A Variation of Mertens’ Theorem for Almost Primes

Paul Kinlaw (Husson University)
Joint work with Jonathan Bayless (Husson University)

Maine/Quebec Number Theory Conference

October 5, 2019
Mertens’ Second Theorem

Mertens’ Second Theorem (1874) gives us an estimate for the partial harmonic sums of primes:

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

where $B = 0.2614972128\ldots$ is the Mertens constant.
Mertens’ Second Theorem

- Mertens’ Second Theorem (1874) gives us an estimate for the partial harmonic sums of primes:

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

where \( B = 0.2614972128 \ldots \) is the Mertens constant.

- D. Popa (2014):

\[
\sum_{pq \leq x} \frac{1}{pq} = (\log \log x + B)^2 - \zeta(2) + O \left( \frac{\log \log x}{\log x} \right).
\]
Mertens’ Second Theorem

- Mertens’ Second Theorem (1874) gives us an estimate for the partial harmonic sums of primes:

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

where \( B = 0.2614972128 \ldots \) is the Mertens constant.

- D. Popa (2014):

\[ \sum_{pq \leq x} \frac{1}{pq} = (\log \log x + B)^2 - \zeta(2) + O \left( \frac{\log \log x}{\log x} \right). \]

Here, the sum is over all ordered pairs \((p, q)\) of primes.
Mertens’ Second Theorem

- Mertens’ Second Theorem (1874) gives us an estimate for the partial harmonic sums of primes:

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

where \( B = 0.2614972128 \ldots \) is the Mertens constant.

- D. Popa (2014):

\[
\sum_{pq \leq x} \frac{1}{pq} = (\log \log x + B)^2 - \zeta(2) + O \left( \frac{\log \log x}{\log x} \right).
\]

Here, the sum is over all ordered pairs \((p, q)\) of primes.

- Proof uses the hyperbola method for primes (the sum is over points in the first quadrant of the \(pq\)-plane with prime coords. \(pq \leq x\)).
Generalizations

- Let $L = \log \log x$ for convenience.
Generalizations

- Let $L = \log \log x$ for convenience.
- D. Popa (2014):
  
  $$S_2(x) \overset{\text{def}}{=} \sum_{pq \leq x} \frac{1}{pq} = (L + B)^2 - \zeta(2) + \epsilon,$$

  where $\epsilon \ll (\log \log x) / \log x$.

- D. Popa (2016):
  
  $$S_3(x) \overset{\text{def}}{=} \sum_{pqr \leq x} \frac{1}{pqr} = (L + B)^3 - 3\zeta(2)(L + B) + 2\zeta(3) + \epsilon,$$

  where $\epsilon \ll (\log \log x)^2 / \log x$. 
Tenenbaum’s Theorem

Tenenbaum (2016):

\[ S_k(x) \overset{\text{def}}{=} \sum_{p_1 \cdots p_k \leq x} \frac{1}{p_1 \cdots p_k} = P_k(\log \log x) + \epsilon, \]

where \( P_k \) is a degree \( k \) polynomial and \( \epsilon \ll (\log \log x)^k / \log x \).
Tenenbaum’s Theorem

- Tenenbaum (2016):

\[ S_k(x) \overset{\text{def}}{=} \sum_{p_1 \cdots p_k \leq x} \frac{1}{p_1 \cdots p_k} = P_k(\log \log x) + \epsilon, \]

where \( P_k \) is a degree \( k \) polynomial and \( \epsilon \ll (\log \log x)^k / \log x \).

- \( P_k(X) = \sum_{0 \leq j \leq k} \lambda_{j,k} X^j, \)

\[
\lambda_{j,k} = \sum_{0 \leq m \leq k-j} \binom{k}{m,j,k-m-j} (B - \gamma)^{k-m-j} \left( \frac{1}{\Gamma} \right)^{(m)} (1).
\]

Here \( \gamma = 0.577215 \ldots \) is Euler’s constant.
Tenenbaum’s proof follows the Selberg-Delange method from complex analysis, writing

\[ S_k(x) = \frac{1}{2\pi i} \int_{c+i\mathbb{R}} P(s + 1)^k x^s \frac{ds}{s}, \]

where \( c > 0, \ x \in \mathbb{R}^+ \setminus \mathbb{N}. \) \( P(s) = \sum p^{-s} \) is the prime zeta function of \( s. \)
Almost Primes

Tenenbaum’s proof follows the Selberg-Delange method from complex analysis, writing

$$S_k(x) = \frac{1}{2\pi i} \int_{c+i\mathbb{R}} P(s + 1)^k x^s \frac{ds}{s},$$

where $c > 0$, $x \in \mathbb{R}^+ \setminus \mathbb{N}$. $P(s) = \sum p^{-s}$ is the prime zeta function of $s$.

