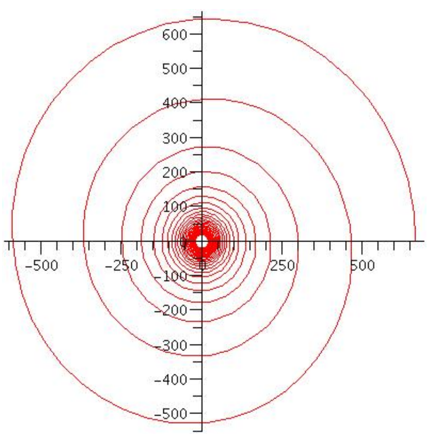


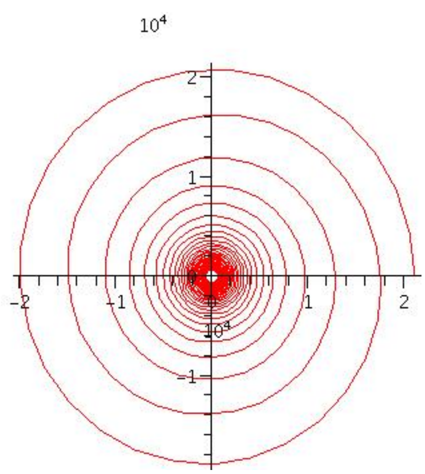
We see no difference in behaviour except the scaling in height which is the result, of course, of the increase of the value of x. The absolute value of integrand (20) on page 36 is a upper- and lower bound on this integrand. The following parametric plots with $a = 3/2$ show that the oscillation of integrand (20) on page 36 between positive and negative values is to be expected :





Here we have the Zeta function on the left and the logarithm of the Zeta function on the right. In both cases the point most to the right on the x-axes corresponds to $t = 0$. We can see for these values of $t \leq 50$ that the absolute value of the functions is indeed bounded by a small real number and that the argument of $\zeta(3/2 + it)$ lies in $(-\pi, \pi]$. Lastly two parametric plots for $x = 100$, $x = 1000$ and $t = 50$ of $\frac{x^{3/2+it}}{3/2+it}$:





2. A New Pair

The idea of *Lagarias* and *Odlyzko* is to construct a new $c(u)$ which deviates from our indicator function (21) on page 41. by a little bit, depending on a new parameter $y \in (1, x)$. Moreover we have $c(u) = 1$ for $0 \leq u \leq x - y$, $c(u) = 0$ for $u \geq x$ and $0 \leq c(u) \leq 1$ for $x - y \leq u \leq x$
$C(u)$ will be constructed in the subinterval $(x - y, x)$ to be smooth, simple to calculate and such that $|F(s)|$ will decrease rapidly for $|\Im(s)| \to \infty$. In this case we

will have that we can choose a T so that integrating up unto this T value will give a good approximation to the correct answer, except that now we will not have that our integral equals the $J^*(x)$ value ! Instead we have, from the general formula,

$$\sum_p \sum_{n=1}^{\infty} \frac{c(p^n)}{n} = \frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} F(s) log\zeta(s) ds$$

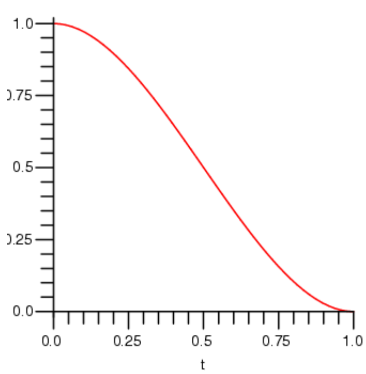
for the general Mellin transform pair $c(u)$ and $F(s)$. aplying this to our new specific $c(u)$ and its corresponding Mellin transform $F(s)$, we get :

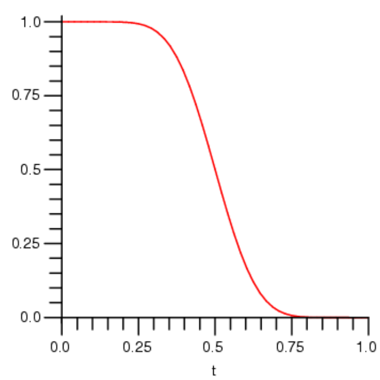
$$J^*(x) - \sum_p \sum_{n=1}^{\infty} \frac{c(p^n)}{n} = \sum_p \sum_{\substack{n=1 \\ x-y<p^n<x}}^{\infty} \frac{1-c(p^n)}{n}.$$

or,

$$J^*(x) = \sum_p \sum_{\substack{n=1 \\ x-y<p^n<x}}^{\infty} \frac{1-c(p^n)}{n} + \frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} F(s) log\zeta(s) ds. \qquad (22)$$

We shall call the sum on the left of (22) $\epsilon_y(x)$ and the integral on the right of (22) $J_y^*(x)$.
The idea for the construction of a new $c(u)$ is to use in the interval $(x-y, x)$ a function which looks like these plots,



