The multiplicative order mod n, on average

Carl Pomerance, Dartmouth College
Hanover, New Hampshire, USA

(Joint work with: Michel Balazard, Pär Kurlberg)

Perfect shuffles

Suppose you take a deck of 52 cards, cut it in half, and perfectly shuffle it (with the bottom card staying on the bottom).

If this is done 8 times, the deck returns to the order it was in before the first shuffle.

But, if you include the 2 jokers, so there are 54 cards, then it takes 52 shuffles, while a deck of 50 cards takes 21 shuffles.

What's going on?

For an odd number n, let $l(n) = l_2(n)$ denote the mutliplicative order of 2 in $(\mathbb{Z}/n\mathbb{Z})^\times$. Note that

$$l(51) = 8, \quad l(53) = 52, \quad l(49) = 21.$$

In fact, it is not hard to prove that the number of perfect shuffles to return a deck of $2n$ cards to it's initial order is $l(2n-1)$.

(Number the cards 0 to $2n-1$, with 0 the bottom card. Then a perfect shuffle takes a card in position i and sends it to $2i$ mod $2n-1$.)

This function $l(n)$ (= the multiplicative order of 2 mod n) appears to be very erratic and difficult to get hold of. It is of interest not only in card shuffling, but in computing the periods of certain pseudo-random number generators, and in other cryptographic contexts.

Further, as a basic and ubiquitous number-theoretic function it seems interesting to study $l(n)$, and more generally $l_a(n)$ (the order of a in $(\mathbb{Z}/n\mathbb{Z})^\times$) from a statistical viewpoint.

What is it normally?

What is it on average?

One elementary result that goes back to Gauss and Carmichael is that $l_a(n) \mid \lambda(n)$.

Here $\lambda(n)$ is the order of the largest cyclic subgroup in $(\mathbb{Z}/n\mathbb{Z})^\times$ and is defined by

$$\lambda([m,n]) = [\lambda(m), \lambda(n)], \qquad \lambda(p^\alpha) = \varphi(p^\alpha)$$

for odd primes p and $p^\alpha = 2$ or 4, and $\lambda(2^\alpha) = 2^{\alpha-2}$ for $\alpha \geq 3$.

For λ, we do have results about it's normal and average order, and they are a far cry from a possible first guess, the normal and average orders of φ.

Erdős, P, Schmutz: *On a set of asymptotic density 1,*

$$\lambda(n) = n/(\log n)^{\log\log\log n + A + o(1)}$$

for a certain explicit positive constant A*.*

Erdős, P, Schmutz: *As* $x \to \infty$*,*

$$\frac{1}{x}\sum_{n\le x}\lambda(n) = \frac{x}{\log x}\exp\left(\frac{(B+o(1))\log\log x}{\log\log\log x}\right)$$

for a certain explicit positive constant B*.*

Further, we know (assuming the Generalized Riemann Hypothesis for the Galois closures of Kummerian fields) that for most n coprime to a, we have $\lambda(n)/l_a(n)$ small. (Results of Li, Kurlberg, and Li & P.)

Thus, one has

$$l_a(n) = n/(\log n)^{\log\log\log n + A + o(1)}$$

for almost all n coprime to a.

Clearly the average order of $\lambda(n)$, which is of greater magnitude than $n/\log n$, is much larger than the normal order, so the average is determined by a thin set of numbers with abnormally large values of λ. Thus, it is unclear what is happening with the average order of $l_a(n)$.

After some numerical experiments, V. I. Arnold recently concluded that on average $l_a(n)$ is $C_a n/\log n$, and he gave a heuristic argument for this based on the physical principle of turbulence. This is in the paper

Number-theoretical turbulence in Fermat–Euler arithmetics and large Young diagrams geometry statistics, *Journal of Fluid Mechanics* **7** (2005), S4–S50.

It also was the subject of one the *Chern Lectures* he gave at UC Berkeley in 2007.

Arnold writes in the abstract:

"Many stochastic phenomena in deterministic mathematics had been discovered recently by the experimental way, imitating Kolmogorov's semi-empirical methods of discovery of the turbulence laws. From the deductive mathematics point of view most of these results are not theorems, being only descriptions of several millions of particular observations. However, I hope that they are even more important than the formal deductions from the formal axioms, providing new points of view on difficult problems where no other approaches are that efficient."

And he asserts that his expression $C_a n/\log n$ for the average order of $l_a(n)$ is in fact supported by *billions* of experiments.

I think we should be a bit suspicious!

First, iterated logarithms grow so slowly that they are difficult to detect numerically.

Second, Arnold did not seem to investigate any of the (admittedly scant) literature dealing with $l_a(n)$. In fact, there are interesting papers on the subject going back to Romanoff (who proved that the sum of $1/(nl_a(n))$ for n coprime to a is convergent), with later papers by Erdős, P, Pappalardi, Li, Kurlberg, Murty, Rosen, Silverman, Saidak, Moree, Luca, Shparlinski, and others.

But...

It's good to have outsiders investigate a field, and if they were expected to first read the literature thoroughly, it might dampen the fresh insight they might bring.

And, his conjecture that the average order of $l(n)$ grows like $n/\log n$ is supported on one side by Hooley's GRH-conditional proof of Artin's conjecture. Thus, assuming the GRH, a positive proportion of primes p have $l(p) = p - 1$, so that just the contribution of primes to the sum of $l(n)$ gives an average order that is $\gg n/\log n$. And perhaps composites do not contribute too much.

However...

