Sieving Positive Integers by Primes

Samir Brahim Belhaouari

Correspondence: Information & Computing Technology, College of Science and Engineering, Hamad Bin Khalifa University, Qatar.

Received: November 8, 2016   Accepted: January 1, 2019   Online Published: January 9, 2019
doi:10.5539/jmr.v11n1p17   URL: https://doi.org/10.5539/jmr.v11n1p17

Abstract

Let $Q$ be a set of primes and let $\Psi(x, y, Q)$ be the number of positive integers less than or equal to $x$ that have no prime factors from $Q$ exceeding the integer $y$. We have enhanced an asymptotic formula for $\Psi(x, y, Q)$ after highlighting some properties of the function $\Psi$.

Keywords: sieving, prime number

1. Introduction

The distribution of positive integers without large prime factors has been deeply investigated, and has linked to several applications and to various problems in number theory like for example, in the study of large gaps between primes, of values of character sums, of Waring’s problem, of primality testing and factoring algorithm, of $S$-unit equations, and of Fermat’s Last Theorem.

This article is an extended version of the papers Warlimont, (Warlimont, 1990), and the paper of Goldston and McCurley (Goldston & McCurley, 1988), where they have worked on the estimation of the number of integers that are co-prime with a certain set of primes. Their results are summarized in the following theorem

**Theorem 1.1.** (see Theorem 1, (Warlimont, 1990)) Let $Q$ be a set of primes with the following property: There is some $\gamma > 0$ and some $A > 0$ such that

$$
\sum_{p \leq x} \ln(p) = \delta x + O\left(x \ln^{-A}(x)\right)
$$

(1)

Denote by $\Psi(x, y, Q)$ the number of all positive integers less than $x$ which do not have a prime divisor which it big than $y$ and belongs to $Q$.

Put $u = \frac{\ln(x)}{\ln(y)}$ and let $R(x)$ be given by

$$
R(x) = \begin{cases} 
\ln^{-A}(x), & \text{if } A < 1; \\
(\ln \ln(x)) \ln^{-1}(x), & \text{if } A = 1; \\
\ln^{-1}(x), & \text{if } A > 1.
\end{cases}
$$

Then

$$
\Psi(x, y, Q) = \frac{e^{xy}}{\Gamma(1 - \delta)} x \prod_{y < p \leq x} \left(1 - \frac{1}{p}\right)\left(1 + O(u^{-1} + R(x))\right)
$$

uniformly in $1.5 \leq y \leq x$. The $O$-constant depends only on $A$, $\delta$ and the $O$-constant in (1).

This result is compared with the only corollary in (Goldston & McCurley, 1988), which states that if $u \to \infty$

$$
\Psi(x, y, Q) \sim \frac{e^{xy}}{\Gamma(1 - \delta)} x \prod_{y < p \leq x} \left(1 - \frac{1}{p}\right)
$$

when the set of primes $Q$ has relative density $\delta$ among the primes, more precisely

$$
\sum_{\substack{p \leq x \\ p \notin Q}} \ln(p) = (1 - \delta)x + x\varepsilon(x), \ x \geq 1, \text{ and } 0 < \delta < 1,
$$
where \( Q^c \) denotes the set of primes that do not belong to \( Q \), we shall assume three existing conditions

\[
\begin{align*}
|e(x)| & \leq B(x); \\
B(x) & \text{ is nondecreasing function for } x \geq 1; \\
B(x) & = O(\ln^{-A}(x)) \text{ as } x \to \infty \text{ for some constant } A > 1.
\end{align*}
\]

In this paper, an estimate of the function \( \Psi(x, y, Q) \) not only for the boundary missing cases, when \( \delta \) equals to 0 or 1, but an estimation of the function \( \Psi(x, y, Q) \) independently of the value \( \delta \) is given as well. The following theorem will summarize our results along with interesting properties of the function \( \Psi \) as it is indicated below.

**Theorem 1.2.** Let \( Q \) be a set of primes that are bigger than certain integer \( y \) strictly bigger than 1. Put \( P_Q := \prod_{p \in Q} p \).

