Proc. Indian Acad. Sci. (Math. Sci.) (2019) 129:84
https://doi.org/10.1007/s12044-019-0512-x

Sparse subsets of the natural numbers and Euler's totient function

MITHUN KUMAR DAS, PRAMOD EYYUNNI
and BHUWANESH RAO PATIL*

Harish-Chandra Research Institute, HBNI, Chhatnag Road, Jhunsi, Allahabad 211 019, India
*Corresponding author.
E-mail: das.mithun3@gmail.com; pramodeyy@gmail.com;
bhuwanesh1989@gmail.com

MS received 3 July 2018; revised 4 November 2018; accepted 25 December 2018; published online 16 August 2019

Abstract. In this article, we investigate sparse subsets of the natural numbers and study the sparseness of some sets associated to the Euler's totient function ϕ via the property of 'Banach density'. These sets related to the totient function are defined as follows: $V := \phi(\mathbb{N})$ and $N_i := \{N_i(m)\colon m \in V\}$ for $i = 1, 2, 3$, where $N_1(m) = \max\{x \in \mathbb{N}\colon \phi(x) \le m\}$, $N_2(m) = \max(\phi^{-1}(m))$ and $N_3(m) = \min(\phi^{-1}(m))$ for $m \in V$. Masser and Shiu (*Pacific J. Math.* **121(2)** (1986) 407–426) called the elements of N_1 as 'sparsely totient numbers' and constructed an infinite family of these numbers. Here we construct several infinite families of numbers in $N_2 \setminus N_1$ and an infinite family of composite numbers in N_3. We also study (i) the ratio $\frac{N_2(m)}{N_3(m)}$ which is linked to the Carmichael's conjecture, namely, $|\phi^{-1}(m)| \ge 2$ for all $m \in V$, and (ii) arithmetic and geometric progressions in N_2 and N_3. Finally, using the above sets associated to the totient function, we generate an infinite class of subsets of $\mathbb{N}$, each with asymptotic density zero and containing arbitrarily long arithmetic progressions.

Keywords. Euler's function; sparsely totient numbers; Banach density.

2010 Mathematics Subject Classification. 11A25, 11B05, 11B25, 11B75.

1. Introduction

Euler's totient function $\phi(n)$, which enumerates the number of positive integers which are co-prime to and less than or equal to n, is a classical arithmetical function. It is a well known fact that the number of solutions to the equation $\phi(x) = m$ is finite for each $m \in \mathbb{N}$ ($\mathbb{N}$ is the set of positive integers). It is natural, then, to ask the following questions:

(i) For a given $m \in \mathbb{N}$, what is the largest integer n such that $\phi(n) \le m$?
(ii) What are the largest and the smallest integers satisfying $\phi(x) = m$?

We denote the set $\{x : \phi(x) = m\}$ by $\phi^{-1}(m)$ and the image of ϕ by V, i.e. $V = \{\phi(m) : m \in \mathbb{N}\}$. The elements of V are called *totients*. For $m \in V$, we define the following quantities with the above questions in mind:

$$N_1(m) = \max\{x : \phi(x) \le m\},$$
$$N_2(m) = \max(\phi^{-1}(m)),$$
$$N_3(m) = \min(\phi^{-1}(m)),$$
$$N_i = \{N_i(m) : m \in V\} \text{ for } i = 1, 2, 3.$$

Note that $N_2(m)$, $N_3(m)$ are defined only on V, whereas $N_1(m)$ can be defined on the whole of $\mathbb{N}$. But this does not contribute any new elements to the image N_1 of $N_1(m)$, since $N_1(m) = N_1(m-1)$ if $m \notin V$. Hence, from here on, we study $N_1(m)$ only for $m \in V$. In 1986, Masser and Shiu [10] studied many properties of N_1 and called its elements as 'sparsely totient numbers'. They gave the following criteria to find examples of sparsely totient numbers.

PROPOSITION 1 [10]

Let $(p_i)_{i=1}^{\infty}$ be the enumeration of the primes in ascending order. Suppose $k \ge 2$, $d \ge 1$ and $l \ge 0$ satisfy conditions $d < p_{k+1} - 1$ and $d(p_{k+l} - 1) < (d+1)(p_k - 1)$. Then $dp_1 \cdots p_{k-1} p_{k+l}$ is a sparsely totient number.

They also found some nice patterns among sparsely totient numbers.

PROPOSITION 2 [10]

For $n \in N_1$, let n' represent the smallest sparsely totient number greater than n. Then
(*i*) $\frac{n'}{n} \to 1$ *as* $n \to \infty$ *and* $n \in N_1$.
(*ii*) *For a given prime* p, $\exists\, m(p) \in \mathbb{N}$ *such that* $N_1(m) \equiv 0 \pmod{p}$ $\forall\, m \ge m(p)$.

This proposition suggests that the distribution of elements of N_1 may be very sparse. To study the notion of sparseness of a subset of integers, we use properties like asymptotic density or Banach density. The asymptotic density gives the fraction of the number of elements of a set in $\mathbb{N}$ whereas Banach density gives an idea about how locally sparse or dense a set is. For example, the set $\cup_{n\in\mathbb{N}}[10^n, 10^n + n]$ has asymptotic density zero but it has, in fact, a maximum Banach density of 1. The notion of Banach density will be defined in section 2. The first theorem in this paper measures the densities of sets V, N_i, etc.

Theorem 3.
(*i*) *The Banach density of V and N_1 is zero.*
(*ii*) *If $f : V \to \mathbb{N}$ is such that $f(m) \in \phi^{-1}(m)$, then the asymptotic density of $f(V)$ is* 0. *In particular, the asymptotic density of N_2 and N_3 is zero.*

More generally, we also look at the Banach density of sets that are images of injective-increasing functions on $\mathbb{N}$.

Theorem 4. *Let $A \subset \mathbb{N}$. Suppose $f : A \to \mathbb{N}$ is an injective and increasing function.*
(*a*) *If the function $\frac{f(n)}{n}$ is increasing on A and $\lim_{\substack{n\to\infty \\ n\in A}} \frac{f(n)}{n} = \infty$, then the Banach density of $f(A)$ is zero.*

(*b*) *For $A = \mathbb{N}$, if there exists $n_0 \in \mathbb{N}$ and positive absolute constants c_1 and c_2 such that $c_1 n \le f(n) \le c_2 n$ for $n \ge n_0$, then the Banach density of $f(\mathbb{N})$ is positive.*

In Theorem 4(a), the hypothesis 'increasing' for $\frac{f(n)}{n}$ is only a sufficient condition. For instance, if $BN_1 = \{m \in V : N_1(m) = N_2(m)\}$, then the function $h\colon BN_1 \to N_1$ given by $h(m) = N_1(m)$ does not satisfy this condition, but nevertheless, the Banach density of N_1 is zero .

In section 3, we observe that $N_2 \supset N_1$ and $N_3 \supset \mathbb{P} \setminus \{2\}$, where $\mathbb{P}$ denotes the set of primes. Therefore, we look for infinite families of elements in $N_2 \setminus N_1$ and an infinite family of composite numbers in N_3. This leads to our next theorem.

For $r, r_1, r_2 \in \mathbb{N}$ and a prime $q \equiv 3 \pmod 4$, define

$$R(r_1, r_2) := 2.3^{r_1}.5^{r_2}, \ K_{q,r} := 2q^{r+1},$$
$$k_{q,r} := \begin{cases} q^r(q-1)+1 & \text{if } q^r(q-1)+1 \in \mathbb{P} \\ q^{r+1} & \text{otherwise.} \end{cases}$$

A prime of the form $2^{2^l} + 1$ with $l \in \mathbb{N} \cup \{0\}$ is called a Fermat prime. We denote the j-th Fermat prime by F_j. The only known Fermat primes are

$$F_1 = 3, \ F_2 = 5, \ F_3 = 17, \ F_4 = 257 \text{ and } F_5 = 65537.$$

The existence of F_6 is not known.

Theorem 5. *$\mathcal{K}_{\max}$, $\mathcal{R}$ and $\mathcal{F}$ are infinite subsets of N_2 in which only finitely many elements are in N_1. Moreover, $\mathcal{K}_{\min}$ is an infinite subset of N_3 in which infinitely many elements are composite. Here,*

$$\mathcal{K}_{\max} = \{K_{q,r} : q \equiv 3 \pmod 4, r \in \mathbb{N}\},$$
$$\mathcal{K}_{\min} = \{k_{q,r} : q \equiv 3 \pmod 4, r \in \mathbb{N}\},$$
$$\mathcal{R} = \{R(r_1, r_2)\colon r_1, r_2 \in \mathbb{N}, r_2 > 2\},$$
$$\mathcal{F} = \left\{ 2^a \prod_{i=1}^{k} F_i : k \in H;\ a \le \log_2(F_{k+1} - 1) \right.$$
$$\left. \text{if } F_{k+1} \text{ exists and } a \in \mathbb{N} \text{ otherwise} \right\},$$

where F_j denotes the j-th Fermat prime and $H = \{k \in \mathbb{N}\colon F_k \text{ exists}\}$.

From this theorem, we observe that N_2 contains infinitely many elements divisible by powers of n, where n is 2, 5 or a prime q with $q \equiv 3 \pmod 4$. The infinite family $\mathcal{F}$ of elements of N_2 shows the importance of Fermat primes to generate many elements in N_2. The family $\mathcal{K}_{\min}$ of elements of N_3 shows that N_3 contains infinitely many elements divisible by powers of some prime q, where $q \equiv 3 \pmod 4$.

Though we have given examples of infinite families in N_2 and N_3, there may still be other elements in these sets. So, we give bounds for general $N_2(m)$ and $N_3(m)$ in the case when $m \not\equiv 0 \pmod 8$. We also study properties of the ratio $\frac{N_2(m)}{N_3(m)}$ and geometric progressions contained inside N_2 and N_3. The ratios $\frac{N_2(m)}{N_3(m)}$ are important in the sense that the statement '$\frac{N_2(m)}{N_3(m)} > 1$ for each $m \in \mathbb{V}$' is equivalent to Carmichael's conjecture which asserts that $|\phi^{-1}(m)| \geq 2$ for all $m \in V$.

Theorem 6. *Let $m \in V$.*

(*i*) *If $m \equiv 2 \pmod 4$ or $4 \pmod 8$, then $m < N_3(m) < 2m$ and $2m < N_2(m) < 4m$.*
(*ii*) *There exist infinitely many m such that $\frac{N_2(m)}{N_3(m)} = 2$. Further, if $m \equiv 2 \pmod 4$, then $2 \leq \frac{N_2(m)}{N_3(m)} \leq 3$.*
(*iii*) *N_2 and N_3 contain an infinite geometric progression.*

In section 4, we discuss about the existence of arithmetic progressions in infinite subsets of natural numbers. The famous Szemerédi's theorem [7] gives a sufficient condition for the existence of arbitrarily long arithmetic progressions in a subset of integers, namely, a positive asymptotic density. But this is not a necessary condition. Therefore we give a class of subsets of the integers having zero asymptotic density and containing arbitrarily long arithmetic progressions. These sets are formed by taking exactly one element from each pre-image $\phi^{-1}(m)$, $m \in V$. Theorem 7 below follows as a consequence by using results due to Green and Tao [6] and Erdős [3, Theorem 4]. Therefore, we have as follows.

