GENERALIZED PRIME NUMBER THEOREM
ON SEMI-GROUPS OF INTEGERS
AND
THE MÖBIUS FUNCTION

by

Haydar Göral

A Thesis Submitted to the

Graduate School of Sciences and Engineering

in Partial Fulfillment of the Requirements for

the Degree of Master of Science

in

Mathematics

Koç University

August 2011

Koç University

Graduate School of Sciences and Engineering

This is to certify that I have examined this copy of a master's thesis

by

Haydar Göral

and have found that it is complete and satisfactory in all respects,

and that any and all revisions required by the final

examining committee have been made.

Thesis Committee Members:

Assoc. Prof. Emre Alkan (Advisor)..........................

Prof.Dr. İlhan İkeda ....

Assoc. Prof. Barış Coşkunüzer.................................

Date: August 2011

Abstract

In this study, we first prove the classical Prime Number Theorem which gives an estimate on the number of primes not exceeding x where x is a given real number. Then, in the third chapter we prove the Wiener-Ikehara Tauberian Theorem and as a result of this theorem, we deduce the Prime Number Theorem just from the non-vanishing of the Riemann Zeta function on the line $\sigma = 1$. In chapter four, we prove Beurling's Generalized Prime Number Theorem on semi-groups of integers and we investigate the boundary condition of this theorem. Also, we consider the partial sums of the Möbius function over such semi-groups and we show the difference between the Generalized Prime Number Theorem and the partial sums of the Möbius function over semi-groups. Based on this difference, in the last part (which is a joint work with my supervisor Assoc. Prof. Emre Alkan) we give quantitative estimates on partial sums of the Möbius function over semi-groups that are also in a given arithmetic progression. Lastly, we apply our results to the fractions.

ÖZET

Bu çalışmada ilk olarak herhangi bir x reel sayısına kadar olan asalların sayısı üzerine sonuç veren Asal Sayı Teoremi kanıtlanacaktır. Daha sonra üçüncü bölümde Wiener-Ikehara Tauberian Teoremini kanıtlayıp, bunun sonucunda Asal Sayı Teoremini kanıtlamak için Riemann Zeta fonksiyonunun $\sigma = 1$ doğrusu üzerinde hiç sıfırının olmamasının yeterli olacağını göstereceğiz. Dördüncü bölümde, Beurling'in genelleştirilmiş Asal Sayı Teoremini tamsayıların sadece çarpma altında kapalı olan kümeleri üzerine kanıtlayacağız ve bu Teoremin sınırlarını araştıracağız. Ayrıca Möbius fonksiyonunun kısmi toplamlarını tamsayıların bu tür alt kümeleri üzerinde düşüneceğiz ve genelleştirilmiş Asal Sayı Teoremi ile Möbius fonksiyonunun kısmi toplamları arasındaki farkı göstereceğiz. Son bölümde de (bu bölüm danışmanım Emre Alkan ile yaptığımız ortak bir çalışmadır.) bu farklılığa dayalı olarak, aynı zamanda tamsayıların sadece çarpma altında kapalı olan bir kümesi ve aritmetik dizi üzerindeki Möbius fonksiyonunun kısmi toplamlarına niceliksel üst sınırlar vereceğiz. Son olarak da, elde ettiğimiz sonuçları kesirlere uygulayacağız.

Acknowledgement

My first and the deepest gratitude is to my advisor Assoc. Prof. Emre Alkan, for his support and tolerance for me. The courses we made together were very fruitful and I learned how to approach to Number Theory from him. Thanks to him, I can add new ideas to my thesis. In short, I owe to him very much.

I would like to thank Prof. Ali Nesin, Prof. Oleg Belegradek and Assoc. Prof. Selçuk Demir, for their support, tolerance and guidance during my education. Their efforts and the time they spent for me are invaluable.

I would like to thank Prof. İlhan İkeda and Assoc. Prof. Barış Coşkunüzer and for their participation in my thesis committee.

I also would like to express my thankfulness to Prof. Ali Ülger, the director of the graduate studies of Mathematics Department.

I thank Burçin Güneş for her help and support during typing my thesis and editing the misprints.

Finally I would like to thank my mother, during my thesis her presence has made the life more bearable.

LIST OF SYMBOLS/ABBREVIATIONS

$\mu(n)$	The Möbius function.
$\varphi(n)$	The Euler's totient function.
$\Lambda(n)$	The Von-Mangoldt function.
$\psi(x)$	$\sum_{n\leq x}\Lambda(n)$; the Chebyshev's ψ function.
$\vartheta(x)$	$\sum_{p\leq x}\log p$; the Chebyshev's ϑ function.
$\pi(x)$	The number of primes $\leq x$.
$\pi(x;q,a)$	The number of primes $\leq x$ which are $\equiv a \pmod q$.
$\psi(x;q,a)$	The sum of $\Lambda(n)$ over $n\leq x$ which are $\equiv a \pmod q$.
$M(x)$	$\sum_{n\leq x}\mu(n)$; the Mertens function.
$\chi(n)$	A Dirichlet character.
$\chi_0(n)$	The principal character.
$L(s,\chi)$	A Dirichlet L-function.
$\zeta(s)$	The Riemann zeta function.
$Li(x)$	$\int_2^x \frac{du}{\log u}$; the logarithmic integral.
$\Re s$	The real part of the complex number s.
$\Im s$	The imaginary part of the complex number s.
$\Gamma(s)$	The Gamma function.
γ	The Euler-Gamma constant.
$\widehat{f}(x)$	$\int_{-\infty}^{\infty} f(t)e^{-2\pi itx}dt$; the Fourier transform of f.
$L^1(\mathbb{R})$	The space of all Lebesgue integrable functions.
$[x]$	The integer part of x.
$\{x\}$	$x-[x]$; the fractional part of x.
$f(x)=O(g(x))$	$\lvert f(x)\rvert\leq Cg(x)$ where C is an absolute constant.
$f(x)=o(g(x))$	$\lim_{x\to\infty} f(x)/g(x)=0$.
$f(x)\ll g(x)$	$f(x)=O(g(x))$.
$f(x)\sim g(x)$	$\lim_{x\to\infty} f(x)/g(x)=1$.
PNT	The Prime Number Theorem.

Contents

1 Preliminaries

This chapter includes the basic information needed to understand the text as we frequently will refer in the following chapters. It consists of five main sections and in each of them, we will present some functions and their properties that we are going to deal with. We also will introduce some theorems and tools that are widely used in Analytic Number Theory and Analysis. All these will be given briefly, without proof since detailed arguments can be found in [7], [26], [27], [30] or [32].

Throughout the text, $\log x$ denotes the natural logarithm and p always denotes a prime number.

1.1 Arithmetic Functions

Definition 1.1. *A complex-valued function defined on the positive integers is called an arithmetic function.*

Now we introduce some arithmetic functions which play an important role on the distribution of prime numbers.

1. Define $I(1) = 1$ and $I(n) = 0$ if $n > 1$.

2. Another aritmetic funciton is u which is defined by $u(n) = 1$ for all $n \geq 1$.

3. The Möbius function μ is defined as follows:

   $\mu(1) = 1$;

   If $n > 1$, write $n = p_1^{a_1} \cdots p_k^{a_k}$. Then

$$\mu(n) = \begin{cases} (-1)^k & \text{if } a_1 = a_2 = \cdots a_k = 1, \\ 0 & \text{otherwise .} \end{cases}$$

4. If $n \geq 1$ the Euler totient $\varphi(n)$ is defined to be the number of positive integers not exceeding n which are relatively prime to n; i.e.,

$$\varphi(n) = \sum_{\substack{m=1 \\ (m,n)=1}}^{n} 1.$$

5. The Von-Mangoldt function $\Lambda(n)$ is defined as:

$$\Lambda(n) = \begin{cases} \log p & \text{if } n = p^m \text{ for some prime } p \text{ and some integer } m \geq 1, \\ 0 & \text{otherwise.} \end{cases}$$

This function seems intricate at a first sight, but as we progress we see that this function is natural and deeply related to the distribution of the primes and we deal with this function throughout this thesis.

Definition 1.2. *Given two arithmetic functions, we define their Dirichlet product as*

$$(f * g)(n) = \sum_{d|n} f(d)g(n/d).$$

It can be shown that $I(n)$ is the identity function of the operation $*$ and $\mu * u = I$.

Definition 1.3. *An arithmetic function f is called multiplicative if*

$$f(mn) = f(m)f(n)$$

whenever $(m, n) = 1$.

f is called completely multiplicative if $f(mn) = f(m)f(n)$ holds for all integers m and n.

Now let $f(n)$ be an arithmetic function. We usually denote by $F(x)$, the summatory function of $f(n)$

$$F(x) = \sum_{n \leq x} f(n).$$

Now we give summatory functions of some important arithmetic functions.

Definition 1.4. *Given $x \geq 0$ define $\pi(x) = \sum_{p \leq x} 1$ and $M(x) = \sum_{n \leq x} \mu(n)$. The function $\pi(x)$ is called the prime counting function and the function $M(x)$ is called the Mertens function.*

In the next chapter we find the asymptotic behavior of the function $\pi(x)$ without an error term. More precisely in Chapter 2, we prove the Prime Number Theorem (PNT) which says that $\lim_{x \to \infty} \frac{\pi(x)}{x/\log x} = 1$. In Chapter 3, we investigate the connection between these two functions and in Chapter 4 we generalize this connection and we also show how they differ from each other. The main idea of this thesis is to study these two functions.

In analytic number theory, we estimate the summatory function $F(x)$ of arithmetic functions because they are expected to behave more regularly whereas an arithmetic function may behave very irregularly even on consecutive integers. So we are interested in tools for evaluating the averages. One of them is Abel's summation formula which is sometimes called the partial summation. We use Abel's summation frequently throughout the text.

Theorem 1.5 (Abel's Summation Formula-The Partial Summation Formula)**.** *Let x and y be real numbers with $0 < y < x$. Let $a(n)$ be an arithmetic function with summatory function $A(x)$ and $f(t)$ be a function with a continuous derivative on $[y, x]$. Suppose also that $A(x) = 0$ when $x < 1$. Then,*

$$\sum_{y<n\leq x} a(n)f(n) = A(x)f(x) - A(y)f(y) - \int_y^x A(t)f'(t)dt. \tag{1.1}$$

In particular, if $x > 1$ and $f(t)$ is continuously differentiable on $[1, x]$, then

$$\sum_{n\leq x} a(n)f(n) = A(x)f(x) - \int_1^x A(t)f'(t)dt. \tag{1.2}$$

This theorem, applied to the functions $a(n) = 1$ and $f(t) = 1/t$ gives

$$\sum_{n\leq x} \frac{1}{n} = \log x + \gamma + r(x) \quad \text{where} \quad |r(x)| < \frac{1}{x}. \tag{1.3}$$

The number γ in (1.3) is called the *Euler-Gamma constant.*

1.2 Elementary Results in the Distribution of Primes

Let us first introduce Chebyshev's functions $\psi(x)$ and $\vartheta(x)$ which have a key role in the study of distribution of primes.

Definition 1.6. *We define Chebyshev's $\psi(x)$ function to be the summatory function of $\Lambda(n)$ by*

$$\psi(x) = \sum_{n\leq x} \Lambda(n).$$

Definition 1.7. *We define Chebyshev's $\vartheta(x)$ function by*

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

where p runs over primes $\leq x$.

Chebyshev has showed that the functions $\psi(x)$ and $\vartheta(x)$ are of order $O(x)$ and their relation gives equivalent forms of the PNT. More precesiley, he has proved the following two theorems:

Theorem 1.8. *There exists $x_0 \in \mathbb{R}$ such that, for all $x \geq x_0$ we have*

$$(0.92)x \leq \vartheta(x) \leq \psi(x) \leq (1.06)x. \tag{1.4}$$

The following theorem states three equivalent forms of PNT (without error term) and its proof is simply based on Abel's summation.

Theorem 1.9. *The following relations are equivalent:*

$$\pi(x) \sim \frac{x}{\log x}. \tag{1.5}$$

$$\vartheta(x) \sim x. \tag{1.6}$$

$$\psi(x) \sim x. \tag{1.7}$$

1.3 Dirichlet Series

Given an arithmetic function $f(n)$, we define the Dirichlet series associated by f as

$$F(s) = \sum_{n=1}^{\infty} \frac{f(n)}{n^s}.$$

A Dirichlet series can be regarded as a function of the complex variable s, defined in the region in which the series converges. We write the variable s as

$$s = \sigma + it, \quad \text{where} \quad \sigma = \Re s, t = \Im s,$$

and we will use this notation throughout the text. Every Dirichlet series has an absciss of convergence σ_c, which means there is a half-plane that the series converges for $\sigma > \sigma_c$. Also for every Dirichlet series there is a number σ_a such that, for $\sigma > \sigma_a$ the series converges absolutely. Moreover, a Dirichlet series constitutes an analytic function in its half-plane of convergence. There is a close relation between the summatory function and the Dirichlet series of an arithmetic function; and we will be considering this in the next chapter.

Another important result about Dirichlet series is the Euler Product Identity when applied to the Dirichlet series.

Theorem 1.10 (Euler Product Identity)**.** *Let f be a multiplicative arithmetic function with Dirichlet series $F(s) = \sum_{n=1}^{\infty} \frac{f(n)}{n^s}$. Assume $F(s)$ converges absolutely for $\sigma > \sigma_a$, then we have*

$$F(s) = \prod_p \left(1 + \frac{f(p)}{p^s} + \frac{f(p^2)}{p^{2s}} + \dots\right) \quad \text{for } \sigma > \sigma_a. \tag{1.8}$$

If f is completely multiplicative, then

$$F(s) = \prod_p \left(1 - \frac{f(p)}{p^s}\right)^{-1} \quad \text{for } \sigma > \sigma_a. \tag{1.9}$$

The most famous Dirichlet series is the one associated with the function $u(n)$, so-called the Riemann zeta function,

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}. \qquad (\sigma > 1)$$

By the Euler product identity (1.9), we have

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_p \left(1 - \frac{1}{p^s}\right)^{-1} \quad (\sigma > 1). \tag{1.10}$$

Logarithmic derivative of the identity (1.10) gives that

$$-\frac{\zeta'(s)}{\zeta(s)} = \sum_p \sum_{n=1}^{\infty} \frac{\log p}{p^{ms}} = \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s} \quad (\sigma > 1). \tag{1.11}$$

It can be shown that if $F(s)$ and $G(s)$ are the Dirichlet series of the arithmetic functions $f(n)$ and $g(n)$ respectively and if $F(s)$ and $G(s)$ converge absolutely in $\sigma > \sigma_a$ then the Dirichlet series of $(f * g)(n)$ is $F(s)G(s)$ which is absolutely convergent in $\sigma > \sigma_a$.

Using the above fact, we can show that

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

when $\sigma > 1$. This and (1.10) both imply that $\zeta(s)$ has no zero in $\sigma > 1$.

1.4 Dirichlet Characters and L-functions

Definition 1.11. *An arithmetic function $\chi(n)$ is called a Dirichlet character modulo q if it satisfies*
(i) $\chi(n) = 0$ for $(n, q) > 1$,
(ii) $\chi(1) = 1$,
(iii) $\chi(n)\chi(m) = \chi(nm)$ for all integers m, n,
(iv) $\chi(n) = \chi(m)$ whenever $n \equiv m \pmod q$, i.e. $\chi(n)$ is q-periodic.

Since $\chi(1) = 1$ and $n^{\varphi(q)} \equiv 1 \pmod q$ when $(n, q) = 1$, $\chi(n)$ must be a $(\varphi(q))$-th root of unity for $(n, q) = 1$. Also, there are $\varphi(q)$ characters to the modulus q. One of them takes the value 1 for all integers relatively prime to q and 0 otherwise, this is called the principal character and denoted by $\chi_0(n)$.

A character $\chi(n)$ modulo q satisfies the following relations:

$$\frac{1}{\varphi(q)} \sum_{n (\mathrm{mod}\ q)} \chi(n) = \begin{cases} 1 & \text{if } \chi = \chi_0, \\ 0 & \text{otherwise.} \end{cases} \tag{1.12}$$

and

$$\frac{1}{\varphi(q)} \sum_{\chi (\mathrm{mod}\ q)} \chi(n) = \begin{cases} 1 & \text{if } n \equiv 1 \pmod q, \\ 0 & \text{otherwise.} \end{cases} \tag{1.13}$$

From the relation (1.13), it is possible to deduce a relation which will be useful when we aim at working on integers belonging to a certain residue class modulo q. If $(a, q) = 1$, then for any n we have

$$\frac{1}{\varphi(q)} \sum_{\chi (\mathrm{mod}\ q)} \bar{\chi}(a)\chi(n) = \begin{cases} 1 & \text{if } n \equiv a \pmod q, \\ 0 & \text{otherwise.} \end{cases} \tag{1.14}$$

This enables us how to catch numbers belonging to a certain residue class modulo q.

Dirichlet also defined L-functions denoted by $L(s,\chi)$ to be the Dirichlet series of $\chi(n)$ for $\sigma > 1$,

$$L(s,\chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s}.$$

All $L(s,\chi)$ converges absolutely in $\sigma > 1$. Moreover if χ is a non-principal character modulo q, then $L(s,\chi)$ converges conditionally in $\sigma > 0$.

By the Euler product identity we have

$$L(s,\chi) = \prod_p \left(1 - \frac{\chi(p)}{p^s}\right)^{-1} \quad (\sigma > 1). \tag{1.15}$$

As for the function $\zeta(s)$, logarithmic differentiation gives that

$$-\frac{L'(s,\chi)}{L(s,\chi)} = \sum_p \sum_{n=1}^{\infty} \frac{\chi(p)\log p}{p^{ms}} = \sum_{n=1}^{\infty} \frac{\Lambda(n)\chi(n)}{n^s} \quad (\sigma > 1). \tag{1.16}$$

1.5 Results from Fourier Analysis

In this subsection, we state some results from Fourier Analysis without proof. The details can be found in [27], [30] or [32].

Theorem 1.12. *Suppose that $T \geq 1$. Let*

$$\Delta_T(x) = T\left(\frac{\sin \pi T x}{\pi T x}\right)^2 \qquad \text{and} \qquad J_T(x) = \frac{3T}{4}\left(\frac{\sin \pi T x/2}{\pi T x/2}\right)^4$$

be the Fejer and Jackson kernels respectively. Then these functions have a peak of height $\asymp T$ and width $\asymp 1/T$ at 0, and have a total mass 1.

Definition 1.13. *Suppose $f \in L^1 = L^1(\mathbb{R})$. Then the function $\widehat{f}(x) = \int_{-\infty}^{\infty} f(t)e(-tx)dt$ is called the Fourier transform of f where $e(x) = e^{2\pi i x}$.*

Theorem 1.14 (Fourier Inversion Theorem)**.** *If $f \in L^1$ and $\widehat{f} \in L^1$, and if $g(x) = \int_{-\infty}^{\infty} \widehat{f}(t)e(tx)dt$, then $f(x) = g(x)$ almost everywhere.*

Theorem 1.15 (Riemann-Lebesgue Lemma)**.** *Suppose $f \in L^1(\mathbb{R})$. Then the Fourier transform of f vanishes at infinity, in other words*

$$\lim_{|x|\to\infty} \widehat{f}(x) = 0.$$

2 The Prime Number Theorem

Recall that $\pi(x)$ = number of primes $p \leq x = |\{2 \leq p \leq x : \text{ p is prime}\}|$. The infinitude of primes was first proved by Euclid. Then Euler also gives another proof of this by showing that the series

$$\sum_{p \leq x} \frac{1}{p}$$

diverges. In the proof, Euler used some analysis and thus this result can be seen as the birth of Analytic Number Theory. Around 1792, Gauss conjectured that a good approximation to $\pi(x)$ is the logarithmic integral

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

Gauss made this observation by finding all primes up to 3.000.000. Almost at the same time, Legendre conjectured that $\pi(x)$ is approximately

$$\frac{x}{\log x - A}$$

where $A = 1.08...$ is some constant. Later it was proved that if

$$\pi(x) = \frac{x}{\log x - A(x)}$$

then $\lim_{x \to \infty} A(x) = 1$. After about 100 years later in 1896, PNT was proved by Jacques Hadamard and de la Vallée Poussin independently. In this chapter, our aim is to prove so-called the Prime Number Theorem that is

$$\pi(x) \sim \frac{x}{\log x}$$

i.e.

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

In fact we will prove an equivalent form of this:

$$\psi(x) = \sum_{n \leq x} \Lambda(n) \sim x \qquad i.e. \qquad \lim_{x \to \infty} \frac{\psi(x)}{x} = 1.$$

There is a close relation between the summatory function of $\Lambda(n)$ $\psi(x) = \sum_{n \le x} \Lambda(n)$ and its Dirichlet series $\sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s} = -\frac{\zeta'(s)}{\zeta(s)}$, which will be given in the Fundamental Formula (2.7). So the Dirichlet series of $\Lambda(n)$ is related to $\zeta(s)$. This is why we prefer $\psi(x)$ and not $\pi(x)$, in other words the Dirichlet series of $\Lambda(n)$ is more familiar than the Dirichlet series of the characteristic function of prime numbers.
Thus the idea is to study the function

$$\sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s} = -\frac{\zeta'(s)}{\zeta(s)}$$

where $\sigma > 1$. $\zeta(s)$ occurs in the denominator, so we have to be careful about its zeros! Therefore we have to analyse $\zeta(s)$ and its zeros as best we can.
By Abel's summation, we have

$$\sum_{k=1}^{n} \frac{\Lambda(k)}{k^s} = \frac{\psi(n)}{n^s} + s \int_1^n \frac{\psi(x)}{x^{s+1}} dx.$$

Since $\frac{\psi(n)}{n^s} \to 0$ as $n \to \infty$ for $\sigma > 1$, we get

$$\sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s} = s \int_1^{\infty} \frac{\psi(x)}{x^{s+1}} dx \qquad i.e. \qquad -\frac{\zeta'(s)}{\zeta(s)} = s \int_1^{\infty} \frac{\psi(x)}{x^{s+1}} dx \tag{2.1}$$

and this is called the Mellin transform of $\psi(x)$.

