MATHEMATICS OF COMPUTATION
Volume 72, Number 244, Pages 2063–2070
S 0025-5718(02)01504-1
Article electronically published on December 18, 2002

COMPUTATIONAL INVESTIGATIONS OF THE PROUHET-TARRY-ESCOTT PROBLEM

PETER BORWEIN, PETR LISONĚK, AND COLIN PERCIVAL

ABSTRACT. We describe a method for searching for ideal symmetric solutions to the Prouhet-Tarry-Escott Problem. We report results of extensive searches for solutions of sizes up to 12. We found two solutions of size 10 that are smaller by two orders of magnitude than the solution found by A. Letac in the 1940s, which was the smallest size 10 solution known before our search.

1. The Prouhet-Tarry-Escott Problem

The *Prouhet-Tarry-Escott Problem* (PTE Problem) is an old unsolved problem in Diophantine number theory. In its most general setting the PTE Problem asks for two distinct multisets of integers $X = \{x_1, \ldots, x_n\}$ and $Y = \{y_1, \ldots, y_n\}$ such that

$$\sum_{i=1}^{n} x_i^e = \sum_{i=1}^{n} y_i^e \quad \text{for } e = 1, 2, \ldots, k \tag{1}$$

for some integer $k \leq n-1$. Any pair X, Y that satisfies (1) is called a *solution* of the PTE Problem; this is denoted by $X =_k Y$. If $k = n-1$, then the solution is called *ideal* and n is called the *size* of this ideal solution. In this paper we restrict our attention to ideal solutions.

If X, Y are multisets satisfying the system (1) and $f(t) = \alpha t + \beta$ is a linear transformation with rational coefficients, then the multisets $f(X), f(Y)$ satisfy (1); we say that X, Y and $f(X), f(Y)$ are *equivalent* solutions. In what follows we consider only integer solutions up to equivalence.

The PTE Problem has a long history and is, in some form, over 200 years old. In this article we do not intend to survey the known results. Extensive accounts of the history and known results (with the exception of the new solutions of sizes 10 and 12 listed in Section 2.6 below) are available in [1, 2, 5]. Many numerical solutions (both ideal and nonideal) can be found at Chen Shuwen's website [4]. Ideal solutions to the PTE Problem are known only for sizes $n = 1, 2, \ldots, 10$ and $n = 12$. Parametric ideal solutions are known for $n = 1, 2, \ldots, 8$ and $n = 10$; in each case they give rise to infinitely many nonequivalent ideal solutions.

Received by the editor November 9, 2001 and, in revised form, March 25, 2002.
2000 *Mathematics Subject Classification.* Primary 11D72, 11Y50; Secondary 11P05.
Research presented in this paper was partially supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) and partially by the National Centre of Excellence MITACS.

1.1. **Ideal symmetric solutions.** The PTE Problem is often considered in its much more restrictive *symmetric version*, which cuts the number of variables by one half. For n odd, the symmetric version takes the form

$$\sum_{i=1}^{n} x_i^e = 0 \quad \text{for } e = 1, 3, \ldots, n-2. \tag{2}$$

If $x_1, \ldots, x_n$ satisfy (2), then

$$\{x_1, \ldots, x_n\} =_{n-1} \{-x_1, \ldots, -x_n\}$$

is an *odd ideal symmetric solution* of the PTE Problem. For n even, the symmetric version of the PTE Problem takes the form

$$\sum_{i=1}^{n/2} x_i^e = \sum_{i=1}^{n/2} y_i^e \quad \text{for } e = 2, 4, \ldots, n-2. \tag{3}$$

If $x_1, \ldots, x_{n/2}$ and $y_1, \ldots, y_{n/2}$ satisfy (3), then

$$\{x_1, \ldots, x_{n/2}, -x_1, \ldots, -x_{n/2}\} =_{n-1} \{y_1, \ldots, y_{n/2}, -y_1, \ldots, -y_{n/2}\}$$

is an *even ideal symmetric solution* of the PTE Problem.

1.2. **An example from history.** Nearly all known solutions were found without using computers, with the exception of the single known size 12 solution found by Chen and others in 1999 and the two small size 10 solutions whose discovery we describe in this article. It is interesting to use modern symbolic computation tools to study the parametric solutions originally obtained by tedious hand computations, mostly in the 1940s. Sometimes that results in an improvement of those solutions.

