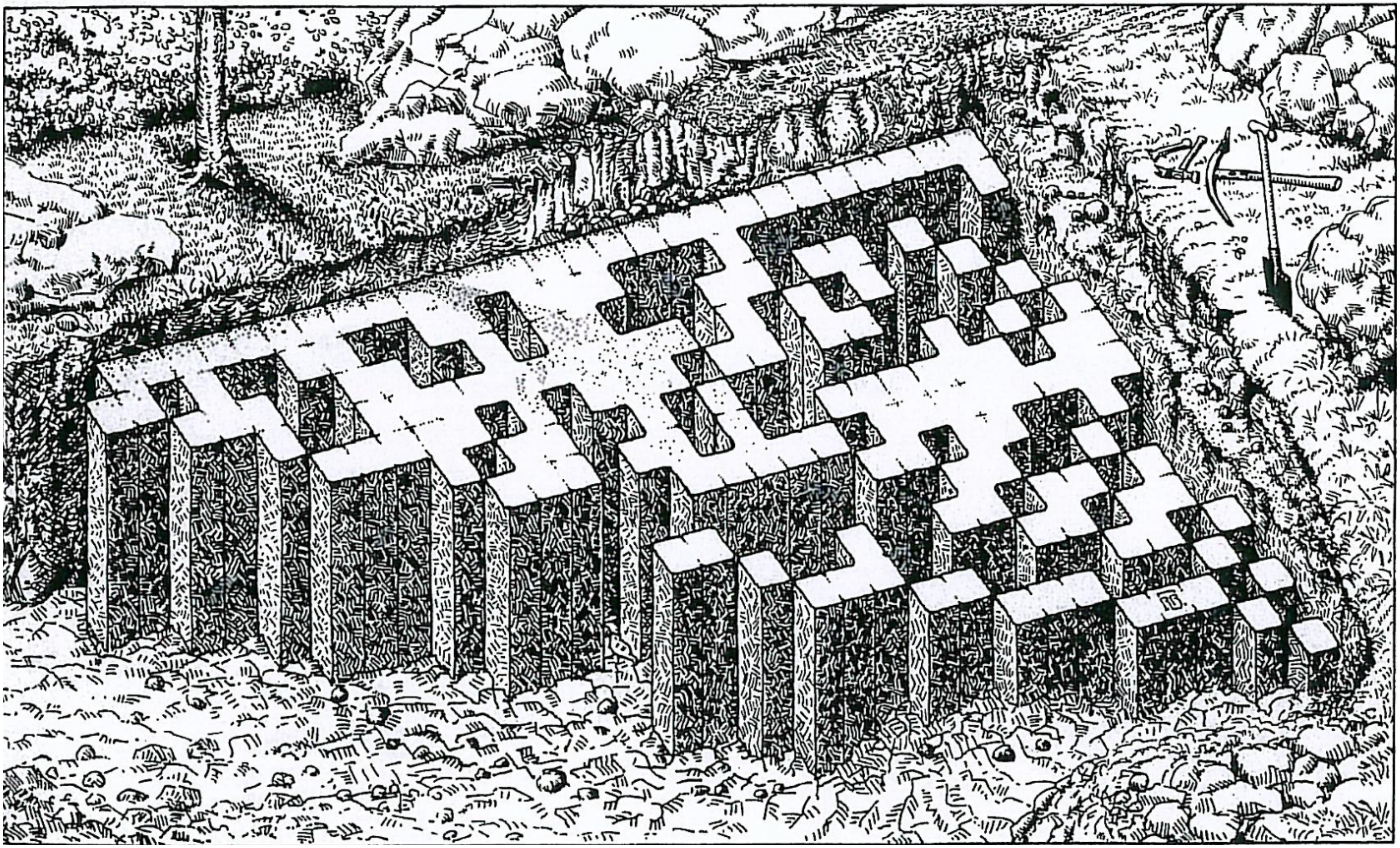
