Two-colouring and density lead to many solutions of $\frac{1}{x} + \frac{1}{y} = \frac{1}{z}$

Wouter van Doorn

Abstract

In 1991, Brown and Rödl [2] proved that for every $k \in \mathbb{N}$ and every r-colouring of $\mathbb{N}$, there exists a monochromatic solution to the equation $\frac{1}{x_0} = \frac{1}{x_1} + \frac{1}{x_2} + \ldots + \frac{1}{x_k}$, answering a question of Erdős and Graham [1]. In this short note we will focus on the case $k = r = 2$ and show that for every $c < \frac{1}{390}$ and for all large enough n, any two-colouring of $\{1, 2, \ldots, n\}$ contains at least cn solutions to $\frac{1}{x} + \frac{1}{y} = \frac{1}{z}$. We will furthermore answer another question of Erdős and Graham by showing that for every $c > \frac{9}{10}$ and all large enough n, any set $S \subseteq \{1, 2, \ldots, n\}$ with $|S| > cn$ contains a solution to $\frac{1}{x} + \frac{1}{y} = \frac{1}{z}$.

1 Two-colouring the positive integers

Let us dive straight in.

Theorem 1. *Let $c = \frac{1}{390} \approx 0.00256$. Then for every $n \in \mathbb{N}$ and every two-colouring of $\{1, 2, \ldots, n\}$, there are at least $cn - \log(n)^3 - 1$ monochromatic triples (x, y, z) with $\frac{1}{x} + \frac{1}{y} = \frac{1}{z}$ and x, y, z distinct.*

Proof. Define $S_1 = \{6, 8, 9, 10, 12, 15, 18, 20, 24, 30, 36, 40, 60, 72, 90, 120\}$. Then it can be checked (either with a computer or with a lot of patience and by hand) that every two-colouring of S_1 leads to a monochromatic solution to $\frac{1}{x} + \frac{1}{y} = \frac{1}{z}$ with x, y, z distinct. More generally, one can then define the set $S_a = \{6a, 8a, 9a, 10a, 12a, 15a, 18a, 20a, 24a, 30a, 36a, 40a, 60a, 72a, 90a, 120a\}$, and by the same argument, every two-colouring of S_a leads to a monochromatic triple. It is therefore sufficient to count the number of pairwise disjoint sets S_a with $a \leq \frac{n}{120}$ to get a lower bound on the total number of monochromatic solutions to $\frac{1}{x} + \frac{1}{y} = \frac{1}{z}$.

Lemma 1. *Define S_a as above for all a with $a = 16^b \cdot 27^c \cdot 25^d \cdot e$, where $b, c, d \geq 0$ and $e \geq 1$ is coprime to $30 = 2 \cdot 3 \cdot 5$. Then all the sets S_a are pairwise disjoint.*

Proof. For given $a = 16^b \cdot 27^c \cdot 25^d \cdot e$ and $a' = 16^{b'} \cdot 27^{c'} \cdot 25^{d'} \cdot e'$ with e and e' coprime to 30, let l be an element of both S_a and $S_{a'}$. Since no element of S_1 is divisible by 16, it follows that l is divisible by 16^b, but not by 16^{b+1}. Since l is similarly divisible by $16^{b'}$, but not by $16^{b'+1}$, we conclude that $b = b'$. The same argument shows $c = c'$ and $d = d'$. It is therefore sufficient to show $e = e'$ as well. Now, since the prime factorization of every element of S_1 only contains the primes $2, 3$ and 5, l can be written as me and as $m'e'$ with m and m' only divisible by the primes $2, 3$ and 5 and both e and e' coprime to $2, 3$ and 5. By unique factorization the equality $me = m'e'$ then implies $m = m'$ and $e = e'$. We therefore conclude $a = a'$. ∎

Lemma 2. *For all $n \in \mathbb{N}$ there are at least $\frac{n}{390} - \log(n)^3 - 1$ sets S_a with a as in Lemma 1 and $a \leq \frac{n}{120}$.*

Proof. For n with $1 \leq n \leq 389$ this lemma trivially holds, so we are free to assume $n \geq 390$. In that case define i, j, k as $\left\lfloor \frac{\log(n)}{\log(16)} \right\rfloor$, $\left\lfloor \frac{\log(n)}{\log(27)} \right\rfloor$ and $\left\lfloor \frac{\log(n)}{\log(25)} \right\rfloor$ respectively. Before we can calculate the number of sets S_a, there are a few things we have to point out.

First we claim that i is smaller than $0.53\log(n) - 1$ and that j and k are smaller than $0.48\log(n) - 1$, by the assumption $n \geq 390$. Indeed, $i \leq \frac{\log(n)}{\log(16)} < 0.361\log(n) = 0.53\log(n) - 1 + \big(1 - 0.169\log(n)\big) < 0.53\log(n) - 1$ and $j \leq k \leq \frac{\log(n)}{\log(25)} < 0.311\log(n) = 0.48\log(n) - 1 + \big(1 - 0.169\log(n)\big) < 0.48\log(n) - 1$.

