Journal of Integer Sequences, Vol. 22 (2019),
Article 19.5.7

On a Family of Functions Defined Over Sums of Primes

Christian Axler
Department of Mathematics
Heinrich-Heine-University
40225 Düsseldorf
Germany
christian.axler@hhu.de

Abstract

Let r and m be real numbers so that the sum $S_{r,m}(x) = \sum_{p \leq x} p^r \log^m p$ diverges as $x \to \infty$. Here p runs over all primes not exceeding x. In this paper, we give an asymptotic formula for each $S_{r,m}(x)$ as $x \to \infty$. The case where x is the nth prime number is of particular interest. Here we use a method developed by Salvy to give an asymptotic formula for $S_{r,m}(p_n)$ as $n \to \infty$, which generalizes, for instance, the previously known one for $S_{1,0}(p_n)$, the sum of the first n prime numbers.

1 Introduction

Let m and r be real numbers and let $S_{r,m} : [0, \infty) \to \mathbb{R}$ be defined by

$$S_{r,m}(x) = \sum_{p \leq x} p^r \log^m p,$$

where p runs over all primes not exceeding x. The sum $S_{r,m}(x)$ diverges as $x \to \infty$ if and only if

$$\text{(i) } r > -1 \text{ and } m \in \mathbb{R} \qquad \text{or} \qquad \text{(ii) } r = -1 \text{ and } m \geq 0.$$

The aim of this paper is to find the asymptotic behaviour of the sum $S_{r,m}(x)$ in the case where m and r satisfy conditions (i) or (ii). First, we study the case (i). Let $\pi(x)$ denote the number of primes not exceeding x. A well-known result (see [10]) concerning this function is the *prime number theorem*, which states that

$$\pi(x) = \mathrm{li}(x) + O(xe^{-a\sqrt{\log x}}) \tag{1}$$

as $x \to \infty$, where a is a positive absolute constant, and the *logarithmic integral* $\mathrm{li}(x)$ is defined for every real $x \geq 0$ as follows:

$$\mathrm{li}(x) = \int_0^x \frac{dt}{\log t} = \lim_{\varepsilon \to 0+} \left\{ \int_0^{1-\varepsilon} \frac{dt}{\log t} + \int_{1+\varepsilon}^x \frac{dt}{\log t} \right\}. \tag{2}$$

Denoting the sum of the first prime numbers not exceeding x by $S(x)$, Szalay [9, Lemma 1] used (1) and Stieltjes integration to find

$$S(x) = \mathrm{li}(x^2) + O(x^2 e^{-a\sqrt{\log x}}) \tag{3}$$

as $x \to \infty$. The case $x = p_n$, where p_n denotes the nth prime number, is of particular interest. Here, $S(x) = \sum_{k \leq n} p_k$ is equal to the sum of the first n prime numbers. Massias and Robin [4, p. 217] found that

$$\sum_{k=1}^{n} p_k = \mathrm{li}((\mathrm{li}^{-1}(n))^2) + O(n^2 e^{-c\sqrt{\log n}}) \tag{4}$$

as $n \to \infty$, where c is a positive absolute constant and $\mathrm{li}^{-1}(x)$ is the inverse function of $\mathrm{li}(x)$. Then they [4, p. 217] used (4) and a result of Robin [7] to derive the asymptotic expansion

$$\sum_{k=1}^{n} p_k = \frac{n^2}{2} \left(\log n + \sum_{i=0}^{N} \frac{A_{i+1}(\log\log n)}{\log^i n} \right) + O_N\left(\frac{n^2 (\log\log n)^{N+1}}{\log^{N+1} n} \right) \tag{5}$$

as $n \to \infty$, where N is a nonnegative integer and the polynomials A_k satisfy the formulas $A_0(x) = 1$ and

$$A'_{i+1} = A'_i - (i-1)A_i. \tag{6}$$

It follows that $\deg(A_0) = 0$, $\deg(A_1) = 1$, and $\deg(A_i) = i - 1$ for every integer $i \geq 2$. Unfortunately, the recursive formula (6) for the derivatives does not yield a description of the polynomials A_i, since the constant coefficient of the polynomials A_i remains undetermined by this equation. This problem was fixed in [3, Theorem 1.4] by applying a method developed by Salvy [8, Theorem 2]. We use the same method to give the following result concerning the sum $S_{r,m}(p_n)$ in the case where $r > -1$ and $m \in \mathbb{R}$. Here, we use the notation

$$\binom{\delta}{0} = 1 \quad \text{and} \quad \binom{\delta}{k} = \frac{\delta(\delta-1)\cdots(\delta-k+1)}{k!}$$

for a real number δ and a positive integer k.

