

Arithmetic of Function Fields

Seewoo Lee

v1.0

Last updated: August 18, 2025

Abstract

This is an introductory note for arithmetic of function fields, used for Berkeley Math REU happened in Summer 2025. Most of the materials are based on the Rosen's book *Number Theory in Function Fields*.

Contents

Introduction	2
1.1 Prerequisites	2
1.2 Notations	2
1.3 SageMath	2
Basic number theory and their analogues for polynomials	3
2.1 Fundamental Theorem of Arithmetic	3
2.2 Chinese Remainder Theorem	3
2.3 Fermat's Little Theorem and Euler's Theorem	5
2.4 Wilson's Theorem	7
Reciprocity Laws	8
3.1 Quadratic Reciprocity for Integers	8
3.2 Quadratic Reciprocity for Polynomials	11
How many primes?	15
4.1 Prime number theorem and Riemann zeta function	15
4.2 Counting irreducible polynomials with zeta function	18
4.3 Counting other things	21
4.4 Sum of arithmetic functions	23
How many primes in arithmetic progressions?	26
5.1 Dirichlet's theorem on primes	26
5.2 Dirichlet's theorem on irreducible polynomials	29
5.3 More precise formula and Chebyshev's bias	33
It is as easy as ABC	36
6.1 Polynomial FLT	36
6.2 Polynomial ABC	37
Appendix	40
A.1 Dictionary between integers and polynomials	40

1 Introduction

The goal of this note is to introduce the theory of function fields (mostly over finite fields), which is the analogue of number theory for polynomials. Especially, our main goal is to study various evidences of the following claim:

A theorem that holds for integers is also true for polynomials (over finite fields), and latter is often easier to prove.

There are tons of evidences of this claim. In Section 2, we will see that all the classical results in number theory, such as Wilson's theorem, have their counterparts in the theory of function fields. Another example is the polynomial version of Fermat's Last Theorem, which only requires few pages to prove (Theorem 6.1). Dictionary between the integers and the polynomials over finite fields can be found in Table 1 of Appendix. I hope this note motivates readers to study the theory of function fields, and also to explore other analogies between integers and polynomials.

1.1 Prerequisites

We assume that the readers are familiar with undergraduate level algebra (groups, rings, fields, etc.), number theory (congruences, prime numbers, etc.), and a bit of complex analysis.


1.2 Notations

Let p be a prime number. We denote by \mathbb{F}_p the finite field of order p , which is the field with p elements. We denote the polynomial ring $\mathbb{F}_p[T]$ by A . For each nonzero polynomial $f \in A$, we denote its norm by $|f| = p^{\deg(f)}$, where $\deg(f)$ is the degree of f , and we set $|0| = 0$.

1.3 SageMath

There are some codes in this note, which are mostly written in Sage. Sage is a free open-source mathematics software system, which is built on top of many existing open-source packages and wrapped in a Python interface. You can run them online in SageMathCell, or install it on your computer. Especially, a lot of number-theoretic functions are implemented in Sage, so it is much easier to experiment with it than writing your own code from scratch. For example, to check if a large number is prime, you can simply run

```
is_prime(10 ^ 9 + 7)
```

Some of the exercises in this note are designed to be solved in Sage (or other programming languages that you are familiar with), and those exercises are marked with a Sage logo .

Exercise 1.3.1. Prove that \mathbb{Z} is not a polynomial ring over a field. In other words, show that there is no field k such that $\mathbb{Z} \cong k[T]$ as rings.

Exercise 1.3.2. Think about your favorite theorems in number theory, and try to find their polynomial analogues. Some of them may appear in this note.

2 Basic number theory and their analogues for polynomials

In this section, we will introduce polynomial analogues of the theorems in number theory, including

- Fundamental Theorem of Arithmetic,
- Chinese Remainder Theorem,
- Fermat's Little Theorem and Euler's Theorem,
- Wilson's Theorem.

Not only the statements of the theorems look similar, but you will also see that the proofs often follow the same arguments.

2.1 Fundamental Theorem of Arithmetic

The Fundamental Theorem of Arithmetic states that every integer greater than 1 can be uniquely expressed as a product of prime numbers, up to the order of the factors. More fancier way to say this is that


Theorem 2.1. \mathbb{Z} is a unique factorization domain (UFD).

The standard proof is based on the following implication:

Theorem 2.2. If R is a Euclidean domain (ED), then R is a principal ideal domain (PID), and hence a UFD.

Recall that R is a Euclidean domain if there exists a function $f : R \setminus \{0\} \rightarrow \mathbb{Z}_{\geq 0}$ such that for any $a, b \in R$ with $b \neq 0$, there exist $q, r \in R$ such that $a = bq + r$ and either $r = 0$ or $f(r) < f(b)$. Intuitively, R is a Euclidean domain if we can perform the division with remainder, and the function f is a measure of the size of the elements in R . For any (not necessarily finite) field k , we can also divide a polynomial by another polynomial over k , where $\deg : R \setminus \{0\} \rightarrow \mathbb{Z}_{\geq 0}$ works as the function f . This shows that:

Theorem 2.3. The polynomial ring $k[T]$ is a ED, where k is any field. Hence it is a PID, and hence a UFD.

Exercise 2.1.1.  For each prime $p \leq 100$, determine if the polynomial $T^2 + 1$ is irreducible over \mathbb{F}_p or not. Can you find a pattern?¹

2.2 Chinese Remainder Theorem

The Chinese Remainder Theorem (CRT) states that if n_1, n_2, \dots, n_k are pairwise coprime integers, then the system of congruences

$$x \equiv a_1 \pmod{n_1}, \quad x \equiv a_2 \pmod{n_2}, \quad \dots, \quad x \equiv a_k \pmod{n_k}$$

has a unique solution modulo $N = n_1 n_2 \cdots n_k$.

¹Answer will be given in the next section.