Our goal is to express $\psi_1(x) = \int_0^x \psi(u) du$ in terms of $-\frac{\zeta'(s)}{\zeta(s)}$ and then we pass to $\psi(x)$ from $\psi_1(x)$ so as to understand the distribution of primes.

2.1 Riemann Zeta Function

Now we study $\zeta(s)$ further since this function is deeply connected with the distribution of primes.

Theorem 2.1 (Analytic continuation of $\zeta(s)$)**.** $\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$ *defined for* $\sigma > 1$ *has an analytic continuation to a function defined on the half plane* $\sigma > 0$*, and that is analytic in this plane exception of a simple pole at* $s = 1$

with residue 1, given by

$$\zeta(s) = \frac{s}{s-1} - s\int_1^\infty \frac{\{x\}}{x^{s+1}}dx.$$

Proof. By Abel's Summation formula ,

$$\sum_{n\le x}\frac{1}{n^s} = \frac{[x]}{x^s} + s\int_1^x \frac{[t]}{t^{s+1}}dt.$$

Therefore, for $\sigma > 1$, we have

$$\begin{aligned}\zeta(s) &= s\int_1^\infty \frac{[x]}{x^{s+1}}dx = s\int_1^\infty \frac{x-\{x\}}{x^{s+1}}dx \\ &= s\int_1^\infty x^{-s}dx - s\int_1^\infty \frac{\{x\}}{x^{s+1}}dx \\ &= \frac{s}{s-1} - s\int_1^\infty \frac{\{x\}}{x^{s+1}}dx.\end{aligned}$$

Given $\epsilon > 0$,

$$\left|\int_1^\infty \frac{\{x\}}{x^{s+1}}dx\right| \le \int_1^\infty \frac{1}{x^{\sigma+1}}dx \le \int_1^\infty \frac{1}{x^{\epsilon+1}}dx = \frac{1}{\epsilon} \qquad \text{when } \sigma \ge \epsilon.$$

Therefore the integral converges absolutely and uniformly in the half plane $\sigma \ge \epsilon$, and represents an analytic function of s for $\sigma \ge \epsilon$. Since $\sigma > 0$ is arbitrary, this function is analytic in $\sigma > 0$. Also, $\frac{s}{s-1} = 1 + \frac{1}{s-1}$ gives a simple pole at $s = 1$ with residue 1. ■

Riemann Hypothesis(RH): We know that $\zeta(s) \neq 0$ for $\sigma > 1$. RH says that if $\zeta(s) = 0$ in $\sigma > 0$ then $\Re(s) = \frac{1}{2}$.

Note that

$$\zeta(s) = 0 \Leftrightarrow \int_1^\infty \frac{\{x\}}{x^{s+1}}dx = \frac{1}{s-1}.$$

Next we bound on the functions $\zeta(s)$ and $\zeta'(s)$ at infinity This will be useful in the proof on the PNT.

Theorem 2.2. *We have*

(i) $|\zeta(s)| \le A\log t$ *(*$\sigma \ge 1,\ t \ge 2$*),*

(ii) $|\zeta'(s)| \le A\log^2 t$ *(*$\sigma \ge 1,\ t \ge 2$*),*

(iii) $|\zeta(s)| \le B(\delta)t^{1-\delta}$ *(*$\sigma \ge \delta,\ t \ge 2,\ 0 < \delta < 1$*)*

where A is an absolute constant and $B(\delta)$ is a constant that may depend on δ.

Proof. By Abel's summation, we obtain

$$\sum_{n\le x}\frac{1}{n^s}=s\int_1^x\frac{[t]}{t^{s+1}}dt+\frac{[x]}{x^s}=\frac{s}{s-1}-\frac{s}{(s-1)x^{s-1}}-s\int_1^x\frac{\{t\}}{t^{s+1}}dt+\frac{1}{x^{s-1}}-\frac{\{x\}}{x^s}.$$

Since

$$\zeta(s)=\frac{s}{s-1}-s\int_1^\infty\frac{\{x\}}{x^{s+1}}dx$$

for $\sigma>0$, we get:

$$\zeta(s)-\sum_{n\le x}\frac{1}{n^s}=\frac{s}{(s-1)x^{s-1}}-\frac{\{x\}}{x^s}-s\int_1^\infty\frac{\{t\}}{t^{s+1}}dt.$$

Hence for $\sigma>0$, $t\ge 1$, $x\ge 1$ we have

$$|\zeta(s)|\le\sum_{n\le x}\frac{1}{n^\sigma}+\frac{1}{tx^{\sigma-1}}+\frac{1}{x^\sigma}+|s|\int_x^\infty\frac{dt}{t^{\sigma+1}}\le\sum_{n\le x}\frac{1}{n^\sigma}+\frac{1}{tx^{\sigma-1}}+\frac{1}{x^\sigma}+\left(1+\frac{t}{\sigma}\right)\frac{1}{x^\sigma}$$

since $|s|<\sigma+t$.

If $\sigma\ge 1$,

$$|\zeta(s)|\le\sum_{n\le x}\frac{1}{n}+\frac{1}{t}+\frac{1}{x}+\frac{1+t}{x}\le(\log x+1)+3+\frac{t}{x},$$

since $t\ge 1$, $x\ge 1$. Taking $x=t$ we obtain (i).

If $\sigma\ge\delta$ where $0<\delta<1$,

$$|\zeta(s)|\le\sum_{n\le x}\frac{1}{n^\delta}+\frac{1}{tx^{\delta-1}}+\left(2+\frac{t}{\delta}\right)\frac{1}{x^\delta}<\frac{x^{1-\delta}}{1-\delta}+x^{1-\delta}+\frac{3t}{\delta x^\delta},$$

since $\sum_{n\le x}\frac{1}{n^\delta}\le\frac{x^{1-\delta}}{1-\delta}$. Taking $x=t$ as before we get

$$|\zeta(s)|\le t^{1-\delta}\underbrace{\left(\frac{1}{1-\delta}+1+\frac{3}{\delta}\right)}_{B(\delta)} \tag{2.2}$$

and this proves (iii).

Lastly we prove (ii). Let $s_0 = \sigma_0 + it_0$ be any point in the region $\sigma \geq 1$, $t \geq 2$ and C a circle with center s_0 and radius $\rho < \frac{1}{2}$. Then by Cauchy's integral formula for $\zeta'(s_0)$, we have

$$|\zeta'(s_0)| = \left| \frac{1}{2\pi i} \int_C \frac{\zeta(s) ds}{(s - s_0)^2} \right| \leq \frac{M}{\rho}$$

where M is the maximum of $|\zeta(s)|$ on C. Now, for all points s on C, we have $\sigma \geq \sigma_0 - \rho \geq 1 - \rho$ and $1 < t < 2t_0$, and hence by (2.2), $(\rho = 1 - \delta)$

$$M \leq (2t_0)^\rho \left(\frac{1}{\rho} + 1 + \frac{3}{1 - \rho} \right) < \frac{10 {t_o}^\rho}{\rho}$$

since $\rho < 1 - \rho < 1$ and $2^\rho < 2$. Hence

$$|\zeta'(s_0)| < \frac{10 {t_o}^\rho}{\rho^2}.$$

Now take

$$\rho = \frac{1}{2 + \log t_0}$$

i.e. ${t_0}^\rho = e^{\rho \log t_0} < e$ and $|\zeta'(s_0)| < 10e(2 + \log t_0)^2$. This implies (ii). ■

2.2 Zeros of $\zeta(s)$

Zeros of $\zeta(s)$ play an important role for the distribution of primes. The following result is the most significant result for the PNT.

Theorem 2.3. *$\zeta(s)$ has no zeros on the line $\sigma = 1$. Furthermore, there is an absolute constant $A > 0$ such that $\frac{1}{\zeta(s)} = O((\log t)^A)$ uniformly for $\sigma \geq 1$, as $t \to \infty$.*

Proof. First observe that

$$3 + 4 \cos \theta + \cos 2\theta = 2(1 + \cos \theta)^2 \geq 0. \tag{2.3}$$

For $\sigma > 1$,

$$\log |\zeta(s)| = \Re(\log(s)) = \Re\left(\sum_p \sum_{m=1}^{\infty} \frac{1}{mp^{ms}} \right)$$
$$= \Re\left(\sum_{n=2}^{\infty} \frac{c_n}{n^{\sigma+it}} \right) = \sum_{n=2}^{\infty} c_n n^{-\sigma} \cos(t \log n)$$

where $c_n = \frac{1}{m}$ if $n = p^m$ for some p prime and 0 otherwise. Hence by (2.3)

$$\log|\zeta^3(\sigma)\zeta^4(\sigma+it)\zeta(\sigma+2it)| = \sum c_n n^{-\sigma}(3+4\cos(t\log n)+\cos(2t\log n)) \geq 0$$

since $c_n \geq 0$. Therefore we have

$$((\sigma-1)\zeta(\sigma))^3\left|\frac{\zeta(\sigma+it)}{\sigma-1}\right|^4|\zeta(\sigma+2it)| \geq \frac{1}{\sigma-1}, \qquad \sigma \geq 1. \tag{2.4}$$

This shows that $1+it$ ($t \neq 0$) cannot be a zero of $\zeta(s)$. Otherwise the left hand side tends to $|\zeta'(1+it)|^4|\zeta(1+2it)|$ but the right hand side tends to infinity as $\sigma \to 1^+$.

For the second part we may assume $1 \leq \sigma \leq 2$ since for $\sigma \geq 2$,

$$\left|\frac{1}{\zeta(s)}\right| = \left|\sum_{n=1}^{\infty}\frac{\mu(n)}{n^s}\right| \leq \sum_{n=1}^{\infty}\frac{1}{n^\sigma} \leq \sum_{n=1}^{\infty}\frac{1}{n^2} = \zeta(2) = \frac{\pi^2}{6}.$$

If $1 < \sigma \leq 2$ and $t \geq 2$ then by (2.4) and Theorem 2.2 (i) we have

$$\begin{aligned}(\sigma-1)^3 &\leq ((\sigma-1)\zeta(\sigma))^3|\zeta(\sigma+it)|^4|\zeta(\sigma+2it)| \\ &\leq A_1^3|\zeta(\sigma+it)|^4A_2\log 2t \\ &\leq A_1^3|\zeta(\sigma+it)|^4A_2 2\log t.\end{aligned}$$

for some constants A_1 and A_2. Thus

$$|\zeta(\sigma+it)| \geq \frac{(\sigma-1)^{3/4}}{A_3(\log t)^{1/4}} \tag{2.5}$$

($1 \leq \sigma \leq 2$, $t \geq 2$) for some constant A_3. Now let $1 < \delta < 2$. For $1 \leq \sigma \leq 2$, $t \geq 2$, by Theorem 2.2 (*ii*) we see that

$$|\zeta(\sigma+it)-\zeta(\delta+it)| = \left|\int_\sigma^\delta \zeta'(u+it)du\right| \leq A_4\log^2 t(\delta-1) \tag{2.6}$$

for some constant A_4. Hence combining (2.5) and (2.6) we get

$$\begin{aligned}|\zeta(\sigma+it)| &\geq |\zeta(\delta+it)| - A_4(\delta-1)\log^2 t \\ &\geq \frac{(\delta-1)^{3/4}}{A_3(\log t)^{1/4}} - A_4(\delta-1)(\log t)^2.\end{aligned}$$

Now, let $\delta = 1 + (2A_3A_4)^{-4}(\log t)^{-9}$ i.e. $\frac{(\delta-1)^{3/4}}{A_3(\log t)^{1/4}} = 2A_4(\delta-1)(\log t)^2$ (when t is large enough) to obtain

$$|\zeta(\sigma+it)| \geq A_4(\delta-1)\log^2 t = A_5(\log t)^{-7},$$

if $1 \leq \sigma \leq 2$ and $t > t_0$ (i.e. when t is large enough) and A_5 is some absolute constant. So we can take $A = 7$.

■

2.3 Fundamental Formula

Working with $\psi(x) = \sum_{n \leq x} \Lambda(n)$ has some convergence problems. So we will work with $\psi_1(x) = \int_0^x \psi(u)du$ instead of $\psi(x)$. This is called the smoothing argument. By Abel's summation, we have

$$\psi_1(x) = \int_0^x \psi(u)du = \sum_{n \leq x} \Lambda(n)(x-n).$$

Since we will prove $\psi(x) \sim x$, we expect $\psi_1(x) \sim \frac{x^2}{2}$. Transition from $\psi_1(x)$ to $\psi(x)$ will be easy. Our aim is to show that the fundamental formula:

$$\psi_1(x) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \frac{x^{s+1}}{s(s+1)}\left(-\frac{\zeta'(s)}{\zeta(s)}\right)ds \qquad \text{when } x > 0, c > 1$$

where the path of integration is the straight line $\sigma = c$.

This formula is significant since it enables us to pass from discrete sum to a continuous sum. Therefore we can use the tools of analysis to determine the asymptotic behavior of a discrete sum. This is the philosophy of Analytic Number Theory.

First we require a lemma in order to obtain the Fundamental Formula.

Lemma 2.4. *For $c > 0$ and $y > 0$, we have*

$$\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \frac{y^s}{s(s+1)} = \begin{cases} 0 & \text{if } y \leq 1, \\ 1 - \frac{1}{y} & \text{if } y \geq 1. \end{cases}$$

Proof. Note that the integral is absolutely convergent since the integrand has modulus less than $y^c|t|^{-2}$ on the line of integration. Denote by J the infinite integral and J_T the integral from $c-iT$ to $c+iT$ (with the factor $\frac{1}{2\pi i}$).

We apply Cauchy's Residue theorem. We replace the line of integration J_T by an arc of the circle c having its centre at $s = 0$ and passing the points $c \pm iT$. If $y \geq 1$, we use the arc c_1 which lies to the left of the line $\sigma = c$, assuming T is large, $R > 2$ where R is the radius of the centre. This gives $J_T = S + J(c_1)$ where S is the sum of residues at $s = 0, -1$ and $J(c_1)$ is the integral along c_1. Now, on c_1 we have $\sigma \leq c$ and thus $|y^s| \leq y^c$ since $y \geq 1$. Moreover $|s|$ and $|s+1| \geq R - 1 > \frac{R}{2}$. This gives

$$|J(c_1)| < \frac{1}{2\pi} \frac{y^c}{(\frac{R}{2})^2} 2\pi R = \frac{4y^c}{R} < \frac{4y^c}{T}.$$

Thus $J_T \to S$ as $T \to \infty$. i.e. $J = S$. But

$$S = Res\left[\frac{y^s}{s(s+1)}, 0\right] + Res\left[\frac{y^s}{s(s+1)}, -1\right] = 1 + \frac{-1}{y} = 1 - \frac{1}{y}.$$

The proof in the case $y \leq 1$ is similar: take c_2 right-hand arc of c and no poles are passed over.

■

Fundamental Formula: For $x > 0, c > 1$, we have

$$\psi_1(x) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{x^{s+1}}{s(s+1)} \left(-\frac{\zeta'(s)}{\zeta(s)} \right) ds. \tag{2.7}$$

Proof. For $x > 0$, by Lemma 15 taking $y = x/n$ we have

$$\frac{\psi_1(x)}{x} = \sum_{n \leq x} \left(1 - \frac{n}{x}\right) \Lambda(n) = \sum_{n=1}^{\infty} \frac{\Lambda(n)}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{(x/n)^s}{s(s+1)} ds.$$

If $c > 1$, the order of summation and integration can be interchanged since

$$\sum_{n=1}^{\infty} \int_{c-i\infty}^{c+i\infty} \left| \frac{\Lambda(n)(x/n)^s}{s(s+1)} \right| ds < x^c \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^c} \int_{-\infty}^{\infty} \frac{dt}{c^2 + t^2} < \infty.$$

Hence

$$\psi_1(x) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{x^{s+1}}{s(s+1)} \left(\sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s} \right) ds,$$

i.e.

$$\psi_1(x) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{x^{s+1}}{s(s+1)} \left(-\frac{\zeta'(s)}{\zeta(s)} \right) ds.$$

■

2.4 Prime Number Theorem

Now we have all tools so as to prove the PNT. Philosophy of the proof is: We will shift the path of integration to the left side of the line $\sigma = 1$ in the Fundamental Formula. This is useful because when shifting the integral we catch the residue at $s = 1$ which contributes to the main term and the power x^s will be small in the left side of the line $\sigma = 1$. Also we need bounds on the function $\zeta(s)$ at infinity and since we deal with $\zeta(s)$, the Theorem 2.3 will be vital.

Theorem 2.5. *We have $\psi_1(x) \sim \frac{x^2}{2}$ when $x \to \infty$.*

Proof. From now on, we assume $x > 1$. Note that the function $-\dfrac{\zeta'(s)}{\zeta(s)}$ has a simple pole at $s = 1$ with residue 1. We will take $c = 1 + \frac{1}{\log x} > 1$ in the Fundamental Formula (2.7) . From (2.7) we know that

$$\frac{\psi_1(x)}{x^2} = \int_{c-i\infty}^{c+i\infty} g(s)x^{s-1}ds$$

where

$$g(s) = \frac{1}{2\pi i}\frac{1}{s(s+1)}\left(-\frac{\zeta'(s)}{\zeta(s)}\right).$$

Moreover $g(s)$ is analytic in $\sigma \geq 1$ except $s = 1$ with residue $\dfrac{1}{4\pi i}$. Also by Theorem 2.2, we have

$$|g(s)| < A_1|t|^{-2}A_2(\log|t|)^2A_3(\log|t|)^7 < |t|^{-3/2}$$

$(\sigma \geq 1, |t| \geq t_0)$.

Let $\epsilon > 0$ be given.

Now we replace the path of integration in (2.7) which is a vertical line by $L = \bigcup_{i=1}^{5} L_i$ where

$$\begin{aligned}
L_1 &= (c - i\infty, c - iT],\\
L_2 &= [c - iT, \alpha - iT],\\
L_3 &= [\alpha - iT, \alpha + iT],\\
L_4 &= [\alpha - iT, c + iT],\\
L_5 &= [c + iT, c + i\infty).
\end{aligned}$$

Choose $T = T(\epsilon)$ and $\alpha = \alpha(\epsilon)$ $(0 < \alpha < 1)$ such that

$$\int_T^\infty |g(c+it)|dt < \frac{\epsilon}{2e} \qquad\qquad (|x^{s-1}| = |x^{c-1}| = x^{\frac{1}{\log x}} = e)$$

and the rectangle $\alpha \le \sigma \le 1$, $-T \le t \le T$ contains no zeros of $\zeta(s)$. This is possible since $\zeta(s)$ has no zero on the line $\sigma = 1$ and such a rectangle can contain at most finitely many zeros of $\zeta(s)$ because otherwise, zeros of $\zeta(s)$ accumulate and $\zeta(s)$ would be zero. By Cauchy's Residue theorem we obtain that

$$\frac{\psi_1(x)}{x^2} = \frac{1}{2} + \int_L g(s)x^{s-1}ds = \frac{1}{2} + J.$$

($1/2$ arises from the pole at $\sigma = 1$).

Note that the integral $\int_L g(s)x^{s-1}ds$ is absolutely convergent.

Write

$$\int_L g(s)x^{s-1}ds = J_1 + J_2 + J_3 + J_4 + J_5,$$

where $J_1, \ldots, J_5$ are the integrals along $L_1, \ldots, L_5$, respectively.

Since $g(\bar{s})x^{\bar{s}-1} = \overline{g(s)x^{s-1}}$, we have

$$|J_1| = |J_5| < \frac{\epsilon}{2e}e = \frac{\epsilon}{2}.$$

Also, if we let M be the maximum of $|g(s)|$ on the finite segments L_2, L_3, L_4 then (since $x > 1$)

$$\begin{aligned}|J_2| = |J_4| &= \left|\int_\alpha^c g(\sigma+it)x^{\sigma+it-1}\right|d\sigma \le M\int_\alpha^c x^{\sigma-1}d\sigma \\ &= M\frac{x^\sigma - 1}{\log x}\bigg|_{\sigma=\alpha}^c = \frac{Me}{\log x} - \frac{Mx^{\alpha-1}}{\log x} \\ &= \left|\int_{\alpha-iT}^{\alpha+iT} g(s)x^{s-1}ds\right| \le Mx^{\alpha-1}2T.\end{aligned}$$

Therefore

$$\left|\frac{\psi_1(x)}{x^2} - \frac{1}{2}\right| < \epsilon + 2\left(\frac{Me}{\log x} - \frac{M}{(\log x)x^{1-\alpha}}\right) + Mx^{\alpha-1}2T.$$

Now choose $x_0 = x_0(\epsilon, T, \alpha, M) = x_0(\epsilon)$ such that if $x \ge x_0$ then

$$\left|\frac{\psi_1(x)}{x^2} - \frac{1}{2}\right| < 3\epsilon.$$

This proves the theorem.

