On the number of co-prime-free sets.

by

Neil J. Calkin and Andrew Granville *

Abstract: For a variety of arithmetic properties P (such as the one in the title) we investigate the number of subsets of the positive integers $\leq x$, that have that property. In so doing we answer some questions posed by Cameron and Erdös.

1. Introduction.

In [CE] Cameron and Erdös investigated subsets of the positive integers $\leq x$ with certain given properties P; in particular, how large such sets can be, and how many there are. The properties P that they were interested in are monotone decreasing, that is, if S has property P, and T is a subset of S, then T has property P. Thus if S is a maximal set of positive integers $\leq x$ with property P then one knows that there are $\geq 2^{|S|}$ such sets. In this paper we improve various estimates in [CE] for the number of sets satisfying certain properties P:

In Theorem 3.5 of [CE], Cameron and Erdös showed that the number of sets of positive integers $\leq x$, in which any two elements have a common factor, lies between $2^{[x/2]}$ and $x2^{[x/2]}$. Here we improve this to

Theorem 1. *The number of sets of integers $\leq x$, with any two elements having a common factor, is*

$$2^{[x/2]} + 2^{[x/2]-N} + O\left(2^{[x/2]-N} \exp\left(-C\frac{x}{\log^2 x \log\log x}\right)\right), \tag{1.1}$$

---

* The second author is supported, in part, by the National Science Foundation (grant number DMS-8610730)

for some absolute constant $C > 0$, where N, which will be defined in the proof, satisfies

$$N = (e^{-\gamma} + o(1))\frac{x}{\log\log x}, \tag{1.2}$$

and γ is the Euler-Mascheroni constant.

In Theorem 3.3 of [CE], Cameron and Erdös showed that the number of sets of positive integers $\leq x$, in which any two elements are coprime, lies between $2^{\pi(x)}e^{(1/2+o(1))\sqrt{x}}$ and $2^{\pi(x)}e^{(2+o(1))\sqrt{x}}$. Here we improve this to

Theorem 2. *The number of sets of integers $\leq x$, with every pair of elements coprime, is*

$$2^{\pi(x)}e^{\sqrt{x}\{1+O(\log\log x/\log x)\}}. \tag{1.3}$$

In Section 4.3 of [CE], Cameron and Erdös conjectured that there are $c(s)^{x+o(x)}$ sets of integers$\leq x$, with sum of reciprocals bounded by s, for some positive constant $c(s)$. We prove a quantitatative form of this conjecture here:

Theorem 3. *The number, $\nu(x)$, of sets $\{a_1, a_2, \ldots a_t\}$ of positive integers $\leq x$, with $\sum_i 1/a_i \leq s$, is*

$$c(s)^x e^{O(x^{3/4})}, \quad \text{where } c(s) = \left(1 + e^{-f(s)}\right), \tag{1.4}$$

and $f(s)$ is defined by

$$s = \int_{f(s)}^{\infty} \frac{\mathrm{d}u}{u(1+e^u)}. \tag{1.5}$$

Cameron and ErdoS̈ observed that $c(s) \geq 2^{1-e^{-s}}$, for all s. We can provide some rather more accurate estimates based on an analysis of (1.4) and (1.5): As $s \to 0$, we have

$$c(s) = 1 + s\left(\log(1/s) + \log\log(1/s) + O(1)\right); \tag{1.6}$$

as $s \to \infty$, we have

$$c(s) = 2 - Ce^{-2s} + O(e^{-4s}), \tag{1.7}$$

for some constant $C \approx .8819384944\ldots$.

