On the Estimate for a Mean Value Relative to $\frac{4}{p} = \frac{1}{n_1} + \frac{1}{n_2} + \frac{1}{n_3}$

Chaohua Jia

Abstract. For the positive integer $n$, let $f(n)$ denote the number of positive integer solutions $(n_1, n_2, n_3)$ of the Diophantine equation

$$\frac{4}{n} = \frac{1}{n_1} + \frac{1}{n_2} + \frac{1}{n_3}.$$  

For the prime number $p$, $f(p)$ can be split into $f_1(p) + f_2(p)$, where $f_i(p) (i = 1, 2)$ counts those solutions with exactly $i$ of denominators $n_1, n_2, n_3$ divisible by $p$.

Recently Terence Tao proved that

$$\sum_{p < x} f_1(p) \ll x \exp\left(\frac{c \log x}{\log \log x}\right)$$

with other results. In this paper we shall improve it to

$$\sum_{p < x} f_1(p) \ll x \log^5 x \log \log^2 x.$$  

1. Introduction

For the positive integer $n$, let $f(n)$ denote the number of positive integer solutions $(n_1, n_2, n_3)$ of the Diophantine equation

$$\frac{4}{n} = \frac{1}{n_1} + \frac{1}{n_2} + \frac{1}{n_3}.$$  

Erdös and Straus conjectured that for all $n \geq 2$, $f(n) > 0$. It is still an open problem now although there are some partial results.

In 1970, R. C. Vaughan[5] showed that the number of $n < x$ for which $f(n) = 0$ is at most $x \exp(-c \log^2 x)$, where $x$ is sufficiently large and $c$ is a positive constant.

Recently Terence Tao[4] studied the situation in which $n$ is the prime number $p$. He gave lower bound and upper bound for the mean value of
Precisely, he split \( f(p) \) into \( f_1(p) + f_2(p) \), where \( f_i(p)(i = 1, 2) \) counts those solutions with exactly \( i \) of denominators \( n_1, n_2, n_3 \) divisible by \( p \). He proved that

\[
x \log^2 x \ll \sum_{p<x} f_1(p) \ll x \exp\left(\frac{c \log x}{\log \log x}\right)
\]

and

\[
x \log^2 x \ll \sum_{p<x} f_2(p) \ll x \log^2 x \log \log x,
\]

where \( p \) denotes the prime number, \( x \) is sufficiently large and \( c \) is a positive constant.

For the progress and some explanation on the estimate in (2), one can see [2]. In this paper we shall improve the upper bound in (1).

**Theorem.** Let \( p \) denote the prime number. Then for sufficiently large \( x \), we have

\[
\sum_{p<x} f_1(p) \ll x \log^5 x \log \log^2 x.
\]

Throughout this paper, let \( p \) denote the prime number, \( c \) denote the positive constant, \( p(n) \) be the least prime factor of \( n \), \( P(n) \) be the largest prime factor of \( n \), \( d(n) \) be the divisor function, \( \varphi(n) \) be the Euler totient function, \( \Omega(n) \) be the number of prime factors of \( n \) with multiplicity.

### 2. Some preliminaries

**Lemma 1.** Let

\[
g(x) = a_n x^n + \cdots + a_1 x + a_0
\]

be the polynomial in integer coefficients, \( G(n) \) be the number of solutions to the congruence equation

\[
g(x) \equiv 0 \pmod{n}.
\]

Then \( G(n) \) is a multiplicative function.

One can see page 34 of [1].

**Lemma 2.** Let \( g(x) \) be the polynomial in integer coefficients. If

\[
g(x) \equiv 0, \quad g'(x) \equiv 0 \pmod{p}
\]

then
have no common solution, then the number of solutions to
\[ g(x) \equiv 0 \pmod{p^l} \]
is equal to that to
\[ g(x) \equiv 0 \pmod{p}. \]

One can see page 36 of [1].

