INTEGERS: ELECTRONIC JOURNAL OF COMBINATORIAL NUMBER THEORY 6 (2006), #A28

AN IMPROVED UPPER BOUND ON THE MAXIMUM SIZE OF k-PRIMITIVE SETS

Peter Hegarty
Chalmers University of Technology, 41296 Göteborg, Sweden
hegarty@math.chalmers.se

Received: 3/7/06, Revised: 9/11/06, Accepted: 9/21/06, Published: 10/13/06

Abstract

For $k \geq 3$ and sufficiently large n depending on k, Sujith Vijay recently provided non-trivial lower and upper bounds for the size of the largest subset of $\{1, 2, \ldots, n\}$ such that no element divides k others. The gap between his lower and upper bounds is, however, substantial, and here we provide a better upper bound which significantly reduces this gap for large k.

1. Introduction

Let n be a positive integer. A subset S of $\{1, \ldots, n\}$ is said to be *primitive* if no member of S divides any other. It is a well-known application of the pigeonhole principle, or if one prefers, of Dilworth's theorem, that $|S| \leq \lceil n/2 \rceil$; the divisibility poset on $\{1, \ldots, n\}$ can be decomposed into $\lceil n/2 \rceil$ disjoint chains $\mathcal{C}_x$, one for each odd number $x \in \{1, \ldots, n\}$, where $\mathcal{C}_x = \{1, \ldots, n\} \cap \{2^k x : k \geq 0\}$. On the other hand, the set $S_{0,n} = \{x : n/2 < x \leq n\}$ is primitive, so $\lceil n/2 \rceil$ is the maximum size of a primitive subset of $\{1, \ldots, n\}$.

Now let $k \geq 2$ be an integer. A subset S of $\{1, \ldots, n\}$ is said to be *k-primitive* if no member of S divides k or more others. Following [5], we denote by $f_k(n)$ the maximum size of a k-primitive subset of $\{1, \ldots, n\}$. No simple formula for $f_k(n)$, for any fixed $k \geq 2$, is known (see [2], Problem B24). A trivial lower bound is found by generalizing the set $S_{0,n}$ above: the set $\{x : n/(k+1) < x \leq n\}$ is k-primitive. Larger k-primitive sets are constructed explicitly in [5], and it is deduced that, for all sufficiently large n, depending on k,

$$\frac{f_k(n)}{n} \geq \frac{k}{k+1} + O\left(\frac{1}{k^4}\right).$$

In the special case $k = 2$, Lebensold [3] had earlier provided an even better construction. He also used the decomposition of the divisibility poset referred to above to provide an upper bound in this special case. He obtains the result that, for sufficiently large n,

$$0.6725\ldots.. \leq \frac{f_2(n)}{n} \leq 0.6736\ldots.$$

and remarks that the gap between these can be improved simply by resort to more computation, rather than to any new ideas. Now as k grows, it is easy to see that the type of argument used in [3] to obtain an upper bound on $f_k(n)/n$ becomes less and less effective. A non-trivial upper bound is provided in [5] for all k but, as we shall see when we quote the result below, it leaves a substantial 'gap' between the lower and upper bounds in a 'non-computational' sense. In this note we will modify the argument in [5] which produces a new upper bound that closes this gap significantly for large k.

For convenience we set $g_k(n) := 1 - \frac{f_k(n)}{n}$ (this reverses the role of upper and lower bounds: we hope no confusion arises in what follows). The following results are proven in [5] :

Theorem A *(i) For each $k \geq 3$ and sufficiently large n,*

$$\frac{1}{8k \ln k} < g_k(n) < \frac{1}{k+1} - \frac{1}{8k^4}. \tag{1}$$

(ii) For each $\epsilon > 0$ there exists $k_0(\epsilon)$ such that, for each $k \geq k_0(\epsilon)$ and all $n \geq n_0(k)$,

$$\frac{1}{(2e^\gamma + \epsilon)k \ln k} < g_k(n) < \frac{1}{k+1} - \frac{1-\epsilon}{k^4}, \tag{2}$$

where γ is Euler's constant.

