On the Erdös-Straus Conjecture: Properties of Solutions to its Underlying Diophantine Equation

Maria Monks and Ameya Velingker

0.1 Introduction

Number theory is a vast and growing branch of mathematics concerned with the properties of integers. It has numerous applications in a variety of fields such as computer science, cryptography, communications, information theory, physics, and numerical mathematics. Number theory has been especially useful in developing algorithms, such as RSA, a public-key cryptography algorithm which relies on the fact that very large numbers are difficult to factor.

One important topic in number theory is the study of Diophantine equations, equations in which only integer solutions are permitted. The type of Diophantine equation discussed in this paper concerns Egyptian fractions, which deal with the representation of rational numbers as the sum of unit fractions [2]:

$$\frac{a}{b} = \frac{1}{x_1} + \frac{1}{x_2} + \cdots + \frac{1}{x_h}. \tag{0.1.1}$$

An unsolved conjecture due to Paul Erdös and Ernst G. Straus states that the equation

$$\frac{4}{p} = \frac{1}{x} + \frac{1}{y} + \frac{1}{z} \tag{0.1.2}$$

can be solved in positive integers x, y, and z for any integer $p > 1$. Much work has been done on the problem. Mordell [3] has proven that the conjecture is true for all p except possibly cases in which p is congruent to 1^2, 11^2, 13^2, 17^2, 19^2, or 23^2 mod 840. Schinzel [4] has demonstrated that if m and n are relatively prime integers and one can express

$$\frac{4}{mt+n} = \frac{1}{x(t)} + \frac{1}{y(t)} + \frac{1}{z(t)}$$

with $x(t)$, $y(t)$, and $z(t)$ being integer polynomials in t with positive leading coefficients, then n cannot be a quadratic residue mod m. In addition, the conjecture has been verified to be true for all $p < 10^8$ by Franceschine [1] and for all $p < 10^{14}$ by Swett [5]. No counterexamples have been found to the conjecture, and Vaughan [6] has established that if $E_a(N)$ is the number of positive integers $b < N$ for which equation (0.1.1) is insoluble with $h = 3$, then

$$E_a(N) \ll N \ \exp\left\{-\frac{(\ln N)^{2/3}}{C(a)}\right\}$$

where $C(a)$ is a positive integer dependent only on a.

If q is a prime dividing p then $(\frac{p}{q}x, \frac{p}{q}y, \frac{p}{q}z)$ is a solution to (0.1.2) if and only if (x, y, z) is a solution to (0.1.2) with $p = q$. Thus, the conjecture is true for all $p > 1$ if and only if it is true for all prime p. Consequently, in this paper we examine the solutions to equation (0.1.2) for the special case in which p is prime. We prove several properties, including lower and upper bounds as well as divisibility relations. We also restate the conjecture in several equivalent forms, revealing new possible approaches to a solution.

0.2 Main Results

In seeking solutions to equation (0.1.2) it would be helpful to narrow down our search by determining properties that any solution must have. Our first theorem reveals several interesting properties of these solutions.

Define $\nu_a(b) = \max\{c : a^c \mid b\}$. We refer to $\nu_a(b)$ as the exponent of a in b.

Theorem 0.2.1 *Let $a \in \mathbb{Z}^+$ with $a \geq 4$, let $p > a$ be a prime, and let (x, y, z) be a solution in positive integers to $\frac{a}{p} = \frac{1}{x} + \frac{1}{y} + \frac{1}{z}$ with $x \leq y \leq z$.*

(a) Let $q \neq p$ be a prime. At least two of $\nu_q(x), \nu_q(y)$, and $\nu_q(z)$ equal $\max\{\nu_q(x), \nu_q(y), \nu_q(z)\}$, i.e. the highest power of q dividing x, y, or z occurs at least twice among the prime factorizations of x, y, and z.

(b) Let q be a positive integer such that $\gcd(q, p) = 1$. If q divides one of x, y, or z, then q divides the product of the remaining two.

(c) $p \nmid x$ and $p \mid z$ and $x < p$.

In parts (d) through (k) we assume $a = 4$.

(d) $p^2 \nmid y$ and $p^2 \nmid z$.

(e) Let $k = \frac{z}{p}$*. Then* $p \nmid y \Leftrightarrow 4k \equiv 1 \pmod{p}$.

(f) $\gcd(y, z) \neq 1$.

(g) If $z = p \cdot lcm(x, y)$ *then* $p^2 + m$ *has a divisor congruent to* $-p \pmod{m}$*, where* $m = 4 \cdot \gcd(x, y)$.

(h) $x = y$ *if and only if* $(x, y, z) = \left(\frac{p+1}{2}, \frac{p+1}{2}, \frac{p(p+1)}{4}\right)$ *and* x *is even. Furthermore, such a solution exists if and only if* $p \equiv 3 \pmod{4}$.

(i) $\lceil \frac{p}{4} \rceil \leq x \leq \frac{p+1}{2} \leq y$.

(j) If $p \mid y$ *then* $x \leq \lfloor \frac{pc}{3c-1} \rfloor$ *where* $c = \left\lceil \sqrt{\lceil \frac{p}{4} \rceil}\, \right\rceil$.

(k) $p \left\lfloor \frac{5+\sqrt{4p-3}}{4} \right\rfloor \leq z \leq \frac{p^2(p+1)\lceil \frac{p}{4} \rceil}{4\lceil \frac{p}{4} \rceil - p}$.

We postpone all proofs until Section 0.3.

The first four parts of Theorem 0.2.1 pertain to the generalized version of the Erdös-Straus Conjecture. The conjecture itself is the case $a = 4$, and we restrict to this case in the remaining parts.

Part (a) of Theorem 0.2.1 reveals an interesting property of the prime factorizations of x, y, and z. The symmetry implied by this result is seen in the example

$$\frac{4}{193} = \frac{1}{2 \cdot 5^2} + \frac{1}{2^2 \cdot 3 \cdot 5 \cdot 23} + \frac{1}{2^2 \cdot 3 \cdot 5^2 \cdot 23 \cdot 193}.$$

Notice that some primes may occur in all three of x, y, and z (2 and 5 in this case), but the highest power of every prime other than p occurs more than once.

