Sums and products

Carl Pomerance, Dartmouth College
Hanover, New Hampshire, USA

International Number Theory Conference in Memory of Alf van der Poorten, AM
12–16 March, 2012
CARMA, the University of Newcastle

Based on joint work with
P. Kurlberg, J. C. Lagarias, & A. Schinzel
Let’s begin with products. 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.

<table>
<thead>
<tr>
<th>$\times$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>4</td>
<td>6</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>6</td>
<td>9</td>
<td>12</td>
<td>15</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>8</td>
<td>12</td>
<td>16</td>
<td>20</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>10</td>
<td>15</td>
<td>20</td>
<td>25</td>
</tr>
</tbody>
</table>

So, $M(5) = 14$. 
So, $M(10) = 42.$
What would you conjecture about $M(N)$ asymptotically?

Maybe

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

Maybe

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

Maybe yes and also

$$\lim_{N \to \infty} \frac{M(N)}{N^2/\log N} = c > 0?$$
Here are some values of $M(N)/N^2$:

<table>
<thead>
<tr>
<th>$N$</th>
<th>$M(N)/N^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0.5600</td>
</tr>
<tr>
<td>10</td>
<td>0.4200</td>
</tr>
<tr>
<td>20</td>
<td>0.3800</td>
</tr>
<tr>
<td>40</td>
<td>0.3231</td>
</tr>
<tr>
<td>80</td>
<td>0.3030</td>
</tr>
<tr>
<td>160</td>
<td>0.2802</td>
</tr>
<tr>
<td>320</td>
<td>0.2671</td>
</tr>
<tr>
<td>640</td>
<td>0.2538</td>
</tr>
<tr>
<td>1000</td>
<td>0.2481</td>
</tr>
<tr>
<td>2000</td>
<td>0.2399</td>
</tr>
<tr>
<td>8000</td>
<td>0.2267</td>
</tr>
<tr>
<td>16000</td>
<td>0.2215</td>
</tr>
<tr>
<td>32000</td>
<td>0.2166</td>
</tr>
</tbody>
</table>
(Calculations by T. D. Noe as reported in the OEIS and by P. Kurlberg.)
Do we have $M(N)$ of the shape $N^{2-c_1}$?

<table>
<thead>
<tr>
<th>$N$</th>
<th>$M(N)/N^2$</th>
<th>$c_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0.5600</td>
<td>0.3603</td>
</tr>
<tr>
<td>10</td>
<td>0.4200</td>
<td>0.3768</td>
</tr>
<tr>
<td>20</td>
<td>0.3800</td>
<td>0.3230</td>
</tr>
<tr>
<td>40</td>
<td>0.3231</td>
<td>0.3063</td>
</tr>
<tr>
<td>80</td>
<td>0.3030</td>
<td>0.2725</td>
</tr>
<tr>
<td>160</td>
<td>0.2802</td>
<td>0.2507</td>
</tr>
<tr>
<td>320</td>
<td>0.2671</td>
<td>0.2289</td>
</tr>
<tr>
<td>640</td>
<td>0.2538</td>
<td>0.2122</td>
</tr>
<tr>
<td>1000</td>
<td>0.2481</td>
<td>0.2018</td>
</tr>
<tr>
<td>2000</td>
<td>0.2399</td>
<td>0.1878</td>
</tr>
<tr>
<td>8000</td>
<td>0.2267</td>
<td>0.1651</td>
</tr>
<tr>
<td>16000</td>
<td>0.2215</td>
<td>0.1557</td>
</tr>
<tr>
<td>32000</td>
<td>0.2166</td>
<td>0.1475</td>
</tr>
</tbody>
</table>
How about $M(N)$ of the shape $N^2/(\log N)^{c_2}$?

<table>
<thead>
<tr>
<th>$N$</th>
<th>$M(N)/N^2$</th>
<th>$c_1$</th>
<th>$c_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0.5600</td>
<td>0.3603</td>
<td>1.2184</td>
</tr>
<tr>
<td>10</td>
<td>0.4200</td>
<td>0.3768</td>
<td>1.0401</td>
</tr>
<tr>
<td>20</td>
<td>0.3800</td>
<td>0.3230</td>
<td>0.8819</td>
</tr>
<tr>
<td>40</td>
<td>0.3231</td>
<td>0.3063</td>
<td>0.8655</td>
</tr>
<tr>
<td>80</td>
<td>0.3030</td>
<td>0.2725</td>
<td>0.8081</td>
</tr>
<tr>
<td>160</td>
<td>0.2802</td>
<td>0.2507</td>
<td>0.7832</td>
</tr>
<tr>
<td>320</td>
<td>0.2671</td>
<td>0.2289</td>
<td>0.7533</td>
</tr>
<tr>
<td>640</td>
<td>0.2538</td>
<td>0.2122</td>
<td>0.7349</td>
</tr>
<tr>
<td>1000</td>
<td>0.2481</td>
<td>0.2018</td>
<td>0.7213</td>
</tr>
<tr>
<td>2000</td>
<td>0.2399</td>
<td>0.1878</td>
<td>0.7038</td>
</tr>
<tr>
<td>8000</td>
<td>0.2267</td>
<td>0.1651</td>
<td>0.6759</td>
</tr>
<tr>
<td>16000</td>
<td>0.2215</td>
<td>0.1557</td>
<td>0.6640</td>
</tr>
<tr>
<td>32000</td>
<td>0.2166</td>
<td>0.1475</td>
<td>0.6539</td>
</tr>
</tbody>
</table>
Or how about $M(N)$ of the shape $N^2/\exp((\log N)^{c_3})$?