The point is that the lower and upper bounds in (1) and (2) have different orders of magnitude as functions of k. We will prove better lower bounds which eliminate this gap. Our result is the following

Theorem B *(i) There exists an absolute constant $C > 0$ such that, for all $k \geq 3$, and all n sufficiently large, depending on k,*

$$\frac{1}{Ck} < g_k(n). \tag{3}$$

In fact, one can take C to be the unique solution in $[2, \infty)$ of the equation

$$\frac{1}{t} = \frac{5}{9}\left[\ln\left(\frac{2t}{t+2}\right) - \frac{11(t-2)}{20t}\right], \tag{4}$$

namely $C \approx 18.1439\ldots$
(ii) Let $\epsilon > 0$. Then there exists $k_0(\epsilon)$ such that, for each $k \geq k_0(\epsilon)$ and all $n \geq n_0(k)$,

$$\frac{1-\epsilon}{k} < g_k(n). \tag{5}$$

2. Proof of Theorem B

Although inequality (5) is a stronger result than (3), except for the explicit value of the constant C, we shall first prove (3) separately. A somewhat technical modification of the argument will then provide (5).

Proof of part (i). For given k and n, let S be a k-primitive subset of $\{1,\ldots,n\}$ and denote its complement inside $\{1,\ldots,n\}$ by S'. Let $c > 2$ be a real number and suppose that $|S'| < n/ck$. We will show that, for sufficiently large n, depending on k, it must hold that

$$\frac{1}{c} \geq \frac{k}{2(k-1)}\left[\ln\left(\frac{2c}{c+2}\right) - \frac{c-2}{2c}\left(1+\frac{1}{k}\right)\right]. \tag{6}$$

This says that $c \leq c_k$ for some $c_k \in [2,\infty)$, where c_{10} is the solution C to (4). For $k \geq 10$ this means that $c \leq c_k \leq c_{10} = C$. This, together with (1) and the fact that $8\ln k < c_{10}$ for all $k < 10$, gives the desired result.

Let $I := \mathbf{Z} \cap \left(\frac{n}{2k}, \frac{n}{k}\right]$. For each $x \in I$ we let $\mathcal{C}_x := \{jx : 1 \leq j \text{ and } jx \leq n\}$. Clearly,

$$k \leq |\mathcal{C}_x| = \lfloor \frac{n}{x} \rfloor \leq 2k-1.$$

Since $|I| = n/2k + O(1)$, it follows by our assumptions that $|I \cap S| > \frac{n}{k}\left(\frac{c-2}{2c}\right) - O(1)$. For each $x \in I \cap S$ let $\mathcal{D}_x := \mathcal{C}_x \cap S'$. Since S is k-primitive we have $|\mathcal{D}_x| \geq |\mathcal{C}_x| - k = \lfloor \frac{n}{x} \rfloor - k$. Thus

$$\begin{aligned}
\sum_{x \in I\cap S} |\mathcal{D}_x| &\geq \sum_{x=\lfloor \frac{n}{k}\rfloor - |I\cap S|+1}^{\lfloor \frac{n}{k}\rfloor} \left\lfloor \frac{n}{x} \right\rfloor - k \\
&\geq \sum_{x=\frac{n}{k}\left(\frac{c+2}{2c}\right)+O(1)}^{\frac{n}{k}+O(1)} \lfloor \frac{n}{x} \rfloor - k \\
&\geq n\cdot\left[\int_{\frac{n}{k}\left(\frac{c+2}{2c}\right)}^{\frac{n}{k}} \frac{dx}{x} + O\left(\frac{k}{n}\right)\right] - (k+1)\cdot\left[\frac{n}{k}\left(\frac{c-2}{2c}\right) + O(1)\right] \\
&= n\cdot\left[\ln\left(\frac{2c}{c+2}\right) - \frac{c-2}{2c}\left(1+\frac{1}{k}\right)\right] + O(k).
\end{aligned}$$

