Dense product-free sets of integers

Carl Pomerance, Dartmouth College
Hanover, New Hampshire, USA

Joint Mathematics Meetings
Boston, January 6, 2012

Based on joint work with
P. Kurlberg, J. C. Lagarias, & A. Schinzel

Analytic number theory abounds with logs, loglogs, logloglogs, etc.

There are jokes about drowning analytic number theorists and other jokes about how Hungarian chickens cluck.

Though the theoreticians assure us that these logs are in truth there, can they really be detected numerically?

It is not so easy.

Here's an example. Take the $N \times N$ multiplication table. It has N^2 entries. It is a symmetric matrix, so most entries appear at least twice. How many distinct entries does it have?

Let $M(N)$ be the number of distinct entries in the $N \times N$ multiplication table.

$\times$	1	2	3	4	5
1	**1**	**2**	**3**	**4**	**5**
2	2	4	**6**	**8**	**10**
3	3	6	**9**	**12**	**15**
4	4	8	12	**16**	**20**
5	5	10	15	20	**25**

So, $M(5) = 14$.

×	1	2	3	4	5	6	7	8	9	10
1	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
2	2	4	6	8	10	**12**	**14**	**16**	**18**	**20**
3	3	6	9	12	**15**	18	**21**	**24**	**27**	**30**
4	4	8	12	16	20	24	**28**	**32**	**36**	**40**
5	5	10	15	20	**25**	30	**35**	40	**45**	**50**
6	6	12	18	24	30	36	**42**	**48**	**54**	**60**
7	7	14	21	28	35	42	**49**	**56**	**63**	**70**
8	8	16	24	32	40	48	56	**64**	**72**	**80**
9	9	18	27	36	45	54	63	72	**81**	**90**
10	10	20	30	40	50	60	70	80	90	**100**

So, $M(10) = 42$.

What would you conjecture about $M(N)$ asymptotically?

Here are some values of $M(N)/N^2$:

N	$M(N)/N^2$
5	0.5600
10	0.4200
20	0.3800
40	0.3231
80	0.3030
160	0.2802
320	0.2671
640	0.2538
1000	0.2481
2000	0.2399
8000	0.2267
16000	0.2215
32000	0.2166

(Calculations to 1000 by T. D. Noe as reported in the OEIS, to 32000 by P. Kurlberg.)

Do we have $M(N)$ of the shape N^{2-c_1}?

N	$M(N)/N^2$	c_1
5	0.5600	.3603
10	0.4200	.3768
20	0.3800	.3230
40	0.3231	.3063
80	0.3030	.2725
160	0.2802	.2507
320	0.2671	.2289
640	0.2538	.2122
1000	0.2481	.2018
2000	0.2399	.1878
8000	0.2267	.1651
16000	0.2215	.1557
32000	0.2166	.1475

How about $M(N)$ of the shape $N^2/(\log N)^{c_2}$?

N	$M(N)/N^2$	c_1	c_2
5	0.5600	.3603	1.2184
10	0.4200	.3768	1.0401
20	0.3800	.3230	.8819
40	0.3231	.3063	.8655
80	0.3030	.2725	.8081
160	0.2802	.2507	.7832
320	0.2671	.2289	.7533
640	0.2538	.2122	.7349
1000	0.2481	.2018	.7213
2000	0.2399	.1878	.7038
8000	0.2267	.1651	.6759
16000	0.2215	.1557	.6640
32000	0.2166	.1475	.6539

Or how about $M(N)$ of the shape $N^2/\exp((\log N)^{c_3})$?

N	$M(N)/N^2$	c_1	c_2	c_3
5	0.5600	.3603	1.2184	1.1453
10	0.4200	.3768	1.0401	.1704
20	0.3800	.3230	.8819	.0300
40	0.3231	.3063	.8655	.0935
80	0.3030	.2725	.8081	.1200
160	0.2802	.2507	.7832	.1482
320	0.2671	.2289	.7533	.1585
640	0.2538	.2122	.7349	.1692
1000	0.2481	.2018	.7213	.1718
2000	0.2399	.1878	.7038	.1755
8000	0.2267	.1651	.6759	.1798
16000	0.2215	.1557	.6640	.1808
32000	0.2166	.1475	.6539	.1817

Paul Erdős studied this problem in two papers, one in 1955, the other in 1960.

