VARIATIONS ON ERDOS-KAC THEOREM

by

Yıldırım Akbal

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

Yıldırım Akbal

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) ....

Assoc. Prof. Mine Çağlar ....

Assoc. Prof. Ünsal Tekir ....

Date: 05 August 2010

ABSTRACT

In this thesis, we study some distribution results concerning prime divisors of arithmetical functions. More precisely, we present two functions of which the distribution of the number of distinct primes divisors obeys a normal law.

ÖZET

Bu çalışmada bazı özel sayı-teoritik fonksiyonların asal bölenlerinin sayısının dağılımı incelenmiştir. Daha kesin olarak, iki farklı aritmetik fonksiyon incelenmiş olup, bu fonksiyonların asal bölenlerinin sayısının normal dağıldığı gösterilmiştir.

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my supervisor Assoc. Dr. Emre Alkan. His guidance through the work, insightful suggestions and his patience mostly made this study come along.

I would like to thank the jury members for serving on the thesis committee.

I would like to thank my family and friends for their encouragements and supports.

Finally, I would like to thank my wife Burcu for her invaluable support through writing this thesis.

Contents

LIST OF SYMBOLS/ABBREVIATIONS

$d(n)$	The number of divisors of n.
$li(x)$	$\int_2^x \frac{du}{\log u}$; the logarithmic integral.
$\omega(n)$	The number distinct prime divisors of n.
$\omega_y(n)$	The number distinct prime divisors of n which are $\leq y$.
$\Omega(n)$	The number distinct prime divisors of n. counted with multiplicity.
$\Omega_y(n)$	The number distinct prime divisors of n which are $\leq y$, counted with multiplicity.
$\phi_k(n)$	The k-fold iterate Euler ϕ function
$J_k(n)$	The Jordan's totient function
$\pi(x)$	The number of primes $\leq x$

ll

$\pi(x;q,a)$	The number of primes $\leq x$ which are $\equiv a \pmod{q}$.
$\Phi(a,b)$	The value of the integral $\frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{t^2}{2}}dt$.
$\Phi(z)$	The value of the integral $\frac{1}{\sqrt{2\pi}}\int_{-\infty}^z e^{-\frac{t^2}{2}}dt$.
$\log_k x$	k-fold iterate of $\log x$.
$f(x) = O(g(x))$	$\lvert f(x)\rvert \leq Cg(x)$ where C is an absolute constant.
$f(x) = o(g(x))$	$\lim f(x)/g(x) = 0$.
$f(x) \ll g(x)$	$f(x) = O(g(x))$.
$f(x) \sim g(x)$	$\lim f(x)/g(x) = 1$.
PNT	The Prime Number Theorem.

1 Introduction and Statement of Results

In 1917 Hardy and Ramanujan wrote a fundamental paper [6] on the distribution of the number of distinct prime divisors of a given integer. Among other things they showed that the number of distinct divisors of a given integer n is about $\log\log n$. It was Turan [16] who refined Hardy's and Ramanujan's result by showing that

$$\frac{1}{x}\sum_{n\leq x}(\omega(n)-\log\log x)^2 \ll \log\log x,$$

from which the result of Hardy and Ramanujan follows immediately.
Underlying ideas behind this theorem form another branch of mathematics so called Probabilistic Number Theory. Subsequently, In 1940 Erdös and Kac, using central limit theorem and Brun Sieve, showed that the quantity

$$\frac{\omega(n)-\log\log n}{\sqrt{\log\log n}}$$

is normally distributed. More precisely they proved the following

$$\lim_{x\to\infty}\frac{1}{x}\sharp\left\{n\leq x \mid \frac{\omega(n)-\log\log n}{\sqrt{\log\log n}}<z\right\}=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z}e^{-\frac{t^2}{2}}dt.$$

Moreover, using probabilistic arguments elaborately, Kubilius generalized their theorem by showing

$$\lim_{x\to\infty}\frac{1}{x}\sharp\left\{n\leq x \mid \frac{f(n)-A(x)}{\sqrt{B(x)}}<z\right\}=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z}e^{-\frac{t^2}{2}}dt,$$

provided that $f(n)$ is strongly additive and satisfies some certain conditions.
Yet another variant of the above theorem was given by Erdös and Pomerance [3] showing that actually, one can take $f(n)=\omega(\phi(n))$.
In this thesis, we generalize their result in two different directions.
In the third chapter, we shall study the article [14], and give full detailed proofs, which basically shows that $\omega(\phi_k(n))$ and $\omega(\phi_k(p-1))$ obey Gaussian distribution, where $\phi_k(n):=\phi_{k-1}(\phi(n))$.
More precisely,

Theorem. *For each fixed integer k, let $a_k = 1/(k+1)!$ and $b_k = 1/\sqrt{2k+1}(k!)$. Then for each real number z*

$$\lim_{x\to\infty} \frac{1}{x}\sharp\left\{n \leq x \mid \frac{\omega(\phi_k(n)) - a_k(\log\log x)^{k+1}}{b_k(\log\log x)^{k+1/2}} < z\right\} = \Phi(z).$$

and

$$\lim_{x\to\infty} \frac{1}{\pi(x)}\sharp\left\{p \leq x \mid \frac{\omega(\phi_k(p-1)) - a_k(\log\log x)^{k+1}}{b_k(\log\log x)^{k+1/2}} < z\right\} = \Phi(z).$$

In the fourth chapter, Inspired by the article [5], we obtain a similar result for $\omega(J_k(n))$.
More precisely, we shall show that;

Theorem. *Let $k \geq 1$ be fixed, then for any real u one has*

$$\lim_{x\to\infty} \frac{1}{x}\left\{n \leq x \mid \frac{\omega(J_k(n)) - \frac{d(k)}{2}(\log\log n)^2}{d(k)(\log\log n)^{3/2}} \leq \frac{z}{\sqrt{3}}\right\} = \Phi(z)$$

as $x \to \infty$.

2 Preliminaries

This chapter includes some basic information and motivation to study distribution of additive functions via probabilistic methods.

2.1 Arithmetic Functions

Definition 1. *A real -or complex-valued function defined on the positive integers is called an arithmetic function.*

We introduce some arithmetic functions which play an important role in this thesis.

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

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

2. if $k \geq 1$ the k-fold iterate Euler totient ϕ_k is defined as follows,

$$\phi_k(n) = \phi_{k-1}(\phi(n))$$

3. The Jordan's totient function $J_k(n)$ of a positive integer n is defined to be the number of k-tuples of positive integers all less than or equal to n that form a co-prime $(k+1)$-tuple together with n; i.e.,

$$J_k(n) = n^k \prod_{p|n} (1 - \frac{1}{p^k}).$$

4. The small omega function which counts the number of distinct prime divisors of a given integer is defined as

$$\omega(n) = \sum_{p|n} 1.$$

5. The big omega function which counts the number of distinct prime divisors of a given integer with multiplicity is defined as

$$\Omega(n) = \sum_{p^\alpha|n} 1.$$

Definition 2. *An arithmetic function f is said to be additive (resp., strongly additive) if it satisfies the following conditions. If $(n, m) = 1$*

1. $f(nm) = f(n) + f(m)$

2. $f(p^\alpha) = f(p)$

for all primes p *and* $\alpha \in \mathbb{N}$.

Note that the function ω is strongly additive and the function Ω is yet additive.

Definition 3. *An arithmetic function* f *is said to be multiplicative (resp., completely multiplicative) if it satisfies the following conditions.*

1. $f(nm) = f(n)f(m)$ *for all* $(n, m) = 1$

2. $f(nm) = f(n)f(m)$ *for all* n, m

2.2 Technical Preparation

In this section, we give some theorems, without proof, that are frequently used in this thesis. The most of the proofs can be found in [2], [9] and [19].

Theorem 2.1 (The Partial Summation Formula)**.** *Let* x *and* y *be real numbers with* $0 < y < x$. *Let* $f(n)$ *be an arithmetic function with summatory function* $F(x)$ *defined by*

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

and $g(t)$ *be a function with a continuous derivative on* $[y, x]$. *Then,*

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

In particular, if $x \geq 2$ *and* $g(t)$ *is continuously differentiable on* $[1, x]$, *then*

$$\sum_{n \leq x} f(n)g(n) = F(x)g(x) - \int_1^x F(t)g'(t)dt.$$

Theorem 2.2. *(Mertens Prime Number Theorem) For* $x \geq 2$,

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

Theorem 2.3. *(PNT) If* $\pi(x)$ *denotes the number of primes* $\leq x$, *then*

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

Although, the error term in the PNT can unconditionally be taken of the form $x\exp(-(c\log x)^{1/2})$ for some $c>0$. However, the above form of PNT will suffice for our purposes.

Theorem 2.4. *(Prime Ideal Theorem) Denote by $\delta_f(p)$ the number of solutions modulo p the congruence $f(x)\equiv 0(\text{mod } p)$, where $f(x)\in\mathbb{Z}[x]$ is a polynomial with k-irreducible components. Then*

$$\sum_{p\le x}\delta_f(p)=k.lix+O(xe^{-c\sqrt{\log x}}).$$

Consequently, by partial summation

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

Theorem 2.5. *(Siegel-Walfisz Theorem) Let A be any real number then there is a constant $C(A)>0$ such that*

$$\pi(x;q,a)=\frac{li(x)}{\phi(q)}+O\left(x\exp\left(-C(A)(\log x)^{\frac{1}{2}}\right)\right)$$

uniformly for all $(a,q)=1$ and $q\le(\log x)^A$.

Theorem 2.6. *(The Brun-Titchmarsh inequality) Let a and k be coprime integers and let x be a positive real number such that $k<x$. then,*

$$\pi(x;k,a)\le\frac{2x}{\phi(k)\log(x/k)}.$$

Theorem 2.7. *(The Bombieri-Vinagradov Theorem) For any $A>1$ there exists $B=B(A)>0$ such that*

$$\sum_{d\le\frac{x^{1/2}}{(\log x)^B}} max_{y\le x}max_{(a,d)=1}|\pi(y;d,a)-\frac{\pi(y)}{\phi(d)}|\ll\frac{x}{(\log x)^B}.$$

Utilizing Siegel-Walfisz theorem as well as partial summation we have

Lemma 2.8. *If $2\le k\le x$, then*

$$\sum_{\substack{p\ge x\\ p\equiv l(\text{mod } k)}}\frac{1}{p}=\frac{\log\log x}{\phi(k)}+O\left(\frac{\log k}{\phi(k)}\right),$$

where the implied constant is uniform.

Lemma 2.9. *Let m be a nonnegative integer and δ a real number with $0<\delta\le 1/2$, then there is a number c depending on m but not on δ so that* *the inequality*

$$\sum_{x^{1-\delta}<p\leq x} p^{m-1}\pi(x;p,1)^m \leq c\left(\frac{x}{\log x}\right)^m$$

holds for sufficiently large values of x.

We have yet another modification of the above lemma.

Lemma 2.10. *Let m be a nonnegative integer, there is c such that, for sufficiently large x we have,*

$$\sum_{x^{1/2}<pq\leq x} pq^{m-1}\pi(x;pq,1)^m \leq c\left(\frac{x}{\log x}\right)^m \log\log x.$$

2.3 A Quick Introduction to Probabilistic Number theory

In this section, we tackle the problem "How probabilistic arguments can be used to treat the value distribution of arithmetical functions?"

To illustrate the problem we begin with the method of Turan. Let us consider the average order of the function $\omega(n)$,

$$\begin{aligned}\sum_{n\leq x}\omega(n) &= \sum_{n\leq x}\sum_{p|n} \\ &= \sum_{pk\leq x} \\ &= \sum_{p\leq x}[x/p] \\ &= x\sum_{p\leq x}\frac{1}{p} + O(\pi(x)) \\ &= x\log\log x + O(x), \end{aligned} \tag{2.1}$$

where the last part follows by Theorem (2.2).

We now consider the second moment of $\omega(n)$.

$$\begin{aligned}
\sum_{n\le x}\omega^2(n) &= \sum_{n\le x}\sum_{\substack{p|n\\ q|n}} \\
&= \sum_{n\le x}\sum_{\substack{p|n\\ q|n\\ p\ne q}} + \sum_{n\le x}\sum_{p|n} \\
&= \sum_{pq\le x}[x/pq] + O(x\log\log x) \\
&= x\sum_{pq\le x}\frac{1}{pq} + O(x\log\log x) \qquad (2.2)
\end{aligned}$$

Now by means of the following elementary estimation,

$$\left(\sum_{p\le\sqrt{x}}\frac{1}{p}\right)^2 \le \sum_{pq\le x}\frac{1}{pq} \le \left(\sum_{p\le x}\frac{1}{p}\right)^2$$

it follows that

$$\sum_{pq\le x}\frac{1}{pq} = (\log\log x)^2 + O(\log\log x). \qquad (2.3)$$

Therefore combining (2.1), (2.2) and (2.3), we have

$$\sum_{n\le x}(\omega(n) - \log\log x)^2 \ll x\log\log x, \qquad (2.4)$$

Using the inequality $(a+b)^2 \ll a^2+b^2$ for all a and b real, one has

$$\begin{aligned}
\sum_{n\le x}(\omega(n)-\log\log n)^2 \ll& \sum_{n\le x}(\omega(n)-\log\log x)^2 + \sum_{n\le x}(\log\log x - \log\log n)^2 \\
=& x\log\log x + \sum_{n\le\sqrt{x}}(\log\log x-\log\log n)^2 \\
&+ \sum_{\sqrt{x}<n\le x}(\log\log x - \log\log n)^2 \\
\ll& x\log\log x + \sqrt{x}(\log\log x)^2 + x \\
\ll& x\log\log x,
\end{aligned}$$

from which it is clear that for fixed $\delta > 0$, the number of integers not satisfying the inequality

$$|\omega(n) - \log\log n| < (\log\log x)^{1/2+\delta} \qquad (2.5)$$

is at most

$$\ll \frac{x}{(\log\log x)^{2\delta}} = o(x).$$

Hence for almost all $n \leq x$ one has

$$|\omega(n) - \log\log n| < (\log\log x)^{1/2+\delta}. \tag{2.6}$$

Without loss of generality we may suppose that $\sqrt{x} < n \leq x$, then for a fixed $\varepsilon > 0$, it follows from (1.6) that

$$(1-\varepsilon)\log\log n < \omega(n) < (1+\varepsilon)\log\log n$$

for almost all $n \leq x$. From the text we see that, the LHS of (2.4) can be viewed as the variance of the variable ω, and the inequality (2.6) resembles Chebyshev's inequality in the theory of probability.

The analogue of the inequality (2.4) for a general additive function is given in Theorem 2.11.

