Mathematical Reflections
Author: Juan José Granier .
February 22, 2025

0690

1 Problem

Let $\mathcal{P}$ be the set of all prime numbers. We want to show that there exists an infinite set S, such that $S \subset \mathcal{P}$ and for every finite subset of S, the sum of its elements is not a perfect power number, i.e. if $A \subset S$ is finite, then $\sum_{p\in A} p \neq b^j$, with $b, j \geq 2$.

2 Solution

First notice that if we start with $p_1 = 2$, then $p_1 \neq b^j, \forall b, j \geq 2$, because p_1 is a prime number. If we add $p_2 = 3$, then we have $p_1, p_2, p_1 + p_2 \neq b^j, \forall b, j \geq 2$, because p_1, p_2 and $p_1 + p_2$ are prime numbers. Then, if we start with $p_1 = 2$, we have the first two prime numbers in a finite construction. If we want to add $p_3 = 5$, we notice that $p_2 + p_3 = 3 + 5 = 8 = 2^3$, then 5 is not a valid prime number for the build. If we keep seeking for the next primes in order, $\{7, 11, 13, 17, 19, ...\}$, we have that the next prime in our construction is $p_3 = 17$, and notice that at that step if we define the index with the order in which the prime numbers appear in our search, we have $p_1 = 2, p_2 = 3$ and $p_7 = 17$ as the prime numbers that we have found, starting from 2, which are valid for the condition we are looking for, because we are looking for each prime number in the canonical ordering. Then, we define $p_j = p_{s_j}$, with s_j the s_j-th prime number in the canonical ordering, such that p_{s_j} is the first prime number that added does not violate the conditions we are looking for, given $p_1, p_2, \cdots, p_{j-1}$. If we keep searching in order for these primes, we have that the fourth and fifth primes such that the condition is accomplished are 37 and 227, respectively. This is possible by hand for until 37 and using an algorithm to find in order the next candidate.

We want to show the result by some arguments. First, we want to show that the size of the restrictions a prime number candidate adds is finite and then the order of growth at each step is subexponential. Suppose we have constructed the set $S_k = \{\overline{p}_1, \overline{p}_2, \cdots, \overline{p}_k\}$, starting from $\overline{p}_1 = 2$, and given $\overline{p}_2 = 3, \overline{p}_3 = 17, \overline{p}_4 = 37, \overline{p}_5 = 227$ and so on, until we reach $\overline{p}_k, k \geq 5$. Let $\mathcal{S}_{\mathcal{R},\mathcal{S}_\|}$, the set of restrictions generated by adding a first candidate of $\overline{p}_{k+1}$, namely $p_* = p_{s_k+1}$, the $s_k + 1-$th prime number. In other words, we define the set

$$\mathcal{S}_{\mathcal{R},S_k}(p_*) = \{b^n - S_A, n \geq 2, A \subseteq S_k\}$$

with $b \geq 2$ and $S_A = \sum_{\overline{p}_i \in A} \overline{p}_i$

Lemma 2.1 *The number of restrictions in $\mathcal{S}_{\mathcal{R},S_k}$ is finite and grows in subexponential order.*

Proof : By Bertrand's Postulate, $p_* \in (\overline{p}_k, 2\overline{p}_k)$. Then, notice that if $n \geq 2$, then $b \geq \left\lceil (S_A + 1)^{\frac{1}{n}} \right\rceil$ and by the election of the candidate, $b^n \leq S_A + 2\overline{p}_k \Rightarrow b \leq \left\lfloor (S_A + 2\overline{p}_k)^{\frac{1}{n}} \right\rfloor$.
Given the fact $p_k \approx k \log k, \forall k \geq 1$, with p_k the $k-$th prime number, we notice that $\overline{p}_1 = 2$ is the first prime number, $\overline{p}_2 = 3$ is the second prime number, and by construction $\overline{p}_3, \cdots, \overline{p}_k$ are the $s_3, \cdots, s_k$-th prime numbers, with $s_3, \cdots, s_k \in \mathbb{N}$.
So, we can approximate $2\overline{p}_k \approx 2s_k \log s_k$ and notice that $S_{S_k} \leq kp_{s_k} \approx ks_k \log s_k \leq \frac{s_k^2}{2} \log s_k$. This 2 bounds and the fact $a^{\frac{1}{n}}$ with $a \geq 1$ is decreasing for $n \geq 1$ implies that

