Sieve Methods: Polynomials, Parity, and Twin Primes

Trinity Term 2026

Word Count: 7499 (provided by Overleaf via TeXcount)

Supervisor: Prof. James Maynard

Master of Mathematics and Philosophy

Mathematical Institute
University of Oxford

Contents

Notation

We use the following conventions, where a symbol carries its analogous meaning from $\mathbb{Z}$ to $\mathbb{F}_q[t]$.

Sets and rings

$\mathbb{N}, \mathbb{Z}, \mathbb{R}, \mathbb{C}$	natural numbers, integers, reals, complex numbers.
$\mathbb{F}_q,\ \mathbb{F}_q[t]$	finite field of q elements and the polynomial ring over it.
$\mathbb{P}$	rational primes (resp. monic irreducibles in $\mathbb{F}_q[t]$).

Divisibility and arithmetic

(a, b)	$\gcd(a, b)$.
$[a, b]$	$\operatorname{lcm}(a, b)$.
$a \mid b$	a divides b.
$p, \mathfrak{p}$	a generic prime (resp. monic irreducible).
$\deg f$	degree of $f \in \mathbb{F}_q[t]$.
$\lvert D\rvert$	the norm of $D \in \mathbb{F}_q[t]$, $\lvert D\rvert = q^{\deg D} = \#\mathbb{F}_q[t]/\langle D\rangle$.
$\lfloor x\rfloor$	floor of x.
$\mathbf{1}_{\{\cdot\}}$	indicator function.

Arithmetic functions

μ	Möbius function.
λ	Liouville function, $\lambda(\cdot) = (-1)^{\Omega(\cdot)}$.
Λ	von Mangoldt function.
Ω	number of prime factors with multiplicity.
υ	number of distinct prime factors.
d	number of divisors.
rad	product of its distinct prime factors.
φ	Euler's totient.
$\zeta(s),\ \zeta_q(s)$	Riemann zeta function and its $\mathbb{F}_q[t]$ analogue.
$f * g$	Dirichlet convolution.

Asymptotic notation

$f \sim g$	$f/g \to 1$.
$f = O(g),\ f \ll g$	$\lvert f\rvert \le C\lvert g\rvert$ for some constant C.
$f = o(g)$	$f/g \to 0$.
$f \asymp g$	$f \ll g$ and $g \ll f$.

1 Introduction

> When you have eliminated the impossible, whatever remains, however improbable, must be the truth.
>
> *Sherlock Holmes*

Counting is perhaps the simplest form of mathematical inquiry we have; "How many of x is there?," is the origin of much of modern number theory. These questions still remain relevant today with questions like, "Are there infinitely many primes of the form $n^2 + 1$," or, "Is there an odd perfect number?" Counting certain objects can be difficult, however, as some objects are more readily described by *what they are not* rather than *what they are* (e.g. primes are numbers that are *not* divisible by anything except 1 and itself). This leads to techniques like inclusion-exclusion-style counting. *Sieve theory* is the systematic refinement of these methods to bound the number of elements in a set. The simplicity of sieve methods, however, is their bane as well. These methods are capped by the parity problem: sieves cannot distinguish between numbers with an even number of prime factors and those with an odd number, and this single obstruction is what stands between the machinery and conjectures like the infinitude of twin primes. The question is thus: where precisely does this limitation land, and what does it take to strengthen these counting methods for accurate results?

1.1 Goals and Directions

Our goal is to understand the weight of the parity problem of sieve techniques. We do so through two major considerations: 1) we leverage the structural similarity of the rings of integers $\mathbb{Z}$ and polynomials $\mathbb{F}_q[t]$ to isolate what analytic and geometric information do or do not permit bypassing parity; 2) we focus on the problem of the *twin prime conjecture*—which hypothesizes there are infinitely many pairs of primes that differ by 2—where the asymmetry between the two settings is sharpest as it is proven in the case of $\mathbb{F}_q[t]$ while still open in $\mathbb{Z}$. These aspects allow $\mathbb{F}_q[t]$ to be the ideal "control experiment" to analyze what sieves lack and what suffices to support them.

We proceed in four major parts: 1) we set up and motivate the basics of sieve theory; 2) we adapt the Selberg sieve to polynomials and give an upper bound on the number of twin irreducible polynomials; 3) we discuss the *parity problem* that is the major hurdle to sieve arguments; and 4) we end with a comparison of the twin prime conjecture over $\mathbb{Z}$ vs. $\mathbb{F}_q[t]$ and what it takes to overcome parity. We provide some novel calculations in translating results from $\mathbb{Z}$ to $\mathbb{F}_q[t]$.

2 The Basic Sieve Set-Up

The principle of sieve theory is to use inclusion-exclusion counting to "sift" away "bad" elements and be left with a count of the "good" elements.

Example (Counting primes)**.** Let $\pi(X) = \#\{\text{primes} \leq X\}$ be the prime-counting function, and say we want to calculate $\pi(100)$. The key characteristic of primes ≤ 100 is that they have no prime factors $\leq \sqrt{100} = 10$, i.e. they are not divisible by $p_1 = 2, p_2 = 3, p_3 = 5, p_4 = 7$. By inclusion-exclusion, we can "sift out" all the composite numbers from $\mathcal{A} = \{n \leq 100\}$ via this divisibility criterion.

$$\pi(100) + 1 - 4 = 100 - \sum_i \left\lfloor \frac{100}{p_i} \right\rfloor + \sum_{i<j} \left\lfloor \frac{100}{p_i p_j} \right\rfloor - \sum_{i<j<k} \left\lfloor \frac{100}{p_i p_j p_k} \right\rfloor + \left\lfloor \frac{100}{p_1 p_2 p_3 p_4} \right\rfloor = 22$$

(noting that we need to count the 4 primes we sifted by, and that 1 is not a prime).

We now generalize this divisibility principle to a more general sifting set. The idea is that by carefully designing an ambient set $\mathcal{A}$ and what prime factors we allow to survive the sieve, we can estimate more complex number-theoretic quantities. We follow [8][3].

Definition 2.1 (Sifted set)**.** Let $\mathcal{A}$ be a finite multiset or sequence from $\mathbb{N}$, $\mathcal{P} \subseteq \mathbb{P} = \{p : p \text{ is prime}\}$, and $z \in \mathbb{R}$. Then the *sifted set* or *sifted function* of $\mathcal{A}$ is

$$\mathcal{S}(\mathcal{A}, \mathcal{P}; z) = \{n \in \mathcal{A} : \forall p \in \mathcal{P} \ \ (p \mid n \Rightarrow p \geq z)\}$$

i.e. the elements of $\mathcal{A}$ that have no prime factor less than z. We sometimes abuse notation and write $\mathcal{S}(\mathcal{A}, \mathcal{P}; z)$ to refer to both the set above, or its cardinality, with context making it clear.

Let $\Pi := \Pi(\mathcal{P}; z) := \prod\{p \in \mathcal{P} : p < z\}$, so in particular $\mathcal{S}(\mathcal{A}, \mathcal{P}; z) = \{n \in \mathcal{A} : (n, \Pi(\mathcal{P}; z)) = 1\}$. If we let $\mathcal{A}_d = \{a \in \mathcal{A} : a \equiv 0 \bmod d\}$ i.e. the elements of $\mathcal{A}$ divisible by d, we obtain:

$$\mathcal{S}(\mathcal{A}, \mathcal{P}; z) = \sum_{\substack{a \in \mathcal{A} \\ (a,\Pi)=1}} 1 = \sum_{a \in \mathcal{A}} \sum_{d \mid (a,\Pi)} \mu(d) = \sum_{d \mid \Pi} \mu(d) \sum_{\substack{a \in \mathcal{A} \\ d \mid a}} 1 = \sum_{d \mid \Pi} \mu(d) \#\mathcal{A}_d$$

using the standard properties $d \mid (a, b) \iff d \mid a, b$ and $\sum_{d \mid n} \mu(d) = \mathbf{1}_{\{n=1\}}$. This is precisely the inclusion-exclusion formula we used earlier.

Remark. The condition $n \in S(\mathcal{A}, \mathcal{P}; z) \iff (n, \Pi(\mathcal{P}; z)) = 1$ is convenient for its expression as a divisor sum of μ, but not the only choice: replacing it with $|\mu(n)| = 1$, for instance, sifts for squarefree elements, and most of the analysis carries through with the corresponding sifting function in place of μ [16].

This formulation is not yet usable for concrete calculation for two reasons. First, the sum has exponentially many terms: taking $\Pi(\mathcal{P}; z)$ as the product of all primes below z gives $2^{\pi(z)}$ divisors. Second, $\#\mathcal{A}_d$ is rarely available exactly, even for $\mathcal{A} = \{\text{primes} \leq x\}$, only asymptotic and on-average results in arithmetic progressions are known (see Theorem 4.4). We adopt instead the heuristic decomposition

$$\#\mathcal{A}_d = |\mathcal{A}| \frac{\omega(d)}{d} + R_d$$

where $\omega : \mathbb{N} \to \mathbb{R}_{\geq 0}$ is multiplicative and R_d absorbs the error; this form must be verified for each $\mathcal{A}$. Heuristically, $\omega(d)$ counts the residue classes modulo d that produce elements of $\mathcal{A}$ divisible by d, and multiplicativity reflects independence of divisibility across coprime moduli—exact in many cases via the Chinese Remainder Theorem.

Example (Prime counting)**.** Let $\mathcal{A} = \{n \in \mathbb{N} : n \leq X\}$, $\mathcal{P} = \mathbb{P}$, $z = \sqrt{X}$. Then $\mathcal{S}(\mathcal{A}, \mathcal{P}; z) = \pi(X) - \pi(z)$. $\omega(p) = 1$ for all p. This was the set we were counting above for $X = 100$.

Example (The twin-prime conjecture)**.** Let $\mathcal{A} = \{n(n+2) : n \leq X\}$, $\mathcal{P} = \mathbb{P}$, $z = \sqrt{X+2}$. Note that if $n(n+2)$ has no prime factor $\leq z$, then both n and $n+2$ are prime. The equation $n(n+2) \equiv 0 \bmod p$ has 1 solution when $p = 2$ and 2 solutions for $p \neq 2$. Hence $\omega(2) = 1$ and $\omega(p) = 2$ for all $p \neq 2$. Thus, $\mathcal{S}(\mathcal{A}, \mathcal{P}; z)$ counts the number of twin-prime pairs in $(\sqrt{X+2}, X)$.

Example (Goldbach's conjecture)**.** Let $\mathcal{A} = \{n(2N-n) : 2 \leq n \leq 2N-2\}$, $\mathcal{P} = \mathbb{P}$, $z = \sqrt{2N}$. $\omega(p) = 2$ for all p. Then $\mathcal{S}(\mathcal{A}, \mathcal{P}; z)$ counts the number of prime pairs that sum to N.

Letting $X = |\mathcal{A}|$, we have by the multiplicativity of μ and ω, a type of Euler-product rewriting:

$$\mathcal{S}(\mathcal{A}, \mathcal{P}; z) = \sum_{d|\Pi} \mu(d) \#\mathcal{A}_d = X \sum_{d|\Pi} \mu(d) \frac{\omega(d)}{d} + \sum_{d|\Pi} \mu(d) R_d = X \prod_{\substack{p \in \mathcal{P} \\ p < z}} \left(1 - \frac{\omega(p)}{p}\right) + \sum_{d|\Pi} \mu(d) R_d$$

noting that $\mu(p) = -1$ for a prime p. The triangle inequality gives us the formal *Sieve of Eratosthenes-Legendre*:

$$\left| \mathcal{S}(\mathcal{A}, \mathcal{P}; z) - X \prod_{\substack{p \in \mathcal{P} \\ p < z}} \left(1 - \frac{\omega(p)}{p}\right) \right| \leq \sum_{d|\Pi} |R_d|$$

Continuing the example of counting primes, we take $\omega(p) = 1$ for all primes p ($\approx 1/p$ numbers are divisible by p). To convert the sieve quantity into a proper estimate, we invoke a theorem of Merten [8] [3]:

Theorem 2.2. $\prod_{p<z} \left(1 - \frac{1}{p}\right) \sim \frac{e^{-\gamma}}{\log z}$ *where* γ *is the Euler-Mascheroni constant.*

Putting it all together, we have:

$$\left| \mathcal{S}(\mathcal{A}, \mathcal{P}; z) - x \frac{e^{-\gamma}}{\log z} \right| \leq \sum_{d|\Pi} |R_d| \leq 2^{\pi(z)}$$

The remainder term on the right can be very large though, making our estimate of $\mathcal{S}(\mathcal{A}, \mathcal{P}; z)$ uncontrolled! To make this useful, ideally the remainder term is small compared to the main term i.e. something like $\sum |R_d| = o(x/\log z)$. A quick calculation with the Prime Number Theorem shows the choice of $z \leq \log x$ is sufficient [8].