Theorem 1. *Let r and m be real numbers with $r > -1$ and let N be a nonnegative integer. As $n \to \infty$, we have*

$$\sum_{k=1}^{n} p_k^r \log^m p_k = \frac{n^{r+1} \log^{r+m} n}{r+1} \left(\sum_{i=0}^{N} \frac{A_{r,m,i}(\log\log n)}{\log^i n} + O_{r,m,N}\left(\frac{(\log\log n)^{N+1}}{\log^{N+1} n} \right) \right), \tag{7}$$

where the polynomials $A_{r,m,i} \in \mathbb{R}[x]$ are defined by

$$A_{r,m,0} = 1, \quad A'_{r,m,i+1} = A'_{r,m,i} + (m+r-i)A_{r,m,i}. \tag{8}$$

The polynomials $A_{r,m,i}$ can be computed explicitly. In particular,

$$A_{r,m,1}(x) = (r+m)x - \frac{m-1}{r+1} - r - 1$$

and

$$A_{r,m,2}(x) = \binom{r+m}{2} x^2 + \frac{(-r^3 - mr^2 + (-2m+3)r - m^2 + 2m)x}{r+1} + \lambda_{r,m},$$

where

$$\lambda_{r,m} = \frac{(m-1)(m-2)}{(r+1)^2} + \frac{r(r-3)}{2} - 2.$$

Remark 2. The polynomials A_i given in (6) and $A_{r,m,i}$ are connected by the formula $A_i = A_{1,0,i}$.

Remark 3. Since $S_{1,0}(p_n) = S(p_n)$, Theorem 1 yields a generalization of (5).

Remark 4. For some alternative asymptotic formulae for $S(p_n)$, see [2, Theorems 1 and 2].

In the second part of this paper, we study the case (ii); i.e. $r = -1$ and $m \geq 0$. If $m = 0$, we see that $S_{r,m}(x)$ is equal to the sum of the reciprocals of all prime numbers not exceeding x. Mertens [5, p. 52] proved that $\log\log x$ is the right order of magnitude for this sum by showing

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

as $x \to \infty$. Here B denotes the Mertens' constant and is defined by

$$B = \gamma + \sum_{p} \left(\log\left(1 - \frac{1}{p}\right) + \frac{1}{p} \right) = 0.261\ldots,$$

where $\gamma = 0.577\ldots$ denotes the Euler–Mascheroni constant. So it suffices to consider the case where $r = -1$ and $m > 0$. Here, we find the following result.

Theorem 5. *Let m be a positive real number and let N be a nonnegative integer. As $n \to \infty$, we have*

$$\sum_{k=1}^{n} \frac{\log^m p_k}{p_k} = \frac{\log^m n}{m} \left(\sum_{i=0}^{N} \frac{B_{m,i}(\log\log n)}{\log^i n} + O_{N,m}\left(\frac{(\log\log n)^{N+1}}{\log^{N+1} n} \right) \right), \tag{9}$$

where the polynomials $B_{m,i} \in \mathbb{R}[x]$ are defined by

$$B_{m,0} = 1, \quad B'_{m,i+1} = B'_{m,i} + (m-i)B_{m,i}.$$

The polynomials $B_{m,i}$ can be computed explicitly. For example, we have

$$B_{m,1}(x) = mx \qquad \text{and} \qquad B_{m,2}(x) = \binom{m}{2}x^2 + mx - m.$$

2 Proof of Theorem 1

In order to prove Theorem 1, we first note *Abel's summation formula*, which can be found in [1].

Lemma 6 (Abel's summation formula). *Let $a : \mathbb{N} \to \mathbb{C}$ be a function, and let $A(x) = \sum_{n \le x} a(n)$, where $A(x) = 0$ if $x < 1$. If g has a continuous derivative on the interval $[y, x]$, where $0 < y < x$, then*

$$\sum_{y<n\le x} a(n)g(n) = A(x)g(x) - A(y)g(y) - \int_y^x A(t)g'(t)\,\mathrm{d}t.$$

Proof. See [1, Theorem 4.2]. □

Using this Lemma, we get the following result.