$$\left\lceil \sqrt{S_A + 1} \right\rceil \leq b \leq \left\lfloor \sqrt{S_A + 2\overline{p}_k} \right\rfloor \leq \left\lfloor \sqrt{\frac{s_k \log s_k}{2} (4 + s_k)} \right\rfloor$$

$$2 \leq n \leq \lfloor \log_2 (S_A + 2\overline{p}_k) \rfloor \leq \left\lfloor \log_2 \left(\frac{s_k \log s_k}{2} (4 + s_k) \right) \right\rfloor$$

and then $|\mathcal{S}_{\mathcal{R},\mathcal{S}_{\|}}| < +\infty$. Furthermore, we notice that

$$|\mathcal{P}\left(S_k\right)| = \mathcal{O}\left(2^k\right)$$
$$b = \mathcal{O}\left(s_k\sqrt{\log s_k}\right) = \mathcal{O}\left(k\sqrt{\log k}\right)$$
$$n = \mathcal{O}\left(\log k\right)$$

The second and third asymptotic equalities are derived from the fact that again by the approximation $p_k \approx k\log k$ and the inequality $t_k \log t_k < p_{t_{k+1}} < 2t_k \log t_k$ with $t_{k+1} > t_k = s_k$ such that the inequality applies by Bertrand's Postulate, we notice that by the Prime Number Theorem $s = \pi\left(p_s\right) \approx \frac{p_s}{\log p_s} \Rightarrow t_{k+1} \approx \frac{p_{t_{k+1}}}{\log p_{t_{k+1}}}$ and then

$$\frac{t_k \log t_k}{\log\left(2t_k \log t_k\right)} < t_{k+1} < \frac{2t_k \log t_k}{\log\left(2t_k \log t_k\right)}$$

and then using the approximation $\log\left(t_k \log t_k\right) \approx \log t_k$, we have

$$t_{k+1} > \frac{t_k \log t_k}{\log t_k} = t_k \Rightarrow t_{k+1} = \Omega\left(t_k\right)$$
$$t_{k+1} < \frac{2t_k \log t_k}{\log t_k} = ts_k \Rightarrow t_{k+1} = \mathcal{O}\left(t_k\right)$$

and this implies that $\exists C > 0 : t_k < t_{k+1} \leq Ct_k$, so $t_{k+1} = \Theta\left(t_k\right)$ and then $t_{k+1} = \mathcal{O}\left(t_k\right)$. So, we have that $t_k \approx k$ is a good approximation for the second asymptotic equality. As we are looking for s_{k+1}, if we keep searching for other candidates in the canonical ordering, if we find the prime $\overline{p}_{k+1} = p_{s_{k+1}}$ in the order we want, then there exists $l \geq 1$, such that $p_{s_{k+1}} \in \left(2^{l-1}p_{s_k}, 2^l p_{s_k}\right)$. Then, as the parameter $l \geq 1$ is a constant w.r.t. s_k, then it follows by the same reasoning that $s_{k+1} = \mathcal{O}\left(s_k\right) \Rightarrow s_k \approx k$, and then we conclude for the second asymptotic equality. For the third asymptotic equality we use an analogous approach, first notice

$$n \leq \log_2\left(\frac{k^2 \log k}{2}\right)$$

and using the property $\log_2\left(ab\right) = \log_2 a + \log_2 b$, we have

$$n \leq \log_2 k^2 + \log_2\left(\log k\right) - 1$$

and asymptotically for big values of k, the term $\log_2\left(\log k\right)$ is small compared to $2\log_2 k$ and then we can bound $n = \mathcal{O}\left(\log k\right)$