Paul Erdős, 1913–1996

In 1955, Erdős proved (in Hebrew) that $M(N)/N^2 \to 0$ as $N \to \infty$ and indicated that it was likely that $M(N)$ is of the basic shape $N^2/(\log N)^c$.

In 1960, at the prodding of Linnik and Vinogradov, Erdős identified (in Russian) the value of "c". Let

$$c = 1 - \frac{1 + \log\log 2}{\log 2} = 0.08607\ldots.$$

Then $M(N^2) = N^2/(\log N)^{c+o(1)}$ as $N \to \infty$.

In work of Tenenbaum progress was made (in French) in nailing down the "$o(1)$".

In 2008, Ford showed (in English) that $M(N)$ is of order of magnitude

$$\frac{N^2}{(\log N)^c (\log\log N)^{3/2}}.$$

No matter the language, we still don't know an asymptotic estimate for $M(N)$, despite this just being about the multiplication table!

So how can the fact that $M(N)$ is small compared to N^2 be explained?

It all comes down to the function $\Omega(n)$, the total number of prime factors of n, counted with multiplicity. For example,

$$\Omega(8) = 3,\ \Omega(9) = 2,\ \Omega(10) = 2,\ \Omega(11) = 1,\ \Omega(12) = 3.$$

In 1917, Hardy and Ramanujan proved that the normal order of $\Omega(n)$ is $\log\log n$. That is, for each $\epsilon > 0$, the set of integers n with

$$|\Omega(n) - \log\log n| < \epsilon \log\log n$$

has asymptotic density 1.

So, this explains the multiplication table. Most products $n_1 n_2$ have both $n_1 > N^{1/2}$ and $n_2 > N^{1/2}$, and most of these have $\Omega(n_1)$ and $\Omega(n_2)$ fairly close to $\log\log N$ (note that $\log\log(N^{1/2})$ differs from $\log\log N$ by less than 1). So most of the products formed have about $2\log\log N$ prime factors, which is an unusual value to have for a number below N^2.

G. H. Hardy

S. Ramanujan

So, $\log\log N$ for integers below N is the center of the distribution. To quantify $M(N)$ one needs to know about estimates for the tail, and that's where the constant c arises.

Product-free sets

How dense can a set of integers be if the set contains none of its products?

For example, take the integers that are 2 (mod 3). The product of any two of them is 1 (mod 3), so is not in the set. And this set has asymptotic density $\frac{1}{3}$.

The set of integers which are a power of 2 times a number that is 3 (mod 4) is product-free, and it has density $\frac{1}{2}$.

Can you do better?

Consider periodic sets, such as the 2 (mod 3) example.