Now, by definition, $\bigcup_{x\in I\cap S} \mathcal{D}_x \subseteq S'$. Let y be a member of this union. Set $A_y := \{x \in I\cap S : y \in \mathcal{D}_x\}$ and $a_y := |A_y|$. We have an injection $A_y \hookrightarrow \{2,\ldots,2k-1\}$ given by $y \mapsto y/x$, hence $a_y \leq 2(k-1)$ for every y. It follows that

$$\left|\bigcup_{x\in I\cap S} \mathcal{D}_x\right| \geq \frac{1}{2(k-1)} \sum_{x\in I\cap S} |\mathcal{D}_x|$$

and hence that

$$\frac{n}{ck} > |S'| \geq \frac{n}{2(k-1)}\left[\ln\left(\frac{2c}{c+2}\right) - \frac{c-2}{2c}\left(1+\frac{1}{k}\right)\right] + O(1).$$

Thus we'll obtain a contradiction for sufficiently large n unless (6) is satisfied, so we're done.

Proof of part (ii). Let $\epsilon \in (0, 1)$ be given. We consider a fixed, sufficiently large k and in turn a sufficiently large n (how large will be determined in due course) and let S be a k-primitve subset of $\{1, \ldots, n\}$ with complement S'. We suppose that $|S'| < (1-\epsilon)\frac{n}{k}$ and aim to derive a contradiction.

Let $I := \mathbf{Z} \cap \left(\frac{\epsilon}{2} \cdot \frac{n}{k}, \frac{n}{k}\right]$. It follows from our assumptions that $|I \cap S| > \frac{\epsilon}{2} \cdot \frac{n}{k} - O(1)$. For each $x \in I \cap S$ define the chain $\mathcal{C}_x$ and its subset $\mathcal{D}_x$ as before. Arguing as before we have that

$$\sum_{x \in I \cap S} |\mathcal{D}_x| \geq n \cdot \left[\int_{\frac{n}{k}\left(1-\frac{\epsilon}{2}\right)}^{\frac{n}{k}} \frac{dx}{x} + O\left(\frac{k}{\epsilon n}\right)\right] - (k+1) \cdot \left[\frac{\epsilon}{2} \cdot \frac{n}{k} + O(1)\right]$$
$$= n \cdot \left[\ln\left(\frac{2}{2-\epsilon}\right) - \frac{\epsilon}{2}\left(1 + \frac{1}{k}\right)\right] + O\left(\frac{k}{\epsilon}\right).$$

Hence for any fixed $k \geq \Theta\left(\frac{1}{\epsilon}\right)$ we have, for all sufficiently large n,

$$\sum_{x \in I \cap S} |\mathcal{D}_x| \geq \Theta(\epsilon^2 n). \tag{7}$$

On the other hand,

$$\left|\bigcup_{x \in I \cap S} \mathcal{D}_x\right| \leq |S'| < \frac{\epsilon n}{k}. \tag{8}$$

For each $y \in \bigcup_{x \in I \cap S} \mathcal{D}_x$ we define the set A_y as before. The map $y \mapsto y/x$ gives an injection of A_y into $\{2, 3, \ldots, \lfloor \frac{2k}{\epsilon} \rfloor\}$, hence $a_y < 2k/\epsilon$ for each y. On the other hand, (7) and (8) imply that, on average, $a_y \geq \Theta(\epsilon k)$. At the very least, we can thus deduce the following :

'There are at least $\Theta\left(\frac{\epsilon^3}{k} n\right)$ numbers in $\{1, \ldots, n\}$, each with at least $\Theta(\epsilon k)$ divisors in $\{1, \ldots, \lfloor \frac{2k}{\epsilon} \rfloor\}$'.