It was the turn of Erdos and Kac to put all these ideas into a more precise probabilistic language. In [3], the main purpose is to approximate the function ω as a sum of independent random variables in order to apply central limit theorem of probability stating that sum of independent random variables tends to Gaussian distribution.

More precisely, we define

$$\omega_y(n) = \sum_{p \leq y} \rho_p(n) \tag{2.7}$$

where

$$\rho_p(n) = \begin{cases} 1 & \text{if } p \mid n \\ 0 & \text{if } p \nmid n. \end{cases}$$

Since the $\rho_p(n)$ are statistically independent the function $\omega_y(n)$ behaves like a sum of independent random variables. Therefore central limit theorem of probability can be applied. Finally using sieve methods elaborately they approximate $\omega(n)$ by $\omega_y(n)$ to deduce the desired result. The general case for strongly additive functions is proven by Kubulius [8] in which the theory of probability is purely served as the main tool.

Moreover, there is another approach which was first discovered by Halberstam [7] so called "Method of Moments", in which he estimates sums of the form:

$$\sum_{p \leq x} (\omega(f(p)) - \log\log x)^k \ , \ \sum_{n \leq x} (\omega(n) - \log\log x)^k$$

where $f(x)$ is an irreducible polynomial with integer coefficients. Among other things, he showed that

$$\lim_{x\to\infty} \frac{1}{\pi(x)} \sharp \left\{ p \le x \mid \frac{\omega(p-1) - \log\log x}{\sqrt{\log\log x}} < z \right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{t^2}{2}} dt.$$

Finally, the last and less probabilistic method is due to Turan and Renyi [17] in which they use Levy's continuity theorem to reduce the problem to an analytic object. More precisely, they estimate the sum

$$\phi_x(\tau) := \frac{1}{x} \sum_{n \le x} exp(\frac{i\tau}{\sqrt{\log\log x}}(\omega(n) - \log\log x)),$$

which is so called characteristic function of the frequencies

$$\frac{1}{x} \sharp \left\{ n \le x \mid \frac{\omega(n) - \log\log x}{\sqrt{\log\log x}} < z \right\}.$$

Using the theory of Riemann zeta function and Perron type formulas, they deduce the following

$$\phi_x(\tau) \sim e^{-\frac{\tau^2}{2}}, \text{ as } x \to \infty \ .$$

Finally, applying Levy's continuity theorem with Berry-Esseen inequality, they deduce the desired result with the best possible error term.
For further details the reader may refer to the monograph of Elliot [15].

2.4 Probabilistic Lemmas

In this section, we give several lemmas which will be frequently used in later chapters. For the sake of completeness, we give the proof of those that can not be found in the references. Throughout this section, we keep the notation used in Theorem 2.12.

Theorem 2.11. *(Turan Kubulius inequality) Let f be a complex valued additive function for all real numbers $x > 0$, set*

$$E(x) = \sum_{p^k \le x} \frac{f(p^k)}{p^k} \left(1 - \frac{1}{p}\right),$$

$$V(x) = \left(\sum_{p^k \le x} \frac{|f(p^k)|^2}{p^k} \right)^{1/2}.$$

Then we have the inequality

$$\sum_{n \le x} |f(n) - E(x)|^2 \le 32xV^2(x). \tag{2.8}$$

Proof. [15] Lemma 4.1. □

Theorem 2.12. *(Kubilius, Shapiro) Let f be a strongly additive function i.e., $f(p^\alpha) = f(p)$. Set*

$$A(x) = \sum_{p \le x} \frac{f(p)}{p}, \tag{2.9}$$

$$B(x) = \left(\sum_{p \le x} \frac{f(p)^2}{p} \right)^{1/2}. \tag{2.10}$$

Assume that f satisfies the following condition

$$\lim_{x \to \infty} \frac{B(x^y)}{B(x)} = 1 \text{ for all } y > 0 \ .$$

Suppose for each $\varepsilon > 0$, we have the following

$$\lim_{x \to \infty} \frac{1}{B^2(x)} \sum_{\substack{p \le x \\ |f(p)| > \varepsilon B(x)}} \frac{f(p)^2}{p} = 0.$$

Then we have

$$\sharp \left\{ n \le x : \alpha \le \frac{f(n) - A(x)}{B(x)} \le \beta \right\} \sim x\Phi(\alpha, \beta) \text{ as } x \to \infty.$$

Proof. See [15] Lemma 12.2. □

Theorem 2.13. *(Barban, Vinagradov, Levin) Let $f(n)$ be a strongly additive function such that*

$$\lim_{x \to \infty} \frac{B(x^y)}{B(x)} = 1 \text{ for all } y > 0 \ ,$$

in order that,

$$\lim_{x \to \infty} \frac{1}{\pi(x)} \sharp \left\{ p \le x : \alpha \le \frac{f(p+1) - A(x)}{B(x)} \le \beta \right\} \sim \Phi(\alpha, \beta)$$

it is sufficient that for each $\varepsilon > 0$

$$\sum_{\substack{p \leq x \\ f(p) > \varepsilon B(x)}} \frac{f^2(p)}{p} = o(B^2(x)).$$

Remark: This theorem is valid if $p+1$ *is replaced by* $p+l$ *for some* l *fixed non-zero integer.*

Proof. [15] Lemma 12.4. □

Lemma 2.14. *Let* f, g *and* h *be arithmetical functions, and suppose that for all* n *we have* $|f(n)-g(n)| \leq h(n)$. *And assume that the following condition is satisfied*

$$\sum_{n \leq x} h(n) = o(xB(x)). \tag{2.11}$$

Suppose that for all constants α *and* β *(with* $\alpha \geq \beta$*) we have*

$$\sharp \left\{ n \leq x : \alpha \leq \frac{f(n) - A(x)}{B(x)} \leq \beta \right\} \sim x\Phi(\alpha, \beta), \tag{2.12}$$

as $x \to \infty$, *then also*

$$\sharp \left\{ n \leq x : \alpha \leq \frac{g(n) - A(x)}{B(x)} \leq \beta \right\} \sim x\Phi(\alpha, \beta) \tag{2.13}$$

as $x \to \infty$.

Proof. Let us define the functions

$$S_f(x, \alpha, \beta) := \sharp \left\{ n \leq x : \alpha \leq \frac{f(n) - A(x)}{B(x)} \leq \beta \right\},$$
$$S_g(x, \alpha, \beta) := \sharp \left\{ n \leq x : \alpha \leq \frac{g(n) - A(x)}{B(x)} \leq \beta \right\}$$

Fix $\varepsilon > 0$, then by the assumption (2.11), we may suppose that $h(n) \leq \varepsilon B(x)$ for almost all $n \leq x$. Thus,

$$a \leq \frac{f(n) - A(x)}{B(x)} \leq b \Rightarrow A(x) + \alpha B(x) \leq f(n) \leq A(x) + \beta B(x).$$

Together with the assumption (2.12) and $g(n) - h(n) \leq f(n) \leq g(n) + h(n)$, for almost all $n \leq x$ we have

$$\alpha - \varepsilon \leq \frac{g(n) - A(x)}{B(x)} \leq \beta + \varepsilon.$$

Therefore for all $\alpha \leq \beta$

$$S_f(x, \alpha, \beta) \leq S_g(x, \alpha - \varepsilon, \beta + \varepsilon) + o(x). \tag{2.14}$$

In other words, replacing α by $\alpha + \varepsilon$ and β by $\beta - \varepsilon$ we have

$$S_f(x, \alpha + \varepsilon, \beta - \varepsilon) \leq S_g(x, \alpha, \beta) + o(x).$$

Similarly,

$$\alpha \leq \frac{g(n) - A(x)}{B(x)} \leq \beta \Rightarrow A(x) + \alpha B(x) \leq g(n) \leq A(x) + \beta B(x),$$

which implies for almost all $n \leq x$

$$\alpha - \varepsilon \leq \frac{f(n) - A(x)}{B(x)} \leq \beta + \varepsilon.$$

Then

$$S_g(x, \alpha, \beta) \leq S_f(x, \alpha + \varepsilon, \beta - \varepsilon) + o(x). \tag{2.15}$$

Let us now consider

$$\begin{aligned}\Phi(\alpha, \beta) - \Phi(\alpha + \varepsilon, \beta - \varepsilon) =& \frac{1}{\sqrt{2\pi}} \int_{\alpha}^{\beta} e^{-\frac{t^2}{2}} dt - \frac{1}{\sqrt{2\pi}} \int_{\alpha+\varepsilon}^{\beta-\varepsilon} e^{-\frac{t^2}{2}} dt \\ =& \int_{\alpha}^{\alpha+\varepsilon} .. + \int_{\beta-\varepsilon}^{\beta} .. \\ \ll& \varepsilon,\end{aligned}$$

which follows by Mean-value theorem for integrals. Hence by (2.14) and (2.15) one has

$$x\Phi(\alpha, \beta) + o(\varepsilon x) \leq S_g(x, \alpha, \beta) \leq x\Phi(\alpha, \beta) + o(\varepsilon x),$$

from which we deduce that

$$\limsup |\frac{S_g(x, \alpha, \beta)}{x} - G(\alpha, \beta)| < \varepsilon.$$

□

Lemma 2.15. *Suppose that* $\mathbb{B} \subset \mathbb{N}$ *and let* $\mathbb{B}(x) = |\{n \in \mathbb{B}, n \leq x\}|$*, let* f,g *and* h *be arithmetical functions, suppose that for all* $n \in \mathbb{B}$ *we have* $|f(n) - g(n)| \ll h(n)$*, and assume that the following condition is satisfied*

$$\sum_{\substack{n \leq x \\ n \in B}} h(n) = o(\mathbb{B}(x)B(x)).$$

Suppose that for all constants α and β (with $\alpha \geq \beta$) we have

$$\sharp\{n \leq x, n \in \mathbb{B}(x) : \alpha \leq \frac{f(n) - A(x)}{B(x)} \leq \beta\} \sim \mathbb{B}(x)\Phi(\alpha, \beta))$$

as $x \to \infty$, then also

$$\sharp\left\{n \leq x, n \in \mathbb{B}(x) : \alpha \leq \frac{g(n) - A(x)}{B(x)} \leq \beta\right\} \sim \mathbb{B}(x)\Phi(\alpha, \beta)$$

as $x \to \infty$.

Proof. The proof is similar to that of previous lemma. We therefore skip it. $\square$

Lemma 2.16. *Let f be an arithmetical function, suppose that we have*

$$\lim_{x\to\infty} \frac{B(x^y)}{B(x)} = 1 \text{ for all } y > 0 \text{ and} \tag{2.16}$$

$$\lim_{x\to\infty} \frac{A(x) - A(x^y)}{B(x)} = 0 \text{ for some } 0 < y < 1. \tag{2.17}$$

Then,

$$\sharp\left\{n \leq x : \alpha \leq \frac{f(n) - A(x)}{B(x)} \leq \beta\right\} \sim x\Phi(\alpha, \beta) \text{ as } x \to \infty$$

implies

$$\sharp\left\{n \leq x : \alpha \leq \frac{f(n) - A(n)}{B(n)} \leq \beta\right\} \sim x\Phi(\alpha, \beta) \text{ as } x \to \infty\ .$$

Proof. Define

$$S_f(x, \alpha, \beta) := \sharp\left\{n \leq x : \alpha \leq \frac{f(n) - A(x)}{B(x)} \leq \beta\right\},$$

$$S'_f(x, \alpha, \beta) := \sharp\left\{n \leq x : \alpha \leq \frac{f(n) - A(n)}{B(n)} \leq \beta\right\}.$$

Suppose for $x^y < n \leq x$ and $0 < y < 1$ we have

$$\alpha \leq \frac{f(n) - A(n)}{B(n)} \leq \beta.$$

Let us write

$$\frac{f(n) - A(x)}{B(x)} = \frac{f(n) - A(n)}{B(n)} + \left(\frac{B(n)}{B(x)} - 1\right)\frac{f(n) - A(n)}{B(n)} - \frac{A(x) - A(n)}{B(x)}.$$

By the assumption (2.17) the last term on RHS tends to zero as $x \to \infty$ with $x^y < n \leq x$. And by the assumption (2.16), it is clear that third term tends to zero.
Consequently,

$$-\varepsilon + \alpha \leq \frac{f(n) - A(x)}{B(x)} \leq \beta + \varepsilon,$$

which implies

$$S'_f(x, \alpha, \beta) \leq S_f(x, \alpha - \varepsilon, \beta + \varepsilon) + o(x).$$

Similarly

$$S_f(x, \alpha, \beta) \leq S'_f(x, \alpha - \varepsilon, \beta + \varepsilon) + o(x).$$

Mimicking the proof of the previous lemma the desired result follows. □

3 Erdös-Kac Theorem for ϕ_k

It was a result of Halberstam [7] that

$$\lim_{x\to\infty} \frac{1}{\pi(x)} \sharp \left\{ p \le x \mid \frac{\omega(p-1) - \log\log x}{\sqrt{\log\log x}} < z \right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{t^2}{2}} dt. \quad (3.1)$$

and a result of Erdös and Pomerance[3] that

$$\lim_{x\to\infty} \frac{1}{x} \sharp \left\{ n \le x \mid \frac{\omega(\phi(n)) - 1/2(\log\log x)^{3/2}}{1/\sqrt{3}(\log\log x)^2} < z \right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{t^2}{2}} dt. \quad (3.2)$$

Their results was generalized by N.L.Bassily, I.Katai & M.Wijsmuller. [14] More precisely, they showed that

Theorem 3.1. *[14] For each fixed integer k, let $a_k = 1/(k+1)!$ and $b_k = 1/\sqrt{2k+1}(k!)$. Then for each real number z*

$$\lim_{x\to\infty} \frac{1}{x} \sharp \left\{ n \le x \mid \frac{\omega(\phi_k(n)) - a_k(\log\log x)^{k+1}}{b_k(\log\log x)^{k+1/2}} < z \right\} = \Phi(z).$$

Theorem 3.2. *[14] For each fixed integer $k \le x$, and a_k and b_k as in Theorem 3.1, then for each real number z*

$$\lim_{x\to\infty} \frac{1}{\pi(x)} \sharp \left\{ p \le x \mid \frac{\omega(\phi_k(p-1)) - a_k(\log\log x)^{k+1}}{b_k(\log\log x)^{k+1/2}} < z \right\} = \Phi(z).$$

First, in order to prove the above theorems we define the following auxiliary functions,

$$\vartheta(p) = p - 1.$$

and extend it to the whole $\mathbb{N}$ as a completely multiplicative function. And furthermore we denote by ϑ_k the k-fold iterate of ϑ with $\vartheta_0(n) = n$. Finally, we define the strongly additive function τ recursively as follows

$$\tau_0(p) = 1 \text{ and } \tau_k(p) = \sum_{q \mid p-1} \tau_{k-1}(q).$$

3.1 Preliminary Lemmas

Lemma 3.3. *For all $k = 0, 1, 2..$ we have*

$$\omega(\vartheta_k(n)) \leq \omega(\phi_k(n)) \leq \omega(n) + \omega(\vartheta_1(n)) + ...\omega(\vartheta_k(n)). \tag{3.3}$$