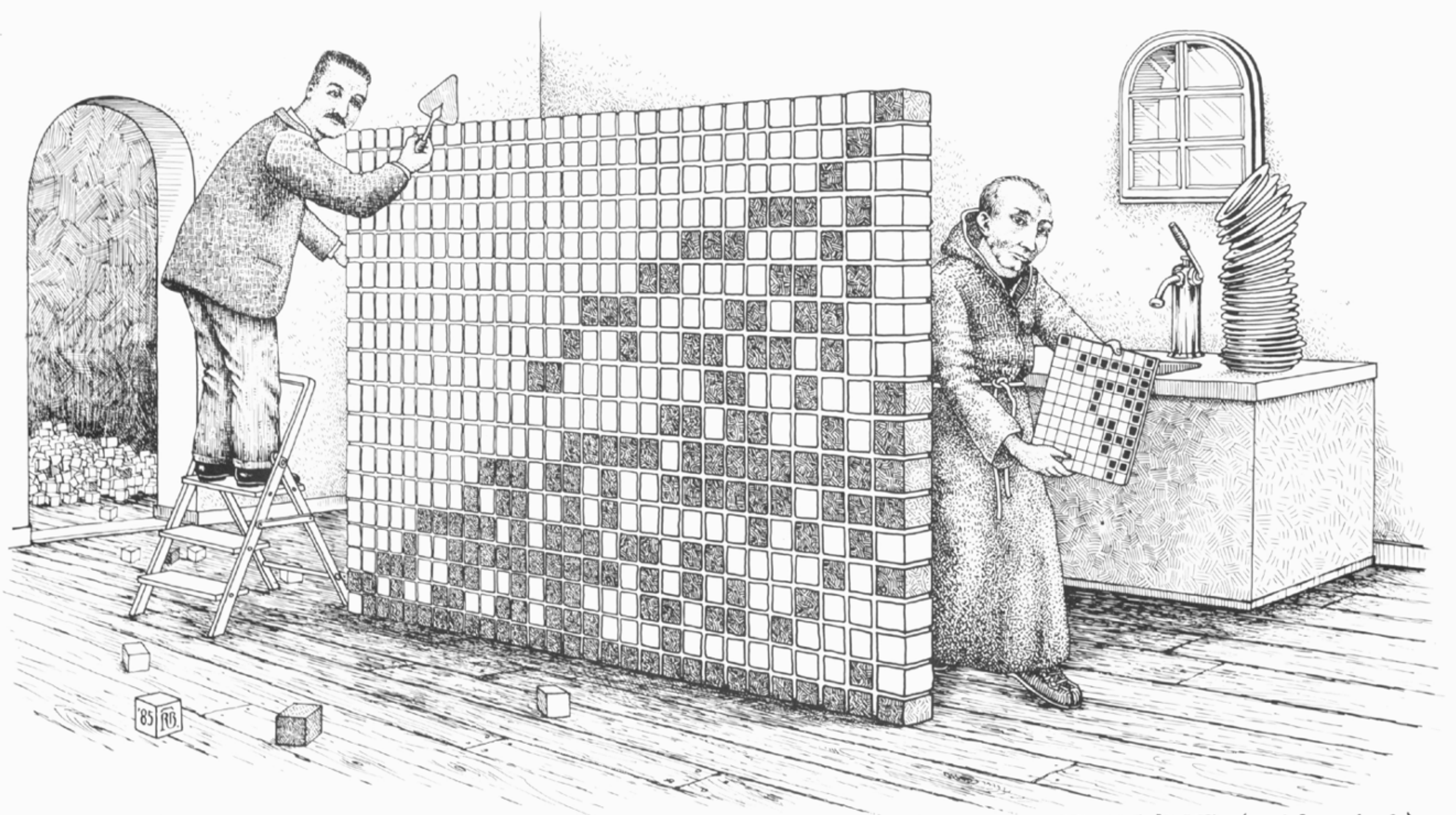
# Cyclotomic primes

