

A NOTE ON THE ERDŐS-STRAUS CONJECTURE

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

Received: 6/1/19, Revised: 9/20/20, Accepted: 2/11/21, Published: 2/23/21

Abstract

This paper makes a fundamental assertion about the Erdős-Straus conjecture. Suppose that for a prime p there exists $x, y, z \in \mathbb{N}$ with $x \leq y \leq z$ so that

$$\frac{4}{p} = \frac{1}{x} + \frac{1}{y} + \frac{1}{z}.$$

The main contribution of this paper is that, under this assumption, the Erdős-Straus conjecture can be reduced by one variable. For example, it is necessarily true that

$$z = \frac{xyp}{\gcd(y,p)\gcd(xy,x+y)}$$

Considering other reductions of the Erdős-Straus conjecture, this paper suggests a method for proof.

1. Preliminaries

The ancient Egyptian society revered unit fractions as the fundamental building blocks of all other fractions. In fact, ancient Egyptians ascribed special unit fractions to the eye of the god Horus. Thus the name "Egyptian fraction" was given to sums of unit fractions with positive denominators. The monumental work of Paul Erdős left a variety of unsolved problems related to Egyptian fractions [7]. This has led to a large body of work [1, 2, 4, 10, 13, 22]. One of the more common areas of study is the Erdős-Straus and related conjectures [3, 5, 6, 8, 9, 11, 12, 14, 15, 16, 17, 18, 19, 20, 21, 23, 24, 25]. In its original form it reduces to the following: given a prime number p there exist natural numbers x, y, and z and without a loss of generality one may assume $x \leq y \leq z$ so that the Erdős-Straus equation is solved. This is expressed as

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

The most famous approach is by Rosati [14] and is outlined in a book by Mordell [11], which described modular identities for specific cases of primes, and an assumption that quadratic residues play a role in the solution of the problem. More recently Elsholtz and Tao [5] attempted to find a solution using arithmetic number theory. Both of these texts motivated a different approach in the paper presented here. This theoretical paper is a continuation of the work in [3], which was more computational in nature. The patterns discovered here help simplify the problem by one dimension but do not solve the conjecture.

Assume that the conjecture is true. Discerning a pattern between a given prime number p and the associated solution values x, y, and z will determine the necessary conditions to show that a solution exists. To prevent any ambiguity, note that this paper does not show that a solution exists. Going forward, reserve x, y, and z as the general solution values for a prime p and insist that $x \leq y \leq z$. The following propositions are taken from [5].

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.

The first lemma serves to further illuminate the nature of the solution values for a given prime p.

Lemma 1. Given a prime number p, necessary conditions for the solution values $x \leq y \leq z$ are

$$\left\lceil \frac{p}{4} \right\rceil \le x \le \left\lfloor \frac{3p}{4} \right\rfloor \quad and \quad \left\lceil \frac{xp}{4x-p} \right\rceil \le y \le \left\lfloor \frac{2xp}{4x-p} \right\rfloor.$$
 (2)

The necessary conditions (6) establish the subsequent lemmata.

Lemma 2. A prime number p is coprime to the smallest solution value x.

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

The fourth lemma guarantees that the solution value z is divisible by p.

Lemma 4. For a prime number p, if gcd(y, p) = p, then gcd(z, p) = p.

Assume that a solution to Equation (1) exists for a prime p. For our purposes, going forward, a Type I solution will have gcd(x, p) = 1, gcd(y, p) = 1, and gcd(z, p) = p. A Type II solution will have gcd(x, p) = 1, gcd(y, p) = p, and gcd(z, p) = p.

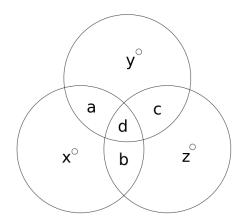


Figure 1: This is a diagram of x, y, and z into its factored parts.

This language is used in [5] to describe types of solutions; however, the current paper insists on an ordering of the solution values. Once classified into types, the solution values must be factored into their smallest relevant components.

Assume that a solution to Equation (1) exists for a prime p. Reserve d = gcd(x, y, z), a = gcd(x, y)/d, b = gcd(x, z)/d, and c = gcd(y, z)/d. Also reserve x°, y° , and z° to be positive integers so that $x = x^{\circ}abd$, $y = y^{\circ}acd$, and $z = z^{\circ}bcd$.