Secondly, by the way we defined i, we get $n \leq 16^{i+1} < 30 \cdot 16^i$, which implies $\frac{2}{n} > \frac{1}{15\cdot 16^i}$. Similarly, $\frac{2}{n} > \frac{1}{26\cdot 27^j}$ and $\frac{2}{n} > \frac{1}{24\cdot 25^k}$.

Finally, since there are 8 integers in every interval $(30m, 30(m+1)]$ that are coprime to 30, we get 8 sets of S_a where all elements are smaller than n, for all $a = 16^b \cdot 27^c \cdot 25^d \cdot e$ with $e \leq \frac{n}{3600\cdot 16^b\cdot 27^c\cdot 25^d}$.

Now we get to calculate.

$$
\begin{aligned}
\#S_a &\geq \sum_{b=0}^{i}\sum_{c=0}^{j}\sum_{d=0}^{k}\sum_{e=1}^{\left\lfloor \frac{n}{3600\cdot 16^b\cdot 27^c\cdot 25^d}\right\rfloor} 8 \\
&\geq \sum_{b=0}^{i}\sum_{c=0}^{j}\sum_{d=0}^{k} \frac{n}{450\cdot 16^b\cdot 27^c\cdot 25^d} - 8 \\
&= \sum_{b=0}^{i}\sum_{c=0}^{j} \frac{n}{450\cdot 16^b\cdot 27^c}\left(\frac{25-25^{-k}}{24}\right) - 8(k+1) \\
&= \sum_{b=0}^{i} \frac{n}{450\cdot 16^b}\left(\frac{27-27^{-j}}{26}\right)\left(\frac{25-25^{-k}}{24}\right) - 8(j+1)(k+1) \\
&= \frac{n}{450}\left(\frac{16-16^{-i}}{15}\right)\left(\frac{27-27^{-j}}{26}\right)\left(\frac{25-25^{-k}}{24}\right) - 8(i+1)(j+1)(k+1) \\
&> \frac{n}{450}\left(\frac{16}{15}-\frac{2}{n}\right)\left(\frac{27}{26}-\frac{2}{n}\right)\left(\frac{25}{24}-\frac{2}{n}\right) - 8(i+1)(j+1)(k+1) \\
&> \frac{n}{450}\left(\frac{15}{13}-\frac{17}{n}\right) - 8\big(0.53\log(n)\big)\big(0.48\log(n)\big)\big(0.48\log(n)\big) \\
&> \frac{n}{390} - \log(n)^3 - 1
\end{aligned}
$$

□

2 Dense sets

Let $S \subseteq \{1, 2, \ldots, n\}$ be a set that does not contain three distinct integers x, y, z with $\frac{1}{x} + \frac{1}{y} = \frac{1}{z}$. By the proof of Theorem 1 we see that S needs to avoid at least one element from every set S_a with $a \leq \frac{n}{120}$. That means $|S| \leq n - (\frac{n}{390} - \log(n)^3 - 1) = \frac{389n}{390} + \log(n)^3 + 1$. Taking the contrapositive we conclude that if S contains more than $\frac{389n}{390} + \log(n)^3 + 1$ elements, then S must contain a solution. We can do better, however.

Theorem 2. *Let n be a positive integer and let $S \subseteq \{1, 2, \ldots, n\}$ be a set with $|S| \geq \frac{9n}{10} + \log(n)^3 + 1$. Then S contains three distinct integers x, y, z with $\frac{1}{x} + \frac{1}{y} = \frac{1}{z}$.*

Define $S_1 = \{2, 3, 6\}$ and $T_1 = \{4, 5, 20\}$. Since $\frac{1}{3} + \frac{1}{6} = \frac{1}{2}$ and $\frac{1}{5} + \frac{1}{20} = \frac{1}{4}$, any set that avoids a solution to $\frac{1}{x} + \frac{1}{y} = \frac{1}{z}$, can contain at most two elements from S_1 and at most two elements from T_1. More generally, define $S_a = \{2a, 3a, 6a\}$ and $T_e = \{4e, 5e, 20e\}$. Similarly to what we did in Section 1, we will find a lower bound on the number of disjoint sets S_a and T_e with all elements smaller than or equal to n. This then provides an upper bound on how many elements a set can have before being forced to contain a solution to $\frac{1}{x} + \frac{1}{y} = \frac{1}{z}$.

Lemma 3. *Define $S_a = \{2a, 3a, 6a\}$ with $a = 4^b \cdot 9^c \cdot d$ and $\gcd(d, 6) = 1$, and define $T_e = \{4e, 5e, 20e\}$ with $e = 16^f \cdot 9^g \cdot 25^h \cdot i$ and $\gcd(i, 30) = 1$. Then $S_a \cap S_{a'} = S_a \cap T_e = T_e \cap T_{e'} = \emptyset$ whenever $a \neq a'$ and $e \neq e'$.*

Proof. The proofs of $S_a \cap S_{a'} = \emptyset$ and $T_e \cap T_{e'} = \emptyset$ are completely analogous to the proof of Lemma 1, so they are left to the reader. To see why $S_a \cap T_e = \emptyset$ as well, note that every element of T_e is exactly divisible by both an even power of two and an even power of three. On the other hand, every element of S_a is either exactly divisible by an odd power of two or by an odd power of three. $\square$

Just like in Section 1, we are left with determining how many subsets S_a and T_e of $\{1, 2, \ldots, n\}$ there are. As this calculation is analogous to the calculation in Lemma 2, we will be a bit more brief.