Part (b), which follows from (a), also relates the divisors of x, y, and z in a very symmetric way. In particular, $x \mid yz$ since x and p are relatively prime by part (c). Also, there is exactly one p in the prime factorization of z and at most one p in the prime factorization of y by parts (c) and (d). Thus $z \mid pxy$ and $y \mid xz$. In this way the first four assertions of Theorem 0.2.1 give an insight into the prime factorizations of x, y, and z in any solution.

All solutions to equation (0.1.2) for a given p have the property that $p \nmid x$ and $p \mid z$. However, verification by computer shows that for each of the first 1000 primes p greater than 4 equation (0.1.2) has at least one solution with y divisible by p and at least one with $p \nmid y$. Therefore all solutions naturally fall into two categories: $p \nmid y$ and $p \mid y$.

When $p \nmid y$, the integer $k = \frac{z}{p}$ shares a common prime factor with y by part (f) of Theorem 0.2.1. The solutions in this category are characterized by the property that $4k \equiv 1 \pmod{p}$ by part (e). Notice that the hypothesis of part (g), $z = p \cdot \operatorname{lcm}(x, y)$, implies that $p \nmid y$ because p^2 does not divide z, so part (g) reveals an additional modular property of this category of solutions. Thus parts (e), (f), and (g) narrow our search for solutions of this type.

It is useful to have bounds on the sizes of x, y, and z in the solutions to equation (0.1.2). Computer algorithms for finding solutions to the equation are faster and more efficient if the ranges over which they check for solutions are smaller. Such bounds are presented in parts (i) and (k) of Theorem 0.2.1. Part (h) tells us that the upper bound on x (and lower bound on y) of $\frac{p+1}{2}$ is the best possible bound for infinitely many primes p, for whenever $p \equiv 3 \pmod{4}$ equation (0.1.2) has solutions with x and y equal to $\frac{p+1}{2}$. For solutions for which p divides y, the bound in part (j) further improves upon this upper bound on x.

The properties of solutions described above are of value in attempting to find a general form for a solution in terms of p. In addition, restatements of the conjecture can provide alternative methods of proof. Our remaining results reformulate the conjecture in several ways and reveal new approaches to solving it.

The following theorem is apparently well-known [5], but was derived independently as part of this research project and we include our proof in Section 0.3.

Theorem 0.2.2 *Let $a, b \in \mathbb{Z}^+$. Then $\frac{a}{b} = \frac{1}{x} + \frac{1}{y}$ for some $x, y \in \mathbb{Z}^+$ if and only if there exist two divisors u and v of b such that $a \mid u + v$.*

Rearranging equation (0.1.1) with $a = 4$, $b = p$, and $h = 3$ as

$$\frac{4x_1 - p}{px_1} = \frac{1}{x_2} + \frac{1}{x_3} \tag{0.2.1}$$

we see that to prove the conjecture we need only show that for every prime p, there is an integer x_1 such that px_1 has two divisors whose sum is a multiple of $4x_1 - p$. Furthermore, if $p \mid x_1$, letting $x_1 = px_1'$ we obtain

$$\frac{4x_1' - 1}{px_1'} = \frac{1}{x_2} + \frac{1}{x_3} \tag{0.2.2}$$

and conclude that we need only show that for every prime p, there is an integer x_1' such that px_1' has two divisors whose sum is a multiple of $4x_1' - 1$.

Note that by part (c) of Theorem 0.2.1, one of $x_1,\ x_2$, and x_3 is divisible by p and one is not. Therefore, if there exists a solution to equation (0.2.1) for a given p then there exists one to equation (0.2.2) and vice versa, so the two sufficient conditions shown above are equivalent. Combining these cases, the following corollary results:

Corollary 0.2.3 *The Erdös-Straus conjecture is true if and only if for every prime p, $\exists n \in \mathbb{Z}^+$ such that pn has two divisors that sum to a multiple of either $4n - p$ or $4n - 1$.*

Thus we have restated a conjecture about solutions to a Diophantine equation as a number theoretic property of primes.

Finally, we approach the Erdös-Straus conjecture in another way: we ask which primes p, if any, have solutions with a specific y and z value, in effect working backwards. In particular, we are interested in the case when y is divisible by p. For the remainder of this section, we let $y = pj$ and $z = pk$.

Define the integer valued function α by $\alpha(r,s) = \frac{4rs-r-s}{d \cdot \gcd(\frac{4rs-r-s}{d}, d)}$ where $d = \gcd(r,s)$.

Theorem 0.2.4 *For every $j, k \in \mathbb{Z}^+$ there is at most one prime p such that $\frac{4}{p} = \frac{1}{x} + \frac{1}{pj} + \frac{1}{pk}$ for some positive integer x. Such a prime p exists if and only if $\alpha(j,k)$ is prime, and in this case $p = \alpha(j,k)$.*

Thus if $\alpha(j,k)$ is prime, then there is a solution to equation (0.1.2) with $p = \alpha(j,k)$, $y = pj$, and $z = pk$. From this we obtain the following corollary:

						k					
		1	2	3	4	5	6	7	8	9	10
	1	**2**	**5**	8	**11**	14	**17**	20	**23**	26	**29**
	2	**5**	**3**	**19**	**13**	33	10	**47**	27	**61**	**17**
	3	8	**19**	10	**41**	52	**7**	74	85	32	107
	4	**11**	**13**	**41**	**7**	**71**	**43**	**101**	**29**	**131**	**73**
j	5	14	33	52	**71**	18	**109**	128	147	166	**37**
	6	**17**	10	**7**	**43**	**109**	**11**	155	**89**	**67**	56
	7	20	**47**	74	**101**	128	155	26	209	236	**263**
	8	**23**	27	85	**29**	147	**89**	209	15	**271**	**151**
	9	26	**61**	32	**131**	166	**67**	236	**271**	34	341
	10	**29**	**17**	107	**73**	**37**	56	**263**	**151**	341	**19**

Table 1: $\alpha(j,k)$

Corollary 0.2.5 *If for all prime p there exist positive integers j and k such that $\alpha(j,k) = p$ then the Erdös-Straus conjecture is true.*

Table 1 shows the values for $\alpha(j,k)$ as j and k range from 1 to 10. Primes are in boldface print. Corollary 0.2.5 shows that if every prime occurs somewhere in the extended table, the Erdös-Straus conjecture is true. (Unfortunately, showing that there exists a prime that does not occur in Table 1 does not disprove the conjecture because this only implies there are no solutions such that $p \mid y$.) Therefore it is useful to understand the distribution of primes among the values of $\alpha(j,k)$.