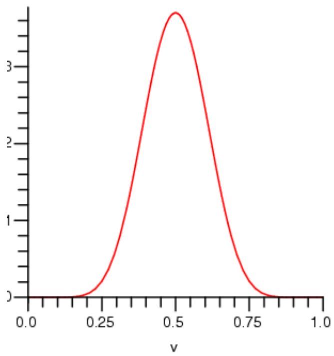


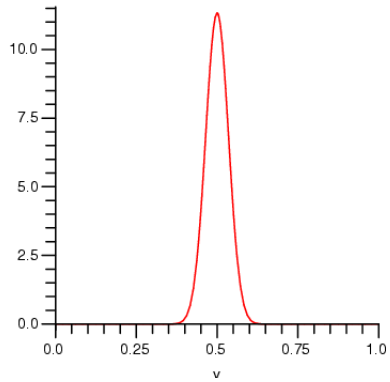
Looking at the plots we see a smooth function which has derivative zero at the endpoints of the interval and it is symetric in the sense that it remains invariant when rotated over an angle of π radians through the vertical line $t = 1/2$. Apart from some scaling parameters we have the following idea. We want the function corresponding to this graph to be a polynomial $G(w)$. Its derivative is then also a polynomial, $g(v)$, with $0 \leq v, w \leq 1$. The point symetry of $G(w)$ implies that $g(1-v) = g(v)$. This in turn implies that we can write $g(v) = h(v - v^2)$ for some polynomial $h(z)$, because $(1-v) - (1-v)^2 = v - v^2$. Furthermore we demand that $g(0) = g(1) = 0$ which is equivalent with $h(0) = 0$. Integrating our polynomial $g(v)$ we get $G(w) = c + \int_1^w g(v)dv$ with c a constant. $G(1) = 0$ gives $c = 0$ and $G(0) = 1$ gives $\int_1^0 g(v)dv = 1$. For a random polynomial $h(z)$ we could thus take $G(w) = \frac{\int_1^w h(v-v^2)dv}{\int_1^0 h(v-v^2)dv}$.

Lagarias and *Odlyzko* start with $h(z) = z^k$ with a positive integer parameter k. They write

$$f_k(v) = \frac{v^k(1-v)^k}{\int_0^1 v^k(1-v)^k} dv$$

which looks like,





where the right graph corresponds to an $f_k(v)$ with greater value of the parameter k than the left graph. With $w = \frac{x-u}{y}$ we get $c(u) = G(\frac{x-u}{y})$, or

$$c(u) = \begin{cases} 0 & (x-u)/y \leq 0 \\ \int_0^{\frac{x-u}{y}} f_k(v)dv & 0 < (x-u)/y < 1 \\ 1 & (x-u)/y \geq 1 \end{cases} \tag{23}$$

3. Calculating $F_{x,y}(s)$

With this new $c(u)$, we obtain the corresponding new $F(s)$:

Lemma 6

$$F_{x,y}(s) = s^{-1} \int_0^1 (x - vy)^s f_k(v) dv. \tag{24}$$

Proof :

Plugging our $c(u)$ into equation 24 we get,

$$F_{x,y}(s) = \int_0^{\infty} \left(\int_0^{\frac{x-u}{y}} f_k(v)dv \right) u^{s-1} du$$

and,

$$\int_0^{\frac{x-u}{y}} f_k(v)dv = \int_0^1 1_{[0,\frac{x-u}{y}]}(v) f_k(v) dv = \int_0^1 1_{[0,x-vy]}(u) f_k(v) dv.$$

So $F_{x,y}(s) =$

$$\int_0^\infty \left(\int_0^1 1_{[0,x-vy]}(u) f_k(v) dv\right) u^{s-1} du = \int_0^1 \left(\int_0^\infty 1_{[0,x-vy]}(u) u^{s-1}\right) f_k(v) dv$$

$$= \int_0^1 \left(\int_0^{x-vy} u^{s-1} du\right) f_k(v) dv = \int_0^1 \frac{(x-vy)^s}{s} f_k(v) dv.$$

End of proof.

We actually have that,

Lemma 7

$$F_{x,y}(s) = (y^k s(s+1)\cdots(s+k))^{-1} \sum_{r=1}^{k+1} \frac{a_r x^{s+k+r} - b_r (x-y)^{s+k+r}}{y^r (s+k+1)\cdots(s+k+r)} \tag{25}$$

with, $a_r = f_k^{(k+r-1)}(0)$ and $b_r = f_k^{(k+r-1)}(1) = (-1)^{k+r-1} a_r$.

Proof :

When we start with the expression (24) for $F_{x,y}(s)$ obtained above and integrate by parts, we get :

$$\int_0^1 \frac{(x-vy)^s}{s} f_k(v) dv = \frac{(x-vy)^{s+1}}{-y(s+1)s} f_k(v)|_0^1 - \int_0^1 \frac{(x-vy)^{s+1}}{-y(s+1)s} f_k^{(1)}(v) dv =$$

$$\int_0^1 \frac{(x-vy)^{s+1}}{y(s+1)s} f_k^{'}(v) dv.$$

Recall that $f_k(v) = c_k v^k (1-v)^k$ is a polynomial with $c_k = \int_0^1 f_k(v) dv$ a constant. In particular this implies $f_k^{(r)}(0) = f_k^{(r)}(1) = 0$ for $0 \le r \le k-1$. Repeating the integration by parts, it follows that

$$F_{x,y}(s) = \frac{1}{y^k s(s+1)\cdots(s+k)} \int_0^1 (x-vy)^{s+k} F_k^{(k)}(v) dv. \tag{26}$$

Again integrating by parts we gives,

$$\int_0^1 (x-vy)^{s+k} f_k^{(k)}(v) dv = \frac{(x-vy)^{s+k+1}}{-y(s+k+1)} f_k^{(k)}|_0^1 + \int_0^1 \frac{(x-vy)^{s+k+1}}{y(s+k+1)} f_k^{(k+1)}(v) dv$$

$$= \frac{f_k^{(k)}(0)x^{s+k+1} - f_k^{(k)}(1)(x-y)^{s+k+1}}{y(s+k+1)}|_0^1 + \int_0^1 \frac{(x-vy)^{s+k+1}}{y(s+k+1)} f_k^{(k+1)}(v)dv$$

after k more steps we get, since $f_k^{(2k+1)} = 0$,

$$\sum_{r=1}^{k+1} \frac{f_k^{k+r-1}(0)x^{s+k+r} - f_k^{k+r-1}(1)(x-y)^{s+k+r}}{y^r(s+k+1)\cdots(s+k+r)}.$$

End of proof.

