3-smooth representations from short intervals

Wouter van Doorn and Anneroos Everts

Abstract
Solving a problem by Erdős, we prove that every positive integer n can be written as a sum

$$n = t_{i_1} + t_{i_2} + \ldots + t_{i_l}$$

of distinct 3-smooth integers with $1 \leq t_{i_1} < t_{i_2} < \ldots < t_{i_l} < 6t_{i_1}$.

1 Introduction

Let $(t_1, t_2, \ldots)$ be the infinite increasing sequence of 3-smooth integers. That is, for every index i there are non-negative integers x_i, y_i for which $t_i = 2^{x_i}3^{y_i}$. In the early 1990s, Erdős conjectured that every positive integer n can be written as a sum of distinct t_i such that no summand divides another. As was quickly realized however (even before it was written down anywhere in the literature, for the first time in [1]), this conjecture actually has a very short induction proof. Indeed, one can make the stronger induction hypothesis that, for all even n, all summands are even as well. For even n we are then done by applying the hypothesis to $\frac{n}{2}$, while for odd n one can apply the induction hypothesis to $n - 3^{\lfloor \log_3 n \rfloor}$.

More generally, sequences such that every large enough integer can be written as a sum of distinct elements where no summand divides another, are called d-complete sequences. They have been studied by Erdős-Lewin [2], Ma-Chen [3], and Chen-Wu [4], mostly in the context of proving, for various fixed p, q, r, that the sequence of integers of the form $p^x q^y r^z$ is d-complete.

Going back to 3-smooth integers, Blecksmith, McCallum and Selfridge show in [5] that the subsequence $(t_{k+1}, t_{k+2}, \ldots)$ is d-complete as well for all k, proving the final conjecture from [2]. Their proof actually provides a fully explicit procedure to write an integer n as a sum of distinct $t_i \geq t_k$, and some variations of this procedure will feature prominently in this paper.

Note that no two summands t_i, t_j can divide one another if $t_i < t_j < 2t_i$. Therefore, a stronger conjecture in this regard would be whether, for some $C \leq 2$, every positive integer n can be written as a sum $n = t_{i_1} + t_{i_2} + \ldots + t_{i_l}$ with $1 \leq t_{i_1} < t_{i_2} < \ldots < t_{i_l} < Ct_{i_1}$. This stronger conjecture turns out to be false however (although it plausibly does hold if we consider 5-smooth integers instead), as was already remarked in [2]. A natural follow-up question is then if a constant $C > 2$ exists for which such a representation always is possible, even though we are by now leaving the realm of d-completeness.

The possible existence of such a C was first considered in [1] where Erdős initially thought that 'surely almost all integers cannot be written in this form'. This question was then repeated, much more neutrally, in [2], and it is now listed as Problem 845 at Bloom's website [8]. There, in the comment section, Cambie suggested that such a constant actually does exist, and checked with a computer that $C = \frac{32}{9} = 3.55\cdots$ works for all $n \le 10^5$.

As we will see shortly, by slightly tweaking the procedure from [5], we are able to show that one can take $C = 6$ for all $n \in \mathbb{N}$. On the other hand, for any $C < 3$ there are infinitely many n for which C is not admissible.

2 Main result

Recalling that $(t_1, t_2, \ldots)$ is the sequence of 3-smooth integers, the goal of this section is to prove the following theorem.

Theorem 1. *Every positive integer n can be written as a sum*

$$n = t_{i_1} + t_{i_2} + \ldots + t_{i_l}$$

with $1 \le t_{i_1} < t_{i_2} < \ldots < t_{i_l} < 6t_{i_1}$. Moreover, in this final inequality one cannot replace 6 by any constant $C < 3$.

Proof. We will first prove the latter claim. So let us assume $C < 3$, choose $\delta > 0$ and $\epsilon > 0$ sufficiently small so that $C(1+\epsilon) < 3-\delta$, and let N be large enough. For ease of reference, let us call a sum $t_{i_1} + t_{i_2} + \ldots + t_{i_l}$ of distinct 3-smooth integers with $t_{i_1} < t_{i_2} < \ldots < t_{i_l} < Ct_{i_1}$ *short*. The goal is to show that the number of short sums with all elements smaller than or equal to N, is smaller than N.

For a non-negative integer j, let x_j be equal to $(1+\epsilon)^j$, and define the interval $I_j = [x_j, x_{j+1})$. For any short sum with $t_{i_1} \in I_j$, we then get

$$\begin{aligned} x_j &\le t_{i_1} \\ &< t_{i_l} \\ &< Ct_{i_1} \\ &< Cx_{j+1} \\ &= C(1+\epsilon)x_j \\ &< (3-\delta)x_j. \end{aligned}$$

In particular, with A_j the number of 3-smooth integers contained in the interval $[x_j, (3-\delta)x_j)$, the number of short sums with $t_{i_1} \in I_j$ is at most 2^{A_j}. From the discussion in Lecture 5 from [6], we moreover have the following lemma bounding A_j.