2. Sets where any two elements have a common factor.

Proof of Theorem 1: Clearly there are $2^{[x/2]}$ such sets that contain only even numbers. We now count such sets that contain at least one odd number:

Let k be the largest integer for which $q = p_1 p_2 \ldots p_k \leq x$ (where p_j is the jth smallest *odd* prime); by the Prime Number Theorem, $k \sim \log x/\log\log x$. Define

$$\mathcal{R} = \{n \leq x \ : \ n \text{ is divisible by } 2p_j \text{ for some } j \leq k\}.$$

Clearly any set of the form $S \cup \{q\}$, where S is a subset of $\mathcal{R}$, has the property that any two elements have a common factor. The number of such subsets S is $2^{|\mathcal{R}|}$, and $|\mathcal{R}| = [x/2] - N$ where N is the number of integers $2m \leq x$ that are coprime to q. From the combinatorial sieve we obtain $N \sim \frac{\phi(q)}{q}\frac{x}{2}$, and then Mertens' Theorem implies (1.2).

Our proof of the rest of Theorem 1 is based on the ideas of Pomerance given in Cameron and Erdös [CE]:

We start by ordering the odd numbers $\leq x$ as $q = m_1, m_2, \ldots$ so that

$$\frac{\phi(m_1)}{m_1} \leq \frac{\phi(m_2)}{m_2} \leq \frac{\phi(m_3)}{m_3} \leq \ldots.$$

Define $f_i(x)$ to be the number of sets of integers $\leq x$, that contain m_i but not $m_{i+1}, m_{i+2}, \ldots$, and for which every pair of integers in the set have a common factor.

If $\frac{\phi(m_i)}{m_i} \geq 2/3$ then

$$\#\{n \leq x : (n, m_i) > 1\} \leq \sum_{p|m_i} \frac{x}{p} \leq x \sum_{p|m_i} \log\left(\frac{p}{p-1}\right) = x\log\left(\frac{m_i}{\phi(m_i)}\right) \leq x\log(3/2),$$

so that $f_i(x) \leq 2^{x\log(3/2)}$, which is part of the error term in (1.1).

Define $A_m(x)$ to be the number of even integers $\le x$, that are coprime to m, and $c_i = A_{m_i}(x)$. Clearly $f_i(x) \le 2^{i-1+[x/2]-c_i}$, and so, in order to complete the proof of Theorem 1, we must show that

$$c_i(x) - i - N \gg \frac{x}{\log^2 x \log\log x}, \tag{2.1}$$

for all $i \ge 2$ for which $\frac{\phi(m_i)}{m_i} \le 2/3$.

Suppose that $m = pr \le m' = p'r \le x$ are odd, squarefree numbers, with $p < x^{1/3}$. Noting that $|A_r(x)| = |A_m(x)| + |A_r(x/p)|$, we see that

$$|A_{m'}(x)| - |A_m(x)| = |A_r(x/p) \setminus A_r(x/p')| \gg \left(\frac{1}{p} - \frac{1}{p'}\right)\frac{x}{\log\log x} \gg \frac{x}{p^2 \log\log x}. \tag{2.2}$$

Thus, given any squarefree integer $r \le x$, $r \ne q$, we form a sequence $r = r_0, r_1, \dots r_j = q$, as follows: If $r_i = p_1 p_2 \dots p_h$ for some $h < k$, then then let $r_{i+1} = r_i p_{h+1}$. Otherwise we construct r_{i+1} by dividing r_i by its largest prime factor, and multiplying it by the smallest odd prime that does not yet divide it.

Now as $|A_q(x)| = N$, and as the prime factors of q are all $\ll \log x$, we find that, if $i \ge 2$ then $c_i - N \gg x/\log^2 x \log\log x$ by (2.2); which implies (2.1) for $i \ll x/\log^2 x \log\log x$.

So we are left with those i for which $\frac{\phi(m_i)}{m_i} \le 2/3$ and $i \gg x/\log^2 x \log\log x$. We first deal with those $i \le 100x/\log\log x$:

If $\mathcal{A}$ is *any* set of $< i$ odd integers $\le x$ then

$$\frac{\phi(m_i)}{m_i} \ge \min_{\substack{n \le x,\ n \text{ odd} \\ n \notin \mathcal{A}}} \frac{\phi(n)}{n}.$$

So we choose $\mathcal{A}$ to be the set of those $[x/\log^3 x]$ odd integers $\le x$, with the most distinct prime factors.