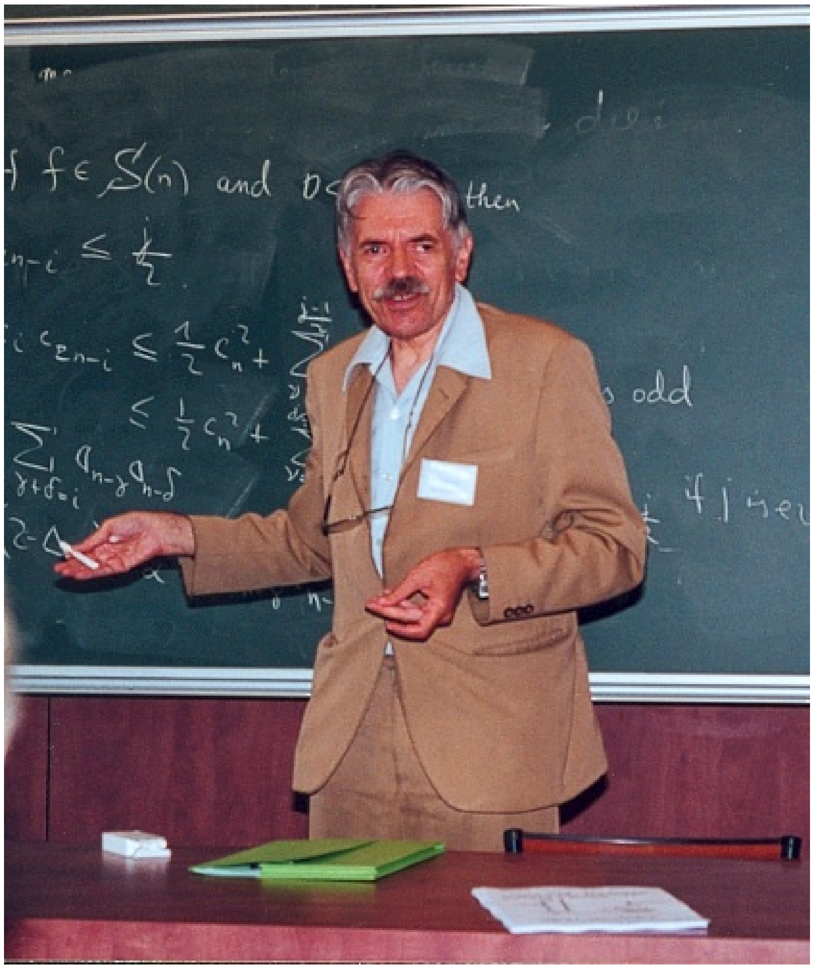
Let $D(n)$ denote the maximal possible density of a product-free set modulo n.

It is not hard to prove that $\liminf_{n\to\infty} D(n) = \frac{1}{2}$.

Do we have $D(n) < \frac{1}{2}$ for all n?

P, Schinzel (2011): *We have* $D(n) < \frac{1}{2}$ *for all* n *except possibly those* n *divisible by the square of a number with at least* 6 *distinct prime factors. Further, the asymptotic density of those* n *divisible by such a square is about* 1.56×10^{-8}. *And the least such number is about* 9×10^{8}.

Moscow Journal of Combinatorics and Number Theory,
1 (2011), 52–66.

