Efficient Algorithms for Decomposing Integers as Sums of Few Tetrahedral Numbers*

Tong-Nong Lin[1], Yu-Cheng Lin[2], Cheng-Chen Tsai[1], Meng-Tsung Tsai[1], and Shih-Yu Tsai[1]

[1] Institute of Information Science, Academia Sinica, Taiwan
{wilsonlin,saiss2022,mttsai,shihyutsai}@iis.sinica.edu.tw

[2] The Affiliated Senior High School of National Taiwan Normal University, Taiwan
lance991144@gmail.com

Abstract. Pollock conjectures that every natural number can be expressed as a sum of at most five tetrahedral numbers. It remains unknown whether this conjecture holds, and Watson proved that sums of at most eight tetrahedral numbers suffice to express all natural numbers.

We devise two algorithms to decompose integers as sums of few tetrahedral numbers. Our first algorithm can decompose any given integer n into a sum of at most eight tetrahedral numbers in $O(\log^3 n/\log\log n)$ time with probability $1-1/n^{\Omega(1)}$, assuming the extended Riemann hypothesis. Our second algorithm can deterministically decompose all integers in $[1,\ell]$ into sums of the fewest possible tetrahedral numbers in $O(\ell)$ time using $O(\ell^{2/3})$ space, assuming a conjecture over the integers in $[1,\ell]$ and the Pollock's conjecture on tetrahedral numbers. While the conjectures for all numbers is unproven, its validity for all numbers in $[1,\ell]$ can be verified in $O(\ell)$ time.

As a result of our second algorithm, we can show that the Pollock's conjecture holds for all natural numbers up to 2.82×10^{21}. This significantly improves upon the previous known bound of 3.77×10^{15}.

Keywords: Pollock's conjecture · image set of cubic congruences · empirical verification

1 Introduction

Pollock conjectured in 1851 that every natural number can be expressed as a sum of at most five ***tetrahedral numbers*** [9], where for each $i \geq 1$ the ith tetrahedral number T_i is defined as

$$\binom{i+2}{3}$$

In 1928, Yang proved that sums of at most nine tetrahedral numbers suffice to express all natural numbers [16], using Legendre's three-square theorem [7]. Subsequently, James proved in 1934 that sums of at most eight tetrahedral numbers suffice to express all

* This paper is not eligible for the best student paper award.

sufficiently large natural numbers [5], using a method due to Landau [6]. Then, in 1935, Hua generalized James' result by showing that, for each integer D, sums of at most eight values in the form of

$$(i+1)D + T_i$$

suffice to express all sufficiently large natural numbers. Watson further strengthened James' and Hua's results in 1952, removing the requirement for the input integers to be sufficiently large [15]. In Watson's proof, the key observation is that the intersection of the image sets of some cubic congruences is non-empty (implicitly) using Hensel's lemma [12].

In this paper, we strengthen Watson's result by showing that this intersection is not only non-empty but also comprises a constant fraction of all possible incongruent values, as detailed in Lemma 4. We then use this finding to devise an algorithm to decompose any given integer n into at most eight tetrahedral numbers, as stated in Theorem 1. The time complexity analyses in this paper are conducted on the standard model, WordRAM, where basic arithmetic operations over $O(\log n)$-bit operands take $O(1)$ time.

Theorem 1. *There exists an algorithm capable of representing an integer n as a sum of at most eight tetrahedral numbers in* $O(\log^3 n/\log\log n)$ *time with probability* $1 - 1/n^{\Omega(1)}$*, assuming the extended Riemann hypothesis (ERH). This assumption can be removed if the decomposition of integers into three squares, a building block of this algorithm, can be done efficiently without assuming ERH.*

In addition to the mathematical attempts to prove the Pollock's conjecture, there have been computational efforts. In what follows, ***k-numbers*** refer to integers that cannot be expressed as sums of fewer than k tetrahedral numbers, but can be expressed as sums of exactly k tetrahedral numbers, following the notations in [3,1].

Salzer and Levine in 1958 [13] verified that the Pollock's conjecture holds for all integers in the closed interval $[1, 10^6]$. They reported that each integer in $[343868, 10^6]$ is a k-number for some $k \leq 4$ and provided a list of 241 5-numbers. Deng and Yang in 1994 [3] improved the upper bound of the computation from 10^6 to 10^9 and reported that no new 5-number was found in their computation. It may be worth noting that, if there exists a k-number n for some $k \geq 6$ and $n - n/k > 343,867$, then there exists a new 5-number. Hence, reporting no new 5-number implies no new k-number for $k \geq 5$ was found. Chou and Deng in 1997 [1] further improved the upper bound of the computation to 4×10^{10}. Again, no new 5-number was found. Given these empirical results, it is conjectured in [3,1] that 343,867 is the largest 5-number. Furthermore, combining the above empirical result and a technique introduced in [3,1], they show that every positive integer up to $3,771,207,667,368,141$ is a k-number for some $k \leq 5$. This gives the known best upper bound of integers for which the Pollock's conjecture holds.