Hardy and Ramanujan [HR] showed that there exists an absolute constant c such that the number of integers $\le x$ with exactly k distinct prime factors is

$$\ll \frac{x}{\log x} \frac{(\log\log x + c)^{k-1}}{(k-1)!};$$

therefore the number of integers $\leq x$ with at least $10\log\log x$ distinct prime factors is $\ll x/\log^{10} x$. Thus, if $n \notin \mathcal{A}$ then n has $\leq 10\log\log x$ distinct prime factors, and so

$$\frac{\phi(n)}{n} \geq \prod_{p \leq (\log\log x)^2} \left(1 - \frac{1}{p}\right) \gg \frac{1}{\log\log\log x}$$

by Mertens' Theorem. Therefore, by the combinatorial sieve,

$$c_i \gg \frac{\phi(m_i)}{m_i} x \gg \frac{x}{\log\log\log x},$$

which implies (2.1) in this range of i.

Finally we come to those i in the range $100x/\log\log x \leq i \leq x/2$, with $\frac{\phi(m_i)}{m_i} \leq 2/3$:

An immediate consequence of Proposition 4 of [PS] is that

$$i \ < \ \frac{3}{20} x \frac{\phi(m_i)}{m_i} \Bigg/ \left(1 - \frac{\phi(m_i)}{m_i}\right) \ \leq \ \frac{9}{20} x \frac{\phi(m_i)}{m_i}.$$

On the other hand, by the combinatorial sieve,

$$c_i \ \geq \ \left\{\frac{1}{2} + o(1)\right\} x \frac{\phi(m_i)}{m_i} \ > \ \left\{\frac{10}{9} + o(1)\right\} i;$$

and thus

$$c_i - i \geq \frac{i}{10} \geq \frac{10x}{\log\log x},$$

and so (2.1) is satisfied.

3. Sets where any two elements are coprime.

Proof of Theorem 2: Let $t = \pi(\sqrt{x})$. For the upper bound, note that the number of composite elements of each set is $\leq t$, as these elements must all have distinct prime factors $\leq x$. All other elements are prime, and so the number of such sets is

$$\leq \ 2^{\pi(x)} \sum_{i=0}^{t} \binom{[x]}{i} \ \ll \ 2^{\pi(x)} \frac{x^t}{t!} \ \ll \ 2^{\pi(x)} \left(\frac{ex}{t}\right)^t,$$

which gives (1.3) by the Prime Number Theorem. (The proof here is the same as in Theorem 3.3 of [CE], except that they made a computational error in the final step.)

To obtain the lower bound, we shall construct (1.3) such sets. Let

$$k = t\left(1 - \frac{\log\log x}{\log x}\right)$$

and let

$$\sqrt{x} < q_1 < q_2 < \ldots < q_k$$

be the k smallest primes larger than $\sqrt{x}$. Note that, using the Prime Number Theorem in the form $\pi(x) = \frac{x}{\log x}(1 + O(1/\log x))$ we have

$$q_j = x^{1/2}(1 \ + j/t \ + \ O(1/\log x)) \tag{3.1}$$

and so $q_k < 2\sqrt{x}$.

We construct our sets as follows:

Each prime in the interval $(2\sqrt{x}, x]$ is in our set or not as desired, giving $2^{\pi(x)-\pi(2\sqrt{x})}$ different options.

We may put any number of the form $p_k q_k$ in the set, where p_k is a prime less than x/q_k (giving $\pi(x/q_k)$ choices).

Then any $p_{k-1}q_{k-1}$ where p_{k-1} is a prime $\leq x/q_{k-1}$ (giving $\pi(x/q_{k-1}\ ) - 1$ choices).

We continue in this fashion, taking, in general, any $p_{k-j}q_{k-j}$, where p_{k-j} is a prime $\leq x/q_{k-j}$, not already used as some p_{k-i}, (giving us $\pi(x/q_{k-j}) - j$ choices), for $j = 0,\ 1, \ldots, k-1$.