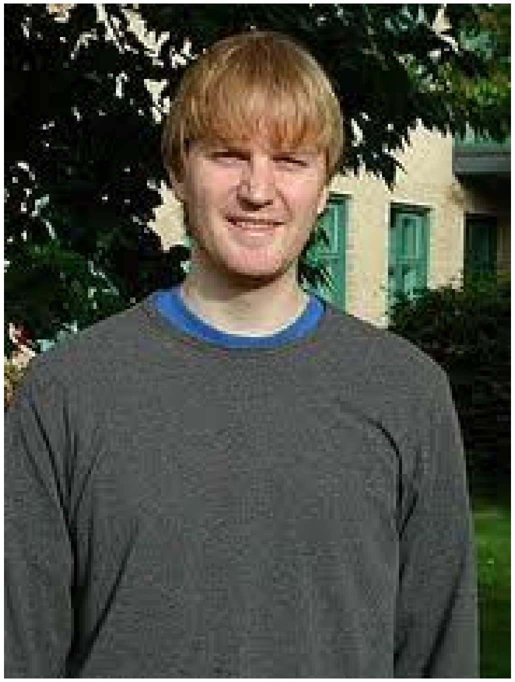


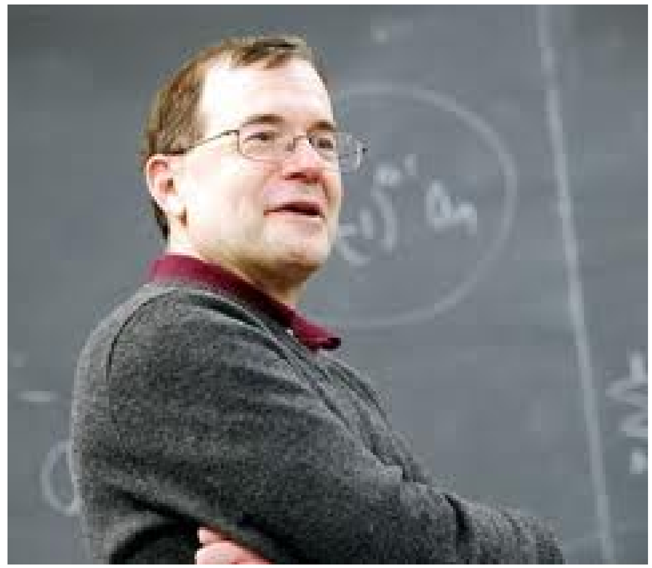
Andrzej Schinzel

However:

Kurlberg, Lagarias, P (2011): *There are infinitely many values of n with $D(n)$ arbitrarily close to* 1*. In particular, there are infinitely many values of n where all of the pairwise products of a subset of* 99% *of the residues* $\pmod n$ *all fall into the remaining* 1% *of the residue classes.*

Acta Arithmetica, to appear in a special issue in honor of Andrzej Schinzel's 75th birthday.





Pär Kurlberg

Jeffrey C. Lagarias

Let's be more modest, just show me one n where $D(n) \geq \frac{1}{2}$.

It's not so easy!

Here's a number. Take the first 10,000,000 primes. For those primes below 1,000,000, take their 14th powers, and for those that are larger, take their squares, and then multiply these powers together to form N. Then $D(N) > 0.5003$. Further, $N \approx 10^{1.61 \times 10^8}$.

Can you find an example with fewer than 100,000,000 decimal digits?

What is behind this construction and proof?

It is actually very similar to the proof of the Erdős multiplication table theorem.

Suppose n is a high power of the product of all of the primes up to x, say the exponent is $\lfloor \log x \rfloor$. Then consider all residues $r \pmod n$ with

$$\frac{2}{3} \log \log x < \Omega(\gcd(r, n)) < \frac{4}{3} \log \log x.$$

Then these residues $r \pmod n$ form a product-free set, and in fact most residues $\pmod n$ satisfy this inequality.

Actually the numbers $\frac{2}{3}$ and $\frac{4}{3}$ are not optimal, but $\frac{e}{4}$ and $\frac{e}{2}$ are. Being especially careful with the estimates leads to the following result:

Kurlberg, Lagarias, P (2011): *There is a positive constant c_1 such that for infinitely many n we have*

$$D(n) > 1 - \frac{c_1}{(\log\log n)^{1-\frac{e}{2}\log 2}(\log\log\log n)^{\frac{1}{2}}}.$$

Note that $1 - \frac{e}{2}\log 2 = 0.0579153\ldots$.

This is optimal for our method of proof, but is this the optimal result? It turns out that yes, apart from the constant c_1, it is optimal:

Kurlberg, Lagarias, P (2011): *There is a positive constant* c_2 *such that for all* n *we have*

$$D(n) < 1 - \frac{c_2}{(\log\log n)^{1-\frac{e}{2}\log 2}(\log\log\log n)^{\frac{1}{2}}}.$$

The idea for this upper bound: use linear programming!

Let me close with another computational problem.

Note that the set S of positive integers that are either 2 or 3 mod 5 is not only product-free, but it is also sum-free: no two members have their sum in the set. Further, S has asymptotic density $\frac{2}{5}$.

Find a numerical example of a product-free, sum-free set with asymptotic density strictly greater than $\frac{2}{5}$. We have proved that such sets exist, in fact with density arbitrarily close to $\frac{1}{2}$, but the least examples are likely to have so many decimal digits, that we would not be able to write down the number of these digits in decimal notation!

Thank You!