■

Transition from $\psi_1(x)$ to $\psi(x)$

Let $0 < \alpha < 1 < \beta$. Since $\psi(x)$ is monotone increasing, we obtain that

$$\psi(x) \leq \frac{1}{\beta x - x}\int_x^{\beta x} \psi(t)dt = \frac{\psi_1(\beta x) - \psi_1(x)}{(\beta - 1)x}.$$

Therefore

$$\frac{\psi(x)}{x} \leq \frac{1}{\beta - 1}\left(\frac{\psi_1(\beta x)}{(\beta x)^2}\beta^2 - \frac{\psi_1(x)}{x^2}\right).$$

Letting $x \to \infty$ and keeping β fixed we have, since $\dfrac{\psi_1(y)}{y^2} \to \dfrac{1}{2}$ as $y \to \infty$,

$$\limsup_{x\to\infty} \frac{\psi(x)}{x} \leq \frac{1}{2}\frac{\beta^2 - 1}{\beta - 1} = \frac{\beta + 1}{2}.$$

Similarly, $\liminf\limits_{x\to\infty} \dfrac{\psi(x)}{x} \geq \dfrac{\alpha + 1}{2}$. Since α, β are arbitrary, we get

$$\frac{\psi(x)}{x} \to 1 \text{ as x } \to \infty.$$

So we obtain (ultimately!)

Prime Number Theorem (PNT): $\pi(x) \sim \dfrac{x}{\log x}$.

Proof. We know that $\psi_1(x) \sim \frac{x^2}{2}$ and so $\psi(x) \sim x$. Now by Theorem 1.9, we get $\pi(x) \sim \frac{x}{\log x}$ as desired. ■

2.5 PNT and $\zeta(s)$ on the line $\sigma = 1$

The proof of the PNT is based on the fact that $\zeta(s)$ has no zeros on the line $\sigma = 1$. Now, we will show the converse. Suppose we have the PNT. Then by (2.1) we have

$$\int_1^{\infty} \frac{\psi(x) - x}{x^{s+1}}dx = -\frac{1}{s}\frac{\zeta'(s)}{\zeta(s)} - \frac{1}{s - 1} = \alpha(s).$$

Let $\epsilon > 0$ be given. Then by PNT $|\psi(x) - x| < \epsilon x$ for $x \geq x_0 = x_0(\epsilon)$. So for $\sigma > 1$,

$$|\alpha(s)| < \int_1^{x_0} \frac{|\psi(x) - x|}{x^2} dx + \int_{x_0}^{\infty} \frac{\epsilon}{x^\sigma} dx < K + \frac{\epsilon}{\sigma - 1}.$$

Thus $|(\sigma-1)\alpha(\sigma+it)| < K(\sigma-1)+\epsilon < 2\epsilon$ for $1 < \sigma \leq \sigma_0 = \sigma_0(\epsilon, K) = \sigma_0(\epsilon)$. Hence for any fixed t, $(\sigma - 1)\alpha(\sigma + it) \to 0$ as $\sigma \to 1^+$. This shows that the point $1 + it$ cannot be a zero of $\zeta(s)$ because otherwise, $(\sigma - 1)\alpha(\sigma + it)$ would tend to a limit different from 0, namely the residue of $\alpha(s)$ at the simple pole $1 + it$.

NOTE: In the next section we will show PNT $\Leftrightarrow \zeta(1+it) \neq 0$ for all $t \in \mathbb{R}$. We already showed $\Leftarrow$. For the converse, we need some Fourier Analysis and the Tauberian Theorem of Wiener and Ikehara.

2.6 Further Works in PNT and Riemann's Memoir

Now we will state some results without proof. All the details can be found in [12] or [26] .

The functional equation of $\zeta(s)$:

Riemann made the greatest contribution to the study of distribution of primes with his memoir in 1859. In his paper he showed that

- The function $\zeta(s)$ satisfies the functional equation

$$\pi^{-\frac{s}{2}} \Gamma\left(\frac{s}{2}\right) \zeta(s) = \pi^{-\frac{1-s}{2}} \Gamma\left(\frac{1-s}{2}\right) \zeta(1-s), \tag{2.8}$$

  where $\Gamma(s)$ is the so-called Gamma function which is analytic in the half plane $\sigma > 0$ with the integal representation $\Gamma(s) = \int_0^{\infty} e^{-x} x^{s-1} dx$.

- $\zeta(s)$ can be continued analytically over the whole plane and $\zeta(s)$ is meromorphic with the simple pole at $s = 1$ with residue 1.

The second can be deduced from the functional equation regarding the properties of the $\Gamma(s)$ function.

To use the functional equation effectively, we define the function

$$\xi(s) = \frac{1}{2} s(s-1) \pi^{-\frac{s}{2}} \Gamma\left(\frac{s}{2}\right) \zeta(s). \tag{2.9}$$

The function $\xi(s)$ is an entire function since it has no poles for $\sigma \geq \frac{1}{2}$ and satisfies $\xi(s) = \xi(1-s)$. Moreover $\xi(s)$ has the product representation

$$\xi(s) = e^{A+Bs} \prod_{\rho} \left(1 - \frac{s}{\rho}\right) e^{s/\rho} \tag{2.10}$$

where A and B are constants and ρ runs through the zeros of $\zeta(s)$ in the *critical strip* $0 < \sigma < 1$. This was proved by Hadamard and lead to improvements in enlarging the zero-free region of $\zeta(s)$

$$\sigma \geq 1 - \frac{c}{\log t}, \tag{2.11}$$

$|t| \geq 2$, which was previously shown as $\sigma \geq 1$.

Improved zero-free region: For $|t| \geq 2$ there exists a positive number c such that $\zeta(s)$ has no zeros in the region

$$\sigma > 1 - \frac{c}{\log t}.$$

Best zero-free region: For $|t| > e^e$ we have $\zeta(s) \neq 0$ for

$$\sigma > 1 - \frac{A}{(\log |t|)^{2/3} (\log \log |t|)^{1/3}}$$

where $A > 1/100$ is an absolute constant. This result was proved independently by Vinogradov [33] and Korobov [23] in 1958.

Explicit Formulas: Explicit formulas give the relation between the summatory functions related to primes and the zeros of $\zeta(s)$.

By the Fundamental Formula, we know that

$$\psi_1(x) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{x^{s+1}}{s(s+1)} \left(- \frac{\zeta'(s)}{\zeta(s)} \right) ds$$

when $x > 0$, $c > 1$. In fact a similar formula exists for $\psi(x)$, for $x \notin \mathbb{Z}$ and $c > 1$,

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

But working with $\psi(x)$ has some convergence problems since the denominator in the integrand is $1/s$ and the integrand is not absolutely convergent in the vertical line of integration.

If we define $\psi_0(x) = \frac{1}{2}\left(\sum_{n\le x}\Lambda(n) + \sum_{n<x}\Lambda(n)\right)$, then we have the two explicit formulas :

$$\psi_0(x) = x - \sum_{\substack{\rho:\text{ nontrival}\\ \text{zero of }\zeta}} \frac{x^\rho}{\rho} - \frac{\zeta'(0)}{\zeta(0)} - \frac{1}{2}\log(1-x^{-2})$$

$$\psi_1(x) = \frac{x^2}{2} - \sum_{\rho}\frac{x^{\rho+1}}{\rho(\rho+1)} - x\frac{\zeta'(0)}{\zeta(0)} + \frac{\zeta'(-1)}{\zeta(-1)} - \sum_{r=1}^{\infty}\frac{x^{1-2r}}{2r(2r-1)}.$$

Moreover, it is known that

$$\sum_{\rho}\frac{1}{|\rho|}$$

diverges but for any $\epsilon > 0$

$$\sum_{\rho}\frac{1}{|\rho|^{1+\epsilon}} < \infty$$

where the summations are taken over non-trivial zeros of $\zeta(s)$.

Improved PNT (de la Vall*é*e Poussin) If we use the improved zero-free region in the proof of PNT, we can obtain

$$\psi(x) = x + O(xe^{-c\sqrt{\log x}}).$$

and

$$\pi(x) = Li(x) + O(xe^{-c_1\sqrt{\log x}}).$$

Riemann Hypothesis (RH):

After the PNT had been proved, the main problem has become obtaining the PNT with an error term as good as possible. Riemann, in his paper in 1859 has conjectured the Riemann Hypothesis which states that all non-trivial zeros of the Riemann zeta-function have real part 1/2. As the error term is related to the zero-free region of $\zeta(s)$, the Riemann hypothesis is equivalent to the both following two form of PNT: $\psi(x) = x + O(\sqrt{x}\log^2 x)$ and $\pi(x) = Li(x) + O(\sqrt{x}\log x)$.

Moreover, $O(\sqrt{x})$ is the best possible error term since it is known that $\zeta(s)$ has a zero on the critical line $\sigma = 1/2$. Unfortunately this problem is still wide open and we are very far from getting what is conjectured. The last progress about the error term has been made by Vinogradov and Korobov in 1958. They have enlarged the zero-free region to $\sigma > 1 - \dfrac{A}{(\log|t|)^{2/3}(\log\log|t|)^{1/3}}$ which resulted in the error term

$$\psi(x) = x + O\bigg(x \exp\bigg(- C\frac{(\log x)^{3/5}}{(\log\log x)^{1/5}}\bigg)\bigg).$$

Another important property of Dirichlet series is that we can relate them to the summatory functions of arithmetic functions. Now let $f(n)$ be an arithmetic function with summatory function $F(x)$. Let $\alpha(s)$ be the Dirichlet series of $f(n)$ with finite abscissa of convergence σ_c. Then we have two versatile theorems that allow us to pass from discrete to continuous and from continuous to discrete. These theorems generalize the Mellin transform of $\psi(x)$ and the Fundamental Formula.

Theorem 2.6 (Mellin Transform Representation of Dirichlet Series)**.**

$$\alpha(s) = s\int_1^\infty F(x)x^{-s-1}dx \quad \sigma > \max(0, \sigma_c). \tag{2.12}$$

Theorem 2.7 (Perron's Formula)**.** *For any* $c > \max(0, \sigma_a)$*, we have*

$$\sideset{}{'}\sum_{n\le x} f(n) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \alpha(s)\frac{x^s}{s}ds \tag{2.13}$$

and

$$F_1(x) = \int_0^x F(u)du = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \alpha(s)\frac{x^{s+1}}{s(s+1)}ds \quad . \tag{2.14}$$

Here, $\sum'$ *indicates that we take the term* $f(x)$ *to be halved in the case when* x *is an integer and the improper integral* $\int_{c-i\infty}^{c+i\infty}$ *is to be interpreted as the symmetric limit* $\lim_{T\to\infty}\int_{c-iT}^{c+iT}$.

Now we state the more generalized version of Mellin transform and Perron's formula.

Theorem 2.8 (Cesàro Weights)**.** *For a positive integer* k*, put*

$$C_k(x) = \frac{1}{k!}\sum_{n\le x} f(n)(x-n)^k.$$

Then $C_k(x) = \int_0^x C_{k-1}(u)du$ *for* $k \geq 1$ *and* $C_0(x) = F(x)$. *Moreover, for* $\sigma > \max(0, \sigma_c)$, *we have*

$$\alpha(s) = s(s+1)...(s+k) \int_1^\infty C_k(x) x^{-s-k-1} dx$$

and for $c > \max(0, \sigma_a)$ *and* $x > 0$, *we have*

$$C_k(x) = \int_{c-\infty}^{c+\infty} \alpha(s) \frac{x^{s+k}}{s(s+1)...(s+k)} ds.$$

Now we state Plancherel Identity which concerns the space $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$. This theorem will be useful in Chapter 4.

Theorem 2.9 (Plancherel Identity). *Suppose that* $\int_0^\infty |\omega(x)| x^{-\sigma-1} dx < \infty$, *and also that* $\int_0^\infty |\omega(x)|^2 x^{-2\sigma-1} dx < \infty$. *Put* $K(s) = \int_0^\infty \omega(x) x^{-s-1} dx$. *Then we have*

$$2\pi \int_0^\infty |\omega(x)|^2 x^{-2\sigma-1} dx = \int_{-\infty}^\infty |K(\sigma+it)|^2 dt. \tag{2.15}$$

One of the significiant application of this identiy is the following. Suppose $f(n)$ is an arithmetic function with the summatory function $F(x)$. Let $\alpha(s)$ be the Dirichlet series of $f(n)$ with a finite abscissa of convergence σ_c. Then we have

$$2\pi \int_0^\infty |F(x)|^2 x^{-2\sigma-1} dx = \int_{-\infty}^\infty \left| \frac{\alpha(\sigma+it)}{\sigma+it} \right|^2 dt$$

for $\sigma > \max(0, \sigma_c)$.

Perron's formula itself is not enough to satisfy an error term for the summatory functions. To estimate the error term for the summatory function of $f(n)$, we need the following result.

Theorem 2.10 (Truncated Perron's Formula). *For any* $c > \max(0, \sigma_a)$, $T > 0$ *and non-integral* x, *we have*

$$\sum_{n \leq x} f(n) = \frac{1}{2\pi i} \int_{c-iT}^{c+iT} \alpha(s) \frac{x^s}{s} ds + R(T), \tag{2.16}$$

where $R(T) \le \dfrac{x^c}{T} \sum_{n=1}^{\infty} \dfrac{|f(n)|}{n^c |\log(x/n)|}$.

PNT for Arithmetic Progression: For a and q positive integers with $(a,q)=1$, define

$$\psi(x,q,a) = \sum_{\substack{n\le x \\ n\equiv a \pmod q}} \Lambda(n)$$

and

$$\pi(x,q,a) = \#\{p \le x: \text{ p is prime and } p \equiv a \pmod q\}.$$

PNT for arithmetic progression states that for a fixed modulus q,

$$\psi(x,q,a) \sim \frac{x}{\varphi(q)}$$

and

$$\pi(x,q,a) \sim \frac{x}{\varphi(q)\log x}.$$

This is expected if we assume that all primes are well distributed to the all residue classes that are coprime to the modulus q. These two estimates comes from tha fact that

$$\prod_{\chi} L(s,\chi) \neq 0, \qquad \text{for} \qquad \sigma \ge 1. \tag{2.17}$$

Moreover, it is known that

$$\psi(x,q,a) = \frac{1}{\varphi(q)} x + O(xe^{-c\sqrt{\log x}}). \tag{2.18}$$

In fact we are allowed to make q grow with x. The following theorem is called the PNT for arithmetic progression with large moduli.

Siegel-Walfisz Theorem: Given $A > 0$ and for $q \le (\log x)^A$, (2.18) holds uniformly with a constant $c = c(A)$ that may depend on A.

This is the best known range for the modulus q for PNT on arithmetic progression. This theorem will be significiant in Chapter 5.

3 Wiener-Ikehara Tauberian Theorem

3.1 Tauberian Theory and an Approximation Lemma

In the previous chapter we derived the Prime Number Theorem $\psi(x) \sim x$ without the error term. For this it is crucial that we have to use

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

for all $t \neq 0$ real numbers. But we also need to bound on $\zeta(s)$, $\frac{1}{\zeta(s)}$ and $\zeta'(s)$ at infinity because we use the Fundamental formula. The natural question is: what is the least information concerning $\zeta(s)$ that would suffice to establish PNT? In this section we only use $\zeta(1+it) \neq 0$ for all real t and the functions $\zeta(s) - \dfrac{1}{s-1}$, $\zeta'(s) + \dfrac{1}{(s-1)^2}$ are continous in the closed half plane $\sigma \geq 1$. Then we prove a general Tauberian theorem of Wiener and Ikehara and as a result we deduce the PNT. Let us make a few words on Tauberian theorems. A Tauberian theorem is one in which the asymptotic behavior of a function is deduced from the behavior of some of its averages or its some generating functions. Generally Tauberian theorems are converses of fairly obvious results, but usually these converses depend on some additional assumptions that are called Tauberian condition. The name Tauberian comes from the Tauber who proved that :

Let $(a_n)_n$ be a sequence of complex numbers and put $s_n = a_0 + ... + a_n$. Suppose also that

(i) $f(r) = \displaystyle\sum_{n=0}^{\infty} a_n r^n$, $0 < r < 1$ and $\displaystyle\lim_{r \to 1^-} f(r) = s$.

(ii) $\displaystyle\lim_{n\to\infty} n a_n = 0$.

Then we have $\displaystyle\lim_{n\to\infty} s_n = s$.

It can be easily shown that $\lim_{n\to\infty} s_n = s$ implies the condition (i). But the condition (i) only is not enough to prove that $\lim_{n\to\infty} s_n = s$. We need the condition (ii) also. The condition (ii) is called the Tauberian condition. Later the Tauberian condition (ii) replaced by $(na_n)_n$ is bounded by Littlewood and that makes the proof much more difficult.

First we require a two sided approximation lemma which is the heart of the Wiener-Ikehara Tauberian Theorem.

Lemma 3.1. *Let $E(x) = e^x$ for $x \leq 0$ and $E(x) = 0$ for $x > 0$. For any given $\epsilon > 0$ there is a T and continuous functions f_+, f_- with $f_\pm \in L^1(\mathbb{R})$ such that*

(i) $f_- \leq E(x) \leq f_+$ *for all* $x \in \mathbb{R}$,

(ii) $\widehat{f_\pm}(t) = 0$ *for all* $|t| \geq T$,

(iii) $\displaystyle\int_{-\infty}^{+\infty} f_+(x)dx < 1+\epsilon, \int_{-\infty}^{+\infty} f_-(x)dx > 1-\epsilon$.

Remark 3.2. *Some remarks are in order. Since $f_\pm \in L^1(\mathbb{R})$, we have that the Fourier transform*

$$\widehat{f_\pm}(t) = \int_{-\infty}^{+\infty} f_\pm(w)e(-tw)dw$$

are uniformly continuous. Therefore from (ii) above it follows that $\widehat{f_\pm}(\pm T) = 0$ so that $\widehat{f_\pm}(t) = 0$ for all $|t| \geq T$. Since $f_\pm(t)$ are also continuous , by the Fourier integral theorem we have

$$\lim_{w\to\infty} \int_{-w}^{w} \left(1 - \frac{|t|}{w}\right) \widehat{f_\pm}(t)e(tx)dt = f_\pm(x)$$

for all x. But the functions $\widehat{f_\pm}$ are supported on the fixed interval $[-T, T]$. From this observation we see that the limit above is $\int_{-T}^{T} \widehat{f_\pm}(t)e(tx)dt$, hence we get

$$f_\pm(x) = \int_{-T}^{T} \widehat{f_\pm}(t)e(tx)dt$$

for all x. Moreover the function $\int_{-T}^{T} \widehat{f_\pm}(t)e^{2\pi itz}dt$ is an entire function of z. Therefore $f_\pm$ is the restriction to the real axis of an entire function.

