A REMARK ON A CONJECTURE OF ERDŐS AND STRAUS

YOUSEF LAZAR

ABSTRACT. The aim of this note is to show that given a positive integer \( n \geq 5 \), the positive integral solutions of the diophantine equation \( \frac{4}{n} = \frac{1}{x} + \frac{1}{y} + \frac{1}{z} \) cannot have solution such that \( x \) and \( y \) are coprime with \( xy < \sqrt{z}/2 \). The proof uses the continued fraction expansion of \( 4/n \).

1. The result

Given a positive integer \( n \) we are interested in the following diophantine equation

\[
\frac{4}{n} = \frac{1}{x} + \frac{1}{y} + \frac{1}{z},
\]

where \( x, y, z \) are three positive integers to be found. The question of finding positive integer solutions \( (x, y, z) \) for this equation was raised by P. Erdős and E. Straus (see e.g. [1]). This problem has attracted a lot of attention and despite many efforts it is still widely open. For an account of the main contributions on this subject we refer the reader to [2].

Usual reductions allow us to assume that \( n \) is an odd prime number, thus one can assume that \( n \) is an odd prime number \( p \). We are looking for integral primitive solutions \( (x, y) \in \mathbb{N} \) of the diophantine problem, when \( z \) is a fixed positive number

\[
\left| \frac{4}{p} - \frac{x + y}{xy} \right| = \frac{1}{z}. \tag{2}
\]

Our main result shows that a solution \( (x, y, z) \) of (2) cannot have its largest coordinate, say \( z \), too far away from the two other coordinates \( x \) and \( y \), provided \( \gcd(x, y) = 1 \). This gives a certain piece of information regarding the localization of the lattice points which are solution of the problem. Let us denote by \( f(p) \) the number of triples of solutions to (1) with \( n = p \) or equivalent to (2). Recently Elsholtz and Tao gave in [2] precise bounds for averages of the form \( \sum_{p \leq N} f(p) \), the study of the counting function \( f(p) \) reduces to count solutions triples of two different types

- Type I solutions of (1) where \( p \) divides \( x \) and \( \gcd(p, yz) = 1 \).
- Type II solutions of (1) where \( p \) divides \( y, z \) and \( \gcd(p, x) = 1 \).

We denote by \( f_I(p) \) (resp. \( f_{II}(p) \)) the number of solutions of (1) of type I (resp. of type II). It was stressed in the same work that for any odd prime number one has the relation

\[
f(p) = 3f_I(p) + 3f_{II}(p).
\]

We introduce a new type which we call the type III which are the solutions of (1) where \( xy < \sqrt{z}/2 \) and \( \gcd(x, y) = 1 \). Analogously we denote by \( f_{III}(p) \) the number of solutions of (1) of type III. Our main result is that for any prime number greater than 3, we have \( f_{III}(p) = 0 \), more precisely

Theorem 1.1. Given an arbitrary prime number \( p \geq 5 \), there are no triple of positive integers \( (x, y, z) \) which is solution of (2) in the range \( xy < \sqrt{z}/2 \) and with \( \gcd(x, y) = 1 \).
Proof of the Theorem. Suppose we fix an arbitrarily large integer \( z_0 > 0 \) in the range \( xy < \sqrt{z}/2 \) and let us try to solve the following diophantine equation with \((x, y) \in \mathbb{N}^2\),

\[
\left| \frac{4}{p} - \frac{x + y}{xy} \right| = \frac{1}{z_0}.
\]

In particular the equation in (3) gives rise to the following inequality

\[
\left| \frac{4}{p} - \frac{x + y}{xy} \right| < \frac{1}{2(xy)^2}.
\]

The conclusion of the theorem will follow from the fact that such \( x \) and \( y \) would never exist. Since we assume that \( \gcd(x, y) = 1 \) then we have \( \gcd(x + y, xy) = 1 \). We need the following classical result of the theory of continued fraction which can be found for instance in [3] (Theorem 19).

Lemma 1.2. Let \( m, n \) be two positive integers and suppose \( \gcd(r, s) = 1 \). If

\[
\left| \frac{m}{n} - \frac{r}{s} \right| < \frac{1}{2s^2}.
\]

Then \( \frac{r}{s} \) is one of the convergents of \( \frac{m}{n} \).