The sum in Tenenbaum’s theorem is over all ordered $k$-tuples of primes $(p_1, \ldots, p_k)$. Thus some terms are counted multiple times.
Tenenbaum’s proof follows the Selberg-Delange method from complex analysis, writing

\[ S_k(x) = \frac{1}{2\pi i} \int_{c+i\mathbb{R}} P(s + 1)^k x^s \frac{ds}{s}, \]

\( c > 0, \ x \in \mathbb{R}^+ \setminus \mathbb{N}. \) \( P(s) = \sum p p^{-s} \) is the prime zeta function of \( s. \)

The sum in Tenenbaum’s theorem is over all ordered \( k \)-tuples of primes \( (p_1, \ldots, p_k). \) Thus some terms are counted multiple times.

\( 1/30 = 1/(2 \cdot 3 \cdot 5) \) is counted 6 times, \( 1/12 = 1/(2^2 \cdot 3) \) is counted three times, and \( 1/8 = 1/2^3 \) is counted only once.
Almost Primes

- Big Omega Function: \( \Omega(p_1^{a_1} \ldots p_m^{a_m}) = a_1 + \ldots + a_m \).
Almost Primes

▶ Big Omega Function: \( \Omega(p_1^{a_1} \ldots p_m^{a_m}) = a_1 + \ldots + a_m \).
▶ A \textit{k-almost prime} is a number \( n \) such that \( \Omega(n) = k \).
Almost Primes

- Big Omega Function: $\Omega(p_1^{a_1} \ldots p_m^{a_m}) = a_1 + \ldots + a_m$.
- A $k$-almost prime is a number $n$ such that $\Omega(n) = k$.
- Let $\mathbb{N}_k = \{ n \in \mathbb{N} : \Omega(n) = k \}$, and let $\tau_k(x) = |\{ n \in \mathbb{N}_k : n \leq x \}|$. 
Almost Primes

- Big Omega Function: \( \Omega(p_1^{a_1} \cdots p_m^{a_m}) = a_1 + \cdots + a_m \).
- A \( k \)-almost prime is a number \( n \) such that \( \Omega(n) = k \).
- Let \( \mathbb{N}_k = \{ n \in \mathbb{N} : \Omega(n) = k \} \), and let \( \tau_k(x) = |\{ n \in \mathbb{N}_k : n \leq x \}|. \)
- Landau (1900): Let \( k \in \mathbb{N} \).

\[
\tau_k(x) = \frac{x}{\log x} \frac{(\log \log x)^{k-1}}{(k - 1)!} \left( 1 + O \left( \frac{1}{\log \log x} \right) \right).
\]
Dusart and others have established explicit bounds for 
\( \tau_1(x) = \pi(x) \).
Dusart and others have established explicit bounds for
\( \tau_1(x) = \pi(x) \).

Bayless, K, Klyve (2019): The squarefree analogue \( \pi_k(x) \) of \( \tau_k(x) \) satisfies
\[
\pi_k(x) < \frac{1.028 x \left( \log \log x + 0.26153 \right)^{k-1}}{\log x (k - 1)!} \quad (k \geq 2, x \geq 3).
\]
Dusart and others have established explicit bounds for $\tau_1(x) = \pi(x)$.

Bayless, K, Klyve (2019): The squarefree analogue $\pi_k(x)$ of $\tau_k(x)$ satisfies

$$\pi_k(x) < \frac{1.028x (\log \log x + 0.26153)^{k-1}}{\log x (k - 1)!} \quad (k \geq 2, x \geq 3).$$

K (2019): We have

$$\tau_3(x) > \frac{x (\log \log x)^2}{\log x^2} \quad (x \geq 500194).$$

This and similar results improve on some of my previous Maine/Quebec talks.
A Combinatorial Formula

\[ R_k(x) \overset{\text{def}}{=} \sum_{n \in \mathbb{N}_k, n \leq x} \frac{1}{n} \sim \frac{(\log \log x)^k}{k!}. \]
A Combinatorial Formula

Let \( R_k(x) \overset{\text{def}}{=} \sum_{n \in \mathbb{N}_k} \frac{1}{n} \sim \frac{(\log \log x)^k}{k!} \).

This is like Tenenbaum’s sum, but it counts each term only once.