Theorem 7. *If $f : V \to \mathbb{N}$ is such that $f(m) \in \phi^{-1}(m)$, then $f(V)$ contains arbitrarily long arithmetic progressions.*

Indeed, we observe that these sets satisfy the hypothesis of the so-called Erdős–Turán conjecture [7, page 4] which asserts that if a set X of positive integers such that the sum of reciprocals of elements of X diverges, then X contains arbitrarily long arithmetic progressions.

Finally, in section 5, we pose some questions about elements of N_2 and Banach density of N_2 and N_3 arising from the present work.

We use the following notation in this paper. Let $\mathbb{N}$, $\mathbb{P}$, $\mathbb{R}^+$ and $\mathbb{Z}$ denote, respectively, the set of positive integers, the set of prime numbers, the set of positive real numbers and the set of integers. p, q will always represent prime numbers unless otherwise mentioned. We write $f(x) = o(g(x))$ if $\frac{f(x)}{g(x)} \to 0$ as $x \to \infty$. $[a, b]$ denotes the set $\{x \in \mathbb{N} : a \leq x \leq b\}$ and similarly for the sets $(a, b]$, $[a, b)$ and (a, b), and finally $W(x)$ denotes the set of prime divisors of x. By convention, we assume empty products and empty sums to take the values 1 and 0 respectively. By 'a divergent sequence (x_n)', we mean that $x_n \to \infty$ as $n \to \infty$.

2. Sparse subsets of natural numbers and sparsely totient numbers

It is well-known that the set of totients V is sparsely distributed, i.e., has asymptotic density zero (see, for example, [4] and references therein).

PROPOSITION 8 [4]

If $V(x)$ is the number of totients less than or equal to x, then

$$V(x) = \frac{x}{\log x} \exp\{(C + o(1))(\log\log\log x)^2\},$$

where $0.81 < C < 0.82$.

Here, we study the sparseness of the set of totients V, the set of sparsely totient numbers N_1 and other subsets of natural numbers using a generalized version of asymptotic density called the Banach density. We will define Banach density using Følner sequences.

DEFINITION 9 (Følner sequence)

A Følner sequence in a countable commutative semigroup $(G, +)$ is a sequence $(F_n)_{n\in\mathbb{N}}$ of finite subsets of G such that for all $g \in G$,

$$\lim_{n\to\infty} \frac{|(g + F_n) \cap F_n|}{|F_n|} = 1.$$

Example 10. In the semigroup $(\mathbb{N}, +)$, let $F_n = [\alpha_n, \beta_n]$ with $\beta_n - \alpha_n \to \infty$ as $n \to \infty$, then $(F_n)_{n\in\mathbb{N}}$ is a Følner sequence.

DEFINITION 11 (Density of a subset of $\mathbb{N}$)

Let $(F_n)_{n\in\mathbb{N}}$ be a Følner sequence in $\mathbb{N}$ and $A \subset \mathbb{N}$. Then the upper density of A with respect to the Følner sequence $(F_n)_{n\in\mathbb{N}}$ is defined by

$$\overline{d}_{F_n}(A) = \limsup_{n\to\infty} \frac{|F_n \cap A|}{|F_n|}$$

and the lower density of A with respect to the Følner sequence $(F_n)_{n\in\mathbb{N}}$ is defined by

$$\underline{d}_{F_n}(A) = \liminf_{n\to\infty} \frac{|F_n \cap A|}{|F_n|}.$$

If the upper density and the lower density are equal, then we say that the density of A with respect to the Følner sequence exists and it equals

$$d_{F_n}(A) = \lim_{n\to\infty} \frac{|F_n \cap A|}{|F_n|}.$$

DEFINITION 12 (Asymptotic density)

The density with respect to the Følner sequence $([1, n])_{n\in\mathbb{N}}$ is called asymptotic density. In this case, the upper asymptotic density, the lower asymptotic density and the density of a subset A are denoted by $\overline{d}(A)$, $\underline{d}(A)$ and $d(A)$ respectively.

DEFINITION 13 (Banach density)

The Banach density $d^*(A)$ of $A \subset \mathbb{N}$ is defined by

$$d^*(A) = \sup\left\{\overline{d}_{F_n}(A) \colon (F_n)_{n\in\mathbb{N}} \text{ is a Følner sequence in } \mathbb{N}\right\}.$$

Example 14. Banach density of the set of primes is zero (see [5, p. 194]).

Using $F_n = [1, n]$ for all $n \in \mathbb{N}$ in the following proposition, one can observe that the Banach density of a subset of $\mathbb{N}$ is equal to density of that subset with respect to the Følner sequence $([t_n + 1, t_n + n])_{n\in\mathbb{N}}$ for some sequence $(t_n)_{n\in\mathbb{N}}$ in $\mathbb{N}$. Therefore, it is enough to consider Følner sequences formed by intervals in $\mathbb{N}$ to evaluate Banach density.

PROPOSITION 15 (p. 418, [1])

Given a subset A of $\mathbb{N}$ *and any Følner sequence* $(F_n)_{n\in\mathbb{N}}$, *there is a sequence* $(t_n)_{n\in\mathbb{N}}$ *such that*

$$d^*(A) = d_{(F_n+t_n)}(A).$$

2.1 *Proof of Theorem* 3

We now show that the Banach density of the set of totients V and that of the set of sparsely totient numbers N_1 is zero. For this, we start with some necessary lemmas.

Lemma 16. *If* $0 < \alpha < \infty$ *and* $x \geq 3$, *then* $\log(1 + \alpha) + \log x \leq (1 + \alpha) \log x$.

Proof. It is enough to show that $f_\alpha(x) = x^\alpha - \alpha - 1 \geq 0$ for $x \geq 3,\ 0 < \alpha < \infty$. Since $f_\alpha(x)$ is increasing in x, it suffices to show that $f_\alpha(3) \geq 0$. Note that $g(\alpha) = f_\alpha(3)$ is a strictly increasing function in α and moreover, $g(0) = 0$. So $g(\alpha) > 0$ for $\alpha > 0$. □

Lemma 17. *If* $0 < \alpha \leq 1$ *and* $z \in \mathbb{R}^+$, *then*

$$\frac{\exp((1+\alpha)^2 z) - \exp(z)}{\alpha} \leq \exp(4z) - 1.$$

Proof. By the definition of the exponential function, we have

$$\frac{\exp((1+\alpha)^2 z) - \exp(z)}{\alpha} = \lim_{k\to\infty} \sum_{n=0}^{k} \left(\frac{(1+\alpha)^{2n} - 1}{\alpha} \right) \frac{z^n}{n!}.$$

Applying the binomial theorem, we get

$$\begin{aligned} \frac{\exp((1+\alpha)^2 z) - \exp(z)}{\alpha} &= \lim_{k\to\infty} \sum_{n=1}^{k} \left(\sum_{m=1}^{2n} \binom{2n}{m} \alpha^{m-1} \right) \frac{z^n}{n!} \\ &\leq \lim_{k\to\infty} \sum_{n=1}^{k} \left(\sum_{m=1}^{2n} \binom{2n}{m} \right) \frac{z^n}{n!}, \end{aligned}$$

as $\alpha \leq 1$. Again, referring to the binomial theorem, one has

$$\frac{\exp((1+\alpha)^2 z) - \exp(z)}{\alpha} \leq \lim_{k\to\infty} \sum_{n=1}^{k} \frac{(4z)^n}{n!} = \exp(4z) - 1.$$

□

Lemma 18. *Suppose that $(F_n)_{n\in\mathbb{N}}$ is a Følner sequence on $\mathbb{N}$ defined by*

$$F_n = (x_n, x_n(1+\alpha_n)],$$

where $(x_n)_{n=1}^{\infty}$ is a sequence in $\mathbb{N}$ and $(\alpha_n)_{n=1}^{\infty}$ is a sequence of positive reals such that $x_n \to \infty$ and $\alpha_n \to 0$ as $n \to \infty$. Then there exist $n_0 \in \mathbb{N}$ and $0 < k < 1$ such that

$$\frac{|V \cap F_n|}{|F_n|} \leq \frac{2}{(\log x_n)^{1-k}} \quad \forall\, n \geq n_0,$$

so that $\overline{d}_{F_n}(V) = 0$.

Proof. Since $x_n \to \infty$ and $\alpha_n \to 0$ as $n \to \infty$, we can choose $n_0 \in \mathbb{N}$ such that $\log\log\log x_n > 5$ and $0 < \alpha_n < 1$ for all $n \geq n_0$. Applying Lemma 16 for each $n \geq n_0$, we get

$$\begin{aligned} \log\log\log((1+\alpha_n)x_n) &\leq \log\log((1+\alpha_n)\log x_n) \\ &\leq \log((1+\alpha_n)\log\log x_n) \\ &\leq (1+\alpha_n)\log\log\log x_n. \end{aligned} \tag{1}$$

Using the estimate of $V(x)$ from Proposition 8 (with the same constant C appearing there), equation (1) and setting $z_n = (C + o(1))(\log\log\log x_n)^2$ for each $n \geq n_0$, we get that for $n \geq n_0$,

$$\begin{aligned} \frac{|V \cap F_n|}{|F_n|} &= \frac{|V \cap [1, x_n(1+\alpha_n)]| - |V \cap [1, x_n]|}{|F_n|} \\ &\leq \frac{1}{\alpha_n \log x_n}((1+\alpha_n)\exp((1+\alpha_n)^2 z_n) - \exp(z_n)). \end{aligned}$$

Since $\alpha_n < 1$ and $z_n \in \mathbb{R}^+$, then Lemma 17 gives

$$\begin{aligned} \frac{|V \cap F_n|}{|F_n|} &\leq \frac{1}{\log x_n}(2\exp(4z_n) - 1) \\ &\leq \frac{2}{\log x_n}(\exp(4(C+o(1))(\log\log\log x_n)^2)). \end{aligned}$$

Applying $4u^2 < e^u$ for all $u > 5$ with $u = \log\log\log x_n$, we get

$$\frac{|V \cap F_n|}{|F_n|} \leq \frac{2}{\log x_n}(\exp((C+o(1))\log\log x_n)) \leq \frac{2(\log x_n)^{C+o(1)}}{\log x_n}.$$

Since $C < 1$, choose a sufficiently large n_1 with $n_1 \geq n_0$ such that $|C + o(1)| < k$ for all $n \geq n_1$ and for some $0 < k < 1$. Therefore,

$$\frac{|V \cap F_n|}{|F_n|} \leq \frac{2}{(\log x_n)^{1-k}} \quad \forall\, n \geq n_1$$

and hence $\bar{d}_{F_n}(V) = 0$. □

Lemma 19. *Suppose that $(F_n)_{n\in\mathbb{N}}$ is a Følner sequence on $\mathbb{N}$ defined by*

$$F_n = (x_n, x_n(1 + \alpha_n)],$$

where $(x_n)_{n=1}^{\infty}$ is a sequence in $\mathbb{N}$ and $(\alpha_n)_{n=1}^{\infty}$ is a sequence of positive reals such that $\alpha_n > \alpha_0 > 0$ for each $n \in \mathbb{N}$. Then $\bar{d}_{F_n}(V) = 0$.