**Lemma 3.** For the fixed integer \( l \), let \( G(n) \) be the number of solutions to the congruence equation
\[ 4lx^2 + 1 \equiv 0 \pmod{n}. \]
Then
\[ G(n) \leq d(n). \]

**Proof.** By Lemma 1, we know
\[ G(p_1^{l_1} \cdots p_s^{l_s}) = G(p_1^{l_1}) \cdots G(p_s^{l_s}). \]
Write
\[ g(x) = 4lx^2 + 1. \]
Then \( g'(x) = 8lx \). It is obvious that
\[ g(x) \equiv 0, \quad g'(x) \equiv 0 \pmod{p} \]
has no common solution. Thus Lemma 2 claims
\[ G(p^l) = G(p). \]
It is easy to see that the congruence equation
\[ 4lx^2 + 1 \equiv 0 \pmod{p} \]
has at most two solutions. Therefore
\[ G(p^l) = G(p) \leq 2 \leq d(p^l). \]
The conclusion of Lemma 3 follows.
Lemma 4. For \( x \geq 2 \), we have
\[
\sum_{n \leq x} \frac{d^2(n)}{n} \ll \log^4 x.
\]

**Proof.** Theorem 2 in [3] asserts that
\[
\sum_{n \leq x} d^2(n) \ll x \log^3 x.
\]

Then
\[
\sum_{n \leq x} \frac{d^2(n)}{n} \leq \sum_{i \leq \log_2 x} \sum_{2^i \leq n < 2^{i+1}} \frac{d^2(n)}{n} \\
\ll \sum_{i \leq \log_2 x} i^3 \\
\ll \log^4 x.
\]

Lemma 5. Let
\[
\Psi(x, y) = \sum_{n \leq x} \sum_{P(n) \leq y} 1.
\]

Then for \( x \geq 10 \), we have
\[
\Psi(x, \log x \log \log x) \ll \exp(\frac{3 \log x}{(\log \log x)^2}).
\]

This is Lemma 1 in [3].

**Lemma 6.** The estimate
\[
\sum_{2^k \leq n \leq x, \ P(n) \leq Z^{\frac{1}{3}}} \frac{d^2(n)}{n} \ll \exp(\sum_{p \leq Z} \frac{d^2(p)}{p} - \frac{r}{10} \log r)
\]
holds true for \( 1 \leq r \leq \frac{\log Z}{\log \log Z} \) uniformly.

This is a special case of Lemma 4 in [3].

**3. The proof of Theorem**

According to the discussion in the beginning of section 3 of [4], in order to estimate
\[
\sum_{p < x} f_1(p),
\]
it is enough to estimate
\[
\sum_{a,l} \frac{xd(4a^2 + 1)}{\varphi(4al) \log(1 + \frac{x}{al})}.
\] (3)

Using the bound
\[
\varphi(n) \gg \frac{n}{\log \log n},
\]
we should estimate
\[
x \log \log x \sum_{a,l} \frac{d(4a^2 + 1)}{al \log(1 + \frac{x}{al})}
\]
or
\[
x \log \log x \sum_{i \leq \log_2 x} \sum_{j \leq \log_2 x-i} \frac{1}{1 + \log_2 x-i-j} \cdot \frac{1}{2^{i+j}} \sum_{2^i < a \leq 2^{i+1}} \sum_{2^j < l \leq 2^{j+1}} d(4a^2 + 1).
\]

Now we consider the estimate for the sum
\[
\sum_{V < l \leq 2V} \sum_{W < a \leq 2W} d(4a^2 + 1). \tag{4}
\]

We shall use some ideas from [3].

Firstly assume that \( V \leq W \). Let
\[
Z = W^{\frac{1}{2V}}. \tag{5}
\]

Write \( n \) uniquely as
\[
n = p_1^{s_1} \cdots p_j^{s_j} p_{j+1}^{s_{j+1}} \cdots p_r^{s_r}, \quad p_1 < \cdots < p_j < p_{j+1} < \cdots < p_r,
\]
where
\[
p_1^{s_1} \cdots p_j^{s_j} \leq Z < p_1^{s_1} \cdots p_j^{s_j} p_{j+1}^{s_{j+1}}.
\]

We can decompose \( 4a^2 + 1 \) as
\[
4a^2 + 1 = (p_1^{s_1} \cdots p_j^{s_j})(p_{j+1}^{s_{j+1}} \cdots p_r^{s_r}) = b(l, a)c(l, a), \tag{6}
\]
where
\[
b(l, a) \leq Z, \quad (b(l, a), c(l, a)) = 1.
\]
We shall discuss in four cases as in [3].

Case I. \( p(c(l, a)) > Z^{\frac{3}{2}} \).