Theorem 2.4 (Chinese Remainder Theorem). Let n_1, n_2, \dots, n_k be pairwise coprime integers and a_1, a_2, \dots, a_k be integers. Then the system of congruences

$$x \equiv a_1 \pmod{n_1}, \quad x \equiv a_2 \pmod{n_2}, \quad \dots, \quad x \equiv a_k \pmod{n_k}$$

has a unique solution modulo $N = n_1 n_2 \cdots n_k$. More precisely, the unique solution is given by

$$x \equiv \sum_{i=1}^k a_i N_i y_i \pmod{N},$$

where $N_i = N/n_i$ and y_i is the multiplicative inverse of N_i modulo n_i , i.e., $N_i y_i \equiv 1 \pmod{n_i}$.

Proof. The proof is essentially hidden in the following isomorphism:

$$\mathbb{Z}/N\mathbb{Z} \cong \mathbb{Z}/n_1\mathbb{Z} \times \mathbb{Z}/n_2\mathbb{Z} \times \cdots \times \mathbb{Z}/n_k\mathbb{Z}.$$

The natural map from the left hand side to the right hand side is given by reducing modulo n_i for each i , and such a map is injective because n_1, n_2, \dots, n_k are pairwise coprime. Since the size of the both sides are equal, the map is also surjective, hence an isomorphism. Finding the solution to the system of congruences is equivalent to finding an element in $\mathbb{Z}/N\mathbb{Z}$ that maps to (a_1, a_2, \dots, a_k) under the isomorphism. It is enough to find solution for the equations

$$x_i \equiv 0 \pmod{n_1}, \quad x_i \equiv 0 \pmod{n_2}, \quad \dots \quad x_i \equiv 1 \pmod{n_i}, \quad \dots, \quad x_i \equiv 0 \pmod{n_k}$$

for each i , then the solution to the original system of congruences is given by the linear combination of the solutions to these equations. Such x_i has to be a multiple of N_i , and if we write $x_i = N_i y_i$, then we have $N_i y_i \equiv 1 \pmod{n_i}$, which means y_i is the multiplicative inverse of N_i modulo n_i . Such y_i exists and can be found by the Euclidean Algorithm, since N_i and n_i are coprime. \square

For example, consider a system of congruences

$$\begin{cases} x \equiv 2 \pmod{3} \\ x \equiv 3 \pmod{5} \end{cases}$$

From Euclidean Algorithm, we can write $1 = 2 \cdot 3 - 1 \cdot 5$, which means $5^{-1} \equiv (-1) \pmod{3}$ and $3^{-1} \equiv 2 \pmod{5}$. Then the formula in Theorem 2.4 gives us $x \equiv 2 \cdot 5 \cdot (-1) + 3 \cdot 3 \cdot 2 \pmod{15}$, which simplifies to $x \equiv -10 + 18 \equiv 8 \pmod{15}$.

What is a polynomial analogue of the Chinese Remainder Theorem? We simply replace integers with polynomials over a finite field, and we get the following theorem.

Theorem 2.5 (Chinese Remainder Theorem for Polynomials). Let $g_1(T), g_2(T), \dots, g_k(T) \in \mathbb{F}_p[T]$ be pairwise coprime polynomials and $a_1(T), a_2(T), \dots, a_k(T) \in \mathbb{F}_p[T]$ be polynomials. Then the system of congruences

$$f \equiv a_1 \pmod{g_1}, \quad f \equiv a_2 \pmod{g_2}, \quad \dots, \quad f \equiv a_k \pmod{g_k}$$

has a unique solution modulo $G = g_1 g_2 \cdots g_k$. More precisely, the unique solution is given by

$$f \equiv \sum_{i=1}^k a_i G_i b_i \pmod{G},$$

where $G_i = G/g_i$ and b_i is the multiplicative inverse of G_i modulo g_i , i.e., $G_i b_i \equiv 1 \pmod{g_i}$.

Using the exact same method, we can solve a system of congruences. For example, consider

$$\begin{cases} f \equiv 1 \pmod{T} \\ f \equiv T \pmod{T^2 + 1} \end{cases}$$

We can write $1 = (-T) \cdot T + 1 \cdot (T^2 + 1)$ (hence T and $T^2 + 1$ are coprime), which means $(T^2 + 1)^{-1} \equiv 1 \pmod{T}$ and $T^{-1} \equiv -T \pmod{T^2 + 1}$. Then the formula in Theorem 2.5 gives us $f \equiv 1 \cdot (T^2 + 1) \cdot 1 + T \cdot T \cdot (-T) \pmod{(T^2 + 1)T}$, which simplifies to $f \equiv T^2 + 1 - T^3 \equiv T^2 + T + 1 \pmod{(T^2 + 1)T}$.

Exercise 2.2.1. sage

1. Find a polynomial $f \in \mathbb{R}_3[T]$ such that²

$$\begin{cases} f \equiv T + 2 \pmod{T^2 + 1} \\ f \equiv T^2 + T \pmod{T^3 - T^2 + T + 2} \end{cases}$$

2. Find a polynomial $f \in \mathbb{R}_7[T]$ such that³

$$\begin{cases} f \equiv 5T^2 + 3T + 6 \pmod{T^3 + 2T^2 + 3T + 2} \\ f \equiv T^3 + 1 \pmod{T^4 + 3T^3 + 2T^2 + 1} \\ f \equiv T^4 - T + 2 \pmod{T^2 + 3T + 1} \end{cases}$$

2.3 Fermat's Little Theorem and Euler's Theorem

Ferma's *Little* (not last!) Theorem states the following:

Theorem 2.6 (Fermat's Little Theorem). Let p be a prime number and a an integer not divisible by p . Then

$$a^{p-1} \equiv 1 \pmod{p}.$$

Here is a proof using group theory.