Proposition 7. *Let r and m be real numbers with $r > -1$ and let N be a nonnegative integer. For a nonnegative integer j, we set*

$$\chi_{r,m,j} = \frac{(-1)^j j!}{(r+1)^j}\binom{m-1}{j}. \tag{10}$$

As $x \to \infty$, we have

$$\int_2^x t^r \log^{m-1} t\,\mathrm{d}t = \frac{x^{r+1}\log^{m-1} x}{r+1} \sum_{j=0}^{N} \frac{\chi_{r,m,j}}{\log^j x} + \chi_{r,m,N+1} \int_2^x t^r \log^{m-2-N} t\,\mathrm{d}t + O(1).$$

Proof. Integration by parts and induction. □

Remark 8. In the case where m is a positive integer and $N \geq m-1$, we get $\chi_{r,m,N+1} = 0$. Here, Proposition 7 gives

$$\int_2^x t^r \log^{m-1} t \,\mathrm{d}t = \frac{x^{r+1}\log^{m-1} x}{r+1} \sum_{j=0}^{N} \frac{\chi_{r,m,j}}{\log^j x} + O(1)$$

as $x \to \infty$.

Proposition 7 implies the following asymptotic formula.

Corollary 9. *Let r and m be real numbers with $r > -1$ and let N be a nonnegative integer. As $x \to \infty$, we have*

$$\int_2^x t^r \log^{m-1} t \,\mathrm{d}t = \frac{x^{r+1}\log^{m-1} x}{r+1} \left(\sum_{j=0}^{N} \frac{\chi_{r,m,j}}{\log^j x} + O_{r,m,N}\left(\frac{1}{\log^{N+1} x} \right) \right).$$

Proof. Using L'Hospital's rule, we see that

$$\int_2^x t^r \log^{m-2-N} t \,\mathrm{d}t = O_{r,m,N}\left(\frac{x^{r+1}}{\log^{N+2-m} x} \right)$$

as $x \to \infty$, and it suffices to apply this equation to Proposition 7. □

Now we apply Corollary 9 to find the following result.

Proposition 10. *Let r and m be real numbers with $r > -1$ and let N be a nonnegative integer. Then*

$$\sum_{p \leq x} p^r \log^m p = \frac{x^{r+1}\log^{m-1} x}{r+1} \left(\sum_{j=0}^{N} \frac{\chi_{r,m,j}}{\log^j x} + O_{r,m,N}\left(\frac{1}{\log^{N+1} x} \right) \right)$$

as $x \to \infty$, where $\chi_{r,m,j}$ is defined by (10).

Proof. For $x \geq 2$, let $R(x) = \pi(x) - \mathrm{li}(x)$. As $x \to \infty$, the asymptotic fomula (1) implies

$$R(x) = O(xe^{-a\sqrt{\log x}}) \tag{11}$$

for some positive absolute constant a. Furthermore, let $y = 3/2$, $g(x) = x^r \log^m x$, and

$$a(n) = \begin{cases} 1, & \text{if } n \text{ is prime;} \\ 0, & \text{otherwise.} \end{cases}$$

We use Lemma 6 to get

$$\sum_{p \leq x} p^r \log^m p = (\mathrm{li}(x) + R(x))x^r \log^m x - \int_2^x (\mathrm{li}(t) + R(t))t^{r-1}(r\log^m t + m\log^{m-1} t)\,\mathrm{d}t.$$

If we combine this with (11) and

$$\int_2^x t^r \log^{m-1} t \,\mathrm{d}t = \mathrm{li}(x)x^r \log^m x - \int_2^x \mathrm{li}(t)t^{r-1}(r\log^m t + m\log^{m-1} t)\,\mathrm{d}t + O(1)$$

as $x \to \infty$, we see that

$$\sum_{p\leq x} p^r \log^m p = \int_2^x t^r \log^{m-1} t\,\mathrm{d}t + O(x^{r+1}e^{-b\sqrt{\log x}}) + O\left(\int_2^x t^r e^{-b\sqrt{\log t}}\,\mathrm{d}t\right) \tag{12}$$

as $x \to \infty$, where the real number b satisfies $0 < b < a$. Using L'Hospital's rule, we get

$$\int_2^x t^r e^{-a\sqrt{\log t}}\,\mathrm{d}t = O(x^{r+1}e^{-a\sqrt{\log x}})$$

as $x \to \infty$. So we can rewrite (12) as

$$\sum_{p\leq x} p^r \log^m p = \int_2^x t^r \log^{m-1} t\,\mathrm{d}t + O(x^{r+1}e^{-b\sqrt{\log x}}) \tag{13}$$

as $x \to \infty$. Finally, we apply Corollary 9 to complete the proof. □

In the following corollary, we give a generalization of (3).