**Carl Pomerance**, **Dartmouth College**

November 19, 2024





Édouard LUCAS (1842–1891) Marin MERSENNE (1588–1648)

(Graphic art by Tobias Baanders, based on some needlepoint of Willemien Lenstra and a concept of Hendrik Lenstra)

These are in fact the sequence of exponents $n$, written in binary, for which $2^n - 1$ is prime.

A prime of the form $2^n - 1$ must have $n$ itself prime. They go back to Pythagoras and Euclid, and are currently known as Mersenne primes.

They have certainly grabbed the public's imagination!

TIME Magazine's 29-th greatest invention of 2008.


The exponent in binary: 10100100011101100010100001.

(While perhaps the 29-th greatest invention, it is the 47-th prime of the form $2^p - 1$.)

The current largest known (Mersenne) prime is

$$2^{136279841} - 1.$$

The exponent in binary is 1000000111110111011100100001.

It was unearthed last month (after a hiatus of 6 years) and is the 52-nd known Mersenne prime.

The current search: One runs through candidate exponents $p$. If $2^p - 1$ survives a search for possible small prime factors $q \equiv 1 \pmod p$ with $(2/q) = 1$, it is checked if

$$3^{2^{p-1}} \equiv -3 \pmod{2^p - 1}.$$

If it is not, the calculation is checked (in much less time than it took to compute the first time). If the congruence holds, the exponent $p$ is then subjected to the Lucas–Lehmer test: Starting with $x = 4$, iterate $x^2 - 2 \pmod{2^p - 1}$ $p - 2$ times. This is 0 if and only if $2^p - 1$ is prime.

The Lucas–Lehmer test "lives" in the quadratic field $\mathbb{Q}[\sqrt{-3}]$. Similar primality tests work for $n$ when $n + 1$ has a fully known (or mostly known) prime factorization. Though in general we have the polynomial time primality test of Agrawal, Kayal, and Saxena (with improvements by Lenstra and P), it is not competitive with Lucas–Lehmer when the latter is appropriate.

Here is a heuristic that there are infinitely many Mersenne primes: The number $2^p - 1$ has least prime factor $> p$. A random number $n$ with all prime factors $> \log n$ is prime with probability $\sim e^\gamma \log\log n / \log n$, where $\gamma$ is Euler's constant. Applying this with $n = 2^p - 1$, the likelihood it is prime is $(e^\gamma / \log 2)(\log p)/p$. It remains to note that the series $\sum (\log p)/p$ diverges. The numerical evidence seems to support this reasoning, even with the special constant $e^\gamma / \log 2$.

We also have the Fermat numbers $F_n = 2^{2^n} + 1$. (An odd prime of the form $2^k + 1$ must have $k$ a power of 2.) Fermat thought all of these numbers are prime, and he was right for $n = 0, 1, 2, 3, 4$. Consider the case $n = 5$. If $p \mid F_5$, then $2^{2^5} \equiv -1 \pmod{p}$, so that $p \equiv 1 \mod 2^6$. Then $(2/p) = 1$, so that $2^{(p-1)/2} \equiv 1 \pmod{p}$, and hence $p \equiv 1 \pmod{2^7}$. The candidates for $p$ are

$$129,\ 257,\ 385,\ 513,\ 641, \ldots .$$

But 129, 385, and 513 are obviously not prime, and 257, which is $F_3$ cannot divide $F_5$ since the Fermat numbers are easily seen to be pairwise coprime. So, the very first candidate is 641, and in fact, as Euler showed, it is indeed a proper factor of $F_5$.

The Fermat numbers have been tested up to $n = 32$, and all of them after $n = 4$ are composite. Many of these were found composite after a nontrivial prime factor was found. We also have Pepin's test: For $n \geq 1$, $F_n = 2^{2^n} + 1$ is prime if and only if $3^{(F_n - 1)/2} \equiv -1 \pmod{F_n}$. The largest $n$ tested this way is $n = 24$. Pepin's test generalizes to the case when the number $m$ to be tested has $m - 1$ completely (or mostly) factored, and is simpler than the Lucas–Lehmer test.

A heuristic suggests there are only finitely many Fermat primes, since $\sum(\log\log F_n)/\log F_n$ converges. In fact it converges so rapidly, it is thought there are no more Fermat primes after $n = 4$.

Here are some possibly easier questions:

Are there infinitely many primes $p$ with $2^p - 1$ composite?

Are there infinitely many $n$ with $2^{2^n} + 1$ composite?

Actually these are also not known!

Perhaps we can follow the maxim: If you can't solve the problem, generalize it.

Let $\Phi_m(x)$ denote the $m$th cyclotomic polynomial. This is the minimum polynomial for $e^{2\pi i/m}$. Some facts:

$$x^n - 1 = \prod_{m \mid n} \Phi_m(x), \quad \Phi_n(x) = \prod_{m \mid n} (x^m - 1)^{\mu(n/m)}.$$

Also, $\deg(\Phi_n) = \varphi(n)$, Euler's function. Some examples:

$$\Phi_p(x) = (x^p - 1)/(x - 1), \quad \Phi_{2^{n+1}} = (x^{2^{n+1}} - 1)/(x^{2^n} - 1) = x^{2^n} + 1,$$

so that

$$\Phi_p(2) = 2^p - 1, \quad \Phi_{2^{n+1}}(2) = 2^{2^n} + 1.$$

So, here are the generalized and supposedly easier questions:

Are there infinitely many $m$ with $\Phi_m(2)$ prime?