This implies

$$|\mathcal{S}_{\mathcal{R},S_k}| = \mathcal{O}\left(2^k k \log^{\frac{3}{2}} k\right)$$

and then the size of the restrictions is big but finite, and the increase in size is bounded. Notice that in the interval $\left(p_{s_k}, 2p_{s_k}\right)$ the growth of perfect power numbers has a law of $\mathcal{O}\left(\left(2p_{s_k}\right)^{\frac{1}{2}}\right)$, that implies that the number of perfect power numbers in the interval is considerably smaller than the total number of numbers in the interval, so the fraction of numbers affected by the restrictions is significantly smaller. More generally, given an interval $[1, x], x \in \mathbb{N}$, the number of perfect power numbers increases at order $\mathcal{O}\left(x^{\frac{1}{2}}\right)$. Thus, combining with the bound we obtained before, we know that the number of subsets of restrictions is considerably smaller than we estimated before. The factor 2^k changes to $2^{\frac{k}{2}}$, and then

$$|\mathcal{S}_{\mathcal{R},S_k}| = \mathcal{O}\left(2^{\frac{k}{2}} k \log^{\frac{3}{2}} k\right) \blacksquare$$

It's easy to notice that the sums $S_A, A \subseteq S_k$ increase at a rate of at least $\mathcal{O}\left(k^2 \log k\right)$, because $p_{s_k} \approx s_k \log s_k \approx k \log k$. Let $\mathcal{S}$ be the set of finite sums of finite subsets of the Prime Numbers, let $\mathcal{N}$ be the set of Perfect Power Numbers, i.e., numbers of the form $b^j, b, j \geq 2$. We prove some additional facts.

Lemma 2.2 *$\mathcal{S}$ is a discrete set in $\mathbb{N}$*

Proof : First, notice that $\mathcal{S}$ is a countable set, because

$$\mathcal{S} = \bigcup_{k=1}^{+\infty} \{ \sum_{p_i \in A} p_i : A \subseteq \{p_1, p_2, \cdots, p_k\}\}$$

with $p_1, p_2, \cdots, p_k$ the first $k \geq 1$ prime numbers in the canonical ordering. Then $\mathcal{S}$ is the countable union of countable sets, because each of these sets is finite.
To see that $\mathcal{S}$ is discrete in $\mathbb{N}$, with minimum gap $\delta = 2$, just notice that given A, B finite subsets of the Prime Numbers, w.l.o.g., $S_A < S_B$, as the prime numbers are strictly increasing, the smallest possible difference occurs when $B = A \cup \{p\}, p \notin A \Rightarrow S_B - S_A = p$. The smallest such p is 2 and is the only even prime number, then $\delta = 2$, and since the set is bounded below by 2 and every change in sum increases by at least 2 or another prime, no accumulation point exists. Thus, $\mathcal{S}$ is discrete in $\mathbb{N}$ ■

Lemma 2.3 *$\mathcal{N}$ is not dense in $\mathbb{N}$*

Proof : s before, we have that the number of perfect power numbers in an interval $[1, x], x \in \mathbb{N}$ is $\mathcal{N}(x) = \mathcal{O}\left(x^{\frac{1}{2}}\right)$, then if we define $\mathbb{N}(x)$ as the number of elements of $\mathbb{N}$ in the interval $[1, x]$, then $\mathbb{N}(x) = x$, and we have

$$\lim_{x \to +\infty} \frac{\mathcal{N}(x)}{\mathbb{N}(x)} = \lim_{x \to +\infty} \frac{\mathcal{O}\left(x^{\frac{1}{2}}\right)}{x} = \lim_{x \to +\infty} \mathcal{O}\left(x^{-\frac{1}{2}}\right) = 0$$

Then, we deduce that $\mathcal{N}$ is not dense in $\mathbb{N}$ ■

Lemma 2.4 *$\mathcal{S}$ is not dense in $\mathcal{N}$*

Proof : Given $b \geq 2$, notice that

$$b^{j+1} - b^j = b^j (b-1) \geq 2^j (b-1) \geq 2^j, \forall j \geq 2$$

Then, the gap between powers of a given number increases at an exponential rate. As we saw before, S_k of the algorithm increases its sums at a linear rate, and the same applies for the set $\mathcal{S}$. Let

$$\Delta = \min\{|s - n| : s \in \mathcal{S}, n \in \mathcal{N}, s \neq n\}$$

Given that $\mathcal{S}$ is discrete and $\mathcal{N}$ grows exponentially, there exists a positive constant $\varepsilon > 0$ such that $\Delta \geq \varepsilon$. That guarantees that there is a gap between $\mathcal{S}$ and $\mathcal{N}$, then we conclude ■

Lemma 2.5 *$\mathcal{S}$ is not dense in $\mathbb{N}$*

Proof : As $\mathcal{S}$ is not dense in $\mathcal{N}$ and $\mathcal{N}$ is not dense in $\mathbb{N}$, we conclude that $\mathcal{S}$ is not dense in $\mathbb{N}$ ■