Proof. : (LHS) First, by induction on k, let us show that

$$a \mid b \text{ implies } \phi_k(a) \mid \phi_k(b). \tag{3.4}$$

The case $k = 1$ is obvious, and follows by the formula $\phi(p^\alpha) = p^{\alpha-1}(p-1)$. Suppose now,

$$a \mid b \text{ implies } \phi_{k-1}(a) \mid \phi_{k-1}(b).$$

Since

$$\phi_k(a) = \phi_{k-1}(\phi(a)),$$

by induction hypothesis, we have

$$a \mid b \Longrightarrow \phi(a) \mid \phi(b) \Longrightarrow \phi_{k-1}(\phi(a)) \mid \phi_{k-1}(\phi(b)) \Longrightarrow \phi_k(a) \mid \phi_k(b).$$

What's more, by induction on k, we will show that

$$\omega(\vartheta_k(p_1^{\alpha_1} p_2^{\alpha_2} ... p_n^{\alpha_n})) = \omega(\vartheta_k(p_1 p_2 ... p_n)). \tag{3.5}$$

It is obvious that the case $k = 1$ is satisfied. Now suppose we have (3.5) for k. Then

$$\omega(\vartheta_{k+1}(p_1^{\alpha_1} p_2^{\alpha_2} ... p_n^{\alpha_n})) = \omega(\vartheta_k((p_1 - 1)^{\alpha_1} (p_2 - 1)^{\alpha_2} ... (p_n - 1)^{\alpha_n})).$$

Let $q_i^{\beta_i}$ be the prime powers appearing in the factorization of $p_i - 1$ for $i = 1, 2..n$. Then, by induction hypothesis, we have

$$\begin{aligned}
\omega(\vartheta_{k+1}(p_1^{\alpha_1} p_2^{\alpha_2} ... p_n^{\alpha_n})) =& \omega(\vartheta_k(q_1^{\beta_1} q_2^{\beta_2} .... q_l^{\beta_l})) \\
=& \omega(\vartheta_k(q_1 q_2 .... q_l)) \\
=& \omega(\vartheta_k((p_1 - 1)(p_2 - 1)...(p_n - 1))) \\
=& \omega(\vartheta_{k+1}(p_1 p_2 ... p_n)).
\end{aligned}$$

Now, the inequality (3.3) is obvious for $k = 1$. Suppose we have the inequality (3.3) for $k-1$. Then by (3.4), (3.5) and the induction hypothesis

$$\begin{aligned}\omega(\phi_k(q_1^{\alpha_1}q_2^{\alpha_2}...q_l^{\alpha_l})) =&\omega(\phi_{k-1}(q_1^{\beta_1-1}(q_1-1)q_2^{\beta_2-1}(q_2-1)....q_l^{\beta_l-1}(q_l-1)))\\ \geq&\omega(\phi_{k-1}((q_1-1)(q_2-1)...(q_l-1))\\ \geq&\omega(\vartheta_{k-1}((q_1-1)(q_2-1)...(q_l-1))\\ =&\omega(\vartheta_k(q_1q_2...q_l))\\ =&\omega(\vartheta_k(n))).\end{aligned}$$

(RHS) We will proceed by induction on k. If we show that

$$p|\phi_k(n) \Rightarrow \vartheta_j(n) \text{ for some } j=0,1...k \ ,$$

then the result will follow.
It is obvious that

$$p|\phi_k(n) \Rightarrow p|\phi_{k-1}(n) \text{ or } p|\vartheta(\phi_k(n)).$$

If $p \mid \phi_{k-1}(n)$, then by induction hypothesis, $p \mid \vartheta_j(n)$ for some $j=0,1...k-1$.
Suppose now, $p \nmid \phi_{k-1}(n)$. Then there is some q dividing $\phi_{k-1}(n)$ such that $p|\vartheta(q)$, and by induction on k, $q|\vartheta_j(n)$ for $j=0,1...k-1$. Therefore, since ϑ is completely multiplicative, this implies $p|\vartheta_j(n)$ for some $j=0,1...k$, which proves the lemma.

□

Lemma 3.4. *For $H \geq 0$, $x \geq 3$ Let $T_k(x,H) := \sharp\{p \leq x | \tau_k(p) \geq H^k\}$. Then, there is an absolute constant $C_2 > 48$ depending on k such that*

$$T_k(x,H) \leq \frac{C_2^k x(\log\log x)^{k-1}}{2^H}. \tag{3.6}$$

Proof. It follows from the definition of $\tau_k(n)$ that $\tau_1(p)=\omega(p-1)$ and since $2^{\omega(n)} \leq d(n)$, we have

$$\sum_{p\leq x} 2^{\omega(p-1)} \leq \sum_{p\leq x} d(p-1).$$

The last sum well known to be $\leq cx$ for all $x \geq 1$ (See The Titchmarsh divisor problem, [12]). Therefore we have

$$T_1(x,H) = \sum_{\substack{\tau_1(p)>H\\ p\leq x}} = \sum_{\substack{\tau_1(p)>H\\ p\leq x}} \frac{2^{\omega(p-1)}}{2^{\omega(p-1)}} < \frac{cx}{2^H}.$$

Thus, (3.6) is true for $k = 1$. Suppose we have (3.6) for $k-1$. Note that if $\tau_k(p) > H^k$, and also if $\omega(p-1) \leq H$ and $\tau_{k-1}(q) \leq H^{k-1}$ for all primes dividing $q-1$, then

$$\tau_k(p) = \sum_{q|p-1} \tau_{k-1}(q) \leq \sum_{q|p-1} H^{k-1} \leq H^k,$$

which contradicts the assumption $\tau_k(p) > H^k$. Therefore either $\omega(p-1) > H$ or $\tau_{k-1}(q) > H^{k-1}$ holds for some $q|p-1$.
Hence it follows that

$$\begin{aligned} T_k(x,H) &\leq \sum_{\substack{\tau_1(p)>H \\ p\leq x}} + \sum_{\substack{p\leq x \\ \tau_{k-1}(q)>H^k \\ q|p-1 \\ \text{for some q}}} \\ &\leq \frac{cx}{2^H} + \sum_{p\leq x} \sum_{\substack{q|p-1 \\ \tau(q)>H^{k-1}}} \\ &= \frac{cx}{2^H} + \sum_{\substack{q\leq x \\ \tau_{k-1}(q)>H^{k-1}}} \pi(x;q,1). \end{aligned}$$

Dividing the interval $2 \leq q \leq x$ into dyadic intervals with $u = |\log x/\log 2| - 1$, using Brun Titchmarsh inequality , the trivial estimate (i.e, $\pi(x;a,q) \leq \frac{x}{q}$) as well as induction hypothesis and the inequality

$$\sum_{n\leq x} \frac{1}{n} \leq 3\log x \text{ for } x \geq 1 \ ,$$

it follows that

$$
\begin{aligned}
T_k(x,H) &\leq \frac{cx}{2^H} + \sum_{\substack{x/2<q\leq x\\ \tau_{k-1}(q)>H^{k-1}}} \pi(x,q,1) + \sum_{j=1}^{u} \sum_{\substack{x/2^{j+1}\leq q\leq x/2^j\\ \tau_{k-1}(q)>H^{k-1}}} \pi(x;q,1)\\
&\leq \frac{cx}{2^H} + \sum_{\substack{x/2<q\leq x\\ \tau_{k-1}(q)>H^{k-1}}} \frac{x}{q} + \sum_{j=1}^{u} \sum_{\substack{x/2^{j+1}\leq q\leq x/2^j\\ \tau_{k-1}(q)>H^{k-1}}} \frac{2x}{q-1\log\frac{x}{q}}\\
&\leq \frac{cx}{2^H} + \sum_{\substack{x/2<q\leq x\\ \tau_{k-1}(q)>H^{k-1}}} \frac{x}{q} + \sum_{j=1}^{u} \frac{2x}{jlog2} \sum_{\substack{x/2^{j+1}\leq q\leq x/2^j\\ \tau_{k-1}(q)>H^{k-1}}} \frac{1}{q-1}\\
&\leq \frac{cx}{2^H} + \frac{C_2^{k-1}x(\log\log x)^{k-2}}{2^H} + \sum_{j=1}^{u} \frac{42^{j+1}}{jlog2} \sum_{\substack{x/2^{j+1}\leq q\leq x/2^j\\ \tau_{k-1}(q)>H^{k-1}}} \frac{1}{q}\\
&\leq \frac{cx}{2^H} + \frac{C_2^{k-1}x(\log\log x)^{k-2}}{2^H} + \sum_{j=1}^{u} \frac{42^{j+1}}{log2} \sum_{\substack{x/2^{j+1}\leq q\leq x/2^j\\ \tau_{k-1}(q)>H^{k-1}}}\\
&\leq \frac{cx}{2^H} + \frac{C_2^{k-1}x(\log\log x)^{k-2}}{2^H} + \sum_{j=1}^{u} \frac{42^{j+1}}{log2} T_{k-1}(\frac{x}{2^j},H)\\
&\leq \frac{cx}{2^H} + \frac{C_2^{k-1}x(\log\log x)^{k-2}}{2^H} + \frac{24C_2^{k-2}x(\log\log x)^{k-1}}{2^H} \sum_{j=1}^{u} \frac{1}{j}\\
&\leq \frac{C_2^{k}x(\log\log x)^{k-1}}{2^H}.
\end{aligned}
$$

□

Remark 3.1. *Let us observe that if $\omega(p-1)=k$, then it is obvious that $2^k \leq p-1$. Therefore, $\omega(p-1)\leq 2\log p$ and by induction on k, one has*

$$\tau_k(p) \leq 2\log p. \tag{3.7}$$

Furthermore, since τ_k is strongly additive, one has

$$\tau_k(n) \leq 2\log n.$$

Let us choose $H \geq 10\log\log x$. Then by partial summation and by (3.7) it follows that

$$\sum_{\substack{q\leq x\\ \tau_k(q)>H^k}} \frac{\tau_k^j(q)}{q} \ll \frac{(C_2\log\log x)^k}{(\log x)^2}$$

for $j = 0, 1, 2, 4$.

3.2 The Moments of τ_k

Let $S_k(x) = \sum_{p\le x} \tau_k(p)$ and $A_k(x) = \sum_{p\le x} \frac{\tau_k(p)}{p}$

Lemma 3.5. *For every k=1,2.. we have*

$$S_k(x) = li(x)A_{k-1}(x) + O(li(x)(C\log\log x)^{k-1}). \tag{3.8}$$

Proof. Let $H = C\log\log x$ as in Lemma (3.4), denote by $\sum'$ the summation over primes with $\tau_k(q) \le H^k$ and by $\sum''$ the summation over primes with $\tau_k(q) > H^k$. Then, using the trivial estimation and Remark (3.1) one has

$$\begin{aligned} S_k(x) &= \sum_{p\le x}\sum_{q|p-1} \tau_{k-1}(q) \\ &= {\sum_{q\le x}}' \tau_{k-1}(q)\pi(x;q,1) + {\sum_{p\le x}}'' \tau_{k-1}(q)\pi(x;q,1) \\ &= {\sum_{q\le x}}' \tau_{k-1}(q)\pi(x;q,1) + O\left(\frac{x(C\log\log x)^{k-1}}{(\log x)^2}\right) \end{aligned} \tag{3.9}$$

We now split the above sum in (3.9) into two parts depending on whether $q > x^{1/3}$ or $q \le x^{1/3}$. Since $\tau_{k-1}(q) \le H^{k-1}$, by Lemma (2.9) and Brun-Titchmarsh inequality it follows that

$$\begin{aligned} {\sum_{q>x^{1/3}}}' \tau_{k-1}(q)\pi(x;q,1) &\ll (C\log\log x)^{k-1} \\ &\quad \left(\sum_{x^{1/3}<q\le x^{1/2}} \pi(x;q,1) + \sum_{q>x^{1/2}} \pi(x;q,1)\right) \\ &\ll (C\log\log x)^{k-1}\frac{x}{\log x}. \end{aligned}$$

Finally, applying Bombieri-Vinagradov Theorem we have

$$\begin{aligned} S_k(x) = &{\sum_{q\le x^{1/3}}}' \tau_{k-1}(q)(\pi(x;q,1) - \frac{li(x)}{\phi(q)}) + li(x){\sum_{q\le x^{1/3}}}' \frac{\tau_{k-1}(q)}{q-1} \\ &+ O\left(\frac{x}{\log x}(C\log\log x)^{k-1}\right) \\ =&li(x){\sum_{q\le x^{1/3}}}' \frac{\tau_{k-1}(q)}{q} + O\left(\frac{x(C\log\log x)^{k-1}}{\log x}\right). \end{aligned}$$

On the other hand, we have

$$\begin{aligned}
A_{k-1}(x) &= \sum_{p\le x} \frac{\tau_{k-1}(p)}{q} \\
&= \sum_{p\le x}{}' \frac{\tau_{k-1}(p)}{q} + \sum_{p\le x}{}'' \frac{\tau_{k-1}(p)}{q} \\
&= \sum_{p\le x}{}' \frac{\tau_{k-1}(p)}{p} + O\left(\frac{(C\log\log x)^{k-1}}{\log^2 x}\right) \\
&= \sum_{p\le x^{1/3}}{}' \frac{\tau_{k-1}(p)}{p} + \sum_{p> x^{1/3}}{}' \frac{\tau_{k-1}(p)}{p} + O\left(\frac{(C\log\log x)^{k-1}}{\log^2 x}\right) \\
&= \sum_{p\le x^{1/3}}{}' \frac{\tau_{k-1}(p)}{p} + O\left((C\log\log x)^{k-1} \sum_{p> x^{1/3}} \frac{1}{p}\right) \\
&= \sum_{p\le x^{1/3}}{}' \frac{\tau_{k-1}(p)}{p} + O((C\log\log x)^{k-1}).
\end{aligned}$$

□

Combining last two estimates gives the desired result.

