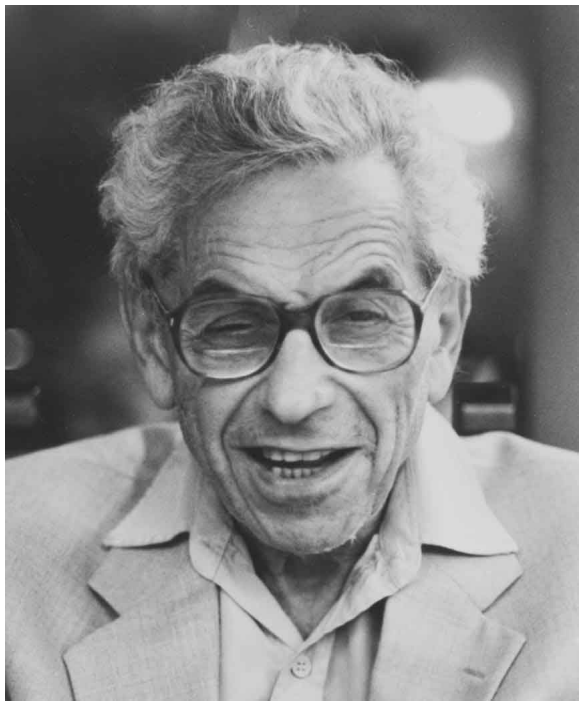
Erdős, van der Corput, and the birth of covering congruences

Carl Pomerance, Dartmouth College



Paul Erdős
1913–1996

Johannes van der Corput
1890–1975

Our story begins with Alphonse de Polignac, 1817–1890.

He is known principally for two conjectures, made in 1849.

(1) *Every positive even number is the difference of two consecutive primes in infinitely many ways.*

(2) *Every odd number is of the form* $2^n + p$, *where* $p = 1$ *or* p *is prime.*

Actually these were announced as theorems. And not only theorems, but the second one was claimed to have been verified up to 3,000,000.

On (1), even leaving out "consecutive" and "infinitely many ways" we still don't know if every even number is the difference of two primes. But probably de Polignac was correct, his conjecture is widely believed, but still unproved. We do know that most even numbers are the difference of two primes at least once, but the only number proved to be the difference of two primes infinitely often is 0.

But conjecture (2) is clearly nonsense. The number 1 is not of the form $2^n + p$. OK, say we start at 3:

$3 = 2^0+2,\ \ 5 = 2^1+3,\ \ 7 = 2^1+5,\ \ 9 = 2^1+7,\ \ 11 = 2^2+7, \ldots.$

But try 127:

$127 - 2^0 = 126,\ \ 127 - 2^1 = 125,\ \ 127 - 2^2 = 123,$
$127 - 2^3 = 119,\ \ 127 - 2^4 = 111,\ \ 127 - 2^5 = 95,\ \ 127 - 2^6 = 63.$

As is not uncommon, de Polignac had been preceded by Euler. And Euler had known about 127 being a counterexample. Since 127 is not only prime, but a Mersenne prime, Euler wondered if that were a clue. But he found the composite number 959 as a counterexample as well, and concluded that perhaps there was not much to be discovered here.

Leonhard Euler

So why might we consider this an interesting problem?

Lets try counting. Let $A(x)$ denote the number of pairs 2^n, p, where $n \geq 1$, p is 1 or an odd prime, and $2^n + p \leq x$. That is, we have a function that sends a pair n, p to an odd integer in $[1, x]$. We know there are (about) $\frac{1}{2}x$ odd numbers to x. What is the size of the domain of this function?

For each odd prime $p < x$, we can count the number of powers of 2 up to $x - p$. This is $\lfloor \log_2(x-p) \rfloor$, where the subscript 2 indicates a base-2 log. And

$$A(x) = \sum_{p<x} \lfloor \log_2(x-p) \rfloor \sim \frac{1}{\log 2} x.$$

Now $\frac{1}{\log 2} = 1.442\cdots > \frac{1}{2}$, so maybe we should conjecture that almost all odds in $[1, x]$ are covered?

Or maybe we should conjecture that the number of odds in $[1, x]$ not of the form $2^n + p$ is $\sim cx$ where

$$c = \frac{1}{2} e^{-2/\log 2} = 0.0279\ldots .$$

(This would be so for a *random* map from a set of size $\frac{1}{\log 2} x$ to a set of size $\frac{1}{2} x$.)

Can we at least prove that a positive proportion of odd numbers are indeed of the form $2^n + p$?

Romanoff (1934). *Yes we can.*

He did this using Cauchy's inequality and a sieve result counting quadruples n, n', p, p' with $2^n + p = 2^{n'} + p' \leq x$.

So, many numbers *can* be represented as $2^n + p$, what about odd numbers like 127 and 959 that cannot be so represented?

Enter Paul Erdős and Johannes van der Corput. In 1950 they published (independent) proofs that in fact a positive proportion of odds are *not* of the form $2^n + p$.

J. van der Corput, Over het vermoeden van de Polignac, *Simon Stevin*, **27** (1950), 99–105.

P. Erdős, On integers of the form $2^k + p$ and some related problems, *Summa Brasil. Math.* **2** (1950), 113–123.

Erdős. *There is an arithmetic progression of odd numbers, no term of which is of the form* $2^n + p$.