Proof. Since $|F_n| = x_n\alpha_n \to \infty$ as $n \to \infty$, we can choose $n_0 \in \mathbb{N}$ such that $\log\log\log(1 + \alpha_n)x_n > 0$ for all $n \geq n_0$. For $n \geq n_0$, we get

$$\frac{|V \cap F_n|}{|F_n|} = \frac{|V \cap [1, x_n(1 + \alpha_n)]| - |V \cap [1, x_n]|}{|F_n|} \leq \frac{|V \cap [1, x_n(1 + \alpha_n)]|}{|F_n|}.$$

Using the estimate of $V(x)$ from Proposition 8 (with the same constant C appearing there), we get

$$\frac{|V \cap F_n|}{|F_n|} \leq \frac{(1 + \alpha_n)}{\alpha_n \log((1 + \alpha_n)x_n)}(\exp((C + o(1))(\log\log\log(1 + \alpha_n)x_n)^2)).$$

Since $\alpha_n > \alpha_0$ for each $n \in \mathbb{N}$, applying the inequality $y^2 < e^y$ for $y > 0$ gives us

$$\frac{|V \cap F_n|}{|F_n|} \leq \frac{(1 + \alpha_0)}{\alpha_0}\left(\frac{(\log((1 + \alpha_n)x_n))^{C+o(1)}}{\log((1 + \alpha_n)x_n)}\right) \to 0 \text{ as } n \to \infty,$$

as $C < 1$. Hence $\bar{d}_{F_n}(V) = 0$. □

PROPOSITION 20

The Banach density of the set of totients is zero.

Proof. Let $(F_n)_{n\in\mathbb{N}}$ be a Følner sequence on $\mathbb{N}$ defined by $F_n = (x_n, x_n(1 + \alpha_n)]$, where $x_n \in \mathbb{N}$, $\alpha_n \in \mathbb{R}^+$. To prove that the Banach density of V is zero, it is enough to show that $\bar{d}_{F_n}(V) = 0$ in the following cases:

Case A. $\alpha_n \to 0$ and $x_n \to \infty$ as $n \to \infty$.
Case B. There exists $\alpha_0 > 0$ such that $\alpha_n > \alpha_0$ for each $n \in \mathbb{N}$.

For Case A, the result follows from Lemma 18 and similarly, Lemma 19 covers Case B. □

Next, we proceed to prove that the Banach density of N_1 is zero.

Lemma 21. *Suppose $A, B \subset \mathbb{N}$ and $g\colon A \to B$ is an injective map satisfying $g(x) \leq x$ for all $x \in A$. Let $(F_n)_{n\in\mathbb{N}}$ be a Følner sequence in $\mathbb{N}$ such that $F_n = (a_n, x_n]$ and $(a_n)_{n=1}^{\infty}$ is bounded. Then $\bar{d}_{F_n}(B) = 0$ implies $\bar{d}_{F_n}(A) = 0$.*

Proof. Since $g\colon A \to B$ is an injective map and $g(x) \leq x$ for all $x \in A$, we have $g : F_n \cap A \to g(A) \cap [1, x_n]$ is injective for each $n \in \mathbb{N}$. It follows that $|F_n \cap A| \leq |g(A) \cap [1, x_n]|$ for all $n \in \mathbb{N}$. Since $g(A) \cap [1, x_n] \subset (g(A) \cap F_n) \cup [1, a_n]$, we get that $|F_n \cap A| \leq |g(A) \cap F_n| + |[1, a_n]|$. Therefore,

$$\frac{|F_n \cap A|}{|F_n|} \leq \frac{|F_n \cap g(A)|}{|F_n|} + \frac{a_n}{|F_n|} \leq \frac{|F_n \cap g(A)|}{|F_n|} + \frac{a}{|F_n|},$$

where a is an upper bound of the sequence $(a_n)_{n=1}^{\infty}$. Since $g(A) \subset B$ and $\bar{d}_{F_n}(B) = 0$, we conclude that $\bar{d}_{F_n}(A) = 0$. □

COROLLARY 22

Let $(F_n)_{n\in\mathbb{N}}$ be a Følner sequence on $\mathbb{N}$ defined by $F_n = (a_n, x_n]$, where $(a_n)_{n=1}^{\infty}$ is bounded. If $f\colon V \to \mathbb{N}$ is such that $f(m) \in \phi^{-1}(m)$, then $\bar{d}_{F_n}(f(V)) = 0$. In particular, the asymptotic density of N_1, N_2 and N_3 is zero.

Proof. Consider $g\colon f(V) \to V$ defined by $g(n) = \phi(n)$. This is an injective map satisfying $g(x) \leq x$ $\forall x \in f(V)$. Since $d^*(V) = 0$, by Proposition 20, it follows that $\bar{d}_{F_n}(f(V)) = 0$, by applying Lemma 21. In particular, $\bar{d}_{F_n}(N_i) = 0$ for $i = 2, 3$. Also, $N_1 \subset N_2$ so that $\bar{d}_{F_n}(N_1) = 0$. □

PROPOSITION 23

The Banach density of N_1 is zero.

Proof. Let $(F_n)_{n\in\mathbb{N}}$ be a Følner sequence on $\mathbb{N}$ defined by $F_n = (x_n, x_n + y_n]$, where $(x_n)_{n=1}^{\infty}$ and $(y_n)_{n=1}^{\infty}$ are sequences in $\mathbb{N}$ with $y_n \to \infty$. To show that the Banach density of N_1 is zero, it is enough to prove that $\bar{d}_{F_n}(N_1) = 0$ for the two cases when (i) $(x_n)_{n=1}^{\infty}$ is bounded, and (ii) $(x_n)_{n=1}^{\infty}$ is divergent.

If $(x_n)_{n=1}^{\infty}$ is a bounded sequence, then Corollary 22 establishes that $\bar{d}_{F_n}(N_1) = 0$. So we can assume that $(x_n)_{n=1}^{\infty}$ is a divergent sequence. If p is a prime number, then Proposition 2(ii) gives the existence of an element $m_0(p) \in \mathbb{N}$ such that $N_1(m) \equiv 0 \pmod{p}$ for each $m \geq m_0(p)$. Since $(x_n)_{n=1}^{\infty}$ is divergent, we can choose $n_0(p) \in \mathbb{N}$ such that $x_n > N_1(m_0(p))$ for each $n > n_0(p)$. Then F_n contains at most $\frac{|F_n|}{p} + 1$ elements of N_1 for $n > n_0(p)$. So, given a prime p, there exists $n_0(p) \in \mathbb{N}$ such that

$$n > n_0(p) \Rightarrow \frac{|F_n \cap N_1|}{|F_n|} \leq \frac{1}{p} + \frac{1}{y_n}.$$

Since $y_n \to \infty$ as $n \to \infty$, this means that $\bar{d}_{F_n}(N_1) \leq \frac{1}{p}$. As this holds for each prime p, we conclude that $\bar{d}_{F_n}(N_1) = 0$ if $(x_n)_{n=1}^{\infty}$ is divergent. Therefore, $d^*(N_1) = 0$. □

Hence, the proof of Theorem 3 is complete by collecting Proposition 23, Proposition 20 and Corollary 22.

2.2 *Some criteria for sparse sets in* $\mathbb{N}$

We have studied the Banach densities of specific sets like V and N_1. Now, we are going to investigate the behavior of sparse sets which are the images of injective, increasing functions on $\mathbb{N}$. We proceed to the proof of Theorem 4.

Proof of Theorem 4(*a*). Let $(a, b]$ be an interval in $\mathbb{N}$ such that $a, b \in f(A)$. Since f is injective and increasing, it follows that $\{x \in A\colon f(x) \in (a, b]\} = \{x \in A\colon x \in (f^{-1}(a), f^{-1}(b)]\}$. Therefore,

$$\frac{|(a,b] \cap f(A)|}{b-a} \leq \frac{f^{-1}(b) - f^{-1}(a)}{b-a} = \frac{1}{b-a}\left(\frac{b}{m_b} - \frac{a}{m_a}\right),$$

where $m_a = \dfrac{a}{f^{-1}(a)}$ and $m_b = \dfrac{b}{f^{-1}(b)}$. Since the function $\dfrac{f(n)}{n}$ is increasing on A, we get $m_a \leq m_b$. It follows that

$$\frac{|(a,b] \cap f(A)|}{b-a} \leq \frac{1}{m_b}.$$

Suppose that $((a_n, b_n])_{n\in\mathbb{N}}$ is a Følner sequence with $|(a_n, b_n] \cap A| > 1$. Then for each $n \in \mathbb{N}$, there exist $a'_n, b'_n \in f(A)$ such that $(a_n, b_n] \cap f(A) = [a'_n, b'_n] \cap f(A)$ and $a_n < a'_n < b'_n \leq b_n$. So,

$$\limsup_{b_n - a_n \to \infty} \frac{|(a_n, b_n] \cap f(A)|}{b_n - a_n} \leq \limsup_{b_n - a_n \to \infty} \frac{|(a'_n, b'_n] \cap f(A)|}{b'_n - a'_n} \leq \limsup_{b_n - a_n \to \infty} \frac{1}{m_{b'_n}}.$$

Note that $b_n - a_n \to \infty \Rightarrow b_n \to \infty$. We claim that $b'_n \to \infty$. If not, there exists a subsequence $\{b'_{n_k}\}$ of $\{b'_n\}$ such that $b'_{n_k} \leq l$ for all $k \in \mathbb{N}$. By the definition of b'_n, we know that $(b'_{n_k}, b_{n_k}] \cap f(A) = \varnothing$ for each k. In particular, $(l, b_{n_k}] \cap f(A) = \varnothing$ for all k. Since $b_{n_k} \to \infty$ as $k \to \infty$, this implies that $f(x) \leq l$ for all $x \in A$. But this is a contradiction since f is strictly increasing and hence grows indefinitely. Thus, $b'_n \to \infty$. Therefore,

$$\limsup_{b_n - a_n \to \infty} \frac{|(a_n, b_n] \cap f(A)|}{b_n - a_n} \leq \limsup_{b'_n \to \infty} \frac{1}{m_{b'_n}} = 0,$$

since $(m_{b'_n})_{n=1}^{\infty}$ is a subsequence of the divergent sequence $\left(\frac{f(n)}{n}\right)_{n\in A}$.