This makes clear that a, b, and c are pairwise coprime. 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 the factorizations of x, y, and z.

Lemma 5. For a prime number p, a Type I solution has $x^{\circ} = y^{\circ} = 1$ and $z^{\circ} = p$, and a Type II solution has $x^{\circ} = y^{\circ} = z^{\circ} = 1$.

These preliminary results assure that the factorizations of x, y, and z are fundamental for both types of solutions. The results in this paper use these factorizations to reveal patterns that reduce the dimensionality of this conjecture.

2. Results

This section outlines the main results of this paper and provides motivation to a method of solution for the conjecture.

Theorem 1. For a prime number p, if x and y are solution values to (1), then necessarily

$$4xy - (x+y)p = \gcd(y,p)\gcd(xy,x+y).$$
(3)

The most relevant implication from (3) is that

$$z = \frac{xyp}{\gcd(y,p)\gcd(xy,x+y)}.$$
(4)

Notice z, by its definition in Equation (4), has to be an integer. For a given p it suffices to find $x, y \in \mathbb{N}$ in the region defined in (6) so that Equation (3) holds. This, in essence, reduces the dimensionality of the problem by one variable. This is also incredibly important because it reveals that the true nature of this problem depends on the greatest common divisor of the product of two numbers and the sum of those same two numbers. This is not immediately understood from the original description, and this paper hopes to inform and motivate mathematicians who have studied this problem. The next theorem is motivated by the results in [3].

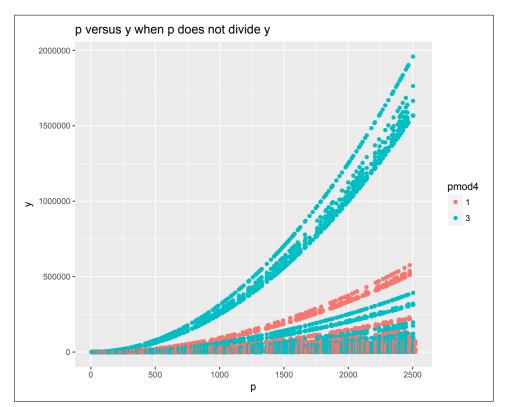


Figure 2: Type I solutions. Red points denote that the prime has remainder 1 after dividing by 4, and blue points denote that the prime has remainder 3 after dividing by 4.

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

$$x = \left\lceil \frac{yp}{4y - p} \right\rceil. \tag{5}$$

The dimensionality of the problem is again reduced by one degree for a Type I solution. It also holds for a vast majority of Type II solutions. Computational evidence suggests that Equation (5) holds for over 97 percent of all solutions [3]. Theorem 2 also helps to reduce the bounds to Lemma 1.

Corollary 1. Given a prime number p, necessary conditions for the solution values $x \le y \le z$ are

$$\left\lceil \frac{p}{4} \right\rceil \le x \le \left\lceil \frac{p}{2} \right\rceil$$
 and $\left\lceil \frac{xp}{4x-p} \right\rceil \le y \le \left\lfloor \frac{2xp}{4x-p} \right\rfloor$. (6)

Further note that for $3 \equiv p \mod 4$, a Type I solution exists when $x = y = \lceil p/2 \rceil$. These are the only types of solutions that exist when $x = \lceil p/2 \rceil$. Considering the bounds in Corollary 1 and the results in Theorems 1 and 2, for a prime p it suffices to find $y \in \mathbb{N}$ so that

$$\left\lceil \frac{p}{2} \right\rceil \le y \le \left\lfloor \frac{2 \left\lceil \frac{p}{4} \right\rceil p}{4 \left\lceil \frac{p}{4} \right\rceil - p} \right\rfloor$$
(7)

and

$$(4y-p)\left\lceil\frac{yp}{4y-p}\right\rceil - yp = \gcd(y,p)\gcd\left(y\cdot\left\lceil\frac{yp}{4y-p}\right\rceil, y+\left\lceil\frac{yp}{4y-p}\right\rceil\right).$$
(8)

For p large enough, it suffices to find a functional expression for y that depends solely on p so that (7) and Equation (8) hold. It is important to note that the functional description for y would have to lie between linear and quadratic behavior in p, although finding the correct description has proven elusive. Figure 2 shows that many patterns exist between y and p for Type I solutions. These patterns are found as modular identities outlined in previous papers [11], but we maintain hope that a general pattern can be found.