Thus the number of different sets constructed is

$$2^{\pi(x)+O(\sqrt{x}/\log x)} \prod_{i=0}^{k} \left\{ \pi\left(\frac{x}{q_i}\right) - (k-i) \right\}.$$

Now, by (3.1),

$$\begin{aligned} \pi\left(\frac{x}{q_i}\right) - (k-i) \ &= \ \pi\left(\frac{x^{1/2}}{1+i/t}\left(1 + O\left(\frac{1}{\log x}\right)\right)\right) - (k-i) \\ &= \ t\left\{\frac{1}{1+i/t} + O\left(\frac{1}{\log x}\right) - 1 + \frac{\log\log x}{\log x} + \frac{i}{t}\right\}, \end{aligned}$$

using the Prime Number Theorem again,

$$= t\left(\frac{\log\log x}{\log x} + \frac{(i/t)^2}{1+i/t} + O\left(\frac{1}{\log x}\right)\right) \ \geq\ \frac{t}{\log x}.$$

Therefore, the number of sets is at least

$$2^{\pi(x)+O(\sqrt{x}/\log x)}\left(\frac{t}{\log x}\right)^k,$$

which gives (1.3).

Remark: With some care it is possible to replace the $\log\log x$ in (1.3) with $(\log\log\log x)^2$, but we are currently unable to do better.

4. Sets whose sum of reciprocals is bounded.

Proof of Theorem 3: Let $y=[x^{1/4}],\ z=[x^{1/2}]$, and $x_j = jx/z$ for $y\leq j\leq z$. A given set of positive integers $\leq x$, whose sum of reciprocals is bounded by s, has, say, b_j integers in the interval $(x_j, x_{j+1}]$ for $y\leq j<z$ (with $0\leq b_j\leq x/z+1$), and so satisfies

$$\sum_{j=y}^{z-1}\frac{b_j}{x_{j+1}} \ \leq\ s. \tag{4.1}$$

So, if the b_j are fixed with these values then the number of sets of integers $\leq x$, with precisely b_j integers from the interval $(x_j, x_{j+1}]$, is

$$\leq\ 2^{xy/z}\prod_{y\leq j<z}\binom{[x/z]+1}{b_j}. \tag{4.2}$$

So define $\alpha_j := b_j/([x/z]+1)$ for each j. Clearly $0\leq\alpha_j\leq 1$ for each j, and, by (4.1), they must satisfy

$$\sum_{j=y}^{z-1}\frac{\alpha_j}{j} \ \leq\ s\left(1+O\left(\frac{1}{y}\right)\right). \tag{4.3}$$

Moreover, by Stirling's formula,

$$\binom{[x/z]+1}{b_j} \ll \exp\Big(O(\log x) \; - \; \frac{x}{z}\left(\alpha_j \log \alpha_j + (1-\alpha_j)\log(1-\alpha_j)\right)\Big).$$

Noting that there are no more than $(x/z+1)^z$ choices for the b_j, we thus see that

$$\nu(x) \ll \exp\left(O(x^{3/4}) \; - \; \frac{x}{z} \min_{\substack{0\le \alpha_j \le 1, y\le j<z \\ (4.3)\ \text{holds}}} \sum_{y\le j<z} \left(\alpha_j \log \alpha_j + (1-\alpha_j)\log(1-\alpha_j)\right) \right).$$

By the method of Lagrange multipliers we find that the minimum occurs when each $\alpha_j = 1/\left(1+e^{A/j}\right)$ for some constant $A>0$.

Now

$$\begin{aligned}\sum_{j=y}^{z-1} \frac{\alpha_j}{j} &= \int_y^z \frac{\mathrm{d}t}{t(1+e^{A/t})} \; + \; O\left(\frac{1}{ye^{A/y}} + \frac{1}{ze^{A/z}}\right) \\ &= \int_{A/z}^{\infty} \frac{\mathrm{d}u}{u(1+e^u)} \; + \; O\left(\frac{1}{ye^{A/y}} + \frac{1}{(A/y)e^{A/y}} + \frac{1}{ze^{A/z}}\right).\end{aligned}$$

