On the Square-free Numbers in Shifted Primes

Zerui Tan
The High School Attached to The Hunan Normal University, China
Advisor : Yongxing Cheng

May 19, 2014

On the Square-free Numbers in Shifted Primes

May 19, 2014

Abstract

In this paper, we prove that for any $A > 2$, $\mathcal{Q}_k(x)$, the number of primes not exceeding x such that $p - k$ is square free, have the following asymptotic formula

$$\mathcal{Q}_k(x) = \prod_{p|k}\left(1 + \frac{1}{p^2 - p - 1}\right)\prod_{p}\left(1 - \frac{1}{p(p-1)}\right)\mathrm{li}x + O\left(\frac{x}{(\log x)^A}\right)$$

with x sufficiently large. Where the implied constant depends only on A.

Contents

[0] *Keywords*: shifted primes, square free numbers

1 Introduction

In this article, We obtain the following results.

Theorem 1.1. *For any $A > 2$, we have*

$$\mathcal{Q}_k(x) = C_k \mathrm{li} x + O_A\left(\frac{x}{(\log x)^A}\right)$$

where the constant

$$C_k = \prod_{p|k}\left(1 + \frac{1}{p^2 - p - 1}\right)\prod_p\left(1 - \frac{1}{p(p-1)}\right)$$

To prove the theorem, we need the following simple ideas. The simplest one is

$$\sum_{d|n}\mu(d) = \delta(n)$$

where

$$\delta(n) := \begin{cases} 1 & n = 1 \\ 0 & \text{otherwise} \end{cases}$$

Therefore we have

$$\mu^2(n) = \sum_{d^2|n}\mu(d)$$

where the sum $\sum_{d^2|n}$ is over all positive divisor d such that $d^2|n$.
We also need a well-known theorem due to Bombieri and A.I.Vinogradov

Theorem 1.2. (Bombieri-Vinogradov)
Let $A > 0$, There is some constant $B = B(A)$ such that

$$\sum_{q \le x^{1/2}/(\log x)^B}\max_{y \le x}\max_{(a,q)=1}\left|\pi(y; q, a) - \frac{\mathrm{li}(y)}{\varphi(q)}\right| \ll_A \frac{x}{(\log x)^A} \quad (x \ge 2)$$

For the proof of the theorem, see [2].

2 Proof of Theorem 1.1

It is easy to see that

$$\mathcal{Q}_k(x) = \sum_{p \le x}\mu^2(p - k)$$

Because any number $n \in \mathbb{N}$ can be written as the form $n = a^2 b$ uniquely, where b is a square free number. n is square free if and only if $b = 1$, by the fact $\delta = \mu * 1$ thus

$$\mu^2(n) = \sum_{d^2|n}\mu(d)$$

We obtain

$$\mathcal{Q}_k(x) = \sum_{p\le x}\sum_{d^2|p-k}\mu(d) = \sum_{d<\sqrt{x}}\sum_{\substack{p\le x\\ p\equiv k (\mathrm{mod}\ d^2)}}\mu(d) = \sum_{d<\sqrt{x}}\mu(d)\pi(x;d^2,k)$$

Let $\mathcal{L} = \log x$, we divide this sum in to two pieces, $[1, x^\alpha\mathcal{L}^{-B_0}), [x^\alpha\mathcal{L}^{-B_0}, x^{1/2}]$

$$\begin{aligned}\mathcal{Q}_k(x) &= \sum_{d<x^\alpha\mathcal{L}^{-B_0}}\mu(d)\pi(x;d^2,k) + \sum_{x^\alpha\mathcal{L}^{-B_0}\le d<x^{1/2}}\mu(d)\pi(x;d^2,k)\\ &= S_\alpha + S_{\alpha,1/2}\end{aligned}$$

For the first part, we can write $\pi(x;d^2,k)$ as

$$\pi(x;d^2,k) = \frac{\rho_k(d)}{\varphi(d^2)}\mathrm{li}x + \left(\pi(x;d^2,k) - \frac{\rho_k(d)}{\varphi(d^2)}\mathrm{li}x\right)$$

