

ON THE SET OF POINTS REPRESENTED BY HARMONIC SUBSERIES

VJEKOSLAV KOVAČ

ABSTRACT. We show that the set of points

$$\left(\sum_{n \in A} \frac{1}{n}, \sum_{n \in A} \frac{1}{n+1}, \sum_{n \in A} \frac{1}{n+2} \right),$$

obtained as A ranges over infinite sets of positive integers, has a non-empty interior. This answers in the affirmative a question of Erdős and Straus (posed in a 1980 problem book by Erdős and Graham).

1. INTRODUCTION

Paul Erdős asked numerous questions on representations of numbers as finite or infinite sums of distinct unit fractions; see [2, §4,6,7] and [1]. Over the years many of these problems have been solved and, more importantly, have motivated the development of new techniques in number theory and combinatorics. The problem that we study in this note is distinctly higher-dimensional and its solution will use virtually no ingredients from number theory, apart from a few funny-looking numerical identities stated in Section 2 below.

Erdős and Graham, in their 1980 monograph on open problems in combinatorial number theory [2], mentioned that

Erdős and Straus [unpublished] proved that if one takes all sequences of integers a_1, a_2, \dots , with $\sum_k 1/a_k < \infty$, then the set

$$\left\{ (x, y) : x = \sum_k \frac{1}{a_k}, y = \sum_k \frac{1}{1+a_k} \right\}$$

contains a [non-empty] open set; [2, p. 65].

It is also understood from the context that the sequence (a_k) is meant to be positive and strictly increasing. (Indeed, allowing the repetition of a_k 's would make this result and the following one significantly less challenging.) Erdős and Graham then asked [2, p. 65]:

Is the same true in three (or more) dimensions, e.g., taking all (x, y, z) with

$$x = \sum_k \frac{1}{a_k}, y = \sum_k \frac{1}{1+a_k}, z = \sum_k \frac{1}{2+a_k}?$$

We can give the positive answer to this three-dimensional problem, which has also been posed on Thomas Bloom's website *Erdős problems* [1, Problem #268]. We state our theorem in the spirit of Bloom's equivalent reformulation of the problem.

Theorem 1. *The set*

$$\left\{ \left(\sum_{n \in A} \frac{1}{n}, \sum_{n \in A} \frac{1}{n+1}, \sum_{n \in A} \frac{1}{n+2} \right) : A \subset \mathbb{N} \text{ is an infinite set with } \sum_{n \in A} \frac{1}{n} < \infty \right\} \subseteq \mathbb{R}^3$$

has a non-empty interior.

2020 *Mathematics Subject Classification.* Primary 11B75; Secondary 11D68, 40A05.

Key words and phrases. harmonic series, unit fraction, number representation, subset sum.

Here \mathbb{N} denotes the set of all positive integers. Theorem 1 clearly implies the aforementioned unpublished two-dimensional result of Erdős and Straus, but we are currently unable to say anything about dimensions larger than 3.

The basic idea of our proof is to construct disjoint finite sets $S_j, T_j \subset \mathbb{N}$, for each of the coordinates $j = 1, 2, 3$, such that replacing the (shifted harmonic) series' terms corresponding to indices nT_j by those corresponding to nS_j , for some $n \in \mathbb{N}$, increases the series' sum in the j -th coordinate by $c_j n^{-j}$, up to an $O(n^{-4})$ error in each of the coordinates. This is achieved after an appropriate linear change of variables that enables the desired asymptotics and after the specification of desired arithmetic constraints on S_j and T_j . Next, we only pick certain positive integers n for which we know that all of the considered harmonic series' terms are mutually different. By choosing n from a sequence that grows roughly like perfect squares we also make all discussed sums convergent. Then we define the point $p = (p_1, p_2, p_3)$ by summing the three series over the set $\cup_n \cup_j nT_j$. For every point $q = (q_1, q_2, q_3)$ in a small rectangular box "a little bit right" from p , we describe an algorithm that makes a sequence of replacements of nT_j with nS_j , for one n at a time, in order to gradually "move" the triple of series' sums from p to q and finally obtain q in the limit.