Hence, $\bar{d}_{F_n}(f(A)) = 0$ for all Følner sequences $(F_n)_{n\in\mathbb{N}}$ with $F_n = (a_n, b_n]$. So, $d^*(f(A)) = 0$. □

Proof of Theorem 4(*b*). Choose $r \in \mathbb{N}\setminus\{1\}$ such that $c_2 < (r-1)c_1$ and consider the Følner sequence $(F_n)_{n\in\mathbb{N}}$ given by $F_n = (n, rn]$. Since $f(\mathbb{N})$ is an infinite set, there exists $k' \in \mathbb{N}$ that for each $n \geq k'$, one can choose integers $a_n, b_n \in f(\mathbb{N})$ such that $(a_n, b_n) \cap f(\mathbb{N}) =$

$(n, rn] \cap f(\mathbb{N})$ and $a_n \le n < rn \le b_n$. Since f is injective and increasing, it follows that $\{x \in \mathbb{N}\colon f(x) \in (a_n, b_n)\} = \{x \in \mathbb{N}\colon x \in (f^{-1}(a_n), f^{-1}(b_n))\}$. We now have

$$\frac{|F_n \cap f(\mathbb{N})|}{|F_n|} = \frac{f^{-1}(b_n) - f^{-1}(a_n) - 1}{(r-1)n}. \tag{2}$$

From the hypothesis, we have

$$c_1 x \le f(x) \le c_2 x \quad \forall\, x \ge n_0. \tag{3}$$

Choose an integer $k \ge k'$ such that $f^{-1}(a_n), f^{-1}(b_n) \ge n_0$ for all $n \ge k$. Inserting the values of $f^{-1}(a_n), f^{-1}(b_n)$ obtained from inequality (3) in equation (2) for $n \ge k$, we get

$$\begin{aligned}\frac{|F_n \cap f(\mathbb{N})|}{|F_n|} &\ge \frac{1}{(r-1)n}\left(\frac{b_n}{c_2} - \frac{a_n}{c_1}\right) - \frac{1}{(r-1)n}\\ &> \frac{1}{(r-1)n}\left(\frac{rn}{(r-1)c_1} - \frac{n}{c_1}\right)\\ &\quad - \frac{1}{(r-1)n} \quad (\text{since } c_2 < (r-1)c_1)\\ &= \frac{1}{c_1(r-1)^2} - \frac{1}{(r-1)n}.\end{aligned}$$

Therefore,

$$\limsup_{n\to\infty} \frac{|F_n \cap f(\mathbb{N})|}{|F_n|} \ge \frac{1}{c_1(r-1)^2}.$$

Hence, $\bar{d}_{F_n}(f(\mathbb{N})) > 0$ and therefore $d^*(f(\mathbb{N})) > 0$. □

We drop the 'increasing function' hypothesis on $\frac{f(n)}{n}$ in Theorem 4(a) and by the two examples given below, we show that the conclusion on Banach density may or may not hold.

Example 24. Given $k, l \ge 2$, we define a map $f_{k,l}\colon \mathbb{N} \to \mathbb{N}$ by

$$f_{k,l}(x) = \begin{cases} k^{2nl} + (l-1)x & \text{if } k^{2n} \le x < k^{2n+1},\\ x^l & \text{if } k^{2n+1} \le x < k^{2n+2}.\end{cases}$$

One can see that $f_{k,l}$ is injective and increasing. The function $\frac{f_{k,l}(n)}{n}$ is divergent but not increasing. Note that $f_{k,l}([k^{2n}, k^{2n+1}))$ is an arithmetic progression of length $k^{2n+1} - k^{2n}$ with a common difference $l-1$. Therefore, $f_{k,l}(\mathbb{N})$ contains arbitrarily long arithmetic progressions with a common difference $l-1$. Hence, it has positive Banach density.

The above example suggests that the 'increasing' hypothesis on $\frac{f(n)}{n}$ is necessary. However, this is not always the case as the next example shows.

Suppose $BN_1 = \{m \in V\colon\ N_1(m) = N_2(m)\}$. Then the function $h\colon BN_1 \to N_1$ defined by $h(m) = N_1(m)$, is both bijective and increasing. By a result of Sanna [11, Lemma 2.1] on the asymptotic of $N_1(m)$, it is readily seen that $\frac{h(m)}{m} = \frac{N_1(m)}{m} \to \infty$ as $m \to \infty$.

Lemma 25 [11, *Lemma* 2.1]. *$N_1(m) \sim e^{\gamma}\, m \log\log m$ as $m \to \infty$, where γ is the Euler–Mascheroni constant.*

The next proposition tells us that $\frac{h(m)}{m}$ is not an increasing function.

PROPOSITION 26

$\dfrac{h(n)}{n}\colon BN_1 \to N_1$ is not an increasing function.

Proof. For $p \in \mathbb{P}$, define

$$X_p = \prod_{q\in\mathbb{P}, q\le p} q.$$

Let p_1 and p_2 be two consecutive primes such that $3 < p_1 < p_2$. Let $a = X_{p_1}, b = \frac{X_{p_2}}{p_1}$, $M_a = \phi(a)$ and $M_b = \phi(b)$. Then by Proposition 1 and the definition of h, we get

$$h(M_a) = N_1(M_a) = a \quad \text{and} \quad h(M_b) = N_1(M_b) = b.$$

The proof of the equation $h(M_b) = b$ uses a result of Nagura which states that $(n, 1.2n) \cap \mathbb{P} \neq \emptyset$ for all $n > 25$. This gives

$$\frac{h(M_a)}{M_a} = \prod_{q\in\mathbb{P}, q\le p_1} \frac{q}{q-1} = \left(\frac{p_1}{p_1-1}\right)\left(\frac{p_2-1}{p_2}\right)\frac{h(M_b)}{M_b}.$$

Since $p_1 < p_2$, it follows that

$$\frac{h(M_a)}{M_a} > \frac{h(M_b)}{M_b} \quad \text{but } M_a = \frac{p_1-1}{p_2-1}M_b < M_b.$$

Therefore, $\frac{h(n)}{n}$ is not an increasing function. □

In contrast to $f_{k,l}$, though $\frac{h(n)}{n}$ is not increasing, we know that $d^*(h(BN_1)) = d^*(N_1) = 0$. Therefore, we observe that if we remove the condition of 'increasing map' on $\frac{f(n)}{n}$, then both the possibilities, namely Banach density is zero or positive may occur.

3. Explicit construction of elements of N_2 and N_3

3.1 *Proof of Theorem* 5

Now we move on to the study of N_2 and N_3. As we know that $N_1 \subsetneq N_2$, we give explicit examples of infinite families of elements in $N_2 \setminus N_1$. Since $\phi(p) = p - 1$ and

$\phi(p-1) < p-1$ for an odd prime p, this implies that $N_3(p-1) = p$. So, $\mathbb{P}\setminus\{2\} \subset N_3$. We are going to show that infinitely many composite numbers also lie in N_3. First, we give the following useful definitions.

DEFINITION 27 ($k_{q,r}$ and $K_{q,r}$)

For $q \in \mathbb{P}$ and $r, r_1, r_2 \in \mathbb{N}$, define

$$k_{q,r} : = \begin{cases} q^r(q-1)+1 & \text{if } q^r(q-1)+1 \text{ is a prime,} \\ q^{r+1} & \text{otherwise .} \end{cases}$$

$$K_{q,r} : = 2q^{r+1}, \qquad R(r_1, r_2) = 2\cdot 3^{r_1}\cdot 5^{r_2}.$$

The following lemma gives a description of the elements of $\phi^{-1}(m)$ for $m \equiv 2 \pmod 4$. This will be useful to construct families of elements in N_2 and N_3 as indicated above.

Lemma 28. *Let $A(m)$ denote the number of solutions to the equation $\phi(x) = m$. Suppose $m > 2$ and $m \equiv 2 \pmod 4$. Then*

(*i*) *every element of $\phi^{-1}(m)$ is of the form p^α or $2p^\alpha$, where $p \equiv 3 \pmod 4$;*
(*ii*) *$A(m) = 0, 2$ or 4;*
(*iii*) *if $A(m) = 2$, then $\phi^{-1}(m) = \{p^\alpha, 2p^\alpha\}$ for some $p \equiv 3 \pmod 4$, $\alpha \geq 1$ and if $A(m) = 4$, then $\phi^{-1}(m) = \{p^\beta, 2p^\beta, q, 2q\}$ for some $p, q \equiv 3 \pmod 4$, with $p < q$, $\beta > 1$.*

Proof. For the proof of (i) and (ii), see [9]. For (iii), if $p \in \mathbb{P}$, $p \equiv 3 \pmod 4$, then $\phi(p^\alpha) = \phi(2p^\alpha)$ for $\alpha \geq 0$. If $A(m) = 2$, then from (i), $\phi^{-1}(m) = \{p^\alpha, 2p^\alpha\}$ for some prime $p \equiv 3 \pmod 4$. On the other hand, if $A(m) = 4$, then $\phi^{-1}(m) = \{p^\beta, 2p^\beta, q^\gamma, 2q^\gamma\}$ for some $p, q \equiv 3 \pmod 4$ and $\beta, \gamma \geq 1$. Now, $p^\beta \neq q^\gamma$ and $\phi(p^\beta) = \phi(q^\gamma) \Rightarrow p \neq q$. Without loss of generality, let us assume that $p < q$. This means that $\beta > \gamma$. Now, $\phi(p^\beta) = \phi(q^\gamma) \Rightarrow p^{\beta-1}(p-1) = q^{\gamma-1}(q-1)$. If $\gamma > 1$, then it means that $q \mid (p-1)$, a contradiction to $p < q$. Thus, $\gamma = 1$.

Therefore, $\phi^{-1}(m) = \{p^\beta, 2p^\beta, q, 2q\}$ for some $p, q \equiv 3 \pmod 4$ with $p < q$, $\beta > 1$ in the case $A(m) = 4$. □

Lemma 29. *Let q be a prime greater than 7. Then there exists a unique odd integer $n \in \{q+2, q+4\}$ such that $n \equiv 0 \pmod 3$, $\gcd(n, q) = 1$ and $\phi(n) < q$.*

Proof. Since q is a prime and $q > 7$, one can choose the unique integer $n \in \{q+2, q+4\}$ such that $n \equiv 0 \pmod 3$. Since q is odd, $\gcd(q, n) = 1$. Let $n = 3^r l$ with $r, l \in \mathbb{N}$ and $3 \nmid l$. Hence $\phi(n) = 2\times 3^{r-1}\phi(l) \leq 2\times 3^{r-1} l = \frac{2n}{3} \leq \frac{2q+8}{3} < q$ if $q \geq 11$. □

PROPOSITION 30

Suppose that $r \in \mathbb{N}$ and q is a prime satisfying $q \equiv 3 \pmod 4$. Then

(*i*) *$N_2(q^r(q-1)) = K_{q,r}$ and $N_3(q^r(q-1)) = k_{q,r}$;*
(*ii*) *$K_{q,r} \notin N_1$ except when $(q, r) = (3, 1)$.*

Proof. Let $m = q^r(q-1)$. Then $q \equiv 3 \pmod 4 \Rightarrow m \equiv 2 \pmod 4$. Since $\phi(q^{r+1}) = q^r(q-1)$, it follows that $\phi^{-1}(m)$ is non-empty. Applying Lemma 28, we get $\phi^{-1}(m) = \{q_1^\alpha, 2q_1^\alpha\}$ or $\{q_2^\beta, 2q_2^\beta, q_3, 2q_3\}$, where $q_2 < q_3$, $m > 2$, $\alpha \geq 1$ and $\beta > 1$.