But notice $\log x \leq \sqrt{x}$, so this choice of z is much coarser for our purposes of trying to count $\pi(x) - \pi(z)$! Our count now additionally includes composite numbers, for example, pq with $\log x \leq p, q \leq \sqrt{x}$. **This is the fundamental tradeoff of sieves**: as we increase z, our main term shrinks (as we multiply by more terms $1 - \omega(p)/p \leq 1$) and makes our count more accurate, but makes our remainder term larger. To get better results out of sieves, different types of refined techniques have emerged. These range from combinatorial ideas (e.g. Brun's sieve) that focus on the behavior of inclusion-exclusion, to analytic ones that have nice proofs of optimality. We focus on an example of the latter to highlight precisely how much we can squeeze out of sieves to see where they lack and need support.

3 The Selberg Sieve over $\mathbb{F}_q[t]$

We have seen how sieves can formalize and make tractable certain number theoretic problems, like counting twin primes. Before we develop deeper, stronger sieve techniques that attempt to tackle such a problem, we first expand to a broader application of sieves to polynomials over finite fields. $\mathbb{Z}$ and $\mathbb{F}_q[t]$ share many of the same algebraic properties that allow sieves to transfer nicely over, but the latter has much stronger analytic and geometric results. This allows $\mathbb{F}_q[t]$ to be a nice testing ground for probing sieves' capabilities and strengths; if sieves alone cannot prove a result in $\mathbb{F}_q[t]$ with access to results like the Riemann hypothesis, there is not much hope it will do better in $\mathbb{Z}$.

3.1 A Recap on Euclidean Domains

There are essentially three key components that sieves require: 1) a set of "primitive" objects to sift by; 2) a notion of size to truncate our sieve counting; and 3) a Chinese Remainder Theorem which underlies the multiplicativity of the density function ω. These fit precisely the qualities of *Euclidean domain*. Recall that a *Euclidean domain* (ED) is an integral domain that admits a division algorithm with respect to a norm $N : R \setminus \{0\} \to \mathbb{N}$, and that primality and irreducibility coincide (as such we use the terms interchangeably).

The most natural ED to work in is $\mathbb{Z}$, where the Fundamental Theorem of Arithmetic gives us our unique factorization into primes and $N(n) = |n|$. The more tacitly understood fact is that the ring of polynomials $K[t]$ is also an ED for a field K: we can factor any polynomial into a product of irreducible polynomials and have a norm of $N(f) = \deg f$. This sets up a correspondence of our sieve formulation for polynomials with irreducible polynomials taking the place of prime numbers. We look at in particular the ring of polynomials over a finite field $\mathbb{F}_q[t]$ as opposed to $\mathbb{Z}[t]$, as that allows us to control the size of the number of residue classes modulo a polynomial. Note the number of monic polynomials of degree n in $\mathbb{F}_q[t]$ is q^n.

To keep the parallels, we use the same notation as before (again allowing context to determine the setting/objects we are working with). Let $\mathcal{A} \subseteq \mathbb{F}_q[t]$, and $\mathcal{P} \subseteq \mathbb{P} = \{\mathfrak{p}(t) \in \mathbb{F}_q[t] : \mathfrak{p} \text{ monic, irreducible}\}$, and $m \in \mathbb{Z}$ an integer that will determine our "prime cut-off" like z did before. Thus, we have

$$\mathcal{S}(\mathcal{A}, \mathcal{P}; m) = \{f(t) \in \mathcal{A} : \forall \mathfrak{p} \in \mathcal{P} \ (\mathfrak{p} \mid f \Rightarrow \deg \mathfrak{p} \geq m)\}$$

That is, $\mathcal{S}(\mathcal{A}, \mathcal{P}; m)$ is the set of polynomials from $\mathcal{A}$ that only have irreducible factors of degree at least m.

Remark. Every polynomial can be written as $f = c \cdot \tilde{f}$ with $\tilde{f}$ monic, so counting monic vs. non-monic polynomials only differs by a factor $q - 1$.

Again, let $\Pi := \Pi(\mathcal{P}; m) := \prod\{\mathfrak{p} \in \mathcal{P} : \deg \mathfrak{p} < m\}$, and $\mathcal{A}_D = \{f(t) \in \mathcal{A} : f(t) \equiv 0 \bmod D(t)\}$. We use the same heuristic $\#\mathcal{A}_D = |\mathcal{A}|\frac{\omega(D)}{|D|} + R_D$.

$$\mathcal{S}(\mathcal{A}, \mathcal{P}; m) = X \sum_{D|\Pi} \mu(D)\frac{\omega(D)}{|D|} + \sum_{D|\Pi} \mu(D) R_D = q^n \sum_{D|\Pi} \mu(D)\frac{\omega(D)}{q^{\deg D}} + \sum_{D|\Pi} \mu(D) R_D$$

where $|D| = \#\mathbb{F}_q[t]/\langle D(t)\rangle = q^{\deg D}$ is the number of residue classes modulo $D(t)$, and $X = |\mathcal{A}| = q^n$. We will also make the assumption that $0 \le \omega(D) < |D| = q^{\deg D}$ so that our sieve is non-trivial.

Example. Counting degree n irreducibles like primes functions exactly the same as before. $\omega(\mathfrak{p}) = 1$ for all $\mathfrak{p}$, and $R_D = 0$ for $\deg D \le n$: a degree n polynomial f divisible by D if and only if it can be written as $f(t) = q(t)D(t)$ with $\deg q = n - \deg D$. Hence there are exactly $q^{n-\deg D} = q^n \cdot 1/q^{\deg D}$ polynomials divisible by D. For $\deg D > n$, then $D \nmid f$ for $\deg f = n$, so $\#\mathcal{A}_D = 0, R_D = -q^n/|D|$. By a similar analysis using more elementary facts of $\mathbb{F}_q[t]$ [11], we have

$$\left|\mathcal{S}(\mathcal{A}, \mathcal{P}; m) - q^n \frac{e^{-\gamma}}{m}\Big(1 + O\big(1/m\big)\Big)\right| \le \sum_{\substack{D|\Pi \\ \deg D > n}} \frac{q^n}{|D|}$$

If we let $x = q^n$, we see this is the same form as the integer case. Ideally, taking $m = \lceil n/2 \rceil$ gives a similar count of irreducibles as we deduced before, but again we need the error small. If we let $m \le \log_q n$, then our remainder is 0 and this is exact, so already we see the beginnings of the strengths of $\mathbb{F}_q[t]$ with much cleaner error control in this simple example.

3.2 The Selberg Sieve and its Fundamental Theorem

We now develop some more sophisticated sieve technology in what is known as the *Selberg sieve*. The idea of this sieve is to approximate sieve quantities with smooth parameters to optimize. We present the sieve and its results as an adaptation of [8], [3], [16], [15].

Remark. Webb [15] formulates the sieve more generally, specifying for each irreducible $\mathfrak{p}_i$ a set of residue classes $\{r_{ij}\}_j$ to remove rather than removing all multiples of $\mathfrak{p}_i$. The two formulations are interchangeable on the problems we consider, so we keep the divisibility setup of [8], [3].

3.2.1 Approximations and Truncations

The core identity that evaluating our sifted set $\mathcal{S}(\mathcal{A}, \mathcal{P}; m)$ relies on is $\sum_{D|N} \mu(D) = 1 \iff N(t) \equiv \mathbf{1}$ is the identically 1 polynomial. By letting $N(t) = (f(t), \Pi(\mathcal{P}; m))$, this gives us

an indicator function for f having no small factors. As noted before, this sum is quantitatively dangerous for obtaining useful estimates summing over $2^{\#\text{irreducible factors of N}}$ which risks blowing up our remainder. To control this remainder, we truncate our sum:

$$\mathcal{S}(\mathcal{A}, \mathcal{P}; m) \approx \sum_{\substack{D | \Pi(\mathcal{P}; m) \\ \deg D \leq y}} \mu(D) \# \mathcal{A}_D$$

y is sometimes referred to as the *sieve level*. However, because of the alternating nature of μ, the choice of y determines whether this truncation is an upper or lower bound (this is the basis for *combinatorial sieves*, like *Brun's sieve*; see [8][3]). To control the behavior of this truncated approximation, we bound our sifted set with another sum that always bounds regardless of truncation.

Thus Selberg opts to find μ^+, μ^- such that

$$\sum_{D|N} \mu^-(D) \leq \sum_{D|N} \mu(D) \leq \sum_{D|N} \mu^+(D)$$

to give the corresponding bound on the sifted set

$$\sum_{D | \Pi(\mathcal{P}; m)} \mu^-(D) \# \mathcal{A}_D \leq \mathcal{S}(\mathcal{A}, \mathcal{P}; m) \leq \sum_{D | \Pi(\mathcal{P}; m)} \mu^+(D) \# \mathcal{A}_D$$

We will focus on the upper bound, as we will see later that lower bounds are almost always trivial. Thus, our goal is to find an appropriate μ^+ and minimize the righthand side of the inequality

$$\mathcal{S}(\mathcal{A}, \mathcal{P}; m) \leq X \overbrace{\sum_{D | \Pi(\mathcal{P}; m)} \mu^+(D) \frac{\omega(D)}{|D|}}^{\mathcal{M}} + \overbrace{\sum_{D | \Pi(\mathcal{P}; m)} |\mu^+(D)||R_D|}^{\mathcal{R}} \tag{1}$$

to get the best results possible.

A natural choice for the upper bound is to take some nonnegative function, like $\mu^+(D) = \lambda_D^2$, for some parameters $\lambda_D \in \mathbb{R}$. This clearly satisfies the constraints if we let $\lambda_{\mathbf{1}} = 1$ (where $\mathbf{1}$ is the constant 1 function), but this is practically useless. To minimize the main term $\mathcal{M}$, we just set $\lambda_D = 0$ for $D(t) \not\equiv \mathbf{1}$, and we are left with $X \cdot \lambda_{\mathbf{1}} \frac{\omega(1)}{|1|} = X$ i.e. we have obtained the trivial bound $\mathcal{S}(\mathcal{A}, \mathcal{P}; m) \leq X$.

To allow for more substantial control, we add some cross-terms. Selberg chooses

$$\mu^+(D) = \sum_{[D_1, D_2] = D} \lambda_{D_1} \lambda_{D_2}$$

If $N(t) \neq \mathbf{1}$, then

$$\sum_{D|N} \mu^+(D) = \sum_{[D_1, D_2] | N} \lambda_{D_1} \lambda_{D_2} = \sum_{D_1, D_2 | N} \lambda_{D_1} \lambda_{D_2} = \Big(\sum_{D|N} \lambda_D \Big)^2 \geq 0$$

using $[f,g] \mid h \iff f,g \mid h$. If $N = 1$, then $\sum_{D|N} \mu^+(D) = \mu^+(1) = \lambda_{\mathbf{1}}^2 = 1 = \sum_{D|N} \mu(D)$. Importantly, this bound μ^+ respects truncation: for any choice of λ_D, (1) holds. Specifically, if we let $\lambda_D = 0$ for $\deg D \geq y/2$, then $\mu^+(D) = 0$ for $\deg D \geq y$: suppose $D = [D_1, D_2]$, so $y \leq \deg D = \deg([D_1, D_2]) \leq \deg(D_1 D_2)$, and hence at least one $\deg D_i \geq y/2$ and so their corresponding $\lambda_{D_i} = 0$, making the sum defining $\mu^+(D) = 0$.

3.2.2 A Minimization Problem

With a truncation on $\mu^+(D)$ imposed by the λ_D as above, the main term $\mathcal{M}$ looks like:

$$\begin{aligned}
\mathcal{M} = \sum_{\substack{D|\Pi(\mathcal{P};m) \\ \deg D<y}} \mu^+(D)\frac{\omega(D)}{|D|} &= \sum_{\substack{D|\Pi(\mathcal{P};m) \\ \deg D<y}} \sum_{[D_1,D_2]=D} \lambda_{D_1}\lambda_{D_2}\frac{\omega(D)}{|D|} \\
&= \sum_{\substack{[D_1,D_2]|\Pi(\mathcal{P};m) \\ \deg D_1,\deg D_2<y/2}} \lambda_{D_1}\lambda_{D_2}\frac{\omega([D_1,D_2])}{|[D_1,D_2]|} \\
&= \sum_{\substack{D_1,D_2|\Pi(\mathcal{P};m) \\ \deg D_1,\deg D_2<y/2}} \frac{\omega(D_1)\lambda_{D_1}}{|D_1|}\frac{\omega(D_2)\lambda_{D_2}}{|D_2|}\frac{|(D_1,D_2)|}{\omega((D_1,D_2))}
\end{aligned}$$

using the fact that ω, $|f| = q^{\deg f}$ are multiplicative and $(f,g)[f,g] = fg$. Let $\xi = y/2$. We have a quadratic minimization problem over the variables $\{\lambda_D \in \mathbb{R} : \deg D \leq \xi\}$, which we can solve with standard methods.

