Revisiting Gauss's analogue of the prime number theorem for polynomials over a finite field[☆]

Paul Pollack

Department of Mathematics
Dartmouth College
Hanover, NH 03755

Abstract

In 1901, von Koch showed that the Riemann Hypothesis is equivalent to the assertion that

$$\sum_{\substack{p \le x \\ p \text{ prime}}} 1 = \int_2^x \frac{dt}{\log t} + O(\sqrt{x} \log x).$$

We describe an analogue of von Koch's result for polynomials over a finite prime field $\mathbf{F}_p$: For each natural number n, write n in base p, say

$$n = a_0 + a_1 p + \cdots + a_k p^k,$$

and associate to n the polynomial $a_0 + a_1 T + \cdots + a_k T^k \in \mathbf{F}_p[T]$. We let $\pi_p(X)$ denote the number of irreducible polynomials encoded by integers $n < X$, and prove a formula for $\pi_p(X)$ valid with an error term analogous to that in von Koch's theorem. Our result is unconditional, and is grounded in Weil's Riemann Hypothesis for function fields. We also investigate an asymptotic expansion for $\pi_p(X)$.

Keywords: prime number theorem, von Koch's theorem, irreducible polynomials, prime polynomials
2000 MSC: 11T06, 11N05

1. Introduction

By 1797, Gauss had already proved[1] one of the foundational results of the theory of finite fields: Letting $\pi(q; d)$ denote the number of one-variable monic irreducible

[☆]This material is based upon work supported by the National Science Foundation under agreement No. DMS-0635607. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

Email address: pppollac@illinois.edu (Paul Pollack)

[1]Actually Gauss stated his result only for prime q, but the argument carries over to the general case without any changes.

polynomials of degree d over $\mathbf{F}_q$, we have (for all $n \geq 1$)

$$\sum_{d|n} d\pi(q;d) = q^n \quad \text{and so by inversion,} \quad \pi(q;n) = \frac{1}{n}\sum_{d|n} q^d \mu(n/d).$$

From these formulas we may easily deduce that

$$\frac{q^d}{d} - 2\frac{q^{d/2}}{d} \leq \pi(q;d) \leq \frac{q^d}{d}. \tag{1}$$

(Slightly sharper estimates are given in [1, Exercises 3.27, 3.28].) In fact, Gauss drafted an entire section 8 of his *Disquisitiones Arithmeticae* devoted to what we now recognize as the theory of finite fields. Owing to considerations of size, this section was cut from the published version; Gauss intended to present this material in a second volume, which unfortunately never appeared (see [2]).

For someone versed in the modern theory of prime numbers, there is a striking resemblance between Gauss's result and the prime number theorem, which asserts that with

$$\pi(x) := \#\{p \leq x : p \text{ prime}\},$$

we have

$$\pi(x) \sim \frac{x}{\log x} \qquad \text{as } x \to \infty. \tag{2}$$

Indeed, (1) implies that $\pi(q;d) \sim q^d/d$ whenever $q^d \to \infty$, and the expression q^d/d has the shape $X/\log_q X$, where $X = q^d$ and $\log_q(\cdot)$ denotes the logarithm to the base q. The purpose of this note is to draw attention to some more subtle and lesser-known analogies between $\pi(x)$ and $\pi(q;d)$.

To motivate what follows, let us recall another early discovery of Gauss. In an 1849 letter to Encke, Gauss describes how (as a boy of around 16 years of age) he observed that the primes near a large number x occur with a 'density' which is roughly $1/\log x$. This observation motivates the conjecture that $\pi(x)$ is approximately given by the *logarithmic sum* of x, defined by

$$\mathrm{ls}(x) := \sum_{2 \leq n \leq x} \frac{1}{\log n}.$$

A straightforward partial summation shows that $\mathrm{ls}(x)$ is asymptotic to $x/\log x$ as $x \to \infty$; thus, by (2), $\mathrm{ls}(x)$ is a good first order approximation to $\pi(x)$. It is now known that in fact $\mathrm{ls}(x)$ is a much better approximation to $\pi(x)$ than $x/\log x$. In 1901, von Koch [3] established the following result, which in particular shows that the Gauss approximation $\mathrm{ls}(x)$ is accurate up to a 'square root error term', provided that the Riemann Hypothesis is true:

Theorem A (von Koch)**.** *The Riemann Hypothesis is equivalent to the estimate*

$$\pi(x) = \mathrm{ls}(x) + O(\sqrt{x}\log x) \qquad \textit{as } x \to \infty.$$

Our first goal in this paper is to establish an analogue of von Koch's result for polynomials over a finite field. For the sake of simplicity, we work initially over $\mathbf{F}_p$, where p is prime. Notice that the nonnegative integers are in bijection with the one-variable polynomials over $\mathbf{F}_p$ via the correspondence

$$a_n p^n + a_{n-1} p^{n-1} + \cdots + a_1 p + a_0 \longleftrightarrow a_n T^n + a_{n-1} T^{n-1} + \cdots + a_1 T + a_0,$$

where the integer represented on the left-hand side is assumed written in its base p expansion (so that $0 \le a_i < p$). If the integer a corresponds to the polynomial A, we will write $\|A\| = a$. For an interval of real numbers I, we define

$$\pi_p(I) := \#\{P \in \mathbf{F}_p[T] : \|P\| \in I \text{ and } P \text{ is irreducible}\},$$

and we set

$$\pi_p(X) := \pi_p([0, X)).$$

Gauss's formula (1) can be read as the assertion that the 'density' of irreducibles among all polynomials of degree d is roughly $1/d$. In analogy with the definition of ls, define

$$\mathrm{ls}_p(X) := \sum_{\substack{\|f\|<X \\ \deg f>0}} \frac{1}{\deg f}.$$

Our main result is the following:

Theorem 1. *Let p be a prime and $X \ge p$. Suppose that $p^n \le X < p^{n+1}$. Then*

$$\pi_p(X) = \mathrm{ls}_p(X) + O(np^{n/2+1}),$$

where the O-constant is absolute.

Notice that the inside of the O-term is $\asymp_p \sqrt{X} \log X$, in exact analogy with von Koch's theorem. Our result is unconditional, owing to Weil's proof of the Riemann Hypothesis for function fields.

The estimate of von Koch alluded to before is usually written in the form

$$\pi(x) = \int_2^x \frac{dt}{\log t} + O(\sqrt{x} \log x).$$

This is equivalent to the preceding formulation, since the right-hand integral, traditionally denoted $\mathrm{li}(x)$, differs by a bounded amount from the sum $\sum_{2 \le n \le x} 1/\log n$. In seeking to approximate $\mathrm{li}(x)$, one is led, via repeated integration by parts, to the approximation

$$\mathrm{li}(x) = \frac{x}{\log x} + 1!\frac{x}{\log^2 x} + 2!\frac{x}{\log^3 x} + \cdots + r!\frac{x}{\log^{r+1} x} + O_r\left(\frac{x}{\log^{r+1} x}\right), \qquad (3)$$

valid for every $r \ge 1$. (This is one of the canonical examples of an *asympotic series*; for background see, e.g., [4, Chapter 1.5].) Since the difference between $\pi(x)$ and $\mathrm{li}(x)$ is known, unconditionally, to be $O_r(x/\log^r x)$ for every r (see [5, Chapter 18]), it follows that $\pi(x)$ has the same asymptotic expansion (3).