This proves that $F_{x,y}(s)$ is a meromorphic function with possibly simple poles at negative integers between 0 and $-2k-1$ and possible including 0. With $\Re(s) = a > 1$, which is needed in order for the Euler product formula to work, we have no poles.

4. Bounds and the approximation of $J^*(x)$

We will be approximating our $J^*(x)$ by integrating from $a-iT$ to $a+iT$ instead of from $a-i\infty$ to $a+i\infty$. We will need to show that we can do that with a negligible error and for that we need to bound $|F_{x,y}(s)|$ for which we start with a lemma :

Lemma 8

$$|f_k^{(k)}(v)| \leq c_k(4k)^k$$

where $c_k = (\int_0^1 f_k(v)dv)^{-1}$.

Proof :

By the binomium of *Newton*, we have

$$(v-v^2)^k = \sum_{n=0}^{k} \binom{k}{n} (-1)^n v^{k+n},$$

differentiating k times we get that the k-th derivative equals

$$\sum_{n=0}^{k} \binom{k}{n} (-1)^n v^n (k+n)\cdots(n+1)$$

the factors $(k+n)\cdots(n+1)$ are all bounded by $2k$ and there are k of them, implying that the absolute value of the k-th derivative on $0 \leq v \leq 1$ is bounded by

$$\sum_{n=0}^{k}\binom{k}{n}(2k)^k = (4k)^k.$$

The last equation is true because $\sum_{n=0}^{k}\binom{k}{n} = 2^k$. Lastly because the constant c_k is invariant under differentiation and we obtain the lemma.

End of proof.

Taking expression (26) on page 48 above we find,

$$|F_{x,y}(s)| \leq |\frac{1}{y^k s(s+1)\cdots(s+k)}| \int_0^1 |(x-vy)^{s+k} f_k^{(k)}(v)| dv$$

$$\leq \frac{c_k(4k)^k max_{[0,1]}|(x-vy)^{s+k}|}{|y^k s(s+1)\cdots(s+k)|} = \frac{c_k(4k)^k x^{a+k}}{|y^k s(s+1)\cdots(s+k)|}.$$

Since also,

$$|\frac{1}{y^k s(s+1)\cdots(s+k)}| \leq |\frac{1}{y^k s^{k+1}}|$$

for $\Re(s) = a > 1$, we finally have

$$|F_{x,y}(s)| \leq c_k(4k)^k x^{a+k} y^{-k} |s|^{-k-1}.$$

Combined with the fact that $|log\zeta(s)| = O(1)$ for $\Re(s) > 1$ fixed, we can bound the upper tailsection of our integral $J_y^*(x)$ 22 on page 46 as function of T :

$$|\frac{1}{2\pi i}\int_{a+iT}^{a+i\infty} F_{x,y}(s) log\zeta(s) ds| = O(x^{a+k} y^{-k} T^{-k})$$

where the constant implied by the O-notation depends on k.
The same estimate is valid for

$$|\frac{1}{2\pi i}\int_{a-i\infty}^{a+iT} F_{x,y}(s) log\zeta(s) ds|.$$

Odlyzko and *Lagarias* fix $a = 3/2$. This has to do with error bounds used on the values needed for the Zeta function, which are, at the time of their paper, only known for specific values of $\Re(s) = a > 1$. If we take a look at the precise error bound for integrating only up to T we have that this bounds equals,

$$2dc_k(4k)^k x^{3/2}(\frac{x}{yT})^k. \tag{27}$$

from which we can see that for the error bound to be small we have to have $yT > x$. Here we have a constant d bounding the absolute value of the Zeta function. We see that for fixed values of x, y and T with $yT > x$ we have that increase in k does not need to mean a smaller error bound because of the growth of the term $2dc_k(4k)^k$. Here we note that c_k also increases with k. If all parameters are held fixed, we see that an increase in the value of $y \leq x$ implies a decrease in the error bound. Of course an increase in the value of T while all other parameters stay fixed results in a decrease of the error bound which was to be expected.
One can take T as function of x, y and k and so get a fixed prescribed error bound. *Lagarias* and *Odlyzko* take $T = y^{-1}x^{1+\delta/10}$ with $k = \lfloor 200\delta^{-1} \rfloor + 1$, this implies $\delta \leq \frac{200}{k-1}$ and

$$x^{3/2+k}y^{-k}T^{-k} \leq x^{3/2+k}y^{-k}y^k x^{-k-(\frac{k}{10}\frac{200}{k-1})} = x^{3/2-\frac{20k}{k-1}}.$$

Note that this nescecitates that $k \geq 2$. With k from 2 to ∞ we get that $\frac{k}{k-1}$ converges from above from a maximum of 2 for $k = 2$ to 1 implying that $x^{-\frac{20k}{k-1}} \leq x^{-20}$, so for this choice of T we certainly have an error of $O(x^{-8})$. Notice the introduction of parameter δ. This parameter is important in the whole algorithm and is to be thought of as small, e.g. in $(0, 1)$, and *Lagarias* and *Odlyzko* take $\delta = \frac{1}{loglog(x)}$. Also as of yet we do not need δ to be small. Finally, due to the oscillatory nature of the absolute value of the integrand (20) on page 36 we expect that the bound we have derived is rather generous.

