Sums and products: What we still don't know about addition and multiplication

Carl Pomerance, Dartmouth College
Hanover, New Hampshire, USA

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

You would think that all of the issues surrounding addition and multiplication were sewed up in third grade!

Well in this talk we'll learn about some things they didn't tell you . . .

Here's one thing they did tell you:

Find 483 × 784.

$$\begin{array}{r} 483 \\ \times\ 784 \\ \hline 1932 \\ 3864 \\ 3381 \\ \hline 378672 \end{array}$$

If instead you had a problem with two 23-digit numbers, well you always knew deep down that math teachers are cruel and sadistic. Just kidding! (*Aside: evil laugh* . . .)

In principle if you really have to, you could work out 23-digits times 23-digits on paper, provided the paper is big enough, but it's a lot of work.

So here's the real question: How much work?

Of course the amount of work depends not only on how long the numbers are, but on what they are. For example, multiplying 10^{22} by 10^{22}, that's 23-digits times 23-digits, but you can do it in your head.

In general, you'll take each digit of the lower number, and multiply it painstakingly into the top number. It's less work if some digit in the lower number is repeated, and there are definitely repeats, since there are only 10 possible digits. But even if it's no work at all, you still have to write it down, and that's 23 or 24 digits. At the minimum (assuming no zeroes), you have to write down $23^2 = 529$ digits for the "parallelogram" part of the product. And then comes the final addition, where all of those 529 digits need to be processed.

So in general if you multiply two n-digit numbers, it would seem that you'd be taking n^2 steps, unless there were a lot of zeroes. This ignores extra steps, like carrying and so on, but that at worst multiplies the n^2 by maybe 2 or 3. We say that the "complexity" of "school multiplication" for two n-digit numbers is of order n^2.

Here is what we don't know:
What is the *fastest* way to multiply?

There's a method known as the *Fast Fourier Transform* that allows you to multiply in about $n \log n$ steps. But we don't know if this is the best possible.

The function "$\log n$" can be thought of as natural log, or common log, or base-2 log, they are all within a constant factor of each other. The takeaway is that $\log n$ grows to infinity as n does, but eventually much more slowly than any root of n. For example, using the natural log, we have

$$\begin{aligned} \log n &< n^{1/2} && \text{for} && n \geq 1 \\ \log n &< n^{1/4} && \text{for} && n \geq 5504 \\ \log n &< n^{1/10} && \text{for} && n \geq 3.431 \times 10^{15} \\ \log n &< n^{1/100} && \text{for} && n \geq 1.286 \times 10^{281} \end{aligned}$$

So, "$n \log n$" is really just barely bigger than "n".

Let’s play **Jeopardy Multiplication!**

Here are the rules: I give you the answer to the multiplication problem, and you give me the problem phrased as a question. And you can’t use “1”.

So, if I say “15”, you say “What is 3×5?”

OK, let’s play.

21

Good. That was easy. Let’s up the ante.

91

Good. That was easy. Let's up the ante.

91

What is 7×13?

Let’s do 8051.

Let's do 8051.

(Thinking, thinking ... . Hmm,

$$8051 = 8100 - 49 = 90^2 - 7^2 = (90 - 7)(90 + 7) = 83 \times 97.$$

Got it!)

What is 83×97?

So, here's what we don't know:
How many steps does it take to come up with the answer, if you are given an n-digit number which *can be* factored?
(A trick problem would be: 17. The only way to write it as $a \times b$ is to use 1, and that was ruled out. So, prime numbers cannot be factored, and the thing we don't know is how long it takes to factor the non-primes.)

The best answer we have so far is about $10^{n^{1/3}}$ steps, and even this is not a theorem, but our algorithm (the *number field sieve*) seems to work in practice.

This is all crucially important for the security of Internet commerce. Or I should say that Internet commerce relies on the premise that we *cannot* factor much more quickly than that.

Here's something else, also related to multiplication.

Let's look at the multiplication table, but not necessarily up to 10×10, but more generally the $N \times N$ multiplication table.