But now we have the desired contradiction, as the above statement cannot possibly be true, for any fixed $\epsilon \in (0, 1]$, any fixed k sufficiently large depending on ϵ, and all sufficiently large n depending on k. This follows immediately from the following lemma :

Lemma *Let $\delta \in (0, 1)$. For each pair l, n of positive integers, set*

$$A(n, l) := \{x \in \{1, \ldots, n\} : x \text{ has at least } \delta \cdot l \text{ divisors in } \{1, \ldots, l\}\}.$$

Then for each $l >>_\delta 0$,

$$\lim_{n \to \infty} \frac{|A(n, l)|}{n} = 0. \tag{9}$$

Proof. The result follows from a number of standard estimates in analytic number theory, but for the sake of completeness we present the argument in detail. First recall some standard notation : for each pair (x, y) of positive real numbers with $1 \leq y \leq x$, we denote by $\Psi(x, y)$ the number of integers in $[1, x]$ all of whose prime divisors lie in $[1, y]$. It was first proven by Dickman [1] that, if $\rho : [0, \infty) \to (0, 1]$ is the unique continuous function satisfying

$$\rho(u) = 1, \quad \text{for } 0 \leq u \leq 1,$$
$$\rho(u-1) + u\frac{d\rho}{du} = 0, \quad \text{for } 1 < u < \infty,$$

then, for each $u \in (1, \infty)$, $\lim_{x\to\infty} \frac{\Psi(x, x^{1/u})}{x} = \rho(u)$. From this and the prime number theorem it follows easily that, for any $\delta' \in (0, 1/\rho^{-1}(\delta))$ and all sufficently large l (depending on δ, δ'), if $\mathcal{A}$ is a subset of $\{1, \ldots, l\}$ of size at least $\delta \cdot l$, then the numbers in $\mathcal{A}$ contain in total at least $l^{\delta'}$ distinct prime factors. Now consider a fixed large l and suppose (9) were false. Then for all large n there would be a positive proportion of the integers in $\{1, \ldots, n\}$ each divisible by at least $l^{\delta'}$ distinct primes in $\{1, \ldots, l\}$. A positive proportion of these will in turn each have at least $\Theta_l\left[(\ln n)^{l^{\delta'}-1}\right]$ distinct divisors. But for all sufficiently large n this will contradict the result of Dirichlet (see any standard number theory text, for example [4], Theorem 8.28) that $\frac{1}{n}\sum_{i=1}^{n} d(i) \sim \ln n$, where $d(i)$ denotes the number of divisors of i. This completes the proof of the lemma, and with it that of Theorem B.

3. Concluding Remarks

The trivial lower bound $g_k(n) \leq \frac{1}{k+1}$ gives a lower bound on the function $k_0(\epsilon)$ in part (ii) of Theorem B of the form

$$k_0(\epsilon) \gtrsim \frac{1}{\epsilon}. \tag{10}$$

The lower bound obtained by the above proof is, however, much worse. This is because, in the proof of the Lemma, we require at the very least that $l^{\frac{1}{\rho^{-1}(\delta)}} = \Omega(1)$. This in turn implies a lower bound of the form

$$k_0(\epsilon) = \Omega\left[\epsilon \cdot \exp\left\{\rho^{-1}(c\epsilon^2)\right\}\right], \tag{11}$$

for some absolute constant c. The gap between (10) and (11) remains a challenging problem.

References

[1] K. Dickman, On the frequency of numbers containing prime factors of a certain relative magnitude, *Ark. Mat. Astron. Fys.* **22A**, no.10, (1930), 1-14.
[2] R.K. Guy, Unsolved Problems in Number Theory (second edition). Springer Verlag, 1994.
[3] K. Lebensold, A divisibility problem, *Studies in Appl. Math.* **56** (1976/77), no.3, 291-294.
[4] I. Niven, H. Zuckerman and H. Montgomery, Introduction to the Theory of Numbers (fifth edition). Wiley, 1991.
[5] S. Vijay, On the largest k-primitive subset of $[1, n]$, *Integers : Electronic Journal of Combinatorial Number Theory* **6** (2006), Paper A1.