Square-free Smooth Polynomials in Residue Classes and Irreducible Polynomial Generators

Christian Bagshaw

January 19, 2023

Euclid's Proof

Euclid's proof of the infinitude of the primes, although typically presented as a proof by contradiction, was worded as a constructive proof.

Given a finite list of primes, Euclid demonstrates how to find a new one, and hence how to generate an infinite sequence of distinct primes:

Proof.

Let $p_1, ..., p_n$ be a finite list of primes.

We define p_{n+1} to be a prime factor of $p_1 p_2 ... p_n + 1$.

The fundamental theorem of arithmetic ensures that p_{n+1} is well-defined, and of course it cannot be a prime already in the list.

This allows us to recursively define an infinite sequence of distinct primes. □

Euclid-Mullin Sequence

If we start with $p_1 = 2$ and take the smallest prime factor of $p_1 p_2 \ldots p_n + 1$ each time, this sequence is called the Euclid-Mullin sequence:

$$2,\ 3,\ 7,\ 43,\ 13,\ 53,\ 5,\ 6221671,\ 38709183810571,\ 139,\ \ldots$$

The natural question is: does every prime occur in this sequence?

We don't know... and in fact there are some who think this problem may be undecidable. But most people think it should be true.

Euclid-Mullin Sequence Variant

If instead we take p_{n+1} to be the largest prime factor of $p_1 p_2 ... p_n + 1$, starting with $p_1 = 2$ the sequence looks like

$$2,\ 3,\ 7,\ 43,\ 139,\ 50207,\ 340999,\ 2365347734339,$$
$$11128622507,\ \ 821384623,\ \ 57088979368630 1597, ...$$

Again we may ask: does every prime occur in this sequence?

Andrew Booker proved that infinitely many primes will not appear!

Moving to $\mathbb{F}_q[T]$

We fix a prime power q, and let $\mathbb{F}_q$ denote the finite field of q elements. We can consider similar questions over $\mathbb{F}_q[T]$.

Sequences of Irreducible Polynomials

We can define an infinite sequence of monic irreducible polynomials as follows: Given a finite list of monic irreducible polynomials $P_1, ..., P_n$, recursively define P_{n+1} to be an irreducible factor of

$$P_1 P_2 ... P_n + 1.$$

Can this sequence be chosen to contain every irreducible polynomial?

We don't know...

What about taking maximums and minimums?

If $q = p$ is prime, then there is a very natural way to define a total order on $\mathbb{F}_q[T]$:

$$f(T) \leq g(T) \iff f(p) \leq g(p)$$

where you consider the coefficients as integers in $\{0, ..., p-1\}$.

Some work of Kurokawa and Satoh [Kurokawa and Satoh, 2008] shows that if you are working in $\mathbb{F}_2[T]$, start with T as the first term in the sequence and take the largest factor each time, then $T^3 + T + 1$ never appears.

If you take the largest factor each time, do you omit infinitely many irreducible polynomials?

Something else we know

Instead of adding 1 each time you could add some $a \in \mathbb{F}_p[T]$.

If you start with T as the first term in the sequence, Kurokawa and Satoh proved that there are infinitely many pairs (p, a) such that $P_1 P_2 ... P_n + a$ is always irreducible, at every stage of the sequence.

Working $(\mathrm{mod}\ P)$

Whether or not an irreducible polynomial P eventually appears in the sequence is of course very closely related to how often

$$P | P_1 P_2 ... P_n + 1$$

for some square-free product of irreducibles $P_1 ... P_n$.

Or equivalently we may want to look at the equation

$$P_1 P_2 ... P_n \equiv -1 \ (\mathrm{mod}\ P).$$

We will talk about this equation for a bit, and then relate it back to these sequences.

For ease of notation, we will call a polynomial k-smooth if all of it's irreducible factors have degree less than k.

Main Result

Theorem (B.)