Proof. Consider the group $G = (\mathbb{Z}/p\mathbb{Z})^\times$. The order of G is $p - 1$, and the order of the subgroup generated by a is a divisor of $p - 1$. Thus, by Lagrange's theorem, we have $a^{p-1} \equiv 1 \pmod{p}$, as desired. \square

Euler's theorem is a generalization of Fermat's Little Theorem, which considers general moduli. To state Euler's theorem, we need to define the *Euler's totient function* $\varphi(n)$, which counts the number of integers from 1 to n that are coprime to n .

Theorem 2.7 (Euler's Theorem). Let n be a positive integer and a an integer coprime to n . Then

$$a^{\varphi(n)} \equiv 1 \pmod{n}.$$

Proof. Proof is similar to the proof of Theorem 2.6, where we consider the group $G = (\mathbb{Z}/n\mathbb{Z})^\times$ which has order $\varphi(n)$. \square

²Answer: $f = T^4 + T^2 + T + 2$

³Answer: $f = T^5 + 3T^2 + 2T + 1$

One may ask how to compute $\varphi(n)$. By CRT again (but for unit groups), we have an isomorphism

$$(\mathbb{Z}/n\mathbb{Z})^\times \cong (\mathbb{Z}/p_1^{k_1}\mathbb{Z})^\times \times (\mathbb{Z}/p_2^{k_2}\mathbb{Z})^\times \times \cdots \times (\mathbb{Z}/p_m^{k_m}\mathbb{Z})^\times,$$

where $n = p_1^{k_1} p_2^{k_2} \cdots p_m^{k_m}$ is the prime factorization of n . Thus, we can compute $\varphi(n)$ as

$$\varphi(n) = \varphi(p_1^{k_1}) \varphi(p_2^{k_2}) \cdots \varphi(p_m^{k_m}),$$

hence we only need to compute $\varphi(p^k)$ for a prime p and a positive integer k . This counts the number of integers from 1 to p^k that are coprime to p^k , or equivalently multiples of p , which is $p^k - p^{k-1} = p^{k-1}(p - 1)$. This gives us the formula

Theorem 2.8. For an integer $n = p_1^{k_1} p_2^{k_2} \cdots p_m^{k_m}$, we have

$$\varphi(n) = n \prod_{i=1}^m \left(1 - \frac{1}{p_i}\right). \quad (1)$$

What is a polynomial analogue of Fermat's Little Theorem and Euler's Theorem? Based on Table 1, prime will be replaced by irreducible polynomial. Unfortunately, it does not make sense to take exponentiation of a polynomial by another polynomial. However, since $\varphi(n)$ was defined as a size of the group $(\mathbb{Z}/n\mathbb{Z})^\times$, we can define a polynomial analogue of $\varphi(g)$ as the size of the group $(A/gA)^\times$. Especially, if g is irreducible, then $(A/gA)^\times = (A/gA) \setminus \{0\}$, which has size $|g| - 1$. The exact same argument as the proof of Theorem 2.6 and Theorem 2.7 works for polynomials, and we get the following theorems.

Theorem 2.9 (Fermat's Little Theorem for Polynomials). Let $f, g \in A$ be polynomials, where g is irreducible and f is not divisible by g . Then

$$f^{|g|-1} \equiv 1 \pmod{g}.$$

Theorem 2.10 (Euler's Theorem for Polynomials). Let $f, g \in A$ be coprime polynomials (g is not necessarily irreducible). Then

$$f^{\varphi(g)} \equiv 1 \pmod{g}.$$

The formula (1) also generalizes to polynomials.

Theorem 2.11. Let $g(T) = p_1^{k_1} p_2^{k_2} \cdots p_m^{k_m}$ be a polynomial in A , where p_1, p_2, \dots, p_m are distinct irreducible polynomials in A . Then

$$\varphi(g) = |g| \prod_{i=1}^m \left(1 - \frac{1}{|p_i|}\right).$$

Exercise 2.3.1. There are several different proofs of Theorem 2.6, e.g. see [wikipedia page](#). Choose your favorite argument (other than the one used in Theorem 2.6) and try to generalize it to prove 2.9.⁴

Exercise 2.3.2. Prove the original version of Fermat's little theorem from the polynomial version.⁵

⁴For example, there's a proof using induction on a . Can you generalize it to polynomials?

⁵Hint: for a prime p and an integer a not divisible by p , consider $f(T) = T + a$ and $g(T) = T$.

2.4 Wilson's Theorem

Another interesting theorem on prime numbers is Wilson's theorem:

Theorem 2.12 (Wilson's Theorem). Let p be a prime number. Then

$$(p-1)! \equiv -1 \pmod{p}.$$

Proof. Consider the group $G = (\mathbb{Z}/p\mathbb{Z})^\times$. The left hand side is the product of all elements in G . Now, we can pair up each element $a \in G$ with its inverse a^{-1} , except for the case when $a = a^{-1}$, which happens if and only if $a^2 \equiv 1 \pmod{p} \Leftrightarrow a \equiv \pm 1 \pmod{p}$. Thus the product of all elements in G is $\equiv 1 \cdot (-1) \equiv -1 \pmod{p}$, as desired. \square

What is a polynomial analogue of Wilson's theorem? Note that the left hand side of Wilson's theorem is the product of all nonzero elements in \mathbb{F}_p , so it might be reasonable to define "factorial" $(g-1)!$ of an irreducible polynomial $g(T) \in \mathbb{F}_p[T]$ as the product of all elements in $(A/gA)^\times$.


Theorem 2.13 (Wilson's Theorem for Polynomials). Let p be a prime number and $A = \mathbb{F}_p[T]$ the polynomial ring over the finite field \mathbb{F}_p . Let $g(T) \in A$ be an irreducible polynomial of degree d . Then

$$\prod_{0 \leq \deg(f) < d} f \equiv -1 \pmod{g}.$$

Note that the left hand side only depends on the degree of $g(T)$. Especially, $(\text{LHS}) + 1$ is divisible by any irreducible polynomial $g(T)$ of degree d .

Exercise 2.4.1. Prove Theorem 2.13.