Proof. We may suppose that $T \geq 1$. Let

$$\Delta_T(x) = T\left(\frac{\sin \pi Tx}{\pi Tx}\right)^2 \qquad \text{and} \qquad J_T(x) = \frac{3T}{4}\left(\frac{\sin \pi Tx/2}{\pi Tx/2}\right)^4$$

be the Fejer and Jackson kernels respectively. These functions have a peak of height $\asymp T$ and width $\asymp 1/T$ at 0, and have a total mass 1 by Theorem

1.12.
Now put

$$f(x) = (E * J_T)(x) = \int_{-\infty}^{+\infty} E(u)J_T(x-u)du.$$

This is a weighted average of the values of $E(u)$ with special emphasis on those u near x.
Next we show that

$$f(x) = E(x) + O(\min(1, 1/(Tx)^2)). \qquad (**)$$

To see this we consider several cases.
If $|x| \le 1/T$, we observe that $0 \le f(x) \le \int_{-\infty}^{+\infty} J_T(u)du = 1$.
If $x \ge 1/T$,

$$0 \le f(x) \ll \frac{1}{T^3}\int_{-\infty}^{0} (x-u)^{-4}du \ll 1/(Tx)^3.$$

Note that

$$f(x) - E(x) = \int_{-\infty}^{+\infty} (E(u) - E(x))J_T(x-u)du,$$

since $\int_{-\infty}^{+\infty} J_T(u)du = 1$.
Now, suppose that $-1 \le x \le -1/T$. If $2x \le u \le 0$ then

$$E(u) - E(x) = e^x(e^{u-x} - 1) = e^x(u - x + O((u-x)^2)).$$

Therefore

$$\int_{2x}^{0} (E(u)-E(x))J_T(x-u)du = -e^x \int_{x}^{-x} uJ_T(u)du + O\Big(\int_{x}^{-x} u^2 J_T(u)du\Big).$$

Note that the first integral on the right above vanishes since the integrand is an odd function, furthermore the second integral is $\ll 1/T^2$. On the other hand

$$\int_{0}^{\infty} (E(u) - E(x))J_T(x-u)du \ll \frac{1}{T^3}\int_{-x}^{\infty} \frac{1}{u^4}du \ll 1/|Tx|^3$$

and similarly $\displaystyle\int_{-\infty}^{2x}(E(u)-E(x))J_T(x-u)du \ll 1/|Tx|^3$, so we have $(**)$ in this case.
Finally, suppose that $x \leq -1$. Then

$$E(u)-E(x) = e^x(u-x+O((u-x)^2))$$

for $x-1 \leq u \leq x+1$, so that

$$\begin{aligned}\int_{x-1}^{x+1}(E(u)-E(x))J_T(x-u)du &= -e^x\int_{-1}^{1} uJ_T(u)du + O\Big(e^x\int_{-1}^{1} u^2J_T(u)du\Big)\\ &\ll e^xT^{-2} \ll 1/(Tx)^2.\end{aligned}$$

Also,

$$\int_{-\infty}^{x-1}(E(u)-E(x))J_T(x-u)du \ll e^xT^{-3}\int_1^{\infty} u^{-4}du \ll 1/(Tx)^2$$

and

$$\int_{x+1}^{\infty}(E(u)-E(x))J_T(x-u)du \ll T^{-3}x^{-4},$$

hence we have $(**)$ again.
Now observe that $\Delta_T(x) \ll T\min(1, 1/(Tx)^2)$, but there is no inequality in the reverse direction since $\Delta_T(x) = 0$ at integral multiples of $1/T$. To overcome this problem we consider a translate of the Fejer kernel. Since

$$\Delta_T(x) + \Delta_T(x+1/(2T)) \gg T\min(1, 1/(Tx)^2)$$

we take $f_{\pm}(x) = f(x) \pm \dfrac{c}{T}\Big(\Delta_T(x) + \Delta\Big(x+\dfrac{1}{2T}\Big)\Big)$.
By $(**)$, we see that if c is large enough, then $f_- \leq E(x) \leq f_+$ for all x. Next we show these functions satisfy the conditions $(i), (ii), (iii)$. By Fubini's theorem, if $f_1, f_2 \in L^1(\mathbb{R})$ then so is $f_1 * f_2$ and $\widehat{f_1 * f_2} = \widehat{f}_1\widehat{f}_2$. Hence, in particular, $f \in L^1(\mathbb{R})$ and $\widehat{f}(t) = \widehat{E}(t)\widehat{J_T}(t)$. But $\widehat{J_T}(t) = 0$ for $|t| \geq T$, thus $\widehat{f}(t) = 0$ for $|t| \geq T$. Similarly $\widehat{\Delta_T}(t) = 0$ for $|t| \geq T$, thus we have (i) and (ii).
Lastly, by Fubini's theorem again, we obtain that

$$\int_{-\infty}^{+\infty} f(x)dx = \Big(\int_{-\infty}^{\infty} E(x)dx\Big)\Big(\int_{-\infty}^{\infty} J_T(u)du\Big) = 1,$$

hence $\displaystyle\int_{-\infty}^{\infty} f_{\pm}(x)dx = 1 \pm \frac{2c}{T}$. Now, to satisfy the (iii), we take $T \geq c/\epsilon$. ∎

3.2 The Theorem of Wiener-Ikehara and Its Corollaries

Now we are ready to prove Wiener-Ikehara Tauberian Theorem.

Theorem 3.3. *Suppose that the function $a(u)$ is non-negative and incresing on $[0, \infty)$, that $\alpha(s) = \displaystyle\int_{0}^{\infty} e^{-us} da(u)$ converges for all s with $\sigma > 1$, and that $r(s) = \alpha(s) - \frac{c}{s-1}$ extends to a continuous function in the closed half-plane $\sigma \geq 1$.*
Then $\displaystyle\int_{0}^{x} 1 da(u) = ce^{x} + o(e^{x})$.

First we prove the theorem and then we deduce its corollaries which consists of the Prime Number Theorem. The theorem was first proved by Ikehara in 1931 who assumed that $\alpha(s) - \frac{c}{s-1}$ is analytic in the closed half-plane $\sigma \geq 1$. Wiener(1932) showed that mere continuity is enough.

Proof. Take $\delta > 0$ and let $E(u)$ be the same function in the Lemma 3.1. Then we can write $\displaystyle\int_{0}^{\infty} e^{-us} da(u) = e^{x} \int_{0}^{\infty} E(u-x) e^{-(1+\delta)u} da(u)$ and by the previous Lemma this is $\leq e^{x} \displaystyle\int_{0}^{\infty} f_{+}(u-x) e^{-(1+\delta)u} da(u)$. Now by the Remark 3.2, this is equal to

$$e^{x} \int_{0}^{\infty} \left(\int_{-T}^{T} \widehat{f_{+}}(t) e(tu - tx) dt \right) e^{-(1+\delta)u} da(u).$$

By Fubini's theorem we are allowed to interchange the order of integration. This will enable us to work with the function $e(x) = e^{2\pi i x}$ in the inner integral. Thus by interchanging the order of integration the above integral becomes

$$\begin{aligned} &e^{x} \int_{-T}^{T} \widehat{f_{+}}(t) e(-tx) \int_{0}^{\infty} e^{-(1+\delta-2\pi i t)u} da(u) dt \\ &\quad = e^{x} \int_{-T}^{T} \widehat{f_{+}}(t) e(-tx) \alpha(1+\delta-2\pi i t) dt. \end{aligned}$$

Now we make an important observation. If $a(u) = e^{u}$ then $\alpha(s) = \frac{1}{s-1}$ and from the calculation above we reach that

$$\int_0^\infty f_+(u-x)e^{-\delta u}du = \int_{-T}^{T} \widehat{f_+}(t)e(-tx)\frac{1}{\delta - 2\pi it}dt.$$

Therefore from the previous observation we obtain that

$$\int_0^\infty e^{-u\delta}da(u) \leq e^x \int_{-T}^{T} \widehat{f_+}(t)e(-tx)r(1+\delta-2\pi it)dt + ce^x \int_0^\infty f_+(u-x)e^{-\delta u}du.$$

Now since the function $r(s)$ is uniformly continuous in the closed rectangle $1 \leq \sigma \leq 1+\delta$, $|t| \leq 2\pi T$, each of the three terms above tends to a limit as $\delta \to 0^+$. Thus we obtain

$$\int_0^\infty 1da(u) \leq e^x \int_{-T}^{T} \widehat{f_+}(t)e(-tx)r(1-2\pi it)dt + ce^x \int_0^\infty f_+(u-x)du.$$

Now we divide both sides with e^x and let $x \to \infty$. By Riemann-Lebesgue lemma, the first integral on the right tends to 0 as $x \to \infty$, and the second integral on the right tends to $c\int_{-\infty}^{\infty} f_+(u)du$ which is less than $c(1+\epsilon)$ by the Lemma 3.1. Thus we obtain that

$$\limsup_{x\to\infty} e^{-x} \int_0^\infty 1da(u) \leq c\int_{-\infty}^{\infty} f_+(u)du \leq c(1+\epsilon).$$

Similarly by using the function f_-, we can also show that

$$\liminf_{x\to\infty} e^{-x} \int_0^\infty 1da(u) \geq c(1-\epsilon).$$

Since $\epsilon > 0$ is arbitrary we have the Theorem. ■

By making the change of variable $a(u) = A(e^u)$ we obtain the following equivalent formulation of the theorem.

Corollary 3.4. *Suppose $A(v)$ is non-negative and increasing function on $[0,\infty)$, that $\alpha(s) = \int_1^\infty v^{-s}dA(v)$ converges for all $\sigma > 1$ and that*

$r(s) = \alpha(s) - \frac{c}{s-1}$ extends to a continuous function in the closed half-plane $\sigma \geq 1$. Then

$$\int_1^\infty 1dA(v) = cx + o(x).$$

By setting $A(v) = \sum_{n \le v} a_n$, we get a useful Tauberian Theorem for Dirichlet series.

Corollary 3.5. *(Wiener-Ikehara) Suppose that $a_n \ge 0$ for all n, that $\alpha(s) = \sum_{n=1}^{\infty} \frac{a_n}{n^s}$ converges for all s with $\sigma > 1$ and that $r(s) = \alpha(s) - \frac{c}{s-1}$ extends to a continuous function in the closed half-plane $\sigma \ge 1$. Then*

$$\sum_{n \le x} a_n = cx + o(x).$$

Corollary 3.6. *(Prime Number Theorem) We have $\psi(x) \sim x$.*

Proof. Taking $a_n = \Lambda(n)$, we have

$$\sum_{n \le x} \Lambda(n) = x + o(x)$$

since $\sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s} = -\frac{\zeta'(s)}{\zeta(s)}$ satisfies the conditions of the Corollary 3.5 with $c = 1$ via Theorem 2.3. ∎

Now first we show directly that $M(x) = \sum_{n \le x} \mu(n) = o(x)$, then we show $M(x) = o(x)$ iff Prime Number Theorem.

Corollary 3.7. $M(x) = o(x)$.

Proof. We take $a_n = 1 + \mu(n)$. Then $a_n \ge 0$ for all $n \ge 1$ and $\alpha(s) = \sum_{n=1}^{\infty} \frac{a_n}{n^s} = \zeta(s) + \frac{1}{\zeta(s)}$ converges for all s with $\sigma > 1$ and $\alpha(s) - \frac{1}{s-1}$ extends to a continuous function in $\sigma \ge 1$ by Theorem 2.3. Thus by Corollary 3.5, we get $\sum_{n \le x} a_n = [x] + M(x) \sim x$. Hence we obtain that $M(x) = o(x)$.

∎

Corollary 3.8. *Prime Number Theorem iff $M(x) = o(x)$.*

Proof. We take $a_n = 1 + \mu(n) + \Lambda(n)$. Then $a_n \ge 0$ for all $n \ge 1$ and $\alpha(s) = \sum_{n=1}^{\infty} \frac{a_n}{n^s} = \zeta(s) + \frac{1}{\zeta(s)} - \frac{\zeta'(s)}{\zeta(s)}$ converges for all s with $\sigma > 1$ and $\alpha(s) - \frac{2}{s-1}$ extends to a continuous function in $\sigma \ge 1$ by Theorem 2.3.

Thus by Corollary 3.5, we get $\sum_{n\le x} a_n = [x] + M(x) + \psi(x) \sim 2x$. Hence we obtain that $M(x) = o(x)$ iff $\psi(x) \sim x$.

■

Now we can prove the Prime Number Theorem for arithmetic progressions by applying Wiener-Ikehara Tauberian Theorem. Similar to the classical Prime Number Theorem , we require that for a given character χ modulo q, $L(s,\chi)$ does not vanish on the line $\sigma = 1$.

Corollary 3.9. *Let $q \ge 1$ be a fixed modulus and a be a positive integer with $(a,q) = 1$. Then we have*

$$\psi(x;q,a) = \sum_{\substack{n\le x \\ n\equiv a \pmod q}} \Lambda(n) = \frac{x}{\varphi(q)} + o(x).$$

Proof. Consider the Dirichlet Series $\alpha(s) = \sum_{\substack{n\le x \\ n\equiv a \pmod q}} \frac{\Lambda(n)}{n^s}$ where $\sigma > 1$. Thus $\alpha(s)$ is analytic in $\sigma > 1$. By using the orthogonality relation of characters we obtain that

$$\begin{aligned}
\alpha(s) &= \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s}\Big(\frac{1}{\varphi(q)}\sum_{\chi}\overline{\chi(a)}\chi(n)\Big) \\
&= \frac{1}{\varphi(q)}\sum_{\chi}\overline{\chi(a)}\sum_{n=1}^{\infty}\frac{\chi(n)\Lambda(n)}{n^s} \\
&= \frac{1}{\varphi(q)}\sum_{\chi}\overline{\chi(a)}\Big(-\frac{L'(s,\chi)}{L(s,\chi)}\Big) \\
&= \frac{1}{\varphi(q)}\Big(-\frac{L'(s,\chi_0)}{L(s,\chi_0)}\Big) + \frac{1}{\varphi(q)}\sum_{\chi\ne\chi_0}\overline{\chi(a)}\Big(-\frac{L'(s,\chi)}{L(s,\chi)}\Big).
\end{aligned}$$

Since $\Big(-\frac{L'(s,\chi_0)}{L(s,\chi_0)}\Big) - \frac{1}{s-1}$ and for any non-principal character χ, $\Big(-\frac{L'(s,\chi)}{L(s,\chi)}\Big)$ can be continuously extended to the half-plane $\sigma \ge 1$ by (2.17), we see that $\alpha(s) - \frac{c}{s-1}$ can be continuously extended to the half-plane $\sigma \ge 1$ where $c = \frac{1}{\varphi(q)}$. Hence by Corollary 3.5 we get that $\psi(x;q,a) \sim \frac{x}{\varphi(q)}$.

■

Using Abel's summation, we have the Prime Number Theorem for arithmetic progressions.

Corollary 3.10. *Let $q \geq 1$ be a fixed modulus and a be a positive integer with $(a, q) = 1$. Then we have*

$$\pi(x; q, a) = \sum_{\substack{p \leq x \\ p \equiv a \pmod{q}}} 1 \sim \frac{x}{\varphi(q) \log x}.$$

4 Beurling's Prime Number Theorem

In this section we prove a generalized version of the Prime Number Theorem. First we define Beurling type of integers that is sometimes called the semi-group on integers. Now let P be a set of primes and B be its complementary set in the set of all primes. Let $\langle P \rangle = N_P$ be the integers that are coprime to all primes in B, thus

$$N_P = \langle P \rangle = \left\{ \prod_{j=1}^{k} p_j^{a_j} \ : \ k \geq 1, \ a_j \geq 0, \ p_j \in P \right\}$$
$$= \{n \geq 1 : (n, p) = 1 \qquad \forall p \in B\}.$$

Also we define $N_P(x)$ and $P(x)$ be the counting functions of N_P and P respectively.
For example if P is the set of all primes then $N_P(x) = [x] = x + O(1)$ and if P is the set of all odd primes then $N_P(x) = \frac{x}{2} + O(1)$.

The question is knowing the asmyptotic behavior of $N_P(x)$, can we determine the asymptotic behavior of $P(x)$? What growth condition must $N_P(x)$ have in order to prove

$$P(x) \sim \frac{x}{\log x}?$$

In this chapter we answer this question. Similar to Chapter 2, we study the function $\Lambda_P(n)$ which is defined by $\Lambda_P(n) = \Lambda(n)$ if $n \in N_P$ and $\Lambda_P(n) = 0$ otherwise.
One can easily show that $P(x) \sim \frac{x}{\log x}$ iff $\psi_P(x) = \sum_{n \leq x} \Lambda_P(n) \sim x$ by applying Abel's summation (partial summation) formula.
Now we define the generating function of N_P, namely

$$\zeta_P(s) = \sum_{n \in N_P} \frac{1}{n^s}$$

for $\sigma > 1$. This function is called the Beurling-Zeta function.
By Euler product formula, for $\sigma > 1$ we have

$$\zeta_P(s) = \prod_{p \in P} (1 - \frac{1}{p^s})^{-1}.$$

Therefore $\zeta_P(s)$ does not have a zero in $\sigma > 1$.

Note also that, for $\sigma > 1$ we have

$$\sum_{n=1}^{\infty} \frac{\Lambda_P(n)}{n^s} = -\frac{\zeta_P'(s)}{\zeta_P(s)}.$$

First we obtain an analytic continuation of this function under some conditions that depend on the asymptotic behavior of $N_P(x)$. Then we show $\zeta_P(s)$ cannot have a zero on the line $\sigma = 1$ under the same conditions. (Because this is significiant so as to prove $P(x) \sim \frac{x}{\log x}$.) Our main tool will be the Wiener-Ikehara Tauberian theorem because otherwise, in order to apply Perron's formula, we have to put bounds on the functions $\zeta_P(s)$ and its derivative at infinity. Moreover we have to pass to the left of the line $\sigma = 1$. But this can be difficult or sometimes impossible since the function depends on the set P.

Lemma 4.1. *Suppose*

$$N_P(x) = cx + O\left(\frac{x}{(\log x)^{\lambda}}\right) \tag{4.1}$$

where c is some positive constant and $\lambda > \frac{3}{2}$. Then we can extend the definition of $\zeta_P(s)$ to the half-plane $\sigma \geq 1$ so that

$$\zeta_P(s) = \frac{c}{s-1} + r_0(s)$$

and $r_0(s)$ is continuous in $\sigma \geq 1$. Furthermore $\zeta_P(s)$ does not vanish on the line $\sigma = 1$.

Proof. Note that $\zeta_P(s)$ is a Dirichlet Series with $\sigma_a = 1$. For $\sigma > 1$ by Abel's summation we have

$$\sum_{\substack{n \leq x \\ n \in \langle P \rangle}} \frac{1}{n^s} = \frac{N_P(x)}{x^s} + s\int_1^x \frac{N_P(u)}{u^{s+1}} du.$$

Therefore

$$\zeta_P(s) = s\int_1^{\infty} \frac{N_P(u)}{u^{s+1}} du = \frac{cs}{s-1} + s\int_1^{\infty} \frac{(N_P(u) - cu)}{u^{s+1}} du.$$

From (4.1) we know that $\displaystyle\int_1^{\infty} \frac{(N_P(u) - cu)}{u^2} du < \infty$. Hence the integral is uniformly convergent in $\sigma \geq 1$ and therefore the integral is continuous in $\sigma \geq 1$. So we can extend the definition of $\zeta_P(s)$ so that

$$\zeta_P(s) = \frac{c}{s-1} + r_0(s)$$

and $r_0(s)$ is continuous in $\sigma \geq 1$.
Also for $\sigma > 1$, $\zeta_P'(s) = -\frac{c}{(s-1)^2} + r_1(s)$ where

$$r_1(s) = r_0'(s) = \int_1^\infty \frac{(N_P(u) - cu)}{u^{s+1}} du - \int_1^\infty \frac{(N_P(u) - cu)\log u}{u^{s+1}} du.$$

Next we show that $\zeta_P(1+it) \neq 0$ when t is real and non-zero. Now we make a crucial observation which is not true if $\lambda \leq \frac{3}{2}$.

First note that

$$\begin{aligned}\int_2^\infty \frac{(\log u)^{1-\lambda}}{u^\sigma} du &= \int_{\log 2}^\infty v^{1-\lambda} e^{-(\sigma-1)v} dv \\ &= (\sigma-1)^{\lambda-2} \int_{(\sigma-1)\log 2}^\infty u^{1-\lambda} e^{-u} du \\ &\ll (\sigma-1)^{-\frac{1}{2}+\delta}\end{aligned}$$

where $\delta = \delta(\lambda) > 0$.

Now combining (4.1) and the observation above we get

$$r_1(s) \ll (\sigma-1)^{-\frac{1}{2}+\delta}.$$

Consequently if t is fixed and non-zero, then

$$\zeta_P(\sigma+it) - \zeta_P(1+it) = \int_1^\sigma \zeta_P'(\alpha+it) d\alpha \ll (\sigma-1)^{\frac{1}{2}+\delta}$$

for $\sigma > 1$ and σ near 1. By Euler product formula for $\sigma > 1$ we have

$$\zeta_P(s) = \prod_{p \in P} \left(1 - \frac{1}{p^s}\right)^{-1}.$$

Taking logarithm of both sides, we obtain that $\log \zeta_P(s) = \sum_{p \in P} \sum_{r=1}^\infty \frac{1}{rp^{rs}}$.
(As in the proof of Theorem 2.3, the trigonometric polynomial $3 + 4\cos\theta + \cos 2\theta$ may not work. So instead of this trigonometric polynomial, we work

with another non-negative trigonometric polynomial $a_0 + \sum_{k=1}^{K} a_k \cos k\theta$ for which the ratio a_1/a_0 is larger.)

Note that $\Delta_K(\theta) = 1 + 2\sum_{k=1}^{K}\left(1 - \frac{k}{K}\right) = \frac{1}{K}\left(\frac{\sin \pi K\theta}{\sin \pi\theta}\right)^2 \geq 0.$

Thus if $\sigma > 1$ then ,

$$\prod_{k=-K}^{K}\left(\zeta_P(\sigma + ikt)\right)^{1-|k|/K} = \exp\left(\sum_{p\in P}\sum_{r=1}^{\infty}\frac{1}{rp^{rs}}\Delta_K(rt(\log p)/(2\pi))\right).$$

Now since $\zeta_P(\sigma - it) = \overline{\zeta_P(\sigma + it)}$, we have $|\zeta_P(\sigma - it)| = |\zeta_P(\sigma + it)|$. From the observation above we see that

$$\zeta_P(\sigma)\prod_{k=1}^{K}\left|(\zeta_P(\sigma + ikt))^{2(1-|k|/K)}\right| \geq 1.$$

Assume $t \neq 0$ is a fixed real number. As $\sigma \to 1^+$, $|\zeta_P(\sigma + ikt)|$ tends to a finite limit for $k = 1, .., K$. Moreover since $\zeta_P(\sigma) \asymp \frac{1}{\sigma - 1}$, we get that

$$|\zeta_P(\sigma + it)| \gg (\sigma - 1)^{K/2(K-1)}$$

as $\sigma \to 1^+$.

Now suppose that $\zeta_P(1+it) = 0$. We know that $\zeta_P(\sigma+it) \ll (\sigma-1)^{\frac{1}{2}+\delta}$. Therefore as $\sigma \to 1^+$ we have

$$(\sigma - 1)^{\frac{1}{2}+\delta} \gg (\sigma - 1)^{\frac{K}{2(K-1)}}.$$

This gives a contradiction if $K > 1 + \frac{1}{2\delta}$.

Hence $\zeta_P(1 + it) \neq 0$ as desired.

■

4.1 Generalized Prime Number Theorem

Now we are ready to prove a generalized version of the prime number theorem. The following theorem was first proved by Beurling in 1937.