Theorem 0.2.6 *Let $j \in \mathbb{Z}^+$.*

(a) The sequence $\alpha(j,1),\ \alpha(j,2),\ \alpha(j,3),\ldots$ contains infinitely many primes.

(b) Let $k \in \mathbb{Z}^+$. If $\alpha(j,k) = p$ for some prime p, then for all $c \in \mathbb{Z}^+$, $\alpha(j,k+cp)$ is divisible by p.

The preceding theorem demonstrates important properties of the occurrence of primes in the values of α as in Table 1. Part (a) shows that every row (and, by symmetry, every column) contains infinitely many primes. Another interesting property of the extended table, stated in part (b) of Theorem 0.2.6, is that the occurrence of primes in each column follows a pattern similar to the Sieve of Eratosthenes. For example, $\alpha(2,2) = 3$, so every third number after that in the second row (e.g. $\alpha(2,5) = 33$, $\alpha(2,8) = 27$, etc.) is divisible by 3 and therefore is not another prime.

0.3 Proofs

In this section we present the proofs of the results of the previous section.

Proof of Theorem 0.2.2. We first prove the "if" direction of the statement. Suppose there exist divisors u and v of b with $a \mid u + v$. Let $u + v = ma$. Then, $\frac{a}{b} = \frac{u}{mb} + \frac{v}{mb}$, and the two fractions summing to $\frac{a}{b}$ can clearly be reduced to unit fractions because u and v are divisors of b.

Now, we prove the other direction of the statement. Suppose $\frac{a}{b} = \frac{1}{x} + \frac{1}{y}$ for positive integers x and y. Without loss of generality, we may assume that a and b are relatively prime. Let $g = gcd(x, y)$, so that $x = gx_0$ and $y = gy_0$. Simple algebra shows that

$$\frac{ag}{b} = \frac{x_0 + y_0}{x_0 y_0}. \tag{0.3.1}$$

We now show that the fraction $\frac{x_0+y_0}{x_0y_0}$ is in reduced form. Assume the contrary and suppose that there exists a prime q which divides both $x_0 + y_0$ and x_0y_0. Then, q also divides $x_0(x_0 + y_0) - x_0y_0 = x_0^2$, and so $q \mid x_0$. Similarly, $q \mid y_0$ also. However, this contradicts the assumption that $g = gcd(x, y)$. Therefore, the fraction $\frac{x_0+y_0}{x_0y_0}$ must be in lowest terms. From equation (0.3.1), we gather that b is a multiple of x_0y_0. Let $b = mx_0y_0$. Equation (0.3.1) now reduces to $ag = m(x_0 + y_0)$. Because a and b are assumed to be relatively prime, a must also be relatively prime to m. This implies that $a \mid x_0 + y_0$. Therefore, x_0 and y_0 are divisors of b summing to a multiple of a, and the theorem has been proven. ■

Proof of Theorem 0.2.1. *Proof of part (a):* The statement is symmetric with respect to x, y, and z, so we may assume without loss of generality that $t = \nu_q(x)$. If $\nu_q(z) = t$, we are done. Since $\nu_q(z) \leq t$ by definition of maximum, we need only consider the case $\nu_q(z) < t$.

For ease of notation, let $s = \nu_q(y)$ and $r = \nu_q(z)$. Then we can write $x = q^t x'$, $y = q^s y'$, $z = q^r z'$ for some $x', y', z' \in \mathbb{Z}^+$. Furthermore, x', y', and z' are relatively prime to q. Substituting these values into equation (0.1.2) and rearranging we obtain:

$$(aq^r z' - p)q^t x'y' = pq^r z'(y' + q^{t-s}x').$$

(Note that $q^{t-s} \in \mathbb{Z}$ because $s \leq t$ by the definition of maximum.) From this we see that $\nu_q(pq^r z'(y' + q^{t-s}x')) = \nu_q((aq^r z' - p)q^t x'y') \geq t$. By assumption, we have $t > r$, so the transitive property yields $\nu_q(pq^r z'(y' + q^{t-s}x')) > r$. This implies that $q \mid pz'(y' + q^{t-s}x')$. By the definition of prime, $\gcd(q, p) = 1$, and it is stated above that $\gcd(q, z') = 1$, so $q \mid y' + q^{t-s}x'$.

Assume $s < t$. Then $q \mid q^{t-s}x'$, so $q \mid y'$. This contradicts the fact that $\gcd(q, y') = 1$, so $s \geq t$. Since $s \leq t$ as well, $s = t$.

Proof of part (b): Let q be an arbitrary divisor of z that is relatively prime to p. Then $z = qt$ for some $t \in \mathbb{Z}^+$, so by substituting into the generalized Erdös-Straus equation and rearranging the terms,

$$(axy - px - py)qt = pxy.$$

Thus $q \mid pxy$. We conclude from the fact that q and p are relatively prime that $q \mid xy$. By a similar argument, any divisor of y relatively prime to p divides xz, and any divisor of x relatively prime to p divides yz.

Proof of part (c): We first show that $x < p$. Since $x \leq y \leq z$, $\frac{1}{z} \leq \frac{1}{y} \leq \frac{1}{x}$.Therefore,

$$\frac{a}{p} = \frac{1}{x} + \frac{1}{y} + \frac{1}{z} \leq \frac{1}{x} + \frac{1}{x} + \frac{1}{x} = \frac{3}{x}.$$

Thus, since $\frac{a}{p} \leq \frac{3}{x}$ and $4 \leq a$,

$$ax \leq 3p < ap.$$

Therefore $x < p$. It immediately follows that $p \nmid x$.

In order to prove that $p \mid z$, note that $axyz = p(xy + yz + xz)$, so $p \mid axyz$. Since $p > a$ is prime, $\gcd(p, a) = 1$, so $p \mid xyz$. Since p is a prime, we conclude that $p \mid x$ or $p \mid y$ or $p \mid z$. But $p \nmid x$, so $p \mid y$ or $p \mid z$.

We now proceed by indirect proof. Assume $p \nmid z$. Then $p \mid y$, i.e. $y = pj$ for some $j \in \mathbb{Z}^+$. By the assumption, z is a divisor of z that is relatively prime to p, and so it follows from part (b) that $z \mid xy$. By substitution, $z \mid pjx$. Since z and p are relatively prime, $z \mid jx$ and so $z \leq jx$. Also, $x < p$, so

$$z \leq jx < jp = y.$$

This contradicts the hypothesis that $y \leq z$, and we conclude that $p \mid z$.