5. New Formula and Parameters

So we have in general as our adapted $J^*(x)$-formula, with $\Re(s) = a > 1$,

$$J^*(x) = \epsilon_{y,k}(x) + J^*_{y,k}(x), \tag{28}$$

with k a positive integer and $y \in (0, x]$, and

$$\epsilon_{y,k}(x) = \sum_p \sum_{\substack{n=1 \\ x-y<p^n<x}}^{\infty} \frac{1 - \int_0^{\frac{x-p^n}{y}} f_k(v)dv}{n}. \tag{29}$$

and

$$J^*_{k,y}(x) = \frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty} \left(\int_0^1 \frac{(x-vy)^s}{s} f_k(v)dv \right) log\zeta(s)ds. \tag{30}$$

In the sum $\epsilon_{y,k}(x)$ one sums over pairs (p, n) such that $x - y < p^n < x$ which implies it is actually a finite sum. So we now have, for x fixed, a by pairs (y, k) parametrised family of Mellin transform pairs, $c(u) = c_{y,k,x}(u)$ (23) on page 47 and $F_{x,y}(s) = F_{y,k,x}(s)$ (25) on page 48, which are at the heart of our formula (28) above. We can consider the parameter y and equivalently $x - y$ as a kind of slide that we can slide from left to right and right to left. The parameter y equals the length of the interval $(x - y, x)$ which contains the prime powers p^n contributing to $\epsilon_{y,k}(x)$. So we see that as the parameter $y = 0$, which is actually precluded based on (23), we have that our error term $\epsilon_{0,k}(x) = 0$, hence,

$$\frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty}\Big(\int_0^1 \frac{x^s}{s} f_k(v)dv\Big)log\zeta(s)ds = \frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty}\frac{x^s}{s}\Big(\int_0^1 f_k(v)dv\Big)log\zeta(s)ds =$$

$$\frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty}\frac{x^s}{s}log\zeta(s)ds,$$

which is the original integral formula $J^*(x)$ 17 on page 34 for $J(x)$. This is obvious since we constructed a $c(u)$ differing from (21) on page 41 only on the interval $(x-y,x)$. If on the other hand we would take $y = x$ we get,

$$J^*(x) = \epsilon_{x,k}(x) + \frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty}\Big(\int_0^1 \frac{(x(1-v))^s}{s} f_k(v)dv\Big)log\zeta(s)ds,$$

which implies, since $J(x) = \sum_{(p,n):p^n\leq x}\frac{1}{n}$, that

$$\sum_p \sum_{\substack{n=1\\ p^n\leq x}}^{\infty} \frac{\int_0^{1-(p^n/x)} f_k(v)dv}{n} = \frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty}\Big(\int_0^1 \frac{(x(1-v))^s}{s} f_k(v)dv\Big)log\zeta(s)ds.$$

Here on the lefthand side we have $J(x)$ weighted by the function $f_k(v)$. On the righthand side we have the Kernel $\int_0^1 \frac{(x(1-v))^s}{s} f_k(v)dv$. When $f_k(v) = 1$ this equals the continiuous mean over the interval $[0,1]$ of our original Kernel $\frac{x^s}{s}$.
The computation of values of $\epsilon_{y,k}(x)$ involves the sieving of the interval $[x-y,x]$ and for every prime power p^n in this interval a calculation of a value of a polynomial of degree $k+1$. We can consider the increase of value of y with fixed value of k as a transport of work from the analytic side to the elementary side. Namely the error in approximating the integral $J^*_{y,k}(x)$ by integrating up to fixed T-value, decreases as y increases as can be seen from bound (27) on page 50. Also increased y value means a larger interval to sieve. Notice however that we do not have for $y = x$, that $\epsilon_{x,k}(x) = J^*(x)$ and $J^*_{x,k}(x) = 0$. On the one hand we want to choose y such that it is small and we have to do as little work as possible on the elementary side. On the other hand, we do not want to make $y = 0$ otherwise we have the original formula (17) on page 34. It is rather difficult to choose a specific value for y based only on this qualitative assessment, we do not know for example what influence the parameters have on the speed of the algorithm.

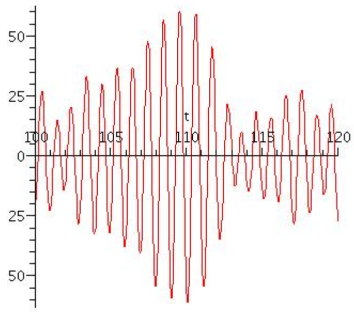
6. Calculations and Figures

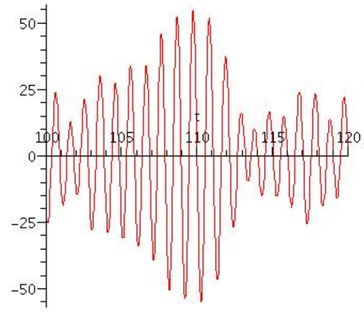
For calculating $\pi(x)$ precisely enough, so that if we round to the nearest integer we get the correct answer, we need multiple values of $J(x)$ by the möbius inversion theorem (19) on page 35. How many values will be needed ? We have $J(x<2) = 0$

which implies that the least n for which values of $J(x^{1/n})$ will be needed is the least n such that $x^{1/n} < 2$. We have $n = \lfloor log_2(x) \rfloor + 1$. So if we need to know $\pi(x)$ to within an error of 0.4 for example, we need for the $J(x)$-values an average precision of $\frac{0.4}{\lfloor log_2(x) \rfloor + 1}$. In practice one will of course have tables of small values of $J(x)$. We have another simple upper bound on the values of $J^*_{y,k}(x)$, namely $J^*(x) - y$, for we can see from (29) on page 51 that $\epsilon_{y,k}(x) \leq y$.
We want to see if we can calculate $\pi(x)$ with this method by using *Mathematica* and *Maple*. So first we need to calculate values of our integral $J^*_{y,k}(x)$ by approximating them by integrating up to T. Using *Mathematica* we have calculated the following tables, where the values in the cells with no specification are the corresponding approximated values of $J^*_{y,k}(x)$. In all cases we have $Re(s) = a = 3/2$ fixed. We briefly note that in *Maple* values of parameter k greater than one results in unworkable time consumption for computation of the integral. We note that we choose x-values which are not prime powers so that there is no difference between the values of $J(x)$ and $J^*(x)$.