Let $q \geq 7$ be a prime power, and let $P \in \mathbb{F}_q[T]$ be an irreducible polynomial of degree at least 2. Then every residue class modulo P has a non-constant, square-free, $\deg P$*-smooth representative.*

That is, for any $a \in \mathbb{F}_q[T]$ with $P \nmid a$, there exists distinct irreducible polynomials $P_1, ..., P_n$, all of degree less than $\deg P$*, such that*

$$P_1 P_2 ... P_n \equiv a \pmod{P}.$$

An integer analogue of this result was proven by Andrew Booker and Carl Pomerance [Booker and Pomerance, 2017].

Why $q \geq 7$?

We believe the result holds for $q \geq 2$, except for the cases

$$P = T^3 + T + 1 \text{ and } P = T^3 + T^2 + 1$$

over $\mathbb{F}_2[T]$.

That is, you could manually check that

$$P_1 \ldots P_n \equiv -1 \pmod{P}, \ P_i \text{ distinct and } \deg P_i \leq 2$$

has no solutions.

So what went wrong? The proof of the main result used analytic techniques: ultimately relying on a bound for character sums that wasn't quite strong enough for small q.

A quick definition

For an irreducible polynomial P in $\mathbb{F}_q[T]$, we let χ denote a non-trivial character modulo P (think Dirichlet character for polynomials).

That is, think of a function

$$\chi : \mathbb{F}_q[T] \to \{z \in \mathbb{C} \ : \ |z| = 1\} \cup \{0\}$$

that is completely multiplicative, periodic modulo P and is equal to zero exactly on multiples of P.

Why $q \geq 7$?

The main ingredient in the analytic proof was utilizing the following [Han, 2020]: for any positive integer t,

$$\left| \sum_{\substack{f \in \mathbb{F}_q[T] \\ \deg f = t \\ f \text{ monic}}} \chi(f) \right| \leq q^{t/2} \binom{\deg P - 1}{t}.$$

Trivial bound via triangle inequality: q^t.
But the binomial coefficients on the right can get close to 2^t, so the right can be close to $q^{t/2}2^t$, which could even be worse than trivial for small q.

The difficulty here is we need something completely explicit (not asymptotic) that is non-trivial for very small q.

A good improvement upon this bound for small prime powers would certainly lead to progress.

The computational side

Because this bound gets worse as q gets small, the analytic proof DID NOT work for any irreducible polynomial P over $\mathbb{F}_q[T]$ satisfying

$$
\begin{aligned}
&q = 2... \\
&q = 3, \quad \deg P \leq 48 \\
&q = 4, \quad \deg P \leq 22 \\
&q = 5, \quad \deg P \leq 12 \\
&q = 7, \quad \deg P \leq 5 \\
&q = 8, \quad \deg P \leq 5 \\
&q = 9, \quad \deg P \leq 4 \\
&q = 11, \quad \deg P \leq 4 \\
&\vdots \\
&q = 61, \quad \deg P = 2
\end{aligned}
$$

The computational side

$$
\begin{aligned}
q = 7, &\quad \deg P \leq 5 \\
q = 8, &\quad \deg P \leq 5 \\
q = 9, &\quad \deg P \leq 4 \\
q = 11, &\quad \deg P \leq 4 \\
&\vdots \\
q = 61, &\quad \deg P = 2
\end{aligned}
$$

All these cases were manually checked with a close-to-trivial algorithm in Sage - just a bit of work to keep track of square-free polynomials already found.
Took my Macbook Air about 5 hours...

The computational side

$$q = 3, \quad \deg P \leq 48$$
$$q = 4, \quad \deg P \leq 22$$
$$q = 5, \quad \deg P \leq 12$$

A very efficient method to compute smooth square-free representatives for each equivalence class of many irreducible polynomials could close up this problem for $q = 3, 4, 5$.