The arithmetic of \mathbb{N} comes into play only when we need to guarantee that such sets S_j, T_j can really be constructed. For example, we need to find two finite sets of distinct unit fractions with mutually equal sums of squares and sums of cubes, which calls for a combination of heuristics and computer search.

2. AN ARITHMETIC LEMMA

As we have already announced, the only arithmetic property of the positive integers needed in the proof is encoded in the following auxiliary result.

Lemma 2. *There exist a matrix $M \in \text{GL}(3, \mathbb{R})$, mutually disjoint finite sets $S_1, S_2, S_3, T_1, T_2, T_3 \subset \mathbb{N}$, and constants $c_1, c_2, c_3 \in (0, \infty)$ such that*

$$\left(\sum_{a \in S_j} - \sum_{a \in T_j} \right) M \begin{pmatrix} 1/(an) \\ 1/(an+1) \\ 1/(an+2) \end{pmatrix} = \frac{c_j}{n^j} e_j + O\left(\frac{1}{n^4}\right) \quad (1)$$

for $1 \leq j \leq 3$.

Here e_1, e_2, e_3 denotes the standard basis of \mathbb{R}^3 . It will always be more convenient to write points and vectors in \mathbb{R}^3 as 3×1 columns.

Proof of Lemma 2. Property (1) relies on the expansions

$$\frac{1}{an+k-1} = \frac{1}{an} - \frac{k-1}{a^2 n^2} + \frac{(k-1)^2}{a^3 n^3} + O\left(\frac{1}{n^4}\right),$$

for $1 \leq k \leq 3$, where the error terms are allowed to depend on a parameter $a \in \mathbb{N}$. If we choose

$$M = \begin{pmatrix} 1 & 0 & 0 \\ 3 & -4 & 1 \\ 1 & -2 & 1 \end{pmatrix},$$

then

$$M \begin{pmatrix} \frac{1}{an} \\ \frac{1}{an+1} \\ \frac{1}{an+2} \end{pmatrix} = \begin{pmatrix} \frac{1}{an} + O\left(\frac{1}{n^4}\right) \\ \frac{2}{a^2 n^2} + O\left(\frac{1}{n^4}\right) \\ \frac{2}{a^3 n^3} + O\left(\frac{1}{n^4}\right) \end{pmatrix}.$$

Thus, to construct (say) S_3 and T_3 , we need to make sure that

$$\sum_{a \in S_3} \frac{1}{a} = \sum_{a \in T_3} \frac{1}{a}, \quad \sum_{a \in S_3} \frac{1}{a^2} = \sum_{a \in T_3} \frac{1}{a^2}, \quad \sum_{a \in S_3} \frac{1}{a^3} \neq \sum_{a \in T_3} \frac{1}{a^3}.$$

The only remaining ingredients of the proof are then elementary identities

$$\begin{aligned} \frac{1}{10} + \frac{1}{30} + \frac{1}{60} &= \frac{1}{12} + \frac{1}{15}, \\ \frac{1}{10^2} + \frac{1}{30^2} + \frac{1}{60^2} &= \frac{1}{12^2} + \frac{1}{15^2}, \\ \frac{1}{16} + \frac{1}{20} + \frac{1}{240} &= \frac{1}{15} + \frac{1}{24} + \frac{1}{120}, \\ \frac{1}{16^3} + \frac{1}{20^3} + \frac{1}{240^3} &= \frac{1}{15^3} + \frac{1}{24^3} + \frac{1}{120^3}, \\ \frac{1}{45^2} + \frac{1}{72^2} + \frac{1}{144^2} + \frac{1}{160^2} + \frac{1}{432^2} + \frac{1}{480^2} &= \frac{1}{48^2} + \frac{1}{60^2} + \frac{1}{120^2} + \frac{1}{720^2} + \frac{1}{1440^2} + \frac{1}{4320^2}, \\ \frac{1}{45^3} + \frac{1}{72^3} + \frac{1}{144^3} + \frac{1}{160^3} + \frac{1}{432^3} + \frac{1}{480^3} &= \frac{1}{48^3} + \frac{1}{60^3} + \frac{1}{120^3} + \frac{1}{720^3} + \frac{1}{1440^3} + \frac{1}{4320^3}. \end{aligned}$$