Proof of part (d): Let $s = \nu_p(y)$ and $t = \nu_p(z)$. We will first prove the following:

$$(t > 1 \text{ or } s > 1) \Rightarrow s = t. \tag{0.3.2}$$

Assume $t > 1$. By the definition of ν, $y = p^s j$ and $z = p^t k$ for some $j, k \in \mathbb{Z}^+$, and furthermore, $p \nmid j$ and $p \nmid k$. We rearrange the generalized Erdös-Straus equation and substitute for y and z to obtain

$$(axp^s j - px - p^{s+1} j)p^{t-1} k = xp^s j.$$

Since $t - 1 > 0$ by assumption, $p \mid (axjp^s - px - p^{s+1}j)p^{t-1}k$ and therefore $p \mid xp^s j$ by substitution. Since p is prime and $p \nmid j, \gcd(j, p) = 1$. Also, by part (c), $\gcd(x, p) = 1$. Therefore $p \mid p^s$, i.e. $s > 0$. Thus

$$(axp^{s-1} j - x - p^s j)p^t k = xp^s j.$$

This gives us $p^t \mid xp^s j$, and therefore $p^t \mid p^s$, so $s \geq t$. By transitivity, $s > 1$, so $p \mid p^{s-1}$. Using modular arithmetic, we have

$$\begin{aligned}(axjp^{s-1} - x - p^s j)k &\equiv (0 - x - 0)k \pmod{p}\\ &\equiv -xk \pmod{p}\\ &\not\equiv 0 \pmod{p}.\end{aligned}$$

In other words, $p \nmid (axp^{s-1}j - x - p^s j)k$. It follows that $t = \nu_p\left((axjp^{s-1} - x - p^s j)p^t k\right) = \nu_p(xp^s j) = s$.

Thus $t > 1 \Rightarrow s = t$. Since this proof did not distinguish between y and z (we did not use the fact that $y \leq z$, only that x was the least), we also have that $s > 1 \Rightarrow s = t$. This proves (0.3.2).

From now on we consider the case $a = 4$.

Since p is a prime greater than 4, p is odd and therefore $p \equiv 1$ or $3 \pmod 4$. We will consider each case separately.

Case 1: $p \equiv 1 \pmod 4$.

Then $p = 4m - 3$ for some $m \in \mathbb{Z}^+$. So $\frac{p}{4} = m - \frac{3}{4}$. By the definition of ceiling, $\lceil \frac{p}{4} \rceil = \lceil m - \frac{3}{4} \rceil = m$.

From equation (0.1.2) we have $\frac{1}{x} < \frac{4}{p}$, which gives us $\frac{p}{4} < x$. But x is assumed to be a positive integer, so

$$\left\lceil \frac{p}{4} \right\rceil \leq x. \tag{0.3.3}$$

By substitution, $m \leq x$. Rewriting this as $\frac{1}{x} \leq \frac{1}{m}$, we obtain:

$$\frac{1}{y} + \frac{1}{z} = \frac{4}{p} - \frac{1}{x} \geq \frac{4}{p} - \frac{1}{m} = \frac{4}{4m-3} - \frac{1}{m} = \frac{3}{m(4m-3)}.$$

Therefore $\frac{1}{y} + \frac{1}{z} \geq \frac{3}{m(4m-3)}$. We now proceed with proof by contradiction.

Assume $p^2 \mid y$ or $p^2 \mid z$. Let $s = \nu_p(y)$ and let $t = \nu_p(z)$. One of s or t is at least 2 by the assumption, and in either case assertion (0.3.2) leads to the conclusion that $s = t$. Therefore p^2 divides both y and z, and we can write $y = p^2 j$ and $z = p^2 k$ for some $j, k \in \mathbb{Z}^+$. By substitution,

$$\frac{1}{p^2 j} + \frac{1}{p^2 k} \geq \frac{3}{m(4m-3)}$$

so

$$\frac{1}{j} + \frac{1}{k} \geq \frac{3p^2}{m(4m-3)} = \frac{3(4m-3)^2}{m(4m-3)} = 12 - \frac{9}{m} \geq 12 - 9 = 3$$

with the last step following from the fact that $m \geq 1$. We now have $3 \leq \frac{1}{j} + \frac{1}{k}$, but also $\frac{1}{j} \leq 1$ and $\frac{1}{k} \leq 1$, so $\frac{1}{j} + \frac{1}{k} \leq 2$. The two inequalities lead to the contradiction $3 \leq 2$.

Thus we conclude that $p^2 \nmid y$ and $p^2 \nmid z$.

Case 2: $p \equiv 3 \pmod 4$.

In this case $p = 4m - 1$ for some $m \in \mathbb{Z}^+$, and an argument similar to Case 1 shows $\frac{1}{y} + \frac{1}{z} \geq \frac{1}{m(4m-1)}$, which again leads to the absurd conclusion that $3 \leq 2$.

Thus we can conclude that $p^2 \nmid y$ and $p^2 \nmid z$. We have now exhausted all cases.

Proof of part (g): Suppose that there exist x, y, z satisfying $z = p \cdot \operatorname{lcm}(x, y)$. Let $g = \gcd(x, y)$ so that $x = ag$ and $y = bg$. Then, $z = abgp$. Substituting these expressions into equation (0.1.2), we obtain

$$g = \frac{(a+b)p + 1}{4ab}. \tag{0.3.4}$$

This implies that $a \mid bp + 1$ and $b \mid ap + 1$. Let

$$bp + 1 = ca$$

$$ap + 1 = db$$

Solving this system, we find that $a = \frac{p+d}{cd-p^2}$ and $b = \frac{p+c}{cd-p^2}$. Substituting these expressions for a and b into equation (0.3.4) yields $g = \frac{cd-p^2}{4}$, implying that $4 \mid cd - p^2$. Clearly, $p \nmid c$ and $p \nmid d$. Since a and b are integers, the above expressions indicate that $p + d$ and $p + c$ are divisible by $m = cd - p^2$. Thus, $c, d \equiv -p \pmod m$, and $m = 4g = cd - p^2$ satisfies the condition that $p^2 + m$ has divisors congruent to $-p \pmod m$, as desired.