If $\phi^{-1}(m) = \{q_1^\alpha, 2q_1^\alpha\}$, then $q_1 = q$ and $\alpha = r+1$ as $\phi(q^{r+1}) = q^r(q-1) = \phi(2q^{r+1})$. Hence $N_2(q^r(q-1)) = K_{q,r}$ and $N_3(q^r(q-1)) = q^{r+1}$ in this case. Since $\phi^{-1}(m)$ does not contain primes, this means that $m+1 = q^r(q-1)+1$ is composite and hence $k_{q,r} = q^{r+1} = N_3(m)$.

If suppose $\phi^{-1}(m) = \{q_2^\beta, 2q_2^\beta, q_3, 2q_3\}$, where $q_2, q_3 \equiv 3 \pmod 4$, $q_2 < q_3$ and $\beta > 1$. If $q^r(q-1)+1$ is a prime, then $q^r(q-1)+1 \in \phi^{-1}(m)$. It follows that $N_3(q^r(q-1)) = k_{q,r}$ in this case. Since the only prime in $\phi^{-1}(m)$ is q_3, we get $q_3 = q^r(q-1)+1 = N_3(q^r(q-1))$. This means that $q_2^\beta > q_3$. Now, note that $q^{r+1} > q^r(q-1)+1$ and q^{r+1} is the only odd composite number in $\phi^{-1}(m)$. Thus, $q_2^\beta = q^{r+1}$, i.e. $q_2 = q$ and $\beta = r+1$. Therefore, $N_2(q^r(q-1)) = 2q^{r+1}$. On the other hand, if $q^r(q-1)+1$ is not a prime, then no element of $\phi^{-1}(m)$ can be prime. But this contradicts the fact that $q_3 \in \phi^{-1}(m)$. Therefore,

$$N_2(q^r(q-1)) = K_{q,r} \quad \text{and} \quad N_3(q^r(q-1)) = k_{q,r}.$$

Coming to the proof of (ii), if $q > 7$, then Lemma 29 gives the odd integer $n \in \{q+2, q+4\}$ such that $n \equiv 0 \pmod 3$, $\gcd(n,q) = 1$ and $\phi(n) < q$. Now we observe that $2nq^{r-1} > 2q^r$ but $\phi(2nq^{r-1}) \leq \phi(2q^r)$. Hence $2q^r \notin N_1$ for all $q > 7$. If $q = 7$, then $2 \times 3^2 \times 7^{r-1} > 2 \times 7^r$ but $\phi(2 \times 3^2 \times 7^{r-1}) \leq \phi(2 \times 7^r)$ and hence $2 \times 7^r \notin N_1$ for all $r \geq 1$. If $q = 5$, then $12 \times 5^{r-1} > 2 \times 5^r$ but $\phi(12 \times 5^{r-1}) \leq \phi(2 \times 5^r)$. Hence $2 \times 5^r \notin N_1$ for all $r \geq 1$. Since $\phi(20 \times 3^{r-2}) < \phi(2 \times 3^r)$ but $20 \times 3^{r-2} > 2 \times 3^r$ for $r \geq 3$, it follows that $2 \times 3^r \notin N_1$ for all $r \geq 3$. □

From Proposition 30, we see that $K_{q,r} \in N_2 \setminus N_1$ for all $r \geq 3$. So, for each $q \equiv 3 \pmod 4$, this gives an infinite family of elements in $N_2 \setminus N_1$. But the proposition does not ensure the presence of infinitely many composite numbers in N_3. For this, we require that $k_{q,r}$ is composite for infinitely many (q, r). If $q = 3$, note that $2 \cdot 3^r + 1$ is divisible by 5 when $r \equiv 3 \pmod 4$. In other words, $k_{3,r}$ is composite for infinitely many r. So, from Proposition 30, we see that N_3 contains infinitely many composite numbers.

Now, we give another infinite family of elements in $N_2 \setminus N_1$. First, we state some definitions, four preliminary lemmas and then prove the two main lemmas which together construct an infinite two-parameter family of elements in $N_2 \setminus N_1$.

DEFINITION 31 $(D(A, B))$

Let A and B be two finite subsets of $\mathbb{P}$. Then $D(A, B)$ is defined by

$$D(A, B) := \left(\prod_{q \in A} \frac{q-1}{q}\right)\left(\prod_{q \in B} \frac{q}{q-1}\right).$$

Lemma 32. *Suppose that* $y, x \in \mathbb{N} \setminus \{1\}$. *If* $\phi(y) \leq \phi(x)$ *and* $y > x$, *then* $D(W(y), W(x)) < 1$.

Proof. We know that

$$\phi(y) = y \prod_{q \in W(y)} \left(\frac{q-1}{q}\right) \quad \text{and} \quad \phi(x) = x \prod_{q \in W(x)} \left(\frac{q-1}{q}\right).$$

This gives

$$1 \geq \frac{\phi(y)}{\phi(x)} = \frac{yD(W(y), W(x))}{x} > D(W(y), W(x)),$$

since $\phi(y) \leq \phi(x)$ and $y > x$. □

Lemma 33. *Let A and B be two finite subsets of $\mathbb{P}$ such that $|B| \leq |A|$. If $\min(B \setminus A) > \max(A)$ or $B \subset A$, then $D(B, A) \geq 1$.*

Proof. If $B \subset A$, then $D(B, A) = \left(\prod_{q \in A \setminus B} \frac{q}{q-1}\right) \geq 1$. In the case when $B \not\subset A$, define an injective map $f : B \to A$ such that $f(x) = x$ for $x \in A \cap B$. If $\min(B \setminus A) > \max(A)$, it follows that $f(x) \leq x$ for all $x \in B$. Therefore,

$$D(B, A) \geq \prod_{x \in B} \left(\frac{(x-1)f(x)}{x(f(x)-1))}\right) = \prod_{x \in B} \left(\frac{xf(x) - f(x)}{xf(x) - x}\right) \geq 1.$$

□

Lemma 34. *Let $a, k \in \mathbb{N} \setminus \{1\}$ and $k \leq a$. Suppose that $x_1, x_2, \ldots, x_k$ are non-negative integers such that at least two of them are postive. Then*

$$\sum_{i=1}^{k} a^{x_i} \leq a^{x_1 + x_2 + \cdots + x_k}.$$

Proof. Since atleast two of the k integers $x_1, x_2, \ldots, x_k$ are positive, it follows that $a^{x_i} \leq a^{x_1 + x_2 + \cdots + x_k - 1}$ for all $i \in [1, k]$. Therefore,

$$\sum_{i=1}^{k} a^{x_i} \leq k a^{x_1 + x_2 + \cdots + x_k - 1} \leq a^{x_1 + x_2 + \cdots + x_k},$$

since $k \leq a$. □

Lemma 35. *Let $x, y, k \in \mathbb{N}$ such that $x, y, k \geq 2$ and $k \leq \min\{x, y\}$. Suppose that for each $i \leq k$, a_i and b_i are non-negative integers such that $a_i + b_i \neq 0$. If $a_1 + a_2 + \cdots + a_k = t$ and $b_1 + b_2 + \cdots + b_k = u$, then*

$$\sum_{i=1}^{k} x^{a_i} y^{b_i} \leq x^t y^u.$$

Proof. Let $k \geq 2$. If possible, suppose that both the sequences $(a_i)_{i=1}^k$, $(b_i)_{i=1}^k$ contain at most one positive integer. If $k \geq 3$, then there exists $j \in [1, k]$ such that $a_j + b_j = 0$, a contradiction. Hence, $k = 2$ and exactly one of $\{a_1, a_2\}$ and one of $\{b_1, b_2\}$ are positive with $a_i + b_i \neq 0$ for $i \in [1, 2]$. Therefore, we can assume $a_1, b_2 > 0$, without loss of generality. We need to show that

$$x^{a_1} + y^{b_2} \leq x^{a_1} y^{b_2}$$

in this case. Since $x, y \geq 2$ and $a_1, b_2 \in \mathbb{N}$, it is enough to show that $v + w \leq vw$ for $v, w \in \mathbb{N} \setminus \{1\}$. This happens iff $v \leq w(v-1)$ iff $v/(v-1) \leq w$ which is true since the left-hand side is not greater than 2.

On the other hand, if one of the sequences, say (a_i), has at least two positive elements, then by Lemma 34, we have

$$\sum_{i=1}^{k} x^{a_i} y^{b_i} \leq \left(\sum_{i=1}^{k} x^{a_i} \right) y^u \leq x^t y^u,$$

since $k \leq \min\{x, y\} \leq x$. □

DEFINITION 36 (Valuation)

Let p be a prime number. Then the p-valuation on the integers $\mathbb{Z}$ is the map $v_p \colon \mathbb{Z} \to \mathbb{N} \cup \{0, \infty\}$ defined by $v_p(0) = \infty$ and $v_p(n) = r$ for $n \neq 0$, where r is the largest non-negative integer such that $p^r \mid n$.

Lemma 37. *Suppose* $r_1, r_2, y \in \mathbb{N}$ *satisfy* $\phi(y) = \phi(R(r_1, r_2))$, $|W(y)| = 4$, $v_2(y) = 1$, $v_3(y) = 0$ *and* $v_5(y) = 0$. *Then* $y \leq R(r_1, r_2)$.

Proof. Since $|W(y)| = 4$, $v_2(y) = 1$, $v_3(y) = 0$ and $v_5(y) = 0$, we can write $y = 2\left(q_1^{v_{q_1}(y)} q_2^{v_{q_2}(y)} q_3^{v_{q_3}(y)}\right)$, where q_1, q_2, q_3 are distinct primes greater than 6 and $v_q(y) \geq 1$ for $q \in \{q_1, q_2, q_3\}$. Since $\phi(y) = \phi(R(r_1, r_2))$, it follows that $v_{q_1}(y) = v_{q_2}(y) = v_{q_3}(y) = 1$ and hence

$$\left(\frac{q_1 - 1}{2}\right)\left(\frac{q_2 - 1}{2}\right)\left(\frac{q_3 - 1}{2}\right) = 5^{r_2 - 1} 3^{r_1 - 1}.$$

Therefore, for each $i \in \{1, 2, 3\}$, we can write $q_i = 2 \cdot 3^{a_i} 5^{b_i} + 1$ such that

$$\begin{aligned} &a_1 + a_2 + a_3 = r_1 - 1,\ b_1 + b_2 + b_3 = r_2 - 1, \\ &a_1 + b_1 \neq 0,\ a_2 + b_2 \neq 0,\ a_3 + b_3 \neq 0, \\ &a_1, a_2, a_3, b_1, b_2, b_3 \geq 0. \end{aligned}$$

Therefore,

$$\begin{aligned} y &= 2(2\cdot 3^{a_1}5^{b_1}+1)(2\cdot 3^{a_2}5^{b_2}+1)(2\cdot 3^{a_3}5^{b_3}+1)\\ &= 2\left(2^3 3^{r_1-1}5^{r_2-1} + 2^2\left(\sum_{i=1}^{3} 3^{r_1-a_i-1}5^{r_2-b_i-1}\right) + 2\left(\sum_{i=1}^{3} 3^{a_i}5^{b_i}\right)+1\right)\\ &= 2\left(2^3 3^{r_1-1}5^{r_2-1} + 2^2 3^{r_1-1}5^{r_2-1}\left(\sum_{i=1}^{3} 3^{-a_i}5^{-b_i}\right) + 2\left(\sum_{i=1}^{3} 3^{a_i}5^{b_i}\right)+1\right). \end{aligned} \tag{4}$$

Since $a_i + b_i \geq 1$ for each $i = 1, 2, 3$, we have

$$\sum_{i=1}^{3} 3^{-a_i}5^{-b_i} \leq \sum_{i=1}^{3} 3^{-(a_i+b_i)} \leq 1. \tag{5}$$

Applying Lemma 35, we have

$$\sum_{i=1}^{3} 3^{a_i}5^{b_i} \leq 3^{r_1-1}5^{r_2-1}. \tag{6}$$

Inserting inequalities (5) and (6) in the right-hand side of (4), we get

$$\begin{aligned} y &\leq 2\left(2^3 3^{r_1-1}5^{r_2-1} + 2^2 3^{r_1-1}5^{r_2-1} + 2\cdot 3^{r_1-1}5^{r_2-1} + 1\right)\\ &\leq 2\cdot 3^{r_1-1}5^{r_2-1}(8+4+2+1) = R(r_1, r_2). \end{aligned}$$

□

Lemma 38. *Suppose $r_1, r_2, y \in \mathbb{N}$ and $r_2 > 2$ satisfy $\phi(y) = \phi(R(r_1, r_2))$, $|W(y)| = 4$, $v_2(y) = 1$, $v_3(y) \geq 1$ and $v_5(y) = 0$. Then $y \leq R(r_1, r_2)$.*

Proof. Since $|W(y)| = 4$, $v_2(y) = 1$, $v_3(y) \geq 1$ and $v_5(y) = 0$, we can write $y = 2\big(3^{v_3(y)} q_1^{v_{q_1}(y)} q_2^{v_{q_2}(y)}\big)$, where q_1, q_2 are distinct primes greater than 6 and $v_q(y) \geq 1$ for $q \in \{q_1, q_2, 3\}$. Since $\phi(y) = \phi(R(r_1, r_2))$, it follows that $v_{q_1}(y) = v_{q_2}(y) = 1$, $v_3(y) \leq r_1$, and hence

$$\left(\frac{q_1-1}{2}\right)\left(\frac{q_2-1}{2}\right) = 5^{r_2-1}3^{r_1-v_3(y)}.$$

Therefore, we can write $q_1 = 2\cdot 3^{a_1}5^{a_2}+1$ and $q_2 = 2\cdot 3^{b_1}5^{b_2}+1$ such that $a_1 + b_1 = r_1 - v_3(y)$, $a_2 + b_2 = r_2 - 1$, $a_1 + a_2 \neq 0$, $b_1 + b_2 \neq 0$ and $a_1, a_2, b_1, b_2 \geq 0$.

Therefore,

$$\begin{aligned} y &= 2\cdot 3^{v_3(y)}\big(2\cdot 3^{a_1}5^{a_2}+1\big)\big(2\cdot 3^{b_1}5^{b_2}+1\big)\\ &= 2\big(2^2\cdot 3^{a_1+b_1+v_3(y)}5^{a_2+b_2} + 2(3^{a_1+v_3(y)}5^{a_2} + 3^{b_1+v_3(y)}5^{b_2}\big) + 3^{v_3(y)}). \end{aligned}$$

Inserting $a_1 + b_1 = r_1 - v_3(y)$ and $a_2 + b_2 = r_2 - 1$ in the above equation, we get

$$y = 2\big(2^2 \cdot 3^{r_1} 5^{r_2-1} + 2\big(3^{r_1-b_1} 5^{a_2} + 3^{r_1-a_1} 5^{b_2}\big) + 3^{v_3(y)}\big). \tag{7}$$

From Lemma 34, we have

$$5^{a_2} + 5^{b_2} \leq \begin{cases} 5^{r_2-1} & \text{if } a_2, b_2 > 0, \\ 1 + 5^{r_2-1} & \text{else.} \end{cases} \tag{8}$$

We are now going to consider the following cases depending on the value of a_1 and b_1.

Case 1. If a_1 and b_1 are positive, then $a_1 + b_1 \geq 2$. Using this along with the conditions $a_1 + b_1 + v_3(y) = r_1$ and $v_3(y) \geq 1$, we get $r_1 \geq 3$ and $v_3(y) \leq r_1 - 2$. Applying these in the right-hand side of equation (7), we have

$$y \leq 2\big(2^2 \cdot 3^{r_1} 5^{r_2-1} + 2 \cdot 3^{r_1-1}\big(5^{a_2} + 5^{b_2}\big) + 3^{r_1-2}\big).$$

Inserting the value of $5^{a_2} + 5^{b_2}$ from (8) in the above inequality, we have

$$\begin{aligned} y &\leq 2 \cdot 3^{r_1-2}\big(36 \cdot 5^{r_2-1} + 6\big(1 + 5^{r_2-1}\big) + 1\big) \\ &\leq 2 \cdot 3^{r_1-2}\big(45 \cdot 5^{r_2-1} - 3 \cdot 5^{r_2-1} + 7\big) \\ &\leq 2 \cdot 3^{r_1} 5^{r_2} = R(r_1, r_2) \ \text{ for each } r_2 \geq 2. \end{aligned}$$

Case 2. If $a_1 = 0$ and $b_1 = 0$, then $v_3(y) = r_1$ due to the fact that $a_1 + b_1 = r_1 - v_3(y)$. Applying these in equation (7), we have

$$y \leq 2 \cdot 3^{r_1}\big(4 \cdot 5^{r_2-1} + 2\big(5^{a_2} + 5^{b_2}\big) + 1\big).$$

Since $a_1 + a_2 \neq 0$ and $b_1 + b_2 \neq 0$, it follows that $a_2, b_2 > 0$. Hence $a_2, b_2 \leq r_2 - 2$ because $a_2 + b_2 = r_2 - 1$. Using this in the previous inequality gives us

$$\begin{aligned} y &\leq 2 \cdot 3^{r_1}\big(4 \cdot 5^{r_2-1} + 4 \cdot 5^{r_2-2} + 1\big) \\ &\leq 2 \cdot 3^{r_1}\big(5^{r_2} - 5^{r_2-2} + 1\big) \leq R(r_1, r_2), \end{aligned}$$

since $r_2 = 1 + a_2 + b_2 \geq 3$.

Case 3. The remaining cases are in which exactly one of a_1 and b_1 is zero. Without loss of generality, assume that $a_1 = 0$ and $b_1 \neq 0$.

If $b_2 \geq 1$, we get $r_2 \geq 2$ and $a_2 \leq r_2 - 2$ because $a_2 + b_2 = r_2 - 1$. Since $a_1 = 0$ and $a_1 + a_2 \neq 0$, we have $a_2 \geq 1$, and hence $b_2 \leq r_2 - 2$. Equation (7) gives

$$\begin{aligned} y &\leq 2\big(2^2 \cdot 3^{r_1} 5^{r_2-1} + 2\big(3^{r_1-1} 5^{r_2-2} + 3^{r_1} 5^{r_2-2}\big) + 3^{r_1}\big) \\ &\leq 2 \cdot 3^{r_1}\big(2^2 \cdot 5^{r_2-1} + 4 \cdot 5^{r_2-2} + 1\big) \\ &\leq 2 \cdot 3^{r_1}\big(2^2 \cdot 5^{r_2-1} + 5^{r_2-1} - 5^{r_2-2} + 1\big) \\ &\leq 2 \cdot 3^{r_1}\big(5^{r_2} - 5^{r_2-2} + 1\big) \\ &\leq 2 \cdot 3^{r_1} 5^{r_2}, \ \text{ since } r_2 \geq 2. \end{aligned}$$

Now, in the case $b_2 = 0$, we have $a_2 = r_2 - 1$. Since $a_1 = 0$ and $a_2 + a_1 \neq 0$, it follows that $a_2 \geq 1$ and hence $r_2 \geq 2$. Then equation (7) gives

$$\begin{aligned} y &\leq 2\big(2^2 \cdot 3^{r_1} 5^{r_2-1} + 2\big(3^{r_1-1} 5^{r_2-1} + 3^{r_1}\big) + 3^{r_1}\big) \\ &\leq 2\big(3^{r_1} 5^{r_2} - 3^{r_1-1} 5^{r_2-1} + 3^{r_1+1}\big) \leq R(r_1, r_2), \ \text{since } r_2 > 2. \end{aligned}$$

Hence $y \leq R(r_1, r_2)$ for $r_1, r_2 \in \mathbb{N},\ r_2 > 2$. □

PROPOSITION 39

$R(r_1, r_2)$ lies in N_2 for each $r_1, r_2 \in \mathbb{N},\ r_2 > 2$.

Proof. Let y be an even number such that $\phi(y) = \phi(R(r_1, r_2))$. Since $v_2(\phi(R(r_1, r_2))) = 3$, it follows that $v_2(\phi(y)) = 3$. This means that y can have atmost 4 prime factors. If $|W(y)| \leq 3$, then

$$D(W(y), W(R(r_1, r_2))) \geq 1,$$

by Lemma 33. This gives $y \leq R(r_1, r_2)$, by Lemma 32. Now we consider the case $|W(y)| = 4$. Since $v_2(\phi(y)) = 3$, it follows in this case that $v_2(y) = 1$.

Suppose $v_5(y) \geq 1$, then $v_2(\phi(y)) \geq 4$. It follows that $v_2(\phi(R(r_1, r_2))) \geq 4$ which contradicts the fact that $v_2(\phi(R(r_1, r_2))) = 3$. Therefore, $v_5(y) = 0$.

If $v_3(y) \geq 1$, Lemma 38 ensures that $y \leq R(r_1, r_2)$. If $v_3(y) = 0$, then Lemma 37 gives $y \leq R(r_1, r_2)$. Therefore $R(r_1, r_2) \in N_2$ in any case. □

Remark 40. From Proposition 2(ii), we get that any element in N_1, all of whose prime factors are less than some prime p, has bounded exponents for its prime factors. But, as seen above from Proposition 39, this is not the case for elements in N_2. In fact, $R(r_1, r_2) \in N_2$ for $r_1, r_2 \in \mathbb{N},\ r_2 > 2$.

So, this raises the following question: For a given odd prime p, do there exist non-negative integers d_q corresponding to each odd prime $q < p$ such that $r_q > d_q$ for each $q < p \Rightarrow 2\prod_{2<q<p} q^{r_q} \in N_2$? The numbers $R(r_1, r_2)$ and $K_{3,r}$ answer this question in the affirmative for $p = 7$ and $p = 5$ respectively.

Now we are going to give another infinite family of elements in N_2 in which the odd prime factors of elements are Fermat primes.

Lemma 41. *If $\phi(x) = 2^r$ for some $x, r \in \mathbb{N}$, then there exist $b, n \in \mathbb{N} \cup \{0\}$ and a sequence of distinct Fermat primes $(F_{i_j})_{j=1}^{j=n}$ such that*

$$x = 2^b \prod_{j=1}^{n} F_{i_j}.$$

Proof. We observe that if $\phi(x) = 2^r$ and if an odd prime $q \mid x$, then $(q-1) \mid 2^r$ which implies that q is of the form $2^l + 1$ for some $l \in \mathbb{N}$. But it is well-known that if $2^l + 1$ is a prime, then $l = 2^\alpha$ for some $\alpha \geq 0$ (see [8, Theorem 17]). Hence, $q = 2^{2^\alpha} + 1$, a Fermat prime. Also, $v_q(x) = 1$ for each such $q \mid x$. If not, then $q \mid \phi(x) = 2^r$, a contradiction since q is odd. Therefore, x will be of the form $x = 2^b \prod_{j=1}^{n} F_{i_j}$, where $i_j \in \mathbb{N},\ b, n \in \mathbb{N} \cup \{0\}$. □

PROPOSITION 42

Let F_j denote the j-th Fermat prime for $j \in \mathbb{N}$. Suppose $F_1, F_2, \ldots, F_k$ exist. If F_{k+1} also exists, then $2^a F \in N_2$, where $F = \prod_{i=1}^{k} F_i$ and $1 \leq a \leq \log_2(F_{k+1} - 1)$. If F_{k+1} does not exist, then $2^a F \in N_2$ for each $a \in \mathbb{N}$.

Proof. Define $y := 2^a \prod_{i=1}^{k} F_i$ with $a \in \mathbb{N}$. To prove $y \in N_2$, it is enough to show that if x is any even integer satisfying $\phi(y) = \phi(x)$, then $x \leq y$. This can be observed using the fact that elements of N_2 are even.

Let x be an even integer satisfying $\phi(x) = \phi(y)$. Since $\phi(y) = 2^r$ for some $r \in \mathbb{N}$, it follows that $\phi(x) = 2^r$ for some $r \in \mathbb{N}$. Then Lemma 41 gives $x = 2^b \prod_{j=1}^{n} F_{i_j}$ for some $b \in \mathbb{N}$ and $n \in \mathbb{N} \cup \{0\}$. Since $\phi(x) = \phi(y)$, we have

$$a = b + \sum_{j=1}^{n} \log_2(F_{i_j} - 1) - \sum_{j=1}^{k} \log_2(F_j - 1). \tag{9}$$

If F_{k+1} exists, then

$$\begin{aligned} |W(x)| > |W(y)| &\Rightarrow \sum_{j=1}^{n} \log_2(F_{i_j} - 1) - \sum_{j=1}^{k} \log_2(F_j - 1) \geq \log_2(F_{k+1} - 1) \\ &\Rightarrow a \geq b + \log_2(F_{k+1} - 1) \ \text{ by equation (9)} \\ &\Rightarrow a \geq 1 + \log_2(F_{k+1} - 1) \ \text{ since } b \in \mathbb{N}. \end{aligned}$$

Therefore,

$$\begin{aligned} a \leq \log_2(F_{k+1} - 1) &\Rightarrow |W(x)| \leq |W(y)| \\ &\Rightarrow D(W(x), W(y)) \geq 1 \ \text{ by using Lemma 33} \\ &\Rightarrow x \leq y \ \text{ by Lemma 32}. \end{aligned}$$

On the other hand, if F_{k+1} does not exist, then $W(x) \subset W(y)$ for each $a \in \mathbb{N}$. It follows that $D(W(x), W(y)) \geq 1$, using Lemma 33. Then Lemma 32 gives $x \leq y$. □

Only five Fermat primes are known till date. From the above proposition, one can see that there exist elements in $N_2 \setminus N_1$ which are divisible by arbitrarily large powers of 2. In all the earlier results, the elements obtained were divisible by 2 but not by 4.

COROLLARY 43

For a positive integer r, there exist infinitely many integers l such that $l \equiv 0 \pmod{2^r}$ and $l \in N_2 \setminus N_1$.

DEFINITION 44

Let F_j denote the j-th Fermat prime for $j \in \mathbb{N}$ and let $H = \{k \in \mathbb{N}\colon F_k \text{ exists}\}$. Define

$$\begin{aligned}
\mathcal{K}_{\max} &= \{K_{q,r} : q \equiv 3 \pmod 4, r \in \mathbb{N}\},\\
\mathcal{K}_{\min} &= \{k_{q,r} : q \equiv 3 \pmod 4, r \in \mathbb{N}\},\\
\mathcal{R} &= \{R(r_1, r_2)\colon r_1, r_2 \in \mathbb{N}, r_2 > 2\},\\
\mathcal{F} &= \Bigg\{ 2^a \prod_{i=1}^{k} F_i : k \in H; a \le \log_2(F_{k+1} - 1)\\
&\qquad \text{if } F_{k+1} \text{ exists and } a \in \mathbb{N} \text{ otherwise} \Bigg\}.
\end{aligned}$$

By collecting Propositions 30, 39 and 42 , we get that (i) $\mathcal{K}_{\max}, \mathcal{R}$ and $\mathcal{F}$ are infinite subsets of N_2 and (ii) $\mathcal{K}_{\min}$ is an infinite subset of N_3 in which infinitely many elements are composite. Proposition 2(ii) gives that only finitely many elements of $\mathcal{K}_{\max}, \mathcal{R}$ and $\mathcal{F}$ belong to N_1.

Theorem 5. *$\mathcal{K}_{\max}, \mathcal{R}$ and $\mathcal{F}$ are infinite subsets of N_2 in which only finitely many elements are in N_1. $\mathcal{K}_{\min}$ is an infinite subset of N_3 in which infinitely many elements are composite.*

3.2 *Proof of Theorem* 6

In the previous results, we looked at several families of elements in N_2 and N_3. Now, we would like to compare the values of $N_2(m)$ and $N_3(m)$. In the following proposition, we are going to give upper and lower bounds for $N_2(m)$ and $N_3(m)$ and we also look at the ratio $N_2(m)/N_3(m)$.

Lemma 45. *Let m be an odd integer. If u is an odd integer satisfying $\phi(u) = 4m$, then*

$$u = \frac{(2z_1 + 1)(2z_2 + 1)}{z_1 z_2} m \quad or \quad \frac{(4z_3 + 1)}{z_3} m,$$

where $z_1 z_2 | m$, $z_3 | m$, $z_1 \ne z_2$ and $2z_1 + 1, 2z_2 + 1, 4z_3 + 1$ are primes. Also, $4m < u \le 7m$.

Proof. Any odd integer u satisfying $\phi(u) = 4m$ can have at most two prime factors. If u has two distinct prime factors q_1 and q_2 such that $q_1 < q_2$ and $q_1, q_2 \equiv 3 \pmod 4$ (since $v_2(\phi(u)) = 2$), then

$$u \left(\frac{(q_1 - 1)(q_2 - 1)}{q_1 q_2} \right) = 4m, \text{ i.e., } u = \frac{4 q_1 q_2}{(q_1 - 1)(q_2 - 1)} m.$$

Since u, q_1, q_2 and m are all odd, we have $q_1 = 2z_1 + 1$, $q_2 = 2z_2 + 1$ for some odd integers z_1, z_2 with $z_1 < z_2$. Moreover, $q_1 q_2 | u \Rightarrow (q_1 - 1)(q_2 - 1) \mid \phi(u) = 4m$, i.e., $z_1 z_2 \mid m$. Using this in the value of u, we have

$$u = \frac{(2z_1+1)(2z_2+1)}{z_1 z_2} m,$$

where $z_1 z_2$ divides m. Clearly, $u > 4m$ and it can take a maximum value of $7m$ if $z_1 = 1, z_2 = 3$. Therefore, $u \leq 7m$. If u has only one prime factor q with $q \equiv 5 \pmod 8$ (since $v_2(\phi(u)) = 2$), then

$$u = \frac{4qm}{q-1}.$$

Now, $q \equiv 5 \pmod 8$, and so $q = 4z_3 + 1$ for some odd integer z_3. Therefore,

$$u = \frac{(4z_3+1)}{z_3} m.$$

Note that $z_3, 4z_3 + 1$ are co-prime integers and hence $z_3 \mid m$. Clearly $u > 4m$ and it can take a maximum value of $5m$ if $z_3 = 1$. □

Now we proceed to prove Theorem 6(i) and 6 (ii).

Proof of Theorem 6(i). If $m \equiv 2 \pmod 4$ and $\phi(x) = m$ is solvable, then by Lemma 28, $m = q^r(q-1)$ for some $q \equiv 3 \pmod 4$. Since $q^r(q-1)+1 \leq q^{r+1} \leq \frac{3m}{2}$, we have $m < N_3(m) \leq \frac{3m}{2}$ and $2m < N_2(m) \leq 3m$, by Proposition 30. If $m \equiv 4 \pmod 8$, firstly, note that the proposition is true for $m = 4$, since $N_3(4) = 5$ and $N_2(4) = 12$. So, we can assume that $m \geq 12$. If $\phi(x) = m$ for $m \equiv 4 \pmod 8$, then $v_2(x) \leq 2$.