Corollary 11. *Let r be a real number with $r > -1$ and let m be a nonnegative integer. As $x \to \infty$, we have*

$$\sum_{p\leq x} \frac{p^r}{\log^m p} = \frac{(r+1)^m}{m!}\left(\mathrm{li}(x^{r+1}) - \sum_{j=0}^{m-1} \frac{j!\,x^{r+1}}{(r+1)^{j+1}\log^{j+1} x}\right) + O(x^{r+1}e^{-a\sqrt{\log x}}).$$

Proof. Induction over m and integration by parts gives

$$\int_2^x \frac{t^r}{\log^{m+1} t}\,\mathrm{d}t = \frac{(r+1)^m}{m!}\left(\mathrm{li}(x^{r+1}) - \sum_{j=0}^{m-1} \frac{j!\,x^{r+1}}{(r+1)^{j+1}\log^{j+1} x}\right) + O(1)$$

as $x \to \infty$. Now it suffices to apply the equation (13). □

To give a proof of Theorem 1, we also need the following result of Salvy [8]. Here we use the notation from Robin [7].

Proposition 12 (Salvy)**.** *Let $y = y(x)$ satisfies $e^y y^{-\alpha} D(1/y) \approx x$ as $x \to \infty$, with $D(u) = \sum_{n\geq 0} d_n u^n$ a formal power series, $\alpha \neq 0$, and $D(0) \neq 0$. Then for any formal power series G with nonzero constant term the following asymptotic expansion hold:*

$$e^{\beta y} y^\gamma G(1/y) \approx \left(\frac{x}{d_0}\right)^\beta (\log x)^{\alpha\beta+\gamma} \sum_{n\geq 0} \frac{Q_n(\log\log x)}{\log^n x} \qquad (x \to \infty).$$

Here Q_n are polynomials with $Q_0 = G(0)$ and $Q'_{n+1}/\alpha = Q'_n + (\alpha\beta + \gamma - n)Q_n$.

Proof. See [8, Theorem 2]. □

Now we can give a proof of Theorem 1.

Proof of Theorem 1. Let N be a nonnegative integer and let $D_N(u) = \sum_{j=0}^{N} j!\, u^j$. We define the formal power series

$$D(u) = \sum_{j=0}^{\infty} j!\, u^j.$$

Then $D(0) = 1$. First, we note that repeated integration by parts in (2) gives

$$\mathrm{li}(x) \approx \frac{x}{\log x} D\left(\frac{1}{\log x}\right) \tag{14}$$

as $x \to \infty$. For $x > 1$, the logarithmic integral $\mathrm{li}(x)$ is increasing with $\mathrm{li}((1,\infty)) = \mathbb{R}$. Thus, we can define the inverse function $\mathrm{li}^{-1} : \mathbb{R} \to (1,\infty)$ by

$$\mathrm{li}(\mathrm{li}^{-1}(x)) = x. \tag{15}$$

We combine (14) and (15) to obtain

$$e^y y^{-1} D(1/y) \approx n \tag{16}$$

as $n \to \infty$, where $y = \log \mathrm{li}^{-1}(n)$. Next, we define $\delta(n)$ by $p_n = \mathrm{li}^{-1}(n) + \delta(n)$. By Massias and Robin [4, p. 217], we have

$$\delta(n) = O(ne^{-c\sqrt{\log n}}) \tag{17}$$

as $n \to \infty$, where c is a positive absolute constant. Since $p_n \sim n \log n$ as $n \to \infty$ (see, for example, [8]), we see that $\mathrm{li}^{-1}(n) \sim n \log n$ as $n \to \infty$. Substituting $x = p_n$ in Proposition 10, we get