Proof of part (h): Assume $x = y$. Then by substituting into equation (0.1.2) and rearranging we obtain

$$2(2x - p)z = px.$$

Since $z \in \mathbb{Z}$, $2(2x - p) \mid px$. Thus $2 \mid px$ and $2x - p \mid px$. Since p is a prime greater than 4, p is odd, so $2 \mid x$. Hence x is even.

It is easily seen that

$$\forall b, c, d \in \mathbb{Z}, \;\; \gcd(b, c) = 1 \Rightarrow \gcd(b, c + bd) = 1. \tag{0.3.5}$$

Also, we know by part (c) that $\gcd(x, p) = 1$, so $\gcd(x, -p) = 1$. Thus by (0.3.5), $\gcd(x, 2x - p) = 1$. Therefore $2x - p \mid p$, but p is prime, implying there are only two possibilities: $2x - p = p$ or $2x - p = 1$. If $2x - p = p$, then $x = p$, which contradicts part (c). We are left with $2x - p = 1$. Solving yields $x = \frac{p+1}{2}$ and thus $y = \frac{p+1}{2}$ and $z = \frac{p(p+1)}{4}$. By substitution, $\frac{p+1}{2}$ is also even, so $4 \mid p + 1$, i.e. $p \equiv 3 \pmod 4$. Hence if a solution with $x = y$ exists, $p \equiv 3 \pmod 4$. This completes the forward directions of both propositions in the theorem.

Assume $p \equiv 3 \pmod 4$. Then $p + 1$ is divisible by 4, so both $\frac{p+1}{2}$ and $\frac{p(p+1)}{4}$ are positive integers. Since $\frac{2}{p+1} + \frac{2}{p+1} + \frac{4}{p(p+1)} = \frac{4}{p}$, we have $(\frac{p+1}{2}, \frac{p+1}{2}, \frac{p(p+1)}{4})$ is a solution in positive integers to equation (0.1.2). It was given that $p > 4$, so $\frac{p}{4} > 1$. Therefore $\frac{p(p+1)}{4} > \frac{p+1}{2}$, so we have a solution (x, y, z) satisfying $x = y \leq z$.

Proof of part (f): We will use the method of indirect proof. Assume $\gcd(y, z) = 1$. Part (c) states that $p \mid z$, which leads to the conclusion that $p \nmid y$, for otherwise y and z would not be relatively prime. Since p is prime, $\gcd(p, y) = 1$. Now y satisfies the conditions of part (b), so $y \mid xz$. Because of the assumption, $y \mid x$, so $y \leq x$. Since it was given that $x \leq y$, $x = y$. Thus $x = y = \frac{p+1}{2}$ and $z = \frac{p(p+1)}{4}$by part (h). Since p is prime and $z \in \mathbb{Z}$, $4 \mid p + 1$. Therefore $p + 1 = 4k$ for some $k \in \mathbb{Z}^+$. By substitution, $z = pk$ and $y = x = 2k$, so $\gcd(y, z) = k$. Furthermore, since $p > 4$, $k > 1$, implying $\gcd(y, z) > 1$, which contradicts our assumption.

We can therefore conclude that $\gcd(y, z) \neq 1$.

Proof of part (e): Part (c) gives us $p \mid z$, so $k \in \mathbb{Z}$. Also, $z = pk$, and substituting this into equation (0.1.2) and rearranging gives us

$$(4k - 1)xy = (x + y)pk. \tag{0.3.6}$$

We will now prove the forward direction of the theorem. Assume $p \nmid y$. From equation

(0.3.6) we know $p \mid (4k-1)xy$. But since p is prime, $\gcd(y,p)=1$ by the assumption, and $\gcd(x,p)=1$ by part (c). Thus $p \mid 4k-1$. It immediately follows that $4k-1 \equiv 0 \pmod{p}$, meaning $4k \equiv 1 \pmod{p}$.

For the reverse direction, we will prove the contrapositive of the statement. Assume $p \mid y$. Then $y = pj$ for some $j \in \mathbb{Z}^+$, so we can substitute into equation (0.3.6) and simplify to obtain

$$(4k-1)xj = (x+y)k.$$

By our assumption, we know $y \equiv 0 \pmod{p}$, and part (c) implies that $x \not\equiv 0 \pmod{p}$, so $x+y \not\equiv 0 \pmod{p}$. Also, $p \nmid k$, for otherwise p^2 would divide z, which contradicts part (d). Therefore $p \nmid (x+y)k$, and by substitution, $p \nmid (4k-1)xj$. It follows that $4k-1 \not\equiv 0 \pmod{p}$, so $4k \not\equiv 1 \pmod{p}$.

We now have $p \mid y \Rightarrow 4k \not\equiv 1 \pmod{p}$, so the contrapositive is true: $4k \equiv 1 \pmod{p} \Rightarrow p \nmid y$. Both directions have been considered, and thus $p \nmid y \Leftrightarrow 4k \equiv 1 \pmod{p}$.

Proof of part (i): The lower bound on x was proven as assertion (0.3.3). We will consider two cases for this proof, $p \mid y$ and $p \nmid y$.

Case 1: $p \mid y$.

In this case, $p \leq y \leq z$, so $\frac{4}{p} = \frac{1}{x} + \frac{1}{y} + \frac{1}{z} \leq \frac{1}{x} + \frac{1}{p} + \frac{1}{p}$. From this we have $\dfrac{2}{p} \leq \dfrac{1}{x}$ and thus

$$x \leq \frac{p}{2} < \frac{p+1}{2} < p \leq y.$$

Therefore $x \leq \frac{p+1}{2} \leq y$.

Case 2: $p \nmid y$.