If \( y < \min \{ p_i \mid p_i \in Q \} \), then

(i) For any positive integers \( k \)

\[
\Psi(x, y, Q) = \text{card} \{ i \in (y + kP_Q, x + kP_Q) \mid \gcd(i, P_Q) = 1 \} + y.
\]

(ii) For any positive integers \( k \)

\[
\Psi(x, y, Q) = \Psi(x + kP_Q, y, Q) - \Psi(y + kP_Q, y, Q) + y,
\]

and if \( x < P_Q \) and \( \gcd(x, P_Q) = 1 \) then

\[
\Psi(x, y, Q) + \Psi(P_Q - x, y, Q) = \Psi(P_Q, y, Q) + 1 = P_Q \prod_{p \in Q} \left( 1 - \frac{1}{p} \right) + 1.
\]

(iii) Uniformly for \( 1.5 \leq y \leq x \)

\[
\Psi(x, y, Q) = (x - y) \prod_{p \in Q} \left( 1 - \frac{1}{p} \right) (1 + O(R(y))) + y,
\]

where \( R(y) = \frac{1}{\ln(y)} \).

(iv) If \( x \) and \( y \) are sufficiently large and \( Q = \{ p_i \mid y < p_i \leq x \} \) then

\[
\Psi(x, y, Q) \sim (x - y) \frac{\ln(y)}{\ln(x)} + y,
\]

and if \( y/x \to 0 \) then

\[
\Psi(x, y, Q) \sim \pi(x, y/\ln(y)) \sim \frac{\ln(y)}{\ln(x)}.
\]

Integers without small prime factors can be viewed as an approximation to prime numbers. While integers without large prime factors, also called smooth integers, are necessarily products of small prime factors and are like the exact opposite of prime numbers. It turns out that these smooth numbers play an important role in many areas of prime number theory, in particular in the construction of large gaps between primes see (Rankin, 1938) and in the analysis of algorithms of factorization see (Pomerance, 1987). The long-standing conjecture of Carmichael (Alford, Garanville, & Pomerance, 1987) was resolved by the help of integers of the form \( p + 1 \) without large prime factors.

The case where the set \( Q \) contains only consecutive prime numbers need to be considered, and denoted by \( Q_{n,m} = \{ p_i \mid n \leq i < n + m \} \). An interesting method of proving the following Lemma is given at the end of this paper.

**Corollary 1.3.** Let \( Q_{n,m} = \{ p_i \mid n \leq i < n + m \} \) and let \( p_N \) be the largest prime less than the integer \( x \). Then for all integers \( x > p_n \) and for all \( m \leq N - n \),

\[
\Psi(x, p_{n-1}, Q_{n,m}) = x \prod_{i=n}^{n+m-1} \left( 1 - \frac{1}{p_i} \right) + O \left( \frac{x}{\ln(x) \ln(n + m)} \right).
\]

If \( n \) sufficiently large, then

\[
\Psi(x, p_{n-1}, Q_{n,N-n}) = x \frac{\ln(n)}{\ln(x)} + O \left( \frac{x}{\ln^2(x)} \right) - x \frac{\ln(n)}{\ln(x)}.
\]
Many authors worked on the estimation of the function \( \Psi \) under different conditions and by using different technique, like "Rankin’s method" (Rankin, 1938), or "Probabilistic and extrapolation’s method" (Halberstam & Richard, 1974). An asymptotic formula for the function \( \Psi \) was first obtained by Dickman (Dickma, 1930), where he established an asymptotic relation between \( \Psi \) and the unique continuous solution to the following differential-difference equation

\[
u(p) = -\nu(u - 1) \quad (u > 1),
\]
satisfying the initial condition

\[
\rho(u) = 1 \quad (0 \leq u \leq 1).
\]

An accurate proof, by modern standards, was later found by Chowla and Buchstab see (Chowla & Buchstab, 1947); The first quantitative estimation results were found by Ramaswami (Ramaswami, 1949) and Buchstab (Buchstab, 1949).