Are there infinitely many $m$ with $\Phi_m(2)$ composite?

The second question has a disappointingly easy answer! In fact two disappointingly easy answers!

Say a prime factor $p$ of $\Phi_m(2)$ is *primitive* if $\ell(p) = m$. Here, $\ell(p)$ denotes the multiplicative order of 2 in $(\mathbb{Z}/p\mathbb{Z})^\times$. Examples: $\ell(5) = 4$, $\ell(7) = 3$, $\ell(17) = 8$, $\ell(31) = 5$.

In fact: If $\ell(p) = m$, then $p \mid \Phi_m(2)$.

Must every prime factor of $\Phi_m(2)$ be primitive?

The answer is "almost." If $m$ is of the form $p^j\ell(p)$ for some prime $p$, with $j \geq 1$, then $p \mid \Phi_m(2)$. In this case $p$ is an *intrinsic* prime factor. It is unique and $p^2 \nmid \Phi_m(2)$.

**Bang** (1886): Each $\Phi_m(2)$ has a primitive prime factor except for $m = 1$ and $m = 6$. (Note that $\Phi_1(2) = 1$ and $\Phi_6(2) = 3$.)

Thus, every $\Phi_m(2)$, where $m = p^j \ell(p)$ for a prime $p > 3$, is composite. For example, $m = 20$, and $\Phi_{20}(2) = 205$, which has the primitive prime factor 41 and the intrinsic prime factor 5.

So, let's reword the problems. Let $\psi_m = \Phi_m(2)/p$ if $m$ is of the form $p^j \ell(p)$, and otherwise let $\psi_m = \Phi_m(2)$.

Are there infinitely many $m$ with $\psi_m$ prime?, composite?

Consider the polynomial $x^4 + 4$. Is it irreducible in $\mathbb{Z}[x]$?

Well, the roots are $\pm 1 \pm i$, which each have degree 2, so it's not irreducible. In fact

$$\begin{aligned} x^4 + 4 = x^4 + 4x^2 + 4 - 4x^2 &= (x^2 + 2)^2 - (2x)^2 \\ &= (x^2 + 2x + 2)(x^2 - 2x + 2). \end{aligned}$$

Similarly,

$$4x^4 + 1 = (2x^2 + 2x + 1)(2x^2 - 2x + 1).$$

Now say $m \equiv 4 \pmod 8$, so $m = 8k+4$. Then $\Phi_m(x) \mid x^{4k+2}+1$, so

$$\Phi_m(2) \mid 4\cdot 2^{4k}+1 = (2\cdot 2^{2k}+2\cdot 2^k+1)(2\cdot 2^{2k}-2\cdot 2^k+1).$$

When $m = 8k+4$, we have $\psi_m = \psi_m^+\psi_m^-$, where

$$\psi_m^+ = \gcd(\psi_m, 2\cdot 2^{2k}+2\cdot 2^k+1), \quad \psi_m^- = \gcd(\psi_m, 2\cdot 2^{2k}-2\cdot 2^k+1).$$

**Schinzel** (1962): For $m \equiv 4 \pmod 8$ and $m > 20$, we have $\psi_m^+ > 1$ and $\psi_m^- > 1$. In particular, $\psi_m$ is composite.

**Theorem** (P, 2024). There are infinitely many $m \not\equiv 4 \pmod 8$ with $\psi_m$ composite. There are infinitely many $m \equiv 4 \pmod 8$ with not both $\psi_m^+$ and $\psi_m^-$ prime.

Sketch of the proof.

Since all primes $p$ with $\ell(p) = m$ divide $2^m - 1$, there are fewer than $m$ of them. Thus, the number of primes $p \le x$ with $\ell(p) \le x^{1/2}/\log x$ is less than

$$\sum_{m \le x^{1/2}/\log x} m < x/(\log x)^2.$$

But the number of all primes $\le x$ is $\sim x/\log x$, so most of these primes $p$ have $\ell(p) > x^{1/2}/\log x$.

Many years ago, Erdős showed that most integers $n$ do not have a divisor near $\sqrt{n}$, where "near" means within a factor $n^\epsilon$. In 1999, **Erdős & Murty** showed the same is true for shifted primes $p - 1$. Thus, most primes $p$ have $\ell(p) > p^{1/2+\epsilon}$.