3.2.3 Diagonalizing $\mathcal{M}$

The problematic factor in the terms are $|(D_1,D_2)|/\omega((D_1,D_2))$, coupling the D_1, D_2. The gcd factors suggests separating them with a divisor relation, so a Möbius inversion will do. Let

$$\frac{1}{g(h(t))} = \left(\mu * \frac{|f(t)|}{\omega(f(t))}\right)(h(t)) = \sum_{L|h} \mu\left(\frac{h}{L}\right)\frac{|L|}{\omega(L)} = \frac{|h(t)|}{\omega(h(t))}\prod_{\mathfrak{p}|h}\left(1 - \frac{\omega(\mathfrak{p})}{|\mathfrak{p}|}\right)$$

with $*$ the Dirichlet convolution. Note that g is multiplicative as ω is, and the product form on the right can be obtained by directly calculating $1/g$ for irreducibles and using multiplicativity. So in particular, $\sum_{L|h} 1/g(L(t)) = |h|/\omega(h)$. Note $1/g$ is multiplicative.

Continuing to rewrite,

$$\begin{aligned}
\mathcal{M} &= \sum_{\substack{D_1,D_2|\Pi(\mathcal{P};m)\\ \deg D_1,\deg D_2<\xi}} \frac{\omega(D_1)\lambda_{D_1}}{|D_1|}\frac{\omega(D_2)\lambda_{D_2}}{|D_2|}\frac{|(D_1,D_2)|}{\omega((D_1,D_2))} \\
&= \sum_{\substack{D_1,D_2|\Pi(\mathcal{P};m)\\ \deg D_1,\deg D_2<\xi}} \frac{\omega(D_1)\lambda_{D_1}}{|D_1|}\frac{\omega(D_2)\lambda_{D_2}}{|D_2|}\sum_{L|(D_1,D_2)}\frac{1}{g(L)} \\
&= \sum_{\substack{L|\Pi(\mathcal{P};m)\\ \deg L<\xi}}\frac{1}{g(L)}\sum_{\substack{D_1,D_2|\Pi(\mathcal{P};m)\\ \deg D_1,\deg D_2<\xi\\ L|D_1,D_2}}\frac{\omega(D_1)\lambda_{D_1}}{|D_1|}\frac{\omega(D_2)\lambda_{D_2}}{|D_2|} \\
&= \sum_{\substack{L|\Pi(\mathcal{P};m)\\ \deg L<\xi}}\frac{1}{g(L)}\left(\sum_{\substack{D|\Pi(\mathcal{P};m)\\ \deg D<\xi\\ L|D}}\frac{\omega(D)\lambda_D}{|D|}\right)^2 \\
&= \sum_{\substack{L|\Pi(\mathcal{P};m)\\ \deg L<\xi}}\frac{1}{g(L)}y_L^2
\end{aligned}$$

where $y_D = \sum \frac{\omega(L)\lambda_L}{|L|}$ over the sum of divisors satisfying the divisibility conditions. Thus the change of variables $\langle\lambda_D\rangle \mapsto \langle y_D\rangle$ diagonalizes the form. This is an invertible linear change in variables, so a minima obtained in y_D is a minima in λ_D. We can easily obtain the inverse change of variables.

Proposition 3.1.

$$\mu(D)\frac{\omega(D)\lambda_D}{|D|} = \sum_{\substack{L|\Pi(\mathcal{P};m)\\ \deg L<\xi\\ D|L}}\mu(L)y_L$$

.

Proof. Note $y_D = \sum \frac{\omega(L)\lambda_L}{|L|} = (\frac{\omega(Lf)\lambda_{Lf}}{|Lf|} * \mathbf{1})(\Pi/L)$, so we get the inverse change of variables by Möbius inversion. The only subtlety being that we combine the divisibility requirements to have the correct shape (i.e. write $D = LE$ for some E). Multiply through by $\mu(D)$ and note $\mu(D)\mu(L/D) = \mu(L)$ for squarefree L (which we have as $L \mid \Pi(\mathcal{P}; m)$). □

3.2.4 Optimizing the Parameters

Our constraint $\lambda_{\mathbf{1}} = 1$ translates to $1 = \frac{\omega(\mathbf{1})\lambda_{\mathbf{1}}}{|\mathbf{1}|} = \sum_{L|\Pi,\ \deg L<\xi} \mu(L)y_L$. By Cauchy-Schwarz, we obtain

$$1 = \left(\sum_{L|\Pi,\ \deg L<\xi} \mu(L)y_L\right)^2 = \left(\sum_{L|\Pi,\ \deg L<\xi} \mu(L)\sqrt{g(L)}\frac{1}{\sqrt{g(L)}}y_L\right)^2$$
$$\leq \left(\sum_{L|\Pi,\ \deg L<\xi} g(L)\right)\left(\sum_{L|\Pi,\ \deg L<\xi} \frac{1}{g(L)}y_L^2\right)$$

as $\mu(L)^2 = 1$ for squarefree L. Let the left sum be $\mathcal{G}(\xi, m) := \sum g(L)$, so we have $\mathcal{M} \geq \mathcal{G}(\xi, m)^{-1}$ with equality if and only if $y_L/\sqrt{g(L)} = c\mu(L)\sqrt{g(L)}$ for some constant c for all L i.e. $y_L = c\mu(L)g(L)$ are the optimal values. Plugging this into our constraint:

$$1 = \sum_{L|\Pi,\ \deg L<\xi} \mu(L)y_L = \sum_{L|\Pi,\ \deg L<\xi} c\mu^2(L)g(L) = c\mathcal{G}(\xi, m)$$

Hence, $y_L^* = \mu(L)g(L)\mathcal{G}(\xi, m)^{-1}$ gives the minimal main term:

$$\min \mathcal{M} = \sum_{\substack{L|\Pi(\mathcal{P};m)\\ \deg L<\xi}} \frac{1}{g(L)}(y_L^*)^2 = \sum_{\substack{L|\Pi(\mathcal{P};m)\\ \deg L<\xi}} g(L)\mathcal{G}(\xi, m)^{-2} = \mathcal{G}(\xi, m)^{-2}\mathcal{G}(\xi, m) = \mathcal{G}(\xi, m)^{-1}$$

Thus, $\mathcal{S}(\mathcal{A}, \mathcal{P}; m) \leq X\mathcal{G}(\xi, m)^{-1} + \sum |\mu^+(D)||R_D|$.

3.2.5 Controlling the Error Term

We have bounded the main term, and now need to check the remainder. Using the inverse change of variables, we tediously obtain the corresponding minimizing λ_D for $\deg D < \xi$ are

$$\lambda_D = \frac{\mu(D)|D|}{\omega(D)} \sum_{\substack{L|\Pi(\mathcal{P};m)\\ \deg L<\xi\\ D|L}} \mu(L)y_L = \frac{\mu(D)|D|}{\omega(D)} \sum_{\substack{L|\Pi(\mathcal{P};m)\\ \deg L<\xi\\ D|L}} \mu(L)^2 g(L)\mathcal{G}(\xi, m)^{-1}$$
$$= \frac{\mu(D)|D|}{\mathcal{G}(\xi, m)\omega(D)} \sum_{\substack{L|\Pi(\mathcal{P};m)\\ \deg L<\xi\\ D|L}} g(L)$$
$$= \frac{\mu(D)|D|}{\mathcal{G}(\xi, m)\omega(D)} \sum_{\substack{DE|\Pi(\mathcal{P};m)\\ \deg DE<\xi}} g(DE)$$
$$= \frac{\mu(D)|D|g(D)}{\mathcal{G}(\xi, m)\omega(D)} \sum_{\substack{DE|\Pi(\mathcal{P};m)\\ \deg E<\xi-\deg D}} g(E)$$

This sum looks almost like $\mathcal{G}(\xi - \deg D, m)$, except that the sum of divisors E is over Π/D instead of Π. Let $\mathcal{G}_D(\xi, m) := \sum_{E|\Pi/D, \deg E < \xi} g(E)$. With this, and recalling the Euler-product form of $g(D)$ continues to expand:

$$\begin{aligned}\lambda_D = \frac{\mu(D)|D|g(D)}{\mathcal{G}(\xi,m)\omega(D)} \sum_{\substack{DE|\Pi(\mathcal{P};m) \\ \deg E < \xi - \deg D}} g(E) &= \frac{\mu(D)|D|g(D)}{\mathcal{G}(\xi,m)\omega(D)} \mathcal{G}_D(\xi - \deg D, m) \\ &= \frac{\mu(D)|D|}{\mathcal{G}(\xi,m)\omega(D)} \mathcal{G}_D(\xi - \deg D, m) \frac{\omega(D)}{|D|} \prod_{\mathfrak{p}|D} \left(1 - \frac{\omega(\mathfrak{p})}{|\mathfrak{p}|}\right)^{-1} \\ &= \mu(D) \frac{\mathcal{G}_D(\xi - \deg D, m)}{\mathcal{G}(\xi, m)} \prod_{\mathfrak{p}|D} \left(1 - \frac{\omega(\mathfrak{p})}{|\mathfrak{p}|}\right)^{-1}\end{aligned}$$

This is messy, but helpful in one way.

Proposition 3.2. *$\mu(D)\mathcal{G}_D(\xi - \deg D, m) \prod_{\mathfrak{p}|D} \left(1 - \frac{\omega(\mathfrak{p})}{|\mathfrak{p}|}\right)^{-1} \leq \mathcal{G}(\xi, m)$, so in particular $|\lambda_D| \leq 1$.*

Proof. Straightforward sum rewriting, grouping elements together by $\gcd(\cdot, D)$:

$$\begin{aligned}\mathcal{G}(\xi,m) = \sum_{\substack{L|\Pi \\ \deg L < \xi}} g(L) = \sum_{K|D} \sum_{\substack{L|\Pi \\ \deg L < \xi \\ (L,D)=K}} g(L) &= \sum_{K|D} \sum_{\substack{KH|\Pi \\ \deg H < \xi - \deg K \\ (H,D)=\mathbf{1}}} g(KH) \\ &= \sum_{K|D} g(K) \sum_{\substack{KH|\Pi \\ \deg H < \xi - \deg K \\ (H,D)=\mathbf{1}}} g(H) \\ &= \sum_{K|D} g(K) \sum_{\substack{\deg H < \xi - \deg K \\ HD|\Pi}} g(H) \\ &\geq \sum_{K|D} g(K) \sum_{\substack{\deg H < \xi - \deg D \\ HD|\Pi}} g(H) \\ &= \sum_{K|D} g(K) \mathcal{G}_D(\xi - \deg D, m)\end{aligned}$$

with the two key steps noting that $(L, D) = K$ implies $(L/K, D) = 1$ as L is squarefree, and $\xi - \deg D \leq \xi - \deg K$ as $K \mid D$ i.e. $\deg K \leq \deg D$. Now, as Π and so D is squarefree, and g is multiplicative, we can do an Euler-product style rewriting:

$$\sum_{K|D} g(K) = \prod_{\mathfrak{p}|D} (1 + g(\mathfrak{p})) = \prod_{\mathfrak{p}|D} \left(1 + \frac{\omega(\mathfrak{p})}{|\mathfrak{p}| - \omega(\mathfrak{p})}\right) = \prod_{\mathfrak{p}|D} \left(\frac{|\mathfrak{p}|}{|\mathfrak{p}| - \omega(\mathfrak{p})}\right) = \prod_{\mathfrak{p}|D} \left(1 - \frac{\omega(\mathfrak{p})}{|\mathfrak{p}|}\right)^{-1} \quad \square$$

We now have all the tools we need to state the most important result of Selberg's sieve. The following allows us to reduce getting sieve estimates to estimating $\mathcal{G}(\xi, m)$—which only depends on ω—and $|R_d|$.

Theorem 3.3 (The Fundamental Theorem of Selberg's Sieve)**.** *Suppose* $0 \le \omega(\mathfrak{p}) < |p| = q^{\deg \mathfrak{p}}$ *for all irreducibles* $\mathfrak{p} \in \mathcal{P}$*. Also, let* $\xi = y/2$*,* $\mathcal{G}(\xi, m)$*, and* $g(L)$ *be as before, and* $v(f) = \sum_{\mathfrak{p}|f} 1$ *be the number of distinct irreducible factors of* $f(t)$*. Then*

$$\mathcal{S}(\mathcal{A}, \mathcal{P}; m) \le X \frac{1}{\mathcal{G}(\xi, m)} + \sum_{\substack{D | \Pi(\mathcal{P}; m) \\ \deg D < y}} 3^{v(D)} |R_D|$$

Proof. The main term is from our result above. Since $|\lambda_D| \le 1$ by Proposition 3.2,

$$|\mu^+(D)| \le \sum_{[D_1, D_2] = D} |\lambda_{D_1}||\lambda_{D_2}| \le \sum_{[D_1, D_2] = D} 1 = 3^{v(D)}$$

where we evaluate the sum combinatorially: since D is squarefree, $[D_1, D_2] = D$ if and only if every irreducible factor of D either 1) divides both D_1 and D_2, 2) only divides D_1 and not D_2, or 3) only divides D_2 and not D_1. There are 3 choices per factor, giving the result. □

3.3 Applications: Irreducibles and Twin Irreducibles

The strength of Theorem 3.3 is that it has done all the hard work in getting optimal bounds for us; all we need to do to apply it is estimate $\mathcal{G}$ by estimating g, which only relies on ω. We provide two original calculations applying Theorem 3.3, inspired by some of the bounding techniques in [15].

