A NOTE ON THE ERDŐS-STRAUS CONJECTURE

Kyle Bradford
Department of Mathematics and Statistics, Georgia Southern University,
Statesboro, GA 30458, USA
kbradford@georgiasouthern.edu

Received: , Revised: , Accepted: , Published:

Abstract
In this paper I make a fundamental assertion about the Erdős-Straus conjecture. Suppose that for some prime \( p \) there exists \( x, y, z \in \mathbb{N} \) with \( x \leq y \leq z \) so that

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

If \( p \nmid y \) then necessarily

\[
z = \frac{xy}{\gcd(xy, x+y)}
\]

This means that the Erdős-Straus conjecture can be reduced by one variable. That is to say that it suffices to show for all primes \( p \) there exist \( x, y \in \mathbb{N} \) with \( p \nmid y \) so that

\[
4xy - (x+y)p = \gcd(xy, x+y)
\]

Considering other reductions of the Erdős-Straus conjecture I suggest a method for proof.

1. Preliminaries

The Erdős-Straus conjecture states that given a prime number \( p \) there exist natural numbers \( x, y \) and \( z \) (w.l.o.g \( x \leq y \leq z \)) so that the Erdős-Straus equation is solved,

\[
\frac{4}{p} = \frac{1}{x} + \frac{1}{y} + \frac{1}{z}
\] (1)

To discern a pattern I am going to assume that the conjecture is true. For a given prime number \( p \) and I am going to determine the necessary properties of the

\[1\text{ any footnote here} \]
associated solution values $x, y$ and $z$. To prevent any ambiguity: to complete a proof of the conjecture one will need to prove that a solution exists, which I do not do in this paper.

**Definition.** Under the assumption that a solution to (1) exists for some prime $p$ reserve $x, y$ and $z \in \mathbb{N}$ as general solution values for that specific prime $p$ and insist that $x \leq y \leq z$

The following propositions follow from an article [2].

**Proposition 1.** A prime number $p$ must divide at least one of its solution values $x, y$ or $z$.

**Proposition 2.** For a given prime number $p$ the solution values $x, y$ and $z$ cannot simultaneously be divisible by $p$.

I introduce an elementary lemma that will help to further illuminate the nature of the solution values for a given prime $p$.

**Lemma 1.** Given a prime number $p$ with solution values $x \leq y \leq z$ we must have

$$\left\lfloor \frac{p}{4} \right\rfloor \leq x \leq \left\lfloor \frac{p}{2} \right\rfloor \quad \frac{xp}{4x - p} \leq y \leq \left\lceil \frac{2xp}{4x - p} \right\rceil$$ \hspace{1cm} (2)

or

$$\left\lfloor \frac{p}{2} \right\rfloor \leq x \leq \left\lfloor \frac{3p}{4} \right\rfloor \quad x \leq y \leq \left\lceil \frac{2xp}{4x - p} \right\rceil$$ \hspace{1cm} (3)

It is also quite easy to show that if $x, y$ and $z$ are solution values for a given $p$ in the regions defined by (2) and (3), then $x \leq y \leq z$. While Lemma 1 is trivial, it helps elucidate the necessary properties of our solution values for a given prime $p$. For example, the following lemmata are direct consequences.

**Lemma 2.** For a prime number $p$ we have that the smallest solution value $x$ is relatively prime from $p$.

**Lemma 3.** For a prime number $p$ we have that $\gcd(y, p^2) \neq p^2$ and $\gcd(z, p^2) \neq p^2$.

The following lemma guarantees that the solution value $z$ is divisible by $p$, but it creates the first dichotomy.

**Lemma 4.** For a prime number $p$ we have that if $\gcd(y, p) = p$ then $\gcd(z, p) = p$.

This allows us make essentially the same definitions as in [2] with the slight distinction being that this paper insists on an ordering of the solutions.

**Definition.** Under the assumption that a solution to (1) exists for some prime $p$ define a Type I solution as one so that $\gcd(x, p) = 1$, $\gcd(y, p) = 1$ and $\gcd(z, p) = p$ and define a Type II solution as one so that $\gcd(x, p) = 1$, $\gcd(y, p) = p$ and $\gcd(z, p) = p$. 
To factor $x, y$ and $z$ into their smallest relevant components I make the following definitions.

**Definition.** Under the assumption that a solution to \([7]\) exists for some prime $p$ we will reserve $d = \gcd(x, y, z)$, $a = \gcd(x, y)/d$, $b = \gcd(x, z)/d$ and $c = \gcd(y, z)/d$. We will also reserve $x^\circ, y^\circ$ and $z^\circ$ so that $x = x^\circ abd$, $y = y^\circ acd$ and $z = z^\circ bcd$.