It is natural to wonder if there is an analogue of formula (3) for $\pi_p(X)$. This is, in fact, the case:

Theorem 2. *Let p be a prime. Let X be an integer with $X \geq p^2$, and let n be the natural number with $p^n \leq X < p^{n+1}$. For each $r \geq 2$, we have*

$$\pi_p(X) = \frac{X}{n} + \sum_{k=2}^{r}(1 - 1/p)A_{p,k}\frac{p^n}{n^k} + O\left(np^{n/2+1} + A_{p,r+2}\frac{p^n}{n^{r+1}} + \frac{p}{n}\sum_{k=1}^{r}A_{p,k}\right).$$

Here the implied constant is absolute, and the $A_{p,k}$ are defined by

$$A_{p,k} := \sum_{m=1}^{\infty}\frac{m^{k-1}}{p^{m-1}}.$$

Since for a fixed k, the constants $A_{p,k}$ are decreasing in p, Theorem 2 implies that whenever $X \geq p^2$,

$$\pi_p(X) = \frac{X}{n} + \sum_{k=2}^{r}(1 - 1/p)A_{p,k}\frac{p^n}{n^k} + O_r\left(\frac{p^n}{n^{r+1}}\right).$$

In particular (taking, say, $r = 2$), we obtain the following result, which should be viewed as a version of the prime number theorem for polynomials that holds with some uniformity:

Corollary. *We have $\pi_p(X) \sim X/\log_p X$ whenever $\log_p X \to \infty$.*

That Theorem 2 is really the correct analogue of (3) will emerge from an estimate for $A_{p,k}$ proved as Lemma 7 below.

2. Preliminary results on irreducibles with prescribed leading coefficients

Let $\mathcal{P}$ be the (multiplicative) monoid of monic polynomials over $\mathbf{F}_q$. For each $l \geq 0$, we define a relation $\mathcal{R}_l$ on $\mathcal{P}$ by saying that $A \equiv B \pmod{\mathcal{R}_l}$ if A and B have the same first l next-to-leading coefficients. Here if $A = T^n + a_{n-1}T^{n-1} + \cdots + a_0$, its *first l next-to-leading coefficients* are $a_{n-1}, \ldots, a_{n-l}$, with the understanding that $a_i = 0$ for $i < 0$. Thus $T^6 + 3T^4 - T^3 + T + 1$ and $T^2 + 3$ are congruent modulo $\mathcal{R}_2$ but not modulo $\mathcal{R}_3$.

In the terminology of [6], $\mathcal{R}_l$ is a *congruence relation* on $\mathcal{P}$, i.e., an equivalence relation satisfying

$$A \equiv B \bmod \mathcal{R}_l \Rightarrow AC \equiv BC \bmod \mathcal{R}_l \quad \text{for all} \quad A, B, C \in \mathcal{P}.$$

Thus there is a well-defined quotient monoid $\mathcal{P}/\mathcal{R}_l$. Moreover, every element of $\mathcal{P}/\mathcal{R}_l$ is invertible: Indeed, if $A(T)$ is any monic polynomial and k is chosen so that $q^k > l$, then

$$A(T)(A(T)^{q^k-1}) = A(T)^{q^k} = A(T^{q^k}) \equiv 1 \pmod{\mathcal{R}_l}.$$

Thus $\mathcal{P}/\mathcal{R}_l$ is an abelian group. Clearly $\#\mathcal{P}/\mathcal{R}_l = q^l$.

The following explicit formula for primes in congruence classes modulo $\mathcal{R}_l$ is a consequence of Weil's Riemann Hypothesis. It is a special case (the case when $M = 1$) of [7, Lemma 1]:

Lemma 1. *Let A be a monic polynomial. Then*

$$q^l \sum_{\substack{Q^j \equiv A \pmod{\mathcal{R}_l} \\ \deg Q^j = n}} \deg Q = q^n - \sum_{\chi} \bar{\chi}(A) \sum_{i=1}^{a(\chi)} \beta_i(\chi)^n,$$

where the left-hand sum is over monic irreducible Q and χ runs over all characters modulo $\mathcal{R}_l$. Moreover, $a(\chi) \le l$ for all χ, and each $|\beta_i(\chi)| \le q^{1/2}$.

