On the Diophantine Equation $\sum X_i = \prod X_i$

By M. L. Brown

Abstract. The diophantine equation $X_1 + \cdots + X_k = X_1 \cdots X_k$ has at least one solution in positive integers for $k \geq 2$. The set of integers $k$ for which this is the only solution are investigated; in particular, this set is conjectured to be a known finite sequence.

The equation $f_k(X) = X_1 + X_2 + \cdots + X_k - X_1 X_2 \cdots X_k = 0$ has the solution, for $k \geq 2$, given by $X_1 = 2, X_2 = k, X_3 = X_4 = \cdots = X_k = 1$. Schinzel showed that there are no other solutions in positive integers, apart from permutations of this given solution, for $k = 6$ and $k = 24$. Misiurewicz [2] states that $k = 2, 3, 4, 6, 24, 144, 174, 444$ are the only values of $k < 1000$ for which $f_k(X) = 0$ has essentially one solution, as above. But the number $144$ in this list (given in both [2] and [1, D24]) is probably a misprint for $114$, for with this correction Misiurewicz’s assertion is then correct (evidently $144$ will not do because of the extra solution $1^{141} \cdot 2 \cdot 4 \cdot 21 = 168 = (141) \cdot 1 + 2 + 4 + 21$). We report here on some further calculations with this equation.

Proposition 1. The equation $f_k(X) = 0$, for $k \geq 4$, has only one solution in positive integers, apart from permutations, if and only if the following conditions hold:

1. $k - 1$ is a prime number.
2. Let $s, n$ be any integers, if any, with $3 \leq s \leq \log_2 k + 1$, $2^{s-2} \leq n \leq (k^{1/s} + 1)^{s-2}$ and $n$ being a product $x_1 \cdots x_{s-2}$ of $s - 2$ integers $x_i \geq 2$. Put $t = x_1 + x_2 + \cdots + x_{s-2}$. Then no factor of $N = (k - s + t)n + 1$ is congruent to $-1$ modulo $n$ except possibly for $n - 1$ and $N/n - 1$.

Proof. Let $f_k(x) = 0$ be a solution in positive integers; we may suppose that $x_1, \ldots, x_k$ are precisely those integers among $x_1, \ldots, x_k$ which do not equal 1. Thus $x_1 \cdots x_s = k - s + x_1 + \cdots + x_s$. Since $x_i \geq 2$ for all $i \leq s$, it follows that $2^s \leq k + s$. It is then elementary to show that for $k \geq 4$ we have $s \leq \log_2 k + 1$.

The case $s = 1$ is easily ruled out, so we next consider the case $s = 2$. A solution $f_k(x) = 0$ with $s = 2$ gives $x_1(x_2 - 1) = k - 2 + x_2$. If this is distinct from the given solution we have that $x_2 - 1$ does not equal $k - 1$ and is a proper factor of $k - 2 + x_2$. It follows that $x_2 - 1$ is a proper factor of $k - 1$. Thus no other solution exists if and only if $k - 1$ is a prime number.

Suppose now $s \geq 3$. Given integers $x_1, \ldots, x_{s-2} \geq 2$, put $n = x_1 x_2 \cdots x_{s-2}$ and $t = x_1 + x_2 + \cdots + x_{s-2}$. Then there are integers $x_{s-1}, x_s \geq 2$ with $f_k(x) = k - s + t + x_{s-1} + x_{s-2} - nx_{s-1} x_{s-2} = 0$ if and only if if $f = x_{s-1} n - 1$ is a factor of.
and not equal to, \( k - s + t + x_s \). Put \( N = (k - s + t)n + 1 \). We deduce \( x_{s - 1}, x_s \geq 2 \) exist as required if and only if \( N \) has a factor \( f \neq n - 1 \) and \( \neq N/(n - 1) \) with \( f \) congruent to \(-1\) modulo \( n \). It remains to bound \( n \).

Let \( x_1, \ldots, x_s \) be a solution, as above, with the \( x_i \) arranged in increasing magnitude. Consider the problem of maximizing \( n \) subject to \( f_k(x) = 0 \), \( x_i \geq 2 \), and the other stated constraints on \( x_i \), except that we now allow nonintegral values. Since \( n \), as a function of \( x_1, \ldots, x_s \), is nonconstant in the given region, it must take its extreme values for extreme values of the variables \( x_i \). It is easy to see that for the maximum value of \( n \) we must have all the \( x_i, i \leq s \), being equal. Thus the maximum value of \( n \) is \( x^{s - 2} \) where \( x \) is the positive root of the equation \( x^s = k - s + sx \). Plainly, \( k^{1/s} \leq x \leq k^{1/s} + 1 \); thus \( n \leq (k^{1/s} + 1)^{s - 2} \) as required.

**Corollary 1.** Suppose \( f_k(X) = 0 \) has only 1 essential solution in positive integers. Then:

1. \( k - 1 \) and \( 2k - 1 \) are prime numbers.
2. \( 6|k \) if \( k > 4 \).
3. \( 4k + 1 \) and \( 4k + 5 \) are sums of two squares for \( k > 4 \).

**Proof.** (1) Take \( s = 3, n = 2 \) in the proposition. (2) follows from (1). (3) Take \( n = 4, s = 3 \) or \( n = 2, s = 4 \) in the proposition and apply Fermat's criterion, noting from (2) that neither \( 4k + 1 \) nor \( 4k + 5 \) is divisible by 3 for \( k > 4 \).

Proposition 1 can be used to give an algorithm for testing if an integer \( k \) has the required property; the number of steps required is at most \( O(k^{3/2 + \epsilon}) \), for all \( \epsilon > 0 \). Using this algorithm, a PET 4032 microprocessor was used to test suitable values of \( k \); this revealed the discrepancy in Misiurewicz's list, though it is easy to check by hand using Proposition 1 that \( k = 114 \) should be in the list. No other values of \( k < 11,000 \) were found for which \( f_k(X) = 0 \) has one solution, this computation taking 40 minutes of computing time. We thus end with:

**Conjecture.** The only values of \( k \) for which \( f_k(X) = 0 \) has one solution are \( k = 2, 3, 4, 6, 24, 114, 174, 444 \).

**Added in Proof.** With a different program, the conjecture has now been verified for all \( k \leq 50,000 \).

Department of Pure Mathematics
University College
P.O. Box 78
Cardiff CF1 1XL, Great Britain