Where

$$\rho_k(n) := \sum_{d|(n,k)}\mu(d)$$

is the characteristic function of the number n that relative prime with k. Thus we get

$$\begin{aligned}S_\alpha &= \mathrm{li}x\sum_{d<x^\alpha\mathcal{L}^{-B_0}}\frac{\rho_k(d)\mu(d)}{d\varphi(d)} + \sum_{d<x^\alpha\mathcal{L}^{-B_0}}\mu(d)\left(\pi(x;d^2,k) - \frac{\rho_k(d)\mathrm{li}x}{\varphi(d^2)}\right)\\ &= C_k\mathrm{li}x + O\left(\mathrm{li}x\sum_{d\ge x^\alpha\mathcal{L}^{-B_0}}\frac{\rho_k(d)\mu(d)}{\varphi(d^2)}\right)\\ &+ \sum_{d<x^\alpha\mathcal{L}^{-B_0}}\mu(d)\left(\pi(x;d^2,k) - \frac{\rho_k(d)}{\varphi(d^2)}\mathrm{li}x\right)\\ &= C_k\mathrm{li}x + O(S_1) + R_\alpha\end{aligned}$$

Where

$$C_k = \sum_{d=1}^{\infty}\frac{\rho_k(d)\mu(d)}{\varphi(d^2)} = \prod_{p|k}\left(1 + \frac{1}{p^2-p-1}\right)\prod_p\left(1 - \frac{1}{p(p-1)}\right)$$

2.1 Upper bound for $S_{\alpha,1/2}$, S_1 and R_α

Now, we are going to find the upper bound for

$$S_{\alpha,1/2} = \sum_{x^\alpha\mathcal{L}^{-B_0}\le d<x^{1/2}}\mu(d)\pi(x;d^2,k)$$

We only need to consider the trivial bound

$$\pi(x;d^2,k) \ll \frac{x}{d^2}$$

So we obtain

$$\begin{aligned}|S_{\alpha,1/2}| &= \left| \sum_{x^{\alpha}\mathcal{L}^{-B_0} \le d < x^{1/2}} \mu(d)\pi(x; d^2, k) \right| \\ &\ll x \sum_{x^{\alpha}\mathcal{L}^{-B_0} \le d < x^{1/2}} d^{-2} \\ &\ll x \int_{x^{\alpha}\mathcal{L}^{-B_0}}^{x^{1/2}} \frac{\mathrm{d}t}{t^2} \\ &\ll x^{1-\alpha}(\log x)^{B_0} - x^{1/2}\end{aligned}$$

Of course we let $1/2 > \alpha > 0$, namely $|S_{\alpha,1/2}| \ll x^{1-\alpha}(\log x)^{B_0}$. Now, we are going to find the upper bound for

$$S_1 = \mathrm{li} x \sum_{n \ge x^{\alpha}\mathcal{L}^{-B_0}} \frac{\rho_k(n)\mu(n)}{\varphi(n^2)}$$

In fact, this is very easy ,since $\varphi(d^2)$ is approximately equals to d^2 for square free number d. Thus S_1 will goes to zero in some sense like $O(x^{1-\alpha+\epsilon})$ as $x \to \infty$.

Lemma 2.1. *If d is a square free number, then we have*

$$\varphi(d) \gg \frac{d}{(\log\log_2 d \log\log\log_2 d)^{c_2}}$$

where $c_2 > 0$ is some constant.

Proof.

$$\begin{aligned}\varphi(d) = \varphi\left(\prod_{p|d} p\right) &= d\prod_{p|d}\left(1 - \frac{1}{p}\right) \ge d \prod_{p \le p_{\omega(d)}} \left(1 - \frac{1}{p}\right) \\ &= d\exp\left(\sum_{p \le p_{\omega(d)}} \log(1 - 1/p)\right) \\ &\ge d\exp\left(\sum_{p \le p_{\omega(d)}} -\frac{2\log 2}{p}\right) \\ &= d\exp\left(-2\log 2\left(\log\log p_{\omega(d)} + c_1 + O(1/\log p_{\omega(d)})\right)\right) \\ &\gg de^{-c_2(\log\log\log_2 d + \log\log\log\log_2 d)} \\ &= \frac{d}{(\log\log_2 d \log\log\log_2 d)^{c_2}}\end{aligned}$$

□

Lemma 2.2.

$$\mathrm{li}x \sum_{n\geq x^\alpha \mathcal{L}^{-B_0}} \frac{\rho_k(n)\mu(n)}{\varphi(n^2)} \ll \frac{x^{1-\alpha}\left(\log\log_2 x \log\log\log_2 x\right)^{c_2}}{(\log x)^{1-B_0}} \ll x^{1-\alpha}\mathcal{L}^{B_0}$$

Proof. Because $\mu(n) = 0$ if n is not square free, we can assume n is square free. We have

$$\varphi(n^2) = n\varphi(n)$$

And

$$\sum_{n\geq x} \frac{1}{n^2} \leq \int_{x-1}^{\infty} \frac{\mathrm{d}t}{t^2} \ll \frac{1}{x}$$