Let $k = \frac{z}{p}$. By part (e), $4k \equiv 1 \pmod{p}$. Thus $k = \frac{pr+1}{4}$ for some $r \in \mathbb{Z}^+$. Since $k \in \mathbb{Z}^+$, $r \geq 1$, so $k \geq \frac{p+1}{4}$, and therefore $z \geq \frac{p(p+1)}{4}$. Using $\frac{1}{z} \leq \frac{4}{p(p+1)}$ and $\frac{1}{y} \leq \frac{1}{x}$ in equation (0.1.2) we obtain:

$$\frac{4}{p+1} = \frac{4}{p} - \frac{4}{p(p+1)} \leq \frac{4}{p} - \frac{1}{z} = \frac{1}{x} + \frac{1}{y} \leq \frac{2}{x}.$$

By transitivity, $\frac{4}{p+1} \leq \frac{2}{x}$, so

$$x \leq \frac{p+1}{2}.$$

Also, $\frac{1}{x} + \frac{1}{y} < \frac{4}{p}$ because $\frac{1}{z}$ is positive, and the fact that $\frac{1}{y} \leq \frac{1}{x}$ implies $\frac{2}{y} \leq \frac{1}{x} + \frac{1}{y}$. Again by transitivity, $\frac{2}{y} < \frac{4}{p}$, so

$$\frac{p}{2} < y.$$

From this, by the definition of ceiling, $\left\lceil \frac{p}{2} \right\rceil \leq y$, and since p is an odd prime, $\left\lceil \frac{p}{2} \right\rceil = \frac{p+1}{2}$. Thus we have $x \leq \frac{p+1}{2} \leq y$. We have now exhausted all cases.

Proof of part (j): Since $p \mid y$ is given and $p \mid z$ by part (c), $y = pj$ and $z = pk$ for some $j, k \in \mathbb{Z}^+$. Part (c) tells us that $\gcd(x, p) = 1$, which in turn implies that $x \mid yz$ by part (b). Therefore $x \mid p^2kj$, implying $x \mid kj$ since x and p are relatively prime. By the definition of divides, $x \leq kj$. We also have the inequality $pj \leq pk$, so $j \leq k$. Multiplying both sides of this inequality by k produces $kj \leq k^2$. Hence $x \leq k^2$. Furthermore, $\frac{1}{x} < \frac{4}{p}$, so $\frac{p}{4} < x$. Since $x \in \mathbb{Z}^+$, $\left\lceil \frac{p}{4} \right\rceil \leq x$. By transitivity, we have $\left\lceil \frac{p}{4} \right\rceil \leq k^2$, so

$$\sqrt{\left\lceil \frac{p}{4} \right\rceil} \leq k$$

and so

$$\left\lceil \sqrt{\left\lceil \frac{p}{4} \right\rceil} \right\rceil \leq k$$

because $k \in \mathbb{Z}^+$. Let $c = \left\lceil \sqrt{\left\lceil \frac{p}{4} \right\rceil} \right\rceil$, so that $c \leq k$. Then $\frac{1}{pk} \leq \frac{1}{pc}$.

Also, $1 \leq j$, so $\frac{1}{pj} \leq \frac{1}{p}$. Using these inequalities, the following results:

$$\begin{aligned} \frac{4}{p} &= \frac{1}{x} + \frac{1}{pj} + \frac{1}{pk} \\ \frac{4}{p} &\leq \frac{1}{x} + \frac{1}{p} + \frac{1}{pc} \\ \frac{3}{p} - \frac{1}{pc} &\leq \frac{1}{x} \\ x &\leq \frac{pc}{3c-1} \end{aligned}$$

To further improve this bound, note that $x \in \mathbb{Z}^+$, so $x \leq \left\lfloor \frac{pc}{3c-1} \right\rfloor$.

Proof of part (k): First we prove that $z \neq p$ by contradiction. Assume $z = p$. Since $\gcd(y, z) \neq 1$ by part (f), $p \mid y$. But $y \leq z$, so $y = p$ and substituting into equation (0.1.2) and solving for x gives us $x = \frac{p}{2}$. Since p is odd but $x \in \mathbb{Z}$, we have a contradiction. Therefore $z \neq p$ and so $z \geq 2p$.

We will next prove that

$$\text{if } p \geq 4n^2 - 10n + 7 \text{ for some positive integer } n, \text{ then } z \geq np. \tag{0.3.7}$$

We use induction on n. The statement is clearly true for $n = 1$ and $n = 2$ because $z \geq 2p$. Now suppose the statement is true for $n = c$, with $c \geq 2$. We demonstrate that the statement is also true for $n = c+1$. Let $p \geq 4(c+1)^2 - 10(c+1) + 7 = 4c^2 - 2c + 1$. By the induction hypothesis, $z \geq cp$. By part (c), z is divisible by p. Thus, in order to prove that $z \geq (c+1)p$, it suffices to show that $z \neq cp$. For the sake of contradiction, assume that there exists a decomposition of $\frac{4}{p}$ in which $z = cp$, i.e. $\frac{4}{p} = \frac{1}{x} + \frac{1}{y} + \frac{1}{cp}$ for some positive integers x and y. Rearranging the equation yields $\frac{1}{x} + \frac{1}{y} = \frac{4c-1}{cp}$. By theorem 0.2.2, there exist two divisors u and v of cp for which $4c - 1 \mid u + v$. We now consider three different cases:

Case 1: Both u and v are divisors of c. Then $u + v \leq 2c < 4c - 1$ which contradicts the statement that $4c - 1 \mid u + v$.

Case 2: Both u and v are divisible by p. Then we may let $u = pd_1$ and $v = pd_2$ so that d_1 and d_2 are divisors of c. Therefore, $4c - 1 \mid p(d_1 + d_2)$. We note that for $c \geq 2$, $4c - 1 < 4c^2 - 2c + 1 \leq p$. Hence, p and $4c - 1$ are relatively prime to each other and so $4c - 1 \mid d_1 + d_2$. However, $d_1 + d_2 \leq 2c < 4c - 1$, which yields a contradiction.

Case 3: Only one of u and v is divisible by p. Without loss of generality assume $p \mid u$. We let $u = pd_1$ and $v = d_2$ where d_1 and d_2 are divisors of c. To calculate values for x and y which correspond to our choice of u and v, note that

$$\frac{4c-1}{cp} = \frac{pd_1 + d_2}{\frac{cp(pd_1+d_2)}{4c-1}} = \frac{pd_1}{\frac{cp(pd_1+d_2)}{4c-1}} + \frac{d_2}{\frac{cp(pd_1+d_2)}{4c-1}} = \frac{1}{\frac{c(pd_1+d_2)}{d_1(4c-1)}} + \frac{1}{\frac{cp(pd_1+d_2)}{d_2(4c-1)}}.$$

By part (c), $p \nmid x$, from which we deduce that $x = \frac{c(pd_1+d_2)}{d_1(4c-1)}$ and $y = \frac{cp(pd_1+d_2)}{d_2(4c-1)}$. Note that

$$\frac{(4c-2)d_2}{d_1} \leq c(4c-2) < 4c^2 - 2c + 1 < p.$$

We observe that

$$y = \frac{cp(pd_1+d_2)}{d_2(4c-1)} > \frac{cp\left(\left(\frac{(4c-2)d_2}{d_1}\right)d_1 + d_2\right)}{d_2(4c-1)} = cp = z.$$

This is clearly a contradiction because we must have $y \leq z$.