From this formula we can easily deduce the following theorem concerning primes in 'progressions' modulo $\mathcal{R}_l$:

Lemma 2. *Let l be a nonnegative integer. The number of monic irreducibles of degree n belonging to a prescribed residue class modulo $\mathcal{R}_l$ is*

$$\frac{1}{n} q^{n-l} + O\left((l+1)\frac{q^{n/2}}{n}\right).$$

Proof. The right-hand side of Lemma 1 differs from q^n by an error which is bounded in absolute value by $q^l \cdot l \cdot q^{n/2}$, so that

$$\sum_{\substack{Q^j \equiv A \pmod{\mathcal{R}_{l,M}} \\ \deg Q^j = n}} \deg Q = q^{n-l} + O(lq^{n/2}).$$

The terms of the sum for which $j > 1$ contribute

$$\le \sum_{\substack{d|n \\ d<n}} d\pi(q;d) \le \sum_{\substack{d|n \\ d<n}} q^d \le 2q^{n/2},$$

where we use the upper bound on $\pi(q;d)$ from (1). Hence

$$n \sum_{\substack{Q \equiv A \pmod{\mathcal{R}_l} \\ \deg Q = n}} 1 = q^{n-l} + O\left((l+1)q^{n/2}\right).$$

Now divide by n. □

3. Proof of Theorem 1

For the proof of Theorem 1 we may (and do) assume that X is an integer. Suppose that $p^n \le X < p^{n+1}$, and write

$$X = a_n p^n + a_{n-1} p^{n-1} + \cdots + a_1 p + a_0,$$

with each $0 \le a_i < p$. Then we have the easy decomposition

$$\pi_p(X) = \pi_p([0, p^n)) + \pi_p([p^n, a_n p^n)) + \sum_{j=1}^{n} \pi_p\left(\left[\sum_{i=j}^{n} a_i p^i, \sum_{i=j-1}^{n} a_i p^i\right)\right). \quad (4)$$

We treat each of the three terms of (4) separately.

Lemma 3. *We have*

$$\pi_p([0, p^n)) = (p-1)\sum_{m=1}^{n-1} \frac{p^m}{m} + O(p^{(n+1)/2}/n).$$

Proof. Clearly $\pi_p([0, p^n)) = (p-1)\sum_{m=1}^{n-1} \pi(p; m)$. Now put in the estimate $\pi(p; m) = p^m/m + O(p^{m/2}/m)$; the lemma follows once we show that

$$\sum_{1 \le m \le n-1} \frac{p^{m/2}}{m} \ll \frac{p^{(n-1)/2}}{n}.$$

This latter estimate is trivial if $n \le 3$, so suppose that $n \ge 4$. The terms with $m < 3$ contribute $\ll p \ll p^{(n-1)/2}/n$ to the sum. For $3 \le m < n-1$, the ratio

$$\frac{p^{m/2}/m}{p^{(m+1)/2}/(m+1)} \le \frac{m+1}{m} p^{-1/2} \le \frac{4}{3} 2^{-1/2} < 1,$$

and so

$$\sum_{3 \le m \le n-1} \frac{p^{m/2}}{m} \ll \frac{p^{(n-1)/2}}{n-1} \ll \frac{p^{(n-1)/2}}{n}.$$

So the estimate holds in this case also. □

Lemma 4. *We have*

$$\pi_p([p^n, a_n p^n)) = (a_n - 1)\frac{p^n}{n} + O(p^{n/2+1}/n).$$

Proof. The left-hand side counts the number of irreducibles of degree n with leading coefficient one of $1, 2, \dots, a_n - 1$, so again by (1),

$$\pi_p((p^n, a_n p^n]) = (a_n - 1)\left(\frac{p^n}{n} + O(p^{n/2}/n)\right) = (a_n - 1)\frac{p^n}{n} + O(p^{n/2+1}/n). \qquad \square$$

Lemma 5. *For every $1 \le j \le n$, we have*

$$\pi_p\left(\left[\sum\nolimits_{i=j}^{n} a_i p^i, \sum\nolimits_{i=j-1}^{n} a_i p^i\right)\right) = a_{j-1}\frac{p^{j-1}}{n} + O\left((n-j+2)\frac{p^{n/2+1}}{n}\right).$$

