Hardy-Ramanujan Journal
Vol.17 (1994) 32-39

ON A SUM OVER PRIMES

J.W. Sander[1]

Abstract. *It will be shown that, for any $\delta > 0$,*

$$\sum_{p\leq n}^{*} \frac{\log p}{p} = \frac{1}{2}\log n + O((\log n)^{\frac{5}{6}+\delta}),$$

where $(*)$ restricts the summation to those primes p, which satisfy $n = kp+r$ for some integers k and $r, p/2 < r < p$. This result is connected with questions concerning prime divisors of binomial coefficients.

1. **Introduction.** In 1975, *Erdős, Graham, Ruzsa* and *Straus* [3] investigated the sum

$$f(n) = \sum_{\substack{p\leq n \\ p \text{ not dividing } \binom{2n}{n}}} \frac{1}{p},$$

where p runs over the primes. They proved that

$$\lim_{x\to\infty} \frac{1}{x}\sum_{n\leq x} f(n) = \sum_{k=2}^{\infty} \frac{\log k}{2^k},$$

but could not decide whether $f(n)$ itself is bounded or not. In connection with this, they conjuectured that

$$\sum_{p\leq n}^{*} \frac{1}{p} = \left(\frac{1}{2} + o(1)\right) \log\log n, \tag{1}$$

[1]Supported by the Tata Institute of Fundamental Research, Bombay

where $(*)$ indicates that the summation is extended over all primes p such that $n = kp + r$ for some integer k and $p/2 < r < p$. Since

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

conjecture (1) roughly says that half of the primes satisfy condition $(*)$. We will prove the following

THEOREM. *For $n > 1$ and any $\delta > 0$,*

$$\sum_{p \leq n}{}^{*} \frac{\log p}{p} = \frac{1}{2} \log n + O((\log n)^{\frac{5}{6}+\delta}), \tag{2}$$

where the constant implied by $O(\ )$ may only depend on δ.

It is well-known by a result of *Mertens* that

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

Hence our theorem shows that, on avarage, indeed half of the primes satisfy $(*)$ as predicted by (1), however with regard to a slightly different weight.

We did not make any effort to obtain the best possible error term in (2). For references to related problems, the reader may consult [2]. Throughout the paper, $c_1, c_2, \cdots$ will denote positive absolute constants.

With pleasure I take the opportunity to thank Professor S. Srinivasan and the editor of this journal Professor K. Ramachandra for their hospitality during my stay at the Tata Institute of Fundamental Research in Bombay.

2. Preliminaries. Two ingredients will be used in the sequel. The first is an exponential sum estimate due to Jutila.

LEMMA 1 ([4]). *For $2 \leq t \leq n$, we have*

$$\sum_{p \leq t} e\left(\frac{n}{p}\right) \ll \left(t^{1-c_1\Lambda(t,n)} + t^{3/2} n^{-1/2}\right) exp(c_2(\log\log t)^2),$$

where

$$\Lambda(t,n) = \left(\frac{\log t}{\log n}\right)^2$$

and $e(x) = exp(2\pi i x)$.

The second tool in our proof is *Vinogradov's* Fourier series method, as it may be found in [5], p. 32, or [1], Lemma 2.1.

LEMMA 2. *Let* $0 < \varepsilon < \frac{1}{8}$. *Then there are periodic functions* $\psi(x)$ *and* $\Psi(x)$ *with period 1 satisfying* $0 \le \psi(x) \le 1, 0 \le \Psi(x) \le 1$ *for all* $x \in \mathbb{R}$ *and*

$$\psi(x) = \begin{cases} 1 & \text{for } \frac{1}{2} + \varepsilon \le x \le 1 - \varepsilon, \\ 0 & \text{for } 0 \le x \le \frac{1}{2}, \end{cases} \tag{3}$$

and

$$\Psi(x) = \begin{cases} 1 & \text{for } \frac{1}{2} \le x \le 1, \\ 0 & \text{for } \varepsilon \le x \le \frac{1}{2} - \varepsilon. \end{cases} \tag{4}$$

Moreover, $\psi(x)$ *and* $\Psi(x)$ *have Fourier expansions in the form*

$$\psi(x) = \frac{1}{2} - \varepsilon + \sum_{0<|m|<\infty} a_m e(mx) \tag{5}$$

and

$$\Psi(x) = \frac{1}{2} + \varepsilon + \sum_{0<|m|<\infty} A_m e(mx), \tag{6}$$

where $a_m, A_m \in \mathbb{C}$ *satisfy for* $m \neq 0$

$$| a_m | \ll \frac{1}{m^2 \varepsilon}, | A_m | \ll \frac{1}{m^2 \varepsilon}. \tag{7}$$