Lemma 1. *There exists an absolute constant c such that $A_j < \frac{\log x_j \log(3-\delta)}{\log 2 \log 3} + c$ for all $j \ge 0$.*

Applying Lemma 1 and using the fact that every subset sum with all elements smaller than or equal to N must have $t_{i_1} \in I_j$ for some $j \leq L := \lfloor \frac{\log N}{\log(1+\epsilon)} \rfloor$, the total number of short sums at most N is upper bounded by

$$\begin{aligned}\sum_{j=0}^{L} 2^{A_j} &< 2^c \sum_{j=0}^{L} x_j^{\frac{\log(3-\delta)}{\log 3}} \\ &\leq 2^c (L+1) x_L^{\frac{\log(3-\delta)}{\log 3}} \\ &\leq 2^c (L+1) N^{\frac{\log(3-\delta)}{\log 3}}.\end{aligned}$$

Since this latter quantity is significantly smaller than N when N is large enough, we conclude that almost all positive integers cannot be represented as a short sum.

We may now focus on the other direction and prove that for $C = 6$ such representations do exist for all n. Hence, let n be any arbitrary, but fixed, positive integer. Following in the footsteps of [5], we define k and s by

$$s = t_1 + t_2 + \ldots + t_k \leq \frac{n}{2} < t_1 + t_2 + \ldots + t_{k+1},$$

and let r be equal to $n - 2s$. It then follows that $r < 2t_{k+1}$, and we write $r = t_{j_1} + t_{j_2} + \ldots + t_{j_{l'}}$ in binary, i.e. as a sum of distinct powers of two. We further define u and v to be the indices for which $t_u = 3t_k$ and $t_v = 6t_k$ respectively. By adding the representations of $2s$ and r we then obtain the following representation of n:

$$n = c_1 t_1 + c_2 t_2 + \ldots + c_k t_k + c_{k+1} t_{k+1} + \ldots + c_v t_v. \tag{1}$$

Here, $c_i \in \{2, 3\}$ for $1 \leq i \leq k$. For $i > k$ we have $c_i \in \{0, 1\}$ with $c_i = 1$ if, and only if, t_i occurred as a power of two in the binary representation of r. By the inequality $r < 2t_{k+1}$, one can check that $c_i = 0$ for all i with $u < i \leq v$.

Using a variation on the procedure laid out as 'the midgame' in [5], we are going to transform the representation from equation (1) into a different representation of n in such a way that, at the end, $c_i \in \{0, 1\}$ for all i, with $c_i = 0$ unless $k < i \leq v$. Since $t_v = 6t_k$, this would finish the proof.

The transformation procedure consists of at most u steps, where in step i we consider c_i. If $c_i \in \{0, 1\}$ (which we recall does not happen for $i \leq k$), we do not do anything and we simply go to step $i + 1$. On the other hand, if $c_i > 1$, define for $m \in \{2, 3\}$ the index j_m such that $t_{j_m} = mt_i$. Now let us first assume $i \leq k$. In that case, if $c_i = m \in \{2, 3\}$, then we lower c_i to 0 and increase c_{j_m} by 1. If $c_i = 4$, we lower c_i to 0 and increase c_{j_2} by 2. And if $c_i = 5$, we lower c_i to 0 and increase both c_{j_2} and c_{j_3} by 1. If, however, $i > k$, then we lower c_i by 2 and increase c_{j_2} by 1. Following these transformations we claim the following result.

Lemma 2. *Throughout the entire procedure we have*

$$c_i \le \begin{cases} 5 & \text{if } 1 \le i \le k \\ 3 & \text{if } k < i \le u \\ 1 & \text{if } u < i \le v \\ 0 & \text{if } i > v \end{cases}$$

Let us point out that Theorem 1 already follows from Lemma 2. Indeed, by the transformation procedure outlined above, $c_i \in \{2, 3, 4, 5\}$ gets reduced to 0 for $i \le k$, while $c_i \in \{2, 3\}$ gets reduced by 2 for $k < i \le u$. In particular, after u steps we have $c_i \in \{0, 1\}$ for all i, with $c_i = 0$ unless $k < i \le v$.

Proof of Lemma 2. When $1 \le i \le u$, we claim that the coefficient c_i can increase by at most 3. And in the special case where t_i is a power of two, we instead claim that the coefficient c_i can increase by at most 2. To see this, all one has to do is use the fact that the only elements that can be responsible for increasing c_i are those corresponding to $\frac{t_i}{2}$ (which increases c_i by at most two) and $\frac{t_i}{3}$ (which is impossible if t_i is a power of two).

Now, before the first step we had, for t_i not a power of two, $c_i \le 2$ for $i \le k$, and $c_i = 0$ for $k < i \le u$. On the other hand, if t_i is a power of two, then we instead had the inequalities $c_i \le 3$ for $i \le k$, and $c_i \le 1$ for $k < i \le u$. In either case the first two inequalities from Lemma 2 follow by the argument from the previous paragraph.