$$\sum_{k=1}^{n} p_k^r \log^m p_k = \frac{p_n^{r+1} \log^{m-1} p_n}{r+1} \left(\sum_{j=0}^{N} \frac{\chi_{r,m,j}}{\log^j p_n} + O_{r,m,N}\left(\frac{1}{\log^{N+1} p_n}\right)\right) \tag{18}$$

as $n \to \infty$. Using the mean value theorem and (17), we deduce

$$\log^{m-1}(p_n) = y^{m-1} + O(e^{-d\sqrt{\log n}}) \tag{19}$$

as $n \to \infty$, where d is a real number satisfying $0 < d < c$. Combined with (18), this gives

$$\sum_{k=1}^{n} p_k^r \log^m p_k = \frac{p_n^{r+1} y^{m-1}}{r+1} \left(\sum_{j=0}^{N} \frac{\chi_{r,m,j}}{\log^j p_n} + O_{r,m,N}\left(\frac{1}{y^{N+1}}\right)\right) \tag{20}$$

as $n \to \infty$. By the binomial theorem, we have

$$p_n^{r+1} = (\mathrm{li}^{-1}(n) + \delta(n))^{r+1} = \mathrm{li}^{-1}(n)^{r+1} + O_r(n^{r+1} e^{-c\sqrt{\log n}})$$

as $n \to \infty$. Applying this to (20), we see that

$$\sum_{k=1}^{n} p_k^r \log^m p_k = \frac{e^{(r+1)y} y^{m-1}}{r+1} \left(\sum_{j=0}^{N} \frac{\chi_{r,m,j}}{\log^j p_n} + O_{r,m,N}\left(\frac{1}{y^{N+1}}\right)\right)$$

as $n \to \infty$. Let $f(x) = 1/\log^k x$, where k is an integer with $0 \leq k \leq m-1$. Again, by the mean value theorem there exists a real $\xi \in (\min\{p_n, \mathrm{li}^{-1}(n)\}, \max\{p_n, \mathrm{li}^{-1}(n)\})$ such that $f(p_n) = f(\mathrm{li}^{-1}(n)) + \delta(n) f'(\xi)$. Since $f'(x) = O(1/(x \log^{k+1} x))$ as $x \to \infty$, we get $f(p_n) = f(\mathrm{li}^{-1}(n)) + O(e^{-c\sqrt{\log n}})$ as $n \to \infty$. Hence

$$\sum_{k=1}^{n} p_k^r \log^m p_k = \frac{e^{(r+1)y} y^{m-1}}{r+1} \left(G(1/y) + O_{r,m,N}\left(\frac{1}{y^{N+1}}\right)\right) \tag{21}$$

as $n \to \infty$, where

$$G(u) = G_{r,m,N}(u) = \sum_{j=0}^{N} \chi_{r,m,j}\, u^j.$$

Since (16) holds, we can apply Proposition 12 with $\alpha = 1, \beta = r+1$, and $\gamma = m-1$ to see that

$$\frac{e^{(r+1)y} y^{m-1} G(1/y)}{r+1} = \frac{n^{r+1} \log^{r+m} n}{r+1} \left(\sum_{i=0}^{N} \frac{A_{r,m,i}(\log_2 n)}{\log^i n} + O_{r,m,N}\left(\frac{(\log_2 n)^{N+1}}{\log^{N+1} n}\right)\right) \tag{22}$$

as $n \to \infty$, where $\log_2 x = \log\log x$ and the polynomials $A_{r,m,i} \in \mathbb{R}[x]$ are defined by (8). If we combine (21) and (22), we arrive at the end of the proof of (7). Again, we apply the symbolic algebra system `Maple` to compute the polynomials $A_{r,m,1}, \ldots, A_{r,m,N}$ from the appendices of [8]. It suffices to write, in `Maple`,

```
'sum'(p^r*log(p)^m) = 1/(r+1)*theorem2_part2(1,r+1,m-1,D_N,G_{r,m,N},n,N);
```

with $N = 2$ to get the polynomials $A_{r,m,1}$ and $A_{r,m,2}$. This completes the proof. □