3. Proof of the theorem. Clearly, the condition $(*)$ in (2) is equivalent to the fact that

$$\frac{n}{p} = k + r'$$

for some integer k and $1/2 < r' < 1$. Since

$$\left\{ \frac{n}{p} \right\} = \{r'\},$$

where $\{x\}$ denotes the fractional part of x, we obviously have

$$\sum_{p \le n}{}^{*} \frac{\log p}{p} = \sum_{\substack{p \le n \\ \{\frac{n}{p}\} > \frac{1}{2}}} \frac{\log p}{p}. \tag{8}$$

We define

$$B = B(n) = n(\log n)^{-2}\exp(-4c_2(\log\log n)^2).$$

Applying Lemma 2, we get by (8), (3) and (4)

$$\sum_{p\le B}\frac{\log p}{p}\psi(\frac{n}{p}) \le \sum_{p\le B}{}^{*}\frac{\log p}{p} \le \sum_{p\le B}\frac{\log p}{p}\Psi(\frac{n}{p}). \tag{9}$$

By (5) and (7),

$$\begin{aligned}\psi(x) &= \tfrac{1}{2} - \varepsilon + \sum_{0<|m|<\varepsilon^{-2}} a_m e(mx) + O\left(\sum_{|m|\ge\varepsilon^{-2}} \tfrac{1}{m^2\varepsilon}\right)\\ &= \tfrac{1}{2} + \sum_{0<|m|<\varepsilon^{-2}} a_m e(mx) + O(\varepsilon).\end{aligned} \tag{10}$$

We set

$$\varepsilon = (\log n)^{-\eta}$$

with $\eta > 0$ to be chosen later. It is well-known that

$$\sum_{p\le n}\frac{\log p}{p} = \log n + O(1). \tag{11}$$

Hence, by (10),

$$\sum_{p\le B}\frac{\log p}{p}\psi(\frac{n}{p}) = \frac{1}{2}\log B + \sum_{0<|m|<\varepsilon^{-2}} a_m \sum_{p\le B}\frac{\log p}{p}e(\frac{mn}{p}) + O(\varepsilon\log n). \tag{12}$$

Similarly, (6) and (7) yield

$$\sum_{p\le B}\frac{\log p}{p}\psi(\frac{n}{p}) = \frac{1}{2}\log B + \sum_{0<|m|<\varepsilon^{-2}} A_m \sum_{p\le B}\frac{\log p}{p}e(\frac{mn}{p}) + O(\varepsilon\log n). \tag{13}$$

Let

$$A = A(n) = \exp((\log n)^{\alpha})$$

where $\alpha > 2/3$. Thus, for any $t \ge A$, constants $c_3 > 0$ and $c_4 > 0$, and sufficiently large n,

$$\log n \le (\log t)^{1/\alpha} \le \left(\frac{c_3}{c_4}\right)^{1/2}\left(\frac{\log t}{(\log\log t)^{2/3}}\right)^{3/2}.$$

Therefore,

$$c_4(\log\log t)^2 \le c_3\frac{(\log t)^3}{(\log n)^2} = c_3\Lambda(t,n)\log t,$$

which implies for suitable $c_4 > c_2$

$$\exp(c_2(\log\log t)^2)(\log t)^3 \le \exp(c_4(\log\log t)^2) \le t^{c_3\Lambda(t,n)}.$$

Hence, for $t \ge A$ and sufficiently large n,

$$t^{1-c_3\Lambda(t,n)}\exp(c_2(\log\log t)^2) \le \frac{t}{(\log t)^3}.$$

Since for $0 < m < \varepsilon^{-2}$

$$\Lambda(t,mn) \ge \Lambda(t,\varepsilon^{-2}n) \ge \Lambda(t,n^3) = \frac{1}{9}\Lambda(t,n),$$

we get for $t \ge A, 0 < m < \varepsilon^{-2}$ and sufficiently large n

$$t^{1-c_1\Lambda(t,mn)}\exp(c_2(\log\log t)^2) \le \frac{t}{(\log t)^3}. \tag{14}$$

By partial summation, Lemma 1 and (14)

$$\begin{aligned}\sum_{A<p\le B}\tfrac{\log p}{p}e\left(\tfrac{mn}{p}\right) &= \tfrac{\log B}{B}\sum_{A<p\le B}e\left(\tfrac{mn}{p}\right) + \int_A^B\left(\sum_{A<p\le t}e\left(\tfrac{mn}{p}\right)\right)\tfrac{\log t-1}{t^2}dt\\ &\ll \tfrac{\log B}{B}\left(\tfrac{B}{(\log B)^3} + B^{3/2}(mn)^{-1/2}\exp(c_2(\log\log n)^2)\right)\\ &\quad+\int_A^B\left(\tfrac{t}{(\log t)^3} + t^{3/2}(mn)^{-1/2}\exp(c_2(\log\log n)^2)\right)\tfrac{\log t}{t^2}dt\\ &\ll \tfrac{1}{(\log B)^2} + \exp(-c_2(\log\log n)^2) + \tfrac{1}{\log A}\\ &\ll \tfrac{1}{\log A} \ll (\log n)^{-\alpha}.\end{aligned}$$