Theorem 4.2. *Suppose* (4.1) *where c is some positive constant and $\lambda > \frac{3}{2}$. Then we have $P(x) \sim \frac{x}{\log x}$.*

Before proving this heorem, note that this Theorem is really generalization of the classical Prime Number Theorem since if P is the set of all primes then $N_P(x) = [x] = x + O(1)$.

Proof. We define $\Lambda_P(n) = \Lambda(n)$ if $n \in N_P$ and $\Lambda_P(n) = 0$ otherwise.
One can easily show that $P(x) \sim \frac{x}{\log x}$ iff $\psi_P(x) = \sum_{n \leq x} \Lambda_P(n) \sim x$ by applying Abel's summation (partial summation) formula.
Moreover note that for $\sigma > 1$,

$$\sum_{n=1}^{\infty} \frac{\Lambda_P(n)}{n^s} = -\frac{\zeta_P'(s)}{\zeta_P(s)}$$

and

$$-\frac{\zeta_P'(s)}{\zeta_P(s)} = \frac{1}{s-1} + r(s)$$

where

$$r(s) = \frac{-r_0(s) + (s-1)r_1(s))}{(s-1)\zeta_P(s)}$$

and $r_0(s)$,$r_1(s)$ are same as in the Theorem 15.
If $\lambda > 2$ then the functions $r_0(s)$ and $r_1(s)$ are continuous in $\sigma \geq 1$ since from (2) the integral

$$\int_1^{\infty} \frac{(N_P(u) - cu)\log u}{u^2} du < \infty.$$

Therefore $r(s)$ is continuous in $\sigma \geq 1$ by Theorem 4.1. Hence by Wiener-Ikehara Tauberian theorem, we obtain that $\psi_P(x) \sim x$ and so $P(x) \sim \frac{x}{\log x}$.
From now on we assume $\frac{3}{2} < \lambda \leq 2$.
Under this condition we cannot ensure that $r_1(s)$ is continuous thus we cannot guarantee that $r(s)$ is continuous. Thus we benefit from the fact that $r_1(s)$ is bounded in mean-square by Plancherel's identity (2.15). So we follow a similar proof of Wiener-Ikehara theorem and we apply Plancherel's identity.
Suppose that $\delta > 0$, that T is a large positive number, and that $E(u)$ is defined as in Lemma 3.1. Then

$$\sum_{\substack{n \leq x \\ n \in N_P}} \Lambda(n) n^{-\delta} = x \sum_{n \in} \Lambda(n) n^{-1-\delta} E(\log n - \log x)$$

which by Lemma 3.1 is

$$\begin{aligned}
&\le x \sum_{n \in N_P} \Lambda(n) n^{-1-\delta} f_+(\log n - \log x) \\
&\le x \sum_{n \in N_P} \Lambda(n) n^{-1-\delta} \int_{-T}^{T} \widehat{f_+}(t) \left(\frac{x}{n}\right)^{-2\pi i t} dt \\
&= -x \int_{-T}^{T} \widehat{f_+}(t) x^{-2\pi i t} \frac{\zeta_P'}{\zeta_P}(1+\delta-2\pi i t) dt. \qquad (4.2)
\end{aligned}$$

Note that, similarly

$$\begin{aligned}
\int_1^\infty u^{-1-\delta} f_+(\log u - \log x) du &= \int_1^\infty u^{-1-\delta} \int_{-T}^{T} \widehat{f_+}(t) \left(\frac{x}{n}\right)^{-2\pi i t} du dt \\
&= \int_{-T}^{T} \widehat{f_+}(t) x^{-2\pi i t} \int_1^\infty u^{-1-\delta+2\pi i t} du dt \\
&= \int_{-T}^{T} \widehat{f_+}(t) x^{-2\pi i t} \frac{1}{\delta - 2\pi i t} dt.
\end{aligned}$$

We multiply both sides of this by x and combine with (4.2) to see that

$$\begin{aligned}
\sum_{\substack{n \le x \\ n \in N_P}} \Lambda(n) n^{-\delta} \le x &\int_1^\infty u^{-1-\delta} f_+(\log u - \log x) du \qquad (4.3) \\
&+ x \int_{-T}^{T} \widehat{f_+}(t) x^{-2\pi i t} \Big(-\frac{\zeta_P'}{\zeta_P}(1+\delta-2\pi i t) - \frac{1}{\delta - 2\pi i t} \Big) dt.
\end{aligned}$$

By using our formula for $r_i(s)$ in terms of integrals we see that we may write

$$r_1(s) = r_0'(s) = -sJ(s + \frac{r_0(s) - c}{s})$$

where

$$J(s) = \int_1^\infty (N_P(u) - cu)(\log u) u^{-s-1} du,$$

and

$$-\zeta_P'(s) = \frac{c}{(s-1)^2} - \frac{r_0(s) - c}{s} + sJ(s).$$

Thus

$$-\frac{\zeta_P'}{\zeta_P}(s) - \frac{1}{s-1} = \frac{c(s-1) + (1-2s) r_0(s)}{s(s-1)\zeta_P(s)} + \frac{s}{\zeta_P(s)} J(s)$$

and by splitting the integral at Ω, where Ω is a large parameter , we have

$$-\frac{\zeta_P'}{\zeta_P}(s) - \frac{1}{s-1} = C(s) + R(s)$$

where

$$R(s) = \int_\Omega^\infty (N_P(u) - cu)(\log u)u^{-s-1}du$$

and $C(u)$ is continuous for $\sigma \geq 1$. We consider first contribution of the remainder $R(s)$ to (4.3). By the Cauchy-Schwartz inequality, we see that

$$\begin{aligned}&\left|\int_{-T}^{T} \widehat{f_+}(t)x^{-2\pi it}R(1+\delta-2\pi it)dt\right|^2 \\ &\leq \left(\int_{-T}^{T}\left|\widehat{f_+}(t)\frac{1+\delta-2\pi it}{\zeta_P(1+\delta-2\pi it)}x^{-2\pi it}\right|^2 dt\right)\left(\int_{-T}^{T}\left|\int_\Omega^\infty \frac{(N_P(u)-cu)(\log u)}{u^{-2+\delta-2\pi it}}du\right|^2 dt\right).\end{aligned} \tag{4.4}$$

In Plancherel's identity (2.15), we take $\sigma = 1+\delta$ and $w(u) = (N_P(u) - cu)\log u$ for $u \geq \Omega$, $w(u) = 0$ otherwise. Thus we see that

$$\int_{-\infty}^{\infty}\left|\int_\Omega^\infty (N_P(u)-cu)(\log u)u^{-2+\delta-2\pi it}du\right|^2 dt = \int_\Omega^\infty (N_P(u)-cu)^2(\log u)^2u^{-3-2\delta}du,$$

which by (4.1) is

$$\ll \int_\Omega^\infty u^{-1}(\log u)^{2-2\lambda}du \ll_\lambda (\log\Omega)^{3-2\lambda}$$

uniformly for $\delta > 0$. The first integral on the right-hand side of (4.4) is also uniformly bounded as $\delta \to 0$, because $\zeta_P(1+it) \neq 0$. Thus the contribution of $R(s)$ to (4.3) is $\ll_\lambda (\log\Omega)^{3/2-\lambda}$, uniformly for $\delta > 0$. Therefore if we let $\delta \to 0$ in (4.2) and divide both sides by x, we obtain that

$$\frac{\psi_P(x)}{x} \leq \int_1^\infty u^{-1}f_+(\log u - \log x)du + \int_{-T}^{T}\widehat{f_+}(t)x^{-2\pi it}C(1-2\pi it)dt + O_\lambda((\log\Omega)^{3/2-\lambda}).$$

Thus as $x \to \infty$, the first integral on the right of the above inequality tends to $\int_{-\infty}^{\infty} f_+(v)dv$. Also since $\widehat{f_+}(t)C(1-2\pi it)$ is continuous, by Riemann-Lebesgue lemma the second integral on the right of the above inequality tends to 0 as $x \to \infty$. Hence

$$\limsup_{x\to\infty} \frac{\psi_P(x)}{x} \leq \int_{-\infty}^{\infty} f_+(v)dv + O_\lambda((\log \Omega)^{3/2-\lambda}).$$

So by Lemma 3.1, we know that the integral on the right is $< 1+\epsilon$ if T is sufficiently large. Since Ω might also be taken arbitrarily large, we conclude that

$$\limsup_{x\to\infty} \frac{\psi_P(x)}{x} \leq 1.$$

Similarly using the function f_-, we can show that

$$\liminf_{x\to\infty} \frac{\psi_P(x)}{x} \geq 1.$$

Thus we have

$$\psi_P(x) \sim x.$$

Hence we have the generalized Prime Number Theorem

$$P(x) \sim \frac{x}{\log x}.$$

■

So far we have discussed generalized Prime Number Theorem without an error term. This results from two the following two reasons. The first reason is that we apply Wiener-Ikehara Tauberian theorem and this theorem gives just an asymptotic result. The second reason is that, in order to get an error term in the Prime Number Theorem we need to apply Perron's formula, thus we have to work with the function $\zeta_P(s)$ in the left side of the line $\sigma = 1$. However if we assume (4.1) with some $\lambda > 0$ we cannot ensure that $\zeta_P(s)$ has an analytic continuation to the left of $\sigma = 1$. Thus we require better error term in (4.1). If we have $N_P(x) = cx + O(x^\theta)$ for some $c > 0$ and $\theta < 1$ then we can guarantee at least that $\zeta_P(s)$ has an analytic continuation to $\sigma > \theta$. Applying the same method in the proof of the Prime Number Theorem with classical error term $O(xe^{-c'\sqrt{\log x}})$, we can prove the following Theorem.

Theorem 4.3. *Suppose we have*

$$N_P(x) = cx + O(x^\theta)$$

for some $c > 0$ *and* $\theta < 1$. *Then we have* $P(x) = Li(x) + O(xe^{-c'\sqrt{\log x}})$ *for some constant* $c' > 0$.

4.2 Möbius Function over Semi-Groups

Now we define the Mobius function over a semi-group on integers and we discuss when the partial sums of the Mobius function over a semi-group has a cancelation.

Define $M_P(x) = \sum_{\substack{n \le x \\ n \in N_P}} \mu(n) = \sum_{n \le x} \mu_P(n)$, where $\mu_P(n) = \mu(n)$ if $n \in N_P$ and $\mu_P(n) = 0$ otherwise. Then observe that, for $\sigma > 1$ we have

$$\sum_{n=1}^{\infty} \frac{\mu_P(n)}{n^s} = \frac{1}{\zeta_P(s)}.$$

Since the Dirichlet series of $\mu_P(n)$ only depends on $\zeta_P(s)$ and not on $\zeta_P'(s)$, we expect that it is easier to work with $M_P(x)$. Indeed this is the case and compared to the Theorem 4.2 , proving $M_P(x) = o(x)$ under the condition (4.1) is much easier.

Theorem 4.4. *Suppose* (4.1) *with* $\lambda > 3/2$. *Then we have* $M_P(x) = o(x)$.

Proof. Since we have (4.1) with $\lambda > 3/2$, by Theorem 4.1 the function $\zeta_P(s)$ does not vanish on the line $\sigma = 1$. Thus the Dirichlet series

$$\alpha(s) = \sum_{\substack{n=1 \\ n \in N_P}}^{\infty} \frac{1 + \mu(n)}{n^s} = \zeta_P(s) + \frac{1}{\zeta_P(s)}$$

is analytic in $\sigma > 1$ and $\alpha(s) - \frac{c}{s-1}$ extends to a continuous function in the closed half-plane $\sigma \ge 1$. Hence by Wiener-Ikehara Tauberian theorem we obtain that $N_P(x) + M_P(x) \sim cx$ and so $M_P(x) = o(x)$.

■

In the previous chapter we showed that Prime Number Theorem is equivalent to $M(x) = o(x)$. In fact by mimicking the proof of the Theorem 4.2, we can prove the following Theorem.

Theorem 4.5. *Suppose we have* (4.1) *with* $\lambda > 3/2$. *Then* $P(x) \sim \frac{x}{\log x}$ *iff* $M_P(x) = o(x)$.

Remarks: Zhang show that (see [34]) if (4.1) holds with $\lambda > 1$ then it is still true that $M_P(x) = o(x)$. This is a bit surprising compared to the classical case $\pi(x) \sim x/\log x$ iff $M(x) = o(x)$, because when $1 < \lambda \leq 3/2$, Prime Number Theorem may fail (see [13]) but we can still have $M_P(x) = o(x)$. Thus it seems that partial sums of the Mobius function have a tendency to make a cancelation on semi-groups that even have not very much prime numbers. Based on this observation in the next chapter we discuss the Mobius function supported on a semi-group of integers and we give quantiative upper bounds for partial sums of the Mobius function supported on a semi-group. The next chapter is the main part of my thesis and it is a joint work with Emre Alkan.

5 Sums over the Möbius function and discrepancy of fractions

In this chapter we obtain quantitative upper bounds on partial sums of the Möbius function over semigroups of integers in an arithmetic progression. Exploiting the cancelation of such sums, we deduce upper bounds for the discrepancy of fractions in the unit interval $[0, 1]$ whose denominators satisfy the same restrictions. In particular, the uniform distribution and approximation of discrete weighted averages of such fractions are established as a consequence.

5.1 Introduction

Let $\mu(n)$ be the Möbius function. Estimating the size of the partial sum

$$M(x) = \sum_{n \leq x} \mu(n)$$

received continuous attention for a long time in the literature since for any fixed but arbitrary $\epsilon > 0$, the collection of estimates $M(x) \ll_\epsilon x^{\frac{1}{2}+\epsilon}$ is known to be equivalent to the Riemann hypothesis. Current unconditional estimates on $M(x)$ are far from being as satisfactory as the conditional ones, and the best such result is of form

$$M(x) \ll x \exp\left(-\frac{C\,(\log x)^{\frac{3}{5}}}{(\log\log x)^{\frac{1}{5}}}\right)$$

for some constant $C > 0$. This estimate is deduced from the strongest zero-free region for the Riemann zeta function due to Vinogradov and Korobov [23], [33]. Even the weaker estimate $M(x) = o(x)$ is equivalent to the prime number theorem and this equivalence is proved in any introductory course of analytic number theory. For detailed accounts of deeper connections between $M(x)$ and the distribution of zeros of the Riemann zeta function, we refer the reader to [12], [31]. Our main concern in this section is to obtain cancelation for partial sums of the Möbius function when $n \leq x$ ranges over certain semigroups of integers. Precisely, if P is a set of primes, then the semigroup generated by all primes in P is the set

$$\langle P\rangle = \left\{ \prod_{j=1}^{k} p_j^{a_j} \ : \ k \geq 1,\ a_j \geq 0,\ p_j \in P \right\}.$$

Alternatively, if B is the set of all primes not in P, then $\langle P\rangle$ can be viewed as the set of all remaining integers after sieving by the primes in B. The asymptotic theory of such semigroups were first investigated by Beurling [8] (although the original motivation can be traced back to the work of Landau [24] on the number of prime ideals in algebraic number fields whose norms are $\leq x$) who proved a vast generalization of the prime number theorem valid for such semigroups (see also [22] for a nice exposition of many different aspects of the theory). In particular, if $N_P(x)$ and $P(x)$ are the counting functions of $\langle P\rangle$ and P respectively, then Beurling proved that

$$N_P(x) = cx + O\left(\frac{x}{(\log x)^{\lambda}}\right)$$

with constants $c > 0$ and $\lambda > \frac{3}{2}$ implies the asymptotic $P(x) \sim \frac{x}{\log x}$. In connection with this paper, disproving a conjecture of Hall [19] and developing a variant of the Halász-Wirsing method, Zhang [34] has particularly shown, among other things, that if

$$N_P(x) = cx + O\left(\frac{x}{(\log x)^{\lambda}}\right)$$

holds with constants $c > 0$ and $\lambda > 1$, then

$$\sum_{\substack{n \leq x \\ n \in \langle P\rangle}} \mu(n) = o(x).$$

This is a surprising phenomenon for such Möbius sums since, unlike the classical case, the corresponding prime number theorem is not in general true when $1 < \lambda \leq \frac{3}{2}$. Motivated by Zhang's result, we give quantitative cancelations for partial sums of the Möbius function over integers in $\langle P\rangle$ that are in a given arithmetic progression.

Theorem 5.1. *Let B be a set of primes such that*

$$B(x) = \#\{p \leq x \ : \ p \in B\} \ll \frac{x}{(\log x)^{\lambda}}$$

with $1 < \lambda < 2$. Let P be the set of all primes that are not in B and denote by $\langle P \rangle$, the set of integers all of whose prime factors are in P. Then for any fixed $k \geq 1$, $(b,k) = 1$ and $x \geq 2$, we have

$$M_{P,b,k}(x) = \sum_{\substack{n \leq x \\ n \in \langle P \rangle \\ n \equiv b \pmod{k}}} \mu(n) = O_{k,B}\left(\frac{x}{(\log x)^{\lambda-1}}\right),$$

where the implied constant depends only on k and B. Moreover, if $B(x) \ll \frac{x}{\log^2 x}$, then for $(b,k) = 1$ and $x \geq 3$

$$M_{P,b,k}(x) = \sum_{\substack{n \leq x \\ n \in \langle P \rangle \\ n \equiv b \pmod{k}}} \mu(n) = O_{k,B}\left(\frac{x \log\log x}{\log x}\right).$$

Finally, if $B(x) \ll \frac{x}{(\log x)^{\lambda}}$ with $\lambda > 2$, then for $(b,k) = 1$ and $x \geq 2$

$$M_{P,b,k}(x) = \sum_{\substack{n \leq x \\ n \in \langle P \rangle \\ n \equiv b \pmod{k}}} \mu(n) = O_{k,B}\left(\frac{x}{\log x}\right).$$

Some remarks are now in order. First of all, the trivial estimate gives

$$|M_{P,b,k}(x)| \leq \#\{n \leq x \ : \ n \in \langle P \rangle\,, \ n \equiv b \pmod{k}\},$$

and it is possible to show by a similar reasoning as in the proof of Lemma 5.6 below that the set of all $n \in \langle P \rangle$ with $n \equiv b \pmod{k}$ has positive lower density. This justifies the cancelation obtained for $M_{P,b,k}(x)$ in Theorem 5.1. Moreover, by the result of Beurling [8], the estimate

$$N_P(x) = cx + O\left(\frac{x}{(\log x)^{\lambda}}\right)$$

with $c > 0$ and $\lambda > \frac{3}{2}$ implies that

$$P(x) = \pi(x) + o\left(\frac{x}{\log x}\right),$$

and

$$B(x) = o\left(\frac{x}{\log x}\right)$$

follows. Therefore, the sieve conditions $B(x) \ll \frac{x}{(\log x)^\lambda}$ with $\lambda > 1$, used in Theorem 5.1 are compatible with this consequence. Our method for proving Theorem 5.1 is indeed flexible and in addition we can even obtain similar cancelation for $M_{P,b,k}(x)$ over arithmetic progressions with large moduli, namely when k is allowed to grow with x.

Theorem 5.2. *Let B and P be complementary sets of primes as in Theorem 5.1. Assuming $B(x) \ll \frac{x}{(\log x)^\lambda}$ with $1 < \lambda < 2$ and $1 \le k \le \log x$, we have for $(b,k) = 1$ and $x \ge e$ that*

$$M_{P,b,k}(x) = \sum_{\substack{n \le x \\ n \in \langle P \rangle \\ n \equiv b \pmod k}} \mu(n) = O_B\left(\frac{x}{(\log x)^{\lambda-1}}\right).$$

Here the implied constant depends only on B. Moreover, assuming $B(x) \ll \frac{x}{\log^2 x}$ and $1 \le k \le \log x$, we have for $(b,k) = 1$ and $x \ge 3$ that

$$M_{P,b,k}(x) = \sum_{\substack{n \le x \\ n \in \langle P \rangle \\ n \equiv b \pmod k}} \mu(n) = O_B\left(\frac{x \log\log x}{\log x}\right).$$

Finally, assuming $B(x) \ll \frac{x}{(\log x)^\lambda}$ with $\lambda > 2$ and $1 \le k \le \log x$, we have for $(b,k) = 1$ and $x \ge e$ that

$$M_{P,b,k}(x) = \sum_{\substack{n \le x \\ n \in \langle P \rangle \\ n \equiv b \pmod k}} \mu(n) = O_B\left(\frac{x}{\log x}\right).$$

Note that, because of the trivial estimate $M_{P,b,k}(x) = O\left(\frac{x}{k}\right)$, only under the conditions $k = o\left((\log x)^{\lambda-1}\right)$, $o\left(\frac{\log x}{\log\log x}\right)$, $o(\log x)$, nontrivial cancelation is obtained in Theorem 5.2. It is possible to make interesting choices for the set B above. As a consequence of Selberg's Λ^2 sieve [28], the number of twin primes that are $\leq x$ is $\ll \frac{x}{\log^2 x}$. Therefore, one could possibly take B as the set of twin primes in Theorems 5.1 and 5.2.