Lemma 3.6. *For every $k = 0, 1, 2...$*

$$A_k(x) = \frac{1}{(k+1)!}(\log\log x)^{k+1} + O((C\log\log x)^k) \qquad (3.10)$$

Proof. We proceed by induction on k. The case $k = 0$ is due to Mertens(lemma (2.2)). Suppose that (3.10) is true for $k-1$ then by the previous Lemma we have

$$S_k(x) = \frac{x}{\log x}\frac{(\log\log x)^k}{k!} + O\left(\frac{x(C\log\log x)^{k-1}}{\log x}\right).$$

Therefore by partial summation

$$\begin{aligned}
A_k(x) &= \frac{S_k(x)}{x} + \int_{x_0}^{x} \frac{S_k(t)}{t^2}dt + O(1) \\
&= \frac{S_k(x)}{x} + \int_{x_0}^{x} \frac{(\log\log t)^k}{k!t\log t}dt + O\left(\int_{x_0}^{x} \frac{(C\log\log t)^{k-1}}{t\log t}dt\right) \\
&= \frac{(\log\log x)^{k+1}}{(k+1)!} + O((C\log\log x)^k)
\end{aligned}$$

and the proof is complete. □

Lemma 3.7. *for any* $k = 1, 2, 3...$ *Let*

$$D_k(x) = \sum_{p \le x} \tau_k^2(p)$$

then,

$$D_k(x) = \frac{li(x)(\log\log x)^{2k}}{(k!)^2} + O(li(x)(C\log\log x)^{2k-1/2})$$

Proof. First observe that, by remark (3.1) and the trivial estimation subject to the inequality $\tau_k(p) \le H^k$, it follows that

$$\begin{aligned} D_k(x) &= \sum_{\substack{p \le x \\ \tau_k(p) \le H^k}} \tau_k^2(p) + \sum_{\substack{p \le x \\ \tau_k(p) > H^k}} \frac{p\tau_k^2(p)}{p} \\ &= \sum_{\substack{p \le x \\ \tau_k(p) \le H^k}} \tau_k^2(p) + \frac{x(C\log\log x)^k}{\log^2 x} \\ &\ll (C\log\log x)^{2k} li(x). \end{aligned} \tag{3.11}$$

Let

$$\tau_k(p) = \sum_{q|p-1} \tau_{k-1}(q) = \sum_{\substack{q|p-1 \\ q \le x^{1/6}}} \tau_{k-1}(q) + \sum_{\substack{q|p-1 \\ q > x^{1/6}}} \tau_{k-1}(q) = f_1(p) + f_2(p)$$

and define $U(x) = \sum_{p \le x} f_1^2(p)$. Therefore

$$\begin{aligned} D_k(x) =& U(x) + \sum_{p \le x} f_1(p) f_2(p) + \sum_{p \le x} f_2^2(p) \\ =& U(x) + \sum_1 + \sum_2 . \end{aligned}$$

Hence we have

$$\sum_2 = \sum_{q\le x} \sum_{\substack{q_1>x^{1/6}\\ q_1|p-1}} \sum_{\substack{q_2>x^{1/6}\\ q_2|p-1}} \tau_{k-1}(q_1)\tau_{k-1}(q_2)$$

$$= \sum_{q\le x} \sum_{\substack{q>x^{1/6}\\ q|p-1}} \tau_{k-1}^2(q) + \sum_{q\le x} \sum_{\substack{q_1>x^{1/6}\\ q_1|p-1}} \sum_{\substack{q_2>x^{1/6}\\ q_2|p-1\\ q_1\neq q_2}} \tau_{k-1}(q_1)\tau_{k-1}(q_2)$$

$$= \sum_{q>x^{1/6}} \tau_{k-1}^2(q)\pi(x;q,1) + \sum_{\substack{q_1,q_2>x^{1/6}\\ q_1\neq q_2}} \tau_{k-1}(q_1)\tau_{k-1}(q_1)\pi(x;q_1q_2,1). \tag{3.12}$$

Splitting the first sum into two sums depending on whether $\tau_{k-1}(q) \gtreqless H^{k-1}$ and using remark (3.1), we have

$$\sum_{\substack{q>x^{1/6}\\ \tau_{k-1}(q)>H^{k-1}}} \tau_{k-1}^2(q)\pi(x;q,1) \le x \sum_{\substack{q>x^{1/6}\\ \tau_{k-1}(q)>H^{k-1}}} \frac{\tau_k^2(q)}{q} \ll \frac{x(C\log\log x)^{k-1}}{\log^2 x}$$

and that,

$$\begin{aligned}
\sum_{\substack{q\ge x^{1/6}\\ \tau_k(q)\le H^{k-1}}} \tau_{k-1}^2(q)\pi(x;q,1) = & \sum_{\substack{x^{1/2}\ge q\ge x^{1/6}\\ \tau_{k-1}(q)\le H^{k-1}}} \tau_{k-1}^2(q)\pi(x;q,1) \\
& + \sum_{\substack{q>x^{1/2}\\ \tau_{k-1}(q)\le H^{k-1}}} \tau_{k-1}^2(q)\pi(x;q,1) \\
\ll & (C\log\log x)^{2k-2} \\
& \times \Big(\sum_{x^{1/2}\ge q\ge x^{1/6}} \pi(x;q,1) + \sum_{q>x^{1/2}} \pi(x;q,1)\Big) \\
\ll & (C\log\log x)^{2k-2} \left(\sum_{x^{1/2}\ge q\ge x^{1/6}} \frac{x}{(q-1)\log\frac{x}{q}} + \frac{x}{\log x} \right) \\
\ll & li(x)(C\log\log x)^{2k-2}.
\end{aligned}$$

Finally, we have

$$\sum_{q>x^{1/6}} \tau_{k-1}^2(q)\pi(x;q,1) \ll li(x)(C\log\log x)^{2k-2}.$$

We now treat the second sum in (3.12) by beginning with the case $\tau_{k-1}(q_i) > H^{k-1}$.

$$\sum_{\substack{q_i \geq x^{1/6} \\ \tau_{k-1}(q_i) > H^{k-1}}} \tau_{k-1}(q_1)\tau_{k-1}(q_2)\pi(x, q_1q_2, 1) \ll x \sum_{\substack{q_i \geq x^{1/6} \\ \tau_{k-1}(q_i) > H^{k-1}}} \frac{\tau_{k-1}(q_1)\tau_{k-1}(q_2)}{q_1q_2}$$
$$\ll x \left(\sum_{\substack{q \geq x^{1/6} \\ \tau_{k-1}(q) > H^{k-1}}} \frac{\tau_{k-1}(q)}{q} \right)^2$$
$$\ll \frac{x(C \log\log x)^{2k-2}}{\log^2 x},$$

and the case $\tau_{k-1}(q_i) \leq H^{k-1}$

$$\sum_{\substack{q_i \geq x^{1/6} \\ \tau_{k-1}(q_i) \leq H^{k-1}}} \tau_{k-1}(q_1)\tau_{k-1}(q_2)\pi(x, q_1q_2, 1) \ll (C \log\log x)^{2k-2} \sum_{q_i > x^{1/6}} \pi(x; q_1q_2, 1)$$

therefore,

$$\sum_{q_i > x^{1/6}} \pi(x; q_1q_2, 1) \ll \sum_{q_1q_2 > x^{1/6}} \pi(x; q_1q_2, 1)$$
$$= \sum_{q_1q_2 > x^{1/2}} \pi(x; q_1q_2, 1) + \sum_{x^{1/2} \geq q_1, q_2 > x^{1/6}} \pi(x; q_1q_2, 1)$$
$$\ll \frac{x \log\log x}{\log x},$$

where the first inequality follows by Lemma (2.9) and the second by Brun-Titchmarsh inequality. Hence, it follows that

$$\sum_{\substack{q_i \geq x^{1/6} \\ \tau_{k-1}(q_i) \leq H^{k-1}}} \tau_{k-1}(q_1)\tau_{k-1}(q_2)\pi(x, q_1q_2, 1) \ll \frac{x(C \log\log x)^{2k-1}}{\log x}.$$

Finally the case $\tau_{k-1}(q_1) > H^{k-1}$ and $\tau_{k-1}(q_2) \le H^{k-1}$,

$$\sum_{\substack{q_i > x^{1/6} \\ \tau_{k-1}(q_1) > H^{k-1} \\ \tau_{k-1}(q_2) \le H^{k-1}}} \tau_{k-1}(q_1)\tau_{k-1}(q_2)\pi(x; q_1 q_2, 1) \ll x(C\log\log x)^{k-1} \sum_{\substack{q_i > x^{1/6} \\ \tau_{k-1}(q_1) > H^{k-1} \\ \tau_{k-1}(q_2) \le H^{k-1}}} \frac{\tau_{k-1}(q_1)}{q_1 q_2}$$
$$\ll x(C\log\log x)^{k-1} \sum_{\substack{q_1 > x^{1/6} \\ \tau_{k-1}(q_1) > H^{k-1}}} \frac{\tau_{k-1}(q_1)}{q_1}$$
$$\ll \frac{x(C\log\log x)^{k-2}}{\log^2 x},$$

by Cauchy-Schwarz inequality and (3.11)

$$\begin{aligned}\sum_1 &= (\sum_{p\le x} f_1^2(p))^{1/2}(\sum_{p\le x} f_2^2(p))^{1/2} \\ &\ll (U(x))^{1/2}(\sum_2)^{1/2} \\ &\ll (D_k(x))^{1/2}(\sum_2)^{1/2} \\ &\ll li(x)(C\log\log x)^{k-1/2}.\end{aligned}$$

To evaluate $U(x)$, recall;

$$U(x) = \sum_{\substack{q_i \le x^{1/6} \\ q_1 \ne q_2}} \tau_{k-1}(q_1)\tau_{k-1}(q_2)\pi(x, q_1 q_2, 1) + \sum_{q \le x^{1/6}} \tau_{k-1}^2(q)\pi(x, q, 1)$$

Invoking Bombieri-Vinagradov theorem, one has

$$\begin{aligned}U(x) = &\sum_{\substack{q_i \le x^{1/6} \\ q_1 \ne q_2}} \tau_{k-1}(q_1)\tau_{k-1}(q_2)(\pi(x, q_1 q_2, 1) - \frac{\pi(x)}{\phi(q_1 q_2)}) \\ &+ \sum_{q \le x^{1/6}} \tau_{k-1}^2(q)(\pi(x, q, 1) - \frac{\pi(x)}{\phi(q)}) + \sum_{\substack{q_i \le x^{1/6} \\ q_1 \ne q_2}} \tau_{k-1}(q_1)\tau_{k-1}(q_2)\frac{\pi(x)}{\phi(q_1 q_2)} \\ &+ \sum_{q \le x^{1/6}} \tau_{k-1}^2(q)\frac{\pi(x)}{\phi(q)} \\ =& li(x)A^2(x^{1/6}) + O(li(x)(\log\log x)^{2k-1})\end{aligned}$$

Since $A_{k-1}(x) - A_{k-1}(x^{1/6}) = O((C\log\log x)^{k-1})$, Lemma (3.7) follows. □

Lemma 3.8. *For every* $k = 0, 1, 2..$ *let*

$$B_k^2(x) = \sum_{p \le x} \frac{\tau_k^2(p)}{p}$$

then

$$B_k^2(x) = \frac{(\log\log x)^{2k+1}}{(2k+1)(k!)^2} + O\left(\frac{(C\log\log x)^{2k+1/2}}{2k+1/2}\right).$$

Proof. The desired result follows easily by partial summation and previous Lemma. □

3.3 The Distribution of τ_k

In this section, we shall show that the normalization,

$$\frac{\tau_k(n) - A_k(x)}{B_k(x)}$$

obeys a normal law, by using Theorem (2.12).
Recall,

$$B_k^2(x) = \frac{1}{(2k+1)(k!)^2}(\log\log x)^{2k+1} + O\left(\frac{(C\log\log x)^{2k+1/2}}{2k+1/2}\right)$$
$$A_k(x) = \frac{1}{(k+1)!}(\log\log x)^{k+1} + O((C\log\log x)^k)$$

Lemma 3.9. *For each fixed $k \geq 1$ and every real number α,*

$$\lim_{x\to\infty} \frac{1}{x}\sharp\left\{n \leq x \mid \frac{\tau_k(n) - A_k(x)}{B_k(x)} < \alpha\right\} = \Phi(\alpha), \tag{3.13}$$
$$\lim_{x\to\infty} \frac{1}{\pi(x)}\sharp\left\{p \leq x \mid \frac{\tau_k(p-1) - A_k(x)}{B_k(x)} < \alpha\right\} = \Phi(\alpha). \tag{3.14}$$

Proof. Let $H = (\varepsilon B_k(x))^{1/k}$. And let us recall that, Remark (3.1) for $H \geq 10\log\log x$ yields

$$\sum_{\substack{q\leq x\\ \tau_k(q)>H^k}} \frac{\tau_k(q)}{q} \ll \frac{(C_2\log\log x)^k}{(\log x)^2}.$$

It is clear that, once we fix ε, then $H = (\varepsilon B_k(x))^{1/k} > 10\log\log x$, for x large. Therefore it follows that

$$\sum_{\substack{p\leq x\\ \tau_k(p)>\varepsilon B_k(x)}} \frac{\tau_k^2(p)}{p} \ll \sum_{\substack{p\leq x\\ \tau_k(p)>(10\log\log x)^k}} \frac{\tau_k^2(p)}{p} = o(B_k^2(x)).$$

Hence (3.13) and (3.14) follows by invoking Theorems (2.12) and (2.13). □

We now attempt to replace τ_k by $\omega(\vartheta_k)$. To accomplish this we introduce the notion of "K-chain".