Mathar and Lichtman computed \( P_k(s) \) and \( \sum_{n \in \mathbb{N}_k} \frac{1}{n \log n} \) to high precision, \( k \leq 20 \), extending work of H. Cohen.
A Combinatorial Formula

- Let \( R_k(x) \equiv \sum_{\substack{n \in \mathbb{N}_k \atop n \leq x}} \frac{1}{n} \sim \frac{(\log \log x)^k}{k!} \).

- This is like Tenenbaum’s sum, but it counts each term only once.

- We build on his estimate via a combinatorial formula for the almost prime zeta function \( P_k(s) = \sum_{n \in \mathbb{N}_k} n^{-s} \) (see preprints of R. J. Mathar, 2009, and J. Lichtman, 2019).
A Combinatorial Formula

Let \( R_k(x) \) define
\[
\sum_{n \in \mathbb{N}_k, \ n \leq x} \frac{1}{n} \sim \frac{(\log \log x)^k}{k!}.
\]

This is like Tenenbaum’s sum, but it counts each term only once.

We build on his estimate via a combinatorial formula for the almost prime zeta function \( P_k(s) = \sum_{n \in \mathbb{N}_k} n^{-s} \) (see preprints of R. J. Mathar, 2009, and J. Lichtman, 2019).

\[
P_2(s) = \frac{(P(s)^2 + P(2s))}{2!},
\]
\[
P_3(s) = \frac{(P(s)^3 + 3P(2s)P(s) + 2P(3s))}{3!},
\]
\[
P_4(s) = \frac{(P(s)^4 + 6P(2s)P(s)^2 + 3P(2s)^2 + 8P(3s)P(s) + 6P(4s))}{4!}, \ldots
\]
A Combinatorial Formula

Let $R_k(x) \overset{\text{def}}{=} \sum_{n \in \mathbb{N}_k, n \leq x} \frac{1}{n} \sim \frac{(\log \log x)^k}{k!}$.

This is like Tenenbaum’s sum, but it counts each term only once.

We build on his estimate via a combinatorial formula for the almost prime zeta function $P_k(s) = \sum_{n \in \mathbb{N}_k} n^{-s}$ (see preprints of R. J. Mathar, 2009, and J. Lichtman, 2019).

$P_2(s) = (P(s)^2 + P(2s))/2!$,
$P_3(s) = (P(s)^3 + 3P(2s)P(s) + 2P(3s))/3!$,
$P_4(s) = (P(s)^4 + 6P(2s)P(s)^2 + 3P(2s)^2 + 8P(3s)P(s) + 6P(4s))/4!$, ....

Mathar and Lichtman computed $P_k(s)$ and $\sum_{n \in \mathbb{N}_k} 1/(n \log n)$ to high precision, $k \leq 20$, extending work of H. Cohen.
A Combinatorial Formula for $R_k(x)$

- $R_k(x) \to \infty$, but we have:

$$R_2(x) = \frac{1}{2!} \left( S_2(x) + \sum_{p^2 \leq x} \frac{1}{p^2} \right)$$
A Combinatorial Formula for $R_k(x)$

$R_k(x) \rightarrow \infty$, but we have:

\[
R_2(x) = \frac{1}{2!} \left( S_2(x) + \sum_{p^2 \leq x} \frac{1}{p^2} \right)
\]

\[
R_3(x) = \frac{1}{3!} \left( S_3(x) + 3 \sum_{p^2 q \leq x} \frac{1}{p^2 q} + 2 \sum_{p^3 \leq x} \frac{1}{p^3} \right)
\]
A Combinatorial Formula for $R_k(x)$

► $R_k(x) \to \infty$, but we have:

$$R_2(x) = \frac{1}{2!} \left( S_2(x) + \sum_{p^2 \leq x} \frac{1}{p^2} \right)$$

$$R_3(x) = \frac{1}{3!} \left( S_3(x) + 3 \sum_{p^2q \leq x} \frac{1}{p^2q} + 2 \sum_{p^3 \leq x} \frac{1}{p^3} \right)$$

$$R_4(x) = \frac{1}{4!} \left( S_4(x) + 6 \sum_{p^2qr \leq x} \frac{1}{p^2q^2r} + 3 \sum_{p^2q^2 \leq x} \frac{1}{p^2q^2} + 8 \sum_{p^3q \leq x} \frac{1}{p^3q} + 6 \sum_{p^4 \leq x} \frac{1}{p^4} \right)$$
A Combinatorial Formula for $R_k(x)$