Exercise 2.4.2. Prove the original version of Wilson's theorem from the polynomial version.⁶

Exercise 2.4.3.  For each p and d , compute $\prod_{0 \leq \deg(f) < d} f + 1$. Can you find any pattern?

Exercise 2.4.4. It is known that the following converse of Wilson's theorem holds: if a natural number n satisfies $(n-1)! \equiv -1 \pmod{n}$, then n is a prime number. We can consider a polynomial analogue of this converse. Here are two suggestions:

1. If a polynomial $g(T) \in \mathbb{F}_p[T]$ of degree d satisfies

$$\prod_{\substack{0 \leq \deg(f) \leq d \\ \gcd(f, g) = 1}} f \equiv -1 \pmod{g},$$

then $g(T)$ is irreducible over \mathbb{F}_p . Here the left hand side is the product of all elements in $(A/gA)^\times$.

2. If a polynomial $g(T) \in \mathbb{F}_p[T]$ of degree d satisfies

$$\prod_{\substack{0 \leq \deg(f) \leq d \\ f \neq 0}} f \equiv -1 \pmod{g},$$

then $g(T)$ is irreducible over \mathbb{F}_p .

Prove or disprove these claims.

⁶Hint: consider $g(T) = T$.

3 Reciprocity Laws

Quadratic reciprocity is one of the most important theorem in number theory, which is the origin of all modern number theory including algebraic number theory and Langlands program. In this section, we will see the analogue of quadratic reciprocity for function fields.

3.1 Quadratic Reciprocity for Integers

CRT tells us how to solve *linear* congruences, i.e. equations of the form $ax + b \equiv 0 \pmod{n}$. Now, let's go one step further and consider *quadratic* congruences, i.e. equations of the form $ax^2 + bx + c \equiv 0 \pmod{n}$. How can we solve such equations? Recall that we solve usual quadratic equations by “completing the square”, and we can do the same for congruences.⁷ Hence we can reduce to a simpler question:

For given integer a and n , when does the congruence $x^2 \equiv a \pmod{n}$ have a solution?

Although the question seems pretty simple, the answer is quite deep and also interesting. To get some intuition, let's first consider the case when $n = p$ is a prime number and $a = -1$. In other words, we want to know when the congruence $x^2 \equiv -1 \pmod{p}$ has a solution (this is equivalent to one of the exercises in Section 2). Let's just try some small primes. Here's a Sage code to check the congruence $x^2 \equiv -1 \pmod{p}$ for primes $p < 100$:

```
with_sol = []
without_sol = []
for p in primes(100):
    for a in range(p):
        if (a ^ 2 + 1) % p == 0:
            with_sol.append(p)
            break
        if a == p - 1:
            without_sol.append(p)
print("Primes with solution", with_sol)
print("Primes without solution", without_sol)
```

Here's the output of the code you should get:

```
Primes with solution [2, 5, 13, 17, 29, 37, 41, 53, 61, 73, 89, 97]
Primes without solution [3, 7, 11, 19, 23, 31, 43, 47, 59, 67, 71, 79, 83]
```

Stare the lists carefully for a while. What do you notice? The answer is:

Theorem 3.1. Let p be a prime number. Then the congruence $x^2 \equiv -1 \pmod{p}$ has a solution if and only if $p = 2$ or $p \equiv 1 \pmod{4}$.

Proof. The case $p = 2$ is trivial, so we assume p is odd. Assume that the equation has a solution x . Then $x^4 \equiv (-1)^2 \equiv 1 \pmod{p}$, so x is an element of order 4 in $(\mathbb{Z}/p\mathbb{Z})^\times$. By Lagrange's theorem, the order of x must divide $p - 1$, so $4 \mid (p - 1) \Leftrightarrow p \equiv 1 \pmod{4}$.

⁷when n is odd and a is invertible modulo n .

Conversely, assume $p \equiv 1 \pmod{4}$ and let $g \in (\mathbb{Z}/p\mathbb{Z})^\times$ be a generator of the group. Then $g^{(p-1)/4}$ is an element of order 4, so we can set $x = g^{(p-1)/4}$ and $(x^2)^2 \equiv x^4 \equiv 1 \pmod{p}$. Since $x^2 \not\equiv 1 \pmod{p}$, we have $x^2 \equiv -1 \pmod{p}$. \square

It is amazing that the solvability of the congruence equation is simply given by a congruence condition on the prime p .

Let's try other a , say $a = 3$. Here's a Sage code to check the congruence $x^2 \equiv 3 \pmod{p}$ for primes $p < 100$:

```
with_sol = []
without_sol = []
for p in primes(100):
    for a in range(p):
        if (a ^ 2 - 3) % p == 0:
            with_sol.append(p)
            break
    if a == p - 1:
        without_sol.append(p)
print("Primes with solution", with_sol)
print("Primes without solution", without_sol)
```

Here's the output of the code you should get:

```
Primes with solution [2, 3, 11, 13, 23, 37, 47, 59, 61, 71, 73, 83, 97]
Primes without solution [5, 7, 17, 19, 29, 31, 41, 43, 53, 67, 79, 89]
```

Again, the question is: what do you notice? The answer is slightly more complicated than the previous case, and you need to observe the numbers modulo 12.

Theorem 3.2. Let p be a prime number. Then the congruence $x^2 \equiv 3 \pmod{p}$ has a solution if and only if $p = 2, 3$ or $p \equiv 1, 11 \pmod{12}$.

To answer the question for general a , we need to introduce the Legendre symbol.

Definition 3.3 (Legendre symbol). Let p be a prime and a be an integer. The Legendre symbol $\left(\frac{a}{p}\right)$ is defined as follows:

$$\left(\frac{a}{p}\right) = \begin{cases} 0 & a \equiv 0 \pmod{p} \\ 1 & \text{if } x^2 \equiv a \pmod{p} \text{ has a solution} \\ -1 & \text{otherwise.} \end{cases} \quad (2)$$

Do not confuse it with fraction. By definition, it tells you whether the congruence $x^2 \equiv a \pmod{p}$ has a solution or not. But it does not tell you how to compute the symbol. The following theorem gives a way to compute the Legendre symbol for odd primes.

