PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 138, Number 3, March 2010, Pages 787–797
S 0002-9939(09)10189-2
Article electronically published on October 23, 2009

THE CARDINALITY OF SOME SYMMETRIC DIFFERENCES

PO-YI HUANG, WEN-FONG KE, AND GÜNTER F. PILZ

(Communicated by Jim Haglund)

Abstract. In this paper, we prove that for positive integers k and n, the cardinality of the symmetric differences of $\{1, 2, \dots, k\}$, $\{2, 4, \dots, 2k\}$, $\{3, 6, \dots, 3k\}$, $\dots$, $\{n, 2n, \dots, kn\}$ is at least k or n, whichever is larger. This solved a problem raised by Pilz in which binary composition codes were studied.

1. Introduction

The symmetric difference of two sets A and B, denoted by $A \Delta B$, is $(A\setminus B)\cup(B\setminus A)$. For any positive integers k and n, the cardinality of the symmetric difference

$$\{1, 2, \dots, k\} \,\Delta\, \{2, 4, \dots, 2k\} \,\Delta \cdots \Delta\, \{n, 2n, \dots, kn\}$$

is of interest in several different situations. Here we mention three of them:

- "To love or not to love". Let us take $k = 3$. Suppose three people with numbers 1, 2, and 3 on their back enter an empty room. Then three more people with numbers 2, 4 and 6 go into this room. Now two people have the same number, namely 2; they fall in love and leave the room. So only numbers 1, 3, 4 and 6 remain. Next, people with numbers 3, 6 and 9 come in. Numbers 3 and 6 find partners, and only the three people with numbers 1, 4 and 9 remain, and so on. The conjecture is: "There will always be at least 3 people in the room." This is easy to show, but it seems considerably harder for a general k greater than 3.
- "Summands in binary polynomials". Over $\mathbb{Z}_2$, consider the sum of polynomials

$$\begin{aligned}(1 + x + \cdots + x^k) \circ x &+ (1 + x + \cdots + x^k) \circ x^2 \\ &\qquad + \cdots + (1 + x + \cdots + x^k) \circ x^n \\ = (1 + x + \cdots + x^k) &+ (1 + x^2 + \cdots + x^{2k}) \\ &\qquad + \cdots + (x^n + x^{2n} + \cdots + x^{kn}).\end{aligned}$$

Received by the editors June 1, 2009.

2010 *Mathematics Subject Classification.* Primary 05A05; Secondary 11N05, 94B05.

The first author was supported by the National Science Council, Taiwan, grant #96-2115-M-006-003-MY3.

The second author was partially supported by the National Science Council, Taiwan, grant #97-2923-M-006-001-MY2.

The third author was supported by grant P19463 of the Austrian National Science Fund (FWF).

Will there always be at least k summands present? Equivalently, has the symmetric difference

$$\{1, 2, \ldots, k\} \,\Delta\, \{2, 4, \ldots, 2k\} \,\Delta \cdots \Delta\, \{n, 2n, \ldots, kn\}$$

always at least k elements?

- "Codes by composition". Encode a binary message $(a_1, a_2, \ldots, a_n)$ of length n as a polynomial by composition as

$$a_1(1 + x + \cdots + x^k) \circ x + a_2(1 + x + \cdots + x^k) \circ x^2 + \cdots + a_n(1 + x + \cdots + x^k) \circ x^n.$$

There is a good reason to do this kind of coding; see [3]. A positive answer in item 2 would give a positive indication that the minimal weight of these codewords is k.

With some extensive experimental data, the following conjecture, with a convenient name, was raised in [3].

1-2-3 Conjecture. *The cardinality of the symmetric difference*

$$\{1, 2, \ldots, k\} \,\Delta\, \{2, 4, \ldots, 2k\} \,\Delta \cdots \Delta\, \{n, 2n, \ldots, kn\}$$

is always at least k.

It was shown in [3] that the conjecture holds true for $k \le 6$. In private communications, E. Fried (Budapest) proved it for $k = 7$ and $k = 8$, and P. Fuchs (Linz) for $k \ge 10^{12}$.

In this paper, we prove a slightly general version of the conjecture. Some notation can be useful for our discussion.

For $k, s \in \mathbb{N}$, let $I_k = \{1, 2, \ldots, k\}$ and $sI_k = \{s, 2s, \ldots, ks\}$. For $1 \le u < v$, put $D_{k\times[u,v]} = uI_k \,\Delta\, (u+1)I_k \,\Delta \cdots \Delta\, vI_k$. When $u = 1$, we use $D_{k\times v}$ instead of $D_{k\times[1,v]}$ and denote by $d_k(n)$ the cardinality of $D_{k\times n}$. It is obvious that if $1 < s < n$, then $D_{k\times n} = D_{k\times s} \,\Delta\, D_{k\times[s+1,n]}$. Also,

$$D_{k\times n} = D_{n\times k} \quad \text{for all } k \text{ and } n. \tag{1:1}$$

Now, we modify the conjecture (but still keep the same name) and shorten it using the prepared notation.