Therefore, (11) implies

$$\begin{aligned}\sum_{p\le B}\tfrac{\log p}{p}e\left(\tfrac{mn}{p}\right) &\le \sum_{p\le A}\tfrac{\log p}{p} + \Big|\sum_{A<p\le B}\tfrac{\log p}{p}e\left(\tfrac{mn}{p}\right)\Big|\\ &\ll (\log n)^{\alpha}.\end{aligned}$$

Using this in (12) resp. (13), we get by (7)

$$\sum_{p\le B}\frac{\log p}{p}\psi\left(\frac{n}{p}\right) = \frac{1}{2}\log B + O(\varepsilon^{-1}(\log n)^{\alpha}) + O(\varepsilon\log n)$$

respectively

$$\sum_{p\le B}\frac{\log p}{p}\Psi\left(\frac{n}{p}\right)=\frac{1}{2}\log B+O(\varepsilon^{-1}(\log n)^{\alpha})+O(\varepsilon\log n).$$

Taking $\eta=\frac{1}{2}(1-\alpha)$, we thus have by (9)

$$\sum_{p\le B}{}^{*}\frac{\log p}{p}=\frac{1}{2}\log B+O\left((\log n)^{\frac{1}{2}(1+\alpha)}\right). \tag{15}$$

By (11),

$$\begin{aligned}\sum_{B<p\le n}\frac{\log p}{p} &= \log n-\log B+O(1)\\ &= 4c_3(\log\log n)^2+2\log\log n+O(1)\\ &= O((\log\log n)^2).\end{aligned}$$

Now (15) yields

$$\sum_{p\le n}{}^{*}\frac{\log p}{p}=\sum_{p\le n}{}^{*}\frac{\log p}{p}+O((\log\log n)^2)=\frac{1}{2}\log n+O\left((\log n)^{\frac{1}{2}(1+\alpha)}\right)$$

which proves the theorem since α was an arbitrary constant satisfying $\alpha>2/3$.

4. **Final remarks.** The used method, i.e. the application of *Jutila's* result, is not sufficient to deal with the sum

$$\sum_{p\le n}{}^{*}\frac{1}{p}. \tag{16}$$

The reason for this is that, contrary to (14), we easily obtain

$$t^{1-c_1\Lambda(t,n)}exp(c_2(\log\log t)^2)\gg t$$

for $t\le exp((\log n)^{2/3})=K(n)$, say, which means that, for $t\le K(n)$, Lemma 1 does not imply a better estimate than the trivial one. In other words, we cannot take advantage of the condition (*) in (16) for $p \le K(n)$. However, it is well-known that

$$\sum_{p \le K(n)} \frac{1}{p} = log\ log\ K(n) + O(1) = \frac{2}{3}\ log\ log\ n + O(1),$$

which is too big already to get the conjectured formula (1). Hence, a non-trivial estimate for the exponential sum of Lemma 1 for "small" values of t (compared with n) would be needed.

On the other hand, the ranges of primes which give essential contributions to (1) aresp. (2) are not completely disjoint. In order to see this, one may observe that for $L(n,\varepsilon) = exp((log\ n)^{1-\varepsilon})$, we have

$$\sum_{L(n,\varepsilon) \le p \le n} \frac{1}{p} = o(log\ log\ n), \tag{17}$$

as $\varepsilon \to 0$. Clearly,

$$\sum_{p \le L(n,\varepsilon)} \frac{log\ p}{p} = o(log\ n) \tag{18}$$

for any $\varepsilon > 0$. So, if we only knew the asymptotic behaviour

$$\sum_{p \le n}{}^{*} \frac{log\ p}{p} = \left(\frac{1}{2} + o(1)\right) log\ n, \tag{19}$$

then, by (17) and (18), the validity of (1) resp. (19) could be regarded as being dependent on disjoint sets of primes, which means that (19) contains no information about (1). However, (2) is more precise than (19) and actually does depend on primes $p \le L(n,\varepsilon)$ relevant for the validity of (1). In this sense, our theorem does give some indication for the conjecture.

References

[1] R.C. Baker, Diophantine Inequalities, Clarendon Press, Oxford, 1986.

[2] P. Erdös, R.L. Graham, *Old and New Problems and Results in Combinatorial Number Theory*, L'Enseign. Math., Geneva, 1980.

[3] P. Erdös, R.L. Graham, I.Z. Ruzsa, E.G. Straus, "*On the Prime Factors of* $\binom{2n}{n}$", Math. Comp. 29 (1975), 83-92.

[4] M. Jutila, "*On the numbers with a large prime factor, II*", J. Indian Math. Soc. 38 (1974), 125-130.

[5] I.M. Vinogradov, The Method of Trigonometrical Sums in the Theory of Numbers, Interscience Publishers, London, O.J.

Institut für Mathematik
Universität Hannover
Welfengarten 1
3000 Hannover 1
Fed. Rep. of Germany