As our second contribution in this paper, we devise an algorithm that can express all integers in $[1, \ell]$ as sums of the fewest possible tetrahedral numbers in $O(\ell)$ time and $O(\ell^{2/3})$ space, assuming a conjecture over integers in $[1, \ell]$ and the Pollock's conjecture on tetrahedral numbers. This result is formally stated in Theorem 2. Though the conjectures are unproven, it can be verified in $O(\ell)$ time. A similar algorithm to express all integers in $[1, \ell]$ as sums of the fewest possible tetrahedral numbers with the same

time complexity was claimed in [1], but the analysis is based on the following heuristic. For each 4-number n, let $r(n)$ be the rank of the smallest tetrahedral number $T_{r(n)}$ such that $n - T_{r(n)}$ is a 3-number. Their linear-time algorithm requires the assumption that $\lim_{n\to\infty} r(n) = O(1)$. They also reported that $r(n) \leq 68$ for all 4-numbers less than or equal to 4×10^{10}, which we found to be incorrect. Indeed, there are 68 4-numbers $n \leq 4 \times 10^{10}$ with $r(n) > 68$, as listed in Table 2, and the largest $r(n)$ among them is 98.

Theorem 2. *There exists a deterministic algorithm capable of representing an integer n as a sum of the fewest possible tetrahedral numbers in $O(\ell)$ time and $O(\ell^{2/3})$ space, assuming that the number of the 4-numbers n in $[1, \ell]$ whose $r(n) = O(\log n)$ is at least $\ell - \ell^{1/3}$ and the Pollock's conjecture on tetrahedral numbers. This assumption over the integers in $[1, \ell]$ is unproven, but it can be empirically verified in $O(\ell)$ time with probability $1 - 1/\ell^{\Omega(1)}$.*

As a result of our second algorithm, we empirically verify that for all integers up to 10^{14} there exists no 5-number other than the known 241 ones. Together with an auxiliary computation, all integers up to 2.82×10^{21} are k-numbers for some $k \leq 5$. This gives a new upper bound on the integers for which the Pollock's conjecture holds and significantly improves upon the previous known bound. As a remark,

Theorem 3. *The Pollock's conjecture on tetrahedral numbers holds for all integers up to*

$$2.82 \times 10^{21}.$$

Paper Organization. In Section 2, we strengthen Watson's result by showing that the intersection of the image sets of some cubic congruences comprises a constant fraction of all possible incongruent values. Based on this findings we devise an efficient randomized algorithm to represent any given integer n as a sum of at most eight tetrahedral numbers. Then, in Section 3, we devise a linear-time algorithm that can express all integers in $[1, \ell]$ as sums of the fewest possible tetrahedral numbers in $O(\ell)$ time and $O(\ell^{2/3})$ space, assuming a conjecture over integers in $[1, \ell]$. We also present how to empirically verify this unproven conjecture over the integers in $[1, \ell]$ in $O(\ell)$ time. Finally, in Section 4, we draw conclusions from our work.

2 Expressing Integers as Sums of At Most Eight Tetrahedral Numbers

We will prove Theorem 1 in this section. Our techniques are mainly based on Watson's techniques [15], Hensel's Lemma [12], and Chinese Remainder Theorem [12]; however, the key lemma (Lemma 4) that we obtain is very different from those lemmas in Watson's paper [15]. Lemma 4 and some simple arguments together yield a proof for Theorem 1.

Lemma 1 (A restatement of [15, Lemma 3]). *Let n, m be positive integers so that $(m, 6) = 1$ and m is a multiple of some square number. If m and n satisfy*

$$\left(\frac{1}{8} + \frac{9}{125}\right) m^3 < n < \frac{1}{4} m^3, \tag{1}$$

then there exist integers x, y, k *so that* $0 \le x, y < 3m/5$, $0 \le k < m^2/8$, *and*

$$6n = x^3 - x + y^3 - y + \frac{1}{8}\left(6m^3 - 24m + 6m(8k+3)\right). \tag{2}$$

By Legendre's three-square theorem [7], $8k+3$ can be expressed as a sum of three squares for any integer $k \ge 1$. Hence, we can rewrite Eq. (2) as, for some $u, v, w \ge 0$,

$$\begin{aligned}
6n &= x^3 - x + y^3 - y + \frac{1}{8}\left(6m^3 - 24m + 6m(u^2+v^2+w^2)\right) \\
&= 6T(x) + 6T(y) + \frac{1}{8}\sum_{x\in\{u,v,w\}} 2m^3 - 8m + 6x^2m \\
&= 6T(x) + 6T(y) + \sum_{x\in\{u,v,w\}} \left(\frac{m+x}{2}\right)^3 - \frac{m+x}{2} + \left(\frac{m-x}{2}\right)^3 - \frac{m-x}{2} \\
&= 6T(x) + 6T(y) + \sum_{x\in\{u,v,w\}} 6T\left(\frac{m+x}{2}\right) + 6T\left(\frac{m-x}{2}\right)
\end{aligned}$$

Because $u^2+v^2+w^2 \equiv 3 \pmod 8$, it follows that u, v, w are odd integers. Additionally, since $0 \le k < m^2/8$, we have $0 \le u, v, w \le m$. Combining that m is an odd integer, both $(m+x)/2$ and $(m-x)/2$ for $x \in \{u, v, w\}$ are non-negative integers. Hence, n can be expressed as a sum of at most eight tetrahedral numbers.

To implement the above decomposition, we need the following building blocks:

1. Computing a proper m.
2. Computing x and y such that $0 \le x, y < 3m/5$ and satisfying the congruence relation
$$6n \equiv x^3 - x + y^3 - y \pmod m$$
3. Representing $8k+3$ as a sum of three squares.

In Sections 2.1 to 2.3, we show how to realize these building blocks in $O(\log^3 / \log\log n)$ time, thereby proving Theorem 1.