x	100	200	500	1000	2000	5000	10000
$\pi(x)$	25	46	95	168	303	669	1229
$J(x)$	28.533	51.095	101.668	176.696	313.629	683.563	1247.098
$\sqrt{x}$	10	14.14	22.36	31.6	44.72	70.71	100
$k=1$, $y=1$, $T=100$	28.527	50.752	100.247	172.201	301.411	723.969	1347.560
$k=1$, $y=10$, $T=100$	27.725	48.783	100.896	173.129	302.066	723.803	1348.75
$k=1$, $y=\sqrt{x}$, $T=100$	27.725	48.3306	100.598	175.887	308.469	712.565	1346.500
$k=5$, $y=1$, $T=100$	28.524	50.754	100.247	172.200	301.410	723.970	1347.560
$k=5$, $y=10$, $T=100$	27.711	48.650	100.913	173.089	302.027	723.829	1348.760
$k=5$, $y=\sqrt{x}$, $T=100$	27.711	48.210	100.990	176.227	308.351	713.643	1348.14

Considering the first table we see little variance in results if we look at a fixed value of x. This could be the result of the fact that $T = 100$. Note that the values of the parameters k, y and T do not produce a small bound on the error using our bound (27) on page 50, so we can not as of yet conclude that the calculations are or are not precise enough. We also see that for increased value of x we have increased absolute value of the difference of the approximated $J^*_{y,k}(x)$-values and the $J(x)$-values . We look at some plots from *Maple* to see if the results seem in accordance with the plots.

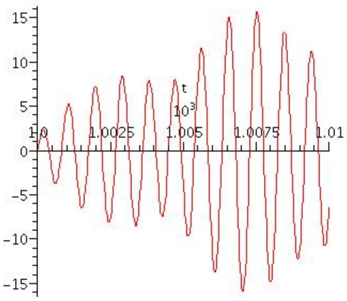


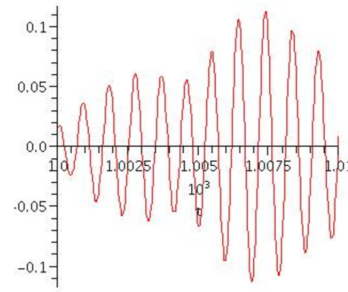


Here we see almost identical figures although the left figure corresponds to parameter values of $x = 1000$, $y = 10$ and $k = 1$. The right figure corresponds to the values $x = 1000$, $y = \sqrt{1000} = 31.6$ and $k = 5$. So the identical looking figures suggest that the seeming invariance of the results in table one, under different parameter values, is confirmed based on this one case. The integrand seems, for the parameter values of the plots above, very symmetric around the T-axis. The absolute value of the integrand in this range however does not seem to imply that the difference between the computations and the true values seen in table one for corresponding parameter values are out of bound. Notice the true value of $J^*_{31.6,5}(1000)$ should be less than $176.696 - 31.6 = 145.096$. We look at a second table of values,

x	100	200	500	1000	2000	5000	10000
$\pi(x)$	25	46	95	168	303	669	1229
$J(x)$	28.533	51.095	101.668	176.696	313.629	683.563	1247.098
$\sqrt{x}$	10	14.14	22.36	31.6	44.72	70.71	100
$k = 1$, $y = 1$, $T = 1000$	1.301	-72.847	523.938	-54.660	-604.951	-1368.560	-24174.1
$k = 1$, $y = 10$, $T = 1000$	28.343	27.755	728.127	-226.470	1329.570	641.424	-11867.2
$k = 1$, $y = \sqrt{x}$, $T = 1000$	28.343	-15.426	35.747	267.889	-1436.92	510.889	-488.096
$k = 5$, $y = 1$, $T = 1000$	-0.228	-75.842	527.406	-55.510	-604.959	-1368.420	-24174.9
$k = 5$, $y = 10$, $T = 1000$	31.667	28.173	781.048	-233.756	1913.320	674.038	-13772.0
$k = 5$, $y = \sqrt{x}$, $T = 1000$	31.667	-2.289	-33.828	345.718	-1410.38	-914.407	-22635.0