To obtain equality in (4.3), we need to select $A = x^{1/2} f(s) + O(x^{1/4})$. Thus

$$\begin{aligned}-\sum_{y\le j<z} \alpha_j \log \alpha_j + (1-\alpha_j)\log(1-\alpha_j) &= \sum_{y\le j<z} \log\left(1+e^{-A/j}\right) \; + \; \frac{A}{j}\frac{1}{\left(1+e^{A/j}\right)} \\ &= \int_y^z \left(\left(1+e^{-A/u}\right) \; + \; \frac{A}{u\left(1+e^{A/u}\right)}\right)\mathrm{d}u \; + \; O(1).\end{aligned}$$

By the substitution $t = A/u$, and the fact that

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\frac{\log(1+e^{-t})}{t} \; = \; \frac{1}{t^2}\left(\log(1+e^{-t}) + \frac{t}{1+e^t}\right),$$

this last line is

$$z\log(1+e^{-A/z}) \; + \; O(1),$$

and so (1.4) is an upper bound for $\nu(x)$.

To obtain a lower bound for $\nu(x)$ note that if we select integers b_j such that

$$\sum_{j=y}^{z-1} \frac{b_j}{x_j} \ \leq \ s, \tag{4.4}$$

then the sum of the reciprocals of any set consisting of b_j integers from the interval (x_j, x_{j+1}), for each $y \leq j < z$, is $\leq s$. Clearly the number of such sets is

$$\geq \prod_{y \leq j < z} \binom{[x/z]}{b_j}. \tag{4.5}$$

Now the idea is to select each $b_j = x/z\left(1 + e^{A/j}\right) \ + \ O(1)$, for some constant $A > 0$, so that (4.4) is satisfied; this can be done by the choice $A = x^{1/2} f(s) + O(x^{1/4})$ (the proof being almost identical to that for the upper bound). Now, when we estimate (4.5) with this value for A, we proceed as in the upper bound, and show that (4.5) is at least (1.4).

Remarks: Cameron and Erdoš observed that any set of integers taken from $[x/e^s, x]$ has sum of reciprocals $\leq s$, and thus $c(s) \geq 2^{1-e^{-s}}$. We will derive (1.6) and (1.7):

If x is 'large' then

$$\int_x^\infty \frac{\mathrm{d}u}{u(1+e^u)} \ = \ \frac{1}{xe^x}\left\{1 + O\left(\frac{1}{x}\right)\right\},$$

which implies (1.6). On the other hand, (1.5) gives that, for $f(s) < 1$ (that is, s 'large'),

$$s = C_1 + \frac{1}{2}\log\left(1/f(s)\right) - \int_0^{f(s)} \left(\frac{1}{1+e^u} - \frac{1}{2}\right)\frac{\mathrm{d}u}{u}, \tag{4.6}$$

where

$$C_1 = \int_1^\infty \frac{\mathrm{d}u}{u(1+e^u)} \ + \ \int_0^1 \left(\frac{1}{1+e^u} - \frac{1}{2}\right)\frac{\mathrm{d}u}{u}.$$

Therefore $f(s) \asymp e^{-2s}$, and so the last term in (4.6) is $\ll f(s) \ll e^{-2s}$. Thus $f(s) = Ce^{-2s} + O(e^{-4s})$, where $C = e^{2C_1}$ (which can be computed explicitly), which implies (1.7).

References

[CE] CAMERON, P.J. & ERDÖS, P., On the number of sets of integers with various properties, Number Theory (ed. R.A. Mollin), (de Gruyter, New York, 1990), 61–79.

[HR] HARDY, G.H. & RAMANUJAN, S,, The normal number of prime factors of a number n, *Quarterly J. Math.*, 48 (1917), 76–92.

[PS] POMERANCE, C. & SELFRIDGE, J.L., Proof of D.J. Newman's coprime mapping conjecture, *Mathematika*, 27 (1980), 69–83.

Neil J. Calkin, School of Mathematics, Georgia Institute of Technology, Altanta, GA 30332, USA.

Andrew J. Granville, Department of Mathematics, University of Georgia, GA 30602, USA.