2.1 Computing a proper *m*

In this section, we present how to compute a proper m in $O(\log n)$ time.

Lemma 2. *For every sufficiently large n, there exists an integer*

$$m = 5^a \cdot 7^b \cdot 11^c \text{ for some } a \ge 2, b \ge 0, c \ge 0$$

so that Eq. (1) *holds.*

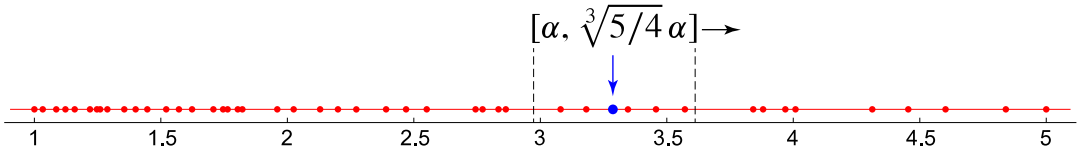

Fig. 1. An illustration of I, in which the dots represent the reals in I.

Proof. Let

$$I := \left\{5^a \cdot 7^b \cdot 11^c : a, b, c \text{ are integers in } [-4, 4]\right\} \cap [1, 5],$$

as depicted in Fig. 1. It can be verified that $1, 5 \in I$ and every two consecutive reals in I have a ratio within the range $(1, \sqrt[3]{5/4})$. Hence, for any real $\alpha \in [1, 5]$, the interval $[\alpha, \sqrt[3]{5/4}\alpha]$ contains at least one real number in I.

To find an m satisfying Eq. (1), it suffices to have an m such that

$$\sqrt[3]{4n} < m < \sqrt[3]{5n}.$$

As n is sufficiently large, we can find some $a \geq 6$ such that $\ell = 5^a \cdot 7^4 \cdot 11^4$ and

$$\sqrt[3]{4n}/\ell \in [1, 5].$$

Let $\alpha = \sqrt[3]{4n}/\ell$. Then $[\sqrt[3]{4n}/\ell, \sqrt[3]{5n}/\ell] = [\alpha, \sqrt[3]{5/4}\alpha]$. By the property of I, $[\alpha, \sqrt[3]{5/4}\alpha]$ contains a real of the form $5^a \cdot 7^b \cdot 11^c$ for some $a, b, c \geq -4$. We are done. □

The m mentioned in Lemma 2 can be obtained by finding the ℓ mentioned in the proof of Lemma 2 followed by checking whether ℓx is a proper m by scanning all reals in $x \in I$. ℓ can be found in $O(\log n)$ time by starting ℓ' with $5^6 \cdot 7^4 \cdot 11^4$ and iteratively multiplying ℓ' with 5 until ℓ' satisfies the need of ℓ. Since I contains only 47 real numbers, checking whether ℓx is a proper m can be done in $O(1)$ time.

Remark. One may wonder why not pick an $m = 5^2 \cdot p$ for some prime p. The reason is that the running time of the above is faster than using m in the form $m = 5^2 \cdot p$.

Here are details. It is shown in [4, p. 494] that, for every real $\varepsilon > 0$,

$$\lim_{n\to\infty} \frac{\pi((1+\varepsilon)n) - \pi(n)}{n/\log n} = \varepsilon.$$

Then, we sample an integer uniformly at random from the interval $[\sqrt[3]{4n}/25, \sqrt[3]{5n}/25]$ and, using Miller-Rabin's algorithm [10], test whether the sampled integer is a prime in $O(\log^2 n)$ time. To have the failure probability bounded by the claimed $1/n^{\Omega(1)}$, we require to sample $O(\log^2 n)$ integers in the interval. Hence, the total running time is $O(\log^4 n)$, exceeding our budget for the time complexity.

2.2 Computing x and y

In this section, we present how to compute x and y that satisfy Lemma 1 in $O(\log^2 n)$ time with failure probability bounded by $1/n^{\Omega(1)}$.

We extend Lemma 1 from [15] to Lemma 3, where Lemma 1 of [15] is equivalent to setting $r = 1$ in Lemma 3.

Lemma 3. *Let p be a prime number at least 5 and r be a positive integer. If $a \not\equiv 0 \pmod p$, the image set $\mathcal{I}_r$ of the congruence*

$$x^3 + ax \pmod{p^r}$$

contains exactly $p^{r-1}|\mathcal{I}_1|$ incongruent values where $|\mathcal{I}_1| = \lfloor (2p+1)/3 \rfloor$; that is,

$$|\mathcal{I}_r := \{x^3 + ax \pmod{p^r} : x \in [p^r]\}| = p^{r-1}|\mathcal{I}_1|.$$

Proof. We will use Hensel's Lemma [12, Theorem 4.15] to have a many-to-one mapping from the image set $\mathcal{I}_r$ to $\mathcal{I}_1$. Let $f_n(x) = x^3 + ax - n$. We study in the following claim whether there exists a solution v of $f_n(x) \equiv 0 \pmod{p}$ such that $f_n'(v) \not\equiv 0 \pmod{p}$.

Claim. For every $n \in \mathcal{I}_1$, there exists a value $v \in [p]$ such that

$$f_n'(v) \not\equiv 0 \pmod{p}.$$

Proof. We classify $n \in \mathcal{I}_1$ into the following three categories, where $\mathcal{I}_{1,t}$ for each $t \in [3]$ is the collection of all the values $n \in [p]$ such that the number of incongruent solutions to $f_n(x)$ is exactly t. Because p is a prime, $f_n(x) \equiv 0 \pmod{p}$ has at most three incongruent solutions. Hence, $\mathcal{I}_1 = \mathcal{I}_{1,1} \cup \mathcal{I}_{1,2} \cup \mathcal{I}_{1,3}$.