1-2-3 Conjecture. *For all $k, n \in \mathbb{N}$, $d_k(n) \ge \max\{n, k\}$.*

It is easy to see that $D_{k\times k} = \{1^2, 2^2, \ldots, k^2\}$, which has k elements. The fact (1:1) tells us that it suffices to show

Restricted 1-2-3 Conjecture. *For all $n > k$, it follows that $d_k(n) \ge n$.*

In this paper, we will show that this conjecture has a positive answer.

2. The case when $k < n \le 2k$

Let k and w be fixed such that $1 \le w \le k$, and let $n = k + w$. First, we make two general observations:

(1) For any positive integer a, the set aI_k contains at most $\lfloor\sqrt{k}\rfloor$ many squares. To see this, we notice that the greatest common divisor u, say, of the squares in aI_k is itself a square and is a multiple of a. Hence $u \in aI_k$. Therefore, the squares contained in aI_k are contained in $\{u, 4u, 9u, \ldots, \lfloor\sqrt{k}\rfloor^2 u\}$, and so there are at most $\lfloor\sqrt{k}\rfloor$ of them.

(2) Let s and t be distinct integers with $k < s \leq w$ and $k < t \leq w$. Then $st \notin sI_k$ and $st \notin tI_k$ since $st > sk$ and $st > tk$. Suppose that ℓ is the greatest common divisor of s and t. Then $\ell \leq |t-s| \leq w-1$, and

$$sI_k \cap tI_k \subseteq \{st/\ell, 2st/\ell, \ldots, (\ell-1)st/\ell\}.$$

Therefore, $|sI_k \cap tI_k| \leq \ell - 1$, and the number of cancellations taking place in sI_k and tI_k is at most $2(\ell-1)$.

Since $D_{k\times k} = \{1^2, 2^2, \ldots, k^2\}$, and the number of squares in $(k+1)I_k$ is less than or equal to $\sqrt{k}$, we have $d_k(k+1) \geq 2k - 2\sqrt{k}$. Also, since $k+1$ and $k+2$ are coprime to each other, $(k+1)I_k$ and $(k+2)I_k$ do not have anything in common. Thus $d_k(k+2) \geq 3k - 4\sqrt{k}$. Now, $k+2$ and $k+3$ are also coprime, and $(k+1)I_k$ and $(k+3)I_k$ have at most one element in common. We have $d_k(k+3) \geq 4k - 6\sqrt{k} - 2$. Finally, counting in the possible cancellations among $(k+1)I_k$, $(k+2)I_k$, $(k+3)I_k$, and $(k+4)I_k$, we have $d_k(k+4) \geq 5k - 8\sqrt{k} - 6$. Thus, for $w = 1, 2, 3, 4$, we want

$$\begin{aligned} 2k - 2\sqrt{k} &\geq k+1, \\ 3k - 4\sqrt{k} &\geq k+2, \\ 4k - 6\sqrt{k} - 2 &\geq k+3, \\ 5k - 8\sqrt{k} - 6 &\geq k+4. \end{aligned}$$

As long as $k \geq 9$, the above inequalities hold.

In the following, we assume that $k \geq 9$ and $w \geq 5$. For $2 \leq \ell \leq w-1$, put

$$C_\ell = \{a \mid \ell \text{ divides } a, \text{ and } k+1 \leq a \leq k+w\}.$$

Then $|C_\ell| \leq \lceil \frac{w}{\ell} \rceil$, and so the number of cancellations among the aI_k's, $a \in C_\ell$, can be no more than

$$\binom{\lceil \frac{w}{\ell} \rceil}{2} \cdot 2(\ell-1) = \lceil \tfrac{w}{\ell} \rceil (\lceil \tfrac{w}{\ell} \rceil - 1)(\ell - 1) \leq (\tfrac{w}{\ell} + 1)\tfrac{w}{\ell}(\ell-1) < (\tfrac{w}{\ell} + 1)w.$$

Therefore, the total number of cancellations occurring in $(k+1)I_k, \ldots, (k+w)I_k$ is at most

$$\begin{aligned} \sum_{\ell=2}^{w-1} \binom{\lceil \frac{w}{\ell} \rceil}{2} \cdot 2(\ell-1) &< \sum_{\ell=2}^{w-1} (\tfrac{w}{\ell} + 1)w < \sum_{\ell=2}^{w-1} (\tfrac{w}{\ell} + \tfrac{w}{\ell})w \\ &= \sum_{\ell=2}^{w-1} \frac{2w^2}{\ell} = 2w^2 \sum_{\ell=2}^{w-1} \frac{1}{\ell} < 2w^2 \cdot \ln(w-1). \end{aligned}$$

After cancellations there are at least $kw - 2w^2 \ln(w-1)$ many elements left in $D_{k\times[k+1,k+w]}$.