It has N^2 entries. It is a symmetric matrix, so most entries appear at least twice. What we don't know:
How many different numbers appear in the table?

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

×	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**

$M(10) = 42.$

It may be too difficult to expect a neat exact formula for $M(N)$.

Instead, we could ask for its order of magnitude, or even approximate order of magnitude.

For example, does $M(N)$ go to infinity like a constant times N^2, or more slowly. That is, maybe

$$\lim_{N\to\infty} \frac{M(N)}{N^2} = c > 0?$$

Or maybe

$$\lim_{N\to\infty} \frac{M(N)}{N^2} = 0?$$

Here are some values of $M(N)/N^2$ (Brent & Kung 1981):

N	$M(N)$	$M(N)/N^2$
1	1	1.0000
3	6	0.6667
7	25	0.5102
15	89	0.3956
31	339	0.3528
63	1237	0.3117
127	4646	0.2881
255	17577	0.2703
511	67591	0.2588
1023	258767	0.2473
2047	1004347	0.2397
4095	3902356	0.2327
8191	15202049	0.2266

And some more values (Brent & Kung 1981, Brent 2012):

N	$M(N)$	$M(N)/N^2$
$2^{14}-1$	59410556	0.2213
$2^{15}-1$	232483839	0.2165
$2^{16}-1$	911689011	0.2123
$2^{17}-1$	3581049039	0.2084
$2^{18}-1$	14081089287	0.2049
$2^{19}-1$	55439171530	0.2017
$2^{20}-1$	218457593222	0.1987
$2^{21}-1$	861617935050	0.1959
$2^{22}-1$	3400917861267	0.1933
$2^{23}-1$	13433148229638	0.1909
$2^{24}-1$	53092686926154	0.1886
$2^{25}-1$	209962593513291	0.1865

And some statistically sampled values (Brent & P 2012):

N	$M(N)/N^2$	N	$M(N)/N^2$
2^{30}	0.1774	2^{100000}	0.0348
2^{40}	0.1644	2^{200000}	0.0312
2^{50}	0.1552	2^{500000}	0.0269
2^{100}	0.1311	$2^{1000000}$	0.0240
2^{200}	0.1119	$2^{2000000}$	0.0216
2^{500}	0.0919	$2^{5000000}$	0.0186
2^{1000}	0.0798	$2^{10000000}$	0.0171
2^{2000}	0.0697	$2^{20000000}$	0.0153
2^{5000}	0.0586	$2^{50000000}$	0.0133
2^{10000}	0.0517	$2^{100000000}$	0.0122
2^{20000}	0.0457	$2^{200000000}$	0.0115
2^{50000}	0.0390	$2^{500000000}$	0.0095

Richard P. Brent

OK, maybe you're convinced that $M(N)/N^2 \to 0$ as $N \to \infty$.

But can you prove it?

And if so, how fast does it tend 0?

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 shape $N^2/(\log N)^E$.

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

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

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

(Here, "$o(1)$" is a function that tends to 0 as $N \to \infty$.)

However, the formula $N^2/(\log N)^E$ doesn't look too good with our numbers. For example, at $N = 2^{5 \cdot 10^8}$, $1/(\log N)^E \approx .1841$, or close to 20 times higher than the experimental value .0095. While at $N = 2^{30}$ it is only 4 times higher.

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

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

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

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,

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

Some higher values: $\Omega(1024) = 10$, $\Omega(1009) = 1$, and $\Omega(2^{17} - 1) = 1,\ \Omega(2^{17}) = 17$.

But what is $\Omega(n)$ *usually*? That is, can $\Omega(n)$ be approximately predicted from the size of n if we throw out thin sets like primes and powers of 2?

Indeed it can.