Definition 4. *A $(k+1)-$tuple of primes $(q_0, q_1, q_2...q_k)$ is called a K-chain if $q_{i-1}|q_i - 1$ for $i = 0, 1, ...k$. A general K-chain is denoted by Q_k, a K-chain with the property that $q_k|n$ is denoted by $Q_k(n)$ and $Q_k(n, q_0)$ denotes those K-chains where q_0 is fixed and $q_k|n$.*

Lemma 3.10. *There exists a K-chain starting at q_0 and ending at q_k if and only if $q_0|\vartheta_k(q_k)$.*

Proof. $\Rightarrow$ Suppose such a K-chain exists. Since ϑ is completely multplicative it follows that

$$\exists q_{k-1}|\vartheta(q_k) \implies \exists q_{k-2}|\vartheta_1(q_k).... \implies \exists q_0|\vartheta_k(q_k)$$

$\Leftarrow$ By induction on k, the assertion is obvious for $k = 1$.
Suppose $q_0|\vartheta_{k-1}(q_{k-1})$ implies existence of a $K-1$-chain, then $q_0|\vartheta_k(q_k)$ implies $q_0|\vartheta_{k-1}(q_k - 1) = q_0|\vartheta_{k-1}(p_1p_2...)$. Therefore $q_0|\vartheta_{k-1}(p_i)$ for some $p_i|q_k-1$. Thus by induction hypothesis, there exists a K-chain $(q_0, q_1, ...p_i, q_k)$. □

Lemma 3.11. *Let $|Q_k(n, q_0)| = \sharp Q_k(n, q_0)$. Then we have*

$$\tau_k(n) = \sum_{q_0|\vartheta_k(n)} Q_k(n, q_0). \tag{3.15}$$

Proof. By the definition of $\tau_k(n)$

$$\tau_k(n) = \sum_{p|n} \tau_k(p) = \sum_{p|n} \sum_{p_0|p-1} \tau_{k-1}(p_0) = \sum_{p|n} \sum_{p_0|p-1} \sum_{p_1|p_0-1} \sum_{p_2|p_1-1} \cdots \sum_{p_k|p_{k-1}-1}, \tag{3.16}$$

which obviously counts the number of K-chains ending at p with $p|n$, on the other hand by Lemma (3.10) the sum on the RHS of (3.15) counts the same K-chains. Therefore they are equal. □

Let us define the following functions

$$L(x) := \sum_{n\leq x}(\tau_k(n) - \omega(\vartheta_k(n))) = L^{(1)} + L^{(2)}$$
$$R(x) := \sum_{p\leq x}(\tau_k(p-1) - \omega(\vartheta_k(p-1))) = R^{(1)} + R^{(2)},$$

where

$$L^{(1)} = \sum_{n\leq x} \sum_{\substack{q_0|\vartheta_k(n) \\ q_0<y}} (|Q_k(n, q_0)| - 1)$$
$$R^{(1)} = \sum_{p\leq x} \sum_{\substack{q_0|\vartheta_k(p-1) \\ q_0<y}} (|Q_k(p-1, q_0)| - 1)$$
$$L^{(2)} = L(x) - L^{(1)}, \text{ and } R^{(2)} = R(x) - R^{(1)}$$

Lemma 3.12. *When* $y = (\log x)^2$

$$L^{(1)} \ll x(C \log\log x)^k \log\log\log x \tag{3.17}$$

$$R^1 \ll li(x)(C \log\log x)^k \log\log\log x \tag{3.18}$$

Proof. Let

$$\begin{aligned}
L^{(1)} &\leq \sum_{n\leq x} \sum_{\substack{q_0|\vartheta_k(n)\\ q_0<y}} |Q_k(n, q_0)| \\
&= \sum_{n\leq x} \sum_{\substack{q_0|\vartheta_k(n)\\ q_0<y}} \sum_{\substack{q_k|n\\ q_k\leq x}} \sum_{\substack{q_{k-1}|q_k-1\\ q_{k-1}\leq x}} \cdots \sum_{\substack{q_0|q_1-1\\ q_0<y}} \\
&\leq x \sum_{q_0<y} \sum_{\substack{q_1\leq x\\ q_0|q_1-1}} \sum_{\substack{q_2\leq x\\ q_1|q_2-1}} \cdots \sum_{\substack{q_k\leq x\\ q_{k-1}|q_k-1}} \frac{1}{q_k}.
\end{aligned}$$

By repeated use of Lemma (2.8) we have

$$L^{(1)} \ll x(C \log\log x)^k \log\log x. \tag{3.19}$$

Moreover,

$$\begin{aligned}
R^{(1)} &\leq \sum_{p\leq x} \sum_{\substack{q_0|\vartheta_k(p-1)\\ q_0<y}} |Q_k(p-1, q_0)| \\
&\leq \sum_{p\leq x} \sum_{\substack{q_k|p-1\\ q_k\leq x}} \sum_{\substack{q_{k-1}|q_k-1\\ q_{k-1}\leq x}} \cdots \sum_{\substack{q_0|q_1-1\\ q_0<y}} \\
&= \sum_{q_0<y} \sum_{\substack{q_1\leq x\\ q_0|q_1-1}} \sum_{\substack{q_2\leq x\\ q_1|q_2-1}} \cdots \sum_{\substack{q_k\leq x\\ q_{k-1}|q_k-1}} \pi(x, q_k, 1) \\
&= \sum_1 + \sum_2 + \sum_3 .
\end{aligned}$$

In $\sum_1$ we consider $q_k \leq x^{1/2}$ for which Brun-Titchmarsh inequality is applicable. Therefore

$$\begin{aligned}
\sum_1 &\ll li(x) \sum_{q_0<y} \sum_{\substack{q_1\leq x\\ q_0|q_1-1}} \sum_{\substack{q_2\leq x\\ q_1|q_2-1}} \cdots \sum_{\substack{q_k\leq x\\ q_{k-1}|q_k-1}} \frac{1}{q_k} \\
&\ll li(x)(C \log\log x)^k \log\log\log x
\end{aligned}$$

In $\sum_2$ we consider $q_k > x^{1/2}$ for which $\tau_k(q_k) \le (C\log\log x)^k = H^k$. We first note that the number of all K-chains ending at q_k equals $\tau_k(q_k)$. Therefore, it follows that

$$\begin{aligned}\sum_2 &\ll \sum_{p\le x}\sum_{q_0|\vartheta_k(p-1)}\sum_{\substack{q_k|p-1\\ q_k>x^{1/2}\\ \tau_k(q_k)\le H^k}} |Q_k(q_k,q_0)|\\ &= \sum_{p\le x}\sum_{\substack{q_k|p-1\\ q_k>x^{1/2}\\ \tau_k(q_k)\le H^k}} \tau_k(q_k)\\ &\ll \sum_{\substack{q_k>x^{1/2}\\ \tau_k(q_k)\le H^k}} \tau_k(q_k)\pi(x;q_k,1)\\ &\ll li(x)(C\log\log x)^k\end{aligned}$$

Finally, in $\sum_3$ we consider $q_k > x^{1/2}$ for which $\tau_k(q_k) > (C\log\log x)^k = H^k$. By the same reasoning, as in the previous case, Theorem (2.10) yields

$$\begin{aligned}\sum_2 &\ll \sum_{\substack{q_k>x^{1/2}\\ \tau_k(q_k)\ge H^k}} \tau_k(q_k)\pi(x;q_k,1)\\ &\ll \left(\sum_{\substack{q_k\le x\\ \tau_k(q_k)>H^k}} \frac{\tau_k^2(q_k)}{q_k}\right)^{1/2}\left(\sum_{x^{1/2}<q_k\le x} q_k\pi(x,q_k,1)^2\right)^{1/2}\\ &\ll \frac{(C\log\log x)^k x}{\log^2 x}.\end{aligned}$$

□

Lemma 3.13. *When* $y = (\log x)^2$

$$L^{(2)} \ll \frac{k^2x(C\log\log x)^{2k+1}}{\log^2 x},\ R^{(2)} \ll \frac{k^2li(x)(C\log\log x)^{2k+1}}{\log^2 x}.$$

Proof. Recall,

$$L^{(2)} = \sum_{n\le x}\sum_{\substack{q_0|\vartheta_k(n)\\ q_0\le y}} (|Q_k(n,q_0)|-1) \tag{3.20}$$

$$R^{(2)} = \sum_{p\le x}\sum_{\substack{q_0|\vartheta_k(p-1)\\ q_0\le y}} (|Q_k(p-1,q_0)|-1). \tag{3.21}$$

We first note that those pairs (n, q_0) for which $|Q_k(n, q_0)| = 1$ do not make contribution to the sum $L^{(2)}$. We therefore consider those pairs (n, q_0) for which $|Q_k(n, q_0)| \geq 2$. In this case, we obviously have

$$|Q_k(n, q_0)| - 1 \leq \binom{|Q_k(n, q_0)|}{2}$$

That is to say that, we obtain an upper bound for (3.20) and (3.21)(the same argument works for this case), if we count number of distinct K-chains starting at q_0 and ending at q_k with $q_k|n$.
Let Q denote the set of distinct K-chain pairs. Define $\mu_i(n, q_o)$ as follows
Let $P_1 = (q_0, p_1, p_2...p_k), P_2 = (q_0, p'_1, p'_2...p'_k) \in |Q_k(n, q_0)|$

$$(P_1, P_2) \in \mu_i(n, q_o) \Longleftrightarrow p_i \neq p'_i \text{ and } p_j = p'_j \text{ for } j > i. \tag{3.22}$$

It is obvious that $\mu_i(n, q_o) \cap \mu_j(n, q_o) = \emptyset$, unless $i = j$.
Therefore, if $(P_1, P_2) \in Q$ and $P_1 \neq P_2$, then $\exists i$ such that $p_{i_0} \neq p'_{i_0}$ then if, $p_j = p'_j$ for all $j > i_0$, then $(P_1, P_2) \in \mu_{i_0}(n, q_o)$. If not choose smallest i_1 such that, $p_{i_1} \neq p'_{i_1}$, then if $p_j = p'_j$ for all for all $j > i_1$, then $(P_1, P_2) \in \mu_{i_1}(n, q_o)$. Since P_1 and P_2 distinct, this process may not continue to the kth step. Therefore, there is some i_m such that, $p_j = p'_j$ for all $j > i_m$. Consequently

$$(P_1, P_2) \in \mu_{i_m}(n, q_o) \text{ for some } i_m\ . \tag{3.23}$$

By inclusion-exclusion principle we have the following inequality

$$|Q| \leq \sum_{i \leq k} \sharp \mu_i(n, q_o). \tag{3.24}$$

Substituting (3.24) in (3.20) we have

$$\begin{aligned} L^{(2)} &\leq \sum_{n \leq x} \sum_{\substack{q_0 \geq y \\ q_0 | \vartheta_k(n)}} \mu_1(n, q_o) + \mu_2(n, q_o) ... \mu_k(n, q_o) \\ &= M_1 + M_2 + ... + M_k, \end{aligned}$$

where $M_i = \sum_{n \leq x} \sum_{\substack{q_0 \geq y \\ q_0 | \vartheta_k(n)}} \mu_i(n, q_o)$.

We begin by evaluating M_k.
Obviously,

$$\mu_k(n,q_o)=\sum_{q_k|n}Q_k(q_0,q_k)\sum_{\substack{q'_k|n\\ q'_k\neq q_k}}Q_k(q_0,q'_k)$$

Therefore,

$$\begin{aligned}
M_k=&\sum_{n\le x}\sum_{\substack{q_0\ge y\\ q_0|\vartheta_k(n)}}\mu_k(n,q_o)\\
=&\sum_{n\le x}\sum_{\substack{q_0\ge y\\ q_0|\vartheta_k(n)}}\sum_{p_k|n}Q_k(q_0,p_k)\sum_{\substack{q'_k|n\\ q'_k\neq q_k}}Q_k(q_0,q'_k)\\
\le&\sum_{n\le x}\sum_{\substack{q_0\ge y\\ q_0|\vartheta_k(n)}}\sum_{\substack{q_kq'_k|n\\ q_k\neq q'_k}}\Big(\sum_{q'_{k-1}|q'_k-1}\ \sum_{q'_{k-2}|q'_{k-1}-1}\cdots\sum_{q_0|q'_1-1}\Big)\Big(\sum_{q_{k-1}|q_k-1}\ \sum_{q_{k-2}|q_{k-1}-1}\cdots\sum_{q_0|q_1-1}\Big)\\
\le&x(C\log\log x)^2\sum_{q_0\ge y}\sum_{\substack{q_0|q_1-1\\ q_0|q'_1-1\\ q'_1,q_1\le x}}\cdots\sum_{\substack{q_{k-2}|q_{k-1}-1\\ q'_{k-2}|q'_{k-1}-1\\ q'_{k-1},q_{k-1}\le x}}\frac{1}{q_{k-1}q'_{k-1}} \qquad (3.25)
\end{aligned}$$

To evaluate further, we distinguish between two cases: $q_{k-1}=q'_{k-1}$ and $q_{k-1}\neq q'_{k-1}$. In the first case let us consider the following sum

$$\begin{aligned}
\sum_{q_{k-1}\ge y}\frac{\tau^2_{k-1}(q_{k-1})}{q^2_{k-1}}=&\sum_{x\ge q_{k-1}\ge y}\frac{1}{q^2_{k-1}}\sum_{\substack{q_{k-2}|q_{k-1}-1\\ q'_{k-2}|q_{k-1}-1}}\ \sum_{\substack{q_{k-3}|q_{k-2}-1\\ q'_{k-3}|q'_{k-2}-1}}\cdots\sum_{\substack{q_0|q_1-1\\ q'_0|q'_1-1}}\\
=&\sum_{\substack{q'_0\le x\\ q_0\le x}}\sum_{\substack{q'_1\le x\\ q_1\le x\\ q_0|q_1-1\\ q'_0|q'_1-1}}\cdots\sum_{\substack{q'_{k-3}\le x\\ q_{k-3}\le x\\ q_{k-3}|q_{k-2}-1\\ q'_{k-3}|q'_{k-2}-1}}\ \sum_{\substack{y\le q_{k-1}\le x\\ q_{k-2}|q_{k-1}-1\\ q'_{k-2}|q'_{k-1}-1}}\frac{1}{q^2_{k-1}}.
\end{aligned}$$

In (3.25) the condition $q_0\ge y$ implies $q_{k-1}\ge y$. Therefore we have

$$\sum_{p_0\ge y}\sum_{\substack{p_0|p_1-1\\ p'_0|p'_1-1\\ p'_1,p_1\le x}}\cdots\sum_{\substack{p_{k-2}|p_{k-1}-1\\ p'_{k-2}|p'_{k-1}-1\\ p'_{k-1},p_{k-1}\le x}}\frac{1}{p_{k-1}p'_{k-1}}\ll\sum_{q_{k-1}\ge y}\frac{\tau^2_{k-1}(q_{k-1})}{q^2_{k-1}}\ll\frac{B^2_{k-1}(x)}{y},$$

and the case $q_{k-1} \neq q'_{k-1}$ can be read

$$\sum_{q_0 \geq y} \sum_{\substack{q_0 | q_1 - 1 \\ q_0 | q'_1 - 1 \\ q'_1, q_1 \leq x}} \cdots \sum_{\substack{q_{k-2} | q_{k-1} - 1 \\ q'_{k-2} | q'_{k-1} - 1 \\ q'_{k-1}, q_{k-1} \leq x \\ q_{k-1} \neq q'_{k-1}}} \frac{1}{q_{k-1} q'_{k-1}} \ll 2^2 (C \log\log x)^2$$

$$\times \sum_{q_0 \geq y} \sum_{\substack{q_0 | q_1 - 1 \\ q_0 | q'_1 - 1 \\ q'_1, q_1 \leq x}} \cdots \sum_{\substack{q_{k-2} | q_{k-1} - 1 \\ q'_{k-3} | q'_{k-2} - 1 \\ q'_{k-2}, q_{k-2} \leq x}} \frac{1}{q_{k-2} q'_{k-2}}.$$