Looking at the second table we see more diversity among the values of $J^*_{y,k}(x)$ approximated by integrating up to $T = 1000$. For every fixed x-value in table two we see difference in parameter configurations having more effect on the results. This could be the result of the oscillatory nature of the integrand. We expect in general however that the increase in T-value should be accompanied by a decrease in the absolute value of the difference between the calculated values and the true values of $J^*_{y,k}(x)$. These should not be greater than the values $J(x) - y$. For $x = 2000$ for example we see that the reverse is the case. We look at some plots from Maple, Here on the left we have a figure of $J^*_{10,5}(2000)$ upto $T = 1000$ and on the right we





have a figure of $J_{44.7,5}(2000)$ up to $T = 1000$. We clearly see that the increase in y-value has a decreasing effect on the absolute value of the integrand. However we see that we should not expect the magnitude of error suggested by table two for these parameter values. Notice that the values of x, y, k and T in tables one and two do not allow us to use the formula (27) on page 50 for the bound on the error, because for these values the bound is not small, so that we can not prove that at least some values in the tables are not correct. This would then suggest that we could not rely on the algorithm *Mathematica* uses to give precise enough results with which we could in practice calculate values of $\pi(x)$ using the analytic method of *Lagarias* and *Odlyzko*. We look at a third table,

x	100	200	500	1000	2000	5000	10000
$\pi(x)$	25	46	95	168	303	669	1229
$J(x)$	28.533	51.095	101.668	176.696	313.629	683.563	1247.098
$0.1x$	10	20	50	100	200	500	1000
$k = 20$, $y = 0.1x$, $T = 100$	27.744	47.737	96.749	166.439	295.422	659.497	1193.470
$k = 20$, $y = 0.1x$, $T = 500$	27.399	24.683	148.637	1209.25	3894.96	4142.29	33466.9
$k = 20$, $y = 0.1x$, $T = 1000$	15.379	16.189	-155.808	159.913	8760.88	1846.09	-15312.8
$k = 20$, $y = 0.1x$, $T = 5000$	57.746	112.835	637.9	1273.38	101.778	21031.7	10108.4
$k = 20$, $y = 0.1x$, $T = 10000$	25.23	52.08	143.43	456.664	1661.49	-24264.2	-90269.8
$k = 20$, $y = 0.1x$, $T = 100000$	56.426	-30.7075	2527.12	-23.698	5138.5	18811.4	-14540

In this table we have configurations of parameter values that enable us to prove, using our formula for the bound (27), that the *Mathematica* results are not precise enough. For instance for $x = 500$ and $T = 5000$ we get that the error must be less than 15.3 or more precisely the corresponding approximated value for $J^*_{50,20}(500)$ should be in $(101.668 - 50)^+_-15.3$ or in $[36.368, 66.968]$. Also the value for the upto $T = 100000$ approximated $J^*_{1000,20}(10000)$-value should be in $(1247.098 - 100)^+_-0.0013$ or in $[1147.0967, 1147.0993]$. We used the option $Method \rightarrow Oscillatory$ on *Mathematica* procedure *NIntegrate* to get these results.

7. Euler-Maclaurin summation formula and Complexity

Using a version of the conversion to a real integral (20) on page 36, we have

$$\frac{1}{2\pi i}\int_{3/2-iT}^{3/2+iT} F_{x,y}(s) log\zeta(s) ds = \frac{1}{\pi}\int_0^T h_{3/2}(t)dt,$$

where

$$h_{3/2}(t) = h_{3/2,x,y}(t) = \Re\{F_{x,y}(3/2+it)log\zeta(3/2+it)\}.$$

Lagarias and *Odlyzko* suggest the use of the so-called $Euler-Maclaurin$ summation formula to calculate the real integral (20). This means that we can write, for any positive integer n,

$$\int_0^T g(t)dt = \frac{T}{n}\left(\sum_{j=0}^{n} g(\frac{jT}{n}) - \frac{g(0)+g(T)}{2}\right) \tag{31}$$

$$-\sum_{i=1}^{m-1}\frac{B_{2i}}{(2m)!}(\frac{T}{n})^{2i+1}(g^{(2i-1)}(T) - g^{(2i-1)}(0)) + R_{2m},$$

where the B_{2i} are the Bernoulli numbers and the remainder R_{2m} is bounded by

$$|R_{2m}| \leq n\frac{B_{2m}}{(2m)!}(\frac{T}{n})^{2m+1} max_{0\leq t\leq T}|g^{(2m)}(t)|. \tag{32}$$

Here $g(t)$ is a complex valued function with its first $2m$ derivatives continious on $[0,T]$. Also use is made of the following bound on the derivates of $h_{3/2}(t)$:

$$\frac{d^m}{dt^m}h_{3/2,x,y}(t)| = O(m!(\frac{1}{4})^{-m}x^{1\frac{1}{4}+k}), \tag{33}$$

where $m \geq 1$ and the constant implied by the O-notation depends on k. In our case $g(t) = h_{3/2}(t)$.
We have

Lemma 9

for the the number of steps $n = \lfloor Tx^{\frac{\delta}{10}} \rfloor = \lfloor y^{-1}x^{1+\frac{\delta}{5}} \rfloor$ and number of derivatives $2m = 2k^2$ we get $|R_{2k^2}| = O(x^{-8})$.

Sketch of proof :

Using the bound (32) on $|R_{2m}|$ and the bound (33) on the derivatives of $h_{3/2}(t)$ we obtain

$$|R_{2k^2}| = O(y^{-1}x^{1+\frac{\delta}{5}}x^{-\frac{\delta}{10}(2k^2+1)}x^{1\frac{1}{4}+k}).$$

Here we used that $n \leq Tx^{\frac{\delta}{10}}$ and all terms not involving powers of x are delegated to the constant implied by the O-notation. Because $k = \lfloor 200\delta^{-1} \rfloor + 1$ we get $\delta \leq \frac{200}{k-1}$ and so,

$$|R_{2k^2}| = O(x^{1+\frac{40}{k-1}-\frac{20(2k^2+1)}{k-1}+1\frac{1}{4}+k}) = O(x^{\frac{-39k^2+1\frac{1}{4}+18\frac{3}{4}}{k-1}}).$$

So $k \geq 2$ and assuming that $\frac{-39k^2+1\frac{1}{4}+18\frac{3}{4}}{k-1} \approx 1\frac{1}{4} + \frac{18\frac{3}{4}}{k} - 39k \leq -19$, this is certainly equal to $O(x^{-8})$. The constant implied by the O-notation depends on k.