Example (Counting Irreducibles of Degree n)**.** With Selberg's sieve, we now improve our calculation from before on counting irreducible polynomials, notably allowing $m = \lfloor n/2 \rfloor$ to get our desired count. Let

- $\mathcal{A} = \{f(t) \in \mathbb{F}_q[t] : f \text{ monic}, \deg f = n\}$
- $\mathcal{P} = \{\mathfrak{p} \in \mathbb{F}_q[t] : \mathfrak{p} \text{ monic, irreducible}\}$
- $m = \lfloor n/2 \rfloor$

So $\forall \mathfrak{p} \in \mathcal{P}$ $\omega(\mathfrak{p}) = 1$, and by the Chinese Remainder Theorem/multiplicativity $\forall D \mid \Pi(\mathcal{P}; m)$ $\omega(D) = 1$. Note that if $\deg D \le n$, the error term disappears: $R_D = 0$ since $\deg D \le n$ so the count is exact as we described before. To make this happen, we let $y \le n$ i.e. $\xi = \lfloor n/2 \rfloor$ (to make the sieve as accurate as possible we make ξ as big as possible). Now we calculate the auxiliary functions.

$$g(\mathfrak{p}) = \frac{\omega(\mathfrak{p})}{|\mathfrak{p}|}\Big(1 - \frac{\omega(\mathfrak{p})}{|\mathfrak{p}|}\Big)^{-1} = \frac{1}{|\mathfrak{p}|}\Big(1 - \frac{1}{|\mathfrak{p}|}\Big)^{-1} = \frac{1}{|\mathfrak{p}| - 1} = \sum_{k=1}^{\infty} |\mathfrak{p}|^{-k}$$

By multiplicativity of g and a divisor $L \mid \Pi$ being squarefree:

$$g(L) = \prod_{\mathfrak{p} | L} \sum_{k=1}^{\infty} |\mathfrak{p}|^{-k} = \sum_{\mathrm{rad}(M) = L} \frac{1}{|M|}$$

where $\mathrm{rad}(M)$ is the product of the distinct irreducible factors of M. Note M is monic as all the irreducible factors $\mathfrak{p}$ are. So

$$\mathcal{G}(\xi, m) = \sum_{\substack{L|\Pi, \\ \deg L<\xi}} g(L) = \sum_{\substack{L|\Pi, \\ \deg L<\xi}} \sum_{\mathrm{rad}(M)=L} \frac{1}{|M|} \geq \sum_{\substack{M \text{ monic} \\ \deg M<\xi}} \frac{1}{|M|} = \sum_{d=0}^{\xi-1} \sum_{\substack{M \text{ monic} \\ \deg M=d}} \frac{1}{q^d} = \sum_{d=0}^{\xi-1} \frac{q^d}{q^d} = \xi$$

Hence, letting $\pi_q(n) = \#\{\text{monic irreducible polynomials } f : \deg f = n\}$ and using $\xi = \lfloor n/2 \rfloor = n/2 + O(1)$,

$$\pi_q(n) = \mathcal{S}(\mathcal{A}, \mathcal{P}; m) \leq \frac{X}{\mathcal{G}(\xi, m)} \leq \frac{q^n}{\xi} = \frac{q^n}{n/2 + O(1)} = \frac{q^n}{\frac{n}{2}(1 + O(1/n))} \leq (2 + o(1))\frac{q^n}{n}$$

Example (Counting Twin Irreducibles of Degree n)**.** Fix a nonzero polynomial K of degree $< n$. We will estimate the number polynomials such that $f, f + K$ are both irreducible. Let

- $\mathcal{A} = \{f(f + K) : f \in \mathbb{F}_q[t] \text{ monic}, \deg f = n\}$
- $\mathcal{P} = \{\mathfrak{p} \in \mathbb{F}_q[t] : \mathfrak{p} \text{ monic, irreducible}\}$
- $m = \lfloor n/2 \rfloor$

Since $\deg(f + K) = n$, it is only reducible if it has a factor of degree $\leq \lfloor n/2 \rfloor$. Hence if $f(f + K)$ does not have any factor of degree $\leq \lfloor n/2 \rfloor$, then both f and $f + K$ must be irreducible. For $\mathfrak{p} \nmid K$, $\omega(\mathfrak{p}) = 2$ as either $f \equiv 0 \bmod \mathfrak{p}$ or $f \equiv -K \bmod \mathfrak{p}$. If $\mathfrak{p} \mid K$, then $\omega(\mathfrak{p}) = 1$ as the divisibility conditions become the same. Similar to before, if $\deg D \leq n$, then $R_D = 0$: say $D \mid f(f + K)$ i.e. $f \equiv \lambda \bmod D$ where λ is one of the $\omega(D)$ residue classes that permits divisibility. Then $f = QD + \lambda$ with $\deg Q = n - \deg D$, and so the number of possible f is equivalent to a choice of the quotient and remainder. Hence $\#\mathcal{A}_d = q^{n-\deg D} \cdot \omega(D) = q^n \omega(D)/|D|$ is exact. Thus, we set $y \leq n$ and so $\xi = \lfloor n/2 \rfloor$ to eliminate the remainder term. Similar to before,

$$g(\mathfrak{p}) = \begin{cases} \frac{2}{|\mathfrak{p}|-2} = \sum_{k=1}^{\infty} (2/|\mathfrak{p}|)^k & \text{if } \mathfrak{p} \nmid K \\ \frac{1}{|\mathfrak{p}|-1} = \sum_{k=1}^{\infty} (1/|\mathfrak{p}|)^k & \text{if } \mathfrak{p} \mid K \end{cases}$$

For $(L, K) = \mathbf{1}$, we have

$$g(L) = \prod_{\mathfrak{p}|L} \sum_{k=1}^{\infty} \frac{2^k}{|\mathfrak{p}|^k} = \sum_{\mathrm{rad}(M)=L} \frac{2^{\Omega(M)}}{|M|}$$

where $\Omega(M)$ is the number of irreducible factors of M with multiplicity. We now lower bound $\mathcal{G}(\xi, m)$ by sieving polynomials coprime to K (i.e. ignoring the irreducibles $\mathfrak{p} \mid K$):

$$\mathcal{G}(\xi, m) = \sum_{\substack{L|\Pi \\ \deg L<\xi}} g(L) \geq \sum_{\substack{L|\Pi \\ \deg L<\xi \\ (L,K)=\mathbf{1}}} g(L) = \sum_{\substack{L|\Pi \\ \deg L<\xi \\ (L,K)=\mathbf{1}}} \sum_{\mathrm{rad}(M)=L} \frac{2^{\Omega(M)}}{|M|} \geq \sum_{\substack{M \text{ monic} \\ \deg M<\xi \\ (M,K)=\mathbf{1}}} \frac{2^{\Omega(M)}}{|M|}$$

The coprimality condition is tedious, so we bound it below with first-order inclusion-exclusion:

$$\begin{aligned}
\mathcal{G}(\xi, m) \geq \sum_{\substack{M \text{ monic} \\ \deg M < \xi \\ (M,K)=1}} \frac{2^{\Omega(M)}}{|M|} &= \sum_{\substack{M \text{ monic} \\ \deg M < \xi}} \frac{2^{\Omega(M)}}{|M|} - \sum_{\substack{M \text{ monic} \\ \deg M < \xi \\ (M,K)\neq 1}} \frac{2^{\Omega(M)}}{|M|} \\
&\geq \sum_{\substack{M \text{ monic} \\ \deg M < \xi}} \frac{2^{\Omega(M)}}{|M|} - \sum_{\mathfrak{p}|K} \sum_{\substack{M \text{ monic} \\ \deg M < \xi \\ \mathfrak{p}|M}} \frac{2^{\Omega(M)}}{|M|} \\
&= \sum_{\substack{M \text{ monic} \\ \deg M < \xi}} \frac{2^{\Omega(M)}}{|M|} - \sum_{\mathfrak{p}|K} \sum_{\substack{N \text{ monic} \\ \deg N < \xi - \deg \mathfrak{p}}} \frac{2^{\Omega(\mathfrak{p}N)}}{|\mathfrak{p}N|} \\
&\geq \sum_{\substack{M \text{ monic} \\ \deg M < \xi}} \frac{2^{\Omega(M)}}{|M|} - \sum_{\mathfrak{p}|K} \frac{2^{\Omega(\mathfrak{p})}}{|\mathfrak{p}|} \sum_{\substack{N \text{ monic} \\ \deg N < \xi}} \frac{2^{\Omega(N)}}{|N|} \\
&= \left(1 - \sum_{\mathfrak{p}|K} \frac{2}{|\mathfrak{p}|}\right) \sum_{\substack{M \text{ monic} \\ \deg M < \xi}} \frac{2^{\Omega(M)}}{|M|}
\end{aligned}$$

noting that $\Omega(\mathfrak{p}) = 1$. Let $c_K = 1 - \sum_{\mathfrak{p}|K} \frac{2}{|\mathfrak{p}|}$. Letting $d(M)$ be the number of divisors of M, we can again bound this sum using $2^x \geq x + 1$. Also, note $\sum_{\deg M = j} d(M) = \#\{(A, B) : AB = M, \deg M = j\} = \sum_{a=0}^{j} q^a q^{j-a} = (j+1)q^j$, giving us the further bound:

$$\mathcal{G}(\xi, m) \geq c_K \sum_{\substack{M \text{ monic} \\ \deg M < \xi}} \frac{2^{\Omega(M)}}{|M|} \geq c_K \sum_{\substack{M \text{ monic} \\ \deg M < \xi}} \frac{d(M)}{|M|} = c_K \sum_{j=0}^{\xi-1} \frac{(j+1)q^j}{q^j} = c_K \frac{\xi(\xi+1)}{2}$$

Thus, we conclude that the number of twin irreducibles $f, f + K$ is bounded above by

$$\mathcal{S}(\mathcal{A}, \mathcal{P}; m) \leq \frac{X}{G(\xi, m)} \leq \frac{q^n}{c_K \xi(\xi+1)/2} \leq \frac{2q^n}{c_k(n/2 + O(1))^2} = (C_K + o(1))\frac{q^n}{n^2}$$

with $C_K = 8/c_K$.

4 The Parity Problem

With nothing but essentially inclusion-exclusion, we obtained some very promising results, especially knowing the true counts we want.

Theorem 4.1 (Prime Number Theorem for Polynomials)**.** $\pi_q(n) = \frac{1}{n} \sum_{d|n} \mu(n/d) q^d \sim q^n/n$.

Proof. We follow [11]. Let $\zeta_q(s) = \sum_{f \in \mathbb{F}_q[t] \text{ monic}} 1/|f|^s$ be the zeta function of $\mathbb{F}_q[t]$. There are q^d monic polynomials of degree d, so $\zeta_q(s) = \sum_d q^d/q^{ds} = 1/(1 - q^{1-s})$ for all complex

numbers with $\Re(s) > 1$. By unique factorization, $\zeta_q(s) = \prod_{\mathfrak{p}}(1 - |\mathfrak{p}|^{-s})^{-1}$. Letting $Z = q^{-s}$ and equating the two formulations,

$$\frac{1}{1-qZ} = \prod_{\mathfrak{p} \text{ irreducible}} (1 - |\mathfrak{p}|^{-s})^{-1} = \prod_{d=1}^{\infty} (1 - Z^d)^{-\pi_q(d)}$$

where we group the irreducibles $\mathfrak{p}$ by degree. Taking the logarithmic derivative with respect to Z of both sides and multiplying through by Z yields $\frac{qZ}{1-qZ} = \sum_{d=1}^{\infty} \frac{d\pi_q(d)Z^d}{1-Z^d}$. Expanding both sides into power series,

$$\sum_{d=1}^{\infty} q^n Z^n = \sum_{d=1}^{\infty} d\pi_q(d) \sum_{k=1}^{\infty} Z^{kd}$$

Comparing and equating the coefficient of Z^n on either side, we get $q^n = \sum_{d|n} d\pi_q(d)$. Möbius inversion gives $n\pi_q(n) = \sum_{d|n} \mu(n/d)q^d$. □

Remark. This exact formula is a consequence of $\zeta_q(s) = (1 - q^{1-s})^{-1}$ having no zeros. In $\mathbb{Z}$, the Prime Number Theorem is proved via the explicit formula, which expresses prime-counting sums up to X as a contour integral of the logarithmic derivative ζ'/ζ. By the Residue Theorem, each zero ρ of ζ contributes an error term of size $X^{\Re(\rho)}$. In $\mathbb{F}_q[t]$, no zeros for $\zeta_q(s)$ means the power series $\sum_{d=1}^{\infty} \frac{d\pi_q(d)Z^d}{1-Z^d} = qZ/(1-qZ)$ has no correction terms, i.e. the coefficient extraction is exact. For more on the case of $\mathbb{Z}$, see *C3.8 Analytic Number Theory* [4]. This is one of the key analytic differences between $\mathbb{Z}$ and $\mathbb{F}_q[t]$.