Proof. The left-hand side represents the number of degree-n primes whose first $n-j+1$ leading coefficients are $a_n, a_{n-1}, \dots, a_j$, and whose T^{j-1}-coefficient is one of the a_{j-1} values $0, 1, \dots, a_{j-1} - 1$. For each fixed value of the T^{j-1}-coefficient, the number of such irreducibles is the same as the number of degree-n monic irreducibles belonging to a certain prescribed congruence class modulo $\mathcal{R}_{n-j+1}$. By Lemma 2, each such congruence class contains

$$\frac{1}{n}p^{j-1} + O\left((n-j+2)\frac{p^{n/2}}{n}\right)$$

such irreducibles. Summing over the a_{j-1} possible coefficients of T^{j-1} yields the lemma. □

Proof of Theorem 1. By (4) and Lemmas 3–5, we have

$$\pi_p(X) = (a_n - 1)\frac{p^n}{n} + \sum_{j=1}^{n} a_{j-1}\frac{p^{j-1}}{n} + (p-1)\sum_{m=1}^{n-1}\frac{p^m}{m} + O(p^{n/2+1}/n) + O\left(\frac{1}{n}p^{n/2+1}\sum_{1\le j\le n}(n-j+2)\right).$$

Since $\sum_{1\le j\le n}(n-j+2) \ll n^2$, we can collect the O-terms into an error of $O(np^{n/2+1})$. So simplifying, we obtain an estimate for $\pi_p(X)$ of

$$\frac{1}{n}\left(\sum_{i=0}^{n} a_i p^i - p^n\right) + (p-1)\sum_{m=1}^{n-1}\frac{p^m}{m} + O(np^{n/2-1}) = \frac{X-p^n}{n} + (p-1)\sum_{m=1}^{n-1}\frac{p^m}{m} + O(np^{n/2-1}).$$

But the main term in this last expression is precisely $\sum_{\substack{\|f\|<X \\ \deg f>0}} \frac{1}{\deg f} = \mathrm{ls}_p(X)$, as we see upon grouping the contributions to this sum according to the degree of f. □

Proof of Theorem 2

We require the following slight variant of a result of Lenskoi [8]:

Lemma 6. *For each $r \ge 1$ and $n \ge 2$, we have*

$$\sum_{m=1}^{n-1}\frac{p^m}{m} = \sum_{k=1}^{r} A_{p,k}\frac{p^{n-1}}{n^k} + O\left(\frac{1}{n}\sum_{k=1}^{r}A_{p,k}\right) + O\left(A_{p,r+2}\frac{p^{n-1}}{n^{r+1}}\right),$$

where the O-constants are absolute and the constants $A_{p,k}$ are defined as in the statement of Theorem 2.

Proof. We largely follow Lenskoi. We have

$$\frac{1}{p^{n-1}}\sum_{m=1}^{n-1}\frac{p^m}{m} = \sum_{m=1}^{n-1}\frac{1}{mp^{n-1-m}} = \sum_{m=1}^{n-1}\frac{1}{(n-m)p^{m-1}} = \sum_{m=1}^{n-1}\frac{1}{p^{m-1}}\sum_{k=1}^{\infty}\frac{m^{k-1}}{n^k} = \sum_{k=1}^{\infty}\frac{1}{n^k}\sum_{m=1}^{n-1}\frac{m^{k-1}}{p^{m-1}}.$$

We split this last expression into three parts:

$$\sum_{k=1}^{\infty}\frac{1}{n^k}\sum_{m=1}^{n-1}\frac{m^{k-1}}{p^{m-1}} = \sum_{k=1}^{r}\frac{1}{n^k}\sum_{m=1}^{n-1}\frac{m^{k-1}}{p^{m-1}} + \sum_{k=r+1}^{\infty}\frac{1}{n^k}\sum_{m=1}^{n-1}\frac{m^{k-1}}{p^{m-1}} = \sum_{k=1}^{r}\frac{1}{n^k}A_{p,k} - \sum_{k=1}^{r}\frac{1}{n^k}\sum_{m=n}^{\infty}\frac{m^{k-1}}{p^{m-1}} + \sum_{k=r+1}^{\infty}\frac{1}{n^k}\sum_{m=1}^{n-1}\frac{m^{k-1}}{p^{m-1}}.$$