It should be clear from the definitions that $a, b$ and $c$ are pairwise relatively prime. For Type I solutions $p | z^\circ$ and for Type II solutions $p | c$. Refer to Figure 1 for clarity. The final lemma reduces the complexity of our new factorizations of $x, y$ and $z$.

**Lemma 5.** For a prime number $p$ we have for Type I solutions $x^\circ = y^\circ = 1$ and $z^\circ = p$ and for Type II solutions $x^\circ = y^\circ = z^\circ = 1$.

From these preliminary results I have expressions for $x, y$ and $z$ that are factored into fundamental parts. The initial goal of this was to try to find a pattern for the relatively prime parts as a function of $p$, but it lead to a way to reduce the dimensionality of this problem and provide a hope of an easily obtainable solution; one that I have yet to obtain.
2. Results

This section outlines the three main results of this paper and provides motivation to the nature of the method of solution for this problem.

**Theorem 1.** For a prime number $p$ the following are true for a Type I solution:

\[ 4xy - (x + y)p = \gcd(xy, x + y) \]  
\[ 4xz - (x + z)p = p^2 \cdot \gcd(xz, x + z) \]  
\[ 4yz - (y + z)p = p^2 \cdot \gcd(yz, y + z) \]

What I found to be the most relevant in this result is the implication from (4) that for a Type I solution:

\[ z = \frac{xy}{\gcd(xy, x + y)} \]  

Notice that $z$ by its definition in (7) has to be an integer. For a given $p$ if we can find $x, y \in \mathbb{N}$ in the regions defined in (2) or (3) with $p \nmid y$ so that $4xy - (x + y)p = \gcd(xy, x + y)$, then we would necessarily know which integer to pick for $z$ using (7). This, in essence, reduces the dimensionality of the problem by one degree. This is also incredibly important because it reveals that the true nature of this problem depends on the gcd of the product of two numbers and the sum of those same two numbers. That it not immediately understood from the original description and in writing in this paper I hope to inform and motivate mathematicians who have studied problems of this nature.

The next theorem was motivated by the results in [1].

**Theorem 2.** For a prime number $p$ with a Type I solution

\[ x = \left\lfloor \frac{yp}{4y - p} \right\rfloor \]  

I find this result to be even more astounding. I have again reduced the dimensionality of the problem by one degree. For a prime $p$ it suffices to find $y \in \mathbb{N}$ so that $p \nmid y$.

\[ \left\lfloor \frac{p}{2} \right\rfloor \leq y \leq \left\lceil \frac{2 \left\lfloor \frac{p}{4} \right\rfloor p}{4 \left\lfloor \frac{p}{4} \right\rfloor - p} \right\rceil \]  

and

\[ (4y - p) \left\lfloor \frac{yp}{4y - p} \right\rfloor - yp = \gcd \left( y \cdot \left\lfloor \frac{yp}{4y - p} \right\rfloor, y + \left\lfloor \frac{yp}{4y - p} \right\rfloor \right) \]
It’s important to note that this is only true if a Type I solution exists for a given prime \( p \), but computational evidence suggests that every prime \( p > 2 \) has a Type I solution. Furthermore, this can be converted to an asymptotic statement. For \( p \) large enough, it suffices to find a functional expression for \( y \) that depends solely on \( p \) so that (9) and (10) hold. It’s important to note that the functional description for \( y \) would have to lie between linear and quadratic behavior in \( p \). Finding the correct description, though, has proven elusive. One can see that many patterns exist between \( y \) and \( p \) for Type I solutions, as in figure 2. These patterns are found as modular identities outlined in previous papers [3], but I hold out hope that a general pattern can be found.

![Figure 2: Type I solutions. Red points denote that the prime has remainder 1 after dividing by 4 and blue points denote that the prime has remainder 3 after dividing by 4.](image)

The final theorem addresses Type II solutions, but I haven’t found it as useful because I can’t find a uniform description for \( x \) in terms of \( y \) and \( p \) as I found in Theorem 2. This does not suggest that it is not entirely useful to somebody
that wants to consider Type II solutions. We again see that the nature of the solutions depend on the gcd of the sum of two numbers and the product of those same numbers.