3. Proofs

Proof of Lemma 1. First, assume that x < p/4. Substitute the bound p/4 for x in the algebraic expression 4xy - (x+y)p to see that $4xy - (x+y)p \le -p^2/4 < 0$. Also notice that xyp > 0 for all solution values. By definition z = xyp/(4xy - (x+y)p). These observations imply that z < 0, but to be a solution value requires that z > 0.

This creates a contradiction negating the assumption and implies that $x \ge p/4$. Recall that $x \in \mathbb{N}$ to conclude that $x \ge \lfloor p/4 \rfloor$.

Next, assume that y > 2xp/(4x - p). The previous result establishes that 4x - p > 0. When either side of the inequality is multiplied by 4x - p, the result is 4xy - yp > 2xp. This can be rewritten as 4xy - (x + y)p > xp, which indicates that 4xy - (x+y)p > 0 because xp > 0. Multiplying either side of the previous inequality by y > 0 implies that (4xy - (x + y)p)y > xyp. Dividing both sides of the inequality by 4xy - (x + y)p shows that y > xyp/(4xy - (x + y)p) = z. Solution values require that $y \le z$. This creates a contradiction that negates the assumption and implies that $y \le 2xp/(4x - p)$. Recall that $y \in \mathbb{N}$ to conclude that $y \le \lfloor 2xp/(4x - p) \rfloor$.

Next, assume that y < xp/(4x - p). Remember that 4x - p > 0. Multiplying either side of the inequality by 4x - p enables the derivation of 4xy - (x + y)p < 0. Also remember that xyp > 0. Dividing xyp by 4xy - (x + y)p shows that z = xyp/(4xy - (x + y)p) < 0. To be a solution value requires that z > 0. This creates a contradiction that negates the assumption and implies that $y \ge xp/(4x - p)$. Recall that $y \in \mathbb{N}$ to conclude that $y \ge \lceil xp/(4x - p) \rceil$.

Finally, assume that x > 3p/4. For solution values the inequality $x \le y \le z$ makes both y > 3p/4 and z > 3p/4. It implies that 4/p > 1/x + 1/y + 1/z. This contradicts the assumption that x, y, and z are solution values. This contradiction negates the assumption and implies that $x \le 3p/4$. Recall that $x \in \mathbb{N}$ to conclude that $x \le \lfloor 3p/4 \rfloor$.

Proof of Lemma 2. Assume that $gcd(x, p) \neq 1$. This implies that p|x. Lemma 1 shows that $x \leq \lfloor 3p/4 \rfloor$. It is clear that x < p, so $p \nmid x$ which creates a contradiction. This contradiction negates the assumption and implies that x and p are coprime. \Box

Proof of Lemma 3. First, assume that $gcd(y, p^2) = p^2$. This implies that $p^2|y$. Recall from Lemma 1 that $y \leq 2xp/(4x-p)$. Start by showing that 2xp/(4x-p) decreases as the integer $x \geq \lceil p/4 \rceil$ increases. Consider that $2xp(4x-p) + 8xp > 2xp(4x-p) + 8xp - 2p^2$ because $0 > -p^2$, and rewrite it as 2xp(4(x+1)-p) > 2(x+1)p(4x-p). Lemma 1 also shows that x > p/4. As a consequence, both 4x-p > 0 and 4(x+1)-p > 0. Dividing both sides of the previous inequality by (4x-p)(4(x+1)-p) shows that 2xp/(4x-p) > 2(x+1)p/(4(x+1)-p). This shows that 2xp/(4x-p) decreases as the integer $x \geq \lceil p/4 \rceil$ increases. Plugging the smallest possible integer value for x into the expression 2xp/(4x-p) maximizes its possible value over integers $x \geq \lceil p/4 \rceil$. This implies that $y \leq 2\lceil p/4 \rceil p/(4\lceil p/4 \rceil - p)$. Notice that $\lceil p/4 \rceil \leq (p+1)/4$ and $4\lceil p/4 \rceil - p \geq 1$ for all primes p. This shows that $y \leq p(p+1)/2 < p^2$. If $y < p^2$, then $p^2 \nmid y$. which creates a contradiction. This contradiction negates the assumption and implies that $gcd(y, p^2) \neq p^2$.