The first sum yields the main term in Lemma 6, and it remains to show that the latter two contribute appropriately bounded error terms. The first double sum is just

$$\sum_{k=1}^{r} \frac{1}{n^k} \frac{1}{p^{n-1}} \sum_{m=1}^{\infty} \frac{(m-1+n)^{k-1}}{p^{m-1}} \leq \frac{1}{n} \frac{1}{p^{n-1}} \sum_{k=1}^{r} \sum_{m=1}^{\infty} \frac{m^{k-1}}{p^{m-1}} = \frac{1}{n} \frac{1}{p^{n-1}} \sum_{k=1}^{r} A_{p,k},$$

using

$$\frac{m-1+n}{n} = 1 + \frac{m-1}{n} \leq 1 + (m-1) = m.$$

This corresponds to the first O-term above. To estimate the remaining double sum, notice that

$$\begin{aligned}
\sum_{k=r+1}^{\infty} \frac{1}{n^k} \sum_{m=1}^{n-1} \frac{m^{k-1}}{p^{m-1}} &= \frac{1}{n^{r+1}} \sum_{s=0}^{\infty} \frac{1}{n^s} \sum_{m=1}^{n-1} \frac{m^{s+r}}{p^{m-1}} \\
&= \frac{1}{n^{r+1}} \sum_{m=1}^{n-1} \frac{m^r}{p^{m-1}} \frac{1}{1-m/n} \\
&= \frac{1}{n^{r+1}} \sum_{m=1}^{n-1} \frac{m^r}{p^{m-1}} \left(1 + \frac{m}{n-m}\right).
\end{aligned}$$

Since $m/(n-m) \leq m$, this is bounded above by

$$\frac{1}{n^{r+1}} \left(\sum_{m=1}^{n-1} \frac{m^r}{p^{m-1}} + \sum_{m=1}^{n-1} \frac{m^{r+1}}{p^{m-1}} \right) \leq \frac{A_{p,r+1} + A_{p,r+2}}{n^{r+1}} \leq 2 \frac{A_{p,r+2}}{n^{r+1}}.$$

Multiplying through by p^{n-1}, we obtain the second O-term in the estimate of the theorem. □

Proof of Theorem 2. We have already seen that Theorem 1 is just the statement that

$$\pi_p(X) = \frac{X - p^n}{n} + (p-1) \sum_{m=1}^{n-1} \frac{p^m}{m} + O(np^{n/2+1}). \tag{5}$$

According to Lemma 6, we have

$$(p-1) \sum_{m=1}^{n-1} \frac{p^m}{m} = \sum_{k=1}^{r} (1-1/p) A_{p,k} \frac{p^n}{n^k} + O\left(\frac{p}{n} \sum_{k=1}^{r} A_{p,k} \right) + O\left(A_{p,r+2} \frac{p^n}{n^{r+1}} \right). \tag{6}$$

Now the $k = 1$ term in the right-hand sum contributes exactly

$$\left((1-1/p) \sum_{m=1}^{\infty} \frac{1}{p^{m-1}} \right) \frac{p^n}{n} = \frac{p^n}{n}.$$

So Theorem 2 follows upon inserting (6) into (5). □

We now make good on our promise to show that Theorem 2 is a genuine analogue of (3).