Since all three cases result in a contradiction, $z \neq cp$ and assertion (0.3.7) follows.

Finally, define $f(a) = 4a^2 - 10a + 7$. Let r be the (greater) positive real number such that $f(r) = p$. Solving this for r, we find that $r = \frac{5+\sqrt{4p-3}}{4}$. Applying assertion (0.3.7), we find that $z \geq rp$ if r is an integer. If r is not an integer, we simply note that $p \geq f(\lfloor r \rfloor)$, since $r > 2$. Hence, $z \geq \lfloor r \rfloor p$. Either way, $z \geq \lfloor \frac{5+\sqrt{4p-3}}{4} \rfloor p$.

For the upper bound, note that since $\gcd(\frac{z}{p}, p) = 1$ by parts (c) and (d), $\frac{z}{p} \mid xy$ by part (b). Thus $z \mid pxy$, and so $z \leq pxy$. Also, $\lceil \frac{p}{4} \rceil \leq x$ by part (i) and $\frac{1}{y} \geq \frac{1}{z}$, so

$$\frac{2}{y} = \frac{1}{y} + \frac{1}{y} \geq \frac{1}{y} + \frac{1}{z} = \frac{4}{p} - \frac{1}{x} \geq \frac{4}{p} - \frac{1}{\lceil \frac{p}{4} \rceil} = \frac{4\lceil \frac{p}{4} \rceil - p}{p\lceil \frac{p}{4} \rceil}.$$

Thus $y \leq \frac{2p\lceil \frac{p}{4} \rceil}{4\lceil \frac{p}{4} \rceil - p}$. But $x \leq \frac{p+1}{2}$ by part (i), so

$$z \leq pxy \leq p\left(\frac{p+1}{2}\right)\left(\frac{2p\lceil \frac{p}{4} \rceil}{4\lceil \frac{p}{4} \rceil - p}\right).$$

Therefore $z \leq \frac{p^2(p+1)\lceil \frac{p}{4} \rceil}{4\lceil \frac{p}{4} \rceil - p}$. ■

Before proving our remaining theorems we require a simple Lemma.

Lemma 0.3.1 *$\forall k, j \in \mathbb{Z}^+, 4kj - k - j > kj$.*

Proof. Assume the negation of the statement, i.e. that $4kj - k - j \leq kj$ for some $j, k \in \mathbb{Z}^+$. Then $3kj - j \leq k$, implying $j \leq \frac{k}{3k-1}$. Since $k \in \mathbb{Z}^+$, it is clear that $\frac{k}{3k-1} < 1$, so $j < 1$. This contradicts the fact that $j \in \mathbb{Z}^+$, and we have $4kj - k - j > kj$ by indirect proof. ■

Proof of Theorem 0.2.4. We begin by proving that if $j, k \in \mathbb{Z}^+$, $g = \gcd(j, k)$, and $p > 4$ is a prime,

$$\left(\exists x \in \mathbb{Z}^+, \frac{4}{p} = \frac{1}{x} + \frac{1}{pj} + \frac{1}{pk}\right) \Leftrightarrow \frac{4kj - k - j}{g} \mid pg. \tag{0.3.8}$$

By the definition of gcd, $j = j'g$ and $k = k'g$ for some j', $k' \in \mathbb{Z}^+$. Furthermore, $\gcd(j', k') = 1$. We now prove the forward implication.

Assume $\exists x \in \mathbb{Z}^+, \frac{4}{p} = \frac{1}{x} + \frac{1}{pj} + \frac{1}{pk}$. Solving for x, we find that

$$x = \frac{pkj}{4kj - k - j}. \tag{0.3.9}$$

Substituting for j and k,

$$x = \frac{pk'j'g^2}{4k'j'g^2 - k'g - j'g} = \frac{pk'j'g}{4k'j'g - k' - j'}.$$

Since $x \in \mathbb{Z}^+$, $4k'j'g - k' - j' \mid pk'j'g$. However, $\gcd(j', k') = 1$, so by statement (0.3.5), $\gcd(j', 4k'j'g - k' - j') = 1$, and similarly $\gcd(k', 4k'j'g - k' - j') = 1$. Thus $4k'j'g - k' - j' \mid pg$. But $4k'j'g - k' - j' = \frac{4kj-k-j}{g}$, so by substitution, $\frac{4kj-k-j}{g} \mid pg$. This completes the forward direction of the proof.

Now, assume $\frac{4kj-k-j}{g} \mid pg$. Then $4kj - k - j \mid pg^2$, and $pg^2 \mid pkj$, so $4kj - k - j \mid pkj$. By the definition of divides, $\exists x \in \mathbb{Z}^+$, $x(4kj - k - j) = pkj$, or equivalently,

$$\frac{1}{x} + \frac{1}{pj} + \frac{1}{pk} = \frac{4}{p}.$$

This completes the proof of assertion (0.3.8).

Note that $\frac{4}{p} = \frac{1}{x} + \frac{1}{pj} + \frac{1}{pk}$ for some positive integer x if and only if $4kj - k - j \mid pkj$ by equation (0.3.9). We wish to show that if there exists a prime p such that $4kj - k - j \mid pkj$, then that is the only prime for which the statement is true.

Assume there exists a prime p such that $4kj - k - j \mid pkj$. It follows from Lemma 0.3.1 that $4kj - k - j \nmid kj$. This implies that $p \mid 4kj - k - j$ because p is prime. By part (d) of Theorem 0.2.1, $\gcd(j, p) = \gcd(k, p) = 1$. Hence $\gcd(g, p) = 1$, so $p \mid \frac{4kj-k-j}{g}$ as well.