Next, assume that $gcd(z, p^2) = p^2$. This implies that $p^2|z$. Let $z^* \in \mathbb{N}$ so that $z = z^*p^2$. Solution values necessarily satisfy the equation y = xzp/(4xz - (x+z)p)

which can be written as $y = xz^*p^2/((4xz^* - z^*p)p - x)$. Lemma 2 shows that gcd(x,p) = 1 which implies that $p \nmid ((4xz^* - z^*p)p - x)$. This implies that $p^2|y$, which further implies that $gcd(y,p^2) = p^2$. This creates a contradiction because it was just shown that $gcd(y,p^2) \neq p^2$. This contradiction negates the assumption and implies that $gcd(z,p^2) \neq p^2$.

Proof of Lemma 4. Let gcd(y, p) = p and for the sake of contradiction assume that gcd(z,p) = 1. Let $y^* \in \mathbb{N}$ so that $y = y^*p$. Solution values necessarily satisfy the equation z = xyp/(4xy - (x + y)p) which can be written as $z = xy^*p/(4xy - (x + y)p)$ $((4y^*-1)x-y^*p)$. Lemma 2 shows that gcd(x,p) = 1 and Lemma 3 shows that $gcd(y^*, p) = 1$. The assumption that gcd(z, p) = 1 requires that $p|(4y^* - 1)$. Recalling the proof of Lemma 3, it was shown that $y \leq p(p+1)/2$. This implies that $y^* \leq (p+1)/2$ which further implies that $4y^* - 1 \leq 2p+1$. If $p|(4y^* - 1)$ then either $4y^* - 1 = p$ or $4y^* - 1 = 2p$. $4y^* - 1$ cannot be even, so $4y^* - 1 \neq 2p$. If $4y^* - 1 = p$ then it is similarly clear that $3 \equiv p \mod 4$. Writing $y^* = (p+1)/4$ implies that z = (p+1)x/(4x-(p+1)) which exists, is positive, and is maximized at the integer value $x = (p+5)/4 \ge \lfloor p/4 \rfloor$. This implies that $z \le ((p+1)/4)((p+5)/4)$. Because (p+5)/4 < p for all primes such that $3 \equiv p \mod 4$, this implies that z < p(p+1)/4which further implies that z < y. This creates a contradiction to $x \leq y \leq z$ being solution values. This contradiction negates the assumption that gcd(z, p) = 1. Conclude that gcd(z, p) = p because p is prime.

Proof of Lemma 5. Rewriting Equation (1) with the new notation and performing some algebra gives the equation $4x^{\circ}y^{\circ}z^{\circ}abcd = (x^{\circ}y^{\circ}a + x^{\circ}z^{\circ}b + y^{\circ}z^{\circ}c)p$. Without loss of generality, suppose that a prime $q \neq p$ divides one of x°, y° , and z° , for example $q|x^{\circ}$. The equation would imply that $q|y^{\circ}z^{\circ}c$; however, by definition, $q \nmid y^{\circ}$ because $gcd(x^{\circ}, y^{\circ}) = 1$, $q \nmid z^{\circ}$ because $gcd(x^{\circ}, z^{\circ}) = 1$, and $q \nmid c$ because $gcd(x^{\circ}, c) = 1$. Therefore $q \nmid x^{\circ}$. It is also true that primes $q \neq p$ has $p \nmid y^{\circ}$ and $p \nmid z^{\circ}$.

For a Type I solution, the prime p divides neither x nor y. This implies $x^{\circ} = y^{\circ} = 1$. It is clear that $p|z, p \nmid \gcd(x, z)$, and $p \nmid \gcd(y, z)$ and this implies $z^{\circ} = p$. For a Type II solution the prime $p \nmid x$. This implies that $x^{\circ} = 1$. Lemma 3 indicates that both $\gcd(y, p^2) \neq p^2$ and $\gcd(z, p^2) \neq p^2$. Considering this along with p|c implies that $p \nmid y^{\circ}$ and $p \nmid z^{\circ}$. This means that $y^{\circ} = z^{\circ} = 1$.