- If $n \in \mathcal{I}_{1,1}$, then for some $v_0 \in [p]$ we have

  $$f_n(x) \equiv (x - v_0)g(x) \pmod{p}$$

  There are two subcases to discuss:
  - Case I: $g(v_0) \equiv 0 \pmod{p}$. In this case, $f_n(x) \equiv (x - v_0)^2(x - t) \pmod{p}$. Since $f_n(x)$ does not have the quadratic term, $-t - 2v_0 \equiv 0 \pmod{p}$. Since $a \not\equiv 0 \pmod{p}$, $v_0 \not\equiv 0 \pmod{p}$. Consequently, $t \not\equiv v_0 \pmod{p}$. Thus, we have $n \in \mathcal{I}_2$, a contradiction. Such a subcase cannot happen.
  - Case II: $g(v_0) \not\equiv 0 \pmod{p}$. Hence, $f_n'(v_0) \equiv g(v_0) + (v_0 - v_0)g'(v_0) \not\equiv 0 \pmod{p}$, as desired.
- If $n \in \mathcal{I}_{1,2}$, then for some $v_0 \not\equiv v_1 \pmod{p}$ we have

  $$f_n(x) \equiv (x - v_0)^2(x - v_1).$$

  Since $f_n(x)$ does not have the quadratic term, $-2v_0 \equiv v_1 \pmod{p}$. Since $a \not\equiv 0 \pmod{p}$, $v_0 \not\equiv 0 \pmod{p}$. Thus, $v_0 \not\equiv -v_1 \pmod{p}$. The incongruent solutions to $f_n'(x) \equiv 0$, if they exist, must be v' and $-v'$ for some $v' \in [p]$. Since $v_0 \not\equiv -v_1$, at least one of them is not a solution to $f_n'(x)$, as desired.
- If $n \in \mathcal{I}_{1,3}$, then for some incongruent values $v_0, v_1, v_2 \in [p]$ we have

  $$f_n(x) \equiv (x - v_0)(x - v_1)(x - v_2).$$

  Since $f_n'(x) \equiv 0 \pmod{p}$ has at most two incongruent solutions, at least one value $v' \in \{v_0, v_1, v_2\}$ satisfies $f_n'(v') \not\equiv 0 \pmod{p}$, as desired. ∎

The above claim ensures that, for every $n \in \mathcal{I}_1$, there exists $v \in [p]$ such that

$$\begin{cases} f_n(v) \equiv 0 \pmod{p} \\ f_n'(v) \not\equiv 0 \pmod{p} \end{cases}$$

By Hensel's Lemma [12, Theorem 4.15], there exists an unique $t \in [p]$ such that $v+tp$ is a solution to $f_n(x) \equiv 0 \pmod{p^2}$. Since $f_n'(v+tp) \equiv f_n'(v) \pmod{p}$. Inductively, we get $n \in \mathcal{I}_r$ as well. Conversely, for every $n' \in \mathcal{I}_r$, $(n' \bmod p) \in \mathcal{I}_1$.

For each $n' \in [p^r]$, its analysis is the same as $n' \bmod p$. Thus, $|\mathcal{I}_r| = |\mathcal{I}_1|p^{r-1}$. In [15, Lemma 1], $|\mathcal{I}_1| = \lfloor (2p+1)/3 \rfloor$ is shown. This completes the proof. □

By applying the Pigeonhole Principle and Chinese Remainder Theorem, we derive the following lemma:

Lemma 4. *Let $m = p_1^{r_1} p_2^{r_2} \cdots p_t^{r_t}$ where p_i are distinct prime numbers at least 5 and r_i are positive integers for all $i \in [t]$. If $(m, a) = 1$, for every $n \in [m]$ the congruence*

$$x^3 + ax + y^3 + ay \equiv n \pmod{m}$$

is soluble by setting $x^3 + ax \equiv n' \pmod{m}$ for every $n' \in \mathcal{I}$ where $\mathcal{I} \subset [m]$ and

$$|\mathcal{I}| \geq \prod_{i \in [t]} \left(2\lfloor (2p_i + 1)/3 \rfloor - p_i\right) p_i^{r_i - 1}.$$

Proof. By Lemma 3 (where we require $a \not\equiv 0 \pmod{p_i^{r_i}}$), for each $i \in [t]$, the image sets of $x^3 + ax \pmod{p_i^{r_i}}$ and $y^3 + ay \pmod{p_i^{r_i}}$ both contain $\lfloor (2p_i + 1)/3 \rfloor p_i^{r_i - 1}$ incongruent values. By the Pigeonhole Principle, for each $i \in [t]$, there exist at least

$$\left(2\lfloor (2p_i + 1)/3 \rfloor - p_i\right) p_i^{r_i - 1} \tag{3}$$

incongruent values s_i in the image set of $x^3 + ax \pmod{p_i^{r_i}}$ such that $y^3 + ay \equiv n - s_i \pmod{p_i^{r_i}}$ is soluble.