Now, $D_{k\times(k+w)}$ has at least $2k$ elements as long as $kw - 2w^2 \cdot \ln(w-1) \geq 3k$, or equivalently,

$$k \geq \frac{2w^2 \cdot \ln(w-1)}{w-3}. \tag{2:1}$$

Note that the function $f(x) = \frac{2x^2 \cdot \ln(x-1)}{x-3}$ is increasing for $x \geq 5$. For each given k, let $w_k = \max\{w \geq 1 \mid w \text{ satisfies (2:1)}\}$. Table 1 gives various k and w_k.

Table 1

k	w_k	k	w_k	k	w_k	k	w_k
35	5	100	15	1000	105	10000	753
50	8	200	27	2000	189	20000	1381
70	11	300	38	4000	341	40000	2548
90	13	600	68	8000	620	80000	4725

From Table 1 we know that, for example, if $k = 70$, then each of the sets $D_{k\times(k+s)}$, $5 \leq s \leq 11$, has at least $2k = 140$ elements in there. As another example, let $k = 40000$. Then each of the sets $D_{k\times(k+s)}$, $5 \leq s \leq 2548$, has at least 80000 elements in there.

We note that if a prime p occurs in $\{k+1, \dots, k+s\}$, $1 \leq s \leq k$, then $D_{k\times(k+s)}$ has at least k elements in it as the elements of pI_k cannot be canceled. To see this, we assume that $ps = qt \in pI_k \cap qI_k$ for some integer $q > k$ with $q \neq p$, and $s, t \in I_k$. Then p divides q since $t \leq k < p$. From $p > k$, we infer that $q > 2k$.

Lemma 2.1. *Suppose that $w_k \geq 5$ and that there are two distinct primes among $k+1, \dots, k+w_k$. Then $D_{k\times(k+s)}$ has at least $2k$ elements for all s with $5 \leq s \leq k$.*

Proof. If $5 \leq s \leq w_k$, then we have seen from $k \geq \frac{2s^2\cdot\ln(s-1)}{s-3}$ that $D_{k\times(k+s)}$ has at least $2k$ elements. On the other hand, if $w_k < s \leq k$, then the two primes between $k+1$ and $k+w_k$ give us what we want. $\square$

Therefore, we assume that $w_k \geq 5$, and we would like to have two distinct primes among $k+1$, $k+2$, $\dots$, $k+w_k$. This brings us to the prime gaps consideration.

A *prime gap* is the difference between two successive prime numbers. The n-th prime gap is the difference between the $(n+1)$-th and the n-th prime number. One writes $g(p)$ for the the gap $q - p$, where q is the next prime to p. A prime gap is said to be *maximal* if it is larger than all gaps between smaller primes. The notation for the n-th maximal prime gap is g_n. Table 2 shows g_n for $1 \leq n \leq 15$. For example, for any prime p less than 9551, the prime gap $g(p)$ is less than 36. That is to say that for any prime p with $p < 9551$, there must be a prime in the set $\{p, p+1, \dots, p+35\}$.

Table 2

n	g_n	p_n	n	g_n	p_n	n	g_n	p_n
1	1	2	6	14	113	11	36	9551
2	2	3	7	18	523	12	44	15683
3	4	7	8	20	887	13	52	19609
4	6	23	9	22	1129	14	72	31397
5	8	89	10	34	1327	15	86	155921

If p is a prime with $p > 2k$ and the maximal prime gap $g_m = g(q) < \frac{w_k}{2}$, where q is the first prime to have $g(q) = g_m$, then there must exist at least two primes among $k+1$, $k+2$, $\dots$, $k+w_k$.

Take $k = 300$; then $w_k = 38$. Now, 601 is the prime just bigger then $2k = 600$, and the maximal prime gap for primes less than 887 is at most $g_7 = 18$. Thus, $g(601) \leq 18 < \frac{w_k}{2}$. By Lemma 2.1, $D_{k\times(k+s)}$ contains at least $2k = 600$ elements for any s with $5 \leq s \leq 300$.

For those k which are less than 300, we can check easily using GAP [1] that $d_k(n) \geq n$ for $k < n \leq 2k$. Actually, a small program in GAP running on a modern PC takes about 90 seconds to verify it.

For those k that are greater than 300, we will argue in the following that $d_k(n) \geq n$ for $k < n \leq 2k$ by using the monotone increasing property of w_k and maximal prime gaps.