For example, Gloden [5] on pages 42–43 describes the derivation of a parametric ideal symmetric solution of size 7. Because of the apparent difficulty of computations that had to be carried out by hand, the final solution, which uses four parameters f, g, k, l, is not even found explicitly; only a numerical example is given at the bottom of page 43. With a computer algebra system such as Maple or Mathematica one can quickly discover that the seven polynomials implicitly given by Gloden have a large common divisor; after dividing it out the parametric solution turns out to depend only on f and k and becomes quite manageable. We list it below: The values $\alpha_1, \ldots, \alpha_7$ satisfy $\sum_{i=1}^{7} \alpha_i^e = 0$ for $e = 1, 3, 5$.

$$\begin{aligned}
\alpha_1 &= -\left(f^2 - kf + k^2\right)\left(-3\,kf^2 + k^3 + f^3\right), \\
\alpha_2 &= -\left(k - f\right)\left(f + k\right)\left(f^2 - 3\,kf + k^2\right) f, \\
\alpha_3 &= \left(-f + 2\,k\right)\left(-f^2 - kf + k^2\right) kf, \\
\alpha_4 &= \left(k - f\right)\left(k - 2\,f\right)\left(-f^2 + kf + k^2\right) k, \\
\alpha_5 &= \left(k - f\right)\left(f^4 - 2\,kf^3 - k^2f^2 + k^4\right), \\
\alpha_6 &= -\left(k^4 - 2\,fk^3 - k^2f^2 + 4\,kf^3 - f^4\right) k, \\
\alpha_7 &= -\left(k^4 - 5\,k^2f^2 + 4\,kf^3 - f^4\right) f.
\end{aligned}$$

For example, plugging in $f = 3$, $k = 1$ yields the following ideal symmetric solution of size 7:

$$\{-7, 24, 33, -50, -38, -13, 51\} =_6 \{7, -24, -33, 50, 38, 13, -51\}.$$

2. Search for ideal symmetric solutions

Our primary interest was a search for ideal symmetric PTE solutions of sizes $9 \le n \le 12$. For $n \le 8$ and $n = 12$, the coordinates of the smallest ideal symmetric solution of size n grow (in absolute value) roughly as n^2. However, the smallest ideal symmetric solution of size 10 that was known before our computations grossly exceeds this trend, and we were curious whether smaller solutions existed. We indeed found two solutions of size 10 that are substantially smaller—by two orders of magnitude—than the solution found by Letac in the 1940s.

2.1. **The constant associated with an ideal solution.**

Lemma 1 ([2])**.** *The following are equivalent:*

$$\sum_{i=1}^{n} x_i^e = \sum_{i=1}^{n} y_i^e \quad \text{for} \quad e = 1, \dots, k,$$

$$\deg\Big(\prod_{i=1}^{n}(x - x_i) - \prod_{i=1}^{n}(x - y_i)\Big) \le n - k - 1,$$

$$(z-1)^{k+1} \mid \sum_{i=1}^{n} z^{x_i} - \sum_{i=1}^{n} z^{y_i}.$$

Proof. This is an easy exercise in differentiation and manipulation with symmetric polynomials. □

Corollary 2. *The pair of multisets* $\{x_1, \dots, x_n\}, \{y_1, \dots, y_n\}$ *is an ideal PTE solution if and only if*

$$\prod_{i=1}^{n}(x - x_i) - \prod_{i=1}^{n}(x - y_i) = C \tag{4}$$

for some real constant C*.*

From now on we will associate with any ideal PTE solution the corresponding constant C (or, equivalently, $-C$) and we will speak about *the constant* C without any further explanation. It is generally a highly composite number, since

$$C = \prod_{\ell=1}^{n}(x_i - y_\ell) = -\prod_{\ell=1}^{n}(x_\ell - y_j) \quad \text{for all } i, j = 1, \dots, n.$$

2.2. **Divisibility results.**

Proposition 3. *Let* $\{x_1, \dots, x_n\} =_{n-1} \{y_1, \dots, y_n\}$ *be two multisets of integers that constitute an ideal PTE solution, and suppose that a prime* p *divides the constant* C *associated with this solution. Then we can reorder the integers* y_i *so that*

$$x_i \equiv y_i \pmod{p} \quad \text{for } i = 1, \dots, n.$$

Proof. Assume that p is a prime dividing C. Let $\mathbb{F}_p$ denote the field with p elements. From (4) we get $\prod_{i=1}^{n}(x - x_i) = \prod_{i=1}^{n}(x - y_i)$ in $\mathbb{F}_p[x]$ (the ring of univariate polynomials over $\mathbb{F}_p$) and, since $\mathbb{F}_p[x]$ is a unique factorization domain, it follows that the multisets $\{x_1, \dots, x_n\}$ and $\{y_1, \dots, y_n\}$ are equal as subsets of $\mathbb{F}_p$. □

TABLE 1. Divisibility of constants associated with ideal symmetric solutions.

n	some primes dividing the constant C
9	$2, \ldots, 13$
10	$2, \ldots, 7, 13, 17, 23$
11	$2, \ldots, 19, 31$
12	$2, \ldots, 11, 17, 19, 29$
13	$2, \ldots, 41$

The relevance of Proposition 3 to a computer search for PTE solutions is obvious: It cuts the size of the search space by the factor of $1/p$ in *each coordinate* because of the pairing property modulo p. It is therefore desirable to find as large as possible primes p such that p must divide C for *any* ideal (or ideal symmetric) PTE solution.