**Theorem 3.** For a prime number $p$ the following are true for a Type II solution:

\[
\begin{align*}
4xy - (x+y)p &= p \cdot \gcd(xy, x+y) \\
4xz - (x+z)p &= p \cdot \gcd(xz, x+z) \\
4yz - (y+z)p &= p^2 \cdot \gcd(yz, y+z)
\end{align*}
\]

**Proofs**

**Proof. Lemma 1:**

First consider the scenario where $x \leq p/4$. This would imply that $4xy - (x+y)p \leq -xp < 0$. Because $xy > 0$ for all $x, y \in \mathbb{N}$ and by definition $z = xyp/(4xy - (x+y)p)$, we see that $z < 0$. But we know that $z > 0$ to be a solution value, so we a contradiction. This guarantees that $x > p/4$. Because $x \in \mathbb{N}$ we can say that $x \geq \lceil p/4 \rceil$.

Next consider the scenario where $x > 3p/4$. Because $x \leq y \leq z$, this would imply that $y > 3p/4$ and $z > 3p/4$. This would make $4/p > 1/x + 1/y + 1/z$. This contradicts our assumption that $x, y$ and $z$ are solution values. This guarantees that $x \leq 3p/4$. Because $x \in \mathbb{N}$ we can say that $x \leq \lfloor 3p/4 \rfloor$.

We now consider the scenario where $y > 2xp/(4x-p)$. This implies that $4xy - yp > 2xp$, or $4xy - (x+y)p > xp$. Because $y > 0$, $xp > 0$ and $xyp > 0$ we see that $4xy - (x+y)p > 0$ and $y > xyp/(4xy - (x+y)p) = z$ which is a contradiction. This guarantees that $y \leq 2xp/(4x-p)$. Because $y \in \mathbb{N}$ we can say that $y \leq \lfloor 2xp/(4x-p) \rfloor$. Because $x \leq y$ by definition, we see that one possibility for our solution values is to have $\lfloor p/2 \rfloor \leq x \leq \lfloor 3p/4 \rfloor$ and $x \leq y \leq \lfloor 2xp/(4x-p) \rfloor$.

We finally consider $\lfloor p/4 \rfloor \leq x \leq \lfloor p/2 \rfloor$ and $y < xp/(4x-p)$. Because $4x-p > 0$ we see that $4xy - (x+y)p < 0$. Because $xyp > 0$ for all $x, y \in \mathbb{N}$ we see that $z < 0$, which is a contradiction. This guarantees that if $\lfloor p/4 \rfloor \leq x \leq \lfloor p/2 \rfloor$, then $y \geq xp/(4x-p)$. Because $y \in \mathbb{N}$ we can say that $y \geq \lfloor xp/(4x-p) \rfloor$. This now shows that our other possibility is that $\lfloor p/4 \rfloor \leq x \leq \lfloor p/2 \rfloor$, and $\lfloor xp/(4x-p) \rfloor \leq y \leq \lfloor 2xp/(4x-p) \rfloor$.

**Proof. Lemma 2:**

Lemma 1 tells us that either $\lfloor p/4 \rfloor \leq x \leq \lfloor p/2 \rfloor$ or $\lfloor p/2 \rfloor \leq x \leq \lfloor 3p/4 \rfloor$. 

It is clear to see that $x < p$. Because $p$ is prime, then by definition we have that $x$ is relatively prime from $p$.

Proof. Lemma 3:

Lemma 4 tells us that $y \leq \frac{2xp}{(4x - p)}$. The largest possible value for $\frac{2xp}{(4x - p)}$ letting $x$ be an integer is when $x = \lceil \frac{p}{4} \rceil$. We see that $\frac{2xp}{(4x - p)}$ will be even larger when $x = \frac{(p + 1)}{4}$. This would tell us that $y \leq \frac{p(p + 1)}{2}$. We can see that $y < p^2$. This will imply that $\gcd(y, p^2) \neq p^2$.

Next we assume that $\gcd(z, p^2) = p^2$. Lemma 2 tells us that $\gcd(x, p) = 1$. Let $z^* \in \mathbb{N}$ so that $z = z^*p^2$. We can write

$$y = \frac{xz^*p^2}{4xz^*p - x - z^*p^2}$$

where $4xz^*p - x - z^*p^2$ has no factor of $p$. This tells us that $p^2$ must divide $y$, which tells us that $\gcd(y, p^2) = p^2$. This is a contradiction, so we have that $\gcd(z, p^2) \neq p^2$.