In [4, p. 368], one finds

$$\text{for } n \geq 2, \qquad p_n \geq n(\ln n + \ln\ln n - 1.0072629) \tag{2:2}$$

and

$$\text{for } n \geq 7022, \qquad p_n \leq n(\ln n + \ln\ln n - 0.9385). \tag{2:3}$$

Therefore, for $n \geq 7022$, we have

$$(n+1)\ln(n+1) - n\ln n = n\ln(\tfrac{1}{n}+1) + \ln(n+1) \leq n\cdot\frac{1}{n} + \ln(n+1) = 1 + \ln(n+1)$$

and

$$\begin{aligned}(n+1)\ln\ln(n+1) - n\ln\ln n &= n\ln(\tfrac{\ln(n+1)}{\ln n}) + \ln\ln(n+1)\\ &\leq n\cdot(\tfrac{\ln(n+1)}{\ln n} - 1) + \ln\ln(n+1)\\ &\leq n\cdot(\tfrac{\ln 7023}{\ln 7022} - 1) + \ln\ln(n+1)\\ &< 0.00002n + \ln\ln(n+1).\end{aligned}$$

The last two inequalities hold since the function $f(x) = \frac{\ln(x+1)}{\ln x}$ is decreasing and $\frac{\ln 7023}{\ln 7022}$ is less than 1.00002. Hence

$$\begin{aligned}g(p_n) &= p_{n+1} - p_n\\ &\leq (n+1)(\ln(n+1) + \ln\ln(n+1) - 0.9385) - n(\ln n + \ln\ln n - 1.0072629)\\ &\leq (1 + \ln(n+1)) + (0.00002n + \ln\ln(n+1)) + 0.0687629n - 0.9385.\end{aligned}$$

As $\ln x + \ln\ln x$ is increasing and concave downward, we have

$$\begin{aligned}&\ln(n+1) + \ln\ln(n+1)\\ &\leq (\ln 7023 + \ln\ln 7023) + (\ln x + \ln\ln x)'|_{7023}\cdot(n - 7022)\\ &< \ln 7023 + \ln\ln 7023 + 0.0001585n - 1.1\\ &< 11 + 0.0001585n,\end{aligned}$$

and so

$$g(p_n) = p_{n+1} - p_n < 12 + 0.069n < 0.071n \quad \text{for } n \geq 7022. \tag{2:4}$$

Now we consider $D_{k\times n}$ with $k > p_{7022} = 70919$ and $k + 5 \leq n = k + w \leq 2k$. If $k \geq \frac{2w^2\ln(w-1)}{w-3}$, then $d_k(n) \geq 2k \geq n$ as we have seen in Lemma 2.1. Thus, suppose that $k < \frac{2w^2\ln(w-1)}{w-3}$. Since $k > 70919$, we certainly have $w > 9$, and so $\ln\ln(3w) = \ln(\ln 3 + \ln w) > \ln(3\ln 3) > 0$. Also, from $w^2 - 9w > 0$, we get $2w^2 < 3w(w-3)$. Combining these with $\ln 3 > 1.098$ (hence $\ln 3 - 1.0072629 > 0$), we get

$$\begin{aligned} 2w^2\ln(w-1) &< 3w(w-3)\ln w \\ &< 3w(w-3)(\ln w + \ln 3 - 1.0072629 + \ln\ln(3w)) \\ &< 3w(w-3)(\ln(3w) + \ln\ln(3w) - 1.0072629) \\ &\leq (w-3)p_{3w} \end{aligned}$$

by (2:2). Therefore,

$$p_{7022} < k < \frac{2w^2\ln(w-1)}{w-3} < p_{3w}.$$

As $g(p_{3w}) < 0.213w$ by (2:4), the prime gaps for the primes between p_{7022} and p_{3w} are all smaller than $0.213w$. This means that there is a prime between k and $k + 0.213w$, and another one between $k + 0.213w$ and $k + 2 \cdot 0.213w$. In particular, there are at least two distinct primes in $[k, k + w]$, and so $D_k(n) \geq 2k \geq n$ again by Lemma 2.1.

Next, suppose that $k \leq p_{7022}$. From the prime gap table above, we see that for any prime p with $k \leq p \leq p_{7022} = 70919 < 155928$, the prime gap $g(p)$ is less than or equal to $g_{14} = 72$. Also, for $2000 < k \leq 70919$, we have $w_k \geq 189 > 144$. Thus, there are at least two primes between k and $k + w_k$.

For any k with $600 < k \leq 2000$, $w_k \geq 68$. For any prime p with $k \leq p \leq 2000 < 9551$, the prime gap $g(p)$ is at most 34. Thus, for any k with $600 < k \leq 2000$, there are at least two primes between k and $k + w_k$.

For k with $300 < k \leq 600$, $w_k \geq 38$, and for any prime p with $k \leq p \leq 600 < 887$, the prime gap $g(p)$ is at most 18. Again, for k with $300 < k \leq 600$, there are at least two primes between k and $k + w_k$.

Therefore, for all k with $300 < k \leq p_{7022}$, there are always two primes between k and $k + w_k$. By Lemma 2.1, $d_k(n) \geq 2k > n$ for any such k and any n with $k + 5 \leq n \leq 2k$.