Lemma 7. *If p is a prime and k is a positive integer, then*

$$A_{p,k} = p\frac{(k-1)!}{(\log p)^k}\left(1 + O\left(\frac{\log p}{\sqrt{k}}\right)\right).$$

Proof. By the Euler–Maclaurin summation formula, we have

$$\frac{A_{p,k}}{p} = \sum_{m=1}^{\infty} m^{k-1}p^{-m} = \int_0^{\infty} t^{k-1}\exp(-t\log p)\,dt + O\left(\int_0^{\infty}\left|\frac{d}{dt}\left(t^{k-1}\exp(-t\log p)\right)\right|\,dt\right).$$

A change of variables gives a main term of precisely

$$\frac{\Gamma(k)}{(\log p)^k} = \frac{(k-1)!}{(\log p)^k},$$

while the unimodality of the original integrand ensures that the error term is

$$\ll \max_{t\geq 0} t^{k-1}\exp(-t\log p) = t^{k-1}\exp(-t\log p)\Big|_{t=(k-1)/\log p} = ((k-1)/e)^{k-1}/(\log p)^{k-1} \ll \frac{(k-1)!}{(\log p)^k}\frac{\log p}{\sqrt{k}}.$$

In the last line we have applied Stirling's formula to estimate $(k-1)!$. □

The analogy between (3) and the result of Theorem 2 is clearest when $X = p^n$ is a power of p. In this case, Theorem 2 asserts that

$$\pi_p(X) = \sum_{k=1}^{r}(1-1/p)A_{p,k}\frac{p^n}{n^k} + O_r\left(\frac{p^n}{n^{r+1}}\right),$$

where $r \geq 2$ is an integer parameter at our disposal. By Lemma 7, the jth term in the sum is

$$(p-1)\frac{(j-1)!}{(\log p)^j}\frac{p^n}{n^j}\left(1 + O\left(\frac{\log p}{\sqrt{j}}\right)\right) = (p-1)(j-1)!\frac{X}{\log^j X}\left(1 + O\left(\frac{\log p}{\sqrt{j}}\right)\right).$$

If j is large compared to $\log p$, then it makes sense to say that the main term here is

$$(p-1)(j-1)!\frac{X}{\log^j X}.$$

This coincides with the jth term in the asymptotic expansion (3) of $\pi(X)$, except for the factor of $p-1$. This factor can be attributed to $\pi_p(X)$ counting all primes irrespective of their leading coefficient, whereas $\pi(X)$ counts only positive primes.

Remark on the case of arbitrary finite fields. When q is not prime, then there is no longer a canonical correspondence between the integers $0, 1, 2, \ldots, q-1$ and the elements of $\mathbf{F}_q$. However, if we pick any labeling of the elements of $\mathbf{F}_q$ by $\{0, 1, \ldots, q-1\}$ in which 0 corresponds to 0, then all of our results remain true, with O-constants uniform in both q and the choice of labeling.

Acknowledgements

I would like to thank Carl Pomerance, who supervised the doctoral dissertation where these results first appeared. I am also grateful to the anonymous reviewers for several helpful suggestions.

References

[1] R. Lidl, H. Niederreiter, Finite fields, vol. 20 of *Encyclopedia of Mathematics and its Applications*, Cambridge University Press, Cambridge, second edn., 1997.

[2] G. Frei, The unpublished section eight: on the way to function fields over a finite field, in: The shaping of arithmetic after C. F. Gauss's *Disquisitiones arithmeticae*, Springer, Berlin, 159–198, 2007.

[3] H. v. Koch, Sur la distribution des nombres premiers, Acta Math. 24 (1901) 159–182.

[4] N. G. de Bruijn, Asymptotic methods in analysis, Dover Publications Inc., New York, third edn., 1981.

[5] H. Davenport, Multiplicative number theory, vol. 74 of *Graduate Texts in Mathematics*, Springer-Verlag, New York, third edn., 2000.

[6] D. R. Hayes, The distribution of irreducibles in GF[q, x], Trans. Amer. Math. Soc. 117 (1965) 101–127.

[7] P. Pollack, A polynomial analogue of the twin prime conjecture, Proc. Amer. Math. Soc. 136 (11) (2008) 3775–3784.

[8] D. N. Lenskoi, On the arithmetic of polynomials over a finite field, Volz. Mat. Sb. 4 (1966) 155–159.