For each $i \in [t]$, let α_i denote a solution to $x^3 + ax \equiv s_i \pmod{p_i^{r_i}}$, and let β_i denote a solution to $y^3 + ay \equiv n - s_i \pmod{p_i^{r_i}}$. Since p_i for all $i \in [t]$ are distinct primes, by the Chinese Remainder Theorem for each $2t$-tuple $(\alpha_1, \alpha_2, \ldots, \alpha_t, \beta_1, \beta_2, \ldots, \beta_t))$, we can find a unique (α, β) such that for all $i \in [t]$

$$\alpha \equiv \alpha_i \text{ and } \beta \equiv \beta_i \pmod{p_i^{r_i}}.$$

Thus, for any choices of s_i for $i \in [t]$, we can find a 2-tuple (α, β) as a solution to $x^3 + ax + y^3 + ay \equiv n \pmod{m}$.

Note that, for each $i \in [t]$, the number of choices of s_i is lower-bounded by Eq. (3). Each possible t-tuple $(s_1, s_2, \ldots, s_t)$ determines a unique $s^\dagger \pmod{m}$ such that

$$s^\dagger \equiv s_i \pmod{p_i^{r_i}} \text{ for } i \in [t].$$

Let $\mathcal{I}$ be the collection of all such $s^\dagger$s. Thus $\mathcal{I}$ has size at least the product of Eq. (3) for all $i \in [t]$. We are done. □

Thus, by Lemma 4, we can sample an $n' \in \mathcal{I}$ with success probability $> (1/3)^3 - \varepsilon$ for some small constant $\varepsilon > 0$ because the m we picked in Lemma 2 has at most three prime factors. Given n', to check whether it is valid it suffices to find a solution to $x^3 + ax \equiv n' \pmod{p}$ and $y^3 + ay \equiv n - n' \pmod{p}$ for each $p \in \{5, 7, 11\}$. The reason is stated in the proof of Lemma 4. As each $p \in \{5, 7, 11\}$ is a constant, checking the validity of n' takes $O(1)$ time for each random guess and $O(\log n)$ time to succeed with probability $1 - 1/n^{\Omega(1)}$.

Given a valid n', we need to find a solution to $x^3 + ax \equiv n' \pmod{m}$ and another to $y^3 + ay \equiv n - n' \pmod{m}$. To accomplish this, we solve $x^3 + ax \equiv n' \pmod{p}$ and

$y^3 + ay \equiv n - n' \pmod{p}$ for each $p \in \{5, 7, 11\}$, the set of all prime factors of the picked m. For each p, this can be done in $O(1)$ time, as $p = O(1)$. Then, we use Hensel's Lemma [12, Theorem 4.15] to lift the roots to the right powers of p. The total number of times to lift the powers is $O(\log n)$ and each takes $O(1)$ time, as $p = O(1)$. Finally, we use an efficient quadratic-time algorithm for Chinese Remainder Theorem [14] to combine the found solutions to $x^3 + ax \equiv n' \pmod{p^r}$ and $y^3 + ay \equiv n - n' \pmod{p^r}$ for each $p \in \{5, 7, 11\}$, for some integer $r \geq 1$. The final step takes $O(\log^2 n)$ time, dominating the total running time. To get $0 \leq x, y < 3m/5$, we require a fix by the following lemma Lemma 5.

Lemma 5. *Let $m = p_1^{r_1} p_2^{r_2} \cdots p_t^{r_t}$ where p_i are distinct prime numbers at least 5 and r_i are positive integers for all $i \in [t]$. If $(m, a) = 1$ and $r_1 \geq 2$, given a solution (x_0, y_0) to the congruence*

$$x^3 + ax + y^3 + ay \equiv n \pmod{m/p_1},$$

then in $O(p_1)$ time one can obtain a solution (x_1, y_1) to the congruence

$$x^3 + ax + y^3 + ay \equiv n \pmod{m}$$

such that $0 \leq x_1, y_1 < 3m/5$.

Proof. Let $x_1 = x_0 + u(m/p_1)$ and $y_1 = y_0 + v(m/p_1)$ for some u, v to be determined. To let (x_1, y_1) satisfy the congruence

$$x^3 + ax + y^3 + ay \equiv n \pmod{m},$$

it suffices to require that

$$(3x_0^2 + a)u + (3y_0^2 + a)v \equiv 0 \pmod{p_1}. \tag{4}$$

As stated in the proof of Lemma 3, x_0 and y_0 are picked so that $3x_0^2 + a \not\equiv 0 \pmod{p_1}$ and $3y_0^2 + a \not\equiv 0 \pmod{p_1}$. Thus, we can rewrite Eq. (4) as

$$u + \lambda v \equiv 0 \text{ for some } \lambda \not\equiv 0 \pmod{p_1}. \tag{5}$$

Thus, by setting u as a value u_0 in $[p]$, there is a unique v_0 such that (u_0, v_0) is a solution to Eq. (5). Hence, by the Pigeonhole Principle, there exists a value u_0 in $[0, (p_1 - 1)/2]$, by setting u as u_0 the corresponding v_0 also in $[0, (p_1 - 1)/2]$. Consequently, $x_1 = x_0 + u_0(m/p_1) < (u_0 + 1)(m/p_1) \leq m(p_1 + 1)/(2p_1) \leq 3m/5$. The last inequality holds because $p_1 \geq 5$.

To find (u_0, v_0) that satisfies Eq. (5) and $u_0, v_0 \in [0, (p_1 - 1)/2]$, one can enumerate all $u_0 \in [p]$, which takes $O(p_1)$ time. □

2.3 Representing $8k + 3$ as a sum of three squares

In this section, we present how to express $8k + 3$ for any $k \geq 1$ as a sum of three squares in $O(\log^3 n / \log\log n)$ time with failure probability bounded by $1/n^{\Omega(1)}$.