Finally, we have

$$M_k \ll x (C \log\log x)^2$$
$$\times \left(\frac{B_{k-1}^2(x)}{y} + 2^2 (C \log\log x)^2 \Big(\sum_{q_0 \geq y} \sum_{\substack{q_0 | q_1 - 1 \\ q_0 | q'_1 - 1 \\ q'_1, q_1 \leq x}} \cdots \sum_{\substack{q_{k-2} | q_{k-1} - 1 \\ q'_{k-3} | q'_{k-2} - 1 \\ q'_{k-2}, q_{k-2} \leq x}} \frac{1}{q_{k-2} q'_{k-2}} \Big) \right).$$

Applying the same methods to the last sums, we arrive at the following inequality

$$\begin{aligned} M_k &\ll x \sum_{j=1}^{k} \frac{2^{2j} (C \log\log x)^{2j} B_{k-j}^2(x)}{y} \\ &\ll x \frac{(\log\log x)^{2k+1}}{y} \sum_{j=1}^{k} \frac{C^{2j}}{(2k-2j+1)((k-j)!)^2}. \\ &\ll k \frac{x (C \log\log x)^{2k+1}}{y}. \end{aligned}$$

Similarly, if $j < k$

$$\begin{aligned}
M_j =& \sum_{n \le x} \sum_{q_0 \ge y} \mu_j(n, q_o) \\
=& x \sum_{q_0 \ge y} \sum_{\substack{q_0 | q_1 - 1 \\ q_0' | q_1' - 1 \\ q_1 \le x \\ q_1' \le x}} .. \sum_{\substack{q_{j-1} | q_j - 1 \\ q_{j-1}' | q_j' - 1 \\ q_j \le x \\ q_j' \le x \\ q_j \ne q_j'}} \sum_{\substack{q_j | q_{j+1} - 1 \\ q_{j+1} \le x}} \cdots \sum_{\substack{q_{k-1} | q_k - 1 \\ q_k \le x}} \frac{1}{q_k} \\
=& x(C \log\log x)^{k-j} \sum_{q_0 \ge y} \sum_{\substack{q_0 | q_1 - 1 \\ q_0' | q_1' - 1 \\ q_1 \le x \\ q_1' \le x}} .. \sum_{\substack{q_{j-1} | q_j - 1 \\ q_{j-1}' | q_j' - 1 \\ q_j \le x \\ q_j' \le x \\ q_j \ne q_j'}} .
\end{aligned}$$

To evaluate further, we use the same method as evaluating M_k. Therefore one has

$$\begin{aligned}
M_j &\ll x(C \log\log x)^{k-j} \sum_{i=1}^{j} \frac{(C \log\log x)^{2i} B_{k-i}^2(x)}{y} \\
&\ll \frac{j(C \log\log x)^{k+j+1}}{y}.
\end{aligned}$$

If we choose $y = (\log x)^2$ it follows that

$$L^{(2)} \ll k^2 \frac{(C \log\log x)^{2k+1}}{(\log x)^2}.$$

Since

$$\sum_{p \le x} \sum_{q_0 | \vartheta_k(p-1)} (Q_k(p-1) - 1) \ll \sum_{n \le x} \sum_{q_0 | \vartheta_k(n)} (Q_k(n) - 1).$$

Therefore combining Lemma 2.12 and 2.13, we conclude that for almost all $n \le x$

$$\tau_k(n) - \omega(\vartheta_k(n)) = o((\log\log x)^{k+1/2}),$$

and that, for almost all $p \le x$

$$\tau_k(p-1) - \omega(\vartheta_k(p-1)) = o((\log\log x)^{k+1/2}).$$

It is now clear that, once again by Lemma 2.14, we have the following

$$\lim_{x\to\infty}\frac{1}{x}\sharp\left\{n\le x \mid \frac{\omega(\vartheta_k(n))-A_k(x)}{B_k(x)}<\alpha\right\}=\Phi(\alpha),$$
$$\lim_{x\to\infty}\frac{1}{x}\sharp\left\{p\le x \mid \frac{\omega(\vartheta_k(p-1))-A_k(x)}{B_k(x)}<\alpha\right\}=\Phi(\alpha).$$

And finally considering Lemma 3.3 one has

$$\begin{aligned}0\le\sum_{n\le x}(\omega(\phi_k(n))-\omega(\vartheta_k(n))) &\le \sum_{n\le x}(\omega(n)+\omega(\vartheta(n))+...+\omega(\vartheta_{k-1}(n)))\\ &\le \sum_{n\le x}\tau_0(n)+...+\tau_{k-1}(n)\\ &\le k\sum_{n\le x}\tau_{k-1}(n)\\ &\le kx\sum_{q\le x}\frac{\tau_{k-1}(p)}{p}+O(k\sum_{p\le x}\tau_{k-1}(p))\\ &\le kxA_{k-1}(x)+O(k\log x)=o(kxB_k(x)),\end{aligned}$$

and that

$$\begin{aligned}\sum_{p\le x}((\phi_k(p-1)-\omega(\vartheta_k(p-1))) &\le k\sum_{p\le x}\tau_{k-1}(p-1)\\ &= k\sum_{p\le x}\tau_k(p)\ \text{(by Lemma 3.5 and 3.6)}\\ &\ll \frac{1}{(k-1)!}\pi(x)(\log\log x)^k=o(\pi(x)B_k(x)).\end{aligned}$$

Therefore by Lemmas 2.12 and 2.13, Theorems 3.1 and 3.2 follows. □

4 Erdös-Kac Theorem for J_k

In this chapter, we generalize a theorem of Erdös and Pomerance [5] which states

$$\sharp\left\{n \le x \mid \frac{\omega(\phi(n)) - 1/2(\log\log n)^2}{(\log\log n)^{3/2}} \le \frac{u}{\sqrt{3}}\right\} \sim x\Phi(u) \tag{4.1}$$

as $x \to \infty$.
More precisely we will show that

Theorem 4.1. *Let $k \ge 1$ be fixed then for any real u one has*

$$\sharp\left\{n \le x \mid \frac{\omega(J_k(n)) - \frac{d(k)}{2}(\log\log n)^2}{d(k)(\log\log n)^{3/2}} \le \frac{u}{\sqrt{3}}\right\} \sim x\Phi(u) \tag{4.2}$$

as $x \to \infty$,
where $J_k(n)$ denotes Jordan's totient function defined as

$$J_k(n) = n^k \prod_{p|n}(1 - \frac{1}{p^k}).$$

Note that taking $k = 1$, we recover (4.1).

4.1 Preliminary Lemmas

Let $A(n) = \{a \in \mathbb{Z}/n\mathbb{Z} \mid a^k - 1 \equiv 0 \pmod{n}\}$
In this section, we will give several lemmas on the arithmetical structure of $A(n)$.

Lemma 4.2. *$\sharp A(q^\alpha) \le 2k$ for all q.*

Proof. Let us handle the case $q \ne 2$. We know that the group $(\mathbb{Z}/q^\alpha\mathbb{Z})^*$ is cyclic for $q \ne 2$. Let us suppose $l \mid \phi(q^\alpha)$ and consider the equation

$$x^l \equiv 1 \pmod{q^\alpha} \tag{4.3}$$

Let a be a primitive root $\pmod{q^\alpha}$, then the elements of the form $a^{\frac{\phi(q^\alpha)m}{l}}$ for $m = 0, 1 \ldots l-1$ are solutions to the equation (4.3).
Conversely, suppose $b \pmod{q^\alpha}$ is a solution to the equation (4.3), then there exists $\varepsilon \pmod{q^\alpha}$ such that

$$a^\varepsilon \equiv b \pmod{q^\alpha},$$

which implies

$$a^{l\varepsilon} \equiv 1 \pmod{q^\alpha},$$

Since a is a primitive root, it follows that $\phi(q^\alpha) \mid \varepsilon l$. Therefore one has

$$\varepsilon = \frac{\phi(q^\alpha)m}{l} \text{ for some } m \pmod{l} \ .$$

Thus, there are exactly l solutions of the equation (4.3), provided that $l|\phi(q^\alpha)$. Now suppose $l \nmid \phi(q^\alpha)$ consider the equation

$$x^{(l,\phi(q^\alpha))} \equiv 1 \pmod{q^\alpha} \tag{4.4}$$

If c is a solution to the equation (4.3), then so is to the equation (4.4). Conversely, suppose c is a solution to the equation (4.4). Let $(l, \phi(q^\alpha)) = d$. Then, there exist integers x_1 and x_2 such that

$$d = lx_1 + \phi(q^\alpha)x_2. \tag{4.5}$$

From the equation (4.4), it follows that

$$c^{lx_1} \equiv 1 \pmod{q^\alpha}.$$

If we can choose x_1 in such way that $(x_1, \phi(q^\alpha)) = 1$, then it follows that $c^l \equiv 1 \pmod{q^\alpha}$. Now the equation (4.5) is equivalent to the following equation

$$1 = \frac{l}{d}x_1 + \frac{\phi(q^\alpha)}{d}x_2,$$

whose solution set can be parametrized by the following formula

$$\begin{aligned} x &= x_1 + \frac{\phi(q^\alpha)}{d}t \\ y &= x_2 - \frac{l}{d}t \end{aligned} \tag{4.6}$$

for all $t \in \mathbb{Z}$. It is clear from (4.6) that $(x_1, \frac{\phi(q^\alpha)}{d}) = 1$. Therefore by Dirichlet's theorem there are infinitely many primes in the arithmetic progression

$$x = x_1 + \frac{\phi(q^\alpha)}{d}t.$$

The desired result follows choosing one that does not divide $\phi(q^\alpha)$. Therefore by the argument given in the beginning of the proof $(l, \phi(q^\alpha))|\phi(q^\alpha)$ so that there are exactly $(l, \phi(q^\alpha)) \le l$ solutions to the equation (4.4). Hence The proof is complete for the case $q \neq 2$.
Now let us suppose $q = 2$ and $\alpha = 1, 2$.
In both cases the group $\mathbb{Z}/2^\alpha\mathbb{Z}$ is cyclic. Thus the above argument works.

Suppose $q = 2$, and $\alpha \geq 3$ and notice that for any $n(\bmod 2^\alpha)$ there are

$$\mu(\bmod 2)$$
$$\varepsilon(\bmod 2^{\alpha-2}),$$

such that

$$n \equiv (-1)^\mu 5^\varepsilon \bmod 2^\alpha. \tag{4.7}$$

Thus if $n^k \equiv (-1)^{k\mu}5^{k\varepsilon} \equiv 1 \bmod 2^\alpha$, then (since the expression (4.7) is unique) one has

$$k\mu \equiv 0 \bmod 2$$
$$k\varepsilon \equiv 0 \bmod 2^{\alpha-2}.$$

Obviously, if k is odd then, there is only one solution. If k is even say $k = 2^\beta d$, where d is odd and $\beta \geq 1$, then

$$2^\beta \mu \equiv 0 \bmod 2$$
$$2^\beta \varepsilon \equiv 0 \bmod 2^{\alpha-2}$$

Now, suppose $\beta \geq \alpha - 2$. in this case ε and μ can be chosen arbitrarily. Consequently, there are $2^{\alpha-1}$ choices in total. Therefore we have $A(2^\alpha) \leq 2k$. Now, suppose $\beta < \alpha - 2$ in this case, μ is arbitrary but there are 2^β choices for ε hence there $2^{\beta+1} \leq 2k$ choices in total.

□

Lemma 4.3. *$\sharp A(n)$ is a multiplicative function.*

Proof. Suppose that $(n, m) = 1$ and define the following function

$$\tau : A(nm) \mapsto A(n) \times A(m)$$
$$a \mapsto (\bar{a} \pmod n) , \tilde{a} \pmod m)),$$

which is obviously well-defined and one-to-one. Now by Chinese Remainder Theorem, let x be the unique solution (mod nm) satisfying the following system of linear congruences.
For any $a \in A(n)$ and $b \in A(m)$

$$x \equiv a(\bmod n),$$
$$x \equiv b(\bmod m),$$

then we have

$$x^k \equiv 1 (\text{mod } n),$$
$$x^k \equiv 1 (\text{mod } m).$$

In other words, $nm | x^k - 1$ which proves that τ is onto. Hence it follows that $\sharp A(nm) = \sharp A(n) \sharp A(m)$. □

Remark 4.1. *From the theory of Cyclotomic polynomials we have*

$$x^k - 1 = \prod_{d|k} \Phi_d(x),$$

where $\Phi_d(x)$ is the d th cyclotomic polynomial which is irreducible(due to Gauss). Therefore $x^k - 1$ is the product of $d(k)$ irreducible polynomials. Moreover by Prime Ideal Theorem (2.4) it follows that

$$\sum_{q \le x} \frac{\sharp A(q)}{q-1} = d(k) \log \log x + O(1).$$

4.2 The Moments of $\Omega(p^k - 1)$

In order to prove Theorem (4.1) we will proceed almost the same as the previous section. We first define the following auxiliary function

$$h(n) = \sum_{p|n} \Omega(p^k - 1).$$

Obviously h is strongly additive. To find the asymptotic behavior of the functions $A(x)$ and $B(x)$, we first estimate the following sums

$$\sum_{p \le x} \Omega_y(p^k - 1)$$
$$\sum_{p \le x} \Omega_y^2(p^k - 1),$$

where $\Omega_y(n)$ denotes number of prime powers q^α dividing n such that $q \le y$. Furthermore, for the possible improvements which will be mentioned at the end of the chapter, we express the dependence of k in O-terms.