Remark. The above can also be proved with an algebraic argument, using inclusion-exclusion counting on the number of elements not contained in maximal subfields of finite fields [2].

So our current application of Selberg's sieve is only off by a factor of 2 of the true count $\pi_q(n)$. There are thus two improvements we would seek from our proof: 1) reduce the leading constant 2 to 1, and 2) find a corresponding lower bound. We did, after all, rely on an arbitrary choice of working with a quadratic formulation of μ^+, and gave bounds on g that are not obviously tight.

Unfortunately, there is a reason why we did not present possible solutions to either, since it is *impossible* to achieve either improvement with sieve methods alone—for Selberg's sieve and in general! The problem lies within the simplicity of sieve methods for while they lend themselves well to a wide array of problems, they do not have access to enough information to strengthen their results.

Recall the original sifted set quantity we seek to estimate is $\mathcal{S}(\mathcal{A}, \mathcal{P}; m) = \sum_{d|\Pi} \mu(d)\#\mathcal{A}_d$. Since this sum is only over divisors of $\Pi(\mathcal{P}; m)$ and hence only has terms including irreducible factors with degree $< m$, it is blind to the divisibility patterns of irreducibles of degree $\geq m$. Consider two polynomials $f_1 = p_1 \dots p_k q_1$ and $f_2 = p_1 \dots p_k q_1 q_2$ where p_i, q_i are irreducible polynomials with $\deg p_i < m$ and $\deg q_i \geq m$. So $(f_1, \Pi) = (f_2, \Pi)$ and thus $f_1 \in \mathcal{A}_d \iff f_2 \in \mathcal{A}_d$ for all $d \mid \Pi$. So they contribute identical counts and thus to a sieve, f_1 and f_2 are functionally identical; if $f_1 \in \mathcal{A}$ was in a set we are sifting from, and we replace it $\mathcal{A}' = \mathcal{A} \cup \{f_2\} \setminus \{f_1\}$ with f_2, then the counts are the same $\mathcal{S}(\mathcal{A}, \mathcal{P}; m) = \mathcal{S}(\mathcal{A}', \mathcal{P}; m)$.

But f_1 and f_2 clearly are not the same! They have different cofactors $q_1 \neq q_1 q_2$, and most importantly for us, their parity differ $\Omega(f_1) \neq \Omega(f_2)$. Since the sieve only parses the divisibility data of the irreducible factors up to degree m, extra information can hide in those larger prime factors. This is the basis for **the parity problem of sieve methods**.

4.1 The Problem Formally

We demonstrate the parity problem more concretely now, extending the pointwise element argument to whole sets, constructing two different sets $\mathcal{A}^+, \mathcal{A}^-$ with the same divisibility data but different sifted counts. The idea is that the sieve must treat these sieves identically and hence produce the same output, even though what can be sifted from them is wildly different. In particular, suppose we have functions μ^+, μ^- that attempt to bound our sifted set:

$$\sum_{D|\Pi} \mu^-(D) \#\mathcal{A}_D \leq \mathcal{S}(\mathcal{A}, \mathcal{P}; m) \leq \sum_{D|\Pi} \mu^+(D) \#\mathcal{A}_D$$

Assume μ^+ and μ^- are both supported on a subset of all the divisors i.e. $\mu^{\pm}(D) = 0$ for $\deg D > y$ like we had in our Selberg sieve above; any reasonable sieve will need to truncate to some level to keep the error of approximations (in the remainder term $\mathcal{R}$) small. We will show there are fundamental limits to how tight of an upper and lower bound we can obtain. As before, we work with monic polynomials over a fixed finite field $\mathbb{F}_q[t]$. For clarity, we focus specifically on the problem of counting degree n irreducible polynomials $\pi_q(n)$ where we sift from $\mathcal{A} = \{f(t) \in \mathbb{F}_q[t] : f \text{ monic}, \deg f = n\}$ (see Example 3.3).

Proposition 4.2. $\left| \sum_{\substack{\deg f = n \\ f \in \mathbb{F}_q[t]}} \lambda(f) \right| \leq q^{(n+1)/2}$.

Proof. We follow the similar strategy as in $\mathbb{Z}$. Let $L_q(n) = \sum_{\deg f = n} \lambda(f)$ be the sum we want to bound, and consider the generating function $\sum_{n=0}^{\infty} L_q(n) Z^n$. Expanding $L_q(n)$ and by the multiplicativity of λ, we obtain the following chain of equalities of formal power series:

$$\sum_{n=0}^{\infty} L_q(n) Z^n = \sum_{f \in \mathbb{F}_q[t]} \lambda(f) Z^{\deg f} = \prod_{\mathfrak{p}} \sum_{k=0}^{\infty} \lambda(\mathfrak{p}^k) Z^{\deg \mathfrak{p}^k} = \prod_{\mathfrak{p}} \sum_{k=0}^{\infty} (-1)^k Z^{k \deg \mathfrak{p}} = \prod_{\mathfrak{p}} \frac{1}{1 + Z^{\deg \mathfrak{p}}}$$

From the Euler product form of the zeta function above, we have $\prod_{\mathfrak{p}} (1 - Z^{\deg \mathfrak{p}}) = 1/\zeta_q(s) = 1 - qZ$ with the change of variables $Z = q^{-s}$. Similarly for $\prod_{\mathfrak{p}} (1 - Z^{2 \deg \mathfrak{p}}) = 1/\zeta_q(2s) = 1 - qZ^2$:

$$\begin{aligned} \sum_{n=0}^{\infty} L_q(n) Z^n = \prod_{\mathfrak{p}} \frac{1}{1 + Z^{\deg \mathfrak{p}}} = \prod_{\mathfrak{p}} \frac{1 - Z^{\deg \mathfrak{p}}}{1 - Z^{2 \deg \mathfrak{p}}} = \frac{1 - qZ}{1 - qZ^2} &= (1 - qZ) \sum_{k=0}^{\infty} q^k Z^{2k} \\ &= \sum_{k=0}^{\infty} q^k Z^{2k} - q^{k+1} Z^{2k+1} \end{aligned}$$

Reading off the coefficients gives $|L_q(n)| \leq q^{(n+1)/2}$. □

Let $\mathcal{A}^{\pm} = \{f \in \mathbb{F}_q[t] : \deg f = n, \lambda(f) = \pm 1\}$ be the sets of polynomials with an odd/even parity number of irreducible factors with multiplicity. The above tells us that the number of polynomials split between $\mathcal{A}^+$ and $\mathcal{A}^-$ are about equal: note $\#\mathcal{A}^+ + \#\mathcal{A}^- = q^n$ and $\#\mathcal{A}^+ - \#\mathcal{A}^- = L_q(n)$, so $\#\mathcal{A}^{\pm} = (q^n \pm L_q(n))/2 = (q^n \pm O(q^{(n+1)/2}))/2$. By multiplicativity of λ, we also obtain $\#\mathcal{A}_D^{\pm} = (\#\mathcal{A}_D \pm \lambda(D)L_q(n - \deg D))/2 = (q^{n-\deg D} \pm \lambda(D)O(q^{(n-\deg D+1)/2}))/2$. These look similar, and since $L_q(n)$ is relatively small, we have

$$\frac{|\#\mathcal{A}_D^+ - \#\mathcal{A}_D^-|}{\#\mathcal{A}_D^{\pm}} \leq \frac{q^{(n-\deg D+1)/2}}{\frac{1}{2}q^{n-\deg D} + O(q^{(n-\deg D+1)/2})} = O(q^{-(n-\deg D-1)/2})$$

This bound is monotonic in $\deg D$. Recall $\mu^{\pm}$ are supported for $\deg D \leq y$ to truncate the inclusion-exclusion counting. So long as $n - y(n) \to \infty$, then

$$\sup_{\deg D \leq y} \frac{|\#\mathcal{A}_D^+ - \#\mathcal{A}_D^-|}{\#\mathcal{A}_D^{\pm}} \leq O(q^{-(n-y-1)/2}) = o(1)$$

uniformly over $\deg D$. Hence $\#\mathcal{A}_D^- = (1 + o(1))\#\mathcal{A}_D^+$ for all $D \mid \Pi(\mathcal{P}; m)$ with $\deg D \leq y$.

Now the problem starts to realize itself. Since $\#\mathcal{A}_D^+ \approx \#\mathcal{A}_D^-$ on the support of $\mu^{\pm}$, they both must induce the same upper and lower bounds for the sifted sets $\mathcal{S}(\mathcal{A}^{\pm}, \mathcal{P}; m)$. But clearly $\mathcal{S}(\mathcal{A}^-; \mathcal{P}; m)$ contains all the irreducibles (with suitably chosen m), and $\mathcal{S}(\mathcal{A}^+; \mathcal{P}; m)$ only contains even parity elements and hence no irreducibles! This manifests in a strict over- and underestimation of $\mathcal{S}(\mathcal{A}, \mathcal{P}; m)$. Noting that $\mathcal{A}_D = \mathcal{A}_D^+ \cup \mathcal{A}_D^-$, we obtain the following bound on upper bounds of sieves:

$$\sum_{d \mid \Pi} \mu^+(d)\#\mathcal{A}_d = (2 + o(1)) \sum_{d \mid \Pi} \mu^+(d)\#\mathcal{A}_d^- \geq (2 + o(1))\mathcal{S}(\mathcal{A}^-, \mathcal{P}; n/3) = (2 + o(1))\frac{q^n}{n}$$

(note the choice of $m = n/3$: since $\mathcal{A}^-$ only contains polynomials of odd parity, the choice of $n/3$ prevents polynomials that have more than 3 irreducible factors to survive the sieve i.e. only irreducibles survive). So for *any* upper bound on $\mathcal{S}(\mathcal{A}, \mathcal{P}; m)$, it must be at least twice as big as the actual sieve count. Our result obtained with the Selberg sieve, while it does not look ideal, is actually the best we can do. On the other hand, even more dishearteningly, non-trivial lower bounds are hopeless:

$$\sum_{d \mid \Pi} \mu^-(d)\#\mathcal{A}_d = (2 + o(1)) \sum_{d \mid \Pi} \mu^-(d)\#\mathcal{A}_d^+ \leq (2 + o(1))\mathcal{S}(\mathcal{A}^+, \mathcal{P}; n/2) = 0$$

since $\mathcal{A}^+$ contains *no* irreducibles i.e. $\mathcal{S}(\mathcal{A}^+, \mathcal{P}; n/2) = 0$. Hence any lower bound on $\mathcal{S}(\mathcal{A}, \mathcal{P}; m)$ must be trivial. For another perspective on this derivation via linear programming, see [8].

This is the parity problem. Since the divisibility data of $\mathcal{A}^+$ and $\mathcal{A}^-$ are identical, without additional supplemented information, the sieve treats them identically, forcing results about sets of specific parity in the full set $\mathcal{A} = \mathcal{A}^+ \cup \mathcal{A}^-$ to overcorrect. Note that lowering the sifting level to include more elements can fix sieve arguments; if we instead sifted at a level, say $m = n/3 < n/2$ for the lower bound argument, then $\mathcal{S}(\mathcal{A}^+, \mathcal{P}; m) > 0$ as it contains semi-irreducible elements with two factors $f = \mathfrak{p}_1\mathfrak{p}_2$ with allowably small degrees; a lower sifting

level that enforces no specific parity so the issue does not arise, at the cost of answering a weaker question. As Tao [13] summarizes: no sieve can achieve non-trivial lower bounds for sets of fixed parity, and any upper bound is loose by at least a factor of 2.

While we applied this specifically to counting irreducibles since we can determine the true value of $\pi_q(n)$ independently of sieve methods and concretely compare this off-by-factor-of-2 problem, the parity problem applies to estimating the size of a sifted set with a specific parity in mind. In the case of twin irreducibles, our goal was to sift from $\mathcal{A} = \{f(f+K) : \deg f = n\}$ to obtain the number of irreducible pairs encoded by the product $\mathfrak{p}(\mathfrak{p}+K)$. These elements all have even parity, and hence suffer the same problem. The impossibility of non-trivial lower bounds is what truly sets apart the parity problem from other limitations, as it instantly kills any hope of proving any conjectures about existence; while sieves alone can show only so many elements survive the process, they cannot guarantee *something* will survive.

Remark. It is worth mentioning that this is a problem of estimating and bounding $\mathcal{S}(\mathcal{A}, \mathcal{P}; m)$. This is not a problem with sieves themselves, in the sense that the full inclusion-exclusion count $\mathcal{S}(\mathcal{A}, \mathcal{P}; m) = \sum_{D|\Pi} \mu(D) \#\mathcal{A}_D$ is correct and agrees with the truth. For small support $\deg D \leq y$, we get the parity-confused data $\#\mathcal{A}_D^+ \approx \#\mathcal{A}_D^-$, but the divisors D with large $\deg D$ in the full count allow for them to be differentiated and the sieve to be accurate again. $\mathcal{S}(\mathcal{A}, \mathcal{P}; m) = \mathcal{S}(\mathcal{A}^+, \mathcal{P}; m) + \mathcal{S}(\mathcal{A}^-, \mathcal{P}; m)$ and the sieve agrees with the set of the parity it targets.