Some general results about the divisibility of C for ideal (not necessarily ideal symmetric) PTE solutions were obtained by Rees and Smyth [6]. Recall that our goal is to search for ideal *symmetric* solutions (over integers). By an exhaustive search for solutions in the ideal symmetric form over some small finite fields we discovered—not surprisingly—that stronger divisibility properties hold in this case. Table 1 summarizes the results that can be obtained this way (i.e., by solving over finite fields) for ideal *symmetric* solutions for problems of sizes $9 \le n \le 13$. More results can probably be proved using other methods.

2.3. **Odd sizes.** Let $n = 2m+1$ denote the size of an odd ideal symmetric solution. Equation (4) takes the form

$$\prod_{i=1}^{2m+1} (x + x_i) = \prod_{i=1}^{2m+1} (x - x_i) + C\,.$$

We observe that

$$\prod_{i=1}^{2m+1} (x_i + x_j) = C, \qquad j = 1, \ldots, 2m+1,$$

and thus

$$\frac{1}{C} \cdot \prod_{i=m+2}^{2m+1} (x_i + x_j) = \prod_{i=1}^{m+1} (x_i + x_j)^{-1}, \qquad j = 1, \ldots, m+1.$$

If $x_1, \ldots, x_{m+1}$ are distinct, then there is a unique polynomial $f(x)$ of degree m satisfying

$$f(x_j) = \prod_{i=1}^{m+1} (x_i + x_j)^{-1}, \qquad j = 1, \ldots, m+1\,.$$

(The cases when $x_i = -x_j$ for some $1 \le i < j \le m+1$ lead to trivial PTE solutions, and thus we do not worry about them.) Given $x_1, \ldots, x_{m+1}$, we can compute the corresponding $f(x)$ quite easily. Now we simply note that

$$f(-x_j) = \frac{1}{C} \cdot \prod_{i=m+2}^{2m+1} (x_i - x_j) = 0, \qquad j = m+2, \ldots, 2m+1,$$

and that we can thereby compute the unique multiset $\{x_{m+2}, \ldots, x_{2m+1}\}$ by solving the equation $f(-x) = 0$.

We solve this equation numerically, using Newton's method. Since it suffices to perform a test which eliminates the vast majority of candidates and perform a more rigorous test later, we use Newton's method to test if the root with largest real part is within 10^{-6} of an integer within the search limit. To do this we start Newton's iteration with a value above the search limit and continue until either the iteration converges, the iteration fails to monotonically decrease, or the 60th iteration is reached. If the iteration converges, the solution is checked to see if it is close to an integer, and the possible solution is reported. If the iteration is not monotonically decreasing, then there are complex roots, and we can discard the candidate. If the 60th iteration is reached without convergence, then it is impossible for the roots to be distinct integers within the search range, and we again discard the candidate.

In order to speed up the convergence, we start by using a modified Newton iteration which takes double the "step" at each iteration. Based on the assumptions that the roots are all real and less than the current value of the iterand, this can only "jump over" at most one root, an event which will be observed through the sign of $f(x)$ changing. Once this is observed, we return to the previous value of the iterand—one which is still greater than all the roots—and continue the classical Newton iteration from that point. As a fortuitous side effect, this provides quadratic convergence to double roots, ensuring that the iteration should always finish within the maximum 60 iterations. (It is impossible for a solution to the Prouhet-Tarry-Escott Problem to contain triple roots. [1])

This reduces the search space from $2m+1$ dimensions to $m+1$ dimensions; to test whether a solution of size $2m+1$ exists with a given subset of size $m+1$, we need merely compute what the remaining m x_j's would be and check if they are integers.

We can further speed up the search by making use of necessary prime factors of the constant C. From Section 2.2 we know that the constant C for a solution of size $2m+1$ must satisfy certain divisibility conditions. We also know from Proposition 3 that the multisets $\{x_1, \ldots, x_n\}$ and $\{-x_1, \ldots, -x_n\}$ are equal modulo any prime p that divides C. That is, modulo p there is one (or more) $x_i \equiv 0$, and the rest pair off, $x_i \equiv -x_j \pmod{p}$.