With all the previous results in mind, we have a first alternative to prove our result, joint with the fact that by the Prime Number Theorem that in every Bertrand Interval of the form $\left(2^{r-1} p_{s_k}, 2^r p_{s_k}\right)$, with p_{s_k} the s_k−th prime number, with $k \geq 1, r \geq 1$ there exists a finite number of primes. We are going to use these results in another way. First, define S^p as a subset of the Prime Numbers, starting with $\overline{p}_1 = p$, with p a prime number and with the same reasoning as before. Then, in our case, we began constructing S^2. Let

Lemma 2.6 *If S^p and S^q are infinite, with $p \neq q$, then $S^p \neq S^q$*

Proof : Let $p_k^{(p)}$ be the $k-$th prime selected in the construction, starting with the prime p. The function

$$f : \mathbb{N} \times \mathcal{P} \longrightarrow \mathcal{P}$$
$$f(k,p) = p_k^{(p)}$$
$$f(1,p) = p, \forall p \in \mathcal{P}$$

is strictly increasing with fixed $p \in \mathcal{P}$, let $g_p(k) = f(k,p)$. If $p \neq q$ and $S^p = S^q$, then $p = g_p(1) = g_q(1) = q$, a contradiction ■

Let $\mathcal{S}_{\mathcal{R},S_k^p}$ be the set of restrictions as before, with $p \in \mathcal{P}$. We say that $m \in \mathbb{N}$ **activates a restriction** in $\mathcal{S}_{\mathcal{R},S_k}$ if there exists $A \subseteq S_k^p$, such that $S_A + m = b^n, b, n \geq 2$

Lemma 2.7 *The restrictions are unique, i.e., the number of restrictions that are activated by $m \in \mathbb{N}$ is at most one.*

Proof : Suppose $m \in \mathbb{N}$ activates 2 or more restrictions, w.l.o.g. there are two restrictions. Then, there are $(b_1, n_1, A_1), (b_2, n_2, A_2)$ for the same m in the restriction set, with $b_1^{n_1} > b_2^{n_2}$. Then $S_{A_1} > S_{A_2}$ and we have:

$$m = b_1^{n_1} - S_{A_1} = b_2^{n_2} - S_{A_2}$$
$$\Rightarrow b_1^{n_1} - b_2^{n_2} = S_{A_1} - S_{A_2} \geq 2$$

As $b_1^{n_1} - b_2^{n_2} \geq b_2^{n_2}(b_1^{n_1-n_2} - 1) \geq b_2^{n_2}(b_1 - 1)$, we have $S_{A_1} - S_{A_2} \geq b_2^{n_2}$, but then we have

$$S_{A_1} \geq b_2^{n_2} + S_{A_2} \geq b_2^{n_2}$$

Then, $m \geq b_2^{n_2}$, but

$$b_1^{n_1} - S_{A_1} \geq b_2^{n_2}$$
$$\Rightarrow b_1^{n_1} - b_2^{n_2} = S_{A_1} - S_{A_2} \geq S_{A_1}$$

A clear contradiction, because $S_{A_2} > 0$ ■

Before we give the main part of the proof, we notice two more results, regarding the construction of the sets and a relation between the prime numbers and the perfect power prime numbers:

Lemma 2.8 *If a restriction is activated by the addition of a new prime p_*, then p must be in A, with $A \subseteq S_k^p \cup \{p_*\}$*

Proof : by construction of the set, at each step we only add sets of the form $\{p\}$ and $A \cup \{p_*\}$, with $p_* > p_k^{(p)}$. Then, by the uniqueness of the activated restrictions, when we add a prime p to the previous set, if there are additional activated restrictions in $S_k^p \cup \{p\}$, then p must belong to A, with $A \subseteq S_k^p \cup \{p_*\}$. As a consequence, if S_k^p is constructed by the inductive process, i.e. there are no subsets whose sum equals a perfect power number, then if appending p_* adds activated restrictions, then there is only one such restriction and p_* must be in A, because $A \subseteq S_k^p \cup \{p_*\}$ ■