For $Q \geq 1$, let $\mathfrak{F}_Q$ denote the Farey fractions of order Q in the unit interval $[0,1]$. It is well known that

$$|\mathfrak{F}_Q| = N(Q) = \frac{3Q^2}{\pi^2} + O(Q \log Q).$$

For $0 \leq \alpha \leq 1$, let $M(\alpha, Q)$ be the number of Farey fractions of order Q that are $\leq \alpha$. Then the local discrepancy of $\mathfrak{F}_Q$ at α is defined by

$$R_{N(Q)}(\alpha) = \left| \frac{M(\alpha, Q)}{N(Q)} - \alpha \right|.$$

The average behavior of the moments of the local discrepancy is a central problem since Franel [18] and Landau [25] showed that the Riemann hypothesis is equivalent to both of the estimates

$$\sum_{j=1}^{N(Q)} R^2_{N(Q)}(\gamma_j) = O_\epsilon\left(Q^{-1+\epsilon}\right) \text{ and } \sum_{j=1}^{N(Q)} R_{N(Q)}(\gamma_j) = O_\epsilon\left(Q^{\frac{1}{2}+\epsilon}\right)$$

for every $\epsilon > 0$, where $\frac{1}{Q} = \gamma_1 < \gamma_2 < ... < \gamma_{N(Q)} = 1$ are the Farey fractions of order Q in increasing order. In another direction, Erdös, Kac, Van Kampen and Wintner [16] proved that $\mathfrak{F}_Q$ is uniformly distributed as $Q \to \infty$, namely that $R_{N(Q)}(\alpha) \to 0$ as $Q \to \infty$ for all $0 \leq \alpha \leq 1$. This shows that the absolute discrepancy of $\mathfrak{F}_Q$ which is defined as

$$D_{N(Q)}(\mathfrak{F}_Q) = \sup_{0 \leq \alpha \leq 1} R_{N(Q)}(\alpha)$$

tends to 0 as $Q \to \infty$. In [5] anf [6], the authors focused on the effect of addition of Farey fractions and addition of torsion points on elliptic curves to pair correlation measures. In recent years there has been increasing interest for the distribution of subsets of $\mathfrak{F}_Q$ that are defined by certain sieve conditions on the denominators of the fractions. Boca, Cobeli and Zaharescu [9] and Haynes [20] studied the distribution of fractions with odd denominators. Then Haynes [21] extended his results to fractions whose denominators are not divisible by a fixed prime p. Let B be a set of primes such that

$$\sum_{p\in B}\frac{1}{p^{\sigma}}<\infty$$

for some $\sigma<1$. For a given modulus k, let $\mathfrak{F}_{Q,b,k,B}$ be the set of all Farey fractions $\frac{a}{q}$ of order Q such that $q\equiv b \pmod{k}$, $(b,k)=1$ and q is not divisible by any prime in B. Let

$$D_{N_{Q,b,k,B}}(\mathfrak{F}_{Q,b,k,B})=\sup_{0\le\alpha\le 1}R_{N_{Q,b,k,B}}(\alpha)$$

be the absolute discrepancy, where similarly as above

$$R_{N_{Q,b,k,B}}(\alpha)=\left|\frac{M(\alpha,Q,b,k,B)}{N_{Q,b,k,B}}-\alpha\right|$$

is the local discrepancy with $N_{Q,b,k,B}=|\mathfrak{F}_{Q,b,k,B}|$ and $M(\alpha,Q,b,k,B)=|\mathfrak{F}_{Q,b,k,B}\cap[0,\alpha]|$. It is proved in [4] that

$$D_{N_{Q,b,k,B}}(\mathfrak{F}_{Q,b,k,B})\asymp\frac{1}{Q},$$

where the implied constants depend only on k and B. The condition

$$\sum_{p\in B}\frac{1}{p^{\sigma}}<\infty$$

with $\sigma<1$ was essential in [4] since the required estimates for $M_{P,b,k}(x)$ depended on an application of Perron's formula and a contour integral type argument. Our approach here differs from that of [4] in the respect that we give upper bounds on the absolute discrepancy of $\mathfrak{F}_{Q,b,k,B}$ even for large modulus k assuming weaker sieve conditions such as $B(x)\ll\frac{x}{(\log x)^{\lambda}}$ with $\lambda>1$. Note that under such an assumption one can only deduce

$$\sum_{p\in B}\frac{1}{p}<\infty, \tag{5.1}$$

and therefore one can not use the approach of [4] to handle such conditions. We remark that the distribution of integers subject to (5.1) was first studied by Erdös [15]. Applications of (5.1) to the non-vanishing of Fourier coefficients of modular forms are given in [1]-[3]. Armed with Theorem 5.1 and

Theorem 5.2 above, we are able to deduce the following consequences for the absolute discrepancy.

Corollary 5.3. *If $B(x) \ll \frac{x}{(\log x)^\lambda}$ with $1 < \lambda < 2$, then for all $Q \geq 2$ with $N_{Q,b,k,B} \geq 1$ and $(b,k) = 1$, we have*

$$D_{N_{Q,b,k,B}}(\mathfrak{F}_{Q,b,k,B}) = O_{k,B}\left(\frac{(\log Q)^{2-\lambda}}{Q}\right),$$

where the implied constant depends only on k and B. If $B(x) \ll \frac{x}{\log^2 x}$, then for all $Q \geq 3$ with $N_{Q,b,k,B} \geq 1$ and $(b,k) = 1$, we have

$$D_{N_{Q,b,k,B}}(\mathfrak{F}_{Q,b,k,B}) = O_{k,B}\left(\frac{(\log\log Q)^2}{Q}\right).$$

If $B(x) \ll \frac{x}{(\log x)^\lambda}$ with $\lambda > 2$, then for all $Q \geq 3$ with $N_{Q,b,k,B} \geq 1$ and $(b,k) = 1$, we have

$$D_{N_{Q,b,k,B}}(\mathfrak{F}_{Q,b,k,B}) = O_{k,B}\left(\frac{\log\log Q}{Q}\right).$$

Corollary 5.4. *If $B(x) \ll \frac{x}{(\log x)^\lambda}$ with $1 < \lambda < 2$, then for all $1 \leq k \leq (\log Q)^{2-\lambda}$, $Q \geq 3$ with $N_{Q,b,k,B} \geq 1$ and $(b,k) = 1$, we have*

$$D_{N_{Q,b,k,B}}(\mathfrak{F}_{Q,b,k,B}) = O_{B}\left(\frac{k(\log Q)^{2-\lambda}}{Q}\right),$$

where the implied constant depends only on B. If $B(x) \ll \frac{x}{\log^2 x}$, then for all $1 \leq k \leq (\log\log Q)^2$, $Q \geq 16 > e^e$ with $N_{Q,b,k,B} \geq 1$ and $(b,k) = 1$, we have

$$D_{N_{Q,b,k,B}}(\mathfrak{F}_{Q,b,k,B}) = O_{B}\left(\frac{k(\log\log Q)^2}{Q}\right).$$

If $B(x) \ll \frac{x}{(\log x)^\lambda}$ with $\lambda > 2$, then for all $1 \leq k \leq \log\log Q$, $Q \geq 16 > e^e$ with $N_{Q,b,k,B} \geq 1$ and $(b,k) = 1$, we have

$$D_{N_{Q,b,k,B}}(\mathfrak{F}_{Q,b,k,B}) = O_{B}\left(\frac{k\log\log Q}{Q}\right).$$

Note that as an immediate consequence of the discrepancy estimates in the above results, we see that the sets of fractions $\mathfrak{F}_{Q,b,k,B}$ are uniformly distributed. Indeed, even under the condition (5.1), using Lemma 5.6 below and the trivial estimate on Möbius sums, one can again deduce the uniform distribution of $\mathfrak{F}_{Q,b,k,B}$ but the corresponding discrepancy estimate would not be as good as above. To mention a further application, let f be a function of bounded variation $V(f)$ on $[0,1]$. Then by Koksma's inequality, we see that

$$\left| \frac{1}{N_{Q,b,k,B}} \sum_{x_j \in \mathfrak{F}_{Q,b,k,B}} f(x_j) - \int_0^1 f(t)\, dt \right| \leq V(f) D_{N_{Q,b,k,B}} \left(\mathfrak{F}_{Q,b,k,B}\right).$$

Consequently, using our estimates on $D_{N_{Q,b,k,B}}(\mathfrak{F}_{Q,b,k,B})$, it is possible to approximate discrete weighted averages over $\mathfrak{F}_{Q,b,k,B}$ by the Riemann-Stieltjes integral on $[0,1]$ of the weight function f with error tending to zero as a function of Q, k and B as Q tends to infinity.

5.2 Preliminaries

We will need the following lemmas for the proof of our results.

Lemma 5.5. *For any nonnegative real number λ and $x \geq 2$, define*

$$S_\lambda(x) = \sum_{n \leq x} \frac{\Lambda(n)}{n(1 + \log(\frac{x}{n}))^\lambda},$$

where $\Lambda(n)$ is the Von-Mangoldt function. If $\lambda > 1$, then for $x \geq 2$, $S_\lambda(x) = O_\lambda(1)$, where the implied constant depends only on λ. If $\lambda = 1$, then for $x \geq 3$, $S_1(x) = O(\log\log x)$. If $0 \leq \lambda < 1$, then for $x \geq 2$, $S_\lambda(x) = O_\lambda((\log x)^{1-\lambda})$.

Proof. First of all note that

$$\begin{aligned} S_\lambda(x) &= \sum_{n \leq x} \frac{\Lambda(n)}{n(1 + \log(\frac{x}{n}))^\lambda} \\ &= \sum_{p \leq x} \frac{\log p}{p(1 + \log(\frac{x}{p}))^\lambda} + \sum_{\substack{p^m \leq x \\ m \geq 2}} \frac{\log p}{p^m(1 + \log(\frac{x}{p^m}))^\lambda}. \end{aligned} \tag{5.2}$$

Clearly, we have

$$\sum_{\substack{p^m \leq x \\ m \geq 2}} \frac{\log p}{p^m(1 + \log(\frac{x}{p^m}))^\lambda} \leq \sum_p \frac{\log p}{p(p-1)} < \infty. \tag{5.3}$$

Combining (5.2) and (5.3), we see that

$$S_\lambda(x) = \sum_{p\le x} \frac{\log p}{p(1+\log(\frac{x}{p}))^\lambda} + O(1). \tag{5.4}$$

Next for $x \ge 2$, consider an e-adic division of the interval $(1, x]$ into intervals of form $(\frac{x}{e^s}, \frac{x}{e^{s-1}}]$ with $1 \le s \le [\log x]+1$. Using Mertens' estimate, we obtain

$$\sum_{\frac{x}{e^s}<p\le\frac{x}{e^{s-1}}} \frac{\log p}{p} = \left(\log\left(\frac{x}{e^{s-1}}\right) + O(1)\right) - \left(\log\left(\frac{x}{e^s}\right) + O(1)\right) = O(1). \tag{5.5}$$

Also assuming $\frac{x}{e^s} < p \le \frac{x}{e^{s-1}}$, one has

$$s^\lambda \le \left(1+\log\left(\frac{x}{p}\right)\right)^\lambda \tag{5.6}$$

for any $\lambda \ge 0$. From (5.5) and (5.6), we may deduce that

$$\sum_{\frac{x}{e^s}<p\le\frac{x}{e^{s-1}}} \frac{\log p}{p(1+\log(\frac{x}{p}))^\lambda} \le \frac{1}{s^\lambda} \sum_{\frac{x}{e^s}<p\le\frac{x}{e^{s-1}}} \frac{\log p}{p} \ll \frac{1}{s^\lambda}, \tag{5.7}$$

where the implied constant is absolute. It follows from (5.7) that

$$\sum_{p\le x} \frac{\log p}{p(1+\log(\frac{x}{p}))^\lambda} = \sum_{s\le[\log x]+1} \sum_{\frac{x}{e^s}<p\le\frac{x}{e^{s-1}}} \frac{\log p}{p(1+\log(\frac{x}{p}))^\lambda} \ll \sum_{s\le[\log x]+1} \frac{1}{s^\lambda}. \tag{5.8}$$

Note that if $\lambda > 1$, then for $x \ge 2$,

$$\sum_{s\le[\log x]+1} \frac{1}{s^\lambda} = O_\lambda(1).$$

Therefore, combining (5.4) and (5.8), $S_\lambda(x) = O_\lambda(1)$ follows in this case, where the implied constant depends only on λ. If $\lambda = 1$, then

$$\sum_{s\le[\log x]+1} \frac{1}{s} = O(\log\log x)$$

and $S_1(x) = O(\log\log x)$ follows from (5.4) and (5.8), when $x \ge 3$ (so that

$\log\log x > 0$). Finally, if $0 \le \lambda < 1$, then

$$\sum_{s\le[\log x]+1} \frac{1}{s^\lambda} \le 1 + \int_1^{1+\log x} \frac{1}{t^\lambda} dt = 1 + \frac{1}{1-\lambda}((1+\log x)^{1-\lambda} - 1) = O_\lambda((\log x)^{1-\lambda}).$$

Consequently, $S_\lambda(x) = O_\lambda\left((\log x)^{1-\lambda}\right)$ follows again from (5.4) and (5.8) when $x \ge 2$. This completes the proof of Lemma 5.5. ■

Lemma 5.6. *Assume that B is a set of primes satisfying*

$$\sum_{p\in B} \frac{1}{p} < \infty.$$

Let $\mathfrak{F}_{Q,b,k,B}$ be the set of all Farey fractions $\frac{a}{q}$ of order Q such that $q \equiv b \pmod{k}$, $(b,k) = 1$ and q is not divisible by any prime in B. Let $f(Q)$ be a monotonically increasing function such that $f(Q) = O(\log Q)$ for all $Q \ge 2$. If $1 \le k \le f(Q)$ and $N_{Q,b,k,B} = |\mathfrak{F}_{Q,b,k,B}|$, then

$$\frac{1}{N_{Q,b,k,B}} = O_B\left(\frac{k}{Q^2}\right)$$

for all $Q \ge 2$ with $N_{Q,b,k,B} \ge 1$, where the implied constant depends only on B.

Proof. Let P be the set of primes not in B. Clearly, we have

$$N_{Q,b,k,B} = \#\left\{\frac{a}{q} \in \mathfrak{F}_{Q,b,k,B} : q \in \langle P\rangle, q \equiv b \pmod{k}\right\} = \sum_{\substack{q\le Q \\ q\equiv b \pmod{k} \\ q\in\langle P\rangle}} \varphi(q).$$

We remark that since B has density zero in the set of primes, the conditions $q \equiv b \pmod{k}$ and $q \in \langle P\rangle$ are not degenerate. In fact, there are infinitely many primes $q \in \langle P\rangle$ such that $q \equiv b \pmod{k}$ (their density being $\frac{1}{\varphi(k)}$ in the set of primes by the Siegel-Walfisz theorem (see [12]) when $1 \le k \le f(Q)$ and $f(Q) = O(\log Q)$). Let $p_1, p_2, \ldots, p_s$ be the first s primes in B. Then we have

$$\sum_{\substack{q\le Q \\ q\equiv b \pmod{k} \\ q\in\langle P\rangle}} \varphi(q) \ge \sum_{\substack{q\le Q \\ q\equiv b \pmod{k} \\ (q,p_j)=1 \text{ for all } j\le s}} \varphi(q) - \sum_{\substack{q\le Q \\ q\equiv b \pmod{k} \\ q\equiv 0 \pmod{p_j} \text{ for some } j>s}} \varphi(q). \tag{5.9}$$

Using the Inclusion-Exclusion principle, the main term on the right side of (5.9) can be written as

$$\sum_{\kappa}(-1)^{|\kappa|}\sum_{\substack{q\le Q\\ q\equiv b \pmod{k}\\ q\equiv 0 \pmod{d_\kappa}}}\varphi(q), \tag{5.10}$$

where κ runs over all subsets of $\{1,2,\dots,s\}$ and $d_\kappa=\prod_{j\in\kappa}p_j$ (empty products are assumed to be 1). We may assume that $(k,d_\kappa)=1$, since otherwise there are no integers q satisfying $q\equiv b \pmod{k}$ and $q\equiv 0 \pmod{d_\kappa}$. Therefore, by the Chinese Remainder theorem, the congruences $q\equiv b \pmod{k}$ and $q\equiv 0 \pmod{d_\kappa}$ reduce to $q\equiv u \pmod{kd_\kappa}$ for a suitable u. Using these observations, one obtains

$$\sum_{\substack{q\le Q\\ q\equiv b \pmod{k}\\ q\equiv 0 \pmod{d_\kappa}}}\varphi(q)=\sum_{\substack{q\le Q\\ q\equiv u \pmod{kd_\kappa}}}\varphi(q)=\sum_{d\le Q}\mu(d)\sum_{\substack{r\le \frac{Q}{d}\\ dr\equiv u \pmod{kd_\kappa}}} r. \tag{5.11}$$

Define

$$N(u,d,k,d_\kappa):=\#\{1\le r\le kd_\kappa : dr\equiv u \pmod{kd_\kappa}\}.$$

Assume that $r\equiv r_j \pmod{kd_\kappa}$, $1\le j\le N(u,d,k,d_\kappa)$ are all solutions of the congruence $dr\equiv u \pmod{kd_\kappa}$. It follows that

$$\sum_{\substack{r\le \frac{Q}{d}\\ dr\equiv u \pmod{kd_\kappa}}} r=\sum_{1\le j\le N(u,d,k,d_\kappa)}\sum_{\substack{r\le \frac{Q}{d}\\ r\equiv r_j \pmod{kd_\kappa}}} r. \tag{5.12}$$

It is elementary to estimate the inner sum on the right side of (5.12) and we obtain

$$\sum_{\substack{r\le \frac{Q}{d}\\ r\equiv r_j \pmod{kd_\kappa}}} r=\frac{Q^2}{2d^2kd_\kappa}+O\left(\frac{Q}{d}\right)+O(kd_\kappa), \tag{5.13}$$

where the implied constants in (5.13) are absolute. Therefore, (5.12) and

(5.13) give

$$\sum_{\substack{r \le \frac{Q}{d} \\ dr \equiv u \pmod{kd_\kappa}}} r = \sum_{1 \le j \le N(u,d,k,d_\kappa)} \left(\frac{Q^2}{2d^2 k d_\kappa} + O\left(\frac{Q}{d}\right) + O(kd_\kappa) \right) \tag{5.14}$$

$$= \frac{N(u,d,k,d_\kappa) Q^2}{2d^2 k d_\kappa} + O\left(\frac{kd_\kappa Q}{d}\right) + O(k^2 d_\kappa^2),$$

where the trivial bound $N(u,d,k,d_\kappa) \le kd_\kappa$ is used. Using (5.14) on the right side of (5.11), we see that

$$\sum_{\substack{q \le Q \\ q \equiv b \pmod{k} \\ q \equiv 0 \pmod{d_\kappa}}} \varphi(q) = \sum_{d \le Q} \mu(d) \left(\frac{N(u,d,k,d_\kappa) Q^2}{2d^2 k d_\kappa} + O\left(\frac{kd_\kappa Q}{d}\right) + O(k^2 d_\kappa^2) \right) \tag{5.15}$$

$$= \frac{Q^2}{2kd_\kappa} \sum_{d \le Q} \frac{\mu(d) N(u,d,k,d_\kappa)}{d^2} + O(kd_\kappa Q \log Q) + O(k^2 d_\kappa^2 Q)$$

$$= \frac{Q^2}{2kd_\kappa} \sum_{d=1}^{\infty} \frac{\mu(d) N(u,d,k,d_\kappa)}{d^2} + O(kd_\kappa Q \log Q) + O(k^2 d_\kappa^2 Q),$$

since

$$\sum_{d=1}^{\infty} \frac{\mu(d) N(u,d,k,d_\kappa)}{d^2} < \infty$$

and

$$\sum_{d > Q} \frac{\mu(d) N(u,d,k,d_\kappa)}{d^2} = O\left(\frac{kd_\kappa}{Q}\right).$$

Finally, gathering (5.10) and (5.15), we may rewrite the main term as

$$\sum_{\kappa} (-1)^{|\kappa|} \left(\frac{M(u,k,d_\kappa) Q^2}{2kd_\kappa} + O(kd_\kappa Q \log Q) + O(k^2 d_\kappa^2 Q) \right) \tag{5.16}$$

$$= \frac{Q^2}{2k} \sum_{\kappa} (-1)^{|\kappa|} \frac{M(u,k,d_\kappa)}{d_\kappa} + O(2^s P_s k \, Q \log Q) + O(2^s P_s^2 k^2 Q),$$

where

$$M(u,k,d_\kappa) := \sum_{d=1}^{\infty} \frac{\mu(d) N(u,d,k,d_\kappa)}{d^2} \text{ and } P_s = \prod_{j=1}^{s} p_j.$$