In the literature, there are two algorithms [11,8] for this task in $O(\text{polylog } n)$ time. We opt to use the algorithm from [8] for this decomposition task, as it relies only on the Extended Riemann Hypothesis. Although its expected running time is $O(\log^2 / \log\log n)$, to ensure a successful outcome with probability $1 - 1/n^{\Omega}(1)$, the running time increases to $O(\log^3 / \log\log n)$, as desired.

3 Expressing Integers as Sums of the Fewest Possible Tetrahedral Numbers

We will prove Theorem 2 in this section by devising a linear-time algorithm for that, given an integer $\ell \geq 1$, for each integer $n \in [1, \ell]$, determine the least integer k such that n is a k-number. Then we report some empirical results obtained from this algorithm, which yields Theorem 3.

Our linear-time algorithm relies on the Pollock's conjecture and the following unproven Conjecture 1. However, its validity over integers in $[1, \ell]$ can be empirically verified in $O(\ell)$ time with probability $1 - 1/\ell^{\Omega(1)}$. See Section 3.3 for details. This conjecture is motivated by an observation in [1] that almost all 4-numbers up to 4×10^{10} have value $r \leq 30$, and, more completely, our observation that the distribution of the r values of all 4-numbers up to 10^{12}, as depicted in Fig. 2. Besides, this conjecture holds for all the ranges in our conducted experiments, as shown in Table 1.

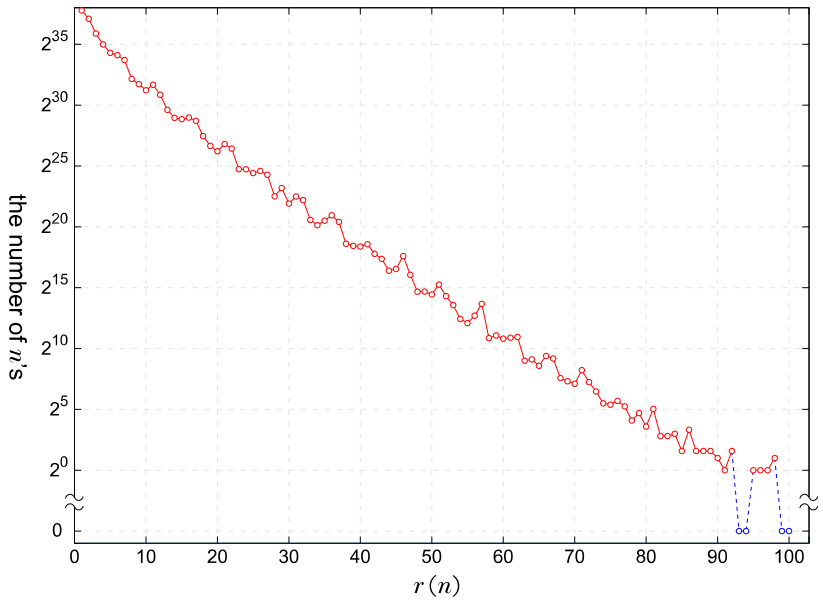

Fig. 2. The distributions of r values for all 4-numbers up to 10^{12}.

Conjecture 1. The number of the 4-numbers n in $[1, \ell]$ whose $r(n) = O(\log n)$ is at least $\ell - \ell^{1/3}$.

3.1 A Simple Linear-Time Algorithm Based on Conjecture 1 and the Pollock's conjecture

Our simple linear-time algorithm works as follows. To determine the least integer k for all $n \in [1, \ell]$ such that n is a k-number, we need only the first $O(\ell^{1/3})$ 1-numbers

	# 1-numbers	# 2-numbers	# 3-numbers	# 5-numbers	# n with $r(n) > 68$
2T	22893	231360135	892533269700	241	983
4T	5950	135938272	892566316867	0	742
6T	4175	114024249	892579370745	0	733
8T	3323	101777580	892586662215	0	725
10T	2806	93543895	892594264156	0	768
12T	2453	87465970	892597268218	0	743
14T	2194	82715101	892600240260	0	763
16T	1993	78854043	892605925192	0	678
18T	1834	75627012	892607810496	0	694
20T	1702	72871243	892611249359	0	684
22T	1592	70477495	892611703794	0	699
24T	1498	68370546	892613608688	0	718
26T	1418	66496077	892615100836	0	685
28T	1346	64810643	892616483208	0	704
30T	1284	63284787	892619064152	0	711
32T	1227	61891868	892620842289	0	705
34T	1178	60616223	892621019193	0	670
36T	1133	59437585	892621494377	0	671
38T	1091	58348054	892623074380	0	660
40T	1053	57331847	892625811511	0	733
42T	1019	56383995	892624729147	0	647
44T	987	55495820	892626082051	0	645
46T	958	54661584	892627506471	0	668
48T	930	53874411	892629429638	0	696
50T	905	53131427	892628890795	0	729
52T	881	52427429	892629999069	0	704
54T	858	51758363	892629240944	0	656
56T	838	51124557	892630608375	0	673
58T	818	50519244	892631146215	0	669
60T	799	49941156	892631761094	0	666
62T	782	49390221	892633870857	0	656
64T	765	48861174	892633446685	0	702
66T	750	48355792	892634312041	0	650
68T	734	47868742	892634383912	0	691
70T	720	47402034	892635181693	0	681
72T	707	46953191	892636150470	0	648
74T	693	46519325	892636792060	0	732
76T	682	46103203	892636618322	0	777
78T	669	45700380	892637949388	0	703
80T	658	45311135	892636248600	0	682
82T	647	44934977	892637894497	0	704
84T	637	44571423	892638390967	0	663
86T	626	44218320	892638667205	0	641
88T	617	43877585	892640243280	0	673
90T	608	43546706	892639526727	0	704
92T	599	43225186	892639595445	0	707
94T	590	42912932	892640997740	0	680
96T	582	42609741	892641631626	0	699
98T	574	42314723	892641152283	0	683
100T	566	42027949	892642642346	0	715