Another progress was made in the 1950s by Bruijn, where he proved a uniform version of Dickman’s result (Bruijn, 1951), then later it was slightly improved by Hensley (Hensley, 1985) and Hildebrand (Hilderbrand, 1986). General Asymptotic results have been given by Specht (Specht, 1949), Hornfeck (Hornfeck, 1959), Beukers (Beukers, 1975), Tenenbaum (Tenenbaum, 1990) and Granville (Granville, 1991). For more details see (Hildebrand & Tenenbaum, 1993).

2. Proof of Theorem 1.2

Let \( \pi(x, n) \) be the co-prime-counting function that give the cardinality of the set of numbers that are co-prime to \( p_i \), \( i > n \), and are less than or equal to the integer \( x \). Now let \( N \) be the rank of largest prime less than \( x \), then we can observe that

\[
\pi(x, 1) = \pi(x) \quad \pi(x, N) = x
\]

where \( \pi(x) \) denotes the function counting the number of prime numbers less than or equal to \( x \) and without losing generality we can assume that \( p_1 = 1 \).

Number of co-prime numbers with \( p_N \) and less than \( x \) is calculated by subtracting all multiples of \( p_N \) from \( x \) integers,

\[
\pi(x, N - 1) = x - \left\lfloor \frac{x}{p_N} \right\rfloor,
\]

by the same way we can find number of co-prime numbers with \( p_N \) and \( p_{N-1} \) can be calculated as

\[
\pi(x, N - 2) = x - \left\lfloor \frac{x}{p_N} \right\rfloor - \left\lfloor \frac{x}{p_{N-1}} \right\rfloor + \left\lfloor \frac{x}{p_N p_{N-1}} \right\rfloor,
\]

where we asymptotically the formula becomes

\[
\pi(x, N - 2) \sim x \left( 1 - \frac{1}{p_N} \right) \left( 1 - \frac{1}{p_{N-1}} \right).
\]

By iteration, we can find \( \pi(x, n) \) as follows

\[
\pi(x, n) = x - \sum_{j=1}^{N} (-1)^j \left( \frac{x}{p_{\pi+1}} \right) \left( \frac{x}{\prod_{i=1}^{j} p_{\pi+i}} \right).
\]

We can generate this to

\[
\Psi(x, y, Q) = x - \sum_{j=1}^{N_Q} (-1)^j \left( \frac{x}{\prod_{i=1}^{j} q_{\pi+i}} \right),
\]

where \( q_i \) are primes inside the set \( Q \) and \( N_Q = Card(Q) \).

2.1 Proof of the First Part of Theorem 1.2

The first part of Theorem 1.2 can be deduced by the fact that for all positive integers \( i \) and for all primes, \( p \in Q \), we have

\[
y + i \equiv (y + kP_Q + i) \mod p,
\]
Thus

\[ \Psi(x, y, Q) = \text{Card} \{ i \in (y + kP_Q, x + kP_Q) \mid \gcd(i, P_Q) = 1 \} + y. \]

2.2 Proof of the Second Part of Theorem 1.2

The second part of Theorem 1.2 can be proved by using the the equation 2 and the fact that for all integers \( x \)

\[ \left\lceil \frac{x + kP_Q}{\prod_{j=1}^{l} q_i} \right\rceil = \left\lceil \frac{x}{\prod_{j=1}^{l} q_i} \right\rceil + \frac{kP_Q}{\prod_{j=1}^{l} q_i} \]

as follows

\[ \Psi(x + kP_Q, y, Q) = x + kP_Q - \sum_{j=1}^{N_Q} (-1)^j \left( \sum_{1 \leq i_1 < i_2 < \cdots < i_l \leq N} \left\lceil \frac{x + kP_Q}{\prod_{j=1}^{l} q_i} \right\rceil \right) \]

\[ = x + kP_Q - \sum_{j=1}^{N_Q} (-1)^j \left( \sum_{1 \leq i_1 < i_2 < \cdots < i_l \leq N} \left( \frac{x}{\prod_{j=1}^{l} q_i} \right) + \frac{kP_Q}{\prod_{j=1}^{l} q_i} \right) \]

and similarly

\[ \Psi(y + kP_Q, y, Q) = y + kP_Q - \sum_{j=1}^{N_Q} (-1)^j \left( \sum_{1 \leq i_1 < i_2 < \cdots < i_l \leq N} \left( \frac{y}{\prod_{j=1}^{l} q_i} \right) + \frac{kP_Q}{\prod_{j=1}^{l} q_i} \right) \]