Consequently, we can restrict our search to $(m+1)$-tuples $(x_1, \ldots, x_{m+1})$ satisfying

$$\begin{aligned}
x_1 &\equiv 0 && \pmod{p_1},\\
(x_1+x_2)\cdot x_1 &\equiv 0 && \pmod{p_2},\\
(x_2+x_3)\cdot(x_1+x_2) &\equiv 0 && \pmod{p_1},\\
(x_3+x_4)\cdot(x_1+x_2+x_3) &\equiv 0 && \pmod{p_2},\\
&\vdots\\
(x_m+x_{m+1})\cdot\sum_{i=1}^{m} x_i &\equiv 0 && \pmod{p_{(m+1) \bmod 2}},
\end{aligned}$$

with p_1 and p_2 being the two largest primes dividing C and the values $j \bmod 2$ taken as 1 or 2. That is, for each $j = 1, \ldots, m$, the value x_{j+1} is paired off with x_j modulo $p_{(j+1) \bmod 2}$ unless all the $x_1, \ldots, x_j$ are already paired off modulo this prime.

2.4. **Even sizes.** Let $n = 2m$ be the size of an even ideal symmetric solution. Equation (4) in this case takes the form

$$\prod_{i=1}^{m}(x^2 - x_i^2) = \prod_{i=1}^{m}(x^2 - y_i^2) + C,$$

and we can observe that

$$\prod_{i=1}^{m}(y_j^2 - x_i^2) = C, \qquad j = 1, \ldots, m,$$

and thus

$$\frac{1}{C} \cdot \prod_{i=m-k+2}^{m}(y_j^2 - x_i^2) = \prod_{i=1}^{m-k+1}(y_j^2 - x_i^2)^{-1}, \qquad j = 1, \ldots, k,$$

for any convenient k. We can now proceed as in the case of odd sizes, except that instead of needing x_1 through x_{m+1}, we need x_1 through x_{m-k+1} and y_1 through y_k.

Again as with odd sizes, we can speed up the search by making use of the divisibility requirements for C. Since $x_i \equiv y_i \pmod{p}$ for any prime p dividing the constant C, we can restrict our search to values satisfying

$$\begin{aligned} x_i^2 &\equiv y_i^2 \pmod{p_1}, \\ (x_{i+1}^2 - y_i^2) \cdot \sum_{j=1}^{i}(x_j^2 - y_j^2) &\equiv 0 \pmod{p_2}, \end{aligned}$$

with p_1 and p_2 again being the two largest primes dividing C.

2.5. **Implementation.** The searches we conducted used a total of roughly 10^{17} floating-point operations. We were able to gain access to unused computing capacity on over 100 Celeron 500 based computers over the course of several months. Given that these computers were vastly overpowered for their primary use, for all practical purposes this meant that we had access to 50 gigaflops of computing power available 24 hours per day.

Searching for solutions to the PTE Problem has the useful property of being "naturally parallel." There is absolutely no need for interprocessor communication: All that is necessary is for processors to be assigned a portion of the search range, and to report back any solutions found.

The actual implementation of the search was surprisingly simple. Windows NT has built-in support for "services" (programs are run even if no user is logged into the machine), including remote management tools. This made it possible for a single service program to be written, placed on the network, and to be installed simultaneously onto all 100+ computers.

To solve the problem of computers being turned off in mid-calculation, we took a simple approach: Only keep track of which search blocks have been finished. Rather than attempting to resume a search part way though, every time a computer finished a block it would report this completion back to the server, and when the server logs were later inspected, any unfinished blocks were re-issued until they were successfully completed.

TABLE 2. Results of search for ideal symmetric PTE solutions.

size	*search limit*	*result*
9	2000	no new solutions found
10	1500	two new solutions found
11	2000	no solutions found
12	1000	no new solutions found

2.6. **Results.** The results of our searches for ideal symmetric solutions, whose coordinates are all bounded by a value that we call the "search limit," are summarized in Table 2.

For size 9, we verified that there are no previously unknown ideal symmetric solutions with coordinates less than 2000. The only primitive solutions found were

$$\{-169, -161, -119, -63, 8, 50, 132, 148, 174\}$$

and

$$\{-98, -82, -58, -34, 13, 16, 69, 75, 99\}.$$

(By a "primitive solution" we mean a solution whose terms do not have a nontrivial common factor. Note that no two primitive integer solutions are equivalent in the sense defined in Section 1.)