We have now shown that $d_k(n) > n$ for all k and n with $k > 300$ and $k < n \leq 2k$. Hence we can announce that *the Restricted 1-2-3 Conjecture is true for any* $n, k \in \mathbb{N}$ *with* $k < n \leq 2k$.

Induction kicks in from here. The starting ground is that $d_1(n) = n$ for all $n > 1$. Let $k > 1$, and assume that we have the Restricted 1-2-3 Conjecture verified up to $k - 1$. That is, assume that $d_s(m) \geq m$ when $s \leq k - 1 < m$. We want to show that $d_k(n) \geq n$ for all $n > k$. From the above, we also know that this is true for n up to $2k$. So, we assume that $n > 2k$ and also that $d_k(m) \geq m$ when $2k \leq m < n$. And we continue. . . .

3. A reduction

The first step is to make certain that we do not need to care too much for n large enough. Namely, we will show that it is sufficient to restrict n to be no larger than $\mathrm{LCM}(I_k)$, the least common multiple of $\{1, 2, \ldots, k\}$.

Let T be a nonempty subset of I_k with $|T| = \ell$. Then we have

$$\begin{aligned}\sum_{\varnothing\neq S\subset T}(-2)^{|S|-1} &= \sum_{i=1}^{\ell}\binom{\ell}{i}(-2)^{i-1}\\ &= (-2)^{-1}\cdot\sum_{i=1}^{\ell}\binom{\ell}{i}(-2)^{i}\\ &= (-2)^{-1}\cdot(((-2)+1)^{\ell}-1)\\ &= (-2)^{-1}\cdot((-1)^{\ell}-1)\\ &= \begin{cases}0, & \text{if } \ell \text{ is even;}\\ 1, & \text{if } \ell \text{ is odd.}\end{cases}\end{aligned}$$

For convenience and by abusing notation, if $T = \varnothing$, we put $\sum_{\varnothing\neq S\subset T}(-2)^{|S|-1} = 0$. Using this identity, we have

Theorem 3.1. *Let $n \geq k$. Then*

$$d_k(n) = \sum_{\varnothing\neq S\subset I_k}\left\lfloor\tfrac{n\cdot\min S}{\mathrm{LCM}(S)}\right\rfloor\cdot(-2)^{|S|-1}. \tag{3:1}$$

Proof. Denote by 2^{I_k} the power set of I_k. Define $\theta : \mathbb{N} \to 2^{I_k}$ by $\theta(m) = \{s \in I_k \mid m \in sI_n\}$. For $\varnothing \neq S \subseteq I_k$, set

$$\widetilde{S} = \theta^{-1}(S) = \{m \in \mathbb{N} \mid \theta(m) = S\}$$

and

$$\overline{S} = \{m \in \mathbb{N} \mid \theta(m) \supseteq S\}.$$

Note that if $\varnothing \neq S \subseteq T$, then $m \in \overline{S}$ whenever $m \in \widetilde{T}$, and $\overline{S}$ is the disjoint union of $\widetilde{T}$ for subsets T of I_k containing S. Therefore, $|\overline{S}| = \sum_{S\subseteq T\subseteq I_k}|\widetilde{T}|$.

An integer $m \geq 1$ will appear in $D_{k\times n}$ if and only if $\theta(m)$ is a nonempty set with an odd number of elements in it. Therefore, we have

$$\begin{aligned}d_k(n) &= \sum_{m\geq 1}\Big(\sum_{\varnothing\neq S\subseteq\theta(m)}(-2)^{|S|-1}\Big)\\ &= \sum_{\varnothing\neq T\subseteq I_k}|\widetilde{T}|\cdot\Big(\sum_{\varnothing\neq S\subseteq T}(-2)^{|S|-1}\Big)\\ &= \sum_{\varnothing\neq T\subseteq I_k}\sum_{\varnothing\neq S\subseteq T}|\widetilde{T}|\cdot(-2)^{|S|-1}\\ &= \sum_{\varnothing\neq S\subseteq I_k}\Big(\sum_{S\subseteq T\subseteq I_k}|\widetilde{T}|\Big)\cdot(-2)^{|S|-1}\\ &= \sum_{\varnothing\neq S\subseteq I_k}|\overline{S}|\cdot(-2)^{|S|-1}.\end{aligned}$$

To finish the proof, we notice that for any nonempty subset S of I_k,

$$\begin{aligned}\overline{S} &= \{m \geq 1 \mid m \in sI_n \text{ for all } s \in S\}\\ &= \{m \geq 1 \mid \mathrm{LCM}(S) \text{ divides } m \text{ and } m \leq sn \text{ for all } s \in S\}.\end{aligned}$$