$R_k(x) \rightarrow \infty$, but we have:

$$R_2(x) = \frac{1}{2!} \left( S_2(x) + \sum_{p^2 \leq x} \frac{1}{p^2} \right)$$

$$R_3(x) = \frac{1}{3!} \left( S_3(x) + 3 \sum_{p^2 q \leq x} \frac{1}{p^2 q} + 2 \sum_{p^3 \leq x} \frac{1}{p^3} \right)$$

$$R_4(x) = \frac{1}{4!} \left( S_4(x) + 6 \sum_{p^2 qr \leq x} \frac{1}{p^2 qr} + 3 \sum_{p^2 q^2 \leq x} \frac{1}{p^2 q^2} + 8 \sum_{p^3 q \leq x} \frac{1}{p^3 q} + 6 \sum_{p^4 \leq x} \frac{1}{p^4} \right)$$

Coefficients are multinomial numbers of integer partitions, see work of R. J. Mathar, and J. Lichtman. See also https://oeis.org/A102189.
We can estimate the sums, for instance,

\[
\sum_{p^2qr \leq x} \frac{1}{p^2qr} = \sum_{p \leq \sqrt[4]{x}} \frac{1}{p^2} \sum_{qr \leq \frac{x}{p^2}} \frac{1}{qr} = \sum_{p \leq \sqrt[4]{x}} \frac{1}{p^2} S_2 \left( \frac{x}{p^2} \right)
\]

\[
= \sum_{p \leq \sqrt[4]{x}} \frac{1}{p^2} \left( \left( \log \log \frac{x}{p^2} + B \right)^2 - \zeta(2) + O \left( \frac{1}{\log \frac{x}{p^2}} \right) \right)
\]

Good bounds are known on \( s(t) \) \( \overset{\text{def}}{=} \sum_{p \leq t} p^{-2} \). Pomerance and Nguyen (2019):

\[
0 < P(2) - \sum_{p \leq t} p^{-2} < (t \log t)^{-1}.
\]
By partial summation,

$$
\sum_{p \leq \sqrt{\frac{x}{4}}} \frac{1}{p^2} \left( \log \log \frac{x}{p^2} \right)^2 = s \left( \sqrt{\frac{x}{4}} \right) (\log \log 4)^2 + \int_2^{\sqrt{\frac{x}{4}}} \frac{4s(t) \log \log \frac{x}{t^2}}{t \log \frac{x}{t^2}} dt
$$

$$
= P(2)(\log \log 4)^2 - \epsilon - \left[ P(2) \left( \log \log \frac{x}{t^2} \right)^2 \right]_{t=2}^{\sqrt{\frac{x}{4}}} + O \left( \int_2^{\sqrt{\frac{x}{4}}} \frac{4 \log \log \frac{x}{t^2}}{t^2 \log t \log \frac{x}{t^2}} dt \right)
$$

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

Here we estimate the integral by splitting the interval at $x^{1/3}$. The technique for this term can be generalized to the others.
Recall that $L = \log \log x$, and let $\epsilon$'s denote small error terms.

\[ R_2(x) = \frac{1}{2} (L + B)^2 + \frac{P(2) - \zeta(2)}{2} + \epsilon, \]
Recall that $L = \log \log x$, and let $\epsilon$’s denote small error terms.

\[
R_2(x) = \frac{1}{2} (L + B)^2 + \frac{P(2) - \zeta(2)}{2} + \epsilon,
\]

\[
R_3(x) = \frac{1}{6} (L + B)^3 + \frac{P(2) - \zeta(2)}{2} (L + B) + \frac{P(3) + \zeta(3)}{3} + \epsilon,
\]
Recall that $L = \log \log x$, and let $\epsilon$'s denote small error terms.

\[ R_2(x) = \frac{1}{2} (L + B)^2 + \frac{P(2) - \zeta(2)}{2} + \epsilon, \]

\[ R_3(x) = \frac{1}{6} (L + B)^3 + \frac{P(2) - \zeta(2)}{2} (L + B) + \frac{P(3) + \zeta(3)}{3} + \epsilon, \]

\[ R_4(x) = \frac{1}{24} (L + B)^4 + \frac{P(2) - \zeta(2)}{4} (L + B)^2 + \frac{P(3) + \zeta(3)}{3} (L + B) \]
\[ + \frac{P(4)}{4} + \frac{\zeta(4)}{16} + \frac{P(2)^2}{8} - \frac{P(2)\zeta(2)}{4} + \epsilon, \ldots. \]
Results

We can also expand in powers of $L = \log \log x$ in place of $L + B$, to determine an estimate which is polynomial in $L$ with a constant term, generalizing Mertens’ constant $B$.

- $B_1 = B$
Results

We can also expand in powers of $L = \log \log x$ in place of $L + B$, to determine an estimate which is polynomial in $L$ with a constant term, generalizing Mertens’ constant $B$.