<table>
<thead>
<tr>
<th>$N$</th>
<th>$M(N)/N^2$</th>
<th>$c_1$</th>
<th>$c_2$</th>
<th>$c_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0.5600</td>
<td>.3603</td>
<td>1.2184</td>
<td>1.1453</td>
</tr>
<tr>
<td>10</td>
<td>0.4200</td>
<td>.3768</td>
<td>1.0401</td>
<td>.1704</td>
</tr>
<tr>
<td>20</td>
<td>0.3800</td>
<td>.3230</td>
<td>.8819</td>
<td>.0300</td>
</tr>
<tr>
<td>40</td>
<td>0.3231</td>
<td>.3063</td>
<td>.8655</td>
<td>.0935</td>
</tr>
<tr>
<td>80</td>
<td>0.3030</td>
<td>.2725</td>
<td>.8081</td>
<td>.1200</td>
</tr>
<tr>
<td>160</td>
<td>0.2802</td>
<td>.2507</td>
<td>.7832</td>
<td>.1482</td>
</tr>
<tr>
<td>320</td>
<td>0.2671</td>
<td>.2289</td>
<td>.7533</td>
<td>.1585</td>
</tr>
<tr>
<td>640</td>
<td>0.2538</td>
<td>.2122</td>
<td>.7349</td>
<td>.1692</td>
</tr>
<tr>
<td>1000</td>
<td>0.2481</td>
<td>.2018</td>
<td>.7213</td>
<td>.1718</td>
</tr>
<tr>
<td>2000</td>
<td>0.2399</td>
<td>.1878</td>
<td>.7038</td>
<td>.1755</td>
</tr>
<tr>
<td>8000</td>
<td>0.2267</td>
<td>.1651</td>
<td>.6759</td>
<td>.1798</td>
</tr>
<tr>
<td>16000</td>
<td>0.2215</td>
<td>.1557</td>
<td>.6640</td>
<td>.1808</td>
</tr>
<tr>
<td>32000</td>
<td>0.2166</td>
<td>.1475</td>
<td>.6539</td>
<td>.1817</td>
</tr>
</tbody>
</table>
Paul Erdős studied this problem in two papers, one in 1955, the other in 1960.
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 $N^2/(\log N)^c$.

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,

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

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 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. For $n_1, n_2 \in [1, N]$, 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.

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

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

has asymptotic density equal to $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-t^2/2} \, dt$. 

16
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

“God may not play dice with the universe, but something strange is going on with the prime numbers.”

—Paul Erdős
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 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). 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) 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 described as those integers in certain residue classes modulo some $n$. Let $D(n)$ denote the maximal possible density of a product-free set modulo $n$.

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

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 $1.56 \times 10^{-8}$.

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 (mod $n$) 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.
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

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

Then these residues $r \pmod{n}$ form a product-free set, and in fact most residues $r \pmod{n}$ 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!
For a product-free set $S$ in $\mathbb{Z}/n\mathbb{Z}$ and for $d \mid n$, let $\alpha_d$ be the proportion of those $s \in S$ with $\gcd(s, n) = d$ among all residues $r \pmod{n}$ with $\gcd(r, n) = d$.

Then each $\alpha_d$ is in $[0, 1]$.

Further, if $|S| \geq n/2$, then $\alpha_1 = 0$ and for all $u, v$ with $uv \mid n$, we have

$$\alpha_u + \alpha_v + \alpha_{uv} \leq 2.$$ 

In some sense, $|S|/n$ is closely modeled by $\sum_{d \mid n} \alpha_d / d$.

So, the LP is to maximize $\sum_{d \mid n} \alpha_d / d$ given the above constraints.
Since we already know that $D(n)$ can be fairly large, we need not prove we have found the maximum of the LP, just some upper bound for it. It is known that any feasible solution to the dual LP gives an upper bound for the primary LP. Thus, we write down the dual LP, find a fairly trivial feasible solution, and then “shift mass” to make it better.

And, voilà, our upper bound for all $n$’s tightly matches our constructed lower bound for champion $n$’s.
Sated now with products, let's move on to sums ... 

No, we're not going to start with addition tables. The analogous problem is trivial, in the addition table for the integers from 1 to $N$ there are precisely $2N - 1$ distinct sums.

But what about *sum-free* sets? Here we have a set of positive integers that contains none of the pairwise sums of its elements. How dense can such a set be?