End of sketch of proof.

If we can also calculate the main term in the $Euler-Maclaurin$ summation formula (31) with error of order $O(x^{-8})$ then we know that we can calculate the integral for $J^*(x)$ (17) on page 34 to within $O(x^{-8})$. If we can calculate the values of $h_{3/2}(t)$ and its derivatives fast, making sure that the main term in (31) has an error of $O(x^{-8})$, we have an efficient analytic algorithm for computing $\pi(x)$.
Assuming we can compute the values of $h_{3/2}(t)$ and its derivatives up to order $2k^2$ with an error of $O(x^{-10})$ fast, we only have to show that the error of the main term of (31) has an error of $O(x^{-8})$.
In the main term of (31) there is a term with only values of $h_{3/2}(t)$,

$$\int_0^T h_{3/2}(t)dt = \frac{T}{n}\left(\sum_{j=0}^{n} h_{3/2}(\frac{jT}{n}) - \frac{h_{3/2}(0)+h_{3/2}(T)}{2}\right)$$

and a term involving values of $\frac{d^m}{dt^m}h_{3/2}(t)$,

$$-\sum_{i=1}^{m-1}\frac{B_{2i}}{(2m)!}(\frac{T}{n})^{2i+1}(h_{3/2}^{(2i-1)}(T) - h_{3/2}^{(2i-1)}(0)).$$

For the first term in (31) we only have to check that $nO(x^{-10}) = O(x^{-8})$. Remember that $n = y^{-1}x^{1+\frac{\delta}{5}} \leq x^{1+\frac{\delta}{5}}$ so $x^{1+\frac{\delta}{5}}x^{-10} = x^{-9+\frac{\delta}{5}}$ and using the fact that $\delta(x) = \frac{1}{loglog(x)}$ which is less than 5 for $x > 1$, we are done. Notice that here we need δ to be small and by definition this means that our k is e.g. greater than 200. These values of k are much larger than what we used in calculations, simply because it becomes unpracticle to calculate without specific tailered numerical integration

algorithms. For example even with our theoretical use of the $Euler-Maclaurin$ formula we assume that we already have the to within prescribed error precomputed values for $h_{3/2}(t)$ and $\frac{d^m}{dt^m}h_{3/2}(t)$. The poor results with our $Mathematica$ computations can probably be ascribed to either parameter-values that are too big and or problems with the approximation of values of the *Zeta*-function and its derivatives.
For the second term involving the derivatives we first state that $m = k^2$ is less than $x^{\frac{\delta}{5}}$ at least for x large enough. Of course by implicit definition of δ, $k = \lfloor \frac{200}{\delta} \rfloor + 1$, k is a growing function of x and thus not a constant but the rate of growth makes it a constant for all practicle purposes. This implies that $m \leq n$, so that if we can proof that $\frac{|B_{2i}|}{(2i)!}x^{\frac{-\delta}{10}(2i-1)}$ with $1 \leq i \leq k^2 - 1$ is no greater than one, we have our error of $O(x^{-8})$ on the main term of the summation formula.

Lemma 10

$$\frac{|B_{2i}|}{(2i)!}x^{(-\delta/10)(2i-11)} \leq 1.$$

for all $i \in [1, k^2 - 1]$. Proof :

$x^{(-\delta/10)(2i-1)})$ is clearly less than one because $\delta > 0$ and $x^{\epsilon} > 1$ for $x > 1$ and any $\epsilon > 0$.
For the factor involving the *Bernoulli* numbers we have the following :

$$B_{2i} = (-1)^{i+1}\frac{2(2i)!}{(2\pi)^{2i}}\left(1 + \frac{1}{2^{2i}} + \frac{1}{3^{2i}} + \frac{1}{4^{2i}} + \cdots\right),$$

with $i \geq 1$. The sequence on the right is bounded by $\zeta(2) = \frac{\pi^2}{6}$. From this we see that $\frac{|B_{2i}|}{(2i)!} \leq 1$, because $\frac{1}{3}\pi^2 < (2\pi)^2$.

End of proof.

The algorithm of *Lagarias* and *Odlyzko* has an elementary part and an analytical part. *Lagarias* and *Odlyzko* state that, using several different sub-algorithms, one can compute the value of $\pi(x)$ to within ± 0.1 using,

$$O(\{x^{1/2} + y + x^{1/5} + T^{3/2-\beta})\}x^{\delta/10}) \tag{34}$$

operations and $O(T^{\beta+\delta})$ space. Here β is a parameter in $[0, 1/2]$ that comes from an algorithm that is needed to compute values of the Zeta function fast. The $x^{1/2} + y$ part comes from the elementary part of the analytic algorithm (29) on page 51 and the part $x^{1/5} + T^{3/2-\beta}$ comes from the analytic part or the integral (30) on page 51. The $x^{\delta/10}$ is the number of operations needed to approximate all values, per specific value, of the integrand (20) on page 36. The part in brackets are the number of operations needed by the elementary- and integral parts of the total algorithm.