In 1917, Hardy and Ramanujan proved that the normal order of $\Omega(n)$ is $\log\log n$. That is, for any given small number, say $\frac{1}{100}$, all but a vanishingly small fraction of numbers have

$$|\Omega(n) - \log\log n| < \frac{1}{100}\log\log n.$$

So, this explains the multiplication table. For $a, b \in [1, N]$, most products ab have both $a > \sqrt{N}$ and $b > \sqrt{N}$, and most of these have $\Omega(a)$ and $\Omega(b)$ fairly close to $\log\log N$ (note that $\log\log\sqrt{N}$ differs from $\log\log N$ by less than 1).
But $\Omega(ab) = \Omega(a) + \Omega(b)$.
So most of the products formed have about $2\log\log N$ prime factors, which is unusual 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.

I should take a small diversion from our progress here and mention one of the most beautiful theorems in number theory, the Erdős–Kac theorem. It says that the "standard deviation" for $\Omega(n)$ for integers up to N is $(\log\log N)^{1/2}$ and that the distribution is Gaussian. Namely, for each real number u, the set

$$\{n : \Omega(n) \le \log\log n + u(\log\log n)^{1/2}\}$$

has "asymptotic density" equal to $\dfrac{1}{\sqrt{2\pi}}\displaystyle\int_{-\infty}^{u} \mathrm{e}^{-t^2/2}\,\mathrm{d}t$.

This impressive looking function gives the area under the Bell curve up to u.

Einstein: “God does not play dice with the universe.”

Einstein: “God does not play dice with the universe.”

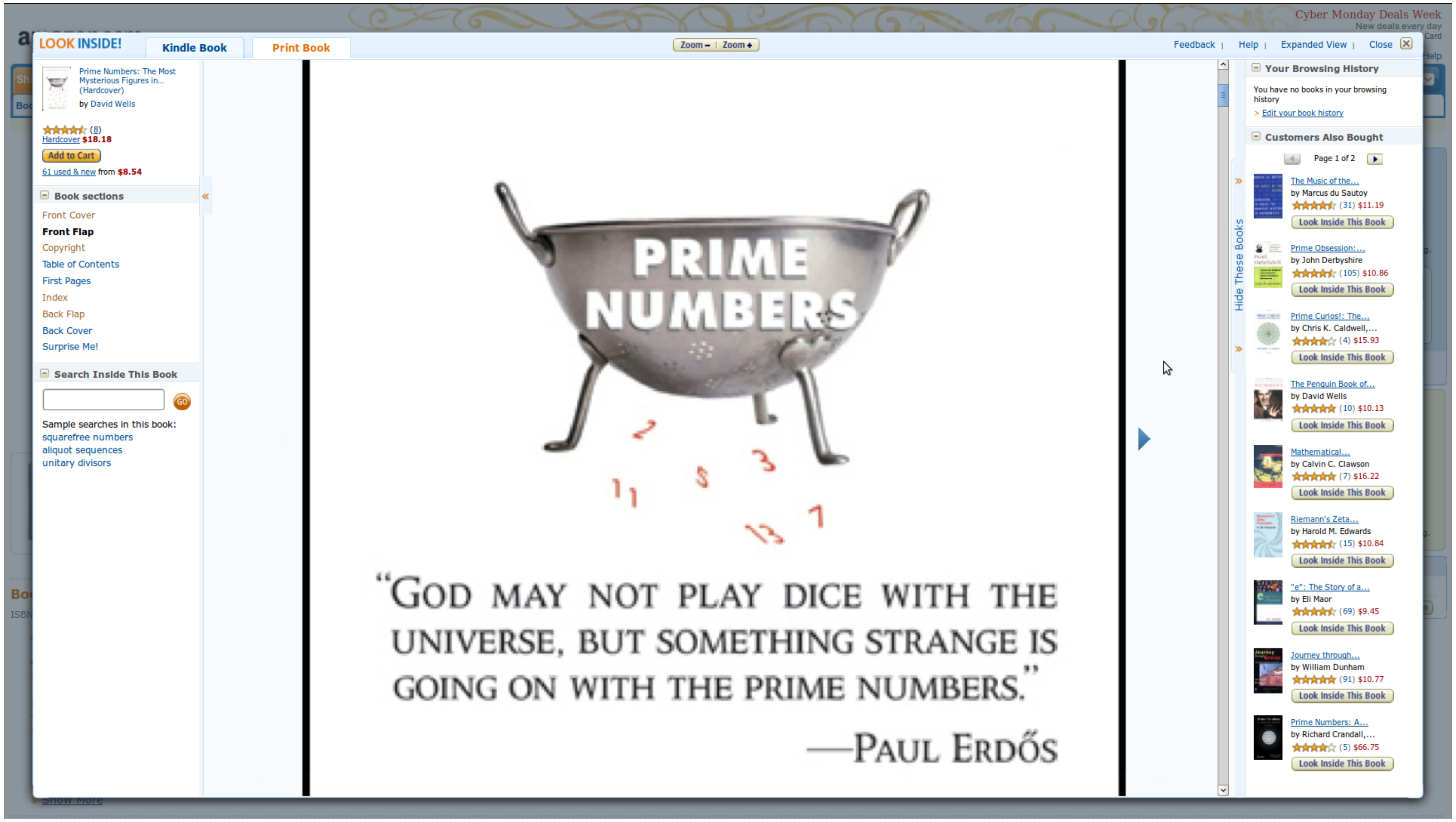
Erdős & Kac: Maybe so but something’s going on with the primes.