Next we show that

$$\sum_{\kappa}(-1)^{|\kappa|}\frac{M(u,k,d_\kappa)}{d_\kappa}$$

is bounded below by a positive constant depending only on B. Recall that $N(u,d,k,d_\kappa)$ counts the number of solutions of the congruence $dr \equiv u \pmod{kd_\kappa}$ or equivalently the number of solutions of the system $dr \equiv b \pmod{k}$ and $dr \equiv 0 \pmod{d_\kappa}$. If $(d,k)>1$, then $(b,k)>1$ so that there are no solutions and $N(u,d,k,d_\kappa)=0$ in this case. If $(d,k)=1$, then $r \equiv bd^{-1} \pmod{k}$ and $r \equiv 0 \pmod{\frac{d_\kappa}{(d,d_\kappa)}}$ give that $N(u,d,k,d_\kappa)=(d,d_\kappa)$ (again assuming $(k,d_\kappa)=1$). Clearly, $N(u,d,k,d_\kappa)$ is a multiplicative function of d and one can write $M(u,k,d_\kappa)$ as an Euler product. Precisely, we have

$$\begin{aligned} M(u,k,d_\kappa) &= \sum_{d=1}^{\infty}\frac{\mu(d)N(u,d,k,d_\kappa)}{d^2} = \prod_p\left(1-\frac{N(u,p,k,d_\kappa)}{p^2}\right) \qquad (5.17)\\ &= \prod_{\substack{p\\(p,k)=1}}\left(1-\frac{(p,d_\kappa)}{p^2}\right) = \frac{6}{\pi^2}\left(\prod_{p|d_\kappa}\frac{p}{p+1}\right)\left(\prod_{p|k}\left(1-\frac{1}{p^2}\right)^{-1}\right)\end{aligned}$$

when $(k,d_\kappa)=1$. Note that if $(k,d_\kappa)>1$, then $N(u,d,k,d_\kappa)=0$ for all $d\geq 1$ and $M(u,k,d_\kappa)=0$. Consequently, using (5.17), we see that

$$\begin{aligned}\sum_{\kappa}(-1)^{|\kappa|}\frac{M(u,k,d_\kappa)}{d_\kappa} &= \sum_{\substack{\kappa\\(k,d_\kappa)=1}}(-1)^{|\kappa|}\frac{M(u,k,d_\kappa)}{d_\kappa} \qquad (5.18)\\ &= \frac{6}{\pi^2}\prod_{p|k}\left(1-\frac{1}{p^2}\right)^{-1}\sum_{\substack{\kappa\\(k,d_\kappa)=1}}\frac{(-1)^{|\kappa|}}{d_\kappa}\left(\prod_{p|d_\kappa}p\right)\left(\prod_{p|d_\kappa}\frac{1}{p+1}\right)\\ &= \frac{6}{\pi^2}\prod_{p|k}\left(1-\frac{1}{p^2}\right)^{-1}\sum_{\substack{\kappa\\(k,d_\kappa)=1}}(-1)^{|\kappa|}\left(\prod_{p|d_\kappa}\frac{1}{p+1}\right)\\ &= \frac{6}{\pi^2}\prod_{p|k}\left(1-\frac{1}{p^2}\right)^{-1}\prod_{\substack{1\leq j\leq s\\(p_j,k)=1}}\left(1-\frac{1}{p_j+1}\right)\\ &\geq \frac{6}{\pi^2}\prod_{\substack{1\leq j\leq s\\(p_j,k)=1}}\left(1-\frac{1}{p_j+1}\right).\end{aligned}$$

Note that since

$$\sum_{p_j \in B} \frac{1}{p_j} < \infty,$$

we have

$$\prod_{p_j \in B} \left(1 - \frac{1}{p_j + 1}\right) > 0.$$

As a result of (5.18), one can deduce that

$$\sum_{\kappa} (-1)^{|\kappa|} \frac{M(u,k,d_\kappa)}{d_\kappa} \geq C_B := \frac{6}{\pi^2} \prod_{p \in B} \left(1 - \frac{1}{p+1}\right) > 0, \tag{5.19}$$

where the positive constant on the right side of (5.19) depends only on B. Combining (5.10), (5.16) and (5.19), we obtain that the main term on the right side of (5.9) is

$$\geq \frac{C_B Q^2}{2k} + O(2^s P_s k \; Q \log Q) + O(2^s P_s^2 k^2 Q). \tag{5.20}$$

It remains to treat the error term on the right side of (5.9). Again we may assume that $(k, p_j) = 1$, since otherwise there are no integers q satisfying $q \equiv b \pmod{k}$ and $q \equiv 0 \pmod{p_j}$ for some $j > s$. If $(k, p_j) = 1$, then these congruences reduce to $q \equiv v_j \pmod{kp_j}$ for a suitable v_j. We also assume that $p_j \leq Q$, since otherwise there are no integers q with $q \leq Q$ and $q \equiv 0 \pmod{p_j}$. In this way one obtains

$$\sum_{\substack{q \leq Q \\ q \equiv b \pmod{k} \\ q \equiv 0 \pmod{p_j} \text{ for some } j > s \\ \text{and } p_j \leq Q}} \varphi(q) \leq \sum_{\substack{q \leq Q \\ q \equiv v_j \pmod{kp_j} \text{ for some } j > s \\ \text{and } p_j \leq Q}} q \tag{5.21}$$

$$\leq Q \sum_{\substack{j > s \\ p_j \leq Q}} \left(\frac{Q}{kp_j} + O(1)\right) \leq \frac{Q^2}{k} \sum_{j > s} \frac{1}{p_j} + O\left(Q \sum_{p_j \leq Q} 1\right).$$

Since

$$\sum_{p \in B} \frac{1}{p} < \infty,$$

B has density zero in the set of primes and consequently

$$\sum_{p_j \le Q} 1 = o\left(\frac{Q}{\log Q}\right)$$

follows. Choose s large enough to satisfy

$$\sum_{j>s} \frac{1}{p_j} \le \frac{C_B}{4}.$$

Therefore, the right side of (5.21) is

$$\le \frac{C_B Q^2}{4k} + o\left(\frac{Q^2}{\log Q}\right). \tag{5.22}$$

Putting the estimates (5.20) and (5.22) into (5.9), we deduce that

$$N_{Q,b,k,B} = \sum_{\substack{q \le Q \\ q \equiv b \pmod k \\ q \in \langle P \rangle}} \varphi(q) \ge \frac{C_B Q^2}{4k} + o\left(\frac{Q^2}{\log Q}\right) + O(2^s P_s k\, Q \log Q) + O(2^s P_s^2 k^2 Q). \tag{5.23}$$

Finally, using the fact that $1 \le k \le f(Q) = O(\log Q)$, we obtain from (5.23) that

$$\frac{1}{N_{Q,b,k,B}} = O_B\left(\frac{k}{Q^2}\right)$$

for all $Q \ge 2$ with $N_{Q,b,k,B} \ge 1$, where implied constant depends only on B. This completes the proof of Lemma 5.6. ■

As was remarked above, replacing $\varphi(q)$ by 1 and repeating the proof of Lemma 5.6, one can show that

$$\sum_{\substack{n \le x \\ n \in \langle P \rangle \\ n \equiv b \pmod k}} 1 \gg_B \frac{x}{k}.$$

5.3 Proof of Theorem 5.1

Proof. For $k \ge 1$, let

$$\Lambda_k(n) = \sum_{d|n} \mu(d) \log^k\left(\frac{n}{d}\right)$$

be the generalized Von-Mangoldt function of order k. Λ_k is supported on integers having at most k prime factors. In particular, the connection between the classical Von-Mangoldt function $\Lambda_1 = \Lambda$ and Λ_2 via Selberg's formula

$$\Lambda_2(n) = \Lambda(n)\log n + \sum_{d|n} \Lambda(d)\Lambda\left(\frac{n}{d}\right)$$

served as the starting point of the first elementary proofs of the Prime Number Theorem by Erdös [14] and Selberg [29]. The importance of the class of functions Λ_k for $k \geq 2$ was fully realized later by the work of Bombieri [11] who obtained, among other things, the asymptotic behavior of sums of the form

$$\sum_{n \leq x} a_n \Lambda_k(n)$$

under an average assumption on the remainders arising from the distribution of the sequence $\{a_n\}$ for all x^ν, $\nu \leq 1$. Bombieri [10] then deduced strong estimates for the number of twin almost-primes. Extreme examples pertaining to the limitations of such asymptotics were constructed by Ford [17]. The basic idea behind the proof of Theorem 5.1 is to exploit a similar connection between Λ and Λ_2 supported on Beurling type integers. To this end, let B and P be complementary sets of primes and let μ_P be the Möbius function supported on $\langle P \rangle$, so that $\mu_P(n) = \mu(n)$ when $n \in \langle P \rangle$ and $\mu_P(n) = 0$ otherwise. It is easy to see that $\sum_{d|n} \mu_P(d) = 0$ when n has a prime divisor in P and $\sum_{d|n} \mu_P(d) = 1$ when n has no prime divisors in P. Similarly, let $\Lambda_P(n)$ be the Von-Mangoldt function supported on $\langle P \rangle$, so that $\Lambda_P(n) = \log p$ if $n = p^m$, $m \geq 1$ with $p \in P$ and $\Lambda_P(n) = 0$ otherwise. Then we have the identity

$$\sum_{d|n} \Lambda_P(d) = \log s(n), \tag{5.24}$$

where $s(n) \in \langle P \rangle$ is the largest such divisor of n. It follows from (5.24) that

$$\Lambda_P(n) = \sum_{d|n} \mu(d) \log s\left(\frac{n}{d}\right) = \log s(n) \sum_{d|n} \mu(d) - \sum_{d|n} \mu(d) \log s(d) \tag{5.25}$$

$$= -\sum_{d|n} \mu(d) \log s(d).$$

Next we introduce the generalized Von-Mangoldt function of order 2 on P as

$$\Lambda_{2,P}(n) := \sum_{d|n} \mu(d) \log^2 s\left(\frac{n}{d}\right). \tag{5.26}$$

Expanding the right side of (5.26) and using (5.25), we arrive at the identity

$$\Lambda_{2,P}(n) = \log^2 s(n) \sum_{d|n} \mu(d) - 2 \log s(n) \sum_{d|n} \mu(d) \log s(d) + \sum_{d|n} \mu(d) \log^2 s(d) \tag{5.27}$$

$$= 2\Lambda_P(n) \log s(n) + \sum_{d|n} \mu(d) \log^2 s(d).$$

Consequently, rewriting (5.27), we have

$$\sum_{d|n} \mu(d) \log^2 s(d) = \Lambda_{2,P}(n) - 2\Lambda_P(n) \log s(n). \tag{5.28}$$

Thus one obtains

$$\mu(n) \log^2 s(n) = \sum_{d|n} \mu(d) F_P\left(\frac{n}{d}\right), \tag{5.29}$$

where

$$F_P(n) := \Lambda_{2,P}(n) - 2\Lambda_P(n) \log s(n). \tag{5.30}$$

In particular, if $n \in \langle P \rangle$, then $s(n) = n$ so that

$$\mu(n) \log^2 n = \sum_{d|n} \mu(d) F_P\left(\frac{n}{d}\right)$$

follows from (5.29) with

$$F_P\left(\frac{n}{d}\right) = \Lambda_{2,P}\left(\frac{n}{d}\right) - 2\Lambda_P\left(\frac{n}{d}\right) \log\left(\frac{n}{d}\right).$$

Let χ be a Dirichlet character modulo k. Multiplying both sides of (5.29) by $\chi(n)$ and summing over all $n \le x$ with $n \in \langle P \rangle$, we deduce that

$$\begin{aligned}\sum_{n\le x}\chi(n)\mu_P(n)\log^2 n &= \sum_{\substack{n\le x\\ n\in\langle P\rangle}}\chi(n)\sum_{d|n}\mu(d)F_P\left(\frac{n}{d}\right) \\ &= \sum_{\substack{d\le x\\ d\in\langle P\rangle}}\chi(d)\mu(d)\sum_{\substack{m\le \frac{x}{d}\\ m\in\langle P\rangle}}\chi(m)F_P(m) \\ &= \sum_{n\le x}\chi(n)\mu_P(n)S_{P,\chi}\left(\frac{x}{n}\right),\end{aligned} \tag{5.31}$$

where

$$S_{P,\chi}(x) := \sum_{\substack{n\le x\\ n\in\langle P\rangle}}\chi(n)\left(\Lambda_{2,P}(n) - 2\Lambda_P(n)\log n\right). \tag{5.32}$$

Moreover, if $n \in \langle P \rangle$, then

$$\Lambda_{2,P}(n) = \sum_{d|n}\mu(d)\log^2 s\left(\frac{n}{d}\right) = \sum_{d|n}\mu(d)\log^2\left(\frac{n}{d}\right) = \Lambda_2(n) \tag{5.33}$$

and $\Lambda_P(n) = \Lambda(n)$. Therefore, (5.32) can be written as

$$S_{P,\chi}(x) = \sum_{\substack{n\le x\\ n\in\langle P\rangle}}\chi(n)\left(\Lambda_2(n) - 2\Lambda(n)\log n\right). \tag{5.34}$$

Using Selberg's formula, (5.34) reduces to

$$S_{P,\chi}(x) = \sum_{\substack{n\le x\\ n\in\langle P\rangle}}\chi(n)(\Lambda*\Lambda)(n) - \sum_{\substack{n\le x\\ n\in\langle P\rangle}}\chi(n)\Lambda(n)\log n, \tag{5.35}$$

where

$$(\Lambda*\Lambda)(n) = \sum_{d|n}\Lambda(d)\Lambda\left(\frac{n}{d}\right)$$

is the convolution of Λ with itself. We estimate both sums on the right side of (5.35). Assume that

$$B(x) \ll \frac{x}{(\log x)^{\lambda}} \tag{5.36}$$

with $1 < \lambda < 2$. Note that

$$\sum_{\substack{n \le x \\ n \in \langle P \rangle}} \chi(n)\Lambda(n) = \sum_{n \le x} \chi(n)\Lambda(n) - \sum_{\substack{n \le x \\ n \notin \langle P \rangle}} \chi(n)\Lambda(n). \qquad (5.37)$$

Since $\Lambda(n)$ is supported only at prime powers, we see that if $n \notin \langle P \rangle$ and $\Lambda(n) \neq 0$, then $n = p^m$ for some $p \in B$ and $m \geq 1$. Consequently,

$$\sum_{\substack{n \le x \\ n \notin \langle P \rangle}} \chi(n)\Lambda(n) = \sum_{\substack{p \le x \\ p \in B}} \chi(p) \log p + \sum_{\substack{p^m \le x \\ m \ge 2 \\ p \in B}} \chi(p) \log p. \qquad (5.38)$$

It is easy to see that

$$\sum_{\substack{p^m \le x \\ m \ge 2 \\ p \in B}} \chi(p) \log p = O(\sqrt{x}), \qquad (5.39)$$

with an absolute implied constant. Moreover, using (5.36), one obtains

$$\sum_{\substack{p \le x \\ p \in B}} \chi(p) \log p = O\left(\sum_{\substack{p \le x \\ p \in B}} \log p \right) = O_B\left(\frac{x}{(\log x)^{\lambda - 1}} \right), \qquad (5.40)$$

where the implied constant depends only on B. Combining (5.38), (5.39) and (5.40), we have

$$\sum_{\substack{n \le x \\ n \notin \langle P \rangle}} \chi(n)\Lambda(n) = O_B\left(\frac{x}{(\log x)^{\lambda - 1}} \right) = O_B\left(\frac{x}{(1 + \log x)^{\lambda - 1}} \right) \qquad (5.41)$$

for all $x \geq 2$. It follows from (5.37) and (5.41) that

$$\sum_{\substack{n \le x \\ n \in \langle P \rangle}} \chi(n)\Lambda(n) = \psi(x, \chi) + O_B\left(\frac{x}{(1 + \log x)^{\lambda - 1}} \right) \qquad (5.42)$$

for $x \geq 2$, where

$$\psi(x, \chi) := \sum_{n \le x} \chi(n)\Lambda(n). \qquad (5.43)$$

Such sums as in (5.43) are naturally encountered in the proof of the prime number theorem over arithmetic progressions. We need to distinguish two cases here. If χ is non-principal and the modulus k is fixed, then it is known that (see [12])

$$\psi(x,\chi) = O\left(xe^{-c_1\sqrt{\log x}}\right) \tag{5.44}$$

for some constant $c_1 > 0$ and all large x in terms of k, where the implied constant is absolute (it turns out that a much weaker estimate such as

$$\psi(x,\chi) = O\left(\frac{x}{\log x}\right)$$

would also suffice for our purposes). Combining (5.42) and (5.44), we deduce that

$$\sum_{\substack{n\leq x \\ n\in\langle P\rangle}} \chi(n)\Lambda(n) = O_{k,B}\left(\frac{x}{(1+\log x)^{\lambda-1}}\right) \tag{5.45}$$

for $x \geq 2$. It follows from (5.45) that

$$\sum_{\substack{n\leq x \\ n\in\langle P\rangle}} \chi(n)\Lambda(n)\log n = O_{k,B}(x(\log x)^{2-\lambda}) \tag{5.46}$$

for $x \geq 2$. Moreover, we have

$$\sum_{\substack{n\leq x \\ n\in\langle P\rangle}} \chi(n)(\Lambda*\Lambda)(n) = \sum_{\substack{n\leq x \\ n\in\langle P\rangle}} \chi(n)\sum_{d|n}\Lambda(d)\Lambda\left(\frac{n}{d}\right) = \sum_{\substack{d\leq x \\ d\in\langle P\rangle}} \chi(d)\Lambda(d)\sum_{\substack{m\leq \frac{x}{d} \\ m\in\langle P\rangle}} \chi(m)\Lambda(m). \tag{5.47}$$

To estimate the inner sum on the right of (5.47), we obtain, using (5.45), that

$$\sum_{\substack{m\leq \frac{x}{d} \\ m\in\langle P\rangle}} \chi(m)\Lambda(m) = O_{k,B}\left(\frac{x}{d(1+\log\left(\frac{x}{d}\right))^{\lambda-1}}\right) \tag{5.48}$$

when $d \leq \frac{x}{2}$. Clearly,

$$\sum_{\substack{\frac{x}{2}<d\leq x \\ d\in\langle P\rangle}} \chi(d)\Lambda(d)\sum_{\substack{m\leq \frac{x}{d} \\ m\in\langle P\rangle}} \chi(m)\Lambda(m) = 0 \tag{5.49}$$

holds for any character χ modulo k. Using (5.48) and (5.49) on the right side of (5.47), we see that

$$\sum_{\substack{n \le x \\ n \in \langle P \rangle}} \chi(n)(\Lambda * \Lambda)(n) = O_{k,B}\left(x \sum_{d \le \frac{x}{2}} \frac{\Lambda(d)}{d(1 + \log\left(\frac{x}{d}\right))^{\lambda-1}} \right). \tag{5.50}$$

Since $0 < \lambda - 1 < 1$, one can apply Lemma 5.5 to get

$$\sum_{d \le \frac{x}{2}} \frac{\Lambda(d)}{d(1 + \log\left(\frac{x}{d}\right))^{\lambda-1}} = O_\lambda\left((\log x)^{2-\lambda}\right) = O_B\left((\log x)^{2-\lambda}\right), \tag{5.51}$$

since dependence of the implied constant on λ can be viewed as dependence on B. Combining (5.50) and (5.51), we deduce that

$$\sum_{\substack{n \le x \\ n \in \langle P \rangle}} \chi(n)(\Lambda * \Lambda)(n) = O_{k,B}\left(x(\log x)^{2-\lambda}\right) \tag{5.52}$$

for $x \ge 2$. Gathering (5.35), (5.46) and (5.52), one has

$$S_{P,\chi}(x) = O_{k,B}\left(x(\log x)^{2-\lambda}\right) \tag{5.53}$$

for $x \ge 2$. Using now (5.31), (5.53) and the fact

$$S_{P,\chi}\left(\frac{x}{n}\right) = 0$$

when $\frac{x}{2} < n \le x$, one obtains that

$$\begin{aligned} \sum_{n \le x} \chi(n)\mu_P(n)\log^2 n &= \sum_{n \le \frac{x}{2}} \chi(n)\mu_P(n) S_{P,\chi}\left(\frac{x}{n}\right) \\ &= O_{k,B}\left(x(\log x)^{2-\lambda} \sum_{n \le \frac{x}{2}} \frac{1}{n} \right) \\ &= O_{k,B}\left(x(\log x)^{3-\lambda}\right). \end{aligned} \tag{5.54}$$

Applying Abel's summation to (5.54), we have

$$\sum_{n \le x} \chi(n)\mu_P(n) = O_{k,B}\left(\frac{x}{(\log x)^{\lambda-1}}\right) + O_{k,B}\left(\int_2^x \frac{1}{(\log t)^\lambda}\, dt\right) = O_{k,B}\left(\frac{x}{(\log x)^{\lambda-1}}\right) \tag{5.55}$$

since

$$\int_2^x \frac{1}{(\log t)^\lambda}\, dt = O\left(\frac{x}{(\log x)^\lambda}\right)$$

with an absolute implied constant. Let us now assume that χ is the principal character modulo k. Then

$$\psi(x,\chi) = x + O\left(xe^{-c_1\sqrt{\log x}}\right) \tag{5.56}$$

for all large x in terms of k, where the implied constant is absolute. Consequently, one obtains

$$\sum_{\substack{n\le x\\ n\in\langle P\rangle}} \chi(n)\Lambda(n) = x + O_{k,B}\left(\frac{x}{(1+\log x)^{\lambda-1}}\right) \tag{5.57}$$

for $x \geq 2$. By Abel's summation on (5.57), we see that

$$\sum_{\substack{n\le x\\ n\in\langle P\rangle}} \chi(n)\Lambda(n)\log n = x\log x + O_{k,B}\left(x(\log x)^{2-\lambda}\right) - \int_2^x \frac{t+R(t)}{t} dt, \tag{5.58}$$

where

$$R(t) = O_{k,B}\left(\frac{t}{(\log t)^{\lambda-1}}\right)$$

for $t \geq 2$. It is now easy to see that

$$\int_2^x \frac{t+R(t)}{t}\, dt = O_{k,B}(x). \tag{5.59}$$

Combining (5.58) and (5.59), we deduce that

$$\sum_{\substack{n\le x\\ n\in\langle P\rangle}} \chi(n)\Lambda(n)\log n = x\log x + O_{k,B}\left(x(\log x)^{2-\lambda}\right) \tag{5.60}$$

for $x \geq 2$. Using (5.49), (5.57) and Lemma 5.5, we have

$$\sum_{\substack{n \leq x \\ n \in \langle P \rangle}} \chi(n)(\Lambda * \Lambda)(n) = \sum_{\substack{d \leq \frac{x}{2} \\ d \in \langle P \rangle}} \chi(d)\Lambda(d) \sum_{\substack{m \leq \frac{x}{d} \\ m \in \langle P \rangle}} \chi(m)\Lambda(m) \tag{5.61}$$

$$= x \sum_{\substack{d \leq \frac{x}{2} \\ d \in \langle P \rangle}} \frac{\chi(d)\Lambda(d)}{d} + O_{k,B}\left(x \sum_{d \leq \frac{x}{2}} \frac{\Lambda(d)}{d\left(1 + \log\left(\frac{x}{d}\right)\right)^{\lambda - 1}} \right)$$

$$= x \sum_{\substack{d \leq \frac{x}{2} \\ d \in \langle P \rangle}} \frac{\chi(d)\Lambda(d)}{d} + O_{k,B}\left(x(\log x)^{2-\lambda} \right).$$