Theorem 3.4 (Euler's criterion). Let p be an odd prime and a be an integer. Then

$$\left(\frac{a}{p}\right) \equiv a^{\frac{p-1}{2}} \pmod{p}. \quad (3)$$

Proof. If $a \equiv 0 \pmod{p}$, then the result is trivial. Recall that the group \mathbb{F}_p^\times is cyclic of order $p-1$. Let g be a generator of the group. Then every nonzero element $a \in \mathbb{F}_p^\times$ can be written as $a \equiv g^k \pmod{p}$ for some integer k . The equation is solvable if and only if k is even modulo $p-1$. \square

Theorem 3.5 (Quadratic reciprocity). Let p and q be distinct odd prime numbers. Then

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{\frac{(p-1)(q-1)}{4}}. \quad (4)$$

Especially, the congruence $x^2 \equiv p \pmod{q}$ has a solution if and only if the congruence $x^2 \equiv q \pmod{p}$ has a solution, except when $p \equiv q \equiv 3 \pmod{4}$.

There are tons of proofs of this theorem, and we will not give a proof here. You can refer to an archive of proofs by Lemmermeyer [19], or a book by Baumgart [3].

We also need the following supplementary results to compute the Legendre symbol completely.

Theorem 3.6. For odd prime p , we have

$$\begin{aligned} \left(\frac{-1}{p}\right) &= (-1)^{\frac{p-1}{2}} = \begin{cases} 1 & p \equiv 1 \pmod{4} \\ -1 & p \equiv 3 \pmod{4} \end{cases} \\ \left(\frac{2}{p}\right) &= (-1)^{\frac{p^2-1}{8}} = \begin{cases} 1 & p \equiv 1, 7 \pmod{8} \\ -1 & p \equiv 3, 5 \pmod{8} \end{cases} \end{aligned}$$

Proof. We already proved the first in Theorem 3.1. For the second here is a proof brought from [32, Theorem]. Let $\alpha \in \mathbb{F}_{p^2}$ be a primitive 8-th root of unity (note that $p^2 - 1$ is divisible by 8 for any odd p). then $\alpha^4 = -1$ and $y = \alpha + \alpha^{-1}$ satisfies $y^2 = 2 + \alpha^2 + \alpha^{-2} = 2$. By binomial expansion, we have $y^p = \alpha^p + \alpha^{-p}$ modulo p , hence $y^p = y$ for $p \equiv \pm 1 \pmod{8}$ and this gives $\left(\frac{2}{p}\right) = 1$ for such p . If $p \equiv \pm 3 \pmod{8}$, then $y^p = -y \Leftrightarrow y^{p-1} = -1$, hence $\left(\frac{2}{p}\right) = y^{p-1} = -1$.⁸ \square

We can use quadratic reciprocity to prove Theorem 3.2.

Proof of Theorem 3.2. By Theorem 3.5, we have

$$\left(\frac{3}{p}\right)\left(\frac{p}{3}\right) = (-1)^{\frac{(p-1)(3-1)}{4}} = (-1)^{\frac{(p-1)}{2}} \Leftrightarrow \left(\frac{3}{p}\right) = (-1)^{\frac{(p-1)}{2}} \left(\frac{p}{3}\right).$$

We can compute each factor on the right-hand side separately, which are

$$\begin{aligned} \left(\frac{p}{3}\right) &= \begin{cases} 1 & p \equiv 1 \pmod{3} \\ -1 & p \equiv 2 \pmod{3} \end{cases}, \\ (-1)^{\frac{(p-1)}{2}} &= \begin{cases} 1 & p \equiv 1 \pmod{4} \\ -1 & p \equiv 3 \pmod{4} \end{cases}. \end{aligned}$$

By combining these two congruences, we can see that


$$\left(\frac{3}{p}\right) = \begin{cases} 1 & p \equiv 1, 11 \pmod{12} \\ -1 & p \equiv 5, 7 \pmod{12} \end{cases}$$

for $p \neq 2, 3$. \square

Exercise 3.1.1. When $p \equiv 1 \pmod{4}$, we know that the congruence $x^2 \equiv -1 \pmod{p}$ has a solution by Theorem 3.1. Show that $x = \left(\frac{p-1}{2}\right)!$ becomes a solution.⁹

⁸It may look random to consider such y , but this can be thought as an analogue of the identity of complex numbers $\sqrt{2} = e^{2\pi i/8} + e^{-2\pi i/8}$ in finite fields.

⁹Hint: Wilson's theorem.

Exercise 3.1.2.  Compute the Legendre symbol $\left(\frac{563}{661}\right)$, by

1. brute-force checking all integers x modulo 661, or
2. using Euler's criterion, or
3. using quadratic reciprocity, or
4. using kronecker in Sage.


Exercise 3.1.3. Prove Theorem 3.2 without using quadratic reciprocity, by adapting the proof of Theorem 3.6. Here's a sketch of a proof:

1. Show that, for any prime $p > 3$, there is a 12-th root of unity β in \mathbb{F}_{p^2} .
2. Show that $z = \beta + \beta^{-1}$ satisfies $z^2 = 3$.
3. Show that $z^p = z$ if and only if $p \equiv 1, 11 \pmod{12}$, and $z^p = -z$ if and only if $p \equiv 5, 7 \pmod{12}$. Conclude that $\left(\frac{3}{p}\right) = 1$ if and only if $p \equiv 1, 11 \pmod{12}$.

It is natural to ask if one can generalize this proof to other primes. This essentially gives a proof of quadratic reciprocity, and the correct analogue of y or z would be the *Gauss sum*.