Remark. The full strength of parity relied on our sum $\sum_{\deg f = n} \lambda(f)$ being small. In cases (i.e. other problems or rings) where it is not, then the sieve data $\#\mathcal{A}_D^+, \#\mathcal{A}_D^-$ is necessarily different and the upper bound factor of 2 may be reduced.

Remark. The phenomenon is specific to $\Omega(f) \bmod 2$: $\sum_{\deg f = n} z^{\Omega(f)}$ has leading coefficient proportional to $1/\Gamma(z)$, which vanishes only at the root of unity $\xi_2 = -1$, where Γ has a pole. Partitions of $\Omega(f)$ modulo $p \geq 3$ remain distinguishable since the series is not small at the corresponding roots of unity; see [14, Thm. II.5.3], [7] for $\mathbb{Z}$.

Remark. While we do not show it here, to actually show that parity truly is a problem for the twin irreducible problem, one would have to show $\sum_{\deg f = n, D | f(f+K)} \lambda(f) = o(\#\mathcal{A}_D^{\pm})$ is small for each $D \mid \Pi$ in the support of the sieve. This requires a bit more work since $D \mid f(f+K) \not\Rightarrow D \mid f$ and so we cannot leverage multiplicativity like we did above.

4.2 How Can We Overcome Parity?

If parity is such a fundamental limitation for sieve methods, they seem almost useless to prove anything. That is mostly true in the case one tries to *only* uses sieve techniques to prove something. What extra information can we give sieves to strengthen them and overcome the parity problem? What additional facts and hypotheses could we need to leverage to overcome this barrier?

It is worth noting what *does not* improve the parity problem. The barrier is robust against the obvious adjustments: the argument made no assumption on $\mu^{\pm}$, on the truncation level y, or on the choice of sieve, so none of these can break it. Increasing y only saturates the

bound at the parity ceiling, and in our examples $y \leq n$ is already maximal. The obstruction is structural.

One route is to avoid parity altogether. Naively, we could in some way change our sieve formulations to adjust for the parity problem. For example, in the case of estimating $\pi_q(n)$, instead of sieving on $\mathcal{A}$, we could sieve directly on $\mathcal{A}^-$ with our knowledge that $\#\mathcal{A}^- \approx \#\mathcal{A}/2$. This would give the correct leading coefficient of 1 instead of 2 in the bound $\pi_q(n) \leq (1+o(1))q^n/n$. Note we have not actually bypassed the parity problem: it is still overshooting by a factor of 2 for what the sieve expects, but because we are sifting on a set half the size, we externally can recognize that this is adjusts for the truth. Alternatively, we can relax our objective and allow for looser objects that are not parity restricted, like looking at *almost irreducibles* $\mathbb{P}_2$ that have at most 2 irreducible factors, but this counts a different set than we originally wanted.

To actually "overcome parity", we need to leverage some additional arithmetic facts with new context beyond typical sieve techniques. This can come in the form of an understanding of a generalization of sieve-like inputs, or something new all together. Sieves have proven only so useful, and we now need something else; there could be a way to use sieve methods to get the stronger asymptotics beyond parity, but as it stands and the theory suggests, it is unlikely.

4.2.1 Generalizing Sieve Information: Type I vs. Type II Sums

In a spirit of returning to the combinatorial roots of sieves with inclusion-exclusion counting, we look at a generalization of the arithmetic facts that sieves leverage. Using the vocabulary of [5] and [3], these come in the form of **Type I** and **Type II** information.

Type I information is, roughly speaking, *linear information* about individual divisibility conditions. Type I information is that which concerns $\#\mathcal{A}_D$ individually.

Definition 4.3 (Type I Information)**.** A *Type I sum* is one of the form $\sum_{\deg D \leq y} \alpha_D \#\mathcal{A}_D$ for $|\alpha_D| \leq 1$. A *Type I estimate or information at level of y for $\mathcal{A}$* is an estimate of this sum that applies uniformly in the bounded weights α_D.

These are quantitative descriptions of how the elements of $\mathcal{A}$ distribute among the divisibility classes. That is, a Type I sum is an instance of a sum that is structurally identical to what a sieve actually computes with inclusion-exclusion. In the Selberg sieve, we had $\alpha_D = \sum_{[D_1,D_2]=D} \lambda_{D_1}\lambda_{D_2}$ (which can be renormalized to be less than 1; bounds by sieves are only up to an $O(1)$ factor anyway). Estimating $\#\mathcal{A}_D = \frac{\omega(D)}{|D|}X + R_D$ as before, we can decompose a Type I sum as

$$\sum_{\deg D \leq y} \alpha_D \#\mathcal{A}_D = X \sum_{\deg D \leq y} \alpha_D \frac{\omega(D)}{|D|} + \sum_{\deg D \leq y} \alpha_D R_D$$

A small remainder means the main term is an accurate asymptotic for the Type I sum, so we thus seek bounds on $\left|\sum_{\deg D \leq y} \alpha_D R_D\right| \leq \sum_{\deg D \leq y} |R_D|$ (note since $|\alpha_D| \leq 1$, this bound is tight taking $\alpha_D = \operatorname{sgn} R_D$).

Example. Recall for $\mathcal{A} = \{f \in \mathbb{F}_q[t] : \deg f = n\}$ and for D with $\deg D \leq n$, we had an exact equality that $\#\mathcal{A}_D = q^{n-\deg D}$ with no remainder term $R_D = 0$, and so $\mathcal{A}$ has Type I information up to level n with $\sum_{\deg D \leq n} \alpha_D \#\mathcal{A}_D = q^n \sum_{\deg D \leq n} \alpha_D / q^{\deg D}$.

In the case of $\mathbb{Z}$, one of the most common pieces of Type I information used is the *Bombieri–Vinogradov theorem* and its description of the level of distribution up to $\theta = 1/2$.

Theorem 4.4 (Bombieri–Vinogradov)**.** *For $x \geq 2$, $q \in \mathbb{N}$, and $(a, q) = 1$, let $\pi(x; q, a) := \#\{p \leq x : p \equiv a \pmod q\}$. Then for every $A > 0$ there exists $B = B(A) > 0$ such that*

$$\sum_{q \leq Q} \max_{y \leq x} \max_{(a,q)=1} \left| \pi(y; q, a) - \frac{\mathrm{Li}(y)}{\varphi(q)} \right| \ll_A \frac{x}{(\log x)^A},$$

uniformly for $Q \leq x^{1/2}(\log x)^{-B}$, where φ denotes Euler's totient function.

Let $\mathcal{A} = \{f(p) : p \leq x \text{ prime}\}$ for some polynomial $f \in \mathbb{Z}[x]$. Then $d \mid f(p) \iff p \bmod d \in \rho(d) = \{a \bmod d : f(a) = 0 \bmod d\}$, and so $\#\mathcal{A}_d = \sum_{a \in \rho(d)} \pi(x; d, a)$. Bombieri–Vinogradov tells us exactly when we can sieve on these types of sets.

Remark. It is worth briefly mentioning two concepts that makes the Bombieri–Vinogradov theorem so widely used, and why the remainder term being $O(x/(\log x)^A)$ for all A is good enough for our purposes. We introduce them in the context of $\mathbb{Z}$ for its later relevance.

Definition 4.5 (Dimension)**.** Let $(\mathcal{A}, \mathcal{P})$ be parameters for a sieve problem on $\mathbb{Z}$ with the multipilcative density function $\omega(d)$. If

$$\sum_{w \leq p < z} \frac{\omega(p) \log p}{p} = \kappa \log\left(\frac{z}{w}\right) + O(1)$$

for $2 \leq w \leq z$, we say κ is the *dimension of the sieve problem.*

κ is essentially the average value of ω over the primes. By partial summation, we obtain $\sum_{p \leq x} \omega(p)/p = \kappa \log\log x + O(1)$, and so with the Sieve of Eratosthenes-Legendre

$$S(\mathcal{A}, \mathcal{P}; z) \leq X \prod_{\substack{p \in \mathcal{P} \\ p < z}} \left(1 - \frac{\omega(p)}{p}\right) + \sum_{d \mid \Pi} |R_d| \asymp \frac{X}{(\log z)^\kappa} + \sum_{\substack{p \in \mathcal{P} \\ p < z}} |R_d|$$

So up to making the remainder small, $X/(\log z)^\kappa$ is expected order of magnitude (note this agrees with the bounds we obtained with the Selberg sieve for counting irreducibles and twin irreducibles). This is the basis and intuition for the *Fundamental Lemma of Sieve Theory* which makes these types of estimates precise. This is particularly relevant since upper bound constants depend and tend to worsen the higher the dimension [8] [3].

This also motivates the complementary notion of the *level of distribution.*

Definition 4.6 (Level of distribution)**.** Given a sieve problem $(\mathcal{A}, \mathcal{P})$ with $|\mathcal{A}| = X$ and remainder terms $R_d = \#\mathcal{A}_d - X\frac{\omega(d)}{d}$, we say the sieve problem has *level of distribution* θ if for every $A > 0$,

$$\sum_{d \leq X^{\theta-\varepsilon}} |R_d| \ll_A \frac{X}{(\log X)^A}$$

for every $\varepsilon > 0$ [3]. More precisely, following [8], we call θ an *admissible level of distribution*, since the optimal value of θ may not be known.

This in tandem with the Fundamental Lemma gives a nice characterization of when we can ignore remainder terms: sift within the level of distribution.

Example. In our polynomial examples above counting irreducibles and twin irreducibles, we achieved maximal level of distribution $\theta = 1$ since $R_D = 0$ for all D with $\deg D \leq n$ i.e. $\sum_{\deg D \leq n\theta} R_D = 0$ (note: although the level of distribution index looks different, it is the same writing the condition as $|D| \leq q^{\theta n}$).

Example. The Bombieri–Vinogradov theorem says we can achieve level of distribution $\theta = 1/2$ for sets built from primes like $\{f(p) : p \leq x \text{ prime}\}$ and $f \in \mathbb{Z}[x]$ a polynomial.

Type I information, already as we have seen, is essentially useless for our purposes of improving parity. After all, the whole point of parity is that it approximately evenly distributes among residue classes—something that Type I information tries to address. What we showed above in our original demonstration of parity was that the $\lambda(f)$ is evenly distributed among the residue classes and thus could hide in these sums. Type I information is limited in that it only looks at individual divisors, one at a time, and as we have seen the divisibility patterns of elements that have odd or even parity can have the same divisors below a given level. So even though we achieve maximal level of distribution $\theta = 1$ in $\mathbb{F}_q[t]$—the strongest Type I information—we still are unable to bypass parity.

What we need instead is an analysis of not just divisibility, but factorization as a whole; if we can look at how an element $f = gh$ factors, we can more effectively determine whether it is truly an irreducible or not. Type II information addresses this, probing elements that might factor nontrivially. Roughly speaking, Type II information is *bilinear information* that gives the ability to control sums of these almost irreducibles.

Definition 4.7 (Type II Information)**.** A *Type II sum* is one of the form

$$\sum_{\deg D \leq y_1} \sum_{\deg E \leq y_2} \alpha_D \beta_E \#\mathcal{A}_{DE}$$

for $|\alpha_D|, |\beta_D| \leq 1$. A *Type II estimate or information on* $[y_1, y_2]$ *for* $\mathcal{A}$ is an estimate of this sum that applies uniformly in the bounded weights α_D, β_E.

These types of sums explicitly probe factorizations DE, and so can be made sensitive to nontrivial factorizations; Type I information tells us how elements distribute among divisibility classes, while Type II information tells us how elements distribute among factorizations.

Creative decompositions, like that provided with Vaughan's identity [12], allow Type II estimates to give stronger results that sieves could not.

Remark. Type II information is only genuinely new beyond Type I when the ranges $[y_1, y_2]$ are chosen so that $y_1 + y_2$ (the maximum degree of DE) exceeds the level at which Type I information is available on $\mathcal{A}$. If instead $y_1 + y_2 \le y$ for some Type I level y, then the bilinear sum collapses to a linear combination of counts $\#\mathcal{A}_F$ with $\deg F \le y_1 + y_2$ via the rearrangement

$$\sum_{\substack{\deg D \le y_1 \\ \deg E \le y_2}} \alpha_D \beta_E \, \#\mathcal{A}_{DE} \;=\; \sum_{\deg F \le y_1 + y_2} \Bigg(\sum_{\substack{DE = F \\ \deg D \le y_1,\, \deg E \le y_2}} \alpha_D \beta_E \Bigg) \#\mathcal{A}_F$$

with the inner weight bounded (up to a divisor factor), reducing to Type I information at level $y_1 + y_2$. Throughout, Type II ranges are understood to be calibrated so that this reduction is unavailable, i.e., so that $y_1 + y_2$ exceeds the operative Type I level on $\mathcal{A}$.