By Lemma 1.2 we infer from (4) that the rational number \( \frac{x + y}{xy} \) must be one of the convergents of \( 4/p \). If we write the continued fraction expansion of \( 4/p = [0; a_1, \ldots, a_l] \), we can say that there exists some \( 1 \leq k \leq l \) such that

\[
\frac{x + y}{xy} = c_k \left( \frac{4}{p} \right) = [0; a_1, \ldots, a_k] = \frac{p_k}{q_k}.
\]

Since \( x, y \) play a symmetric role, we can assume that \( x \leq y \). Our fractions are reduced thus we deduce that we might have \( x + y = p_k \) and \( xy = q_k \). The fact that such \( x \) and \( y \) might exist relies on the solvability in \( \mathbb{N} \) of the following quadratic equation

\[
X^2 - p_k X + q_k = 0.
\]

The discriminant \( D_k = p_k^2 - 4q_k \) cannot vanish otherwise \( p_k \) and \( q_k \) will fail to be coprime. If \( D_k = p_k^2 - 4q_k > 0 \), then a couple \((x, y)\) of rational solutions of the equation

\[
\frac{4}{p} = \frac{1}{x} + \frac{1}{y} + \frac{1}{z_0}
\]

is given by

\[
x_k = \frac{p_k - \sqrt{D_k}}{2} \quad \text{and} \quad y_k = \frac{p_k + \sqrt{D_k}}{2}.
\]

Necessarily \( z_0 = \frac{p}{4 - pc_k} \). Indeed,

\[
\frac{4}{p} - \frac{1}{x_k} - \frac{1}{y_k} - \frac{1}{z_0} = \frac{4}{p} - c_k - \frac{4 - pc_k}{p} = 0.
\]

Regarding \( x_k \) and \( y_k \), both they are in the quadratic field \( \mathbb{Q}[\sqrt{D_k}] \), so these are not necessarily integers. In order to obtain integral solutions we are forced to assume that \( D_k \) has a square root which is an odd integer. In others words, \( D_k = a^2 \) where \( a \) is an odd integer. In this case, we obtain that a triple of solutions \((x_k, y_k, z_0)\) which is given by

\[
x_k = \frac{p_k - a}{2} \quad \text{and} \quad y_k = \frac{p_k + a}{2}.
\]
Note that necessarily
\[ z_0 = \left| \frac{1}{4/p - c_k} \right| = \frac{1}{r_k} \]
where \( r_k = |4/p - c_k| \) is the \( k \)-th error term in the continued fraction approximation. It is well known (see e.g. ) that
\[ r_k = \frac{1}{q_k(x_{k+1} + q_{k-1})} \]
where
\[ x_{k+1} = [a_{k+1}; a_{k+2}, \ldots, a_l]. \]
Hence, the only possible triples of solutions of (3) in the range given above with a fixed \( z_0 \) must take the following form
\[ x_k = \frac{p_k - a}{2}, \quad y_k = \frac{p_k + a}{2} \quad \text{and} \quad z_0 = q_k(x_{k+1}q_k + q_{k-1}) \]
where \( \frac{p_k}{q_k} \) is one of the convergents of the continued fraction expansion \( 4/p = [a_1; a_2, \ldots, a_l] \) and provided \( p_k^2 - 4q_k = a^2 \) with \( a \) being an odd integer. We will show that the latter condition can never be fulfilled.

To proceed we take advantage from the fact that the convergents of \( 4/p \) can only assume specific values which are given in the following lemma,

**Lemma 1.3.** For any prime number \( p \geq 5 \) set
\[ a_1 = \begin{cases} \frac{p - 1}{4} & \text{if } p \equiv 1 \pmod{4} \\ \frac{p - 3}{4} & \text{if } p \equiv 3 \pmod{4} \end{cases} \]

We have two cases,

(a) if \( p \equiv 1 \pmod{4} \), then \( \frac{4}{p} = [0; a_1, 4] \) and the convergents are \( \{0, \frac{4}{p-1}, \frac{4}{p}\} \).

(b) If \( p \equiv 3 \pmod{4} \), then \( \frac{4}{p} = [0; a_1, 1, 3] \) and the convergents are \( \{0, \frac{4}{p-3}, \frac{4}{p+1}, \frac{4}{p}\} \).

**Proof.** The continued fraction of a rational number is entirely determined by the Euclidean algorithm between 4 and \( p \).

(a) Suppose \( p \equiv 1 \pmod{4} \), we perform the division algorithm
\[
\begin{align*}
4 &= p(0) + 4 \\
p &= 4(a_1) + 1 \\
4 &= 1(4) + 0.
\end{align*}
\]
Here \( a_1 = \left\lfloor \frac{p}{4} \right\rfloor = \frac{p - 1}{4} \), thus \( \frac{4}{p} = [0; a_1, 4] \). The successive convergents are given by
\[ c_0 = 0, \quad c_1 = \frac{1}{a_1} = \frac{4}{p-1} \quad \text{and} \quad c_2 = \frac{1}{a_1 + \frac{1}{a_1}} = \frac{4}{p}. \]

(b) Suppose \( p \equiv 3 \pmod{4} \), the division algorithm again shows that
\[
\begin{align*}
4 &= p(0) + 4 \\
p &= 4(a_1) + 3 \\
4 &= 3(1) + 1 \\
3 &= 1(3) + 0.
\end{align*}
\]
Therefore $\frac{4}{p} = [0; a_1, 1, 3]$. The corresponding convergents are $c_0 = 0$, $c_1 = \frac{4}{p-3}$, $c_2 = \frac{1}{a_1 + 1} = \frac{1}{\frac{p-3}{4} + 1} = \frac{4}{p + 1}$ and $c_3 = \frac{1}{a_1 + \frac{1}{\frac{3}{4}}} = \frac{1}{a_1 + \frac{3}{4}} = \frac{4}{p}$. This proves the Lemma.