Finally, assume there is another prime $q \neq p$ such that $\frac{4}{q} = \frac{1}{x} + \frac{1}{qj} + \frac{1}{qk}$ for some $x \in \mathbb{Z}^+$. Assertion (0.3.8) implies $\frac{4kj-k-j}{g} \mid qg$. However, this implies $p \mid qg$, which is a contradiction since both q and g are relatively prime to p. Therefore there exists at most one prime p such that $\frac{4}{p} = \frac{1}{x} + \frac{1}{pj} + \frac{1}{pk}$ for some positive integer x.

We now show that when such a prime p exists, $p = \alpha(j,k)$. Since $p \mid \frac{4kj-k-j}{g}$,

$$\frac{4kj-k-j}{g} = pm \tag{0.3.10}$$

for some $m \in \mathbb{Z}^+$. Since $\frac{4kj-k-j}{g} \mid pg$, we have $m \mid g$. This implies m is a common divisor of g and $\frac{4kj-k-j}{g}$ and thus $m \mid \gcd(g, \frac{4kj-k-j}{g})$. We now prove by contradiction that $m = \gcd(g, \frac{4kj-k-j}{g})$.

Assume $m \neq \gcd(g, \frac{4kj-k-j}{g})$. Then $\frac{\gcd(g,\frac{4kj-k-j}{g})}{m}$ is a divisor of g greater than 1. We know $p = \frac{4kj-k-j}{gm}$ by (0.3.10), and $\frac{4kj-k-j}{g} = w \cdot \gcd(g, \frac{4kj-k-j}{g})$ for some $w \in \mathbb{Z}^+$ by the definition of gcd. Thus $p = \frac{4kj-k-j}{gm} = w \cdot \frac{\gcd(g,\frac{4kj-k-j}{g})}{m}$, which is a multiple of a divisor of g greater than 1, as stated above. This implies $\gcd(g,p) \neq 1$, but this contradicts part (d) of Theorem 0.2.1.

Therefore $m = \gcd(g, \frac{4kj-k-j}{g})$ and by substitution, $p = \alpha(j,k)$. ■

Proof of Theorem 0.2.6. *Proof of part (a):* Let $A_k = \alpha(j,k)$ for all $k \in \mathbb{Z}^+$. Let $k \in \mathbb{Z}^+$ be arbitrary. We will begin by proving indirectly that $A_k \neq 1$.

Assume $A_k = 1$. Then $\frac{4kj-k-j}{g} = \gcd(\frac{4kj-k-j}{g}, g)$ by the definition of A_k. From this we have $\frac{4kj-k-j}{g} \mid g$, so $4kj-k-j \mid g^2$. Also, $g^2 \mid kj$, so $4kj-k-j \mid kj$ and thus

$$4kj-k-j \leq kj.$$

This contradicts Lemma 0.3.1, and we can conclude $\forall k \in \mathbb{Z}^+, a_k \neq 1$.

Let B be a sequence defined by $B_k = (4j-1)k - j$. Then B is an arithmetic sequence with first term $3j-1$ and difference $4j-1$, which are relatively prime because their difference, j, is relatively prime to $4j-1$ by statement (0.3.5). By Dirichlet's Theorem, the sequence $B_1,\ B_2, ...$ contains infinitely many primes.

Note that $\forall k \in \mathbb{Z}^+, A_k \mid B_k$. Let $q \in \mathbb{Z}^+$ such that B_q is prime. Then $A_q \mid B_q$, so $A_q = 1$ or $A_q = B_q$. Since $A_q \neq 1$, we have $A_q = B_q$ and thus A_q is also prime. Therefore the set of all q such that B_q is prime is a subset of the set of all q such that A_q is prime, and so the latter set is also infinite. Thus the sequence $\alpha(j,1),\ \alpha(j,2), ...$ contains infinitely many primes.

Proof of part (b): Let $p = \alpha(j,k)$ and assume p is prime. Let $c \in \mathbb{Z}^+$ be arbitrary, let $g_2 = \gcd(j, k+cp)$, and for ease of notation let $b = g_2 \cdot \gcd(g_2, \frac{\beta(j,k+cp)}{g_2})$. We know by part (d) of Theorem 0.2.1 that $p \nmid j$, so $p \nmid g_2$ and therefore $p \nmid b$. Now, $\alpha(j, k+cp) = \frac{4j(k+cp)-j-k-cp}{b} = \frac{4kj-k-j+(4jc-c)p}{b}$. Because $\alpha(j,k) \mid 4kj-k-j$, we have $p \mid 4kj-k-j$. Thus $p \mid 4kj-k-j+(4jc-c)p$, and since $p \nmid b$, we conclude that $p \mid \alpha(j, k+cp)$. ■

0.4 Conclusion

There are many interesting properties of the solutions to the underlying equation of the Erdös-Straus conjecture, a famous and long-standing open problem in number theory. Motivated by the desire to prove this conjecture, our research has revealed several new such properties, involving symmetry in the prime factorizations of solutions, limitations and bounds on their sizes, and modular restrictions. One application of these properties is to improve the efficiency of computer algorithms which search for solutions and/or test for other properties. Furthermore, these results may aid in finding a proof of the conjecture by narrowing the possibilities.

Additionally, our restatements of the Erdös-Straus conjecture open up several new paths to attack the problem. In particular, the result on the existence of decompositions of a rational number into two unit fractions provides a sufficient condition to prove the conjecture. In addition, a complete understanding of the occurrence of primes among the values of the α function would also be sufficient to resolve the problem. It is our hope that future research using either of these two approaches will eventually result in a proof of the Erdös-Straus conjecture itself.

Bibliography

[1] Franceschine, N. "Egyptian Fractions." MA Dissertation, Sonoma State Coll. CA, 1978.

[2] Guy, R. K. "Egyptian Fractions." §D11 in *Unsolved Problems in Number Theory, 2nd ed.* New York: Springer-Verlag, pp. 158-166, 1994.

[3] Mordell, L. J. *Diophantine Equations.* London: Academic Press, pp. 287-290, 1969.

[4] Schinzel, A. "Sur Quelques Propriétés des Nombres $3/n$ et $4/n$ où n est un Nombre Impair." *Matheseis* **65**, 219-222, 1956.

[5] Swett, A. "The Erdos-Strauss Conjecture." Rev. 10/28/99. http://math.uindy.edu/swett/esc.htm

[6] Vaughan, R. C. "On a Problem of Erdös, Straus, and Schinzel." *Mathematika* **17**, 193-198, 1970.