Exercise 3.1.4. Characterize the prime p such that the congruence $x^2 \equiv 5 \pmod{p}$ has a solution.¹⁰

Exercise 3.1.5. For a prime p and integers a, b, c , give a necessary and sufficient condition for the congruence $ax^2 + bx + c \equiv 0 \pmod{p}$ to have a solution.

Exercise 3.1.6. 

1. Find all prime $p < 100$ where the congruence $x^3 - 3x + 1 \equiv 0 \pmod{p}$ has a solution. Can you find a pattern? If you are brave enough, try to prove it.¹¹
2. Find all prime $p < 100$ where the congruence $x^3 - x - 1 \equiv 0 \pmod{p}$ has a solution. Can you find a pattern?
3. Consider a power series

$$A(T) = T \prod_{n \geq 1} (1 - T^n)(1 - T^{23n}) = \sum_{n \geq 1} a_n T^n \in \mathbb{Z}[[T]].$$

Compute the first 100 coefficients of $A(T)$, and check if there's any relation between the p -th coefficients a_p and the congruence $x^3 - x - 1 \equiv 0 \pmod{p}$.¹²

3.2 Quadratic Reciprocity for Polynomials

Let's move on to the polynomial case. We are going to ask the same question as before:

¹⁰Hint: Use quadratic reciprocity. Answer will be given modulo 5.

¹¹Hint: modulo 9. To prove it, we notice the fact that the roots of $x^3 - 3x + 1$ are $2 \cos(2\pi/9), 2 \cos(4\pi/9), 2 \cos(8\pi/9)$, by double angle formula. Then try to express them in terms of 9-th roots of unity.

¹²Answer: it has a solution in \mathbb{F}_p if and only if $a_p \in \{0, 2\}$, when $p \neq 23$. Buzzwords: Artin representation and modular form of weight 1, Nonabelian class field theory.

For given polynomial $f \in A$ and irreducible polynomial P , when does the congruence $X^2 \equiv f \pmod{P}$ have a solution?

We are going to adapt the exact same strategy as before. **For this section, we assume that p is an odd prime.** Characteristic 2 case is slightly more complicated, and we will not discuss it here. You can refer to Keith Conrad's note [7] for more details for $p = 2$.

Define the *Legendre symbol* for polynomials as follows:

Definition 3.7 (Legendre symbol for polynomials). Let $P \in A$ be an irreducible polynomial and $f \in A$ be a polynomial. The *Legendre symbol* $\left(\frac{f}{P}\right)$ is defined as follows:

$$\left(\frac{f}{P}\right) = \begin{cases} 0 & f \equiv 0 \pmod{P} \\ 1 & \text{if } X^2 \equiv f \pmod{P} \text{ has a solution in } A, \\ -1 & \text{otherwise.} \end{cases} \quad (5)$$

We have Euler's criterion for polynomials as well.

Theorem 3.8 (Euler's criterion for polynomials). Let $P \in A$ be an irreducible polynomial and $f \in A$ be a polynomial. Then

$$\left(\frac{f}{P}\right) = f^{\frac{|P|-1}{2}} \pmod{P}. \quad (6)$$

Exercise 3.2.1. Prove Theorem 3.8 by adapting the proof of Theorem 3.4.¹³

One interesting corollary of Euler's criterion is the following:

Corollary 3.9. Let $P \in A$ be an irreducible polynomial of even degree. Then $\left(\frac{c}{P}\right) = 1$ for any nonzero constant $c \in \mathbb{F}_p$.

Proof. Write

$$\frac{|P|-1}{2} = \frac{p^{\deg P} - 1}{2} = (p-1) \cdot \frac{(p^{\deg P-1} + p^{\deg P-2} + \dots + 1)}{2}$$

where the second factor of the last equation is an integer since $\deg P$ is even and p is odd. By Euler's criterion and $c^{p-1} = 1$, we have

$$\left(\frac{c}{P}\right) = c^{\frac{|P|-1}{2}} = c^{(p-1) \cdot \frac{(p^{\deg P-1} + p^{\deg P-2} + \dots + 1)}{2}} = 1.$$

□

Exercise 3.2.2. Let $p = 7$ and $P(T) = T^2 + T + 3$, which is irreducible modulo 7. Find a polynomial $f \in \mathbb{F}_7[T]$ such that $f^2 \equiv 3 \pmod{P}$.

Theorem 3.10 (Quadratic reciprocity for polynomials). Let $P, Q \in A$ be distinct irreducible polynomials. Then

$$\left(\frac{P}{Q}\right) \left(\frac{Q}{P}\right) = (-1)^{\frac{|P|-1}{2} \cdot \frac{|Q|-1}{2}}. \quad (7)$$

¹³Hint: Use the fact that A/P is a finite field of order $|P| = p^{\deg P}$.

Proof. This proof is due to Carlitz, which can be found in Rosen's book [26, p. 25, Theorem 3.3] or Conrad's note [8, Section 3]. Let k/\mathbb{F}_p be a (finite) field contains all roots of P and Q . Let $d_P = \deg P$ and $d_Q = \deg Q$. We will first show that (7) is equivalent to:

$$\left(\frac{P}{Q}\right) = (-1)^{\frac{(p-1)d_P d_Q}{2}} \left(\frac{Q}{P}\right). \quad (8)$$

From $|P| = p^{d_P}$,

$$\frac{|P| - 1}{2} = \frac{p^{d_P} - 1}{2} = \frac{p - 1}{2} \cdot (p^{d_P-1} + p^{d_P-2} + \dots + 1) \equiv \frac{(p-1)d_P}{2} \pmod{2},$$

since p is odd. Similar congruence holds for Q , hence

$$\frac{|P| - 1}{2} \cdot \frac{|Q| - 1}{2} \equiv \frac{(p-1)d_P}{2} \cdot \frac{(p-1)d_Q}{2} \equiv \frac{p-1}{2} \cdot d_P d_Q \pmod{2}.$$

Note that $n^2 \equiv n \pmod{2}$ for any integer n .