Example (Vaughan's identity)**.** Here is a useful combinatorial identity that highlights the utility of Type II estimates [3]. Let $y, z \ge 1$, and Λ the von Mangoldt function. Then for any $n > z$,

$$\Lambda(n) = \sum_{\substack{b \mid n \\ b \le y}} \mu(b) \log \frac{n}{b} - \sum_{\substack{bc \mid n \\ b \le y,\, c \le z}} \mu(b)\Lambda(c) \;+\; \sum_{\substack{bc \mid n \\ b > y,\, c > z}} \mu(b)\Lambda(c)$$

(and a similar formula exists for μ with a similar derivation using $\mu * \mathbf{1} = \delta$). Summing $\Lambda(n)$ over $n \in \mathcal{A}$ gives

$$\sum_{n \in \mathcal{A}} \Lambda(n) = \underbrace{\sum_{b \le y} \mu(b) \sum_{\substack{n \in \mathcal{A} \\ b \mid n}} \log \tfrac{n}{b}}_{\text{Type I}} \underbrace{- \sum_{\substack{b \le y \\ c \le z}} \mu(b)\Lambda(c) \#\mathcal{A}_{bc} + \sum_{\substack{b > y \\ c > z}} \mu(b)\Lambda(c) \#\mathcal{A}_{bc}}_{\text{Type II}}$$

exhibiting $\sum_{n \in \mathcal{A}} \Lambda(n)$ as a Type I functional at level y (with a logarithmic weight) plus two Type II functionals. With both ingredients controlled, the sum is evaluated asymptotically rather than bounded above by a sieve majorant—a reduction Vaughan's identity and its generalizations make available for many problems.

4.2.2 Other Ingredients?

Type II is not the only route. In $\mathbb{F}_q[t]$, geometric inputs are also available; sieves' analytic intent do not preclude the success of other methods which we will see below.

5 The State of the Twin Prime Conjecture

The twin-prime question for $\mathbb{F}_q[t]$—are there infinitely many irreducible pairs $f, f + K$?—is fully resolved, while its $\mathbb{Z}$-counterpart remains open. The closest available integer result is *Chen's theorem*, producing infinitely many prime-semiprime pairs; the gap between Chen's

theorem and the twin prime conjecture is precisely the parity problem. We now sketch the proof strategies for $\mathbb{Z}$ and $\mathbb{F}_q[t]$, focusing on which sieve-external inputs do or do not break parity. The arguments below draw on sieve constructions beyond Selberg's; we give them in outline rather than in full.

5.1 The Case of $\mathbb{Z}$: Chen's Theorem

Unfortunately, the twin prime conjecture is still undecided for $\mathbb{Z}$. The best we have thus far is *Chen's theorem* that says there are at least infinitely many prime-semiprime pairs.

Theorem 5.1 (Chen). *There are infinitely many primes p such that $\Omega(p+2) \leq 2$.* [1]

There are three critical details to Chen's proof; see [10] [1] [8] [3] for more details. We briefly sketch out the broad strokes in the proof. The first is directly sifting on primes $\mathcal{A} = \{p+2 : p \leq N \text{ prime}\}$ and leveraging the Bombieri–Vinogradov theorem. This drops the dimension of the sieve problem from 2 to 1, and more importantly fixes one of the numbers in our pair p to be prime; statements about the factors of the product of $n(n+2)$ in our original suggestion cannot be isolated to either term.

Secondly, Chen forms the weighted sum

$$S_0 := S(\mathcal{A}, \mathcal{P}; z) - \frac{1}{2} \sum_{z \leq p < N^{1/3}} S(\mathcal{A}_p, \mathcal{P}; z) - \frac{1}{2} \sum_{\substack{n \in \mathcal{A} \\ (n, \Pi(\mathcal{P}, z))=1}} a(n)$$

where $a(n) = 1$ if $n = p_1 p_2 p_3$ is the product of 3 primes such that $p_1 < N^{1/3} \leq p_2 \leq p_3$, and else $a(n) = 0$. The decomposition is inspired to analyze the factorization of each element n, and assign a non-negative weight to n when $\Omega(n) \leq 2$ and assign a non-positive weight otherwise. So the positivity of S_0 implies $\#\{n \in \mathcal{A} : \Omega(n) \leq 2\}$, and suffices to lower bound it, which translates to lower bounding $\mathcal{S}(\mathcal{A}, \mathcal{P}; z)$ and *upper bounding* the latter terms. Upper bounding sieve counts are comparatively easy as we did before, but it might seem weird that we are asking to lower bound the sifted set $\mathcal{S}(\mathcal{A}, \mathcal{P}; z)$ already, but the upshot is that Chen was able to use a much looser and more forgiving choice of small $z = N^{1/10}$ and we instead use this weighted sieve to refine the result, subtracting out counts that would have more than 2 factors; this is a roundabout way to avoid directly dealing with sieves.

The final key is Chen's "switching" or "reversal of roles" trick to bound the final sum involving $a(n)$. In particular, the final sum

$$\sum_{\substack{n \in \mathcal{A} \\ (n, \Pi(\mathcal{P}, z))=1}} a(n) = \#\{p+2 : p+2 = p_1 p_2 p_3, z \leq p_1 < N^{1/3} \leq p_2 \leq p_3\}$$

counts the number of $p+2$ as a product of 3 primes in the given range. This is a bit of a strange set to directly sieve on, as typical sieve machinery cannot filter for factorizations;

[1] The original source of Chen's theorem is hard to find, but there are many discussions surrounding the ideas, especially since his proof technique works identically for proving a similarly weaker version of Goldbach's conjecture.

Type I information cannot alone have us cleanly distinguish this specific structure. However, since we do know the exact structure of the elements it counts, we can exploit this to sieve on a new set that gives the same bounds. Consider instead the set

$$\mathcal{B} = \{p_1p_2p_3 - 2 : z \leq p_1 < N^{1/3} \leq p_2 \leq p_3,\ p_1p_2p_3 \leq N + 2\}$$

Now we can sift on this instead and ask the more familiar question how many elements of $\mathcal{B}$ are prime? With an adapted version of Bombieri–Vinogradov, we can get similar upper bound estimates on $\mathcal{S}(\mathcal{B}, \mathcal{P}, N^{1/2})$ to upper bound the final sum. Of course, the parity problem remains operative in each application of a sieve bound and the smaller sifting level needed.

While not proven, there are some known reductions of the twin prime conjecture. The following is a kind of Type II generalization of the Bombieri–Vinogradov theorem that has been shown if true implies the twin prime conjecture, further reinforcing that sieve inputs are too simple for their goals [10]. Especially in the case of an object $n(n+2)$ that cleanly separate in multiplicative functions, bilinear data seems to be the needed one.

Shifted Möbius Elliott–Halberstam Conjecture $\mathrm{EH}_{\mu_h}(x^\theta)$**:** For any $A > 0$, we have

$$\sum_{q \leq x^\theta} \max_{y \leq x} \max_{(a,q)=1} \left| \sum_{\substack{n \leq y \\ n \equiv a \,(\mathrm{mod}\, q)}} \Lambda(n)\mu(n+h) - \frac{1}{\phi(q)} \sum_{n \leq y} \Lambda(n)\mu(n+h) \right| \ll_A \frac{x}{(\log x)^A}.$$

While not obvious, via Vaughan's identity, this is equivalent to a kind of Type II content in bounding a bilinear sum of Λ and μ. We will see successfully bounding these types of shifted, bilinear (and more general multilinear) sums in μ is what ultimately closes the twin prime conjecture in $\mathbb{F}_q[t]$.

5.2 The Case of $\mathbb{F}_q[t]$

For $\mathbb{Z}$, we had a single parameter to consider in trying to show there are infinitely many primes with taking $n \to \infty$. For the polynomial case, we have two: the choice of ambient ring $\mathbb{F}_q[t]$ and the degree of polynomial we seek to find twin irreducible pairs. As such, we have two distinct results to address both.

5.2.1 Pollack: Twin Irreducibles of Fixed Degree n

Pollack proved the first of a twin prime analogue for polynomials [9].

Theorem 5.2 (Pollack)**.** *Let $n \geq 2$ and let $0 \leq k < n$. Let $M \in \mathbb{F}_q[T]$ have degree k, let p denote the least prime divisor of n, and let $R(n; M, q)$ be the number of polynomials $\mathfrak{p} \in \mathbb{F}_q[T]$ of degree n (not necessarily monic) such that both $\mathfrak{p}$ and $\mathfrak{p} + M$ are irreducible. Then*

$$-\frac{q^n}{n} - 4\,\frac{|M|}{\varphi(M)}\,\frac{q^{n/p+1}}{n^2} \leq R(n; M, q) - \frac{|M|}{\varphi(M)}\frac{q^{n+1}}{n^2} \leq q^n - \frac{q^n}{n} + 2\frac{q^{n/p}}{n}.$$

From this, it is clear from the dominant term we get infinitely many twin irreducibles pairs across multiple fields $\mathbb{F}_q[t]$ as $q \to \infty$.

Corollary 5.3. *Fix $n \geq 2$ and $0 \leq k < n$. Then for every $M \in \mathbb{F}_q[T]$ of degree k,*

$$R(n; M, q) \sim \frac{|M|}{\varphi(M)} \frac{q^{n+1}}{n^2} \qquad \text{as } q \to \infty.$$

In particular, for all sufficiently large q (depending only on n), there exist pairs $(\mathfrak{p}, \mathfrak{p}+M)$ of irreducibles of degree n in $\mathbb{F}_q[T]$; the number of such pairs grows without bound. Explicitly, $R(n; M, q) > 0$ whenever q exceeds an absolute constant multiple of n^2.

The key intermediate result is the identity

$$R(n; M, q) \ = \sum_{h \in (\mathcal{M}/\mathbf{R}_{n-1-k,M})^\times} N_h^2 - \pi_q(n)$$

where $(\mathcal{M}/\mathbf{R}_{l,M})^\times$ denotes the unit group of the congruence relation $\mathbf{R}_{l,M}$ on the monoid $\mathcal{M}$ of monic polynomials, under which $A \equiv B$ iff A and B agree modulo M *and* share their first l next-to-leading coefficients [9]; and N_h denotes the number of monic degree-n irreducibles in the class h. The derivation is direct: two monic degree-n irreducibles Q, Q' satisfy $Q' - Q \in \mathbb{F}_q \cdot M$ precisely when they agree modulo $\mathbf{R}_{n-1-k,M}$, so $\sum_h N_h^2$ counts ordered pairs (Q, Q') with $Q' - Q = \alpha M$ for some $\alpha \in \mathbb{F}_q$. Note such a pair corresponds to the pair $(\alpha^{-1}Q, \alpha^{-1}Q')$ that differ by M. Subtracting the diagonal contribution $\pi_q(n)$ (where $\alpha = 0$) yields the above.

The key lower bound now follows relatively elementarily with the Cauchy-Schwarz inequality and the PNT for $\mathbb{F}_q[t]$, noting $\sum_h N_h = \pi_q(n)$ and $\sum_h 1 = q^{n-1-k}\varphi(M)$ (using the fact that $[f] \in (\mathcal{M}/\mathbf{R}_{l,M})^\times \iff [f] \in (\mathbb{F}_q[t]/\langle M \rangle)^\times$ and $\#(\mathbb{F}_q[t]/\langle M \rangle)^\times = \varphi(M)$ and we have q^l choices of lead coefficients [6]):

$$\sum_h N_h^2 \geq \frac{\left(\sum_h N_h\right)^2}{\sum_h 1^2} = \frac{\pi(n; q)^2}{q^{n-1-k}\varphi(M)} \geq \frac{1}{q^{n-1-k}\varphi(M)} \left(\frac{q^n}{n} - \frac{2q^{n/p}}{n}\right)^2$$

Combined with the key identity gives

$$\begin{aligned} R(n; M, q) \ = \ \sum_h N_h^2 - \pi(n; q) &\geq \frac{1}{q^{n-1-k}\varphi(M)} \left(\frac{q^n}{n} - \frac{2q^{n/p}}{n}\right)^2 - \pi(n; q) \\ &\geq \frac{|M|}{\varphi(M)} \left(\frac{q^{n+1}}{n^2} - 4\frac{q^{n/p+1}}{n^2}\right) - \frac{q^n}{n} \end{aligned}$$

This gives $R(n; M, q) \to \infty$ as $q \to \infty$. The corresponding upper bound is more involved. Pollack instead bounds each N_h via an explicit character-sum expansion to obtain an equidistribution result of irreducibles across the elements of $(\mathcal{M}/\mathbf{R}_{n-1-k,M})^\times$, and applies Weil's Riemann Hypothesis over $\mathbb{F}_q$ (in Rhin's extension to characters of $(\mathcal{M}/\mathbf{R}_{l,M})^\times$) to control the resulting error.

Pollack overcomes parity for twin primes for a similar reason it was when it came to the Prime Number Theorem: $\mathbb{F}_q[t]$ allows for much simpler combinatorial arguments. The stronger analytic results available allows him to get tighter estimates than otherwise expected.