Straightforward (preferably computer assisted) computation then verifies (1) with

$$\begin{aligned} S_1 &= \{45, 72, 144, 160, 432, 480\}, & T_1 &= \{48, 60, 120, 720, 1440, 4320\}, \\ S_2 &= 11 \cdot \{16, 20, 240\}, & T_2 &= 11 \cdot \{15, 24, 120\}, \\ S_3 &= 7 \cdot \{10, 30, 60\}, & T_3 &= 7 \cdot \{12, 15\} \end{aligned}$$

and

$$\begin{aligned} c_1 &= \frac{1}{45} + \frac{1}{72} + \frac{1}{144} + \frac{1}{160} + \frac{1}{432} + \frac{1}{480} - \frac{1}{48} - \frac{1}{60} - \frac{1}{120} - \frac{1}{720} - \frac{1}{1440} - \frac{1}{4320} = \frac{1}{180}, \\ c_2 &= \frac{2}{11^2} \left(\frac{1}{16^2} + \frac{1}{20^2} + \frac{1}{240^2} - \frac{1}{15^2} - \frac{1}{24^2} - \frac{1}{120^2} \right) = \frac{1}{348480}, \\ c_3 &= \frac{2}{7^3} \left(\frac{1}{10^3} + \frac{1}{30^3} + \frac{1}{60^3} - \frac{1}{12^3} - \frac{1}{15^3} \right) = \frac{1}{1029000}. \end{aligned}$$

The matrix M is regular since $\det M = -2$. □

Our application of Lemma 2 in the proof below does not use that the sets S_j and T_j are finite: they could just have finitely many prime factors.

3. PROOF OF THEOREM 1

Let m be the product of all prime factors of the numbers in the set

$$U := S_1 \cup S_2 \cup S_3 \cup T_1 \cup T_2 \cup T_3.$$

The proof of Lemma 2 actually gives

$$m = 2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 = 2310.$$

Every positive integer has at most one representation as $a(k^2m + 1)$ for $a \in U$ and $k \in \mathbb{N}$. The series

$$\sum_{\substack{a \in U \\ k \in \mathbb{N}}} \frac{1}{a(k^2m + 1)}, \quad \sum_{\substack{a \in U \\ k \in \mathbb{N}}} \frac{1}{a(k^2m + 1) + 1}, \quad \sum_{\substack{a \in U \\ k \in \mathbb{N}}} \frac{1}{a(k^2m + 1) + 2}$$

converge, so even after multiplying their terms arbitrarily by -1 , 0 , or 1 , we obtain absolutely convergent series, the terms of which can be freely permuted and grouped. Let $C \in (0, \infty)$ be a constant such that each component of the vector

$$\left(\sum_{a \in S_j} - \sum_{a \in T_j} \right) M \begin{pmatrix} 1/(an) \\ 1/(an+1) \\ 1/(an+2) \end{pmatrix} - \frac{c_j}{n^j} \mathbb{e}_j$$

is at most C/n^4 in the absolute value, for every $n \in \mathbb{N}$ and every $1 \leq j \leq 3$; it exists by Lemma 2. Next, let K be a sufficiently large positive integer such that

$$\underbrace{\sum_{l=k}^{\infty} \frac{3C}{(l^2m+1)^4}}_{\Theta(k^{-7})} < \underbrace{\frac{c_j}{(k^2m+1)^j}}_{\Omega(k^{-6})}, \quad (2)$$