Now, let $\alpha, \beta \in k$ be roots of P and Q , respectively. Then all the zeros of P and Q are given by the p -powers of α and β (images of Frobenius map): they factor as

$$P(T) = \prod_{i=0}^{d_P-1} (T - \alpha^{p^i}), \quad Q(T) = \prod_{j=0}^{d_Q-1} (T - \beta^{p^j}).$$

By Euler's criterion for polynomials and properties of finite fields, we have

$$\begin{aligned} \left(\frac{P}{Q}\right) &\equiv P(T)^{\frac{|Q|-1}{2}} \pmod{Q(T)} \\ &\equiv P(T)^{\frac{p-1}{2}(1+p+\dots+p^{d_Q-1})} \pmod{Q(T)} \\ &\equiv (P(T)P(T)^p \dots P(T)^{p^{d_Q-1}})^{\frac{p-1}{2}} \pmod{Q(T)} \\ &\equiv (P(T)P(T^p) \dots P(T^{p^{d_Q-1}}))^{\frac{p-1}{2}} \pmod{Q(T)} \end{aligned}$$

If we set $T = \beta$ for the last congruence, it becomes

$$\prod_{i=0}^{d_P-1} \prod_{j=0}^{d_Q-1} (\beta^{p^j} - \alpha^{p^i})^{\frac{p-1}{2}} \in \{\pm 1\}$$

Now we can swap α and β , which gives $\left(\frac{Q}{P}\right)$ up to a sign

$$(-1)^{\frac{(p-1)d_P d_Q}{2}}$$

which is exactly matches with the exponent of (-1) in (8). \square

Note that the proof is very specific for polynomials over finite fields, and it does not work for integer case. There are *some* proofs of similar taste, which writes the product $\left(\frac{p}{q}\right)\left(\frac{q}{p}\right)$ as other function that is symmetric in p and q .


One can generalize it to higher order reciprocity laws. Since the proof is very similar to the quadratic case, we only state the result here. Details can be found in [26, Chapter 3].

Theorem 3.11 (d -th power reciprocity law). Let $d \geq 2$ be an integer dividing $p - 1$. For $f \in A$ and irreducible polynomials $P \in A$, define $\left(\frac{f}{P}\right)$ be a unique element of \mathbb{F}_p^\times such that

$$\left(\frac{f}{P}\right)_d = f^{\frac{|P|-1}{d}} \pmod{P}. \quad (9)$$

Then for distinct irreducible polynomials $P, Q \in A$, we have

$$\left(\frac{P}{Q}\right)_d = (-1)^{\frac{p-1}{d} \deg P \deg Q} \left(\frac{Q}{P}\right)_d. \quad (10)$$

Exercise 3.2.3.  Decide if the following congruence equations in X have a solution.

1. $X^2 = T^2 + 3T + 3 \pmod{T^3 + T + 1}$ in $\mathbb{F}_5[T]$.
2. $X^2 \equiv 6T^2 + T + 1 \pmod{T^3 + T^2 + T + 5}$ in $\mathbb{F}_7[T]$.
3. $X^2 + TX + T + 1 \equiv 0 \pmod{T^3 + T^2 - 1}$ in $\mathbb{F}_3[T]$.

Exercise 3.2.4. Let p be an odd prime and $P \in A$ be an irreducible polynomial of degree d . Show that $X^2 \equiv -1 \pmod{P}$ has a solution if and only if d is even or $p \equiv 1 \pmod{4}$. Can you give an explicit formula for a solution in terms of p and P ?

Exercise 3.2.5. For an irreducible polynomial $P \in A$ and polynomials $f, g, h \in A$, give a necessary and sufficient condition for the congruence $fX^2 + gX + h \equiv 0 \pmod{P}$ to have a solution.

Exercise 3.2.6. Let $f \in A$ be a nonsquare polynomial. Show that $\left(\frac{f}{P}\right) = -1$ for infinitely many irreducible polynomial $P \in A$.¹⁴

¹⁴See [8, Theorem 4.5].

4 How many primes?

In this section, we will study several questions on counting numbers and polynomials, e.g. number of prime numbers / irreducible polynomials up to certain bounds. Over integers, many of such problems are wide open. However, the polynomial analogue of these problems are often much easier to prove. Especially, both of the sides uses certain *zeta functions* or *L-functions* to count certain objects, which becomes much simpler in the polynomial case.

4.1 Prime number theorem and Riemann zeta function

We start with the most important fact about prime numbers, proven by Euclid few hundred years ago.

Theorem 4.1 (Euclid). There are infinitely many prime numbers.

Proof. Suppose that there are finitely many prime numbers p_1, p_2, \dots, p_n . Consider the number $N = p_1 p_2 \cdots p_n + 1$. Then N is not divisible by any of the prime numbers p_i , so it must be either prime or divisible by some other prime number. This contradicts our assumption that there are only finitely many prime numbers. \square

Once we know that there are infinitely many prime numbers, we can ask the following question:

For given real number $x > 0$, how many prime numbers are there less than or equal to x ?


We follow the standard notation and denote the number of prime numbers less than or equal to x by $\pi(x)$. Hadamard and de la Vallée-Poussin proved independently in 1896 that

Theorem 4.2 (Prime number theorem). We have

$$\pi(x) \sim \frac{x}{\log x},$$

where \sim means that the ratio of the two sides tends to 1 as x tends to infinity.

In other words, the density of prime numbers among integers is approximately $\frac{1}{\log x}$, which tends to 0 as x tends to infinity.

Exercise 4.1.1.  Plot the graphs of the functions $\pi(x)$ and $\frac{x}{\log x}$ for $x \in [1, 10^6]$.¹⁵

Although there are some elementary proofs by Selberg [31] and Erdős [14], the original proofs and the most of other proofs (e.g. Newman's proof [23, 39]) uses *Riemann zeta function*. In [25], Riemann defined the zeta function:

Definition 4.3 (Riemann zeta function). The *Riemann zeta function* is defined by

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$$

for complex numbers s with $\Re(s) > 1$.