Let

$$L_d = \{p \in (x/2, x] : \ell(p) = (p-1)/d,\ p \equiv 3 \pmod 4\}.$$

Then

$$\sum_{d<x^{1/2-\epsilon}} \#L_d \sim \frac{x}{4\log x}.$$

Thus, there is some $d < x^{1/2-\epsilon}$ with $\#L_d > x^{1/2}$.

Now consider the numbers $m = (p-1)/d$ for $p \in L_d$. There are more than $x^{1/2}$ of them and they are not $\equiv 4 \pmod 8$. Each $\psi_m$ is divisible by $p = dm + 1$. Further, $\Phi_m(2)$ is of magnitude $2^{\varphi(m)}$, and since $p > x/2$, we have $m > x^{1/2}$, so $\psi_m$ is exponentially larger than $x$. Since $p \mid \psi_m$ and $p \le x$, it follows that $\psi_m$ is composite.

Thus, there are more than $x^{1/2}$ values of $m \le x$ with $m \not\equiv 4 \pmod 8$ with $\psi_m$ composite.

By changing $p \equiv 3 \pmod 4$ to $p \equiv 5 \pmod 8$, the numbers $m$ created are $\equiv 4 \pmod 8$ (since $(2/p) = -1$). Thus, the same argument shows there are more than $x^{1/2}$ values of $m \le x$ with $m \equiv 4 \pmod 8$ and not both $\psi_m^+$ and $\psi_m^-$ are prime.

Additional problems:

Can we do better than $x^{1/2}$ values of $m \leq x$ with $\psi_m$ composite?

Are the standard conjectures, like the abc-conjecture, the Riemann Hypothesis, the Generalized Riemann Hypothesis, and the prime $k$-tuples conjecture helpful for the problems we're considering?

The answer is "yes" to all of these.

First, we can use a variant of the proof presented to get $> x^\theta$ values of $m \le x$ with $m \not\equiv 4 \pmod 8$ and $\psi_m$ composite, where $\theta = 3/5$ or a little larger, and similarly for $\psi_m^+$ and $\psi_m^-$ when $m \equiv 4 \pmod 8$. The key here is a result of Baker and Harman that a positive proportion of primes $p$ have a prime factor of $p-1$ that is $> p^\theta$. We need a version where $p \equiv 3 \pmod 4$ and a version where $p \equiv 5 \pmod 8$, which takes some effort.

Assuming the Elliott–Halberstam conjecture in analytic number theory would allow for $\theta$ to be arbitrarily close to 1.

An important issue that I've not mentioned: A composite number need not be divisible by 2 different primes! I don't know how to unconditionally prove that there are infinitely many $m \not\equiv 4 \pmod 8$ with $\psi_m$ divisible by 2 different primes, and similarly for the 4 (mod 8) case. However, using the abc-conjecture, this issue disappears. Even so, it takes some work.

One deals with the "$a+b=c$" equation: $1+(2^m-1)=2^m$. Let rad($n$) denote the product of the distinct primes dividing $n$. Then $\mathrm{rad}(abc)=2\,\mathrm{rad}(2^m-1)$. Now $\psi_m \mid 2^m-1$, and an averaging argument shows that most of the time $\psi_m > 2^{\epsilon m}$. If $\psi_m$ is a prime power, then $\mathrm{rad}(\psi_m) \le \psi_m^{1/2}$, so that $\mathrm{rad}(\psi_m) < 2^{(\epsilon/2)m}$. This implies that $\mathrm{rad}(abc) < 2c^{1-\epsilon/2}$, contradicting the abc-conjecture. Thus, we may assume $\psi_m$ is not a prime power, and since it is composite, it must be divisible by at least 2 distinct primes.

Hooley has shown, using the Generalized Riemann Hypothesis (actually, the RH for the zeta functions of Kummerian fields) that a positive proportion of primes $p$ have 2 as a primitive root. This then gives $\gg x/\log x$ values of $m \not\equiv 4 \pmod 8$ with $\psi_m$ composite, and we can do even a little better than this. Similarly for the other case.