Remark 13. We define $\mathrm{lc}(P)$ to be the leading coefficient of a polynomial $P \in \mathbb{R}[x]$. If $r + m \in \mathbb{N}$, we can use (8) to see that the polynomials $A_{r,m,0}, \ldots, A_{r,m,N} \in \mathbb{R}[x]$ satisfy

$$\deg(A_{r,m,i}) = \begin{cases} i, & \text{if } i \leq r+m; \\ i-1, & \text{otherwise,} \end{cases}$$

and

$$\mathrm{lc}(A_{r,m,i}) = \begin{cases} \binom{r+m}{i}, & \text{if } i \leq r+m; \\ (-1)^{i-1-r-m}\binom{i-1}{r+m}^{-1}, & \text{otherwise.} \end{cases}$$

In the case where $r+m \in \mathbb{R} \setminus (\mathbb{N} \cup \{0\})$, we deduce from (8) that

$$\deg(A_{r,m,i}) = i \quad \text{and} \quad \mathrm{lc}(A_{r,m,i}) = \binom{r+m}{i}.$$

If $r+m=0$, the equation in (8) implies that $A'_{r,m,1} = 0$ and we are not able to say anything in general about the degree or the leading coefficient of the polynomials $A_{r,m,i}$, where $i \geq 1$.

Example 14. Let $\log_2 x = \log\log x$. We write

```
'sum'(p^r*log(p)^m) = 1/(r+1)*theorem2_part2(1,r+1,m-1,D_N,G_{r,m,N},n,N);
```

with $(r,m,N) = (1,-1,2)$, $(r,m,N) = (1,-2,2)$, and $(r,m,N) = (1,-3,2)$ respectively. As $n \to \infty$, this gives the asymptotic formulae

$$\sum_{k=1}^{n} \frac{p_k}{\log p_k} = \frac{n^2}{2}\left(1 - \frac{1}{\log n} + \frac{\log_2 n - 3/2}{\log^2 n} + O\left(\frac{(\log_2 n)^2}{\log^3 n}\right)\right),$$
$$\sum_{k=1}^{n} \frac{p_k}{\log^2 p_k} = \frac{n^2}{2\log n}\left(1 - \frac{\log_2 n + 1/2}{\log n} + \frac{(\log_2 n)^2}{\log^2 n} + O\left(\frac{(\log_2 n)^3}{\log^3 n}\right)\right),$$
$$\sum_{k=1}^{n} \frac{p_k}{\log^3 p_k} = \frac{n^2}{2\log^2 n}\left(1 - \frac{2\log_2 n}{\log n} + \frac{3(\log_2 n)^2 - 2\log_2 n + 2}{\log^2 n} + O\left(\frac{(\log_2 n)^3}{\log^3 n}\right)\right),$$

respectively.

Remark 15. We have $A_{1,1,0} = 1$. The last example implies that $\mathrm{lc}(A_{1,1,1}) = 1$. Now we can use (8) and induction to get $\mathrm{lc}(A_{1,1,i}) = 1$ for all integers i with $i \geq 0$.

Chebyshev's ϑ-function is defined by $\vartheta(x) = \sum_{p \leq x} \log p$, where p runs over primes not exceeding x. Notice that the prime number theorem (1) is equivalent to

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

as $x \to \infty$, where c_1 is a positive absolute constant. Applying a well-known asymptotic expansion for the nth prime number (see [8]), we see that

$$\vartheta(p_n) = n\left(\log n + \log_2 n - 1 + \frac{\log_2 n - 2}{\log n} - \frac{(\log_2 n)^2 - 6\log_2 n + 11}{2\log^2 n} + O\left(\frac{(\log_2 n)^3}{\log^3 n}\right)\right)$$

as $n \to \infty$, where $\log_2 x = \log\log x$, which gives the asymptotic expansion for $S_{r,m}(p_n)$ in the case $r=0$ and $m=1$. For some other cases, we apply again the method developed by Salvy [8, Theorem 2] to get the following further results.

Example 16. Let $\log_2 x = \log\log x$. Again, it suffices to write

```
'sum'(p^r*log(p)^m) = 1/(r+1)*theorem2_part2(1,r+1,m-1,D_N,G_{N,r,m},n,N);
```

with $(r, m, N) = (0, 2, 2)$, $(r, m, N) = (1, 1, 2)$, and $(r, m, N) = (1, 2, 1)$, respectively. As $n \to \infty$, this gives the asymptotic formulae

$$\sum_{k=1}^{n} \log^2 p_k = n \log^2 n \left(1 + \frac{2 \log_2 n - 2}{\log n} + \frac{(\log_2 n)^2 - 2}{\log^2 n} + O\left(\frac{(\log_2 n)^2}{\log^3 n} \right) \right),$$