Table 1. The distribution of k-numbers for all positive integers up to 10^{14}. The unreported numbers are 4-numbers.

because any 1-number appears later is greater than ℓ. To obtained the first $O(\ell^{1/3})$ numbers, one can use the formula of tetrahedral numbers, which takes $O(\ell^{1/3})$ time. Then, the computation of all 2-numbers and 3-numbers in $[1, \ell]$ can be done by enumerating the sum of any three of the first $O(\ell^{1/3})$ 1-numbers, which can be done in $O(\ell)$ time.

For the computation of all the 4-numbers in $[1, \ell]$, we proceed with the following two steps:

- Step 1. Let A be the indicator bit-array of all the k-numbers in $[1, \ell]$ for all $k \leq 3$. Compute the logical-OR of A with shifts by δ for all $\delta \in \{0\} \cup \{T_i : i \in O(\log \ell)\}$. The logical OR of these $O(\log \ell)$-many $O(\ell)$-bit arrays can be accomplished in $O(\ell)$ time because the model of computation is WordRAM where basic arithmetic operations (including logical OR and shift) over $O(\log \ell)$-bit operands take $O(1)$ time. Let B be the result of the logical OR. By definition, B is an indicator bit-array of all the 3-numbers in $[1, \ell]$ and all the 4-numbers in $[1, \ell]$ with $r = O(\log \ell)$.
- Step 2. By Conjecture 1, we know that B has $O(\ell^{1/3})$ 0-bits, each corresponding to a 4-number with $r > \Delta$ for some $\Delta = O(\log \ell)$ or a k-number for some $k \geq 5$. There are $O(\ell^{1/3})$ such numbers, and for each of them we check whether it can be expressed as a sum of two 2-numbers. Given the sorted list of 2-numbers, determining whether $k < 5$ can be done in $O(\ell^{1/3}\ell^{2/3}) = O(\ell)$ time by the standard 2-sum algorithm. See Young tableau [2], for example. Finally, given the sorted list of 2-numbers and A, determining whether $k < 6$ can be done in $O(\ell^{1/3}\ell^{2/3}) = O(\ell)$ time by checking if the tested number subtracted by any 2-number is a 3-number. If any number passes all of the above tests, then it is a k-number for some $k \geq 6$. This refutes the Pollock's conjecture.

To sum up, we have an $O(\ell)$-time algorithm as claimed in Theorem 2, assuming Conjecture 1 and the Pollock's conjecture.

3.2 Reducing the Space from $O(\ell)$ to $O(\ell^{2/3})$

The number of all 2-numbers in $[1, \ell]$ is $O(\ell^{2/3})$ and the length of the indicator bit-array A is $O(\ell)$. Our approach is to divide A into subintervals, each of length $O(\ell^{2/3})$. Then the algorithm Section 3 operates on A subinterval by subinterval. Constructing A subinterval by subinterval does not incur more running time than constructing A entirely at once. Because A can be constructed by enumerating all x in the list of 1-numbers in $[1, \ell]$ and all y in the list of 2-numbers in $[1, \ell]$ and summing each pair of x and y. While constructing A subinterval by subinterval, for each x, we use binary search to locate the lower- and upper-bounds of y in the list of 2-numbers in $[1, \ell]$ such that $x+y$ falls within the current subinterval of A. Therefore, each pair of x and y joins at most one calculation of one subinterval. Therefore, the total running time is $O(\ell^{1/3} \log \ell + \ell) = O(\ell)$.

3.3 Empirical Verification of Conjecture 1

The task is to verify Conjecture 1 over integers in $[1, \ell]$ in $O(\ell)$ time. Let $\Delta = O(\log \ell)$ be an integer. Let X be a random variable indicating that

$$X = \begin{cases} 1 & \text{if an integer } x \text{ sampled uniformly at random from } [1, \ell] \text{ is a 4-number with } r(x) > \Delta \\ 0 & \text{otherwise} \end{cases}$$

Thus, it suffices to show that

$$E[X] \leq \ell^{-2/3}.$$

Let S be the sum variable of $\Omega(\ell^{2/3} \log \ell)$ independent copies of X. By Chernoff bound, we have

$$\Pr\{|S - E[S]| \geq \varepsilon E[S]\} = e^{\Omega(\varepsilon^2 E[S]/(2+\varepsilon))}.$$

Thus, if $E[X] = \Omega(\ell^{-2/3})$, Chernoff bound yields a constant approximation for $E[X]$ with probability $1 - 1/\ell^{\Omega(1)}$. Otherwise $E[X] = \ell^{-2/3}/k$ for some $k = \omega(1)$, setting $\varepsilon E[S] = \log \ell$ also suffices to tell whether $E[X] < \ell^{-2/3}$ with probability $1 - 1/\ell^{\Omega(1)}$.