Einstein: “God does not play dice with the universe.”

Erdős & Kac: Maybe so but something’s going on with the primes.

(Note: I made this up, it was a joke . . .)

Prime numbers, the most mysterious figures in math, D. Wells



Prime Numbers: The Most Mysterious Figures in... (Hardcover)
by David Wells

(8)
Hardcover $18.18

61 used & new from $8.54

Sample searches in this book:
squarefree numbers
aliquot sequences
unitary divisors

"GOD MAY NOT PLAY DICE WITH THE UNIVERSE, BUT SOMETHING STRANGE IS GOING ON WITH THE PRIME NUMBERS."

—PAUL ERDŐS

You have no books in your browsing history
> Edit your book history

The Music of the...
by Marcus du Sautoy
(31) $11.19

Prime Obsession:...
by John Derbyshire
(105) $10.86

Prime Curios!: The...
by Chris K. Caldwell,...
(4) $15.93

The Penguin Book of...
by David Wells
(10) $10.13

Mathematical...
by Calvin C. Clawson
(7) $16.22

Riemann's Zeta...
by Harold M. Edwards
(15) $10.84

"e": The Story of a...
by Eli Maor
(69) $9.45

Journey through...
by William Dunham
(91) $10.77

Prime Numbers: A...
by Richard Crandall,...
(5) $66.75

Keeping with the theme of multiplication, what can be said about sets of positive integers that are *product-free*? This means that for any two members of the set, their product is not in the set. It is as far away as you can get from being closed under multiplication.

It is easy to find such sets, for example the set of primes. But how dense can such a set be?

For example, take the integers that are 2 (mod 3); that is, the numbers 2, 5, 8, 11, ... The product of any two of them is 1 (mod 3), so is not in the set. And this set has asymptotic density $\frac{1}{3}$.

Can you do better?

Well, the set of integers that are 2 or 3 (mod 5), mamely, 2, 3, 7, 8, 12, 13, ... is product-free and has density $\frac{2}{5}$.

The set of integers that are 3, 5, or 6 (mod 7) is product-free with density $\frac{3}{7}$.