By Subtracting the last equations from the first we find that

\[ \Psi(x + kP_Q, y, Q) - \Psi(y + kP_Q, y, Q) = x - \sum_{j=1}^{N_Q} (-1)^j \left( \sum_{1 \leq i_1 < i_2 < \cdots < i_l \leq N} \left( \frac{x}{\prod_{j=1}^{l} q_i} \right) \right) - y \]

\[ = \Psi(x, y, Q) - y, \]

where \( \sum_{j=1}^{N_Q} (-1)^j \left( \sum_{1 \leq i_1 < i_2 < \cdots < i_l \leq N} \left( \frac{x}{\prod_{j=1}^{l} q_i} \right) \right) = 0 \) since \( y < \min(Q) \).

We can find also if \( x < P_Q \) and \( \gcd(x, P_Q) = 1 \) then

\[ \Psi(x, y, Q) + \Psi(P_Q - x, y, Q) = \Psi(P_Q, y, Q) + 1 = P_Q \prod_{p \in Q} \left( 1 - \frac{1}{p} \right). \]

Thus using same the equation 2 we can write the following equations

\[ \Psi(P_Q - x, y, Q) + \Psi(y, y, Q) = P_Q - \sum_{j=1}^{N_Q} (-1)^j \left( \sum_{1 \leq i_1 < i_2 < \cdots < i_l \leq N} \left( \frac{P_Q - x}{\prod_{j=1}^{l} q_i} \right) \right) \]

\[ = - \sum_{j=1}^{N_Q} (-1)^j \left( \sum_{1 \leq i_1 < i_2 < \cdots < i_l \leq N} \left( \frac{x}{\prod_{j=1}^{l} q_i} \right) \right) + P_Q - \sum_{j=1}^{N_Q} (-1)^j \left( \sum_{1 \leq i_1 < i_2 < \cdots < i_l \leq N} \left( \frac{P_Q}{\prod_{j=1}^{l} q_i} \right) \right) \]

\[ = P_Q - \sum_{j=1}^{N_Q} (-1)^j \left( \sum_{1 \leq i_1 < i_2 < \cdots < i_l \leq N} \left( \frac{P_Q}{\prod_{j=1}^{l} q_i} \right) \right) + 1 \]

\[ = P_Q \prod_{p \in Q} \left( 1 - \frac{1}{p} \right) + 1 \]
2.3 Proof of the Third Part of Theorem 1.2

The third part of Theorem 1.2 will be proved by improving Theorem 1 in (Warlimont, 1990) and the proposition 3 in (Goldston & McCurley, 1988), where it stated that: There is some \( \delta, 0 \leq \delta < 1 \), and some \( A > 0 \) such that

\[
\sum_{p \leq x} \ln(p) = \delta x + O(x \ln^{-A}(x)).
\]

Then

\[
\Psi(x, y, Q) = \frac{e^{y/\delta}}{\Gamma(1-\delta)} x \prod_{\substack{p \leq x \\cap \mathbb{P} \\leq Q}} \left(1 - \frac{1}{p}\right) \left(1 + O(D)\right),
\]

where \( D := \frac{\ln(y)}{\ln(x)} + \ln^{-B}(y) \) and \( B := \min(A, 1) \).

It is possible to choose the integer \( k \) large enough in order to verify the following equality

\[
\sum_{p \leq x} \ln(p) = O((y + kP_{Q}) \ln^{-A}(y + kP_{Q})),
\]

where the value \( \delta \) equals to zero and \( A > 1 \).