Proof. Lemma 4:

Let $\gcd(y, p) = p$ and for sake of contradiction assume that $\gcd(z, p) = 1$. We already know from Lemma 2 that $\gcd(x, p) = 1$. Let $y^* \in \mathbb{N}$ so that $y = y^*p$. Lemma 3 tells us that $\gcd(y^*, p) = 1$. We can now write

$$z = \frac{xy^*p}{(4y^* - 1)x - y^*p}$$

Our assumption that $\gcd(z, p) = 1$ requires that $p|\left((4y^* - 1)\right)$. If you recall the proof of Lemma 3 we showed that $y \leq \frac{p(p + 1)}{2}$. This would tell us that $y^* \leq \frac{(p + 1)}{2}$ or $4y^* - 1 \leq 2p + 1$. If $p|(4y^* - 1)$ then either $4y^* - 1 = p$ or $4y^* - 1 = 2p$. It should be clear that $4y^* - 1$ cannot be even, so $4y^* - 1 \neq 2p$. If $4y^* - 1 = p$ then it should be clear that $p \neq 2$. If I use $y^* = \frac{(p + 1)}{4}$, we see that

$$z = \frac{(p + 1)x}{4x - (p + 1)}$$

which is maximized if we select $x = \lceil \frac{p}{4} \rceil$. If $x = \frac{(p + 1)}{4}$, then we have an undefined $z$ and we see that $x, y$ and $z$ are not solution values. Because $p \neq 2$ we cannot have $x = \frac{(p + 2)}{4}$. This implies that $z$ is maximized if $x = \frac{(p + 3)}{4}$. We see that

$$z \leq \frac{p + 1}{4} \cdot \frac{p + 3}{2}$$


with a strict inequality if \( p = 3 \). Because \((p + 3)/2 < p\) for all primes \( p > 3\) we have that \( z < y \) for \( p > 3\). This is a contradiction, so it implies that \( \gcd(z, p) \neq 1 \). Because \( p \) is prime we see that \( \gcd(z, p) = p \).

**Proof. Lemma 5:**

We can rewrite equation (1) with our new notation and perform some algebra to express the equation as follows:

\[
4x^2y^2zabcd = (x^2 y^2 a + x^2 z^2 b + y^2 z^2 c)p
\]

(14)

Without loss of generality, suppose that a prime \( q \neq p \) divides one of \( x^o, y^o \) and \( z^o \), for example \( q | x^o \). (14) would imply that \( q | y^o z^o c \). But definitionally \( q \nmid y^o \) because \( \gcd(x^o, y^o) = 1 \) and \( q \nmid c \) because \( \gcd(x^o, c) = 1 \). Therefore \( q \) cannot divide \( x^o \). The same will be true that primes \( q \neq p \) cannot divide \( y^o \) and \( z^o \).

For a Type I solution the prime \( p \) cannot divide \( x \) or \( y \). This will imply that \( x^o = y^o = 1 \). We see that \( p \) divides \( z \) and \( p \) does not divide \( \gcd(x, z) \) and \( \gcd(y, z) \). This would imply that \( z^o = p \).

For a Type II solution the prime \( p \) cannot divide \( x \). This will imply that \( x^o = 1 \). We see that \( p \) divides both \( y \) and \( z \), so \( p \) divides \( \gcd(y, z) \). Because \( \gcd(y, p^2) = p \) and \( \gcd(z, p^2) = p \) we see that \( p \) cannot divide \( y^o \) and \( z^o \). This means that \( y^o = z^o = 1 \).

**Proof. Theorem 1:**

Using Definition 1 and Lemma 5 we have for Type I solutions that \( x = abd, y = acd, z = bcdp \) and \( p = (4abcd - a)/(b + c) \). We see that

\[
4xy - (x + y)p = 4a^2bcd^2 - ad(4abcd - a)
\]

\[
= a^2 d
\]

Because \( p(b + c) = a(4bcd - 1) \) and \( \gcd(a, p) = 1 \) we see that \( a | (b + c) \). Suppose a prime \( q | ((b + c)/a) \). We have then that \( q | (4bcd - 1) \). If \( q | bcd \), then \( q | 1 \). This implies that \( \gcd(bcd, (b + c)/a) = 1 \). We have then that
\[\text{gcd}(xy, x + y) = \text{gcd}(a^2bcd^2, abd + acd)\]
\[= a^2d \cdot \text{gcd} \left( bcd, \frac{b + c}{a} \right)\]
\[= a^2d\]

This shows that \(4xy - (x + y)p = \text{gcd}(xy, x + y)\). We also see that

\[4xz - (x + z)p = p(4ab^2cd^2 - abd - bcdp)\]
\[= p(bd(b + c)p - bcdp)\]
\[= p^2b^2d\]

Because \(\text{gcd}(acd, p) = 1\) we have that \(\text{gcd}(acd, 4acd - p) = 1\). We have then that

\[p^2 \cdot \text{gcd}(xz, x + z) = p^2bd \cdot \text{gcd}(abcdp, a + cp)\]
\[= p^2b^2d \cdot \text{gcd}(acd, 4acd - p)\]
\[= p^2b^2d\]