Assuming the prime $k$-tuples conjecture, there are infinitely many primes $p$ with $2^p - 1$ composite. Indeed, that conjecture gives infinitely many primes $p \equiv 3 \pmod 4$ with $q = 2p + 1$ prime. Then $q \equiv 7 \pmod 8$ so that $(2/q) = 1$, which implies that $q \mid 2^p - 1$. When $p > 3$ we have $q < 2^p - 1$, so $2^p - 1$ is composite. In fact, it is divisible by at least 2 different primes, since otherwise $2^p - 1$ and $2^p$ would be two consecutive powers, contradicting Catalan's conjecture (= Mihăilescu's theorem).

The Bateman–Horn conjecture implies there are $\gg x/(\log x)^2$ such primes $p \leq x$.

We showed earlier by a simple argument that $\ell(p) > x^{1/2}/\log x$ for almost all primes $p \le x$. For each such $p \equiv 3 \pmod 4$ we have $\psi_{\ell(p)}$ composite, and for each such $p \equiv 5 \pmod 8$ we have that not both $\psi^{+}_{\ell(p)}$ and $\psi^{-}_{\ell(p)}$ are prime. We worked harder to show that there are in fact many values of $m \not\equiv 4 \pmod 8$ with $\psi_m$ composite, and similarly for the other case.

However, there are $\gg x/\log x$ values of $p$ in play. Surely there should be many distinct values of $\ell(p)$ among all these $p$'s! This seems hard to prove other than by the means mentioned above. It is amusing that if there are not so many distinct values of $\ell(p)$, then there are values of $m$ with many distinct primitive prime factors of $\psi_m$. So, we have the following result.

**Theorem**: Either there are $\gg x/\log x$ values of $m \le x$ with $m \not\equiv 4 \pmod 8$ and $\psi_m$ composite or the number of primitive prime factors of $\psi_m$ is unbounded as $m$ varies.

Of course, both must be true!

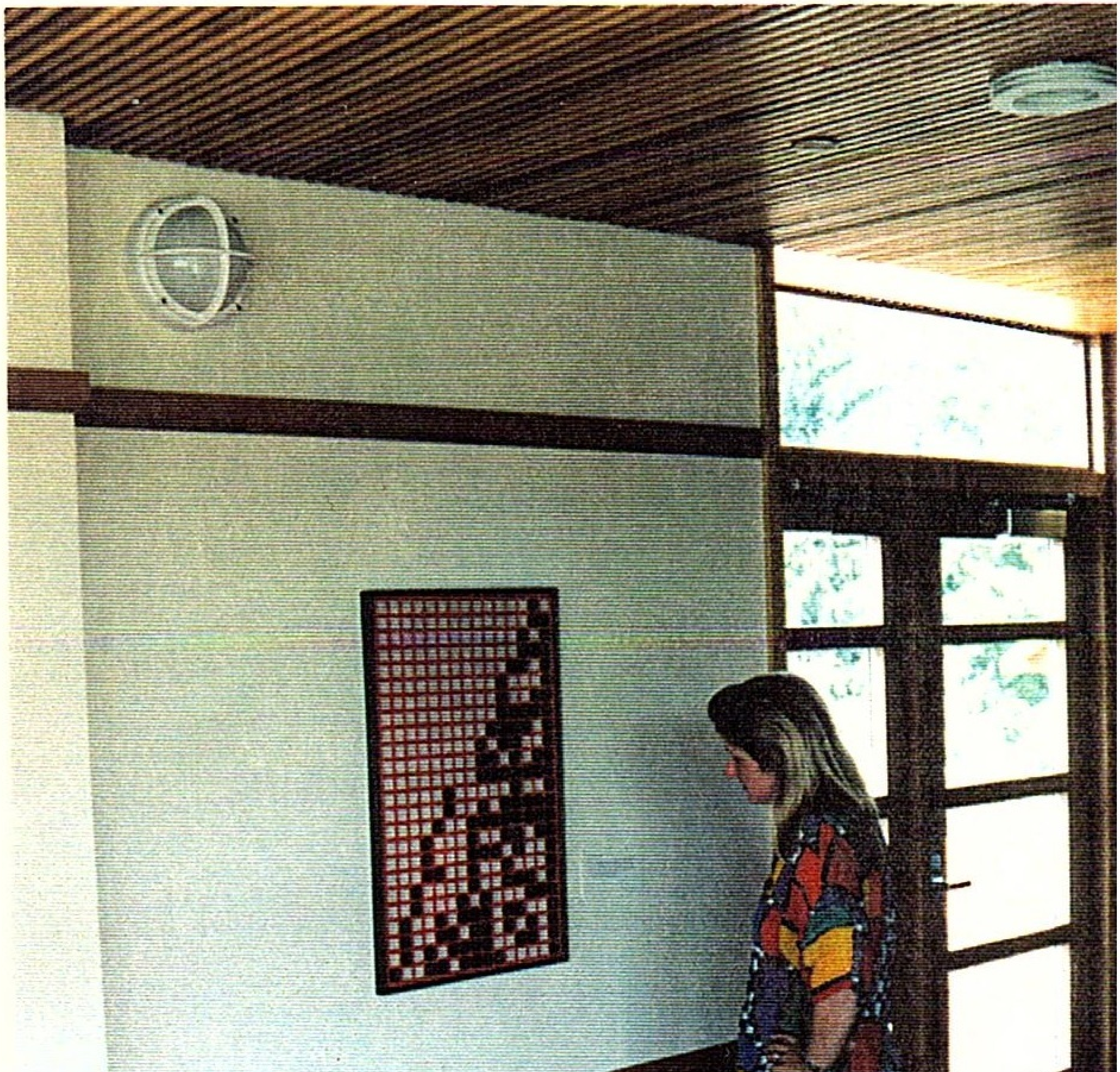
Additional problems, some tractable:

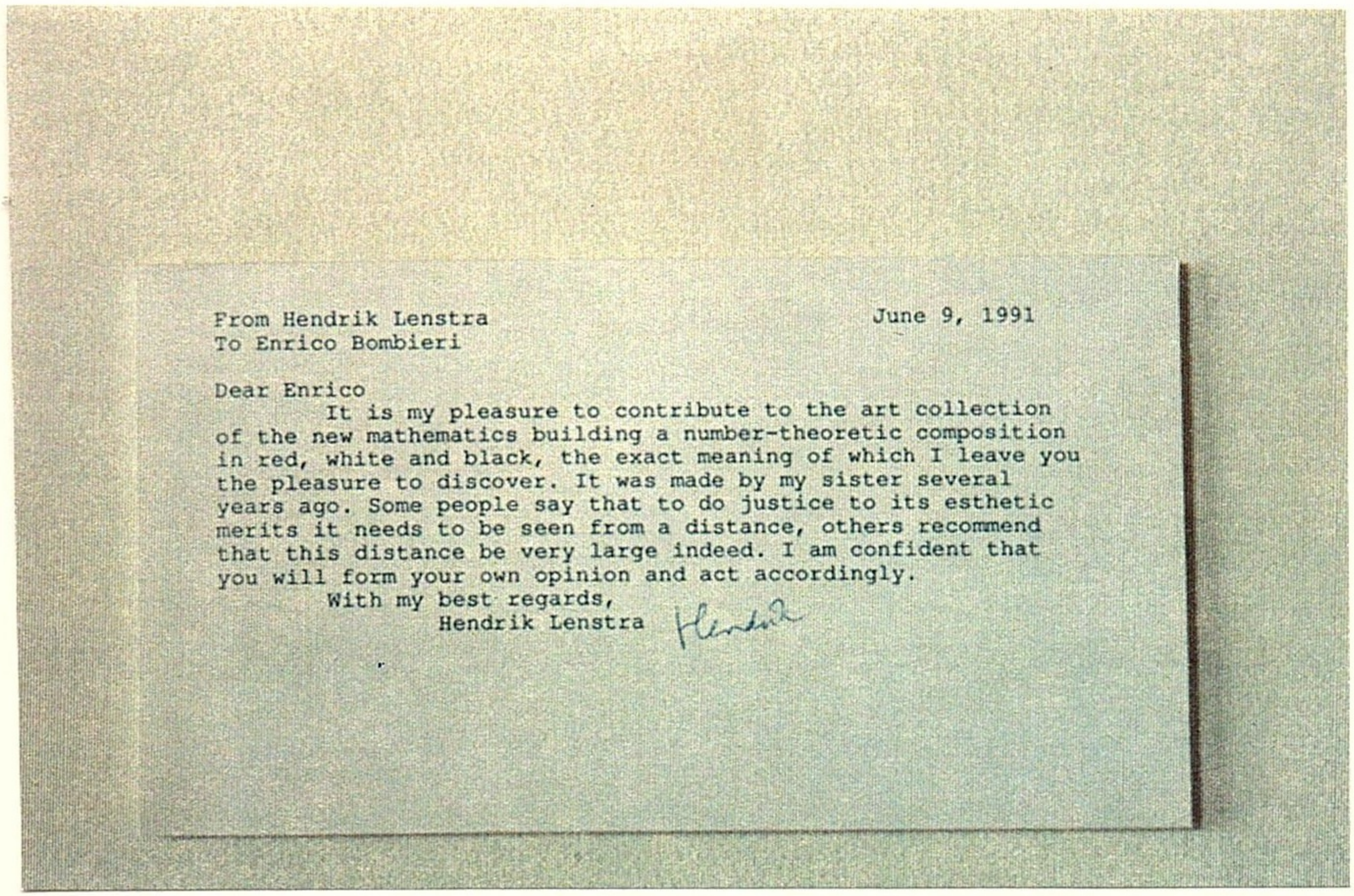
Show there are infinitely many $m \equiv 4 \pmod 8$ with both $\psi_m^+$ and $\psi_m^-$ composite.