Therefore, it implies that

\[
\Psi(x + kP_{Q}, y, Q) = (x + kP_{Q}) \prod_{\substack{y < p \leq x \\cap \mathbb{P} \leq Q}} \left(1 - \frac{1}{p}\right) \left(1 + O(D_{1})\right),
\]

and

\[
\Psi(y + kP_{Q}, y, Q) = (y + kP_{Q}) \prod_{\substack{y < p \leq x \\cap \mathbb{P} \leq Q}} \left(1 - \frac{1}{p}\right) \left(1 + O(D_{2})\right),
\]

where \( D_{1} \) and \( D_{2} \) are functions of same order as the function \( D \), i.e., \( D_{i} \sim D \), for \( i = 1, 2 \). Then we can find for \( k \) large enough that

\[
\Psi(x + kP_{Q}, y, Q) - \Psi(y + kP_{Q}, y, Q) = (x - y) \prod_{\substack{y < p \leq x \\cap \mathbb{P} \leq Q}} \left(1 - \frac{1}{p}\right) \left(1 + O\left(\frac{1}{\ln(y)}\right)\right)\]

\[
= \Psi(x, y, Q) - y.
\]

Then we conclude that

\[
\Psi(x, y, Q) = (x - y) \prod_{\substack{y < p \leq x \\cap \mathbb{P} \leq Q}} \left(1 - \frac{1}{p}\right) \left(1 + O\left(\frac{1}{\ln(y)}\right)\right) + y.
\]

2.4 Proof of the Fourth Part of Theorem 1.2

To Prove the last point of the Theorem 1.2, it suffices to notice that \( p_{i} \sim i \ln(i) \) by Prime Number Theorem in order to quantify

\[
\prod_{\substack{y < p \leq x \\cap \mathbb{P} \leq Q}} \left(1 - \frac{1}{p}\right)
\]

by simple integration as follows
\[
\prod_{\begin{subarray}{c} y < p < x \\ p \in Q \end{subarray}} \left(1 - \frac{1}{p}\right) = \exp \left( \sum_{\begin{subarray}{c} y < p < x \\ p \in Q \end{subarray}} \ln \left(1 - \frac{1}{p}\right) \right) \\
= \exp \left( \sum_{\begin{subarray}{c} y < p < x \\ p \in Q \end{subarray}} \ln \left(1 - \frac{1}{\ln(p) + o(\ln(n))}\right) \right), \text{ see (G. H. Hardy and E. M. Wright, 1979)} \\
\sim \exp \left( \sum_{\begin{subarray}{c} y < p < x \\ p \in Q \end{subarray}} \left(-\frac{1}{\ln(i)} + O\left(\frac{1}{i^2 \ln(i)}\right)\right) \right) \text{ from Taylor series of } \ln(1 + x) \\
\sim \exp \left( \int_{\ln(y)}^{\ln(x)} \left(-\frac{1}{w \ln(w)}\right) \text{d}w \right) \\
\sim \frac{\ln(x)}{\ln(y)}. \\
\]

3. Proof of Corollary 1.3

In this section, a description of an approach in how to estimate the function \(\Psi\) when the set \(Q\) contains only consecutive prime numbers. Suppose that for all integers \(x > p_n\) we have for \(Q_m = \{p_i \mid n \leq i < n + m\}\) and \(p_N\) representing the largest prime less than \(x\)

\[
\Psi(x, p_{n-1}, Q_{n,m}) = x \prod_{i=n}^{n+m-1} \left(1 - \frac{1}{p_i}\right) + O(f(x, m, n)),
\]

for all \(m \leq N - n\) and some error function \(f(x, m, n)\)

Using a proof by induction, the function \(f(x, m)\) can be estimated as a linear function with \(\ln^{-1}(m + n)\).

For \(m = 1\), we have

\[
\Psi(x, p_{n-1}, Q_{n,1}) = x \prod_{i=n}^{n} \left(1 - \frac{1}{p_i}\right) + O(f(x, 1)) \\
= x - \frac{x}{p_n} + O(f(x, 1)),
\]

The error function, \(O(f(x, 1))\), can be calculated easily as follows

\[
O(f(x, 1)) = x - \left\lfloor \frac{x}{p_n} \right\rfloor - \left(x - \frac{x}{p_n}\right) \\
= \frac{x}{p_n} - \left\lfloor \frac{x}{p_n} \right\rfloor \\
\leq 1.
\]

Now assume that

\[
\Psi(x, p_{n-1}, Q_{n,m}) = x \prod_{i=n}^{n+m-1} \left(1 - \frac{1}{p_i}\right) + O(f(x, m))
\]

then we need to show
\[
\Psi(x, p_{n-1}, Q_{n,m}) = x \prod_{i=n}^{n+m} \left(1 - \frac{1}{p_i}\right) + O(f(x, m + 1)).
\]