Since \( p(c(l, a)) > Z^{\frac{3}{2}} \), \( d(c(l, a)) = O(1) \). Thus

\[
d(4la^2 + 1) = d(b(l, a))d(c(l, a)) \ll d(b(l, a)).
\]

Lemmas 3 and 4 yield that

\[
\sum_l \ll \sum_{b \leq Z} \frac{d(b)}{b} \sum_{V < l \leq 2V} \sum_{W < a \leq 2W} \frac{1}{4la^2 + 1 \equiv 0 \text{ (mod } b)}
\]

\[
\ll \sum_{b \leq Z} \frac{d(b)}{b} \sum_{V < l \leq 2V} \frac{1}{b} \sum_{a = 1}^{b} \frac{1}{4la^2 + 1 \equiv 0 \text{ (mod } b)}
\]

\[
\ll VW \sum_{b \leq Z} \frac{d^2(b)}{b}
\]

\[
\ll VW \log^4(2W).
\]

Case II. \( p(c(l, a)) \leq Z^{\frac{3}{2}}, b(l, a) \leq Z^{\frac{3}{2}} \).

Write \( p = p(c(l, a)) \). Then \( p^s \| 4la^2 + 1, p \leq Z^{\frac{3}{2}} \). The fact that \( b(l, a) \leq Z^{\frac{3}{2}} \), \( b(l, a)p^s > Z \) yields \( p^s > Z^{\frac{3}{2}} \). Let \( s_p \) be the smallest \( s \) such that \( p^s > Z^{\frac{1}{2}} \). Thus \( s_p \geq 2 \). On the other hand, \( p^{2s_p} \leq p^{s_p - 1} \leq Z^{\frac{1}{2}} \implies p^{s_p} \leq Z \). Now we have

\[
\frac{1}{p^{s_p}} \leq \min\left(\frac{1}{Z^{\frac{3}{2}}}, \frac{1}{p^2}\right).
\]

Hence

\[
\sum_{p \leq Z^{\frac{1}{2}}} \frac{1}{p^{s_p}} \leq \sum_{p \leq Z^{\frac{1}{2}}} \frac{1}{Z^{\frac{s_p}{2}}} + \sum_{Z^{\frac{1}{2}} < p} \frac{1}{p^2}
\]

\[
\ll Z^{-\frac{1}{4}}.
\]
Lemmas 3 yields that

\[
\sum_{\|} \ll W^\varepsilon \sum_{p \leq Z^\frac{1}{2}} \sum_{V < l \leq 2V} \sum_{W < a \leq 2W} \frac{1}{4a^2 + 1 \equiv 0 \pmod{p^s}}
\]

\[
\ll W^\varepsilon \sum_{p \leq Z^\frac{1}{2}} \sum_{V < l \leq 2V} \frac{W}{p^s} \sum_{a=1}^{p^s} \frac{1}{4a^2 + 1 \equiv 0 \pmod{p^s}}
\]

\[
\ll VW^{1+\varepsilon} \sum_{p \leq Z^\frac{1}{2}} \frac{1}{p^s}
\]

\[
\ll VW^{1-\frac{1}{m}+\varepsilon} \ll VW,
\]

where \(p^s \leq Z\) works.

Case III. \(p(c(l, a)) \leq \log W \log \log W, b(l, a) > Z^\frac{1}{2}\).

We have \(p(c(l, a)) \leq \log W \log \log W \implies P(b(l, a)) < \log W \log \log W\).

Then Lemmas 3 and 5 yield that

\[
\sum_{\|} \ll W^\varepsilon \sum_{Z^\frac{1}{2} < b \leq Z} \sum_{P(b) < \log W \log \log W} \sum_{V < l \leq 2V} \sum_{W < a \leq 2W} \frac{1}{4a^2 + 1 \equiv 0 \pmod{b}}
\]

\[
\ll W^\varepsilon \sum_{Z^\frac{1}{2} < b \leq Z} \frac{d(b)}{b} VW
\]

\[
\ll VW^{1+2\varepsilon} Z^{-\frac{1}{2}} \sum_{P(b) < \log W \log \log W} 1
\]

\[
\ll VW^{1-\frac{1}{m}+2\varepsilon} \Psi(W, \log W \log \log W)
\]

\[
\ll VW^{1-\frac{1}{m}+3\varepsilon} \ll VW.
\]

Case IV. \(\log W \log \log W < p(c(l, a)) \leq Z^\frac{1}{2}, b(l, a) > Z^\frac{1}{2}\).