Therefore, $|\overline{S}| = \lfloor\frac{n\cdot\min S}{\mathrm{LCM}(S)}\rfloor$. □

Suppose that $n = a + b \cdot \mathrm{LCM}(I_k)$, where $a, b \in \mathbb{N}$ with $a \leq \mathrm{LCM}(I_k)$. For any nonempty subset S of I_k, since $\mathrm{LCM}(S)$ divides $\mathrm{LCM}(I_k)$, $\frac{b \cdot \mathrm{LCM}(I_k) \cdot \min S}{\mathrm{LCM}(S)}$ is an integer, and we have

$$\begin{aligned}\left\lfloor \tfrac{n \cdot \min S}{\mathrm{LCM}(S)} \right\rfloor &= \left\lfloor \tfrac{(a + b \cdot \mathrm{LCM}(I_k)) \cdot \min S}{\mathrm{LCM}(S)} \right\rfloor = \left\lfloor \tfrac{a \cdot \min S}{\mathrm{LCM}(S)} + \tfrac{b \cdot \mathrm{LCM}(I_k) \cdot \min S}{\mathrm{LCM}(S)} \right\rfloor \\ &= \left\lfloor \tfrac{a \cdot \min S}{\mathrm{LCM}(S)} \right\rfloor + \tfrac{b \cdot \mathrm{LCM}(I_k) \cdot \min S}{\mathrm{LCM}(S)} = \left\lfloor \tfrac{a \cdot \min S}{\mathrm{LCM}(S)} \right\rfloor + b \cdot \left\lfloor \tfrac{\mathrm{LCM}(I_k) \cdot \min S}{\mathrm{LCM}(S)} \right\rfloor.\end{aligned}$$

This makes $d_k(a + b \cdot \mathrm{LCM}(I_k)) = d_k(a) + b \cdot d_k(\mathrm{LCM}(I_k))$. If we can show that $d_k(n) \geq n$ for all n with $n \leq \mathrm{LCM}(I_k)$, then we are done.

With the above preparation, we make the assumption that

$$2k < n \leq \mathrm{LCM}(I_k)$$

for the rest of the paper, and move on.

4. The case when $n > 2k$

We start with an easy observation.

Lemma 4.1. *Let p and q be distinct primes which are greater than $\max\{k, \sqrt{n}\}$ and less than or equal to n. Then for any $s, t \in \mathbb{N}$ with $s \leq \lfloor \frac{n}{p} \rfloor$ and $t \leq \lfloor \frac{n}{q} \rfloor$, we have $(sp)I_k \cap (tq)I_k = \varnothing$.*

Proof. Suppose that $spa = tqb$ for some $a, b \in I_k$. Then $p \mid tb$. Since p is a prime larger than k, we must have $p \mid t$. From $t \leq \lfloor \frac{n}{q} \rfloor$, we reach a contradiction that $pq \leq tq \leq n$. □

Remark 4.2. Suppose that p is a prime such that $\max\{k, \sqrt{n}\} < p \leq n$. Then

$$pI_k \,\Delta\, 2pI_k \,\Delta \cdots \Delta\, \lfloor \tfrac{n}{p} \rfloor pI_k = p(I_k \,\Delta\, 2I_k \,\Delta \cdots \Delta\, \lfloor \tfrac{n}{p} \rfloor I_k),$$

which has at least $\max\{k, \lfloor \frac{n}{p} \rfloor\}$ many elements by the induction hypothesis. Combining this with Lemma 4.1, our goal is then to show that

$$\sum_{p \in \mathcal{P}} \max\{k, \lfloor \tfrac{n}{p} \rfloor\} \geq n,$$

where $\mathcal{P}$ is defined to be

$$\mathcal{P} = \{p \mid p \text{ is a prime and } \max\{k, \sqrt{n}\} < p \leq n\}.$$

We will use the following results from number theory, where $\pi(x)$ denotes the number of primes less than or equal to x.

Proposition 4.3 ([2, Theorem 2 and Corollary to Theorem 3] and [5, (3.5), (3.6), (3.8)])**.**

(1) $2^k \leq \mathrm{LCM}(I_k) \leq 4^k$.
(2) $\pi(x) > x/\ln x$ *for* $x \geq 17$.
(3) $\pi(x) < 1.25506x/\ln x$ *for* $x > 1$.
(4) $\pi(2x) - \pi(x) > 3x/(5 \ln x)$ *for* $x > 20.5$.

4.1. **The case when $\mathbf{2k < n \leq k^2}$.** Let $\lceil \frac{n}{k} \rceil = m$. Thus, $m \geq 3$ and $km \geq n$. Since each prime in $\mathcal{P}$ contributes at least k elements to the set $D_{k\times n}$ (see Remark 4.2), we aim to show that $|\mathcal{P}| \geq m$.