Theorem 4.4. *if $3 \le y \le x^k$, then*

$$\sum_{p \le x} \Omega_y(p^k - 1) = d(k)\pi(x) \log \log y + O(k \frac{x}{\log x}). \tag{4.8}$$

Proof.

$$\begin{aligned}
\sum_{p\le x}\Omega_y(p^k-1) =& \sum_{q\le y}\sum_{\substack{p\le x\\ p^k\equiv 1 (\mathrm{mod}\ q^\alpha)}} \\
=& \sum_{q\le y}\Big(\sum_{\substack{p\le x\\ p^k\equiv 1 (\mathrm{mod}\ q^\alpha)\\ p\equiv 1 (\mathrm{mod}\ q^\alpha)}} +\ldots+ \sum_{\substack{p\le x\\ p^k\equiv 1 (\mathrm{mod}\ q^\alpha)\\ p\equiv q^\alpha-1 (\mathrm{mod}\ q^\alpha)}}\Big) \\
=& \sum_{q\le y}\sum_{l\in A(q^\alpha)}\sum_{\substack{p\le x\\ p\equiv l (\mathrm{mod}\ q^\alpha)}} \\
=& \sum_{q\le y}\sum_{l\in A(q)}\sum_{\substack{p\le x\\ p\equiv l (\mathrm{mod}\ q)}} + \sum_{q\le y}\sum_{l\in A(q^\alpha)}\sum_{\substack{p\le x\\ p\equiv l (\mathrm{mod}\ q^\alpha)\\ \alpha\ge 2}} \\
=& A_1 + A_2
\end{aligned}$$

We first treat A_1 splitting the range of q into two parts as follows

$$\begin{aligned}
A_1 =& \sum_{q\le \min(y,x^{1/3})}\sum_{l\in A(q)}\sum_{\substack{p\le x\\ p\equiv l (\mathrm{mod}\ q)}} + \sum_{q> \min(y,x^{1/3})}\sum_{l\in A(q)}\sum_{\substack{p\le x\\ p\equiv l (\mathrm{mod}\ q}} \\
=& A_1' + A_2'.
\end{aligned}$$

Considering Remark (4.1) and Bombieri-Vinagradov theorem (2.7) it follows that,

$$\begin{aligned}
A_1' =& \sum_{q\le \min(y,x^{1/3})}\sum_{l\in A(q)}\frac{\pi(x)}{\phi(q)} + \sum_{q\le \min(y,x^{1/3})}\sum_{l\in A(q)}(\pi(x;q,l)-\frac{\pi(x)}{\phi(q)}) \\
=& \pi(x)\sum_{q\le \min(y,x^{1/3})}\frac{\sharp A(q)}{q-1} + O\left(k\sum_{q\le \min(y,x^{1/3})}\max_{(l,q)=1}(\pi(x;q,l)-\frac{\pi(x)}{\phi(q)})\right), \\
=& d(k)\pi(x)\log\log y + O(k\pi(x)). \qquad (4.9)
\end{aligned}$$

and that

$$A_2' = \sum_{q> \min(y,x^{1/3})}\sum_{l\in A(q)}\sum_{\substack{p\le x\\ p\equiv l (\mathrm{mod}\ q)}} \ll \sum_{p\le x}\sum_{\substack{q\mid p^k-1\\ q> x^{1/3}}} \ll k\pi(x). \qquad (4.10)$$

We now invoke Remark 4.1 and split the sum A_2 into two sums,

$$A_2 = \sum_{\substack{q\le y\\ q^\alpha \le x^{1/3}\\ \alpha\ge 2}} \sum_{l\in A(q^\alpha)} \sum_{\substack{p\le x\\ p\equiv l(\text{mod } q^\alpha)}} + \sum_{\substack{q\le y\\ q^\alpha > x^{1/3}\\ \alpha\ge 2}} \sum_{l\in A(q^\alpha)} \sum_{\substack{p\le x\\ p\equiv l(\text{mod } q^\alpha)}} \tag{4.11}$$

$$\ll \left(\pi(x) \sum_{\substack{q\le y\\ q^\alpha \le x^{1/3}\\ \alpha\ge 2}} \frac{\sharp A(q^\alpha)}{\phi(q^\alpha)} + \sum_{\substack{q\le y\\ q^\alpha > x^{1/3}\\ \alpha\ge 2}} \frac{\sharp A(q^\alpha)}{q^\alpha} \right)$$

$$\ll k(\pi(x) + x^{5/6})$$

$$\ll k\pi(x) \tag{4.12}$$

In (4.11), for the first sum we used Brun-Titchmarsh inequality and for the second sum we used the trivial estimate.
Therefore combining (4.9), (4.10) and (4.12), the desired result follows. □

Theorem 4.5. *if* $3 \le y \le x^k$ *then,*

$$\sum_{p\le x} \Omega_y^2(p^k - 1) = d^2(k)\pi(x)(\log\log y)^2 + O(k^2\pi(x)\log\log y)$$

Proof.

$$\sum_{p\le x} \Omega_y^2(p^k - 1) = \sum_{p\le x} \sum_{\substack{q_1^{\alpha_1} | p^k - 1\\ q_2^{\alpha_2} | p^k-1,\\ q_1,q_2\le y}}$$

$$= \sum_{\substack{q_1,q_2\le y\\ q_1\ne q_2}} \sum_{\substack{q_1^{\alpha_1} | p^k - 1\\ q_2^{\alpha_2} | p^k-1,\\ p\le x}} + O(d(k)\pi(x)\log\log y)$$

$$= \sum_{\substack{q_1,q_2\le y\\ q_1\ne q_2}} \sum_{\substack{q_1 | p^k - 1\\ q_2 | p^k-1\\ p\le x}} + \sum_{\substack{q_1,q_2\le y\\ q_1\ne q_2}} \sum_{\substack{q_1^{\alpha_1} | p^k - 1\\ q_2^{\alpha_2} | p^k-1\\ p\le x\\ \alpha_1\alpha_2>1}} + O(d(k)\pi(x)\log\log y)$$

$$= \beta_1 + \beta_2 + O(d(k)\pi(x)\log\log y) \tag{4.13}$$

as before, we split the sum β_1 into two sums.

$$\beta_1 = \sum_{\substack{q_1,q_2\le \min(y,x^{1/6})\\ q_1\ne q_2}} \sum_{l\in A(q_1q_2)} \sum_{\substack{p\equiv l(\text{mod } q_1q_2)\\ p\le x}} + \sum_{\substack{q_1,q_2\\ q_1\ne q_2}} \sum_{l\in A(q_1q_2)} \sum_{\substack{p\equiv l(\text{mod } q_1q_2)\\ p\le x}}$$

$$= \beta_1' + \beta_2'$$

Hence, by Bombieri-Vinagradov theorem (2.7), Prime-Ideal theorem (2.4) and Remark (4.1) one has

$$\begin{aligned}
\beta_1' = & \sum_{\substack{q_1,q_2 \le \min(y,x^{1/6}) \\ q_1 \ne q_2}} \sum_{l \in A(q_1q_2)} \Big(\sum_{\substack{p \equiv l (\bmod\ q_1q_2) \\ p \le x}} - \frac{\pi(x)}{\phi(q_1q_2)} \Big) \\
& + \sum_{\substack{q_1,q_2 \le \min(y,x^{1/6}) \\ q_1 \ne q_2}} \sum_{l \in A(q_1q_2)} \frac{\pi(x)}{\phi(q_1q_2)} \\
= & \pi(x) \sum_{\substack{q_1,q_2 \le \min(y,x^{1/6}) \\ q_1 \ne q_2}} \frac{\sharp A(q_1q_2)}{(q_1-1)(q_2-1)} \\
& + O\left(k^2 \sum_{q_1q_2 \le x^{1/3}} \max_{(l,q_1q_2)=1} (\pi(x,q_1q_2,l) - \frac{\pi(x)}{\phi(q_1q_2)}) \right) \\
= & \pi(x) \left(\sum_{q \le min(y,x^{1/6})} \frac{\sharp A(q)}{q} \right)^2 + O(k^2\pi(x)) \\
= & d^2(k)\pi(x)(\log\log y)^2 + O(k^2\pi(x), d(k)\pi(x)\log\log y). \qquad (4.14)
\end{aligned}$$

In the sum β_2' there are 3 main cases either $q_1, q_2 > min(y, x^{1/6})$ or $q_1 > min(y, x^{1/6})$ and $q_2 \le min(y, x^{1/6})$ or $q_1 \le min(y, x^{1/6})$ and $q_2 > min(y, x^{1/6})$. In the first case,

$$\begin{aligned}
\sum_{\substack{q_1,q_2 > \min(y,x^{1/6}) \\ q_1 \ne q_2}} \sum_{\substack{q_1q_2 | p^k-1 \\ p \le x}} & \le \sum_{\substack{q_1q_2 > x^{1/6} \\ q_1 \ne q_2}} \sum_{\substack{q_1q_2 | p^k-1 \\ p \le x}} \\
& = \sum_{p \le x} \sum_{\substack{q_1q_2 | p^k-1 \\ q_1,q_2 > x^{1/6}}} \\
& \ll k^2\pi(x) \qquad (4.15)
\end{aligned}$$

In the second case(which is symmetric with the third),

$$\sum_{\substack{q_1 > \min(y,x^{1/6}) \\ q_2 \le \min(y,x^{1/6}) \\ q_1 \ne q_2}} \sum_{\substack{q_1q_2 | p^k-1 \\ p \le x}} \le \sum_{p \le x} \sum_{\substack{q_1 | p^k-1 \\ q_1 > \min(y,x^{1/6})}} \sum_{\substack{q_2 | p^k-1 \\ q_2 \le \min(y,x^{1/6})}} \ll k^2\pi(x)\log\log y$$

The sum β_2 is similarly shown to be

$$\beta_2 \ll k^2\pi(x). \qquad (4.16)$$

Therefore combining (4.13), (4.14), (4.15) and (4.16) the desired result follows.

□

Theorem 4.6. *if* $3 \leq y \leq x^k$ *then,*

1.
$$\sum_{p\leq x} \frac{\Omega_y(p^k-1)}{p} = d(k)(\log\log x \log\log y - \frac{1}{2}(\log\log y)^2) + + O(k \log k \log\log x),$$

2.
$$\sum_{p\leq x} \frac{\Omega_y^2(p^k-1)}{p} = d^2(k)(\log\log x (\log\log y)^2 - \frac{2}{3}(\log\log y)^3) + O(k^2 \log k \log\log x \log\log y).$$

Proof. (1)

$$\begin{aligned}
\sum_{p\leq x} \frac{\Omega_y(p^k-1)}{p} &= \frac{\sum_{p\leq x} \Omega_y(p^k-1)}{x} + \int_2^x \frac{1}{t^2} \sum_{p\leq t} \Omega_y(p^k-1)dt \\
&= \frac{\sum_{p\leq x} \Omega_y(p^k-1)}{x} + \int_2^y \frac{1}{t^2} \sum_{p\leq t} \Omega_y(p^k-1)dt \\
&\quad + \int_y^x \frac{1}{t^2} \sum_{p\leq t} \Omega_y(p^k-1)dt \\
&= \frac{d(k)\log\log y}{\log x} + O\left(\frac{k}{\log x}\right) + \int_2^y \frac{1}{t^2} \sum_{p\leq t} \Omega_t(p^k-1)dt \\
&\quad + \int_y^x \frac{1}{t^2} \sum_{p\leq t} \Omega_y(p^k-1)dt \\
&= d(k)\frac{1}{2}(\log\log y)^2 + O(k\log k \log\log x) \\
&\quad + \int_y^x \frac{1}{t^2} \sum_{p\leq t} \Omega_y(p^k-1)dt \qquad (4.17)
\end{aligned}$$

In order to evaluate the last integral suppose $y \leq x$ then,

$$\int_y^x \frac{1}{t^2} \sum_{p\leq t} \Omega_y(p^k-1)dt = d(k)(\log\log x \log\log y - (\log\log y)^2) + O(k \log\log x). \qquad (4.18)$$

and if $x \leq y \leq x^k$ then,

$$\begin{aligned}\int_y^x \frac{1}{t^2} \sum_{p \leq t} \Omega_y(p^k-1)dt =& d(k)(\log\log x \log\log y - (\log\log y)^2) \\ &+ O(|\log\log x - \log\log y|) \\ =& d(k)(\log\log x \log\log y - (\log\log y)^2) + O(k \log k).\end{aligned} \tag{4.19}$$

Combining (4.17) and (4.18) and (4.19) result follows. □

Proof. (2)

$$\begin{aligned}\sum_{p \leq x} \frac{\Omega_y^2(p^k-1)}{p} =& \frac{d^2(k)\log\log y}{\log x} + O\left(\frac{k^2 \log\log y}{\log x}\right) + \int_2^x \frac{1}{t^2} \sum_{p \leq t} \Omega_y^2(p^k-1)dt \\ =& d(k)(\log\log x(\log\log y)^2 - \frac{2}{3}(\log\log y)^2 \\ &+ O\left(\log\log y \int_y^x \frac{k^2}{t \log t} dt\right) + O\left(k^2 \int_2^y \frac{\log\log t}{t \log t} dt\right)\end{aligned} \tag{4.20}$$

In fact the first error term is at most

$$\begin{aligned}\log\log y \int_y^x \frac{k^2}{t \log t} dt \ll& k^2 \log\log y(|\log\log x - \log\log y|) \\ \ll& k^2 \log k \log\log y \log\log x,\end{aligned} \tag{4.21}$$

and the second error term is at most

$$k^2 \int_2^y \frac{\log\log t}{t \log t} dt \ll k^2(\log\log y)^2 \ll k^2 \log k \log\log y \log\log x. \tag{4.22}$$

Combining (4.20), (4.21) and (4.22) the desired result follows. □

4.3 The Distribution of $\Omega(p^k - 1)$

In this section, we deduce the following theorem which will almost complete the proof of Theorem (3.1)

$$\sharp\left\{n \leq x \mid \frac{h(n) - \frac{d(k)}{2}(\log\log x)^2}{d(k)(\log\log x)^{3/2}} \leq \frac{u}{\sqrt{3}}\right\} \sim x\Phi(u) \text{ as } x \to \infty\ . \tag{4.23}$$

To accomplish this let us take $y = x^k$ in Theorems (4.4) and (4.5) then we have,

$$\sum_{p\le x}(\Omega(p^k-1)-d(k)(\log\log x))^2 \ll_k \pi(x)\log\log x \tag{4.24}$$

Therefore (4.23) follows, if we show that

$$\sum_{\substack{p\le x\\ \Omega(p^k-1)>\varepsilon B(x)}} \frac{\Omega^2(p^k-1)}{p} = o(B^2(x)). \tag{4.25}$$

To do this, we define the following function:

$$\alpha(p)=\begin{cases}1 & \text{if } \Omega(p^k-1)>\varepsilon B(x)\\ 0 & \text{otherwise.}\end{cases}$$

By (4.24) we have

$$\sum_{\substack{p\le x\\ \Omega(p^k-1)>\varepsilon B(x)}} (\Omega(p^k-1)-d(k)(\log\log x))^2 \ll_k \pi(x)\log\log x.$$

For a fixed $\varepsilon>0$, It follows that

$$\sum_{\substack{p\le x\\ \Omega(p^k-1)>\varepsilon B(x)}} (\varepsilon B(x)-d(k)(\log\log x))^2 \ll \pi(x)\log\log x. \tag{4.26}$$

Therefore we have for $x\ge x_0(k,\varepsilon)$,

$$\sum_{p\le x}\alpha(p)\ll \frac{\pi(x)}{(\log\log x)^2}.$$

Lemma 4.7. *The sum* $\sum_{p\le x}\frac{\alpha(p)}{p}$ *converges.*

Proof. By partial summation

$$\begin{aligned}\sum_{p\le x}\frac{\alpha(p)}{p} &\ll \frac{1}{\log x(\log\log x)^2}+\int_2^x \frac{1}{x\log x(\log\log x)^2}dx\\ &\ll \frac{1}{\log x(\log\log x)^2}-\frac{1}{\log\log x}\ |_2^x\\ &\le\infty.\end{aligned}$$