- $B_1 = B$
- $B_2 = \frac{1}{2} B^2 + \frac{P(2) - \zeta(2)}{2}$
We can also expand in powers of $L = \log \log x$ in place of $L + B$, to determine an estimate which is polynomial in $L$ with a constant term, generalizing Mertens’ constant $B$.

\[
\begin{align*}
B_1 &= B \\
B_2 &= \frac{1}{2} B^2 + \frac{P(2) - \zeta(2)}{2} \\
B_3 &= \frac{1}{6} B^3 + \frac{P(2) - \zeta(2)}{2} B + \frac{P(3) + \zeta(3)}{3}
\end{align*}
\]
We can also expand in powers of $L = \log \log x$ in place of $L + B$, to determine an estimate which is polynomial in $L$ with a constant term, generalizing Mertens’ constant $B$.

- $B_1 = B$
- $B_2 = \frac{1}{2} B^2 + \frac{P(2) - \zeta(2)}{2}$
- $B_3 = \frac{1}{6} B^3 + \frac{P(2) - \zeta(2)}{2} B + \frac{P(3) + \zeta(3)}{3}$
- $B_4 = \frac{1}{24} B^4 + \frac{P(2) - \zeta(2)}{4} B^2 + \frac{P(3) + \zeta(3)}{3} B$
  
  \[+ \frac{P(4)}{4} + \frac{\zeta(4)}{16} + \frac{P(2)^2}{8} - \frac{P(2)\zeta(2)}{4}, \ldots \]
Heuristics

Heuristic argument for the general case. We should have

\[ R_k(x) = \sum_{n+2n_2+\ldots+kn_k=k} S_n(x) \frac{k}{n!} \prod_{j=2}^{n} \frac{1}{n_j!} \left( \frac{P(j)}{j} \right)^{n_j} + O \left( \frac{(\log \log x)^k}{\log x} \right), \]

where the sum is over integer partitions of \( k \). This would lead to

\[ R_k(x) = T_k(\log \log x + B) + O \left( \frac{(\log \log x)^k}{\log x} \right), \]

\[ T_k(X) = \sum_{n+2n_2+\ldots+kn_k=k} \sum_{0 \leq i \leq n} \frac{a_{i,n}}{n!} \prod_{j=2}^{n} \frac{1}{n_j!} \left( \frac{P(j)}{j} \right)^{n_j} X^i, \]

\[ a_{i,n} = \sum_{0 \leq m \leq n-i} \binom{n}{m, i, n - m - i} (-\gamma)^{n-m-i} \left( \frac{1}{\Gamma} \right)^{(m)} \] (1).
Heuristics

Assuming the above, the following characterization of $T_k$ provides a practical recursive formula.

**Corollary**

$T_k$ satisfies $T'_k = T_{k-1}$, and

$$T_k(0) = \sum_{n+2n_2+\ldots+kn_k=k} \frac{a_{0,n}}{n!} \prod_{j=2}^{k} \frac{1}{n_j!} \left( \frac{P(j)}{j} \right)^{n_j},$$

where

$$a_{0,n} = \sum_{0 \leq m \leq n} \binom{n}{m} (-\gamma)^{n-m} \left( \frac{1}{\Gamma} \right)^{(m)},$$

(1).
Some Remarks

- We believe we are close to completing details for the estimates of $R_k(x)$ for all $k \in \mathbb{N}$.
- A similar version gives estimates for the squarefree sums $R_k^*(x)$.
- We have explicit bounds for the error term in $R_2(x)$ as well as the squarefree counterpart $R_2^*(x)$ and this can be extended, as well as estimates for similar sums.

G. Robin (1983): The error term in Mertens' second theorem changes sign infinitely often.

This isn't the case for $R_2(x)$. The error is between $0.8/\log x$ and $2.2/\log x$, $x \geq 4$. 

Paul Kinlaw
Mertens Estimate for Almost Primes
Some Remarks

- We believe we are close to completing details for the estimates of $R_k(x)$ for all $k \in \mathbb{N}$.
- A similar version gives estimates for the squarefree sums $R_k^*(x)$.
- We have explicit bounds for the error term in $R_2(x)$ as well as the squarefree counterpart $R_2^*(x)$ and this can be extended, as well as estimates for similar sums.
- G. Robin (1983): The error term in Mertens’ second theorem changes sign infinitely often.
- This isn’t the case for $R_2(x)$. The error is between $0.8/\log x$ and $2.2/\log x$, $x \geq 4$. 
Thank You!

Some Sources:

