On modular k-free sets

Victor Lambert
Centre de Mathmatiques Laurent Schwartz
École Polytechnique
Palaiseau, France
victor.lambert@polytechnique.edu

Submitted: Sep 24, 2014; Accepted: Apr 17, 2015; Published: Apr 29, 2015
Mathematics Subject Classifications: 05D05, 11B75, 11P99

Abstract

Let n and k be integers. A set $A \subset \mathbb{Z}/n\mathbb{Z}$ is k-free if for all x in A, $kx \notin A$. We determine the maximal cardinality of such a set when k and n are coprime. We also study several particular cases and we propose an efficient algorithm for solving the general case. We finally give the asymptotic behaviour of the minimal size of a k-free set in $[\![1, n]\!]$ which is maximal for inclusion.

1 Introduction

Let $k \geqslant 1$ be an integer. A set $A \subset \mathbb{N}$ is said to be k-free if $x \neq ky$ for all x, y in A. Wang first investigated in 1989 the problem of 2-free sets in the integers and, using elementary tools, he proved in [8] that the maximal density of a 2-free set in $[\![1, n]\!] := \{1, \dots, n\}$ is $2/3$. More recently, Wakeham and Wood studied in [7] a generalisation of 2-free sets into $\{a, b\}$-multiplicative sets ($ax \neq by$ for all $x, y \in A$). Notice that k-free sets are the particular case of $\{1, k\}$-multiplicative sets. They studied this problem through graph theory to get the maximal size of such a set. In particular, they showed that the maximal density of a k-free set in $[\![1, n]\!]$ is $k/(k + 1)$.

Beyond their own interest, k-free sets are useful for the study of k-fold Sidon sets. Those sets were first introduced by Lazebnik and Verstraëte in [3] through a work on the generalize Turán number.

Definition 1. *A set $A \subset \mathbb{Z}$ is a k-fold Sidon set if A has only trivial solutions to each equation of the form $c_1x_1+c_2x_2+c_3x_3+c_4x_4 = 0$ where $0 \leqslant |c_i| \leqslant k$, and $c_1+c_2+c_3+c_4 = 0$.*

In this definition, up to reordering c_i's, we consider solutions as "trivial" in the following cases :

(i) $\{x, x, x, x\}$ is always a trivial solution,

(ii) if $c_1 = c_2 = -c_3 = -c_4$, $\{x, y, y, x\}$ is a trivial solution,

(iii) if $c_1 = -c_3$ and $c_2 = -c_4$, $\{x, y, x, y\}$ is a trivial solution.

A 1-fold Sidon set is a Sidon set in the usual sense ($x_1 + x_2 = x_3 + x_4$ has only trivial solutions). If we denote by $D^*(A) = \{a_1 - a_2, a_1 \neq a_2 \in A\}$ the set of differences from A, without 0, a 2-fold Sidon set A is a Sidon set which has also the property that $D^*(A)$ is a 2-free set. More generally, for a k-fold Sidon set A, $D^*(A)$ is a k'-free set, for each $k' \leqslant k$. Using only this fact, Cilleruelo and Timmons proved in [2] that for any integer $k \geqslant 1$, a k-fold Sidon set $A \subset [\![0, n]\!]$ has at most $(n/k)^{1/2} + O((nk)^{1/4})$ elements.

We only know that the main term $(n/k)^{1/2}$ is optimal for $k = 1$. Indeed, Sidon sets have been widely studied (see [4] for a survey) and there exist three constructions of maximal Sidon sets in $\mathbb{Z}/n\mathbb{Z}$ for some n. Bose and Chowla proved in [1] the existence of a Sidon set of size $q+1$ in $\mathbb{Z}/\left(q^2+q+1\right)\mathbb{Z}$ (Singer's sets, see also [6]) and q in $\mathbb{Z}/\left(q^2-1\right)\mathbb{Z}$ (Bose's sets) where q is a power of a prime. Ruzsa also made an optimal construction in [5] for $\mathbb{Z}/\left(p^2-p\right)\mathbb{Z}$ where p is a prime number. For $k = 2$, if $n = 2^{2^t+1} + 2^t + 1$ with t a positive integer, we can extract (see [3]) from a Singer's set a 2-fold Sidon set in $\mathbb{Z}/n\mathbb{Z}$ of size

$$|A| \geqslant \frac{n^{1/2}}{2} - 3.$$

For $k \geqslant 3$, we do not even know if there exists a constant $c_k > 0$ such that for all integers $n \geqslant 1$, there is a k-fold Sidon set $A \subset [\![0, n]\!]$ with $|A| \geqslant c_k n^{1/2}$.

In all these problems, we see that it is important and useful to study the case of modular sets. In this paper, we will study k-free sets in $\mathbb{Z}/n\mathbb{Z}$. Notice that we cover the case of $\{a, b\}$-multiplicative set in $\mathbb{Z}/n\mathbb{Z}$ for some a, b and n. Indeed, if $\gcd(a, n) = 1$, an $\{a, b\}$-multiplicative set in $\mathbb{Z}/n\mathbb{Z}$ is a ba^{-1}-free set.

We denote

$$R_k(n) = \max\{|A|, A \text{ is a } k\text{-free set in } \mathbb{Z}/n\mathbb{Z}\}$$

and we show in this article how to compute this quantity recursively in n (Theorems 1, 2, 3 and 4). Proofs also give a way to construct a k-free set of maximal size.

The study of this quantity strongly depends on the arithmetical relative properties of n and k, that is why we split the results in four theorems. We first deal with the case where k and n are coprime, which is actually the most important case. Indeed, when we define k-fold Sidon sets in $\mathbb{Z}/n\mathbb{Z}$, we must add the condition that n is relatively prime to all integers in $[\![1, k]\!]$. Otherwise, one could have $c_i(a_1 - a_2) = 0$ with $a_1 \neq a_2$ for some $|c_i| \leqslant k$, which leads to a nontrivial solution to $c_i(x_1 - x_2) + x_3 - x_4 = 0$ for example.

For k and d integers, we denote by $l_k(d)$ the multiplicative order of k in $(\mathbb{Z}/d\mathbb{Z})^*$. We also use the notations I for the indicator function of odd numbers and φ for the Euler's function. Let see now with the first result below how to compute $R_k(n)$ in the case $\gcd(n, k) = 1$.

Theorem 1. *If* $\gcd(n,k)=1$*,*

$$R_k(n) = \frac{n-1}{2} - \sum_{d|n, d\neq 1} \frac{\varphi(d) I(l_k(d))}{2 l_k(d)}.$$

For the problem of upper bound for the size of a 2-fold Sidon set, we are interested in small $R_2(n)$. Indeed, if $n = 2^m - 1$ is a Mersenne prime number, which implies m prime, then $l_2(n) = m$. Hence, with the notation $\log_r(x) = \ln(x)/\ln(r)$,

$$R_2(n) = \frac{n-1}{2} - \frac{n-1}{2\log_2(n+1)}.$$

For a 2-fold Sidon set A, since $D^*(A)$ is a 2-free set, we have

$$2\binom{|A|}{2} \leqslant R_2(n)$$

which leads to

$$|A| \leqslant \sqrt{\frac{n-1}{2} - \frac{n-1}{2\log_2(n+1)} + \frac{1}{4}} + \frac{1}{2}.$$

Moreover, we prove in Section 3 that for fixed k the error term is $o(n)$. Thus $R_k(n) = (n-1)/2 - o(n)$.

When k divides n, the problem becomes easier and we have the two following results.

Theorem 2. *If* m *is not divisible by* k*, then*

$$R_k(km) = (k-1)m.$$

When k^2 divides n, we get a recursive formula. That is the purpose of Theorem 3.

Theorem 3. *Let* $k, m,$ *and* n *be integers. Then, we have :*

$$R_k(k^2 m) = R_k(m) + \left(k^2 - k\right) m.$$

Notice that Theorems 1, 2 and 3 cover all cases when k is prime. Moreover, recall that the maximal density of a k-free set in $[\![1, n]\!]$ is $k/(k+1)$. In the modular case, applying Theorem 3 we get

$$R_k(k^{2m}) = \frac{k}{k+1}\left(k^{2m} - 1\right)$$

which lead to the next proposition.

Proposition 1. *Let* k *be an integer,* $k \geqslant 1$*, we have*

$$\limsup_n \frac{R_k(n)}{n} = \frac{k}{k+1}.$$

Now, to illustrate the two last theorems, let consider an example. We compute $R_{15}(826875)$:

$$\begin{aligned}R_{15}(826875) &= R_{3.5}(3^3.5^4.7^2)\\ &= R_{3.5}(3.5^2.7^2) + (15^2-15).3.5^2.7^2\\ &= (15-1)5.7^2 + (15^2-15).3.5^2.7^2\\ &= 775180.\end{aligned}$$

We will consider again this example in Section 5.

In the general case, we cannot obtain a closed formula, but in Section 4 we propose an efficient algorithm to compute $R_k(n)$.

Theorem 4. *There exists an algorithm which provides the maximal size of a k-free set in $\mathbb{Z}/n\mathbb{Z}$ and a method to construct one in $O((\log(n))^2)$ operations.*

Here, an "operation" is an addition, a multiplication, a comparison or an assignment. To get this complexity, we must assume that we know the prime factorization of k and n, which is unfortunately hard to obtain in general. However, we can easily apply the algorithm to compute our function R_k for new types of k and n. That is the purpose of the theorem below.

Theorem 5. *Let p and q be prime numbers, α, β and u be integers.*

1. *If* $\gcd(u,p)=1$,

$$R_{up}(p^\alpha) = \sum_{i=0}^{\lfloor\frac{\alpha-1}{2}\rfloor} \varphi(p^{\alpha-2i}).$$

2. *If* $\gcd(u,p)=1$,

$$R_{up^2}(p^\alpha) = \sum_{i=0}^{\lfloor\frac{\alpha-1}{4}\rfloor} \left(\varphi(p^{\alpha-4i}) + \varphi(p^{\alpha-4i-1})\right).$$

3. *If* $\gcd(u,p)=\gcd(u,q)=1$,

$$R_{up}(p^\alpha q^\beta) = \sum_{j=0}^{\beta}\sum_{i=0}^{\lfloor\frac{\alpha-1}{2}\rfloor} \varphi(p^{\alpha-2i}q^{\beta-j})$$

$$R_{up^2}(p^\alpha q^\beta) = \sum_{j=0}^{\beta}\sum_{i=0}^{\lfloor\frac{\alpha-1}{4}\rfloor} \left(\varphi(p^{\alpha-4i}q^{\beta-j}) + \varphi(p^{\alpha-4i-1}q^{\beta-j})\right).$$

In the same way, we could obviously go further and study the case $k = up^3$ or $n = p^\alpha q^\beta r^\gamma$ for instance, but that would give very unpleasant formulas.

Next, we study k-free sets in the set of integers, and not in modular sets anymore. We wonder what is the minimal size of a k-free set in $[\![1, n]\!]$ which is maximal for inclusion, and we answer it in the following theorem, where we define

$$\tilde{R_k}(n) = \min \{|A|, A \subset [\![1, n]\!] \text{ a k-free set which is maximal for inclusion}\}.$$

Theorem 6.

$$\tilde{R_k}(n) = \frac{k^2}{k^2 + k + 1} n + O(\log_k^2(n)).$$

In the next section, we introduce some notations and give two lemmas. Section 3 contains the proof of Theorems 1, 2 and 3. In Section 4, we study and prove the algorithm for the general case, which we use in Section 5. We conclude by the proof of Theorem 6 in the last section.

2 Preparatory lemmas

Let introduce some useful notations for our study. We define $\mathcal{O}^k(x) := \{k^j x, j \in \mathbb{N}\}$ and we call it the orbit of x (by the multiplication by k). We use it in a different context (in $\mathbb{Z}/n\mathbb{Z}$ or in $\mathbb{N}$, Section 6) with the same notation. We denote by $k \cdot A := \{ka, a \in A\}$ the dilated set of A and by A_m the subset of $\mathbb{Z}/n\mathbb{Z}$

$$A_m := \{x, \gcd(x, n) = m\} = \left\{x = mu, \gcd\left(u, \frac{n}{m}\right) = 1\right\}$$

and we have $|A_m| = \varphi(n/m)$.

To study k-free sets, it is important to know more about $\mathcal{O}^k(x)$ for each $x \in \mathbb{Z}/n\mathbb{Z}$. That is the purpose of our first two lemmas.

Lemma 1. *If we write*

$$k = u \prod_{i=1}^{r} p_i^{k_i} \text{ and } n = \prod_{i=1}^{r} p_i^{n_i} \prod_{i=r+1}^{s} p_i^{n_i}$$

with $\gcd(u, p_i) = 1, \forall i \in [\![1, s]\!]$*, then for* m *which divides* n*,* m *has the form*

$$m = \prod_{i=1}^{r} p_i^{m_i} \prod_{i=r+1}^{s} p_i^{m_i}$$

with $m_i \leqslant n_i, \forall i \in [\![1, s]\!]$*, then we have*

$$k \cdot A_m = A_{m'} \text{ where } m' = m \prod_{i=1}^{r} p_i^{\min(k_i, n_i - m_i)}.$$

Proof. Let $x \in A_m$, then $x = mv$ with $\gcd(v, n/m) = 1$. Thus, we get

$$\begin{aligned}\gcd(kx, n) &= m \gcd\left(kv, \frac{n}{m}\right) \\ &= m \gcd\left(k, \frac{n}{m}\right) \\ &= m \gcd\left(\gcd(k, n), \frac{n}{m}\right) \\ &= m \gcd\left(\prod_{i=1}^{r} p_i^{k_i}, \prod_{i=1}^{r} p_i^{n_i - m_i} \prod_{i=r+1}^{s} p_i^{n_i - m_i}\right).\end{aligned}$$

What we get is that $k \cdot A_m \subset A_{m'}$.

Conversely, there exists now $y \in A_{m'}$ such that $y = kx$ with $x \in A_m$ (since $A_m \neq \emptyset$). But for all z in $A_{m'}$, there exists w, $\gcd(w, n) = 1$ and $z = wy$. Clearly, xw belongs to A_m and $z = kxw$, which concludes the proof.

□

Lemma 2. *Let m be a divisor of n, k be an integer such that $\gcd(k, n/m) = 1$ and $x \in A_m$. Then*

$$\left|\mathcal{O}^k(x)\right| = l_k\left(\frac{n}{m}\right).$$

Proof. Since $x \in A_m$, if we denote by $< x >$ the subgroup generated by x, we have

$$< x > \cong \mathbb{Z}/\left(\frac{n}{m}\right)\mathbb{Z}.$$

Then, since k is invertible in this subgroup

$$\mathcal{O}^k(x) \cong \mathcal{O}^k(1) =< k > \subset \mathbb{Z}/\left(\frac{n}{m}\right)\mathbb{Z}$$

and the size of $< k >$ in this subgroup is exactly $l_k(n/m)$.

□

3 Proof of Theorems 1, 2 and 3

We first deal with the Theorem 1, the case $\gcd(n, k) = 1$.

Proof. As mentioned in the introduction, let $l_k(d)$ be the order of k in $(\mathbb{Z}/d\mathbb{Z})^*$ and I be the indicator function of odd numbers.

By lemma 1, $\mathcal{O}^k(x) \subset A_m$, for all x in A_m. Therefore, we consider the suitable partition

$$(\mathbb{Z}/n\mathbb{Z}) \setminus \{0\} = \bigsqcup_{m|n, m<n} A_m,$$

where the notation $\bigsqcup$ means that it is a disjoint union. Notice that this partition is trivial if n is prime. By lemma 2, if $x \in A_m$, we have

$$\left|\mathcal{O}^k(x)\right| = l_k\left(\frac{n}{m}\right).$$

Hence, we can make a partition of A_m in $\varphi(n/m)/l_k(n/m)$ distinct orbits of length $l_k(n/m)$. In each orbit, to get an optimal k-free set, we have to take the most possible elements without taking two consecutive elements. But the orbits are cyclic, that's why if the length l of an orbit is even, we can take $l/2$ elements, whereas if l is odd, we can take only $(l-1)/2$ elements. We finally get the formula

$$\begin{aligned} R_k(n) &= \sum_{d|n, d\neq 1} \frac{\varphi(d)}{l_k(d)} \left(\frac{l_k(d) - I(l_k(d))}{2} \right) \\ &= \frac{n-1}{2} - \sum_{d|n, d\neq 1} \frac{\varphi(d) I(l_k(d))}{2 l_k(d)}. \end{aligned}$$

□

Actually, if we fix k, $R_k(n)$ is asymptotically $(n-1)/2 - o(n)$. Indeed, for all $\varepsilon > 0$, there exists d_0 such that $\log_k d_0 \geqslant 1/\varepsilon$ and there exists n such that $d_0^2/6 \leqslant \varepsilon n/2$. Thus,

$$\begin{aligned} \sum_{d|n, d\neq 1} \frac{\varphi(d) I(l_d)}{2 l_d} &= \sum_{d|n, d\neq 1, d\leqslant d_0} \frac{\varphi(d) I(l_d)}{2 l_d} + \sum_{d|n, d\neq 1, d > d_0} \frac{\varphi(d) I(l_d)}{2 l_d} \\ &\leqslant \sum_{d|n, d\neq 1, d\leqslant d_0} \frac{\varphi(d)}{6} + \sum_{d|n, d\neq 1, d > d_0} \frac{\varphi(d)}{2 log_2 d} \\ &\leqslant \frac{d_0^2}{6} + \frac{\varepsilon n}{2} \\ &\leqslant \varepsilon n. \end{aligned}$$

Now, we consider the case $n = k^2 m$, for which we have a suitable partition of $\mathbb{Z}/n\mathbb{Z}$:

Lemma 3. *In this case, we have*

$$\mathbb{Z}/n\mathbb{Z} = \left(k^2\mathbb{Z}/n\mathbb{Z}\right) \bigsqcup \left(\bigcup_{h \not\equiv 0 \pmod k} \{h, kh\} \right).$$

Proof. Indeed, if $x \not\equiv 0 \pmod{k^2}$ and $x \equiv 0 \pmod k$, then $x = kh$ with $h \not\equiv 0 \pmod k$. Thus, we have all the elements in this union. Moreover, if we have $h \not\equiv 0 \pmod k$, then $kh \not\equiv 0 \pmod{k^2}$, which shows that the first union is disjoint. □

Let see now why this is a good repartition of elements for our problem, through the proof of Theorem 3 :

Proof. We remark two main things :

- If $x \in k^2\mathbb{Z}/n\mathbb{Z}$, $kx \in k^2\mathbb{Z}/n\mathbb{Z}$.
- If $h \not\equiv 0 \pmod k$, we can not write $h = ku$ in $k^2\mathbb{Z}/n\mathbb{Z}$.

We consider now A a k-free set in $\mathbb{Z}/n\mathbb{Z}$. First, for each $h \not\equiv 0 \pmod{k}$, at most one of $\{h, kh\}$ lies in A. Furthermore, by the first remark, $A \cap k^2\mathbb{Z}/n\mathbb{Z}$ is also a k-free set, which can be easily seen equivalent to a k-free set in $\mathbb{Z}/m\mathbb{Z}$. This leads to

$$R_k(k^2m) \leqslant R_k(m) + |\{h \not\equiv 0 \pmod{k}\}| = R_k(m) + \left(k^2 - k\right) m.$$

Let see now the construction of an optimal k-free set. By the second remark, we can take every $h \not\equiv 0 \pmod{k}$ in A, and we now that $kh \notin k^2\mathbb{Z}/n\mathbb{Z}$, so we can take $R_k(m)$ elements from $k^2\mathbb{Z}/n\mathbb{Z}$ in A. Thus, we get

$$R_k(k^2m) = R_k(m) + \left(k^2 - k\right) m$$

and that concludes the proof.

□

Finally, we consider $n = km$ with $m \not\equiv 0 \pmod{k}$. In this case, we have :

Lemma 4.

$$\mathbb{Z}/n\mathbb{Z} = \bigcup_{h \not\equiv 0 \pmod{k}} \{h, kh\}.$$

Proof. If $x \equiv 0 \pmod{k}$, there exists u such that $x = ku$. If $u \not\equiv 0 \pmod{k}$, x is in the right form. Else, $u \equiv 0 \pmod{k}$, then there exists v, $u = kv$ and we have $x = x + n = x + km = k^2v + km$. But $m \not\equiv 0 \pmod{k}$ by hypothesis, then we can write $m = lk + a$ with $a \not\equiv 0 \pmod{k}$. We get

$$x + km = k(kv + lk + a).$$

Since $h = kv + lk + a \not\equiv 0 \pmod{k}$, we have written $x = x + km = kh$ with $h \not\equiv 0 \pmod{k}$, which concludes the lemma. □

We can now easily prove Theorem 2.

Proof. If A is a k-free set, for each $h \not\equiv 0 \pmod{k}$, at most one of $\{h, kh\}$ lies in A, then $|A| \leqslant (k-1)m$. If $h \not\equiv 0 \pmod{k}$, we can not write $h = ku$ in $\mathbb{Z}/n\mathbb{Z}$ since $n = km$. Thus $\{h \not\equiv 0 \pmod{k}\}$ is a k-free set and we get

$$R_k(km) = (k-1)m.$$

□

4 Theorem 4 : the general case

The situation, for the general case, is much more difficult. Indeed, if $x \in A_m$, we do not have necessarily $\mathcal{O}^k(x) \subset A_m$ anymore. To deal with it, our strategy is to build a graph with divisors of n as vertices. Then, we connect m and m' if and only if they are distinct and $k \cdot A_m = A_{m'}$. Moreover, we have to consider separately divisors m such that

$k \cdot A_m = A_m$. Actually, they are going to be the roots, as soon as we will interpret our graph as a forest (a disjoint union of rooted trees). Then, to obtain an optimal k-free set B, we would like to take some A_m, not connected in our graph, maximizing the size of B. That is why we need a result about specific rooted trees. That is the purpose of the next subsection.

4.1 An algorithm on rooted trees

Let T be a rooted tree where the set of nodes is $V = \{v_i\}_{i \in I}$ with I a finite set and E is the set of edges. We associate a value $\alpha_i \geqslant 0$ to each v_i and we denote by l_i its level (recall that the level of a node is defined by 1+ the minimal number of connections between the node and the root). Assume that T has the following property :

$$\text{If } v_i \text{ is the parent of } v_j, \text{ then } \alpha_i < \alpha_j. \tag{1}$$

In other words, α is strictly increasing in each branche. Notice that this condition implies that if v_i is not the root of T, $\alpha_i > 0$. We search a subset A of I satisfying :

$$\forall (i,j) \in A^2, (v_i, v_j) \notin E \text{ and } \alpha_i \neq 0 \tag{2}$$

which maximizes the quantity

$$\Lambda_A = \sum_{i \in A} \alpha_i.$$

We denote by l the maximal level in T and we construct a set B by the following algorithm : Initialization : $B = \{v_i | l_i = l\}$. Then for k from 1 to $l-1$: for all i such that $l_i = l - k$, we add v_i to B if and only if $\alpha_i \neq 0$ and there is no child of v_i in B.

It is clear that B satisfies (2). Actually, B is the required set for our problem.

Lemma 5. *B is the unique subset of I maximizing Λ_A among the sets A satisfying (2).*

Proof. We proceed by induction on the size of I. If $|I| = 1$ there is nothing to say.

Let n be an integer and assume that the lemma holds for all k less than n. Let $|I| = n+1$, B the set from the algorithm applied to T and C be a subset of I maximizing Λ_A among the sets A satisfying (2). We denote by v_0 the root of T, and v_i, $i \in [\![1, K]\!]$ the childs of v_0. We also define T_i the rooted subtree of T of root v_i for all i in $[\![1, K]\!]$, $B_i = B \cap T_i$ and $C_i = C \cap T_i$. By induction, for all i, $\Lambda_{C_i} \leqslant \Lambda_{B_i}$ with equality if and only if $B_i = C_i$.

If $v_0 \in B$ and $v_0 \in C$, we have

$$\Lambda_C - \alpha_0 = \sum_{i=1}^{K} \Lambda_{C_i} \leqslant \sum_{i=1}^{K} \Lambda_{B_i} = \Lambda_B - \alpha_0.$$

Thus, by the definition of C, this is an equality, and finally $B = C$.

If $v_0 \notin B$ and $v_0 \notin C$, we have

$$\Lambda_C = \sum_{i=1}^{K} \Lambda_{C_i} \leqslant \sum_{i=1}^{K} \Lambda_{B_i} = \Lambda_B$$

which ensure that $B = C$ for the same reason.

If $v_0 \in B$ and $v_0 \notin C$, since $\alpha_0 > 0$ (otherwise, v_0 is not in B according to the algorithm), we have

$$\Lambda_C = \sum_{i=1}^{K} \Lambda_{C_i} \leqslant \sum_{i=1}^{K} \Lambda_{B_i} = \Lambda_B - \alpha_0 < \Lambda_B$$

which leads to a contradiction.

If $v_0 \notin B$ and $v_0 \in C$, $\alpha_0 > 0$ (since C satisfies (2)) and it means that there exists i_0 in $[\![1, K]\!]$ such that $v_{i_0} \in B$. Then, we consider the branche from v_0 which contains v_{i_0}. If $K > 1$, its size is strictly less than $n+1$ and we can apply the induction hypothesis to get $\Lambda_{C_{i_0}} + \alpha_0 < \Lambda_{B_{i_0}}$. Thus,

$$\Lambda_C = \sum_{i \neq i_0} \Lambda_{C_i} + \Lambda_{C_{i_0}} + \alpha_0 < \sum_{i \neq i_0} \Lambda_{B_i} + \Lambda_{B_{i_0}} = \Lambda_B$$

and we have a contradiction. If $K = 1$, we denote by v_1 the only child of v_0, $v_1 \in B$ and $v_1 \notin C$, and considering now the subtrees whose the roots are every nodes of level 3, we get

$$\Lambda_C = \sum \Lambda_{C'_i} + \alpha_0 < \sum \Lambda_{B'_i} + \alpha_1 = \Lambda_B$$

where we use $\alpha_1 > \alpha_0$. This is a contradiction.

Finally, $B = C$ in every cases and the lemma is proved.

□

4.2 Proof of Theorem 4

The aim is to define a forest with valuation on vertices for which each tree (a connected component of the forest) satisfies (1) and such that the previous algorithm applied to each tree gives an optimal k-free set. Recall that if m divides n, we denote by A_m the subset of $\mathbb{Z}/n\mathbb{Z}$

$$A_m := \{x, \gcd(x, n) = m\}$$

and we have $|A_m| = \varphi(n/m)$. The disjoint union of A_m is a partition of $\mathbb{Z}/n\mathbb{Z}$.

Let $G = (V, E)$ be a graph with the set of vertices $V = \{m\}_{m|n}$ and we define the set of edges E by :

$$(m, m') \in E \text{ if and only if } m < m' \text{ and } k \cdot A_m = A_{m'}. \tag{3}$$

We first need to well understand this graph, then we will associate suitable values to vertices for our problem. We write

$$k = u \prod_{i=1}^{r} p_i^{k_i} \text{ and } n = \prod_{i=1}^{r} p_i^{n_i} \prod_{i=r+1}^{s} p_i^{n_i}$$

with $\gcd(u, p_i) = 1, \forall i \in [\![1, s]\!]$ and $k_i > 0$ for all i. Let denote by $\mathcal{M}$ the set of divisors of n of the form

$$m = \prod_{i=1}^{s} p_i^{m_i}$$

with $m_i \leqslant n_i, \forall i \in [\![1, s]\!]$, such that there exists $i_0 \leqslant r$ satisfying $m_{i_0} < \min(k_{i_0}, n_{i_0})$. The next proposition gives the structure of G.

Proposition 2. *G is a disjoint union of rooted trees. Furthermore :*

(i) A connected component (a tree) of G is entirely defined by the choice of $\{d_i\}_{i=r+1}^{s}$ with $d_i \leqslant n_i$. We mean :

a) in each tree, all vertices have same $\{d_i\}_{i=r+1}^{s}$,

b) conversely, if m and m' have same $\{d_i\}_{i=r+1}^{s}$, they are in the same tree.

(ii) The leaves are exactly the elements of $\mathcal{M}$.

(iii) The root of the tree defined by $\{d_i\}_{i=r+1}^{s}$ is

$$m = \prod_{i=1}^{r} p_i^{n_i} \prod_{i=r+1}^{s} p_i^{d_i}.$$

(iv) The level of m is $j_m + 1$ where

$$j_m = \min \{j | jk_i \geqslant n_i - m_i, \forall i \in [\![1, r]\!]\}.$$

Proof. We define

$$k^j * m = m \prod_{i=1}^{r} p_i^{\min(jk_i, n_i - m_i)}.$$

By lemma 1, $A_{k*m} = k \cdot A_m$, then if (m, m') is an edge, we have $m_i = m'_i$ for all i in $[\![r+1, s]\!]$. Thus, if there exists a path between two vertices, they have the same $\{d_i\}_{i=r+1}^{s}$.

The next lemma shows that a vertice is either in $\mathcal{M}$ or has a descendant in $\mathcal{M}$.

Lemma 6. *Let $m' = \prod_{i=1}^{s} p_i^{m'_i}$ be a divisor of n which is not in $\mathcal{M}$, then there exists $t > 0$ and m in $\mathcal{M}$ such that $m' = k^t * m$.*

Proof. Let t defined by

$$t = \min \left\{ j | \exists i_0 \leqslant r, m'_{i_0} - jk_{i_0} < \min(k_{i_0}, n_{i_0}) \right\}$$

and define $\alpha_i = \max(0, m'_i - tk_i)$ and

$$m = \prod_{i=1}^{r} p_i^{\alpha_i} \prod_{i=r+1}^{s} p_i^{m'_i}$$

which belongs to $\mathcal{M}$ by definition of t. Notice that $t > 0$ since $m' \notin \mathcal{M}$. Thus, we have

$$\begin{aligned} k^t * m &= m \prod_{i=1}^{r} p_i^{\min(tk_i, n_i - m_i)} \\ &= \prod_{i=1}^{r} p_i^{\alpha_i + \min(tk_i, n_i - \alpha_i)} \prod_{i=r+1}^{s} p_i^{m'_i}. \end{aligned}$$

We have to study three cases :

- $\alpha_i = 0$ and $k_i < n_i$: $m'_i \geqslant k_i$ since $m' \in M$, then $m'_i = tk_i$ by the definition of t and we get in this case $\alpha_i + \min(tk_i, n_i - \alpha_i) = m'_i$.
- $\alpha_i = 0$ and $n_i \leqslant k_i$: $m'_i = n_i$ since $m' \in M$ and we have $\alpha_i + \min(tk_i, n_i - \alpha_i) = n_i = m'_i$.
- Otherwise, $n_i - \alpha_i = n_i - m'_i + tk_i \geqslant tk_i$, then $\alpha_i + \min(tk_i, n_i - \alpha_i) = m'_i$.

We finally get $k^t * m = m'$, as we expected. □

Conversely, if $m \in M$ and $t > 0$, $k^t * m \notin M$, and we get that vertices in $\mathcal{M}$ have no child. Moreover, if we consider

$$m = \prod_{i=1}^{r} p_i^{n_i} \prod_{i=r+1}^{s} p_i^{d_i}$$

it is clear that $k * m = m$, it means that m have no parent. Finally, if m' have the same $\{d_i\}_{i=r+1}^{s}$, we have $k^{j_m} * m' = m$ and $k^{j_m - 1} * m' \neq m$ by the definition of j_m.

Through those observations, we get the conclusions of the proposition.

□

Now, we need to see how to give valuations for vertices. The main problem comes from roots, which are the m satisfying $k \cdot A_m = A_m$. The next lemma computes the maximal size of a k-free set in A_m when m is a root of our graph.

Lemma 7. *If m is a root of our graph (which is tantamount to $\gcd(k, n/m) = 1$), the maximum size of a k-free set included in A_m is*

$$R_k(A_m) := \frac{\varphi(n/m)}{l_k(n/m)} \left(\frac{l_k(n/m) - I(l_k(n/m))}{2} \right).$$

Proof. We have the isomorphism :

$$A_m \cong A'_1 := \left\{ x \in \mathbb{Z}/(n/m)\mathbb{Z},\ \gcd\left(x, \frac{n}{m}\right) = 1 \right\}.$$

But we are in the case $\gcd(k, n/m) = 1$, and if we look through the proof of Theorem 1, we get immediately the result.

□

Thus, we define the valuation of vertices for all m which divides n :

$$\alpha_m = \begin{cases} R_k(A_m) \text{ if } m \text{ is a root} \\ \varphi\left(\dfrac{n}{m}\right) \text{ otherwise.} \end{cases}$$

Notice that our graph has the property (1), which we recall here :

$$\text{If } v_i \text{ is the parent of } v_j, \text{ then } \alpha_i < \alpha_j.$$

When we apply the algorithm of Section 2, we get a set B of vertices. To construct a k-free set, we can join A_m for m in B and not a root, and for the roots m in B we can take K_m a maximal k-free set in A_m. More precisely, we define

$$\overline{B} := \left(\bigsqcup_{\substack{m \in B \\ \gcd(k, n/m) \neq 1}} A_m \right) \bigsqcup \left(\bigsqcup_{\substack{m \in B \\ \gcd(k, n/m) = 1}} K_m \right)$$

which is clearly a k-free set since B satisfies 2 and by the definition of K_m.

Proposition 3. *$\overline{B}$ is an optimal k-free set in $\mathbb{Z}/n\mathbb{Z}$ and has size*

$$\sum_{\substack{m \in B \\ \gcd(k, n/m) \neq 1}} \varphi\left(\frac{n}{m}\right) + \sum_{\substack{m \in B \\ \gcd(k, n/m) = 1}} R_k(A_m).$$

Proof. Assume that C is a k-free set in $\mathbb{Z}/n\mathbb{Z}$ with $|C| > \left|\overline{B}\right|$. Let x be an element in $C \backslash \overline{B}$ of maximal level t, m the integer such that $x \in A_m$ and T_i the rooted tree which contains m.

First case : $t = 1$ and $m \notin B$. Thus, m is a root but not in B, which means that there is a child m' of m in B (otherwise $\alpha_m = R_k(A_m) = 0$ and C could not be a k-free set). Then, the set $k^{-1}(x) = \{y \in A_{m'} | y = kx\}$ has no element in C but has size

$$\left|k^{-1}(x)\right| = \frac{\varphi(n/m')}{\varphi(n/m)} > 1$$

and by substituting $\{x\}$ by $k^{-1}(x)$, we get a k-free set (since t is the maximal level of an element of $C \backslash \overline{B}$) of size strictly greater than C.

Second case : $t > 1$. By the construction of $\overline{B}$, m does not belong to B and we can do as in the previous case.

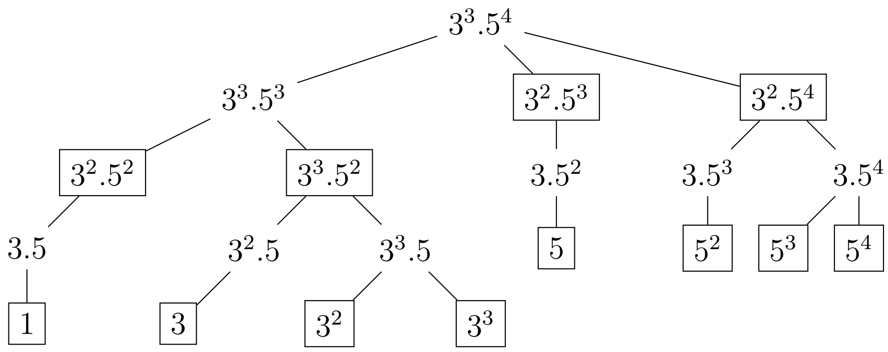
The two cases lead to a contradiction, then all elements x in $C \backslash \overline{B}$ satisfy $t = 1$ and m belongs to B. Thus, m is a root and we can substitute $C \cap A_m$ by K_m for each root, and we get $|C| \leqslant \left|\overline{B}\right|$. We finally get the result by counting the size of $\overline{B}$.

□

Thus, to get $R_k(n)$, if the prime factorization of k and n is known, we need to construct the graph ($O(\log(n))$ operations), to apply the algorithm ($O(\log(n))$), to compute α_m for m in B ($O((\log(n))^2)$ operations since we have the prime factorization of m) and finally add those values.

5 Applications of Theorem 4

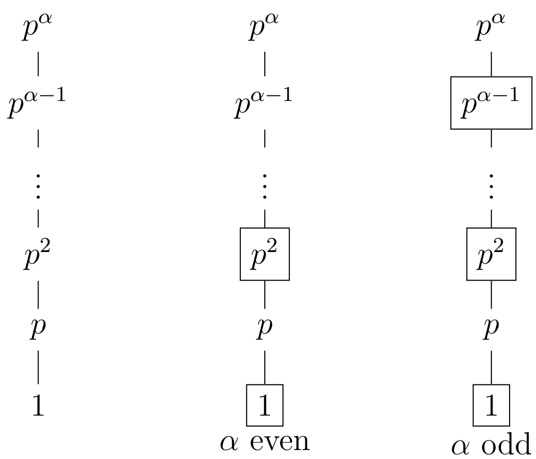
Now, to illustrate the method in a particular case, we deal with the example mentioned in introduction, which is $n = 3^3.5^4.7^2 = 826875$ and $k = 3.5 = 15$. In this case, we get a forest with roots $3^3.5^4$, $3^3.5^4.7$ and $3^3.5^4.7^2$. We just represent below one of those trees. To get the second, we have to multiply each vertice by 7, and for the third, by 7^2. Applying the algorithm, we get :



To get the maximal size of a 15-free set in $\mathbb{Z}/826875\mathbb{Z}$ we have to sum all $\varphi(n/m)$ for all m choosen by the algorithm in each tree. And we get $R_{15}(826875) = 775180$ as we deduced from Theorems 2 and 3.

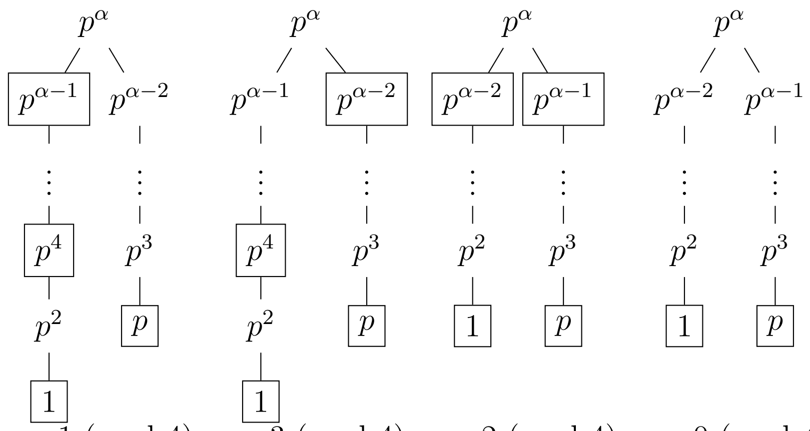
This way to compute $R_k(n)$ does not give a general formula, that is why we study in Theorem 5 theorem several cases, which we prove here.

Proof. 1. The first graph below is the one we get in this particular case ($n = p^\alpha$, $k = up$ with $\gcd(u, p) = 1$), then we apply the algorithm and we obtain a set of vertices, which are the one with a box around. We get the second graph when α is even and the third if α is odd :



Since $A_{p^\alpha} = \{0\}$, $R_{up}(p^\alpha) = 0$, that is why p^α is never considered by the algorithm. By applying the Proposition 3, we get the result.

2. We give below the results of the algorithm ($n = p^\alpha$, $k = up^2$ with $\gcd(u,p) = 1$), which depends on the value of α modulo 4 (notice that $R_{up^2}(p^\alpha) = 0$) :



$\alpha \equiv 1 \pmod 4$ $\alpha \equiv 3 \pmod 4$ $\alpha \equiv 2 \pmod 4$ $\alpha \equiv 0 \pmod 4$

And we compute $R_k(n)$ again thanks to Proposition 3.

3. If $k = up$ and $n = p^\alpha q^\beta$ with $\gcd(u,p) = gcd(u,q) = 1$, we get a forest of $\beta + 1$ rooted trees $(T_j)_{j=0\cdots\beta}$ with T_j :

$$\begin{array}{c} p^\alpha q^j \\ | \\ p^{\alpha-1}q^j \\ | \\ \vdots \\ | \\ p^2 q^j \\ | \\ pq^j \\ | \\ q^j \end{array}$$

Then, the algorithm gives, as in the first case, the size of an optimal k-free set in T_j:

$$\sum_{i=0}^{\lfloor \frac{\alpha-1}{2} \rfloor} \varphi(p^{\alpha-2i} q^{\beta-j}).$$

We just need to add the contribution of all T_j's to get the result.

For the case $k = up^2$, this time, it is a consequence of the second case.

□

6 Proof of Theorem 6

Now, we want to study k-free sets in $[\![1,n]\!]$, and a good way is to consider the partition

$$[\![1,n]\!] = \bigsqcup_{i \not\equiv 0 \pmod k} \left(\mathcal{O}^k(i) \bigcap [\![1,n]\!] \right).$$

Indeed, to be a k-free set is equivalent to not have consecutive elements in such orbits (we abusively call orbit of i the set $\mathcal{O}^k(i) \bigcap [\![1,n]\!]$). Let see now what being maximal for inclusion means in term of orbits. Actually, we can clearly assume that A is a maximal k-free set (for inclusion) if and only if for each orbits of $i \not\equiv 0 \pmod k$, exactly one of the two first elements is in A, exactly one of the two last elements is in A, there is not consecutive elements, and for all three consecutives elements, there is at least one which is in A. That leads us to study the following combinatorial problem :

A set $E \subset [\![1,l]\!]$ satisfies $(\mathcal{P})$ if :

- $1 \in E$ or $2 \in E$.
- $l-1 \in E$ or $l \in E$.
- $i \in E \Rightarrow (i-1) \notin E$ and $(i+1) \notin E$.
- $\forall i \in [\![2, l-1]\!], \{i-1, i, i+1\} \cap E \neq \emptyset$.

We denote by $h(l)$ the minimal size of a set which satisfies $(\mathcal{P})$ in $[\![1,l]\!]$.

Lemma 8.

$$h(l) = \left\lceil \frac{l}{3} \right\rceil.$$

Proof. First case : $l = 3u$. $B = \{2, 5, \cdots, 2+3(u-1)\}$ satisfies $(\mathcal{P})$ and has a size $u = l/3$. Since we have to take one element among $\{3i+1, 3i+2, 3i+3\}, \forall i \in [\![0, u-1]\!]$, $h(l) \geqslant u$. Then, $h(3u) = u$.

Second case : $l = 3u-1$. We consider the following partition :

$$[\![1, 3u-1]\!] = \{1,2\} \bigcup \left(\bigcup_{i \in [\![1,u-1]\!]} \{3i, 3i+1, 3i+2\} \right).$$

Since we must have at least one element from each subsets , we have $h(3u-1) \geqslant u$. But $B = \{2, 5, \cdots, 2+3(u-1)\}$ has still the good size. Then $h(3u-1) = u$.

Third case : $l = 3u-2$, We consider the following partition :

$$[\![1, 3u-2]\!] = \{1,2\} \bigcup \left(\bigcup_{i \in [\![1,u-2]\!]} \{3i, 3i+1, 3i+2\} \right) \bigcup \{3u-3, 3u-2\}.$$

Since we must have at least one element from each subsets, we have $h(3u-2) \geqslant u$. But $B = \{1, 4, \cdots, 1+3(u-1)\}$ satisfies $(\mathcal{P})$. Then $h(3u-2) = u$.

□

We are now able to prove the Theorem 6.

Proof. If we denote $A_i := \rrbracket \frac{n}{k^{i+1}}, \frac{n}{k^i} \rrbracket$, we have

$$\llbracket 1, n \rrbracket = \bigcup_{i=0}^{d} A_i$$

where $d = [\log_k(n)]$. Moreover,

$$|A_i| = \frac{n}{k^i} - \frac{n}{k^{i+1}} + \alpha(i)$$

with $|\alpha(i)| \leqslant 1$. And the numbers of $j \not\equiv 0 \pmod{k}$ in A_i is

$$\left(1 - \frac{1}{k}\right)\left(\frac{n}{k^i} - \frac{n}{k^{i+1}} + \alpha(i)\right) + \epsilon(i)$$

with $|\epsilon(i)| \leqslant 1$. Each element in A_i has an orbit of size $i + 1$, then we deduce from the Lemma 8:

$$\begin{aligned}
\tilde{R}_k(n) &= \sum_{i=0}^{d} \left\lceil \frac{i+1}{3} \right\rceil \left(\left(1 - \frac{1}{k}\right)\left(\frac{n}{k^i} - \frac{n}{k^{i+1}} + \alpha(i)\right) + \epsilon(i)\right) \\
&= \sum_{i=0}^{d} \left\lceil \frac{i+1}{3} \right\rceil \left(1 - \frac{1}{k}\right)\left(\frac{n}{k^i} - \frac{n}{k^{i+1}}\right) + O(\log_k^2(n)).
\end{aligned}$$

To study this sum, we group together by three the terms with same integer part, in order to get a telescopic behaviour. Thus, we get :

$$\begin{aligned}
\tilde{R}_k(n) = &\left(1 - \frac{1}{k}\right) \sum_{i=0}^{\lfloor \frac{d}{3} \rfloor} (i+1)\left(\frac{n}{k^{3i}} - \frac{n}{k^{3i+1}} + \frac{n}{k^{3i+1}} - \frac{n}{k^{3i+2}} + \frac{n}{k^{3i+2}} - \frac{n}{k^{3i+3}}\right) \\
&+ \beta(n) + O(\log_k^2(n)) \\
= &\left(1 - \frac{1}{k}\right) \sum_{i=0}^{\lfloor \frac{d}{3} \rfloor} (i+1)\left(\frac{n}{k^{3i}} - \frac{n}{k^{3i+3}}\right) + \beta(n) + O(\log_k^2(n))
\end{aligned}$$

with

$$\beta(n) \leqslant \left(1 - \frac{1}{k}\right) \times 2\left(\frac{d}{3} + 1\right)\left(\frac{n}{k^{d-1}} - \frac{n}{k^{d+1}}\right) = O(\log_k(n)).$$

Finally,

$$\begin{aligned}
\tilde{R}_k(n) &= \left(1 - \frac{1}{k}\right) \sum_{i=0}^{\lfloor \frac{d}{3} \rfloor} \frac{n}{k^{3i}} + O(\log_k^2(n)) \\
&= \frac{k^2}{k^2 + k + 1} n + O(\log_k^2(n)).
\end{aligned}$$

□

Acknowledgements

This work is part of the PhD of the author. He wishes to thank his advisor Alain Plagne for his support during the preparation of this paper. He also thanks Jeremy Le Borgne for his help and careful reading. The author was supported by an ANR grant Csar, number ANR 12 - BS01 - 0011.

References

[1] R. C. Bose, S. Chowla, *Theorems in the additive theory of numbers* , Commentarii Mathematici Helvetici, 37, 141-147, 1962/1963.

[2] J. Cilleruelo and C. Timmons, *k-fold Sidon sets*, Electronic Journal of Combinatorics, 21(4), 2014, #P12.

[3] F. Lazebnik, J. Verstraëte, *On hypergraphs of girth five*, Electronic Journal of Combinatorics, 10(1), 2003, #R25.

[4] K. O'Bryant, *A complete annotated bibliography of work related to Sidon sequences*, Electronic Journal of Combinatorics, #DS11, 2004.

[5] I. Ruzsa, *Solving a linear equation in a set of integers I*, Acta Arithmetica, 65, 259-282, 1993.

[6] J. Singer, *A theorem in finite projective geometry and some applications to number theory*, Transactions of the American Mathematical Society 43, 377-385, 1938.

[7] D. Wakeham and D.R.Wood, *On Multiplicative Sidon Sets*, Integers 13, Paper No. A26, 2013.

[8] E.T.H. Wang, *On Double-Free Sets of Integers*, Ars Combinatoria, 28, 97-100, 1989.