As for the third inequality, if $u < i \le v$ then we note that $\frac{t_i}{3} > t_k$, so that the only way to possibly increase c_i is through $\frac{t_i}{2}$, implying $c_i \le 1$. And this immediately implies the final inequality as well, as for $i > v$ there is no way for c_i to increase beyond its initial value of 0 at all. □

3 A general procedure for multiplicatively closed complete sequences

Let $A = (a_1, a_2, \ldots)$ be a sequence of positive integers with $1 < a_1 < a_2 < \ldots$ which is closed under multiplication, and such that, for some $n_0 \ge 2$, every $n \ge n_0$ can be written as a sum of distinct elements of A. Similarly to what we have seen before, define $k(n)$ to be the largest integer for which

$$a_1 + a_2 + \ldots + a_{k(n)} \le \frac{n}{n_0} - 1.$$

We can then generalize Theorem 2 from [5] in the following way, which is perhaps of independent interest.

Theorem 2. *Every integer $n \geq n_0$ can be written as a sum*

$$n = d_1 a_{i_1} + d_2 a_{i_2} + \ldots + d_l a_{i_l}$$

with $a_{k(n)} < a_{i_1} < \ldots < a_{i_l}$ and $d_i < n_0$ for all i.

Note that we lose the property that the elements are distinct. On the other hand, every coefficient is smaller than n_0, which is a bound independent of n. Moreover, what we gain is that all elements are larger than $a_{k(n)}$, so that, in particular, they are all contained in the interval $(a_{k(n)}, n]$. As for the value of $a_{k(n)}$, one can check that for every sequence for which the conditions of Theorem 2 are satisfied, there are positive constants C_1 and C_2 such that the inequalities $C_1\sqrt{n} < a_{k(n)} < \frac{C_2 n}{\log n}$ hold for all large enough n.

Proof of Theorem 2. In this setup we define $s = a_1 + a_2 + \ldots + a_{k(n)}$ and $r = n - n_0 s$. By noting that $r \geq n_0$, we may write $r = a'_1 + a'_2 + \ldots + a'_{l'}$ as a sum of distinct elements of A. We then add the representations of $n_0 s$ and r to get

$$n = c_1 a_1 + c_2 a_2 + \ldots + c_{k(n)} a_{k(n)} + c_{k(n)+1} a_{k(n)+1} + \ldots + c_m a_m. \tag{2}$$

Here, $c_i \geq n_0$ for all i with $1 \leq i \leq k(n)$ and m is the largest index with $a_m \leq n$. We then start an m step procedure, very similar to what we have seen before.

In step i, we consider c_i again. If $c_i < n_0$, simply continue on to the next step. On the other hand, if $c_i \geq n_0$ (which, again, happens for the first $k(n)$ values of i), we may write $c_i = a_{i,1} + a_{i,2} + \ldots + a_{i,L}$ as a sum of distinct elements of A. The term $c_i a_i$ in equation (2) then equals $(a_{i,1} + a_{i,2} + \ldots + a_{i,L})a_i$. If we now expand the brackets, we see that all of these terms $a_{i,j}a_i$ are equal to an $a_{i'}$ with $i' > i$, because A is assumed to be multiplicatively closed. Increasing the corresponding coefficient $c_{i'}$ in equation (2) by 1 for all $1 \leq j \leq L$, while reducing c_i to 0, finishes step i.

After m steps we find that all coefficients are smaller than n_0, with a coefficient c_i non-zero precisely when, at the start of step i, we had $1 \leq c_i < n_0$. In particular, c_i ends up being 0 for all $i \leq k(n)$ and the theorem is proven. □

As one application, Birch [7] proved that for any two coprime integers p and q, the sequence of integers of the form $p^x q^y$ with $x, y \in \mathbb{N}$ is complete. As it is also multiplicatively closed, Theorem 2 may indeed be applied to it.

References

[1] P. Erdős, *Some of my favourite problems in various branches of combinatorics.* Le Matematiche, Volume 47, No. 2, 231–240, 1992.

[2] P. Erdős, L. Mordechai, *d-complete sequences of integers.* Mathematics of Computation, Volume 64, No. 214, 837–840, 1996.

[3] M. M. Ma, Y. G. Chen, *On d-complete sequences of integers*. Journal of Number Theory, Volume 164, 1–12, 2016.

[4] Y. G. Chen, W. X. Yu, *On d-complete sequences of integers, II*. Acta Arithmetica, Volume 207, 861–181, 2023.

[5] R. Blecksmith, M. McCallum, J.L. Selfridge, 3-*smooth representations of integers*. The American Mathematical Monthly, Volume 105, No. 6, 529–543, 1998.

[6] G. H. Hardy, *Ramanujan: Twelve lectures on subjects suggested by his life and work*. Chelsea Publishing Company, 1940.

[7] B. J. Birch, *Note on a problem of Erdős*. Mathematical Proceedings of the Cambridge Philosophical Society, Volume 55, 370–373, 1959.

[8] T. F. Bloom, Erdős Problem #845, https://www.erdosproblems.com/845, accessed 2025-10-25.