$$\sum_{k=1}^{n} p_k \log p_k = \frac{n^2 \log^2 n}{2} \left(1 + \frac{2 \log_2 n - 2}{\log n} + \frac{(\log_2 n)^2 - 3}{\log^2 n} + +O\left(\frac{(\log_2 n)^2}{\log^3 n} \right) \right),$$

$$\sum_{k=1}^{n} p_k \log^2 p_k = \frac{n^2 \log^3 n}{2} \left(1 + \frac{3 \log_2 n - 5/2}{\log n} + O\left(\frac{(\log_2 n)^2}{\log^2 n} \right) \right),$$

respectively.

3 Proof of Theorem 5

In 1857, de Polignac [6, part 3] stated without proof that $\log x$ is the right asymptotic behaviour for $\sum_{p \leq x} \log p/p$ as $x \to \infty$, where p runs over primes not exceeding x. A rigorous proof of this statement was given by Mertens [5]. He showed that

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

as $x \to \infty$. We find the following generalization of this asymptotic formula.

Proposition 17. *Let m be a positive real number. As $x \to \infty$, we have*

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

Proof. Similar to Proposition 10. □

Finally, we use Proposition 17 to find the following proof of Theorem 5.

Proof of Theorem 5. We apply Proposition 17 with $x = p_n$ and use (19) to see that

$$\sum_{k=1}^{n} \frac{\log^m p_k}{p_k} = \frac{y^m}{m} + O(1)$$

as $n \to \infty$, where $y = \log \mathrm{li}^{-1}(n)$. Since (16) holds, we can apply Proposition 12 with $\alpha = 1, \beta = 0$, and $\gamma = m$ to get the asymptotic expansion (9). In order to compute the polynomials $B_{m,1}, \ldots, B_{m,N}$, we use the `Maple` code given in the appendices of [8]. It suffices to write

```
'sum'(log(p)^m/p) = 1/(r+1)*theorem2_part2(1,0,m,D_N,1,n,N);.
```

In particular, we get the desired polynomials $B_{m,1}$ and $B_{m,2}$. □

4 Acknowledgement

I would like to express my sincere thanks to Jean-Louis Nicolas for his support. Without his helpful comments this paper would not have appeared in this form. I would also like to thank R. for being a never-ending inspiration. Further, I would like to thank the anonymous referees for useful suggestions to improve the quality of this paper.

References

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

[2] C. Axler, On a sequence involving prime numbers, *J. Integer Sequences* **18** (2015), Article 15.7.6.

[3] C. Axler, On the sum of the first n prime numbers, *J. Théor. Nombres de Bordeaux*, to appear.

[4] J.-P. Massias and G. Robin, Bornes effectives pour certaines fonctions concernant les nombres premiers, *J. Théor. Nombres de Bordeaux* **8** (1996), 213–238.

[5] F. Mertens, Ein Beitrag zur analytischen Zahlentheorie, *J. Reine Angew. Math.* **78** (1874), 42–62.

[6] A. de Polignac, Recherches sur les nombres premiers, *Comptes Rendus Acad. Sci. Paris* **45**, 406–410, 431–434, 575–580, 882–886.

[7] G. Robin, Permanence de relations de récurrence dans certains développements asymptotiques, *Publ. Inst. Math.* **43** (57) (1988), 17–25.

[8] B. Salvy, Fast computation of some asymptotic functional inverses, *J. Symbolic Comput.* **17** (1994), 227–236.

[9] M. Szalay, On the maximal order in S_n and S_n^*, *Acta Arith.* **37** (1980), 321–331.

[10] C.-J. de la Vallée Poussin, Sur la fonction $\zeta(s)$ de Riemann et le nombre des nombres premiers inférieurs à une limite donnée, *Mem. Couronnés de l'Acad. Roy. Sci. Bruxelles* **59** (1899), 1–74.

2010 *Mathematics Subject Classification*: Primary 11N37; Secondary 11A41.
Keywords: asymptotic formula, logarithmic integral, sum of primes.

(Concerned with sequences A000040 and A007504.)

Received December 17 2018; revised versions received December 19 2018; March 29 2019; July 8 2019. Published in *Journal of Integer Sequences*, August 23 2019.

Return to Journal of Integer Sequences home page.