$$\underbrace{\sum_{l=k+1}^{\infty} \frac{c_j}{(l^2m+1)^j}}_{\Theta(k^{-2j+1})} > \underbrace{\frac{4c_j}{(k^2m+1)^j}}_{\Theta(k^{-2j})} \quad (3)$$

both hold for every $k \geq K$ and $1 \leq j \leq 3$. Introduce the particular point

$$p = \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} := \sum_{l=K}^{\infty} \sum_{j=1}^3 \sum_{a \in T_j} M \begin{pmatrix} \frac{1}{a(l^2m+1)} \\ \frac{1}{a(l^2m+1)+1} \\ \frac{1}{a(l^2m+1)+2} \end{pmatrix} \in \mathbb{R}^3.$$

For every

$$q = \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix} \in \mathcal{Q} := p + \prod_{j=1}^3 \left[\frac{c_j}{(K^2m+1)^j}, \frac{2c_j}{(K^2m+1)^j} \right] \quad (4)$$

we are going to construct a sequence $(x_k)_{k=K}^{\infty}$,

$$x_k = \begin{pmatrix} x_{k,1} \\ x_{k,2} \\ x_{k,3} \end{pmatrix} \in \mathbb{R}^3,$$

that converges to the point q . We first set $x_K := p$ and then we define recursively

$$x_{k+1} := x_k + \sum_{j=1}^3 \epsilon_{k,j} \left(\sum_{a \in S_j} - \sum_{a \in T_j} \right) M \begin{pmatrix} \frac{1}{a(k^2m+1)} \\ \frac{1}{a(k^2m+1)+1} \\ \frac{1}{a(k^2m+1)+2} \end{pmatrix}$$

for every $k \geq K$, where the coefficients $\epsilon_{k,j} \in \{0, 1\}$ are determined from

$$\begin{aligned} \epsilon_{k,j} &= 1 && \text{if } x_{k,j} + \frac{3c_j}{(k^2m+1)^j} \leq q_j, \\ \epsilon_{k,j} &= 0 && \text{otherwise} \end{aligned}$$

for $1 \leq j \leq 3$. From this recurrence relation and the definition of the constant C , we have

$$\text{if } \epsilon_{k,j} = 1, \quad \text{then } \left| x_{k+1,j} - x_{k,j} - \frac{c_j}{(k^2m+1)^j} \right| \leq \frac{3C}{(k^2m+1)^4}, \quad (5)$$

$$\text{if } \epsilon_{k,j} = 0, \quad \text{then } |x_{k+1,j} - x_{k,j}| \leq \frac{3C}{(k^2m+1)^4}. \quad (6)$$

We need two auxiliary claims in the proof that $(x_k)_{k=K}^\infty$ really converges to q . The main idea is that the sequence was defined to make a step from x_k towards q only if their distance decreases at most by roughly a factor of $1/3$. Thus, we are not using an entirely greedy algorithm and we never allow x_k to approach too quickly to its tentative limit q in order to avoid possible overshoots coming from the error terms, either in the current step, or in any of the future steps.

Claim 1. *For each $1 \leq j \leq 3$ there exist infinitely many $k \geq K$ such that $\epsilon_{k,j} = 1$.*

Proof of Claim 1. We can immediately rule out the possibility $\epsilon_{k,j} = 0$ for each $k \geq K$, as this would imply

$$x_{N,j} + \frac{3c_j}{(N^2m+1)^j} > q_j \quad \text{for } N \geq K \quad (7)$$

and

$$x_{k+1,j} \stackrel{(6)}{\leq} x_{k,j} + \frac{3C}{(k^2m+1)^4} \quad \text{for } k \geq K,$$

so that

$$x_{N,j} \leq p_j + \sum_{k=K}^{N-1} \frac{3C}{(k^2m+1)^4} \quad (8)$$

for $N > K$. A combination of (7) and (8) then yields

$$q_j \leq p_j + \sum_{k=K}^{N-1} \frac{3C}{(k^2m+1)^4} + \frac{3c_j}{(N^2m+1)^j}$$

and letting $N \rightarrow \infty$ we conclude

$$q_j \leq p_j + \sum_{k=K}^{\infty} \frac{3C}{(k^2m+1)^4} \stackrel{(2)}{<} p_j + \frac{c_j}{(K^2m+1)^j},$$

which contradicts the choice of q , namely (4).