Taking $y = x^{(3-2\beta)/(5-2\beta)}$ and remembering that we choose $T = y^{-1}x^{1+\frac{\delta}{10}}$, we get

$$T^{3/2-\beta} = y^{\beta-3/2}x^{(3/2-\beta)(1+\frac{\delta}{10})} = x^{\frac{3-2\beta}{2}(1+\frac{\delta}{10}-\frac{(3-2\beta)}{(5-2\beta)})} =$$

$$x^{\frac{3-2\beta}{2}(\frac{5-2\beta-(3-2\beta)+(5-2\beta)\delta/10}{5-2\beta})} = x^{\frac{3-2\beta}{5-2\beta}(1+\frac{\delta}{4}-\frac{\beta\delta}{10})}.$$

Because $\beta \in [0, 1/2]$ we have

$$\frac{3-2\beta}{5-2\beta}(\frac{\delta}{4}-\frac{\beta\delta}{10}) \leq \frac{3}{5}\frac{\delta}{4} \leq \frac{\delta}{5}$$

hence $T = x^{(3-2\beta)/(5-2\beta)+\delta/5}$ and with the fact that $\beta \leq 1/2$ we can see that this term dominates all the others in (34) above. Incorperating the term $\delta/10 < \delta/5$ gives an algorithm with time complexity of,

$$O(x^{(3-2\beta)/(5-2\beta)+2\delta/5}).$$

A similar calculation shows that the space complexity is,

$$O(x^{2\beta/(5-2\beta)+\delta}).$$

The constant implied in these O-symbols depends on δ. Choosing $\beta = 5b/(2+2b)$ with $b \in [0, 1/4]$ gives an algorithm $A(b, \delta)$ with $O(x^{(3-2b)/5+\delta)})$ time- and $O(x^{b+\delta})$ space complexity. These algorithms can be combined, with δ going to zero sufficiently slowly, to an algorithm $A(b)$ with $O(x^{(3-2b)/5+\epsilon})$ time complexity and $O(x^{b+\epsilon})$ space complexity. *Lagarias* and *Odlyzko* note, based on an analysis which they omit, that $\delta(x) = \frac{1}{loglog(x)}$ satisfies the requirement such that we get the stated complexities.

Notice that for $\beta = 0$ we have $y = x^{3/5} = x^{0.6}$ and with $\beta = 1/2$ we have $y = x^{1/2} = x^{0.5}$. So it is interesting to see that in order for the speed to be more efficient we still need to sieve an interval of length about $\sqrt{x}$. This is of course less than $x^{2/3}$ as we had with the elementary methods but it is interesting to see that a not insignificant part still has to be elementary.

CHAPTER 5

Conclusions and Remarks

In chapters one and two we could see a clear evolution of the ideas used in the elementary or so-called *Meissel-Lehmer* methods for computing values of $\pi(x)$. With hindsight it is clear that starting with the sieve of *Eratosthenes* we get *Legendre*'s formula. By insight of *Meissel*'s this leads to a formula for $\pi(x)$ which uses the first $\pi(x^{1/3})$ primes to sieve the interval $[1, x^{2/3}]$ and then add the hard part, the *Legendre* sum (5). This *Legendre* sum is equivalent to the sum of the leaves of the binary recursion tree for the partial sieving function $\phi(x, a)$. Using this equivalence and close analysis of the structure of the recursion tree for $\phi(x, a)$, reveals that the *Legendre* sum can be more efficiently computed using truncation rules on the growth of the tree. These insights lead to the methods of *Lagarias*, *Miller* and *Odlyzko* and of *Deglise* and *Rivat*. As can be seen from the time and space complexities the methods from chapter two are a real improvement over those of chapter one.

Is it possible to find even faster elementary ways for computing $\pi(x)$? It might not be possible to reduce the power $x^{2/3}$ in the time complexities of elementary methods. We saw that even with a more detailed analysis of possible forms of node pairs (n, p_{b+1}) of the recursion tree for $\phi(x, a)$ (7) on page 12, *Deglise* and *Rivat* were not able to get the power of x in the time complexity down from 2/3. Of course it is difficult to see what could be improved but this may not be because of exhaustion of the elementary methods.

What can be seen by the analytic algorithm of *Lagarias* and *Odlyzko* is that a faster way to compute $\pi(x)$ exists in theory. The complex contour integral representation for the prime counting function $J(x)$ (17) on page 34 can not be used as approximation to $J(x)$ using numerical integration up to some value of T. The adaptation of *Lagarias* and *Odlyzko* seems to work in theory but is based on lots of error estimates that would have to be made precise before implementation. Also it is based on other algorithms used to calculate for instance the values of $\zeta(s)$ fast. All of these algorithms would also have to be implemented.

The *Mathematica* computations show that good control over all approximated values used in the algorithm is essential. This is allways true yet the highly irregular nature of the Zeta function makes it especially difficult in this algorithm. The current implementation of numerical integration algorithms used in the software packages *Maple* and *Mathematica* do not allow us to compute the integrals, needed in the analytic algorithm, with sufficient accuracy.

The analysis of *Lagarias* and *Odlyzko* suggests that we still need to sieve an interval of length about $\sqrt{x}$ if we want a choice of parameters resulting in a fast analytic algorithm as presented here. Of course if an ingenious way to calculate the integral for the prime power counting function $J(x)$ (17) could be found, then this could result in an even faster algorithm.

Bibliography

[1] Edwards, H.M., *Riemann's Zeta Function*. Dover Publications, New York, 2001.

[2] Deglise, M., and Rivat, J., *Computing* $\pi(x)$*: The Meissel, Lehmer, Lagarias, Miller, Odlyzko Method*. Math. of Comp., **213**, 235-245 (1996).

[3] Lagarias, J.C., Miller, V.S., and Odlyzko, A.M., *Computing* $\pi(x)$*: The Meissel-Lehmer Method*. Math. of Comp., **44**,537-560 (1985).

[4] Lagarias, J.C., and Odlyzko, A.M., *Computing* $\pi(x)$*: An Analytic Method*. Journal of Algorithms, **8** (1987).

[5] Riesel, H., *Prime Numbers and Computer Methods for Factorization*. Birkhäuser, Boston, 1994.