Remark. The non-monic counting in $R(n; M, q)$ is essential: $\sum_h N_h^2 - \pi_q(n)$ counts monic pairs differing by *any* scalar multiple of M, and the scaling $(Q, Q') \mapsto (\alpha^{-1}Q,\ \alpha^{-1}Q)$ collapses this $\mathbb{F}_q^\times$-orbit into the single shift M at the cost of allowing non-monic output. A monic analogue of Theorem 5.2 is, as Pollack notes, substantially harder.

5.2.2 Sawin-Shusterman: Twin Irreducibles in Fixed Field $\mathbb{F}_q$

Much more recently, Sawin and Shusterman proved the alternative, closer analogy of the twin prime conjecture for certain fixed fields [12].

Theorem 5.4 (Sawin–Shusterman)**.** *Let p be an odd prime and let q be a power of p satisfying $q > 685090\, p^2$. Then for any nonzero $h \in \mathbb{F}_q[T]$, we have*

$$\#\big\{ f \in \mathbb{F}_q[T] : |f| = X,\ f \text{ and } f+h \text{ are irreducible} \big\} \ \sim\ \mathfrak{S}_q(h)\, \frac{X}{\log_q^2 X}$$

as $X \to \infty$ through powers of q, with a power-saving error term depending on q. Here the singular series is

$$\mathfrak{S}_q(h) \ =\ \prod_{\mathfrak{p}} \big(1 - |\mathfrak{p}|^{-1}\big)^{-2} \big(1 - |\mathfrak{p}|^{-1} - |\mathfrak{p}|^{-1} \mathbf{1}_{\mathfrak{p} \nmid h}\big)\,,$$

where the product runs over all monic irreducibles $\mathfrak{p} \in \mathbb{F}_q[T]$.

That is, for a fixed field $\mathbb{F}_q$ satisfying the hypothesis, there are infinitely many twin irreducibles $f, f+h$ letting $\deg f \to \infty$. Their key intermediate result is resolving the *k-Chowla conjecture.*

Theorem 5.5 (k-Chowla, [12], Thm. 1.1)**.** *Let p be an odd prime, $k \geq 1$, and q a power of p with $q > p^{2k^2} e^2$. Then for every fixed choice of distinct $h_1, \dots, h_k \in \mathbb{F}_q[t]$,*

$$\sum_{\substack{f \in \mathbb{F}_q[t]^+ \\ |f| \leq X}} \mu(f+h_1)\, \mu(f+h_2) \cdots \mu(f+h_k) \ =\ o(X), \qquad X \to \infty$$

with a power-saving error term, where $\mathbb{F}_q[t]^+ = \{f \in \mathbb{F}_q[t] : f \text{ monic}\}$.

While not the same statement, k-Chowla and EH_{μ_h} both concern correlations of μ at shifted arguments—the Type II input that sieve methods cannot supply and that, via Vaughan's identity over $\mathbb{Z}$ or the convolution identity $\Lambda = -\mathbf{1} * (\mu \cdot \deg)$ over $\mathbb{F}_q[t]$, carries the parity information the twin prime problem demands.

Theorem 5.5 is the principal input to Theorem 5.4. The reduction in fact uses a shift-uniform strengthening ([12], Thm. 4.5), but the mechanism is that of Theorem 5.5. The key

observation is that for any polynomial r, the function $\mu(r+s^p)$ is essentially $\chi_r(s+c_r)$ for a quadratic Dirichlet character χ_r and shift c_r depending only on r ([12], Lem. 3.2)—a consequence of Pellet's formula and the vanishing of the derivative $(r+s^p)' = r'$ in characteristic p. Restricting the Chowla sum to p-th power slices and collapsing the product via Chinese remainder reduces it to a short sum of a single Dirichlet character, which is bounded by a geometric refinement of Burgess's bound ([12], Thm. 1.4) via ℓ-adic sheaf cohomology and Deligne's Weil II. There is no obvious analogous integer strategy.

The second ingredient is a level-of-distribution estimate for μ.

Theorem 5.6 (Level of distribution for μ, [12], Thm. 1.7)**.** *Fix $\eta > 0$. For an odd prime p and q a power of p sufficiently large in terms of η, the following holds. For coprime $a, M \in \mathbb{F}_q[t]$ and $X \in \mathbb{R}$ with $X^{1-\eta} \geq |M|$,*

$$\sum_{\substack{f \in \mathbb{F}_q[t]^+ \\ |f| \leq X \\ f \equiv a \bmod M}} \mu(f) = o\left(\frac{X}{|M|}\right), \qquad |M| \to \infty.$$

Remark. Note this is level of distribution for μ itself—a type of cancellation statement—as distinct from the Type I level of distribution for a sieve sequence $\mathcal{A}$ as discussed earlier. The latter was shown to be unhelpful for parity, whereas the former is precisely the kind of μ-input parity-breaking requires. They are not completely unrelated, though. For example, an equivalent formulation of the Bombieri–Vinogradov theorem can be thought of as stating the level of distribution of Λ is $\theta = 1/2$.

So μ over $\mathbb{F}_q[t]$ achieves level of distribution $\theta = 1 - \eta$ for arbitrarily small η—beyond Bombieri–Vinogradov's $\theta = 1/2$, and past what the (original, not shifted) Elliott–Halberstam conjecture predicts over $\mathbb{Z}$. The proof routes through the same Dirichlet-character observation ([12], Lem. 3.2): μ in an arithmetic progression modulo M reduces to a short character that is handled by the Burgess bound.

The two inputs are combined through the convolution identity

$$\Lambda = -\mathbf{1} * (\mu \cdot \deg), \qquad \Lambda(f) = -\sum_{g \mid f} \mu(g)\, \deg(g),$$

the function-field analogue of $\Lambda = \mu * \log$. A hyperbola-truncated form of the above convolution applied to $\Lambda(a + gM)$ splits at the diagonal: Theorem 5.6 controls the short-divisor piece, the Chowla input controls the long-divisor piece. The output ([12], Cor. 6.1) is a power-saving bound on

$$\sum_{\substack{g \in \mathbb{F}_q[t]^+ \\ \deg g = d}} \Lambda(a + gM) \prod_{i=1}^{n} \mu(a_i + gM_i)$$

This is the direct polynomial analogue of EH_{μ_h}, and is in fact stronger, since the previous conjecture is an averaged statement over moduli with log-saving error, while the above holds for each fixed M with power-saving error and for multiple linear-form shifts simultaneously.

Theorem 5.4 then follows from the above strengthened EH_{μ_h} via the reduction shown by [10]. In short: what Sawin–Shusterman supply is the Type II μ-cancellation parity demands, produced by a characteristic-p identification of μ with a Dirichlet character that has no integer analogue.

6 Conclusion

We have looked at sieve techniques with a focus on the twin prime conjecture, comparing the difference between $\mathbb{Z}$ and $\mathbb{F}_q[t]$ to highlight and isolate where sieves fail and where other methods succeed. Our discussion had three major arcs. First, we adapted sieve methodologies—specifically the Selberg sieve—to $\mathbb{F}_q[t]$ to compute an upper bound on the number of twin irreducible polynomials. Second, we discussed the parity problem as the primary limitation to sieve techniques producing not only asymptotics, but non-trivial lower bounds on sets that have constant parity. Notably, it seemed to be simplicity of the Type I inputs to sieves that prevented them from seeing and working around parity. Finally, using sieves as a baseline, we looked at the state of the twin prime conjecture in $\mathbb{Z}$ and $\mathbb{F}_q[t]$, and some of the key facts at the heart of the conjecture to bypass sieve shortcomings, whose availability differs in the two settings. The strong combinatorial, analytic, and geometric data of $\mathbb{F}_q[t]$ ultimately is what made the twin prime conjecture tractable; lacking integer analogues, sieves remain at the forefront for it in the case of $\mathbb{Z}$.

7 References

[1] Yingchun Cai. "On Chen's theorem (II)". en. In: *Journal of Number Theory* 128.5 (May 2008), pp. 1336–1357. ISSN: 0022314X. DOI: 10.1016/j.jnt.2007.04.010. URL: https://linkinghub.elsevier.com/retrieve/pii/S0022314X07001217 (visited on 04/19/2026).

[2] Sunil K. Chebolu and JáN Mináč. "Counting Irreducible Polynomials over Finite Fields Using the Inclusion-Exclusion Principle". en. In: *Mathematics Magazine* 84.5 (Dec. 2011), pp. 369–371. ISSN: 0025-570X, 1930-0980. DOI: 10.4169/math.mag.84.5.369. URL: https://www.tandfonline.com/doi/full/10.4169/math.mag.84.5.369 (visited on 04/05/2026).

[3] John Friedlander and Henryk Iwaniec. *Opera de Cribro*. en. Vol. 57. Colloquium Publications. Providence, Rhode Island: American Mathematical Society, June 2010. ISBN: 978-0-8218-4970-5 978-1-4704-1766-6. DOI: 10.1090/coll/057. URL: http://www.ams.org/coll/057 (visited on 02/02/2026).

[4] Ben Green. "ANALYTIC NUMBER THEORY". en. In: (). URL: https://courses.maths.ox.ac.uk/pluginfile.php/118621/mod_resource/content/3/C3.8-notes-2025.pdf.

[5] Glyn Harman. *Prime-detecting sieves*. en. First paperback printing, 2020. London Mathematical Society Monograph series Vol. 33. Princeton Oxford: Princeton University Press, 2020. ISBN: 978-0-691-12437-7 978-0-691-20299-0.

[6] David R Hayes. "The distribution of irreducibles in GF[q,x]". en. In: *Transactions of the American Mathematical Society* 117 (May 1965), pp. 101–127. DOI: 10.2307/1994199.

[7] Xiaoyu He. *Why is there a Parity Problem in Sieve Theory and not a Mod p problem for any other p?* URL: https://mathoverflow.net/questions/233240/why-is-there-a-parity-problem-in-sieve-theory-and-not-a-mod-p-problem-for-any-ot.

[8] D. R. Heath-Brown. *Lectures on sieves*. en. arXiv:math/0209360. Sept. 2002. DOI: 10.48550/arXiv.math/0209360. URL: http://arxiv.org/abs/math/0209360 (visited on 02/08/2026).

[9] Paul Pollack. "A polynomial analogue of the twin prime conjecture". en. In: *Proceedings of the American Mathematical Society* 136.11 (May 2008), pp. 3775–3784. ISSN: 0002-9939, 1088-6826. DOI: 10.1090/S0002-9939-08-09351-9. URL: https://www.ams.org/proc/2008-136-11/S0002-9939-08-09351-9/ (visited on 01/19/2026).

[10] M. Ram Murty and Akshaa Vatwani. "Twin primes and the parity problem". en. In: *Journal of Number Theory* 180 (Nov. 2017), pp. 643–659. ISSN: 0022314X. DOI: 10.1016/j.jnt.2017.05.011. URL: https://linkinghub.elsevier.com/retrieve/pii/S0022314X17302135 (visited on 04/19/2026).

[11] Michael Rosen. *Number Theory in Function Fields*. en. Vol. 210. Graduate Texts in Mathematics. New York, NY: Springer New York, 2002. ISBN: 978-1-4419-2954-9 978-1-4757-6046-0. DOI: 10.1007/978-1-4757-6046-0. URL: http://link.springer.com/10.1007/978-1-4757-6046-0 (visited on 01/19/2026).

[12] Will Sawin and Mark Shusterman. "On the Chowla and twin primes conjectures over $\mathbb{F}_q[T]$". en. In: *Annals of Mathematics* 196.2 (Sept. 2022). ISSN: 0003-486X. DOI: 10.4007/annals.2022.196.2.1. URL: https://projecteuclid.org/journals/annals-of-mathematics/volume-196/issue-2/On-the-Chowla-and-twin-primes-conjectures-over-mathbbF_qT/10.4007/annals.2022.196.2.1.full (visited on 01/19/2026).

[13] Terry Tao. *Open question: The parity problem in sieve theory — What's new*. en. July 2007. URL: https://terrytao.wordpress.com/2007/06/05/open-question-the-parity-problem-in-sieve-theory/.

[14] Gérald Tenenbaum. *Introduction to Analytic and Probabilistic Number Theory*. en. Trans. by Patrick Ion. Vol. 163. Graduate Studies in Mathematics. Providence, Rhode Island: American Mathematical Society, July 2015. ISBN: 978-0-8218-9854-3 978-1-4704-2223-3 978-1-4704-2237-0. DOI: 10.1090/gsm/163. URL: https://www.ams.org/gsm/163 (visited on 04/11/2026).

[15] William A Webb. "Sieve methods for polynomial rings over finite fields". en. In: *Journal of Number Theory* 16.3 (June 1983), pp. 343–355. ISSN: 0022314X. DOI: 10.1016/0022-314X(83)90062-8. URL: https://linkinghub.elsevier.com/retrieve/pii/0022314X83900628 (visited on 01/19/2026).

[16] Denis Xavier Charles. *Sieve Methods*. Thesis. July 2000. URL: https://pages.cs.wisc.edu/~cdx/Sieve.pdf.