□

Now by Cauchy-Schwarz inequality one has

$$\sum_{p\le x}\alpha(p)\frac{\Omega^2(p^k-1)}{p}\le\left(\sum_{p\le x}\frac{\alpha(p)}{p}\right)^{1/2}\left(\sum_{p\le x}\frac{\Omega^4(p^k-1)}{p}\right)^{1/2}. \qquad (4.27)$$

By Lemma 3.7 the first sum on RHS converges and by similar methods one can show that

$$\sum_{p\le x}\frac{\Omega^4(p^k-1)}{p}\ll(\log\log x)^5=o(B^4(x)).$$

We therefore have shown that (4.23) indeed holds. We now attempt to replace $h(n)$ by $\Omega(J_k(n))$. First notice that the function defined by

$$F(n)=\Omega(J_k(n))-h(n)$$

is additive. Therefore Turan-Kubilius inequality can be applied to the function $F(n)$ with

$$B_1(x)=\sum_{p^\alpha\le x}\frac{F(p^\alpha)}{p^\alpha}\left(1-\frac{1}{p}\right)\text{ and }B_2^2(x)=\sum_{p^\alpha\le x}\frac{F^2(p^\alpha)}{p^\alpha}.$$

To see that both $B_1(x)$ and $B_2(x)$ remain bounded as $x\to\infty$

$$\begin{aligned}B_1(x)&\ll\sum_{p^\alpha\le x}\frac{\Omega(p^k-1)-\Omega(p^{(\alpha-1)k}(p^k-1)}{p^\alpha}\\&\ll\sum_{\substack{p^\alpha\le x\\ \alpha\ge2}}\frac{\alpha-1}{p^\alpha}\\&\ll\sum_{p\le x}\sum_{\alpha\ge2}\frac{\alpha-1}{p^\alpha}\\&\ll\sum_{p\le x}\frac{1}{p^2}\sum_{\alpha\ge2}\frac{\alpha-1}{p^{\alpha-2}}\\&\ll\sum_{p\le x}\frac{1}{p^2}=O(1)\end{aligned}$$

similarly

$$\begin{aligned}
B_2(x) &= \sum_{p^\alpha \le x} \frac{(\Omega(p^k-1) - \Omega(p^{(\alpha-1)k}(p^k-1)))^2}{p^\alpha} \\
&\ll \sum_{\substack{p^\alpha \le x \\ \alpha \ge 2}} \frac{(\alpha-1)^2}{p^\alpha} \\
&\ll \sum_{p\le x} \frac{1}{p^2} \sum_{\alpha \ge 2} \frac{(\alpha-1)^2}{p^{\alpha-2}} = O(1)
\end{aligned}$$

The last inequality follows by differentiating the following identity

$$z + 2z^2 + 3z^3 + ... + nz^n ... = \frac{z}{(1-z)^2} \text{ for } |z| < 1 \ .$$

Therefore, one has

$$\sum_{n\le x} (h(n) - \Omega(J_k(n)))^2 \ll \sum_{n \le x} (h(n) - \Omega(J_k(n)) - B_1(x))^2 + \sum_{n\le x} B_1(x)^2 \quad \ll x$$

For a fixed $\varepsilon > 0$, it follows that for almost all $n \le x$ we have

$$|h(n) - \Omega(J_k(n))| < \varepsilon(\log\log x)^{3/2}.$$

Therefore, by Theorem (2.14) we may replace $h(n)$ by $\Omega(J_k(n))$.
Finally, we want to replace $\Omega(J_k(n))$ by $\omega(J_k(n))$. To do this, once again we will make use of Turan-Kubilius inequality (2.11). Let us take $y = (\log\log x)^2$, then the function Ω_y is additive. Let us consider the following function:

$$\begin{aligned}
E(x) &= \sum_{p^\alpha \le x} \frac{\Omega_y(J_k(p^\alpha))}{p^\alpha} \left(1 - \frac{1}{p}\right) \\
&= \sum_{p^\alpha \le x} \frac{(\alpha-1)k + \Omega_y(p^k-1)}{p^\alpha} \left(1 - \frac{1}{p}\right) \\
&= \sum_{p \le x} \frac{\Omega_y(p^k-1)}{p} \\
&\quad + O\left(\sum_{\substack{p^\alpha \le x \\ \alpha \ge 2}} \frac{k(\alpha-1) + \Omega(p^k-1)}{p^\alpha} + \sum_{p^\alpha \le x} \frac{k(\alpha-1) + \Omega(p^k-1)}{p^{\alpha+1}} \right)
\end{aligned} \tag{4.28}$$

Now suppose $n = p_1^{r_1} p_2^{r_2} ... p_n^{r_n}$, then

$$\Omega(n) = r_1 + r_2 ... + r_3$$

On the other hand

$$n \geq 2^{r_1 + r_2 ... + r_n}$$

it follows that

$$r_1 + r_2 ... + r_n \ll \log n.$$

Therefore, one has $\Omega(n) \ll \log n$, for all $n \in \mathbb{N}$.When $n = p^k - 1$, one has $\Omega(p^k - 1) \ll k \log p$.
Using the above argument and by similar methods in which Turan-Kubilius inequality previously is applied, it follows that the error term in (4.28) $\ll k$.
And the function,

$$\begin{aligned} V(x) &= \sum_{p^\alpha \leq x} \frac{\Omega_y^2(J_k(p^\alpha))}{p^\alpha} \\ &= \sum_{p \leq x} \frac{\Omega_y^2(p^k - 1)}{p} + O\left(\sum_{\substack{p^\alpha \leq x \\ \alpha \geq 2}} \frac{\Omega_y^2(J_k(p^\alpha))}{p^\alpha} \right) \end{aligned} \tag{4.29}$$

Using similar arguments it follows that the error term in (4.29) $\ll k^2$.
With the choice of $y = (\log \log x)^2$, we have the following inequality

$$\sum_{n \leq x} (\Omega_y(J_k(n)) - E_y(x))^2 \ll x D_y^2(x),$$

with

$$E_y(x) = d(k) \log \log x \log \log \log \log x + O(k \log k \log \log x)$$
$$D_y^2(x) = d(k) \log \log x (\log \log \log \log x)^2 + O(k^2 \log k \log x \log x \log \log \log \log x).$$

Using the inequality $(a + b)^2 \ll a^2 + b^2$, we have

$$\sum_{n \leq x} (\Omega_y(J_k(n)) - d(k) \log \log x \log \log \log \log x)^2 \ll x d^2(k) \log \log x (\log \log \log \log x)^2.$$

Now let S be the set of integers satisfying

$$\Omega_y(J_k(n)) > 2d(k) \log \log x \log \log \log \log x,$$

then the density of S is at most

$$\frac{1}{x}\sum_{\substack{n\le x\\ n\in S}} \ll \frac{1}{\log\log x} = o(x).$$

Therefore, for almost all $n \le x$ we have

$$\Omega_y(J_k(n)) \le 2d(k)\log\log x \log\log\log\log x$$

and that,

$$\Omega_y(J_k(n)) - \omega_y(J_k(n)) \le 2d(k)\log\log x \log\log\log\log x. \tag{4.30}$$

Now our aim is to show that the following equality holds for almost all integers $n \le x$

$$\Omega(J_k(n)) - \Omega_y(J_k(n)) = \omega(J_k(n)) - \omega_y(J_k(n)). \tag{4.31}$$

First of all notice that there are two main cases

1. if $q^2|J_k(n)$ then $q^2|J_k(p_1^{\alpha_1})$ for some $p_1^{\alpha_1}|n$,
2. if $q^2|J_k(n)$ then $q|J_k(p_1^{\alpha_1})$ and $q|J_k(p_2^{\alpha_2})$ for some distinct $p_1^{\alpha_1}|n$ and $p_2^{\alpha_2}|n$.

Suppose that $q^2|J_k(p_1^{\alpha_1})$ which implies either $q^2|p_1^{(\alpha_1-1)k}$ or $q^2|p_1^k-1$, If $q^2|p_1^{(\alpha-1)k}$, then it follows that, $q^2|n$
In this case the number of integers divisible by a square of a prime $> y$ is at most

$$\sum_{q>y}\frac{1}{q^2} = o(x).$$

If $q^2|p_1^k-1$, then the number of integers having prime factor p_1 such that $q^2|p_1^k-1$ with $q>y$ is at most

$$\begin{aligned}
\sum_{q>y}\sum_{\substack{q^2|p^k-1\\ p\le x}}\frac{x}{q} \ll& x\sum_{q>y}\sum_{l\in A(q^k)}\sum_{\substack{p\le x\\ p\equiv l(\text{mod } q)}}\frac{1}{q}\\
\ll&\sum_{q>y}\sum_{l\in A(q^2)}(\frac{\log\log x}{q(q-1)}+O\left(\frac{\log q}{q^2}\right)\\
\ll& k^2x\log\log x\sum_{q>y}\frac{1}{q^2}+k^2x\sum_{q>y}\frac{\log q}{q^2}\\
=&O\left(\frac{k^2x\log\log x}{y\log\log y}\right)+O\left(\frac{k^2x}{y}\right) = o(x).
\end{aligned}$$

Suppose $\exists p_1, p_2$ such that $p|J_k(p_1^{\alpha_1})$ and $p|J_k(p_2^{\alpha_2})$ for some distinct $p_1^{\alpha_1}|n$ and $p_2^{\alpha_2}|n$ In this case,

$$q|p_1^{(\alpha_1-1)k}(p_1^k-1)$$
$$q|p_2^{(\alpha_2-1)k}(p_2^k-1)$$

If $q = p_1$, then $q^2|n$ and $q|(p_2^k-1)$, then the number of all such integers is at most,

$$\sum_{q>y}\frac{1}{q^2} = o(x).$$

if $q|(p_1^k-1)$ and $q|(p_2^k-1)$, then the number such $n \leq x$ is at most

$$x\sum_{q>y}\sum_{\substack{q|(p_1^k-1)\\ q|(p_2^k-1)\\ p_i\leq x}}\frac{1}{q_1q_2} \ll \sum_{q>y}(\sum_{\substack{q|(p_1^k-1)\\ p_1\leq x}}\frac{1}{p_1p_2})^2$$

$$\ll x\sum_{q>y}(\sum_{l\in A(q)}\sum_{\substack{p_1\equiv l \pmod q\\ p_1\leq x}}\frac{1}{p_1})^2$$

$$\ll xk^2\sum_{q>y}(\frac{\log\log x}{q}+O(\frac{\log q}{q}))^2$$

$$\ll xk^2(\log\log x)^2\sum_{q>y}\frac{1}{q^2}+O(xk^2\sum_{q>y}\frac{(\log p)^2}{p^2})$$

$$\ll xk^2\frac{\log\log x}{y\log\log y}+xk^2\frac{\log\log x}{y} = o(x).$$

Therefore combining (4.30) and (4.31), the desired result follows.

.

5 Concluding Remarks

1. It would be fruitful to investigate, if one can take k as an increasing function of x in Theorem 4.1.

2. It is also probable to combine Theorems (4.1) and (3.1) to deduce that the function $\omega(J_k^m(n))$ where $J_k^m(n) = J_k^{m-1}(J_k(n))$ obeys a normal law, if one estimates the sums of the form

$$\sum_{p_0 \le x} \sum_{p_0 | p_1^k - 1} \sum_{p_1 | p_2^k - 1} \sum_{p_2 | p_3^k - 1} \cdots \sum_{p_{m-1} | p_m^k - 1} .$$

3. Another variation of Theorem (3.1) may be given for the function $\omega(\rho(n) - 1)$, where $\rho(n)$ denotes the least prime divisor of n, provided that one has a powerful analogue of the sieve of Eratosthenes. Since it provides asymptotic formula for the number of integers having the least prime divisor p for a poor range of p(for instance $p \le \log x$). For the present case we are capable of proving

$$\sum_{n \le x} \omega(\rho(n) - 1) \asymp x(\log\log x)^2$$

using the sieve of Selberg.

References

[1] Apostol, T. M., -Introduction to Analytic Number Theory, Springer, 1976.

[2] Davenport, H., -Multiplicative number theory, 3rd edition, Springer, 2000.

[3] Erdös, P.&Kac, M, -Gaussian law of errors in the theory of additive number theoretic functions. Amer.J.Math.,62,,738-742, 1940

[4] Erdös, P. - On the normal order of prime factors of $p-1$ and some related problems concerning Euler's ϕ function, Quart.Journ. Math (Oxford),6,205-213, 1935

[5] Erdös, P.& Pomerance, C. -On the normal number of prime factors $\phi(n)$, Rocky Mountain Journal of Math. Soc, 15,343-352, 1985

[6] Hardy, G.H. & Ramanujan, S. -The normal number of prime factors of a number n, Quart.J. Pure Appl.Math., 48,76-97, 1917

[7] Halberstam, H. -On the distribution of additive number-theoritic functions (I,II & III) , J. London Math. Soc., 30,43-53, 1955, and 31, 1-14, and 31- 15.27, 1956

[8] Kubulius, J, -Probabilistic methods in number theory, Translations of Math. Monographs (No.11), Rhode Island, 1964

[9] Montgomery, H. L. and Vaughan R., -Multiplicative number theory, I. Classical theory. Cambridge, 2007.

[10] Murty,R, -Problems in Analytic Number Theory, Springer Verlag, Vol.206, New York, 2001

[11] Murty,R, -Problems in Algebraic Number Theory, Springer Verlag, Vol.206, New York, 2001

[12] Murty, R,& Alina Carmen Cojocaru -An Introduction to Sieve Methods and Their Applications (London Mathematical Society Student Texts), 2004

[13] Murty,K. & Murty, R.M.- An Analogue of the Erdös-Kac theorem for Fourier coefficients of modular forms Indian J.pure & App.Math., 15, 1090-1101, 1984

[14] N.L.Bassily, I.Katai & M.Wijsmuller, Number of Prime Divisors of $\phi_k(n)$, Where ϕ_k is the k-fold iterate of ϕ Journal of Number Theory 65, 226-239(1997) NT972169

[15] P.D.T.A. Elliot, -Probabilistic Number Theory I, II, Springer Verlang, Newyork, 1980

[16] P. Turan, On a theorem of Hardy and Ramanujan, J. London Math. Soc. 9 (1934), 274-276

[17] Renyi, A. and Turan, P . On a theorem of Erdös and Kac. Acta Aritmetica, 1, 71-84 (1957).

[18] Saidak,F. - Non-Abelian generalizations of Erdös-Kac Theorem Canad. J. Math. Vol.56 (2), 356-372, 2004

[19] Tenenbaum, G., -Introduction to Analytic and Probabilistic Number Theory, Cambridge, 1995.