For size 10, two previously unknown solutions were found. Before this paper, the only known solutions of size 10 were large solutions derived from rational points on an elliptic curve [2], with the smallest (found by A. Letac in the 1940s) being

$$\begin{aligned}\{&\pm 436, \pm 11857, \pm 20449, \pm 20667, \pm 23750\}\\ =_9\quad &\{\pm 12, \pm 11881, \pm 20231, \pm 20885, \pm 23738\}.\end{aligned}$$

We found the two solutions

$$\{\pm 71, \pm 131, \pm 308, \pm 180, \pm 307\} =_9 \{\pm 99, \pm 100, \pm 301, \pm 188, \pm 313\}$$

and

$$\{\pm 18, \pm 245, \pm 331, \pm 471, \pm 508\} =_9 \{\pm 103, \pm 189, \pm 366, \pm 452, \pm 515\}.$$

These (and multiples of them) are the only even ideal symmetric solutions of size 10 whose coordinates are less than 1500.

For size 11, we searched up to 2000 and did not find any solutions. At present no solutions of size 11 are known.

For size 12 only the solution

$$\{\pm 22, \pm 61, \pm 86, \pm 127, \pm 140, \pm 151\} =_{11} \{\pm 35, \pm 47, \pm 94, \pm 121, \pm 146, \pm 148\},$$

first discovered by Nuutti Kuosa, Jean-Charles Meyrignac and Chen Shuwen in 1999 [4], was found.

3. An open problem

We have also computed many ideal symmetric solutions for sizes up to 8. Interestingly, more than 85% of all nonequivalent ideal symmetric solutions $\{x_1, \ldots, x_7\}$ of size 7 that we computed (some of them with coordinates exceeding 6000) are subject to a relation of the form

$$x_1 + x_2 + x_3 = x_4 + x_5 + x_6 + x_7 = 0.$$

A similar observation holds for the two known ideal symmetric solutions of size 9 (see the previous section), which are both subject to a relation of the form $x_1 + x_2 + x_3 + x_4 = x_5 + x_6 + x_7 + x_8 + x_9 = 0$; the larger of the two solutions of size 9 is in fact subject to as many as four independent relations of this form. It would be interesting to understand the nature of this phenomenon, which is sometimes consciously used to construct ideal symmetric solutions (such as for example Gloden's parametric solution of size 7 introduced in Section 1.2). A similar approach is taken by Bremner [3], who finds rational solutions of $x^5 + y^5 + z^5 = u^5 + v^5 + w^5$ by introducing two auxiliary linear equations $x + y + z = u + v + w$ and $x - y = u - v$.

One possible cause for this phenomenon is the divisibility properties; given that such relations must exist modulo any prime dividing the constant C associated with the solution, it is not altogether surprising that such relations so often occur over the integers. Any statistical treatment of this is however made impossible by a lack of understanding of the "statistical distribution" of solutions.

4. Acknowledgments

We are indebted to Academic Computing Services of Simon Fraser University, in particular to Paul Geenen and Bunny Penn Tan, for allowing us to access the unused capacity of personal computer labs and for installing our software on the lab networks.

References

1. P. Borwein, *Excursions in Computational and Diophantine Number Theory.* Springer-Verlag, New York (to appear).
2. P. Borwein, C. Ingalls, The Prouhet-Tarry-Escott Problem revisited. *Enseign. Math.* **40** (1994), 3–27. MR **95d:**11038
3. A. Bremner, A geometric approach to equal sums of fifth powers. *J. Number Theory* **13** (1981), no. 3, 337–354. MR **83g:**14017
4. Chen Shuwen, The Prouhet-Tarry-Escott Problem. `http://member.netease.com/~chin/eslp/-TarryPrb.htm`
5. A. Gloden, *Mehrgradige Gleichungen.* Second Edition. P. Noordhoff, Groningen, 1944. MR **8:**441f
6. E. Rees, C. Smyth, On the constant in the Tarry-Escott Problem. Cinquante ans de polynômes (Paris, 1988), 196–208, Lecture Notes in Math., 1415, Springer, Berlin, 1990. MR **91g:**11030

Department of Mathematics, Simon Fraser University, Burnaby, British Columbia, Canada
E-mail address: `pborwein@cecm.sfu.ca`

Department of Mathematics, Simon Fraser University, Burnaby, British Columbia, Canada
E-mail address: `lisonek@cecm.sfu.ca`

Department of Mathematics, Simon Fraser University, Burnaby, British Columbia, Canada
Current address: Wadham College, Oxford University, Oxford, England
E-mail address: `cperciva@sfu.ca`