Suppose that there exists an integer z such that $v_2(z) = 2$ and $\phi(z) = m$, where $m = 4m_0,\ m_0$ being an odd integer. If $z = 4y$, then y is an odd integer satisfying $\phi(x) = 2m_0$. Then by Lemma 28, $y = p^\alpha$ for some $\alpha \geq 1$ and prime $p \equiv 3 \pmod 4$. Therefore, $z = 4p^\alpha$. If $p > 3$, then $3p^\alpha < z < 6p^\alpha$ and $\phi(3p^\alpha) = \phi(z) = \phi(6p^\alpha)$. If $p = 3$ and $\alpha > 1$, then $7 \times 3^{\alpha-1} < z < 14 \times 3^{\alpha-1}$ and again they have the same ϕ value. Hence, if $\phi(z) = m,\ m \equiv 4 \pmod 8$ and $v_2(z) = 2$, then $z \notin \{N_2(m), N_3(m)\}$. Moreover, an odd integer $l \in \phi^{-1}(m)$ iff $2l \in \phi^{-1}(m)$. Therefore, $v_2(N_3(m)) = 0$ and $v_2(N_2(m)) = 1$. Now, note that $N_3(m)$ and $\frac{N_2(m)}{2}$ are odd integers satisfying $\phi(x) = m = 4m_1$, where m_1 is odd. Therefore, by Lemma 45, we have $m < N_3(m)$ and $\frac{N_2(m)}{2} \leq \frac{7m}{4}$ and the result follows. □

Proof of Theorem 6(ii). By Lemma 28, if $m \in V$, then $m \equiv 2 \pmod 4 \iff m = q^r(q-1)$ for some prime $q \equiv 3 \pmod 4,\ r \geq 0$. Now, by Proposition 30, if $q^r(q-1)+1$ is composite, then

$$\frac{N_2(q^r(q-1))}{N_3(q^r(q-1))} = 2.$$

Else, if $q^r(q-1)+1$ is prime, then

$$\frac{N_2(q^r(q-1))}{N_3(q^r(q-1))} = \frac{2q^{r+1}}{q^r(q-1)+1} = \frac{2q}{q-1+\frac{1}{q^r}}.$$

It is readily seen that the rightmost quantity lies between 2 and 3 since $r \geq 0,\ q \geq 3$. □

To prove Theorem 6(iii), we give examples of infinite length geometric progressions in N_2 and N_3. For each prime $q \equiv 3 \pmod 4$, note that $\{K_{q,r}\}_{r\in\mathbb{N}}$ is a geometric progression in N_2 with common ratio q. Also, we see that $\{R(r_0, r)\}_{r=3}^{\infty}$ is a geometric progression in N_2 with common ratio 5.

Now, we turn our attention to geometric progressions in N_3. We construct an infinite geometric progression in N_3 with the help of the following lemma.

Lemma 46. *Let q be a prime satisfying $q \equiv 3$ or $7 \pmod{20}$. Then the set $\{r : q^r(q-1)+1$ is composite$\}$ contains an infinite arithmetic progression.*

Proof. Suppose that $q \equiv 3 \pmod{20}$. We observe that $q^r(q-1)+1$ is divisible by 5 for $r \equiv 3 \pmod 4$. Now, if $q \equiv 7 \pmod{20}$, we see that $q^r(q-1)+1$ is divisible by 5 for $r \equiv 2 \pmod 4$. So, in any case, the set $\{r : q^r(q-1)+1$ is composite$\}$ contains an infinite arithmetic progression. □

If $q \equiv 3$ or $7 \pmod{20}$, then $k_{q,r} \in N_3$ for each $r \in \mathbb{N}$. As the set $S = \{r : q^r(q-1)+1$ is composite$\}$ contains an infinite arithmetic progression, the set $\{k_{q,r} : r \in S\}$ contains an infinite geometric progression. So, corresponding to each such q, there is an infinite geometric progression. This implies an infinite family of such geometric progressions in N_3 due to the following result of Dirichlet [8, Theorem 15, page 16].

PROPOSITION 47 (Dirichlet's theorem)

Suppose $(a, q) = 1$. Then there are infinitely many primes p satisfying $p \equiv a \pmod q$.

4. Arithmetic progressions in sparse sets

Szemerédi's theorem assures the existence of arbitrarily long arithmetic progressions in a set having positive asymptotic density. But the converse is not necessarily true. For example, the set of prime numbers has zero asymptotic density but contains arbitrarily long arithmetic progressions, as proved by Green and Tao [6].

DEFINITION 48

Let $A \subset \mathbb{P}$. Then $Rd(A) = \limsup_{N\to\infty} \frac{|A\cap[1,N]|}{\pi(N)}$ defines the relative upper density of A with respect to $\mathbb{P}$, where $\pi(N)$ denotes the number of primes less than or equal to N.

PROPOSITION 49 [6]

Let A be any subset of the prime numbers of positive relative upper density $Rd(A)$. Then A contains infinitely many arithmetic progressions of length k for all k.

Let $f : V \to \mathbb{N}$ be such that $f(m) \in \phi^{-1}(m)$. By Corollary 22, $\bar{d}(f(V)) = 0$. We observe that it satisfies the hypothesis of the following famous conjecture of Erdős and Turán.

Conjecture 1 (*Erdős and Turán*). If $A \subset \mathbb{N}$ is such that $\sum_{n \in A} n^{-1}$ diverges, then A contains arbitrarily long arithmetic progressions.

PROPOSITION 50

Let $f : V \to \mathbb{N}$ be such that $f(m) \in \phi^{-1}(m)$. Then $\sum_{m \in V} \frac{1}{f(m)}$ diverges.

Proof. We have

$$\sum_{n \in N_2} \frac{1}{n} \le \sum_{m \in V} \frac{1}{f(m)} \le \sum_{n \in N_3} \frac{1}{n}.$$

If $m \equiv 2 \pmod 4$, then $N_2(m) \le 3N_3(m)$, by Proposition 6(ii). So, it follows that

$$\sum_{m \in V} \frac{1}{N_2(m)} \ge \sum_{\substack{m \equiv 2 \pmod 4 \\ m \in V}} \frac{1}{3N_3(m)} \ge \sum_{p \equiv 3 \pmod 4} \frac{1}{3p},$$

since N_3 contains each odd prime. The right most series is divergent (see [2, Chapter 4]) and the result follows. □

Hence we may expect arbitrarily long arithmetic progressions in $f(V)$. Indeed, the following result due to Erdős [3, p. 15] and Proposition 49 confirm our intuition.

PROPOSITION 51 [3]

Suppose $m \in V$ with $|\phi^{-1}(m)| = k$ for $k \ge 2$. Then there exists a set $P \subset \mathbb{P}$ such that $Rd(P) > 0$ and for each $p \in P$, $\phi^{-1}(m(p-1)) = p\phi^{-1}(m)$.

Proof of Theorem 7. Consider the set $V_1 = \{m \in V : |\phi^{-1}(m)| = 2\}$ and let $m' \in V_1$. Then, by Proposition 51, there exists $P_1 \subset \mathbb{P}$ such that $Rd(P_1) > 0$ and $m'(p-1) \in V_1$ for each $p \in P_1$. Now, consider the sets $P_2 = \{p \in P_1 : pN_2(m') \in f(V)\}$ and $P_3 = \{p \in P_1 : pN_3(m') \in f(V)\}$. By the definition of the set $f(V)$, we have $P_2 \cap P_3 = \varnothing$ and $P_2 \cup P_3 = P_1$. Therefore, at least one of the sets P_2 or P_3 has positive relative density in $\mathbb{P}$ and thus, by Proposition 49, contains arbitrarily long arithmetic progressions. Therefore, one of the subsets $N_2(m')P_2$ or $N_2(m')P_3$ of $f(V)$ contains arbitrarily long arithmetic progressions and the result follows. □

5. Questions

As discussed in Remark 40, we raise the following question about elements in N_2.

Question 1. For a given odd prime p, do there exist non-negative integers d_q for each prime $q < p$ such that $r_q > d_q$ for each prime $q < p \Rightarrow 2\prod_{2<q<p} q^{r_q} \in N_2$?

We have observed that the Banach densities of V and N_1 are zero. Also, the asymptotic density $\bar{d}(N_2)$ of N_2 is zero. Since $N_1 \subset N_2$, $d^*(N_1) = 0$, and there is a bijection

$f_{\max} : V \to N_2$ given by $f_{\max}(m) = N_2(m)$, it may seem that $d^*(N_2) = 0$. But, on the other hand, consider the function $f : \mathbb{N}^2 \to \mathbb{N}$ (where $\mathbb{N}^2$ is the set of square numbers $\{1, 4, 9, 16, \ldots\}$) defined by $f(y) = \sqrt{y}$, $\sqrt{y}$ being the unique positive square root of y. By Theorem 4(a), we see that $d^*(\mathbb{N}^2) = 0$. So f is a bijection from a zero Banach density set to a set with positive Banach density. Now, the function $f_{\max}$ is not increasing and hence, it suggests that N_2 may have a positive Banach density. So, it is interesting to ask the following question.

Question 2. What is the Banach density of N_2 and N_3?

In fact, except for Følner sequences of type $(a_n, a_n(1+\alpha_n)]$, where $\alpha_n \to 0,\ a_n \to \infty$, and $\alpha_n a_n \to \infty$, one can see that the upper density with respect to other Følner sequnces is zero.

Acknowledgements

The authors would like to thank the Department of Atomic Energy, Government of India for the financial support. In addition, the research of the second and third authors (PE and BRP) was also supported by the 'INFOSYS scholarship for senior students'. They would also like to thank Harish-Chandra Research Institute for the excellent facilities. They thank their supervisors Prof. D S Ramana and Prof. Gyan Prakash for their comments on the final draft of this paper. They sincerely thank K Mallesham for his insightful remarks and constant encouragement, and Samrat Kadge for help with the computation and programming. They would also like to thank the referee for a very careful reading of this paper and for his/her suggestions.

References

[1] Beiglböck M, Bergelson V and Fish A, Sumset phenomenon in countable amenable groups, *Adv. Math.* **223(2)** (2010) 416–432
[2] Davenport H, Multiplicative number theory, third edition, revised and with a preface by Hugh L Montgomery, Graduate Texts in Mathematics, 74 (2000) (Springer-Verlag)
[3] Erdős P, Some remarks on Euler's ϕ function, *Acta Arith.* **4** (1958) 10–19
[4] Ford K, The distribution of totients, *Ramanujan J.* **2(1–2)** (1998) 67–151, cited version available online as arXiv:1104.3264v1
[5] Geroldinger A and Ruzsa I Z, Combinatorial number theory and additive group theory, Advanced Courses in Mathematics (2009) (CRM Barcelona: Birkhäuser Verlag)
[6] Green B and Tao T, The primes contain arbitrarily long arithmetic progressions, *Ann. Math.* **167(2)** (2008) 481–547
[7] Gowers W T, A new proof of Szemerédi's theorem, *Geom. Funct. Anal.* **11(3)** (2001) 465–588
[8] Hardy G H and Wright E M, An Introduction to the Theory of Numbers, 6th ed., revised by D R Heath-Brown and J H Silverman, with a foreword by Andrew Wiles (2008) (Oxford: Oxford University Press) xxii+621 pp.
[9] Klee V L Jr, On the equation $\phi(x) = 2m$, *Amer. Math. Monthly* **53** (1946) 327–328
[10] Masser D W and Shiu P, On sparsely totient numbers, *Pacific J. Math.* **121(2)** (1986) 407–426
[11] Sanna C, On the sum of digits of the factorial, *J. Number Theory* **147** (2015) 836–841

COMMUNICATING EDITOR: Sanoli Gun