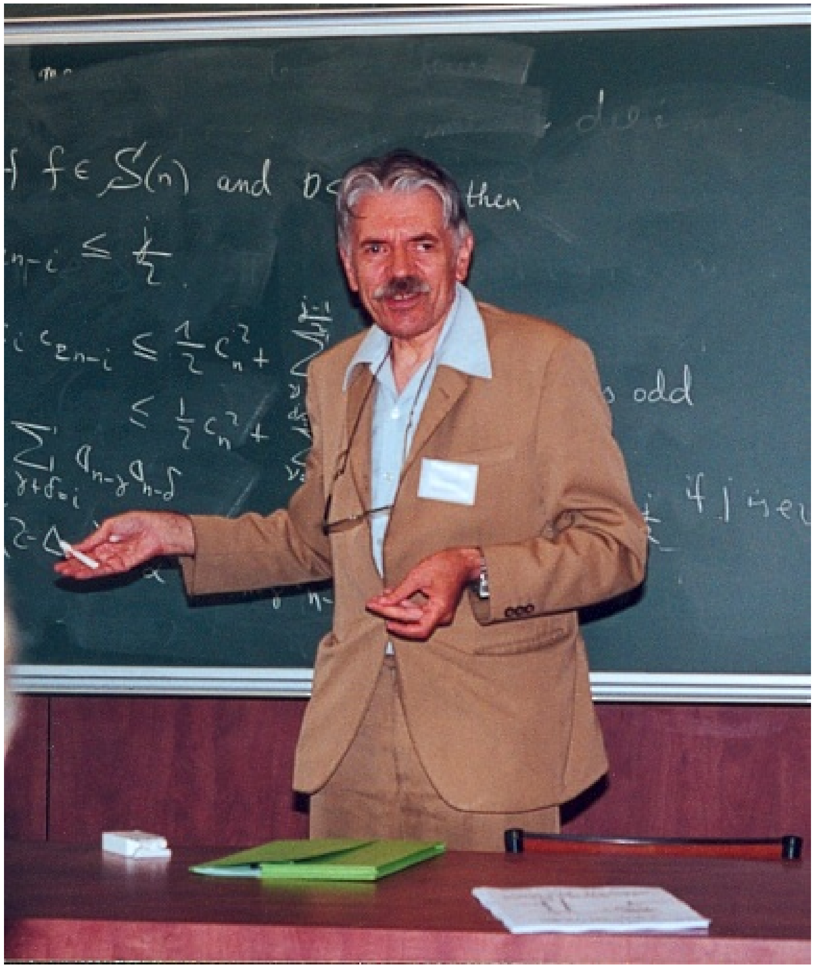
These sets are all periodic with some period n.

For all product-free sets that are periodic with period n, let $D(n)$ denote the maximal possible density. So, $D(5) = \frac{2}{5}$ and $D(7) = \frac{3}{7}$.

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}.

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



Andrzej Schinzel

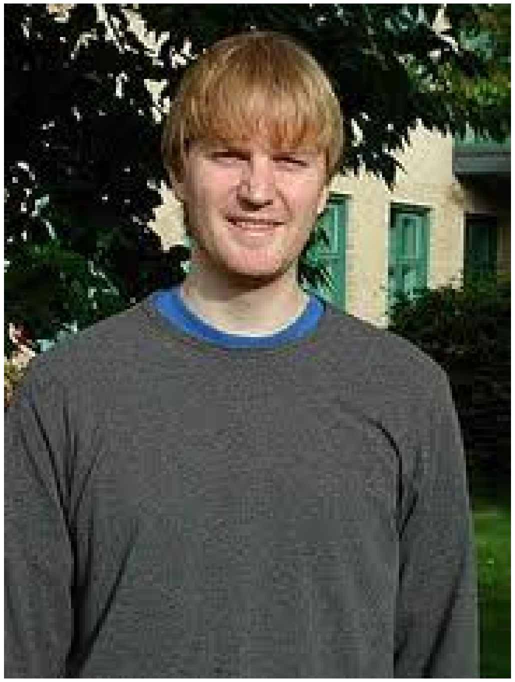
Surely that cements it, and $D(n) < \frac{1}{2}$ for all n, right?

Surely that cements it, and $D(n) < \frac{1}{2}$ for all n, right?