Proof of Theorem 1. Using the factorizations of x, y, and z with Lemma 5 makes x = abd, y = acd, z = bcdp, and p = (4abcd - a)/(b + c) for Type I solutions. This makes

$$4xy - (x+y)p = 4a^{2}bcd^{2} - ad(4abcd - a)$$
$$= a^{2}d.$$

Let $c^* \in \mathbb{N}$ so that $c = c^*p$. Using the factorizations of x, y, and z with Lemma 5 makes x = abd, y = acd, z = bcd, and $p = 4abd - (a + b)/c^*$ for Type II solutions. This makes

$$4xy - (x + y)p = p(4a^{2}bc^{*}d^{2} - abd - ad(4abc^{*}d - (a + b)))$$

= $pa^{2}d$.

This implies that, regardless of the type of solution, $4xy - (x+y)p = \gcd(y, p)a^2d$. For a Type I solution p(b+c) = a(4bcd-1) and $\gcd(a, p) = 1$ implies that a|(b+c). Suppose a prime q|((b+c)/a). This implies that q|(4bcd-1). If q|bcd, then q|1. This implies that $\gcd(bcd, (b+c)/a) = 1$. Conclude that

$$gcd(y,p) gcd(xy, x + y) = gcd(a^{2}bcd^{2}, abd + acd)$$
$$= a^{2}d \cdot gcd\left(bcd, \frac{b+c}{a}\right)$$
$$= a^{2}d.$$

For a Type II solution $pc^* = a(4bc^*d - 1) - b$ which implies that if $p|(4bc^*d - 1)$, then p|b. gcd(p,b) = 1 implies that $gcd(p,4bc^*d - 1) = 1$. Also note that $gcd(bc^*d,4bc^*d - 1) = 1$, so it is clear that $gcd(bc^*dp,4bc^*d - 1) = 1$. Conclude that

$$gcd(y,p) gcd(xy, x + y) = p gcd(a^2bc^*d^2p, abd + ac^*dp)$$
$$= pa^2d \cdot gcd (bc^*dp, 4bc^*d - 1)$$
$$= pa^2d.$$

~

This implies that regardless of the type of solution, $gcd(y, p) gcd(xy, x + y) = gcd(y, p)a^2d$. Finally conclude that regardless of the type of solution 4xy - (x+y)p = gcd(y, p) gcd(xy, x + y).

Proof of Theorem 2. Theorem 1 implies that 4xy - (x + y)p = gcd(xy, x + y) for any Type I solutions. Dividing both sides by 4y - p shows that

$$\begin{aligned} x - \frac{yp}{4y - p} &= \frac{\gcd(xy, x + y)}{4y - p} \\ &< \frac{x + y}{4y - p} \\ &< \frac{2y}{2y + (2y - p)}. \end{aligned}$$

By definition y > p/2, so

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

Noting that

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

and $x \in \mathbb{N}$, conclude

$$x - \left\lceil \frac{yp}{4y - p} \right\rceil = 0.$$

Proof of Corollary 1. Consider possible solutions with $x > \lceil p/2 \rceil$. Recall from the proof of Lemma 3, 2xp/(4x-p) decreases as the integer value $x \ge \lceil p/4 \rceil$ increases. This implies $2xp/(4x-p) < 2 \lceil p/2 \rceil p/(4 \lceil p/2 \rceil - p) \le p$ for integers $x > \lceil p/2 \rceil$. Lemma 1 then suggests $y \le \lfloor 2xp/(4x-p) \rfloor < p$. Type II solutions require p|y. It is impossible to have Type II solutions when $x > \lceil p/2 \rceil$ because y < p.

Solution values also require $x \leq y$. If $x > \lfloor p/2 \rfloor$, then $y > \lfloor p/2 \rfloor$. Similarly, it can be inferred from the proof of Lemma 3, yp/(4y-p) decreases as the integer value $y \geq \lfloor p/4 \rfloor$ increases. This implies $yp/(4y-p) < \lfloor p/2 \rfloor p/(4 \lfloor p/2 \rfloor - p) \leq \lfloor p/2 \rfloor$ for integers $y > \lfloor p/2 \rfloor$. Theorem 2 makes clear that for Type I solutions $x = \lfloor yp/(4y-p) \rfloor$, implying that $x \leq \lfloor p/2 \rfloor$ for $y > \lfloor p/2 \rfloor$. Note that this cannot occur, because $x > \lfloor p/2 \rfloor$. Conclude that $x \leq \lfloor p/2 \rfloor$, because there are are no Type I or Type II solutions for $x > \lfloor p/2 \rfloor$. Combining this with Lemma 1 finishes the corollary.