This too is easy. The odd numbers form a sum-free set of asymptotic density $\frac{1}{2}$. And one cannot do better.
Here’s the proof. Say $\mathcal{A}$ is a sum-free set of positive integers and $a \in \mathcal{A}$. Then the set $a + \mathcal{A}$ is disjoint from $\mathcal{A}$. If $\mathcal{A}$ has $N = N(x)$ members in $[1, x]$, then $a + \mathcal{A}$ has $N + O(1)$ numbers here, so $x \geq 2N + O(1)$. Hence for all $x$, we have $N(x) \leq \frac{1}{2}x + O(1)$. We conclude that the upper density of a sum-free set $\mathcal{A}$ of positive integers is at most $\frac{1}{2}$.

Let us look at a somewhat more subtle problem. How dense can a sum-free subset of $\mathbb{Z}/n\mathbb{Z}$ be?

If $n$ is even, then take the odd residues, and this is best possible.

But what if $n$ is odd?
Diananda & Yap (1969), Green & Ruzsa (2005):
If $n$ is solely divisible by primes that are 1 (mod 3), then the maximal density of a sum-free set in $\mathbb{Z}/n\mathbb{Z}$ is $\frac{1}{3} - \frac{1}{3n}$. If $n$ is divisible by some prime that is 2 (mod 3), then the maximal density of a sum-free set in $\mathbb{Z}/n\mathbb{Z}$ is $\frac{1}{3} + \frac{1}{3p}$, where $p$ is the least such prime. Otherwise, the maximal density of a sum-free set in $\mathbb{Z}/n\mathbb{Z}$ is $\frac{1}{3}$.

This problem has been considered in general finite abelian groups and also for non-abelian groups. A survey article by recent Jeopardy contestant Kiran Kedlaya:

After hearing a shorter version of this talk at a conference in Georgia last October, several graduate students asked me the following question: What if you consider both sums and products?

Well, there is a famous and seminal problem here in which the Erdős multiplication-table theorem plays a role:

Among all sets \( A \) of \( N \) positive integers what is the minimum value of \( |A + A| + |A \cdot A| \)?

If one takes \( A = \{1, 2, \ldots, N\} \), then \( |A + A| = 2N - 1 \) and \( |A \cdot A| = N^2/(\log N)^c + o(1) \), so for large \( N \),

\[
|A + A| + |A \cdot A| > N^{2-\epsilon}.
\]
If on the other hand we take $A = \{1, 2, \ldots, 2^{N-1}\}$, then 
$|A \cdot A| = 2N - 1$ and $|A + A| = \frac{1}{2}N^2 + \frac{1}{2}N$, so that again 

$$|A + A| + |A \cdot A| > N^{2-\epsilon}. \quad (1)$$

**Erdős & Szemerédi** asked in 1983: Is (1) true for any set $A$ of $N$ positive integers?

There has been a parade of results getting better and better lower bounds, with game players being the posers **Erdős & Szemerédi**, then Nathanson, Chen, Elekes, Bourgain, Chang, Konyagin, Green, Tao, Solymosi, ...
Seeing a couple of Fields medalists in this list, with the problem still not solved, is a bit daunting!

But what the grad students asked was about dense sets $\mathcal{A}$ that are simultaneously 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. We cannot do better than $\frac{1}{2}$ for the density (considering only the sum-free property), but can we beat $\frac{2}{5}$ for both sum-free and product-free?
Kurlberg, Lagarias, P (2011): Say $A$ is sum-free and product-free with upper density $D(A)$.

1. If $A \subset \mathbb{Z}_{>0}$ with least element $a$, then $D(A) \leq \frac{1}{2} \left(1 - \frac{1}{5a}\right)$.

2. There is a constant $\kappa_1 > 0$, such that if $A \subset \mathbb{Z}/n\mathbb{Z}$, then
   \[ D(A) \leq \frac{1}{2} - \frac{\kappa_1}{(\log \log n)^{1 - \frac{e}{2} \log 2} (\log \log \log n)^{\frac{1}{2}}}. \]

3. There is a constant $\kappa_2$ and infinitely many $n$ such that for some $A \subset \mathbb{Z}/n\mathbb{Z}$,
   \[ D(A) \geq \frac{1}{2} - \frac{\kappa_2}{(\log \log n)^{1 - \frac{e}{2} \log 2} (\log \log \log n)^{\frac{1}{2}}}. \]
There remains a numerical problem: find an example of a number $n$ and a sum-free, product-free subset $A$ of $\mathbb{Z}/n\mathbb{Z}$ with $D(A) > \frac{2}{5}$.

For $D(A) = \frac{2}{5}$, we have $n = 5$. Back-of-the-envelope calculations suggest that there is some $n$ that beats $n = 5$ around

$$10^{10^{500,000}},$$

a number so large that we cannot even write the number of its digits in ordinary decimal notation.

But we haven’t looked at this problem too closely and there may be a much more modest example.
Thank You!