Let

\[
r_0 = \left[ \frac{\log Z}{\log(\log W \log \log W)} \right].
\]

Since

\[
\log W \log \log W > Z^{r_0 + 1},
\]

for \(2 \leq r \leq r_0\), we consider these \((l, a)\) which satisfy

\[
Z^{\frac{1}{r_0+1}} < p(c(l, a)) \leq Z^\frac{1}{r}
\]

7
so that

\[ P(b(l, a)) < p(c(l, a)) \leq Z^{1/2}. \]

We have

\[ \Omega(c(l, a)) \leq \frac{3 \log W}{\log p(c(l, a))} \leq \frac{(r + 1) \log W}{\log Z} \leq 60(r + 1) \leq 120r \]

so that

\[ d(c(l, a)) \leq A^r, \]

where \( A \) is a positive constant.

Lemmas 3, 4 and 6 yield that

\[
\sum_{IV} \ll \sum_{2 \leq r \leq r_0} A^r \sum_{\substack{Z^{1/2} < b \leq Z \\ P(b) < Z^{1/2}}} d(b) \sum_{V < l \leq 2V} \sum_{W < a \leq 2W} \frac{1}{\prod_{4a^2 + 1 \equiv 0 (\text{mod} \ b)}}
\]

\[
\ll VW \sum_{2 \leq r \leq r_0} A^r \sum_{\substack{Z^{1/2} < b \leq Z \\ P(b) < Z^{1/2}}} \frac{d^2(b)}{b}
\]

\[
\ll VW \sum_{2 \leq r \leq r_0} A^r \exp\left(\sum_{p \leq Z} \frac{d^2(p)}{p} - \frac{r}{10} \log r\right)
\]

\[
= VW \sum_{2 \leq r \leq r_0} A^r \exp\left(\sum_{p \leq Z} \frac{4}{p} - \frac{r}{10} \log r\right)
\]

\[
\ll VW \sum_{2 \leq r \leq r_0} A^r \exp\left(4 \log \log Z - \frac{r}{10} \log r\right)
\]

\[
\ll VW \log^4 Z \sum_{r=2}^{\infty} A^r \exp\left(-\frac{r}{10} \log r\right)
\]

\[
\ll VW \log^4 (2W),
\]

where the series is convergent.

Then assume that \( W < V \). In this situation, we shall change the role of \( l \) and \( r \) and shall consider the linear congruence equation

\[ 4a^2l + 1 \equiv 0 (\text{mod} \ n) \]

for the fixed \( a \). This situation is simpler than previous one. We can get the similar estimate as above.
Combining all of above, we get
\[ \sum_{V < l \leq 2V} \sum_{W < a \leq 2W} d(4la^2 + 1) \ll VW \log^4(2W). \]

Hence,
\[
x \log \log x \sum_{i \leq \log_2 x} \sum_{j \leq \log_2 x} \frac{1}{1 + \log_2 x - i - j} \cdot \frac{1}{2^i + j} \sum_{2^i < a \leq 2^{i+1}} \sum_{2^j < l \leq 2^{j+1}} d(4la^2 + 1) \\
\ll x \log^4 x \log \log x \sum_{i \leq \log_2 x} \sum_{j \leq \log_2 x - i} \frac{1}{1 + \log_2 x - i - j} \\
\ll x \log^4 x \log \log x \sum_{i \leq \log_2 x} \sum_{1 \leq h \leq \log_2 x - i + 1} \frac{1}{h} \\
\ll x \log^4 x \log \log x \sum_{i \leq \log_2 x} \log(\log_2 x - i + 2) \\
\ll x \log^5 x \log \log^2 x.
\]

So far the proof of Theorem is finished.

References

[1] Lookeng Hua, *An Introduction to the Number Theory*, Science Press, Beijing, 1995, in Chinese.

[2] Chaohua Jia, *A note on Terence Tao’s paper “On the number of solutions to \( \frac{4}{p} = \frac{1}{n_1} + \frac{1}{n_2} + \frac{1}{n_3} \)”*, available at http://arxiv.org/abs/1107.5394

[3] P. Shiu, *A Brun-Titchmarsh Theorem for multiplicative functions*, J. Reine Angew. Math., 313(1980), 161-170.

[4] Terence Tao, *On the number of solutions to \( \frac{4}{p} = \frac{1}{n_1} + \frac{1}{n_2} + \frac{1}{n_3} \)“, available at http://arxiv.org/abs/1107.1010

[5] R. C. Vaughan, *On a problem of Erdös, Straus and Schinzel*, Mathematika, 17(1970), 193-198.