References

- [1] Abdulrahman A. Abdulaziz, On the Egyptian method of decomposing $\frac{2}{n}$ into unit fractions, *Historia Math.* **35** (2008), 1-18.
- [2] M. Bello-Hernández, M. Benito and E. Fernández, On Egyptian fractions, arXiv: 1010.2035v2, (2012).
- [3] K. Bradford and E. Ionascu, A geometric reduction of the Erdős-Straus conjecture, Adv. Mod. and Optim. 1 (2015), vol. 17, 41-54.
- [4] E. S. Croot III, Egyptian Fractions, Ph. D. Thesis, (1994).
- [5] C. Elsholtz and T. Tao, Counting the number of solutions to the Erdős-Straus equation on unit fractions, J. Aust. Math. Soc. 94 (2013), vol. 1, 50-105.
- [6] P. Erdős, $Az 1/x_1 + \cdots + 1/x_n = a/b$ egyenlet egész számú megoldásairól, Mat. Lapok 1 (1950).
- [7] R. Guy, Unsolved Problems in Number Theory, Springer, New York, 2004.
- [8] E. J. Ionascu and A. Wilson, On the Erdős-Straus conjecture, Rev. Roumaine Math. Pures Appl., 56(1) (2011), 21-30.

- [9] D. Li On the equation 4/n = 1/x + 1/y + 1/z, J. Number Theory, 13 (1981), 485-494.
- [10] G. G. Martin, The distribution of prime primitive roots and dense egyptian fractions, Ph. D. Thesis, (1997).
- [11] L. G. Mordell, Diophantine equations, Academic Press, London, (1969).
- [12] M.R. Obláth, Sur l' équation diophantienne $4/n=1/x_1+1/x_2+1/x_3,\ Mathesis$ 59 (1950), 308-316.
- [13] Y. Rav, On the representation of rational numbers as a sum of a fixed number of unit fractions, J. Reine Angew. Math. 222 (1966), 207-213.
- [14] L. A. Rosati, Sull'equazione diofantea $4/n = 1/x_1 + 1/x_2 + 1/x_3$, Boll. Unione Mat. Ital., serie III, Anno IX, (1954), No. 1.
- [15] J.W. Sander, On 4/n = 1/x + 1/y + 1/z and Rosser's sieve, Acta Arith. 49 (1988), 281-289.
- [16] J.W. Sander, On 4/n = 1/x + 1/y + 1/z and Iwaniec' Half Dimensional Sieve, Acta Arith. 59 (1991), 183-204.
- [17] J.W. Sander, Egyptian fractions and the Erdős-Straus Conjecture, Nieuw Arch. Wiskd. (4) 15 (1997), 43-50.
- [18] A. Schinzel, On sums of three unit fractions with polynomial denominators, Funct. Approx. Comment. Math. 28 (2000), 187-194.
- [19] D. G. Terzi, On a conjecture by Erdős-Straus, Nordisk Tidskr. Info. (BIT) 11 (1971), 212-216.
- [20] R.C. Vaughan, On a problem of Erdos, Straus and Schinzel, Mathematika, 17 (1970), 193-198.
- [21] W. Webb, On 4/n = 1/x + 1/y + 1/z, Proc. Amer. Math. Soc. 25 (1970), 578-584.
- [22] W. Webb, On a theorem of Rav concerning Egyptian fractions, Canad. Math. Bull. 18 (1975), no. 1, 155-156.
- [23] W. Webb, On the diophantine equation $k/n = a_1/x_1 + a_2/x_2 + a_3/x_3$, Časopis pro pěstováni matematiy, roč 10 (1976), 360-365.
- [24] K. Yamamoto, On the diophantine equation 4/n = 1/x + 1/y + 1/z, Memoirs of the Faculty of Science, Kyushu University, Ser. A, Vol. **19** (1965), No. 1, 37-47.
- [25] X.Q. Yang, A note on 4/n = 1/x + 1/y + 1/z, Proc. Amer. Math. Soc., 85 (1982), 496-498.