Similarly, suppose that there exists the largest integer $L \geq K$ such that $\epsilon_{L,j} = 1$. This means

$$x_{L,j} + \frac{3c_j}{(L^2m+1)^j} \leq q_j$$

and

$$x_{k,j} + \frac{3c_j}{(k^2m+1)^j} > q_j \quad \text{for } k \geq L+1,$$

implying

$$x_{N,j} - x_{L,j} \geq \frac{3c_j}{(L^2m+1)^j} - \frac{3c_j}{(N^2m+1)^j} \quad (9)$$

for every $N > L$. As opposed to that, the recurrence relation gives

$$|x_{L+1,j} - x_{L,j}| \stackrel{(5)}{\leq} \frac{c_j}{(L^2m+1)^j} + \frac{3C}{(L^2m+1)^4}$$

and

$$|x_{k+1,j} - x_{k,j}| \stackrel{(6)}{\leq} \frac{3C}{(k^2m+1)^4} \quad \text{for } k \geq L+1,$$

which yields

$$|x_{N,j} - x_{L,j}| \leq \frac{c_j}{(L^2m+1)^j} + \sum_{k=L}^{N-1} \frac{3C}{(k^2m+1)^4} \stackrel{(2)}{\leq} \frac{2c_j}{(L^2m+1)^j}. \quad (10)$$

Note that (9) and (10) together give

$$\frac{c_j}{(L^2m+1)^j} \leq \frac{3c_j}{(N^2m+1)^j}.$$

We arrive at a clear contradiction in the limit as $N \rightarrow \infty$, so the claim follows. \square

Claim 2. *For each $1 \leq j \leq 3$ there exist infinitely many $k \geq K$ such that $\epsilon_{k,j} = 0$.*

Proof of Claim 2. We certainly have $\epsilon_{K,j} = 0$ because

$$x_{K,j} + \frac{3c_j}{(K^2m+1)^j} = p_j + \frac{3c_j}{(K^2m+1)^j} > q_j$$

by (4). Suppose that the claim fails and that $L \geq K$ is the largest integer such that $\epsilon_{L,j} = 0$. We can read this as

$$x_{L,j} + \frac{3c_j}{(L^2m+1)^j} > q_j$$

and

$$x_{k,j} + \frac{3c_j}{(k^2m+1)^j} \leq q_j \quad \text{for } k \geq L+1.$$

Thus,

$$x_{N,j} - x_{L,j} \leq \frac{3c_j}{(L^2m+1)^j} \tag{11}$$

for every $N > L$. The sequence (x_k) has been defined in a way that

$$x_{L+1,j} \stackrel{(6)}{\geq} x_{L,j} - \frac{3C}{(L^2m+1)^4}$$

and

$$x_{k+1,j} \stackrel{(5)}{\geq} x_{k,j} + \frac{c_j}{(k^2m+1)^j} - \frac{3C}{(k^2m+1)^4} \quad \text{for } k \geq L+1,$$

implying

$$\begin{aligned} x_{N,j} - x_{L,j} &\geq \sum_{k=L+1}^{N-1} \frac{c_j}{(k^2m+1)^j} - \sum_{k=L}^{N-1} \frac{3C}{(k^2m+1)^4} \\ &\stackrel{(2)}{\geq} \sum_{k=L+1}^{N-1} \frac{c_j}{(k^2m+1)^j} - \frac{c_j}{(L^2m+1)^j}. \end{aligned} \tag{12}$$