Lemma 4. *For all $n \in \mathbb{N}$ with $n > 1000$ there are more than $\frac{n}{12} - \frac{1}{6}\log(n)^3 - \frac{1}{2}$ sets S_a with a as in Lemma 3 and $a \leq \frac{n}{6}$, and more than $\frac{n}{60} - \frac{5}{6}\log(n)^3 - \frac{1}{2}$ sets T_e with e as in Lemma 3 and $e \leq \frac{n}{20}$.*

Proof. To determine a lower bound on the number of sets S_a, first define j and k as $\left\lfloor \frac{\log(n)}{\log(4)} \right\rfloor$ and $\left\lfloor \frac{\log(n)}{\log(9)} \right\rfloor$ respectively. Since $n > 1000$ by assumption, we obtain the inequalities $\log(n) > 6.5$, $j < 0.9\log(n) - 1$ and $k < 0.6\log(n) - 1$. We then get the following:

$$
\begin{aligned}
\#S_a &\geq \sum_{b=0}^{j}\sum_{c=0}^{k}\sum_{d=1}^{\left\lfloor \frac{n}{36\cdot 4^b\cdot 9^c}\right\rfloor} 2 \\
&\geq \sum_{b=0}^{j}\sum_{c=0}^{k}\frac{n}{18\cdot 4^b\cdot 9^c}-2 \\
&> \frac{n}{18}\left(\frac{4}{3}-\frac{2}{n}\right)\left(\frac{9}{8}-\frac{2}{n}\right)-2(j+1)(k+1) \\
&> \frac{n}{12}-\frac{1}{6}\log(n)^3-\frac{1}{2}
\end{aligned}
$$

As for T_e, now define j,k,l as $\left\lfloor\frac{\log(n)}{\log(16)}\right\rfloor$, $\left\lfloor\frac{\log(n)}{\log(9)}\right\rfloor$ and $\left\lfloor\frac{\log(n)}{\log(25)}\right\rfloor$ respectively. It can furthermore be checked that $\frac{n}{60}-\frac{5}{6}\log(n)^3-\frac{1}{2}<0$ for $n\leq 60000$, so let us assume $n>60000$. We then have $j<0.46\log(n)-1$, $k<0.55\log(n)-1$ and $l<0.41\log(n)-1$.

$$
\begin{aligned}
\#T_e &\geq \sum_{f=0}^{j}\sum_{g=0}^{k}\sum_{h=0}^{l}\sum_{i=1}^{\left\lfloor \frac{n}{600\cdot 16^f\cdot 9^g\cdot 25^h}\right\rfloor} 8 \\
&\geq \sum_{f=0}^{j}\sum_{g=0}^{k}\sum_{h=0}^{l}\frac{n}{75\cdot 16^f\cdot 9^g\cdot 25^h}-8 \\
&> \frac{n}{75}\left(\frac{16}{15}-\frac{2}{n}\right)\left(\frac{9}{8}-\frac{2}{n}\right)\left(\frac{25}{24}-\frac{2}{n}\right)-8(j+1)(k+1)(l+1) \\
&> \frac{n}{60}-\frac{5}{6}\log(n)^3-\frac{1}{2} \qquad \square
\end{aligned}
$$

Proof of Theorem 2. Since $\frac{9n}{10}+\log(n)^3+1>n$ for all $n\leq 1000$, we are free to assume $n>1000$, in which case Lemma 4 applies. We then get from Lemmas 3 and 4 that there are more than $\frac{n}{12}-\frac{1}{6}\log(n)^3-\frac{1}{2}+\frac{n}{60}-\frac{5}{6}\log(n)^3-\frac{1}{2}=\frac{n}{10}-\log(n)^3-1$ disjoint sets S_a and T_e with largest element smaller than or equal to n. Every subset of $\{1,2,\ldots,n\}$ that does not contain any solution to $\frac{1}{x}+\frac{1}{y}=\frac{1}{z}$ must avoid at least one element from all sets S_a and T_e, and must therefore have less than $\frac{9n}{10}+\log(n)^3+1$ elements. $\square$

References

[1] P. Erdős, R.L. Graham, *Old and New Problems and Results in Combinatorial Number Theory.* Enseign. Math. (2), vol. 28, Enseignement Math., Geneva, 1980. Also available here.

[2] T.C. Brown, V. Rödl, *Monochromatic solutions to equations with unit fractions.* Bull. Aust. Math. Soc. 43 387–392, 1991. Also available here.