Proof. Every integer satisfies at least one of

$$0 \pmod 2, \quad 0 \pmod 3, \quad 1 \pmod 4,$$
$$3 \pmod 8, \quad 7 \pmod{12}, \quad 23 \pmod{24}.$$

So, if m is in the residue class determined by

$$1 \pmod 3, \quad 1 \pmod 7, \quad 2 \pmod 5,$$
$$2^3 \pmod{17}, \quad 2^7 \pmod{13}, \quad 2^{23} \pmod{241},$$

then for any n, we have that $m - 2^n$ is divisible by one of the primes 3, 5, 7, 13, 17, 241. □

It is difficult to make out van der Corput's proof, at least for me, it being largely expository, floridly written, and in Dutch. There also appears to be a mistake in the argument, which is only carefully worked out in a base case, and generalized with a flourish. It is possible to generalize the argument correctly, and it can be summed up as using a combination of the two ideas we've looked at: the pigeon-hole principle and a covering.

We saw the pigeon-hole failed because the number of pairs n, p that map to odd numbers $2^n + p \leq x$, namely $\frac{1}{\log 2} x$, is larger than $\frac{1}{2} x$. Van der Corput notes that if $2^n + p \equiv 1 \pmod 3$, then either n is even and $p = 3$ (a negligible case) or n is odd and $p \equiv 2 \pmod 3$. Thus we've cut down the domain by a factor 4, and the range by a factor 3.

This is not enough, but it's progress.

Say we also ask for $2^n + p \equiv 2 \pmod 5$. We are already restricting to n odd, so $n \equiv 1$ or $3 \pmod 4$. If $n \equiv 1 \pmod 4$, then $p = 5$, a negligible case, so that we can reduce to $n \equiv 3 \pmod 4$ and $p \equiv -1 \pmod 5$. We've cut down the domain by a further factor of $2 \times 4 = 8$, and the range by a factor 5.

Is this enough?

We're now looking at pairs n, p where $n \equiv 3 \pmod 4$ and $p \equiv -1 \pmod{15}$, and these map to odd integers in the residue class 7 (mod 15).

No, $\frac{1}{\log 2} \cdot \frac{1}{4} \cdot \frac{1}{8} = .045\ldots \ > \ \frac{1}{2} \cdot \frac{1}{15} = .033\ldots$.

Lets also bring in the prime 17, and say we try and hit odd numbers of the form $2^n + p \equiv 127 \pmod{255}$.

We have seen that but for some negligible cases, we must have $n \equiv 3 \pmod 4$ and $p \equiv -1 \pmod{15}$. But $127 \equiv 8 \pmod{17}$, so if $n \equiv 3 \pmod 8$, then $p = 17$, which is neglible. We may thus assume $n \equiv 7 \pmod 8$ and $p \equiv -1 \pmod{17}$.

Our domain now consists of pairs n, p with $n \equiv 7 \pmod 8$ and $p \equiv -1 \pmod{255}$, while the range consists of odd numbers that are 127 (mod 255).

And $\frac{1}{\log 2} \cdot \frac{1}{8} \cdot \frac{1}{\varphi(255)} = .001408\ldots \;<\; \frac{1}{2} \cdot \frac{1}{255} = .001960\ldots$.

So this does it, via a partial covering and the pigeon-hole principle:

Van der Corput: *There is a positive proportion of odd numbers that are not of the form* $2^n + p$.

We now ask about this proportion. Van der Corput tells us that it is at least

$$\frac{1}{2} \cdot \frac{1}{255} - \frac{1}{\log 2} \cdot \frac{1}{8} \frac{1}{\varphi(255)} \approx 5.52 \times 10^{-4}.$$

While Erdős tells us it is at least

$$\frac{1}{2} \cdot \frac{1}{3} \cdot \frac{1}{5} \cdot \frac{1}{7} \cdot \frac{1}{13} \cdot \frac{1}{17} \cdot \frac{1}{241} \approx 8.94 \times 10^{-8}.$$

Both arguments can be improved to achieve a larger density.

First, in the Erdős argument, using the same primes, there are several ways of choosing the residue classes. It is the exponents n that are covered, and this can be done with 2 choices mod 2, 2 further choices mod 4, 2 further choices mod 8, 3 choices mod 3, 2 choices mod 12, and 1 choice mod 24. That is, 8.94×10^{-8} is to be multiplied by 48, giving a density of 4.29×10^{-6}.

There are also other coverings that might be used, but it seems the van der Corput argument is the way to go.

With van der Corput, there are also multiple ways to achieve the same ends, and with modulus 255, there are 8 ways, so that density is increased to 4.41×10^{-3}.

By bringing in the primes 7 and 13 that Erdős used as part of his argument, we can increase the density to 7.47×10^{-3}.

Then bringing in 11, 31, and 41 (so that we start to consider the exponent n (mod 5)), we increase further to 8.49×10^{-3}.

And bringing in 19, 37, and 73, we increase further to 8.84×10^{-3}.

In 2006, **L. Habsieger & X.-F. Roblot** rediscovered van der Corput's method, and with the help of a large computer, improved the above estimate of 8.84×10^{-3}, by about 2.4% to 9.05×10^{-3}.

Can one achieve density 10^{-2}?

And what can we say about the density of numbers that can be represented?

Recall that Romanoff showed that a positive proportion of numbers are indeed of the form $2^n + p$.

Romanoff did not work out a numerical value for the density, but in the last decade, several papers addressed this issue:

Y. G. Chen & X.-G. Sun (2004): .0868

L. Habsieger & X.-F. Roblot (2006): .0933

J. Pintz (2006): .09368

G. Lü (2007): .09322

We conclude that the density of odd numbers represented in the form $2^n + p$ is between .09368 and .49095.

Actually we don't even know the density exists, so these are bounds for the lower density and the upper density, respectively.

So, what is the truth? A small numerical experiment indicates that perhaps the density is about .455.

It is interesting to reflect that even a totally wrong conjecture (of de Polignac) can spur interesting work, and in this case, it gave birth to the entire field of covering congruences.

THANK YOU