As a result, one can use $O(\ell^{2/3} \log \ell)$ independent copies of X to verify Conjecture 1. Evaluating a copy of X takes $O(\Delta)$ time if A is given, which takes $O(\ell)$ time to prepare.

3.4 Verification of the Pollock's Conjecture Up to 10^{21}

Given our result in Table 1, we know that there are only 241 5-numbers in $[1, 10^{14}]$. Thus, for every two consecutive tetrahedral numbers T_i and T_{i+1} whose difference is at most 10^{14}, one can conclude that all numbers between T_i and T_{i+1} are k-numbers for some $k \leq 5$ except for $T_i + x$ for all x in the set of 241 5-numbers. However, these exceptional numbers all fall within $[T_{i-1}, T_{i-1}+10^{14}]$ (but not exceptional numbers w.r.t. T_{i-1}) for each sufficiently large i. This yields a proof of Theorem 3. This technique is introduced in [1].

4 Conclusion

In this paper, we devise two algorithms to decompose integers as sums of few tetrahedral numbers. Our results leave some interesting directions to further explore.

The algorithm in Section 2 is the first one for decomposing integers into sums of at most eight tetrahedral numbers in time polynomial in the input size, assuming the extended Riemann hypothesis. This assumption can be removed if the decomposition of integers into three squares, a building block of this algorithm, can be done efficiently without assuming ERH. A natural question to ask is whether ERH is truly necessary for this task.

Our second algorithm relies on an assumption on the distribution of r values. We also wonder whether the correctness of assumption can be rigorously proved.

Acknowledgments

We want to thank Dr. Pangfeng Liu for introducing us to the Pollock's conjecture on tetrahedral numbers.

References

1. Chou, C., Deng, Y.: Decomposing 40 billion integers by four tetrahedral numbers. Math. Comput. **66**(218), 893–901 (1997)
2. Cormen, T.H., Leiserson, C.E., Rivest, R.L., Stein, C.: Introduction to Algorithms, 3rd Edition. MIT Press (2009)
3. Deng, Y.F., Yang, C.N.: Waring's problem for pyramidal numbers. Science in China (Scientia Sinica) Series A **37**(3), 277–283 (1994)
4. Hardy, G.H., Wright, E.M.: An introduction to the theory of numbers (6th ed.). Oxford University Press (2008)
5. James, R.D.: The representation of integers as sums of pyramidal numbers. Mathematische Annalen **109**, 196–199 (1934)
6. Landau, E.: Über eine anwendung der primzahltheorie auf das waringsche problem in der elementaren zahlentheorie. Math. Ann. **66**, 102–105 (1908)
7. Legendre, A.M.: Essai sur la Théorie des Nombres. Cambridge Library Collection - Mathematics, Cambridge University Press, 2 edn. (2009)
8. Pollack, P., Schorn, P.: Dirichlet's proof of the three-square theorem: An algorithmic perspective. Math. Comput. **88**(316), 1007–1019 (2019)
9. Pollock, J.F.: On the extension of the principle of Fermat's theorem of the polygonal numbers to the higher orders of series whose ultimate differences are constant. with a new theorem proposed, applicable to all the orders. Proc. Roy Soc. London **5**, 922–924 (1851)
10. Rabin, M.O.: Probabilistic algorithm for testing primality. Journal of Number Theory **12**(1), 128–138 (1980)
11. Rabin, M.O., Shallit, J.O.: Randomized algorithms in number theory. Comm. Pure Appl. Math. **39**, 239–256 (1986)
12. Rosen, K.: Elementary Number Theory and Its Applications. Pearson (2011)
13. Salzer, H.E., Levine, N.: Table of integers not exceeding 1,000,000 that are not expressible as the sum of four tetrahedral numbers. Math. Comput. **12**, 141–144 (1958)
14. Selianinau, M.: Computationally efficient approach to implementation of the chinese remainder theorem algorithm in minimally redundant residue number system. Theory of Computing Systems **65**, 1117–1140 (2021)
15. Watson, G.L.: Sums of eight values of a cubic polynomial. Journal of the London Mathematical Society **s1-27**(2), 217–224 (04 1952)
16. Yang, K.C.: Various generalization of Waring's problem. thesis, chicago university (1928)

A More Tables

n	$r(n)$	n	$r(n)$)	n	$r(n)$)
837293	96	14780388803	69	30404017737	81
1751787	81	16212264602	81	31249750702	71
468164933	69	16490211457	71	31391571118	84
725334878	70	18131378533	71	31413964447	71
726409283	69	18685629958	75	31720955863	69
1872385653	79	21042275158	69	31820242053	72
1999255043	74	21282661867	71	32783549982	76
2390056433	72	21400673177	71	33723523927	78
3281447262	70	21909670998	76	33792979677	73
4269476262	70	22491659283	69	34001374367	70
5631140023	72	23084057267	69	34085489982	72
6240760377	70	23547651727	71	34118155573	71
6349723882	70	24041633103	69	34492104467	73
8798752737	71	24218940783	81	35040177258	71
9074616777	76	25902476693	69	35248597663	69
9518151603	71	26867883863	69	35780935678	72
10873443037	69	27429950322	70	37306165122	71
11332666483	80	28488454738	70	38970420417	71
12115559363	71	28954173022	71	38978879802	98
12636741707	72	29020094637	70	39282652547	74
12709829722	71	29375497387	71	39389678018	75
13646925297	73	30106439547	81	39981386443	69
14533010753	75	30224854987	69		

Table 2. Errata.