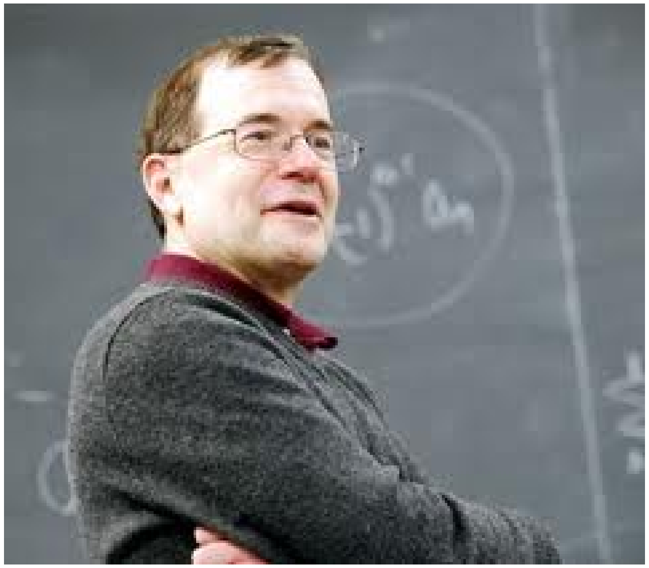
Well, no.

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 **155** (2012), 163–173.



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 power, and for those that are larger, take their square, 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 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

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

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 such that for infinitely many n we have*

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

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

We also showed that apart from “c” this is best possible.

International Mathematical Research Notices, to appear.

We have seen there is a lot we don't about multiplication, but what about addition?

Here's a famous problem due to Erdős & Szemerédi that involves both concepts, in fact, their interaction:

Among all sets $\mathcal{A}$ of N positive integers what is the minimum value of

$$|\mathcal{A}+\mathcal{A}| \ + \ |\mathcal{A}\cdot\mathcal{A}|?$$

Here $\mathcal{A}+\mathcal{A}$ is the set of all numbers $a+b$ where $a, b \in \mathcal{A}$, and $|\mathcal{A}+\mathcal{A}|$ is the number of elements in this set. Similarly for $|\mathcal{A}\cdot\mathcal{A}|$.

To repeat:
Among all sets $\mathcal{A}$ of N positive integers what is the minimum value of

$$|\mathcal{A}+\mathcal{A}| \ + \ |\mathcal{A}\cdot\mathcal{A}|?$$

We don't know. The conjecture is that this minimum value exceeds $N^{1.99}$ for all large N, and "1.99" can be replaced with any number smaller than "2".

Since Erdős & Szemerédi asked this in 1983, and got the result that there is some positive number σ such that

$$|\mathcal{A}+\mathcal{A}| \ + \ |\mathcal{A}\cdot\mathcal{A}| \ > \ N^{1+\sigma}$$

for all sets $\mathcal{A}$ of N integers and N sufficiently large, people have tried to do better with "σ".

The game players with the sum/product problem:
Erdős, Szemerédi, Nathanson, Chen, Elekes, Bourgain, Chang, Konyagin, Green, Tao, Solymosi, . . .

The best that they can do is $\sigma = \frac{1}{3}$.

Seeing a couple of Fields medalists, a Wolf Prize winner, an Abel Prize winner, four Salem Prize Winners, and two Crafoord Prize winners in this list, with the problem still not solved, is a bit daunting!

So far, all of the problems we've looked at have been fairly new, as far as Mathematics goes. Here's a very old problem that we still haven't solved and involves both sums and products, liberally interpreted.

A prime number, as we saw earlier, is a trick problem in Jeopardy Multiplication. It is a number larger than 1 that cannot be factored into two smaller (positive) whole numbers. Dating to correspondence in 1742 between Goldbach and Euler, it is conjectured that every even number starting at 4 can be represented as the sum of two primes.

271 years later: **We still don't know if Goldbach's conjecture is true.**

I'd like to close with one last problem that involves both sums and products, and it's a new one.

Suppose you have a set of residues mod n that is both sum-free and product-free.

For example, take the numbers that are 2 or 3 (mod 5). It is a set of asymptotic density $\frac{2}{5}$ and is both sum-free and product-free. It is easy to see that we cannot do better than $\frac{1}{2}$, and in another paper of Kurlberg, Lagarias, & P we showed that one can get arbitrarily close to $\frac{1}{2}$.

Here's the problem: **Give a numerical example that beats $\frac{2}{5}$. We know one exists, can one be explicitly described?**

Thank You!