Shparlinski (2007): *Let* $|a| > 1$. *Assuming the GRH, there is some* $C_a > 0$ *with*

$$\frac{1}{x} \sum_{\substack{n \le x \\ (a,n)=1}} l_a(n) \gg \frac{x}{\log x} \exp\left(C_a (\log\log\log x)^{3/2}\right).$$

(On some dynamical systems in finite fields and residue rings, *Discrete and continuous dynamical systems, Series A* **17** (2007), 901–917.)

And he suggests that with more work, the exponent "3/2" might possibly be replaced with "2".

Balazard, Kurlberg, P: *Let $|a| > 1$. Assuming the GRH,*

$$\frac{1}{x} \sum_{\substack{n \le x \\ (a,n)=1}} l_a(n) = \frac{x}{\log x} \exp\left(\frac{(B + o(1)) \log\log x}{\log\log\log x} \right).$$

Here "B" is the same constant that appears in the average order of $\lambda(n)$, namely

$$B = e^{-\gamma} \prod_p \left(1 - \frac{1}{(p-1)^2(p+1)} \right) = 0.3453720641\ldots .$$

In particular, the upper bound in the theorem holds unconditionally.

The proof is a bit intense, borrowing heavily from the structure of the proof in Erdős, P, & Schmutz of the corresponding result for $\lambda(n)$.

However, the following lemma is also used:

Kurlberg & P (2005): *For* $1 \le y \le \log x / \log\log x$

$$\{p \le x : l_a(p) < p/y\} \ll \frac{\pi(x)}{y}.$$

This result follows essentially from the the Hooley GRH conditional proof of Artin's primitive-root conjecture. (Pappalardi (1996) had this result in a wider range for y, but it has been retracted. Kurlberg (2003) had this result in the range $y \le (\log x)^{1-\varepsilon}$.)

Probably better suited for presentation in a talk is a proof of the following result from Balazard, Kurlberg, & P:

Assume the GRH. The average order of $l(p)$ *is* $\frac{159}{160}cp$*, where*

$$c = \prod_p \left(1 - \frac{p}{p^3 - 1}\right).$$

(Note that $\frac{159}{160}c = 0.57236022\ldots$, so that on average, $l(p) > \frac{4}{7}p$.)

Luca (2002) has shown that for averaging the orders of all elements of $(\mathbb{Z}/p\mathbb{Z})^\times$, this statistic has average order cp.

For an odd prime p, let $i(p) = (p-1)/l(p)$, namely the index of the subgroup $\langle 2 \rangle$ in $(\mathbb{Z}/p\mathbb{Z})^\times$. Let z be some parameter tending to infinity that we will specify later. We have

$$\sum_{p \le x} l(p) = \sum_{\substack{p \le x \\ i(p) \le z}} l(p) + \sum_{\substack{p \le x \\ i(p) > z}} l(p) = A + B,$$

say. Further, if $z \le \log x / \log\log x$, then the Kurlberg–P lemma (or alternately, the Brun–Titchmarsh inequality) implies

$$B \ll \frac{x}{z} \cdot \frac{\pi(x)}{z} \ll \frac{x\pi(x)}{z^2},$$

which is $o(x\pi(x))$ provided $z \to \infty$.

Now, for the main term A we might use known results for the distribution of primes p where $i(p)$ is fixed at some number, but it seems simpler to use an inclusion–exclusion:

$$A = \sum_{\substack{p \le x \\ i(p) \le z}} (p-1) \sum_{uv \mid i(p)} \frac{\mu(v)}{u}.$$

We write this as $C - D$, where in C we drop the condition $i(p) \le z$, but assume $uv \le z$, while in D we assume that $i(p) > z$ and $uv \le z$.

For D we majorize trivially (replace $\mu(v)$ with 1) and use the Kurlberg–P lemma. For $z < i(p) < z^2$, we get $x\pi(x)/z^{1-o(1)}$ and for $i(p) \ge z^2$, we get $x\pi(x)(\log\log x)/z$.

This leaves the main term

$$C = \sum_{p \le x} (p-1) \sum_{\substack{uv \mid i(p) \\ uv \le z}} \frac{\mu(v)}{u} = \sum_{uv \le z} \frac{\mu(v)}{u} \sum_{\substack{p \le x \\ uv \mid i(p)}} (p-1).$$

For a given value of $uv \le z$, we can compute the inner sum by partial summation and a GRH-conditional result in Hooley, getting

$$\frac{x\pi(x)}{2uv\varphi(uv)} + O\left(\frac{x^2}{\log^2 x}\right)$$

if $8 \nmid uv$ and twice this if $8 \mid uv$. (If $z \le (\log x)^{1/7}$, say, we have this unconditionally, but we needed the GRH to estimate the error term D.)

Thus, the main term is

$$C = \frac{1}{2} x\pi(x) \left(\sum_{uv \le z} \frac{\mu(v)}{u^2 v \varphi(uv)} + \sum_{\substack{uv \le z \\ 8 | uv}} \frac{\mu(v)}{u^2 v \varphi(uv)} \right) + O\left(\frac{x^2 z \log z}{\log^2 x} \right).$$

So we see that a convenient choice of z is say $(\log x)^{1/2}$. For the main term, we consider the expressions as infinite sums, estimate the errors in truncation, and then rewrite the infinite sums as Euler products. We get:

$$\frac{1}{\pi(x)} \sum_{2 < p \le x} l(p) = \frac{159}{320} cx + O\left(\frac{x}{(\log x)^{1/2-\varepsilon}} \right).$$

□