This shows that \(4xz - (x + z)p = p^2 \cdot \text{gcd}(xz, x + z)\). We finally see that

\[4yz - (y + z)p = p\left(4abc^2d^2 - acd - bcdp\right)\]
\[= p(cd(b + c)p - bcdp)\]
\[= p^2c^2d\]

Because \(\text{gcd}(abd, p) = 1\) we have that \(\text{gcd}(abd, 4abd - p) = 1\). We have then that

\[p^2 \cdot \text{gcd}(yz, y + z) = p^2cd \cdot \text{gcd}(abcdp, a + bp)\]
\[= p^2c^2d \cdot \text{gcd}(abd, 4abd - p)\]
\[= p^2c^2d\]

This shows that \(4yz - (y + z)p = p^2 \cdot \text{gcd}(yz, y + z)\). \(\square\)
Proof. Theorem 2:

From Theorem 1 we see that $4xy - (x+y)p = \gcd(xy, x+y)$ for any Type I solutions. Dividing both sides by $4y - p$ we see that

$$\frac{x - \frac{yp}{4y-p}}{4y-p} = \frac{\gcd(xy, x+y)}{4y-p}$$

$$< \frac{x+y}{4y-p}$$

$$< \frac{2y}{2y+(2y-p)}$$

By definition we have that $y > p/2$, so we see that

$$x - \frac{yp}{4y-p} < 1$$

Because

$$\frac{\gcd(xy, x+y)}{4y-p} > 0$$

and $x \in \mathbb{N}$ we have that

$$x - \left\lfloor \frac{yp}{4y-p} \right\rfloor = 0$$

Proof. Theorem 3:

Using Definition 1 and Lemma 5 we have for Type II solutions that $x = abd, y = acdp, z = bcdp$ and $p = 4abd - (a+b)/c$. We see that

$$4xy - (x+y)p = p(4a^2bcd^2 - abd - ad(4abcd - (a+b)))$$

$$= pa^2d$$

Because $pc = a(4bcd - 1) - b$ we can see that if $p|(4bcd - 1)$, then $p|b$. Because $\gcd(p, b) = 1$ we have that $\gcd(p, 4bcd - 1) = 1$. We also see that $\gcd(bcd, 4bcd - 1) = 1$, so it should be clear that $\gcd(bcdp, 4bcd - 1) = 1$. We have then that

$$p \cdot \gcd(xy, x+y) = p \cdot \gcd(a^2bcd^2p, abd + acdp)$$

$$= pa^2d \cdot \gcd(bcdp, 4bcd - 1)$$

$$= pa^2d$$
This shows that $4xy - (x + y)p = p \cdot \text{gcd}(xy, x + y)$. We also see that

$$4xz - (x + z)p = p(4ab^2cd^2 - abd - bcdp)$$
$$= p(bdcp + b^2d - bcdp)$$
$$= pb^2d$$

Because $\text{gcd}(acd, p) = 1$ we have that $\text{gcd}(acdp, 4acd - p) = 1$. We have then that

$$p \cdot \text{gcd}(xz, x + z) = pbd \cdot \text{gcd}(abcdp, a + cp)$$
$$= pb^2d \cdot \text{gcd}(acdp, 4acd - 1)$$
$$= pb^2d$$

This shows that $4xz - (x + z)p = p \cdot \text{gcd}(xz, x + z)$. We finally see that

$$4yz - (y + z)p = p^2 \left(4abc^2d^2 - acd - bcd\right)$$
$$= p^2 \left(pc^2d\right)$$
$$= p^3c^2d$$

Because $\text{gcd}(abd, p) = 1$ we have that $\text{gcd}(abdp, 4abd - p) = 1$. We have then that

$$p^2 \cdot \text{gcd}(yz, y + z) = p^3cd \cdot \text{gcd}(abcdp, a + b)$$
$$= p^3c^2d \cdot \text{gcd}(abcdp, 4abd - p)$$
$$= p^3c^2d$$

This shows that $4yz - (y + z)p = p^2 \cdot \text{gcd}(yz, y + z)$. \qed

References

[1] K. Bradford and E. Ionascu, A Geometric reduction of the Erdős-Straus Conjecture, *Advanced Modeling and Optimization* 1 (2015), vol. 17, 41-54.
[2] C. Elsholtz and T. Tao, Counting the number of solutions to the Erdős-Straus Equation on Unit Fractions, *Journal of the Australian Mathematical Society* **94** (2013), vol. 1, 50-105.

[3] L. G. Mordell, *Diophantine equations*, London-New York, Acad. Press, 1969