Therefore, we want $\pi(n) - \pi(k) \geq m$. It is easy to verify that $d_k(n) \geq n$ for k and n with $k \leq 20$ and $2k < n \leq k^2$. So we assume that $k \geq 21$.

Let $v = m - 1$. Then $2 \leq v \leq k - 1$ and $n \geq vk + 1$, and the goal is to show that $\pi(vk + 1) - \pi(k) \geq v + 1$. For $v = 2$, 3 or 4, we have

$$\pi(vk + 1) - \pi(k) \geq \pi(2k) - \pi(k) \geq 3k/(5 \ln k) \geq 3 \cdot 21/(5 \ln 21) > 4.1.$$

Since $\pi(vk + 1) - \pi(k)$ is an integer, it is at least 5, which is greater than $v - 1$.

On the other hand, for $v \geq 5$ we have

$$\begin{aligned}
\pi(vk+1) - \pi(k) \geq \pi(vk) - \pi(k) &\geq \frac{vk}{\ln(vk)} - \frac{1.25506k}{\ln k} \\
&> \frac{vk}{\ln(k^2)} - \frac{1.3k}{\ln k} > \frac{vk}{2\ln k} - \frac{1.3k}{\ln k} \\
&= \frac{k}{\ln k} \cdot \big(\frac{v}{2} - 1.3\big) \geq \frac{21}{\ln 21} \cdot \big(\frac{v}{2} - 1.3\big) \\
&> 6 \cdot \big(\frac{v}{2} - 1.3\big) > v + 1.
\end{aligned}$$

Thus the case when $2k < n \leq k^2$ is done.

4.2. **The case when $\mathbf{n > k^2}$.** In this case, $\max\{k, \sqrt{n}\} = \sqrt{n}$. Let $a_0 = n$, $a_i = \frac{n}{k+i}$ for $i = 1, \ldots, \ell$, and $a_{\ell+1} = \sqrt{n}$, where ℓ is the largest integer such that $\frac{n}{k+\ell+1} \leq \sqrt{n} < \frac{n}{k+\ell}$. Here we have $\lfloor \frac{n}{p} \rfloor \leq k$ if $a_1 < p \leq a_0$, and $\lfloor \frac{n}{p} \rfloor = k + i$ if $a_{i+1} < p \leq a_i$ $(i = 1, 2, \ldots, \ell)$; therefore,

$$\begin{aligned}
\sum_{p\in\mathcal{P}} \max\{k, \lfloor \tfrac{n}{p} \rfloor\} &= \sum_{\substack{\sqrt{n}<p\leq n \\ p \text{ is prime}}} \max\{k, \lfloor \tfrac{n}{p} \rfloor\} \\
&= \sum_{a_1<p\leq a_0} \max\{k, \lfloor \tfrac{n}{p} \rfloor\} + \sum_{a_2<p\leq a_1} \max\{k, \lfloor \tfrac{n}{p} \rfloor\} + \cdots + \sum_{a_{\ell+1}<p\leq a_\ell} \max\{k, \lfloor \tfrac{n}{p} \rfloor\} \\
&= \sum_{a_1<p\leq a_0} k + \sum_{a_2<p\leq a_1} \lfloor \tfrac{n}{p} \rfloor + \cdots + \sum_{a_{\ell+1}<p\leq a_\ell} \lfloor \tfrac{n}{p} \rfloor \\
&= \sum_{a_1<p\leq a_0} k + \sum_{a_2<p\leq a_1} (k+1) + \cdots + \sum_{a_{\ell+1}<p\leq a_\ell} (k+\ell) \\
&= \sum_{i=0}^{\ell} (k+i)(\pi(a_i) - \pi(a_{i+1})) \\
&= \sum_{i=0}^{\ell} k(\pi(a_i) - \pi(a_{i+1})) + \sum_{i=1}^{\ell} i(\pi(a_i) - \pi(a_{i+1})) \\
&= \Big(k\pi(a_0) - k\pi(a_{\ell+1})\Big) + \Big(\Big(\sum_{i=1}^{\ell} \pi(a_i)\Big) - \ell\pi(a_{\ell+1})\Big) \\
&= k\pi(n) + \Big(\sum_{i=1}^{\ell} \pi(a_i)\Big) - (k+\ell)\pi(\sqrt{n}).
\end{aligned}$$

Using Proposition 4.3, we have

$$\pi(\sqrt{n}) < 1.25506 \cdot \frac{\sqrt{n}}{\ln\sqrt{n}} < 2.52 \cdot \frac{\sqrt{n}}{\ln n},$$

and if $n \geq 17$,

$$\begin{aligned}\sum_{i=1}^{\ell} \pi(\tfrac{n}{k+i}) > \sum_{i=1}^{\ell} \frac{\frac{n}{k+i}}{\ln(\frac{n}{k+i})} > \sum_{i=1}^{\ell} \frac{\frac{n}{k+i}}{\ln n} = \frac{n}{\ln n} \sum_{i=1}^{\ell} \tfrac{1}{k+i} \\ > \frac{n}{\ln n}(\ln(k+\ell+1) - \ln(k+1)).\end{aligned}$$