Another question

Before relating this all back to generating sequences of irreducible polynomials, we will consider one more question:

Given some a, we could also ask:

How many non-constant, square-free, deg P-smooth representatives does a have modulo P?

Lets additionally only count those such that

$$\deg(P_1 P_2 ... P_n) < k$$

for some integer k.

So we are really counting solutions (say $N(a, k, P)$) to

$$P_1 ... P_n \equiv a \pmod{P}, \quad \begin{array}{l} \deg P_i < \deg P \\ \deg(P_1 ... P_n) < k \\ P_1 ... P_n \text{ square-free.} \end{array}$$

Another question

What might we expect $N(a, k, P)$ to be?

Maybe we would say that when $\deg P$ is really big, square-free polynomials should distribute themselves somewhat uniformly across all the residue classes modulo P. So a has its fair share, but not too much less.
So maybe something like

$$\begin{aligned} N(a, k, P) &\approx E(P)\left(\frac{\text{number of polynomials of degree less than } k}{\text{number of residue classes modulo } P}\right) \\ &= E(P)q^{k-\deg P} \end{aligned}$$

where $E(P)$ is some error term, taking into account that not all polynomials are square-free and $\deg P$-smooth. And maybe if k isn't too big nor small, then $E(P)$ isn't too big.

Another question

Theorem (B. and Shparlinski)

If $k \in [3 \deg P/2, 3 \deg P]$ *then*

$$N(a, k, P) \geq q^{k - \deg P + o(\deg P)}$$

as $\deg P \to \infty$.

As we let $\deg P \to \infty$, $N(a, k, P)$ is at least about $q^{k - \deg P}$ as long as k isnt too big nor too small, and the error is something like $q^{-\epsilon \deg P}$.

Applying this to generate irreducibles

Lets move on from that stuff, and talk about how we can use the main result to generate some polynomials in a similar fashion to Euclid.

First relaxation of Euclid-Mullin

The first corollary to the main result is the following, in the case that $q = p$ is prime.

- We begin with just T as the first term in the sequence.
- At each stage, we choose some of the terms in our sequence $P_1, ..., P_n$ and take as the next term an irreducible factor of

$$P_1 ... P_n + 1.$$

This sequence can be chosen to contain every irreducible polynomial.

Instead of multiplying all the terms in the sequence together and adding 1, we just multiply some of them together and add 1.

Second relaxation of Euclid-Mullin

The second corollary to the main result is the following (for arbitrary q).

- We begin with all of the degree 1 polynomials as the first terms in our sequence.
- At each stage, we split the terms in our sequence into two sets $\{P_i\}$ and $\{P_j\}$ and take as the next term an irreducible factor of

$$\prod_i P_i + \prod_j P_j.$$

This sequence can be chosen to contain every irreducible polynomial.

Instead of multiplying all the terms in the sequence together and adding 1, we split them into two sets and add the products.

Do we need to start with all linear polynomials?

The way I stated this result was that we need to start with all the degree 1 (linear) polynomials already in the sequence.

> Do we need to start with all degree 1 polynomials in the sequence for this to work? Can we just start with "T" in the sequence, for example, and still generate all the irreducible polynomials this way?

Even to computationally check this for a given prime power is difficult. Working in $\mathbb{F}_3[T]$, again my MacBook Air tells me there are a total of 1397132 possibilities for the first 7 terms in the sequence, none of which contain more than 3 degree 1 polynomials (of course there are 6 linear polynomials in $\mathbb{F}_3[T]$).

Thank you!

References


Booker, A. and Pomerance, C. (2017).
Squarefree smooth numbers and Euclidean prime generators.
Proc. Am. Math. Soc, 145:5035–5042.


Han, D. (2020).
A note on character sums in function fields.
Finite Fields and Their Applications, 68:101734.


Kurokawa, N. and Satoh, T. (2008).
Euclid prime sequences over unique factorization domains.
Experimental Mathematics, 17:145–152.