Lemma 2.9 *The quantity of prime numbers is much greater than the quantity of perfect power numbers at greater intervals. As a consequence, the Perfect Power Numbers are not dense in the Prime Numbers.*

Proof : Given $x \in \mathbb{N}$, we have that $\mathcal{N}(x) = \mathcal{O}\left(x^{\frac{1}{2}}\right)$, while $\pi(x) = \mathcal{O}\left(\frac{x}{\log x}\right)$, with $\pi(x)$ the number of prime numbers in the interval $[1, x]$, and this last result is given by the Prime Number Theorem. Then, we have that $\mathcal{N}(x)$ is at most polynomial, while the growth of prime numbers is superpolynomial, because as $x \to +\infty$, the growth of $x \log x$ is greater than any polynomial function of x, with $p_x \approx x \log x$, where p_x is the $x-$th prime number. Then, we have

$$0 \leq \lim_{x \to +\infty} \frac{\mathcal{N}(x)}{\pi(x)} = \lim_{x \to +\infty} \frac{\mathcal{O}\left(x^{\frac{1}{2}}\right)}{\mathcal{O}\left(\frac{x}{\log x}\right)}$$
$$\leq \lim_{x \to +\infty} \frac{\mathcal{O}(\log x)}{\mathcal{O}\left(x^{\frac{1}{2}}\right)} = 0$$

Then, there exists $N \in \mathbb{N}$ such that $\pi(x) \ggg \mathcal{N}(x), \forall x \geq N, x \in \mathbb{N}$ ■

Henceforth, we define S^2 as the initial construction. Previously, we knew the values of $p_i^{(2)}, 1 \leq i \leq 5$, then $k \geq 5$. By contradiction, suppose that S^p is finite for every $p \in \mathcal{P}$. Then, for every prime, there exists $k_p \in \mathbb{N}$ such that $S^p = S^p_{k_p}$, with $k = k_2 \geq 5$ as we know. Let $p > p_k^{(2)}$, then as S^2 is finite, $\exists!\,(b, n, A) : b^n - S_A = p, A \subseteq S^2, b, n \geq 2$.

Lemma 2.10 *There exists an interval* $[x_0, 2x_0]$ *such that if we add primes without considering correctness in strictly increasing order, using the canonical order, then* $|\hat{\mathcal{S}}_{\mathcal{R},S}| > \mathcal{N}(x_0)$ *with* $\hat{\mathcal{S}}_{\mathcal{R},S}$ *the restriction of* $\mathcal{S}_{\mathcal{R},S}$ *to* $[x_0, 2x_0]$

Proof : We have proven that the number of active restrictions generated at each step meets the order $\mathcal{O}\left(2^{\frac{k}{2}} k \log^{\frac{3}{2}} k\right)$. On the other hand, we have $\mathcal{N}(x) = \mathcal{O}\left(x^{\frac{1}{2}}\right)$. Then, if we consider adding primes without considering active restrictions, then if we consider $\hat{\mathcal{S}}_{\mathcal{R},S}$ and $\mathcal{N}(x, 2x)$, with the last being the number of perfect power numbers in the interval $[x, 2x]$, noticing that it is no more than the shift of the interval $[0, x]$ by adding x to both ends and if we define $\mathcal{N}(x)$ in the interval $[0, x]$, it is the same as before, then we have that in both cases they meet the same bound, i.e. the same order. Then, given the bounds, we notice that the number of restrictions grows at an exponential rate w.r.t. the prime numbers that we add and the number of perfect power numbers in the same interval grows only with order $x^{\frac{1}{2}}$, then there exists $x_0 \in \mathbb{N}$ such that $|\hat{\mathcal{S}}_{\mathcal{R},S}| > \mathcal{N}(x_0)$ ■