Generalize these results to $\Phi_m(a)$ where $a > 2$.

Generalize to the Fibonacci numbers, and similar Lucas sequences.

There must be an elliptic curve analogue ...





From Hendrik Lenstra June 9, 1991
To Enrico Bombieri

Dear Enrico

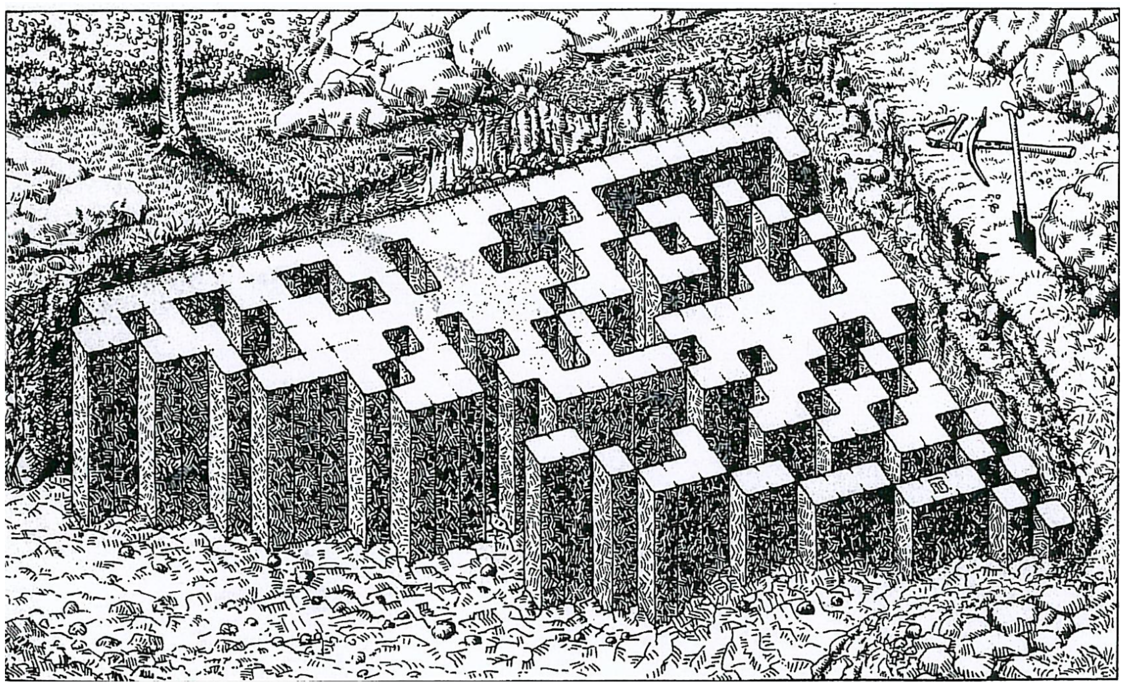
It is my pleasure to contribute to the art collection of the new mathematics building a number-theoretic composition in red, white and black, the exact meaning of which I leave you the pleasure to discover. It was made by my sister several years ago. Some people say that to do justice to its esthetic merits it needs to be seen from a distance, others recommend that this distance be very large indeed. I am confident that you will form your own opinion and act accordingly.

With my best regards,

Hendrik Lenstra *Hendrik*

Princeton, koffiekamer
van 't Inst. of Adv. Technology
priemgetallen v. Mersenne

The digs at Mersenneachus



# Thank you!