Combining (11) and (12) we get

$$\sum_{k=L+1}^{N-1} \frac{c_j}{(k^2m+1)^j} \leq \frac{4c_j}{(L^2m+1)^j},$$

so letting $N \rightarrow \infty$ we obtain an estimate that contradicts (3). \square

Now we can return to the proof of Theorem 1. From $|x_{k+1} - x_k| = O(1/k^2)$ it follows that (x_k) is a Cauchy sequence in \mathbb{R}^3 , so it converges. By Claim 1 there exist infinitely many indices $k(j,1) < k(j,2) < \dots$ such that

$$x_{k(j,t),j} + \frac{3c_j}{(k(j,t)^2m+1)^j} \leq q_j,$$

which gives

$$\lim_{k \rightarrow \infty} x_{k,j} = \lim_{t \rightarrow \infty} x_{k(j,t),j} \leq q_j.$$

From Claim 2 we have infinitely many indices $k'(j, 1) < k'(j, 2) < \dots$ such that

$$x_{k'(j,t),j} + \frac{3c_j}{(k'(j,t)^2 m + 1)^j} > q_j,$$

which implies

$$\lim_{k \rightarrow \infty} x_{k,j} = \lim_{t \rightarrow \infty} x_{k'(j,t),j} \geq q_j.$$

It follows that (x_k) converges precisely to the point q .

On the other hand, by the definition of the sequence (x_k) , we have

$$\lim_{k \rightarrow \infty} M^{-1}x_k = \sum_{\substack{k \geq K \\ 1 \leq j \leq 3 \\ \epsilon_{k,j}=1}} \sum_{a \in S_j} \left(\frac{\frac{1}{a(k^2 m + 1)}}{\frac{1}{a(k^2 m + 1) + 1}} \right) + \sum_{\substack{k \geq K \\ 1 \leq j \leq 3 \\ \epsilon_{k,j}=0}} \sum_{a \in T_j} \left(\frac{\frac{1}{a(k^2 m + 1)}}{\frac{1}{a(k^2 m + 1) + 1}} \right).$$

Therefore, for every q as in (4), the point $M^{-1}q$ is of the desired form

$$\left(\sum_{z \in A} \frac{1}{z}, \sum_{z \in A} \frac{1}{z+1}, \sum_{z \in A} \frac{1}{z+2} \right)$$

for the set

$$A := \bigcup_{\substack{k \geq K \\ 1 \leq j \leq 3 \\ \epsilon_{k,j}=1}} \{a(k^2 m + 1) : a \in S_j\} \cup \bigcup_{\substack{k \geq K \\ 1 \leq j \leq 3 \\ \epsilon_{k,j}=0}} \{a(k^2 m + 1) : a \in T_j\}.$$

We are done since $M^{-1}\mathcal{Q}$ is a non-degenerate parallelepiped, so it has a non-empty interior. The proof of Theorem 1 is complete.

As Erdős and Graham mentioned, one could study further higher-dimensional generalizations of the same problem. The corresponding generalizations of Lemma 2 would require more involved variants of ad-hoc arithmetic identities used in its proof, but perhaps one could find a different proof of Theorem 1 that avoids such identities altogether.

ACKNOWLEDGMENT

This work was supported in part by the Croatian Science Foundation under the project HRZZ-IP-2022-10-5116 (FANAP).

REFERENCES

- [1] Thomas Bloom. Erdős problems. <https://www.erdosproblems.com/>. Accessed: May 12, 2024.
- [2] Paul Erdős and Ronald Lewis Graham. *Old and new problems and results in combinatorial number theory*, volume 28 of *Monographies de L'Enseignement Mathématique*. Université de Genève, L'Enseignement Mathématique, Geneva, 1980.

DEPARTMENT OF MATHEMATICS, FACULTY OF SCIENCE, UNIVERSITY OF ZAGREB, BIJENIČKA CESTA 30,
10000 ZAGREB, CROATIA

Email address: vjekovac@math.hr