It follows, by Abel's summation on (5.57), that

$$\sum_{\substack{d \leq x \\ d \in \langle P \rangle}} \frac{\chi(d)\Lambda(d)}{d} = O_{k,B}(1) + \int_2^x \frac{t + R(t)}{t^2}\, dt, \tag{5.62}$$

where

$$R(t) = O_{k,B}\left(\frac{t}{(\log t)^{\lambda-1}} \right)$$

for $t \geq 2$. Therefore, we have

$$\int_2^x \frac{R(t)}{t^2}\, dt = O_{k,B}\left(\int_2^x \frac{1}{t(\log t)^{\lambda-1}}\, dt \right) = O_{k,B}\left((\log x)^{2-\lambda} \right) \tag{5.63}$$

for $x \geq 2$. Combining (5.62) and (5.63), one obtains

$$\sum_{\substack{d \leq x \\ d \in \langle P \rangle}} \frac{\chi(d)\Lambda(d)}{d} = \log x + O_{k,B}\left((\log x)^{2-\lambda} \right). \tag{5.64}$$

Replacing x by $\frac{x}{2}$ and gathering (5.61) and (5.64), we see that

$$\sum_{\substack{n \leq x \\ n \in \langle P \rangle}} \chi(n)(\Lambda * \Lambda)(n) = x \log x + O_{k,B}\left(x(\log x)^{2-\lambda} \right). \tag{5.65}$$

From (5.35), (5.60) and (5.64), we again have

$$S_{P,\chi}(x) = O_{k,B}\left(x(\log x)^{2-\lambda} \right)$$

for $x \geq 2$, when χ is the principal character modulo k. Consequently as above, we obtain

$$\sum_{n \leq x} \chi(n)\mu_P(n) = O_{k,B}\left(\frac{x}{(\log x)^{\lambda-1}}\right) \tag{5.66}$$

for $x \geq 2$. Lastly, using (5.55) and (5.66), we deduce that

$$\sum_{\substack{n \leq x \\ n \in \langle P \rangle \\ n \equiv b \pmod k}} \mu(n) = \frac{1}{\varphi(k)} \sum_{\chi} \chi(b^{-1}) \sum_{n \leq x} \chi(n)\mu_P(n) = O_{k,B}\left(\frac{x}{(\log x)^{\lambda-1}}\right)$$

for $x \geq 2$. If $B(x) \ll \frac{x}{\log^2 x}$, then by a similar argument as above, using Lemma 5.5 with $\lambda = 1$ to get $S_1(x) = O(\log\log x)$, we see for any Dirichlet character χ modulo k that

$$S_{P,\chi}(x) = O_{k,B}(x \log\log x)$$

with $x \geq 3$. Noting that $S_{P,\chi}(\frac{x}{n}) = O(1)$, when $\frac{x}{3} < n \leq x$, one may deduce that

$$\sum_{n \leq x} \chi(n)\mu_P(n) = O_{k,B}\left(\frac{x \log\log x}{\log x}\right) \tag{5.67}$$

for $x \geq 3$. It follows from (5.67) that

$$\sum_{\substack{n \leq x \\ n \in \langle P \rangle \\ n \equiv b \pmod k}} \mu(n) = O_{k,B}\left(\frac{x \log\log x}{\log x}\right)$$

for $x \geq 3$. Finally, if $B(x) \ll \frac{x}{(\log x)^{\lambda}}$ with $\lambda > 2$, then using Lemma 5.5 with $\lambda - 1 > 1$ and $S_\lambda(x) = O_\lambda(1)$, we obtain for any Dirichlet character χ modulo k that $S_{P,\chi}(x) = O_B(x)$ with $x \geq 2$. It follows as above that

$$\sum_{\substack{n \leq x \\ n \in \langle P \rangle \\ n \equiv b \pmod k}} \mu(n) = O_{k,B}\left(\frac{x}{\log x}\right)$$

for $x \geq 2$. This completes the proof of Theorem 5.1.

■

5.4 Proof of Theorem 5.2

Proof of Theorem 5.2 is similar to the proof of Theorem 5.1, the only significant modification is needed for obtaining the required uniformity for the modulus k of the arithmetic progression and this is accomplished easily as a consequence of the Siegel-Walfisz theorem (see [12]). Precisely, if χ is a Dirichlet character modulo k and $1 \le k \le (\log x)^A$, where A is a positive constant, then there exists a constant $c_1 > 0$ such that

$$\psi(x,\chi) = \sum_{n \le x} \chi(n)\Lambda(n) = O(xe^{-c_1\sqrt{\log x}}) \tag{5.68}$$

holds with an absolute implied constant when χ is non-principal and $x \ge x_0(A)$. If χ is principal, then

$$\psi(x,\chi) = x + O(xe^{-c_1\sqrt{\log x}}) \tag{5.69}$$

holds when $x \ge x_0(A)$. Taking $A = 1$ and using our assumption $1 \le k \le \log x$ (note that this forces $x \ge e$), we can see that (5.68) and (5.69) hold for $x \ge x_0(1)$, where $x_0(1)$ is an absolute constant. Therefore, the required estimates for $M_{P,b,k}(x)$ can be obtained similarly as in the proof of Theorem 5.1 for $x \ge x_0(1)$, where the implied constants depend only on B. But obviously, by adjusting the constants depending on B, these estimates also hold for smaller values of x as well. $\square$

5.5 Proof of Corollary 5.3

First of all, observing that

$$M(\alpha, Q, b, k, B) = \sum_{\substack{q \le Q \\ q \in \langle P \rangle \\ q \equiv b \pmod{k}}} \#\{1 \le a \le \alpha q \ : \ (a,q) = 1\},$$

we have

$$M(\alpha, Q, b, k, B) - \alpha N_{Q,b,k,B} = \sum_{\substack{q \le Q \\ q \in \langle P \rangle \\ q \equiv b \pmod{k}}} \big(\#\{1 \le a \le \alpha q \ : \ (a,q) = 1\} - \alpha\varphi(q)\big)$$

$$= \sum_{\substack{q\le Q\\ q\in\langle P\rangle\\ q\equiv b \pmod k}} \left(\sum_{1\le a\le \alpha q} \sum_{\substack{m|a\\ m|q}} \mu(m) - \alpha \sum_{m|q} \frac{q\mu(m)}{m} \right) = \sum_{\substack{q\le Q\\ q\in\langle P\rangle\\ q\equiv b \pmod k}} \sum_{m|q} \mu(m) \left(\left[\frac{\alpha q}{m}\right] - \frac{\alpha q}{m} \right)$$

$$= - \sum_{\substack{q\le Q\\ q\in\langle P\rangle\\ q\equiv b \pmod k}} \sum_{m|q} \mu\left(\frac{q}{m}\right) \{\alpha m\} = - \sum_{\substack{m\le Q\\ m\in\langle P\rangle\\ (m,k)=1}} \{\alpha m\} \sum_{\substack{n\le \frac{Q}{m}\\ n\in\langle P\rangle\\ n\equiv m^{-1}b \pmod k}} \mu(n),$$

where $mm^{-1} \equiv 1 \pmod k$ and $\{x\} = x - [x]$ is the fractional part of x. It follows that

$$|M(\alpha, Q, b, k, B) - \alpha N_{Q,b,k,B}| \le \sum_{\substack{m\le Q\\ m\in\langle P\rangle\\ (m,k)=1}} \left| \sum_{\substack{n\le \frac{Q}{m}\\ n\in\langle P\rangle\\ n\equiv m^{-1}b \pmod k}} \mu(n) \right| = \sum_{\substack{m\le Q\\ m\in\langle P\rangle\\ (m,k)=1}} \left| M_{P,m^{-1}b,k}\left(\frac{Q}{m}\right) \right|.$$

Assuming $B(x) \ll \frac{x}{(\log x)^\lambda}$ with $1 < \lambda < 2$ and using Theorem 5.1, it is easy to see that

$$M_{P,m^{-1}b,k}(x) = O_{k,B}\left(\sum_{d\le x} \frac{1}{(1+\log d)^{\lambda-1}} \right)$$

for $x \ge 1$. Therefore, we obtain

$$|M(\alpha, Q, b, k, B) - \alpha N_{Q,b,k,B}| \le C_{k,B} \sum_{m\le Q} \sum_{d\le \frac{Q}{m}} \frac{1}{(1+\log d)^{\lambda-1}}$$

$$= C_{k,B} \sum_{d\le Q} \frac{1}{(1+\log d)^{\lambda-1}} \sum_{m\le \frac{Q}{d}} 1 \le C_{k,B}\, Q \sum_{d\le Q} \frac{1}{d(1+\log d)^{\lambda-1}} = O_{k,B}\left(Q(\log Q)^{2-\lambda} \right)$$

for any $Q \ge 2$, where $C_{k,B} > 0$ is a constant depending only on k and B. If $N_{Q,b,k,B} \ge 1$, then using Lemma 5.6, we see that

$$R_{N_{Q,b,k,B}}(\alpha) = \left| \frac{M(\alpha, Q, b, k, B)}{N_{Q,b,k,B}} - \alpha \right| = O_{k,B}\left(\frac{(\log Q)^{2-\lambda}}{Q} \right)$$

holds uniformly in α. Taking supremum over all $\alpha \in [0, 1]$,

$$D_{N_{Q,b,k,B}}(\mathfrak{F}_{Q,b,k,B}) = O_{k,B}\left(\frac{(\log Q)^{2-\lambda}}{Q} \right)$$

follows for any $Q \geq 2$ with $N_{Q,b,k,B} \geq 1$. The proofs of the other statements in Corollary 5.3 are entirely similar. Therefore, we omit the details. $\square$

5.6 Proof of Corollary 5.4

Assume that $B(x) \ll \frac{x}{(\log x)^\lambda}$ with $1 < \lambda < 2$ and $1 \leq k \leq (\log Q)^{2-\lambda}$. We again have, for any $\alpha \in [0,1]$, that

$$|M(\alpha, Q, b, k, B) - \alpha N_{Q,b,k,B}| \leq \sum_{\substack{m \leq Q \\ m \in \langle P \rangle \\ (m,k)=1}} \left| M_{P,m^{-1}b,k}\left(\frac{Q}{m}\right) \right|. \tag{5.70}$$

Split the range of the sum on the right side of (5.70) taking into account the m's that are close to Q. Note that if

$$m \leq \frac{Q}{e^{(\log Q)^{2-\lambda}}},$$

then $1 \leq k \leq (\log Q)^{2-\lambda} \leq \log\left(\frac{Q}{m}\right)$. Therefore, we may apply Theorem 5.2 to get

$$M_{P,m^{-1}b,k}\left(\frac{Q}{m}\right) = O_B\left(\frac{Q}{m\left(\log\left(\frac{Q}{m}\right)\right)^{\lambda-1}}\right) = O_B\left(\sum_{d \leq \frac{Q}{m}} \frac{1}{(1+\log d)^{\lambda-1}}\right). \tag{5.71}$$

It follows from (5.71) that

$$\begin{aligned} \sum_{\substack{m \leq \frac{Q}{\exp((\log Q)^{2-\lambda})} \\ m \in \langle P \rangle \\ (m,k)=1}} \left| M_{P,m^{-1}b,k}\left(\frac{Q}{m}\right) \right| &\leq C_B \sum_{m \leq \frac{Q}{\exp((\log Q)^{2-\lambda})}} \sum_{d \leq \frac{Q}{m}} \frac{1}{(1+\log d)^{\lambda-1}} \\ &\leq C_B\, Q(\log Q)^{2-\lambda} = O_B\left(Q(\log Q)^{2-\lambda}\right) \end{aligned} \tag{5.72}$$

for any $Q \geq 3$, where $C_B > 0$ is a constant depending only on B. If

$$\frac{Q}{e^{(\log Q)^{2-\lambda}}} < m \leq Q,$$

then we estimate trivially to get

$$\left| M_{P,m^{-1}b,k}\left(\frac{Q}{m}\right)\right| \leq \frac{Q}{m}.$$

In this way one obtains

$$\sum_{\substack{\frac{Q}{\exp\left((\log Q)^{2-\lambda}\right)}<m\leq Q \\ m\in\langle P\rangle \\ (m,k)=1}} \left| M_{P,m^{-1}b,k}\left(\frac{Q}{m}\right)\right| \leq Q \sum_{\frac{Q}{\exp\left((\log Q)^{2-\lambda}\right)}<m\leq Q} \frac{1}{m}. \tag{5.73}$$

Using the well known asymptotics of the harmonic series, we have

$$\sum_{\frac{Q}{\exp\left((\log Q)^{2-\lambda}\right)}<m\leq Q} \frac{1}{m} = \sum_{m\leq Q}\frac{1}{m} - \sum_{m\leq \frac{Q}{\exp\left((\log Q)^{2-\lambda}\right)}} \frac{1}{m} \tag{5.74}$$

$$= \log Q - \log\left(\frac{Q}{e^{(\log Q)^{2-\lambda}}}\right) + O\left(\frac{e^{(\log Q)^{2-\lambda}}}{Q}\right) = O\left((\log Q)^{2-\lambda}\right).$$

Combining (5.73) and (5.74), we see that

$$\sum_{\substack{\frac{Q}{\exp\left((\log Q)^{2-\lambda}\right)}<m\leq Q \\ m\in\langle P\rangle \\ (m,k)=1}} \left| M_{P,m^{-1}b,k}\left(\frac{Q}{m}\right)\right| = O\left(Q(\log Q)^{2-\lambda}\right), \tag{5.75}$$

where the implied constant in (5.75) is absolute. As a result of (5.70), (5.72) and (5.75), we deduce that

$$|M(\alpha,Q,b,k,B) - \alpha N_{Q,b,k,B}| = O_B\left(Q(\log Q)^{2-\lambda}\right)$$

for $Q \geq 3$. Therefore, if $Q \geq 3$ and $N_{Q,b,k,B} \geq 1$, then since $1 < \lambda < 2$ and $1 \leq k \leq (\log Q)^{2-\lambda} = O(\log Q)$, we may apply Lemma 5.6 to obtain

$$R_{N_{Q,b,k,B}}(\alpha) = \left|\frac{M(\alpha,Q,b,k,B)}{N_{Q,b,k,B}} - \alpha\right| = O_B\left(\frac{k(\log Q)^{2-\lambda}}{Q}\right)$$

uniformly in α. Taking supremum over all $\alpha \in [0,1]$, one finally arrives at the desired estimate

$$D_{N_{Q,b,k,B}}(\mathfrak{F}_{Q,b,k,B}) = O_B\left(\frac{k(\log Q)^{2-\lambda}}{Q}\right)$$

for the absolute discrepancy. The proofs of the other statements in Corollary 5.4 are similar. In the case when $B(x) \ll \frac{x}{\log^2 x}$ and $1 \le k \le (\log\log Q)^2$, we split the range $m \le Q$ as $m \le \frac{Q}{e^{(\log\log Q)^2}}$ and $\frac{Q}{e^{(\log\log Q)^2}} < m \le Q$. When $B(x) \ll \frac{x}{(\log x)^\lambda}$ with $\lambda > 2$ and $1 \le k \le \log\log Q$, we split as $m \le \frac{Q}{\log Q}$ and $\frac{Q}{\log Q} < m \le Q$. This completes the proof of Corollary 5.4. □

References

[1] E. Alkan, *Nonvanishing of Fourier coefficients of modular forms*, Proc. Amer. Math. Soc. **131** (2003), 1673-1680.

[2] E. Alkan, *On the sizes of gaps in the Fourier expansion of modular forms*, Canad. J. Math. **57** (2005), 449-470.

[3] E. Alkan, *Average size of gaps in the Fourier expansion of modular forms*, Int. J. Number Theory **3** (2007), 207-215.

[4] E. Alkan, A. H. Ledoan, M. Vajaitu, A. Zaharescu, *Discrepancy of fractions with divisibility constraints*, Monatsh. Math. **149** (2006), 179-192.

[5] E. Alkan, M. Xiong, A. Zaharescu, *Pair correlation of torsion points on elliptic curves*, J. Math. Anal. Appl. **356** (2009), 752-763.

[6] E. Alkan, M. Xiong, A. Zaharescu, *Pair correlation of sums of rationals with bounded height*, J. Reine. Angew. Math. **641** (2010), 21-67.

[7] T. M. Apostol, Introduction to Analytic Number Theory, Undergraduate Text in Mathematics, Springer-Verlag, New York, 1976.

[8] A. Beurling, *Analyse de la loi asymptotique de la distribution des nombres premiers généralisés, I*, Acta Math. **68** (1937), 255-291.

[9] F. P. Boca, C. Cobeli, A. Zaharescu, *On the distribution of the Farey sequence with odd denominators*, Michigan Math. J. **51** (2003), 557-574.

[10] E. Bombieri, *On twin almost primes*, Acta. Arith. **28** (1975), 177-193.

[11] E. Bombieri, *The asymptotic sieve*, Mem. Acad. Naz. dei XL, **1/2** (1976), 243-269.

[12] H. Davenport, Multiplicative Number Theory, Graduate Texts in Mathematics, Third edition, Springer, New York, 2000.

[13] H. G. Diamond, *A set of generalized numbers showing Beurling's theorem to be sharp*, Illinois J. Math. **14** (1970), 29-34.

[14] P. Erdös, *On a new method in elementary number theory which leads to an elementary proof of the prime number theorem*, Proc. Natl. Acad. Sci. USA **35** (1949), 374-384.

[15] P. Erdös, *On the difference of consecutive terms of sequences defined by divisibility properties*, Acta Arith. **12** (1966), 175-182.

[16] P. Erdös, M. Kac, E. R. Van Kampen, A. Wintner, *Ramanujan sums and almost periodic functions*, Studia Math. **9** (1940), 43-53.

[17] K. Ford, *On Bombieri's asymptotic sieve*, Trans. Amer. Math. Soc. **357** (2005), 1663-1674.

[18] J. Franel, *Les suites de Farey et les problèmes des nombres premiers*, Nachr. Ges. Wiss. Göttingen, Math. Phys. KL. (1924), 198-201.

[19] R. S. Hall, *Theorems about Beurling's generalized primes and the associated zeta function*, Ph. D. Thesis, Univ. of Illinois, Urbana, Illinois, 1967.

[20] A. Haynes, *A note on Farey Fractions with odd denominators*, J. Number Theory **98** (2003), 89-104.

[21] A. Haynes, *The disribution of special subsets of the Farey sequence*, J. Number Theory **107** (2004), 95-104.

[22] J. Knopfmacher, Abstract Analytic Number Theory, Second Edition, New York, Dover, 1990.

[23] N. M. Korobov, *Weyl's estimates of sums and the distribution of primes*, Dokl. Akad. Nauk SSSR **123** (1958), 28-31.

[24] E. Landau, *Neuer Beweis des Primzahlsatzes und Beweis des Primidealsatzes*, Math. Ann. **56** (1903), 645-670.

[25] E. Landau, *Bemerkungen zu derr vorstehenden Abhandlung von Herrn Franel*, Nachr. Ges. Wiss. Göttingen, Math. Phys. KL. (1924), 202-206.

[26] H. L. Montgomery, R. C. Vaughan, Multiplicative Number Theory, I. Classical Theory, Cambridge University Press, 2006.

[27] W.Rudin, Real and Complex Analysis, Third Edition, McGraw-Hill International Editions, Singapore, 1986.

[28] A. Selberg, *On an elementary method in the theory of primes*, Norske Vid. Selsk. Forh., Trandhjem **19** (1947), 64-67.

[29] A. Selberg, *An elementary proof of the prime-number theorem*, Ann. Math. **50** (1949), 305-313.

[30] E. C. Titchmarsh,The theory of Functions, Oxford University Press, Second Edition, 1997.

[31] E. C. Titchmarsh, The theory of the Riemann zeta-function, Second Edition, Oxford University Press, New York, 1986.

[32] E. C. Titchmarsh, Introduction to the theory of Fourier integrals, Oxford Clarendon Press, Second Edition, 1962.

[33] I. M. Vinogradov, *A new evaluation of* $\zeta(1+it)$, Izv. Akad. Nauk SSSR **22** (1958), 161-164.

[34] W. B. Zhang, *A generalization of Halász theorem to Beurling's generalized integers and its application*, Illinois J. Math. **31** (1987), 645-664.