Let \( A_i \) be the set of multiple of the prime number \( p_{n+i-1} \) and are less than \( x \), e.i., \( A_i = \{ i \mid \gcd(i, p_{n+i-1}) = p_{n+i-1} \text{ and } i \leq x \} \), then

\[
\Psi(x, p_{n-1}, Q_{n,m+1}) = x - \text{Card} \left( \bigcup_{i=1}^{m+1} A_i \right)
\]

\[
= x - \text{Card} \left( \bigcup_{i=1}^{m} A_i \right) \bigcup A_{m+1}
\]

\[
= x - \text{Card} \left( \bigcup_{i=1}^{m} A_i \right) - \text{Card}(A_{m+1}) + \text{Card} \left( \bigcup_{i=1}^{m} A_i \bigcap A_{m+1} \right)
\]

Note that \( A_i \cap A_{m+1} = \{ i \mid \gcd(i, p_{n+m}p_{n+i-1}) = p_{n+m}p_{n+i-1} \text{ and } i \leq x \} \) and \( \Psi(x, p_{n-1}, Q_{n,m}) = x - \text{Card} \left( \bigcup_{i=1}^{m} A_i \right) \) then

\[
\Psi(x, p_{n-1}, Q_{n,m+1}) = \Psi(x, p_{n-1}, Q_{n,m}) - \frac{x}{p_{n+m}} - \frac{x}{p_{n+m}} - \Psi \left( \frac{x}{p_{n+m}}, p_{n-1}, Q_{n,m} \right)
\]

\[
= x \prod_{i=n}^{n+m-1} \left(1 - \frac{1}{p_i}\right) - \frac{x}{p_{n+m}} + \frac{x}{p_{n+m}} - \frac{x}{p_{n+m}} \prod_{i=n}^{n+m-1} \left(1 - \frac{1}{p_i}\right)
\]

\[
+ O(f(x, m) - f(x/p_{n+m}, m))
\]

\[
= x \prod_{i=n}^{n+m} \left(1 - \frac{1}{p_i}\right) + O(f(x, m) - f(x/p_{n+m}, m) + 1)
\]

In order to force the last formula to be correct we have to find the function, \( f \), that verifies the following equality

\[
f(x, m) + f(x/p_{n+m}, m) = f(x, m + 1) + O(1).
\]

For \( n \) a large enough integer, we obtain that

\[
-f(x/p_{n+m}, m) = f(x, m + 1) - f(x, m)
\]

\[
\sim \frac{\partial f(x, m)}{\partial m}
\]

Since \( p_{n+m} \sim (n + m) \ln(n + m) \) then the following differential equation needs to be solved

\[
\frac{\partial f(x, m)}{\partial m} = -f \left( \frac{x}{(n + m) \ln(n + m)}, m \right).
\] (3)

The function

\[
f(x, m) \sim \frac{x}{\ln(x) \ln(n + m)}
\]

is an asymptotic solution of the differential equation 3 because of the following
where \( c \) is a fixed constant bigger than one.

The existence of better asymptotic solution function \( f \), \( f = o \left( \frac{x}{\ln(x) \ln(n+m)} \right) \), is still remained.

Thus, for all \( m \leq N - n \).

\[
\Psi(x, p_{n-1}, Q_{n,m}) = x \prod_{i=m}^{n+m-1} \left( 1 - \frac{1}{p_i} \right) + O \left( \frac{x}{\ln(x) \ln(n+m)} \right),
\]

And if \( n \) is large enough, we can write

\[
\Psi(x, p_{n-1}, Q_{n,N-n}) = x \frac{\ln(n)}{\ln(x)} + O \left( \frac{x}{\ln^2(x)} \right).
\]

References


Hensley, D. (1985). The number of positive integers \( \leq x \) and free of prime divisors \( > y \), *J. Number Theory*, 21(21), 286-298.


**Copyrights**

Copyright for this article is retained by the author(s), with first publication rights granted to the journal. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/4.0/).