We are ready to conclude. In both case, Lemma 1.3 shows that all the non-trivial convergents of $\frac{4}{p}$ (i.e. other that 0 and $\frac{4}{p}$) are egyptian fractions, in particular $p_1 = 1$ (case (a)) and $p_1 = p_2 = 1$ (case (b)). It follows that the $X^2 - 4q_kX + q_k = 0$ is not solvable in $\mathbb{N}$ since $D_k = 1 - 4q_k < 0$ in all the cases. Hence in the range $xy < \sqrt{z/2}$ there are no solution of (2) with $\gcd(x, y) = 1$. This finishes the proof of Theorem 1.1.

2. Concluding remarks

Let $p$ be a prime number, and let us introduce the subset

$$E_p := \{(x, y, z) \in \mathbb{N}^3 \mid \frac{4}{p} = 1/x + 1/y + 1/z\}.$$ 

Then $p$ is solution of (2) if and only if $E_p \neq \emptyset$. So the validity of the Erdos-Straus conjecture amounts to prove that $E_p \neq \emptyset$ for every prime $p$. The Theorem 1.1 tells us that for any prime $p \geq 5$ the set

$$R_p := E_p \cap \{(x, y, z) \in \mathbb{N}^3 \mid \gcd(x, y) = 1, xy < \sqrt{z/2}\}$$

is empty, in other words $\mu_{III}(p) = 0$. Now suppose we are looking for solutions of (1) with $0 < x, y, z \leq N$, so we have a total number of possibilities equal to $(N - 1)^3$ from which we have to remove the elements of $\{(x, y, z) \in [1, N]^3 \mid \gcd(x, y) = 1, xy < \sqrt{z/2}\}$. Let us denote by $A_N$ this subset and put $a_N = |A_N|$. An estimate of $a_N$ can be obtained. Indeed by slicing we get,

$$a_N = \sum_{1 \leq z \leq N} |\{1 \leq x, y \leq \lfloor\sqrt{z/2}\rfloor \mid \gcd(x, y) = 1, xy < \sqrt{z/2}\}|.$$

The inner sum counts the number of primitive lattice points under the hyperbola of equation $y = \sqrt{z/2} x^{-1}$. Thus we can write

$$a_N = \sum_{1 \leq z \leq N} \sum_{x \leq \lfloor\sqrt{z/2}\rfloor} \sum_{xy \leq \lfloor\sqrt{z/2}\rfloor} \varphi(y).$$

We need the following estimate which can be found in [5] (Thm 3.4) as $X \to \infty$

$$\sum_{n \leq X} \varphi(n) = \frac{6}{\pi^2} X^2 + O(X \ln X).$$

This asymptotics leads us to

$$a_N = \frac{3}{\pi^2} \sum_{1 \leq z \leq N} \sum_{x \leq \lfloor\sqrt{z/2}\rfloor} \frac{x}{x^2} + O\left(\frac{\sqrt{z}}{2} \left(\ln \frac{z^{1/2}}{x}\right)\right).$$

Using crude bounds we get

$$\frac{3}{\pi^2} N (1 + o(1)) \ll a_N \ll \frac{1}{2\sqrt{2\pi^2}} N^{5/2} + O(N^{3/2} \ln N).$$
As a conclusion, the previous discussion shows that the number of lattices points to be discarded from $E_p \cap [1, N]^3$ is at an order of magnitude of at least $N$ and at most $N^{5/2}$ lattices points inside the cube $[1, N]^3$.

Further questions. It would be interesting to find an analog of the main result for $5/p$ instead of $4/p$, which is also a conjecture due to Sierpinski. More generally, one can also try to check what is happening for fractions of the form $a/p$ where $a$ is a integer which does not divide $p$. The continued fraction expansion will depend on the class of $a \mod p$ and in this case Lemma 1.3 might be less trivial than the $4/p$-case.

References

[1] P. Erdős, Az $1/x_1 + 1/x_2 + \ldots + 1/x_n = a/b$ egyenlet egész számú megoldásairól, Mat. Lapok 1 (1950), 192-210h.
[2] C. Elsholtz and T. Tao, Counting the number of solutions of the Erdos-Straus equation on unit fractions, J. Aust. Math. Soc. 94 (2013), 50-105.
[3] A. Ya. Khinchin, Continued Fractions, Dover, 100 (1997).
[4] D. Li, On the equation $4/n = 1/x + 1/y + 1/z$, J. Number Theory, 13 (1981), 485-494.
[5] G. Tenenbaum, Introduction to Analytic and Probabilistic Number Theory, Graduate Texts in Mathematics, 3rd Ed. Amer. Math. Soc. 163 (2015).

E-mail address: ylazar77@gmail.com