Lemma 2.11 *Given* $p_i^{(2)} \in S^2, 1 \leq i \leq k$, *there exists a prime* $p > p_k^{(2)}$ *such that*

$$b^n - S_A = p \Rightarrow p_i^{(2)} \notin A$$

Proof : If not, then the family $I = \{A \subseteq S^2 : p_i^{(2)} \notin A\}$ is not activated by any $p > p_k^{(2)}$. Consider the set $B = \{p_{i+1}^{(2)}, p_{i+2}^{(2)}, \cdots, p_k^{(2)}\}$. Let p be a prime number, $p > p_k^{(2)}$. We have that $\mathcal{S}_{\mathcal{R},B}(p)$ is empty by the construction of B; then we can extend B appending p, let $B_1 = B \cup \{p\}$. Suppose that there exists a finite $j \leq 2$, such that in $B_{j-1} = B \cup \{p_1, p_2, \cdots p_{j-1}\}$ we have $\mathcal{S}_{\mathcal{R},B_{i-1}}(p_i)$ are empty and $\mathcal{S}_{\mathcal{R},B_j}(p) \neq \emptyset$, for every prime $p_j > p_{j-1} > \cdots > p_1$ with $B_0 = B$ and $p_1 = p$. Then, by construction, we have that if (b, n, A) is such that $b^n - S_A = p_j, A \subseteq B_{j-1} \cup \{p_j\}$, then $p, p_2, \cdots, p_j$ must be in A. If we keep adding new primes p_{j+l}, with $l = 1, 2, \cdots$, then by **Lemma** 2.9 , there exists j_* such that $\pi(l) \ggg \mathcal{N}(l), \forall l \geq j_*$, then the number of activated restrictions in $\mathcal{S}$ grows faster than the number of perfect power numbers available.If we keep adding primes and activated restrictions, then by **Lemma** 2.10, there will be an interval $[x, 2x]$ where $|\mathcal{S}_{\mathcal{R},B_l}| > \mathcal{N}(x)$. Then, we know that every $S_A \in \mathcal{S}, \forall A \subseteq B_j, \forall j \geq 0$, and we notice that there are more active restrictions than perfect power numbers in the interval $[x, 2x]$, meaning that some restrictions are forced to form within ever smaller ranges. Formally speaking, there exists $s_0 \in \mathcal{S}$ and an infinite sequence $(s_k)_{k \in \mathbb{N}} \subset \mathcal{S}$, with $s_k \neq s_0$, such that

$$\lim_{k \to +\infty} |s_k - s_0| = 0$$

Then, s_0 is an accumulation point in $\mathcal{S}$. However, we have that $\mathcal{S}$ is discrete in $\mathcal{N}$, a contradiction ■

To conclude the contradiction, using this last result, we can find a prime number $p > p_k^{(2)}$ such that $b^n - S_A = p \Rightarrow 2, 17 \notin A$. Then, noticing that $17 + 19 = 36 = 6^2$ is the only activated restriction generated in the construction when considering $p_* = 19$ as a candidate for $p_4^{(2)}$, then if we define

$$\hat{S}^2 = S^2 \setminus \{17\} \cup \{19, p\}$$

we have that there are no activated restrictions in $\hat{S}^2$ and $|\hat{S}^2| > |S^2| = k$, a contradiction with the maximality of S^2. Then there exists a prime number p such that, starting from p, we have an infinite subset of prime numbers $S = S^p$ such that the sum of every finite subset of S is not a perfect power number, this finishes the proof ■

3 An extension

Let $T \subseteq \mathbb{N}$ an infinite set. Let $x \in \mathbb{N}$ and define $T(x)$ as the number of elements of T in $[1, x]$..

Then, maybe a more interesting question is if the following is true:

Question : Let $B \subset \mathbb{N}$ be an **allowed Set**, i.e., an infinite set in $\mathbb{N}$. Let $F \subset \mathbb{N}$ a **forbidden set**, i.e. an infinite set $F \neq B$ such that for $x \in \mathbb{N}, F(x) = o(B(x))$. Then, does there exist an algorithmic construction that allows us to find an infinite set $S \subseteq B$ such that no sum of subsets of size at least $z^k_{B,F}$ of S belongs to F at the k-th step of the algorithm, with

$$1 \leq z^k_{B,F} \leq z_{B,F}, \forall k \geq 1$$
$$z_{B,F} = \sup_{k \geq 1} z^k_{B,F} < +\infty?$$

Notice that in the original case $z_{\mathcal{B},\mathcal{N}} = z^k_{B,\mathcal{N}} = 1, \forall k \geq 1$, with $\mathcal{B} = \mathcal{P}$

Notice that we could extend the question to other sets (ex. $\mathbb{Z}$), or other sets in higher dimensions, (ex. $\mathbb{N}^d, \mathbb{Z}^d, d \geq 2$), and even with groups or semi-groups, the geometric or topological setting would be crucial.