The last inequality came from $\sum_{i=1}^{\ell} \frac{1}{k+i} > \int_{k+1}^{k+\ell+1} \frac{1}{x} dx$. Also, from $\frac{n}{k+\ell+1} \leq \sqrt{n} < \frac{n}{k+\ell}$ we have $k+\ell < \sqrt{n} \leq k+\ell+1$, and so $\frac{1}{2}\ln n = \ln(\sqrt{n}) \leq \ln(k+\ell+1)$. This yields

$$\begin{aligned}\sum_{p\in\mathcal{P}} \max\{k, \lfloor \tfrac{n}{p} \rfloor\} > k\frac{n}{\ln n} + \frac{n}{\ln n}(\ln(k+\ell+1) - \ln(k+1)) - (k+\ell)\frac{\sqrt{n}}{\ln n} \cdot 2.52 \\ \geq \frac{n}{2} + \frac{n}{\ln n} \cdot (k - \ln(k+1) - 2.52).\end{aligned}$$

To finish the task, we need only to have $\frac{k-\ln(k+1)-2.52}{\ln n} \geq \frac{1}{2}$. To this end, we use the fact that $n \leq \mathrm{LCM}(I_k) < 4^k$. So

$$\frac{k - \ln(k+1) - 2.52}{\ln n} \geq \frac{k - \ln(k+1) - 2.52}{\ln(\mathrm{LCM}(I_k))} > \frac{k - \ln(k+1) - 2.52}{k \cdot \ln 4}.$$

Now,

- $\dfrac{k - \ln(k+1) - 2.52}{k \cdot \ln 4}$ is increasing for all k and is more than $\frac{1}{2}$ when $k = 18$.
- $\dfrac{k - \ln(k+1) - 2.52}{\ln(\mathrm{LCM}(I_k))} \geq \dfrac{1}{2}$ for $k = 8, 9, \ldots, 17$.

So, we see that, indeed,

$$\sum_{p\in\mathcal{P}} \max\{k, \lfloor \tfrac{n}{p} \rfloor\} \geq n \text{ when } n \geq k^2 \geq 64.$$

For $k < 8$, we have to check $d_k(n)$ for $k^2 < n \leq \mathrm{LCM}(I_k)$. Since $k^2 > \mathrm{LCM}(I_k)$ for $k = 2, 3, 4$, this amounts to checking the cases $k = 5$ with $25 < n \leq \mathrm{LCM}(\mathbf{5}) = 60$, $k = 6$ with $36 < n \leq 60$, and $k = 7$ with $49 < n \leq 420$. Again, a simple computer routine verifies that these are all fine. Therefore, we conclude that *the Restricted 1-2-3 Conjecture is true for all k and n with* $n \geq k^2$, and as well conclude our proof for the *Restricted 1-2-3 Conjecture*.

Finally, for the "Codes by composition" (see the introduction), one needs even more. Here, we mention as an open problem:

Extended 1-2-3 Conjecture. For every finite, nonempty subset I of the natural numbers, the symmetric difference of the sets iI_k, $i \in I$, has at least k elements.

Note that it is not true if one changes "at least k elements" to "at least k or n elements, whichever is larger"!

References

[1] The GAP Group, GAP—Groups, Algorithms, and Programming, Version 4.4.12, 2008, http://www.gap-system.org.

[2] M. Nair, *On Chebyshev-type inequalities for primes.* Amer. Math. Monthly **89** (1982), 126–129. MR643279 (83f:10043)

[3] G. Pilz, *On polynomial near-ring codes*, in Contributions to General Algebra, 8, Hölder-Pichler-Tempsky, Vienna, 1992, 233–238. MR1281844 (95e:11131)

[4] G. Robin, *Estimate of the Chebyshev function θ on the kth prime number and large values of the number of prime divisors function $\omega(n)$ of n.* Acta Arith. **42** (1983), 367–389. MR736719 (85j:11109)

[5] J. B. Rosser and L. Schoenfeld. *Approximate formulas for some functions of prime numbers.* Illinois J. Math. **6** (1962), 64–94. MR0137689 (25:1139)

Department of Mathematics and National Center for Theoretical Sciences (South), National Cheng Kung University, 1 University Road, Tainan 701, Taiwan

E-mail address: pyhuang@mail.ncku.edu.tw

Department of Mathematics and National Center for Theoretical Sciences (South), National Cheng Kung University, 1 University Road, Tainan 701, Taiwan

E-mail address: wfke@mail.ncku.edu.tw

Department of Algebra, Johannes Kepler Universität Linz, Altenberger Strasse 69, 4040 Linz, Austria

E-mail address: guenter.pilz@jku.at