Thus

$$\left|\sum_{n\geq x^\alpha \mathcal{L}^{-B_0}} \frac{\rho_k(n)\mu(n)}{\varphi(n^2)}\right| \leq \sum_{n\geq x^\alpha \mathcal{L}^{-B_0}} \frac{1}{n\varphi(n)} \ll \frac{\left(\log\log_2 x \log\log\log_2 x\right)^{c_2}}{x^\alpha (\log x)^{-B_0}}$$

And notice that $\mathrm{li}x \ll x/\log x$, which completes the proof. □

Now we are going to consider the R_α.

$$R_\alpha = \sum_{d<x^\alpha \mathcal{L}^{-B_0}} \mu(d)\left(\pi(x;d^2,k) - \frac{\rho_k(d)}{\varphi(d^2)}\mathrm{li}x\right)$$

where the constant $B_0 = B(A)/2$ is the constant in the Bombieri-Vinogradov Theorem. Now we let $\alpha = 1/4$, by the Bombieri-Vinogradov Theorem we simply obtain

$$\begin{aligned} |R_1/4| &= \left|\sum_{d<x^{1/4}/(\log x)^{B_0}} \mu(d)\left(\pi(x;d^2,k) - \frac{\rho_k(d)}{\varphi(d^2)}\mathrm{li}x\right)\right| \\ &\ll \left|\sum_{\substack{d^2<x^{1/2}\mathcal{L}^{-2B_0} \\ \rho_k(d)=1}} \pi(x;d^2,k) - \frac{\mathrm{li}x}{\varphi(d^2)}\right| + \left|\sum_{\substack{d^2<x^{1/2}\mathcal{L}^{-2B_0} \\ \rho_k(d)=0}} \pi(x;d^2,k)\right| \\ &\ll \frac{x}{(\log x)^A} \end{aligned}$$

2.2 Completion

Combining the results we got, for any $A \geq 2$, $k \in \mathbb{N}^+$ and x sufficiently large, we have

$$\begin{aligned}
\mathcal{Q}_k(x) &= C_k \mathrm{li}x + O(S_1) + S_{1/4,1/2} + R_1/4 \\
&= C_k \mathrm{li}x + O\left(x^{3/4}(\log x)^{B_0} + x^{3/4}(\log x)^{B_0} + \frac{x}{(\log x)^A}\right) \\
&= \mathrm{li}x \prod_{p|k}\left(1 + \frac{1}{p^2 - p - 1}\right)\prod_p\left(1 - \frac{1}{p(p-1)}\right) + O\left(\frac{x}{(\log x)^A}\right)
\end{aligned}$$

Therefore

$$\mathcal{Q}_k(x) = \mathrm{li}x \prod_{p|k}\left(1 + \frac{1}{p^2 - p - 1}\right)\prod_p\left(1 - \frac{1}{p(p-1)}\right) + O\left(\frac{x}{(\log x)^A}\right)$$

This is the result we desire. As a consequence, we have

$$\lim_{x\to\infty} \frac{\mathcal{Q}_k(x)}{\mathrm{li}x} = C_k > 0$$

That is, the primes that $p - k$ square-free has the positive density in the primes.

Appendix Table of $\mathcal{Q}_k(x)$ For $k = 1, 2, 3$

We made a program using C++ on our computer to calculate some numerical value of $\mathcal{Q}_k(x)$ (for $k = 1, 2, 3$ and $x \leq 10^7$).

x	$Q_1(x)$	$C_1 \mathrm{li} x$	${C_1}^{-1} Q_1(x)/\mathrm{li} x$	$Q_2(x)$	$Q_3(x)$
10	3	1.9148	1.5667	3	1
50	8	6.5156	1.2278	11	6
100	13	10.875	1.1954	20	10
500	40	37.676	1.0617	74	41
1000	68	66.027	1.0299	127	74
5000	255	255.50	0.99804	506	295
1×10^4	467	465.61	1.0030	925	548
5×10^4	1943	1931.7	1.0058	3841	2280
1×10^5	3599	3600.7	0.99953	7175	4292
5×10^5	15602	15559	1.0028	31020	18603
1×10^6	29397	29403	0.99980	58653	35153
5×10^6	130391	130375	1.0001	260381	156249
1×10^7	248518	248650	0.99947	496848	298075

References

[1] G. Tenenbaum, *Introduction to Analytic and Probabilistic Number Theory*

[2] Dimitris Koukoulopoulos, *Sieve Methods*

[3] Melvyn B. Nathanson, *Elementary Methods in Number Theory*

[4] Kenneth Ireland and Michael Rosen, *A Classical Introduction to Modern Number Theory (Second Edition)*

[5] Karatsuba, *Basic Analytic Number Theory*