¹⁵Hint: Sage has `prime_pi`.

Exercise 4.1.2. Why it converges for $\Re(s) > 1$?

By writing $\zeta(s)$ as an integral, Riemann showed that it can be analytically (more precisely, meromorphically) continued to the whole complex plane, except for a simple pole at $s = 1$.

Theorem 4.4. 1. $\zeta(s)$ can be analytically continued to a meromorphic function on the whole complex plane, with a simple pole at $s = 1$ with residue 1.

2. Define the *completed zeta function* as

$$\xi(s) = \pi^{-\frac{s}{2}} \Gamma\left(\frac{s}{2}\right) \zeta(s), \quad (11)$$

where $\Gamma(s)$ is the gamma function. Then $\xi(1-s) = \xi(s)$ for all $s \in \mathbb{C} \setminus \{0, 1\}$.

Proof. See Apostol [1] or Stein–Shakarchi [34]. □

This allow us to state the infamous *Riemann Hypothesis*:

Conjecture 4.5 (Riemann Hypothesis). All *non-trivial* zeros of $\zeta(s)$ lie on the critical line $\Re(s) = \frac{1}{2}$.

Here the *trivial* zeros are the negative even integers, i.e. $s = -2, -4, -6, \dots$

But why Riemann zeta function is related to prime numbers? The following theorem by Euler is the key.

Theorem 4.6 (Euler factorization). The Riemann zeta function can be expressed as an infinite product over all prime numbers p :

$$\zeta(s) = \prod_p \frac{1}{1 - p^{-s}}, \quad (12)$$

for $\Re(s) > 1$.

Proof. This essentially follows from the fundamental theorem of arithmetic (Theorem 2.1). More precisely, using geometric series we can rewrite the right hand side as

$$\prod_p \frac{1}{1 - p^{-s}} = \prod_p \left(1 + \frac{1}{p^s} + \frac{1}{p^{2s}} + \dots \right)$$

and Theorem 2.1 implies that each term $\frac{1}{n^s}$ in the left hand side in (12) appears exactly once in the right hand side. □

Let's return to the prime number theorem. The main idea of the proof(s) is in two steps:

1. Reduce the problem to the following claim: $\zeta(s)$ is non-vanishing for $\Re(s) = 1$ and $s \neq 1$.
2. Prove the above claim.

Here we give a sketch of the proof of 2. By taking logarithmic derivative of (12), we get

$$\log \zeta(s) = - \sum_p \log(1 - p^{-s}) = \sum_p \sum_{k \geq 1} \frac{p^{-ks}}{k}.$$

If $s = x + iy$, we get

$$|\zeta(x + iy)| = \exp \left(\sum_p \sum_{k \geq 1} \frac{\cos(ky \log p)}{kp^{kx}} \right)$$

and

$$|\zeta(x)^3 \zeta(x + iy)^4 \zeta(x + 2iy)| = \exp \left(\sum_p \sum_{k \geq 1} \frac{3 + 4 \cos(ky \log p) + \cos(2ky \log p)}{kp^{kx}} \right) \quad (13)$$

for all $x > 1$. Now, one can prove the inequality

$$3 + 4 \cos \theta + \cos(2\theta) \geq 0 \quad (14)$$

for any $\theta \in \mathbb{R}$, which shows that the right hand side of (13) is at least 1. Assume that $\zeta(1 + iy) = 0$ for some y . Then one can derive a contradiction from (13) by considering the limit as $x \rightarrow 1^+$, where the left hand side tends to 0 while the right hand is bounded below by 1.

Exercise 4.1.3. Prove (14).

So we can approximate $\pi(x)$ by $\frac{x}{\log x}$, but how good is this approximation? More precisely, what is the growth of the error term? First of all, there's a better approximation than $\frac{x}{\log x}$ suggested by Dirichlet, which is *logarithmic integral*:

$$\text{Li}(x) := \int_2^x \frac{dt}{\log t}. \quad (15)$$

In terms of growth, $\frac{x}{\log x}$ and $\text{Li}(x)$ are asymptotically equivalent (see Exercise 4.1.4 below). However, $\text{Li}(x)$ gives much better approximation than $\frac{x}{\log x}$. In [12], de la Vallée-Poussin proved that the error term $\pi(x) - \text{Li}(x)$ is bounded by

$$|\pi(x) - \text{Li}(x)| = O(xe^{-c\sqrt{\log x}})$$


for some constant $c > 0$. Koch [37] showed that, *assuming the Riemann Hypothesis*, we obtain the improved bound

$$|\pi(x) - \text{Li}(x)| = O\left(\sqrt{x} \log x\right) \quad (16)$$

which in fact turned out to be equivalent to the Riemann Hypothesis.

Exercise 4.1.4. Prove that $\text{Li}(x) \sim \frac{x}{\log x}$ as $x \rightarrow \infty$. More precisely, show that

$$\text{Li}(x) \sim \frac{x}{\log x} \sum_{k \geq 0} \frac{k!}{(\log x)^k} = \frac{x}{\log x} + \frac{x}{(\log x)^2} + \frac{2x}{(\log x)^3} + \dots \quad (17)$$

Exercise 4.1.5.  Plot the graphs of the functions $\pi(x)$, $\text{Li}(x)$, and $\frac{x}{\log x}$ for $x \in [1, 10^6]$, and observe that $\text{Li}(x)$ is much closer to $\pi(x)$ than $\frac{x}{\log x}$. Also, compare $\pi(x)$ and $\text{Li}(x)$ - do you have any conjecture on the difference?¹⁶

¹⁶You may find that $\pi(x) < \text{Li}(x)$ when x is not too small, and it is tempting to conjecture that $\pi(x) < \text{Li}(x)$ for sufficiently large x . However, this is not true. In fact, Littlewood [20] proved that the sign of $\pi(x) - \text{Li}(x)$ changes infinitely many times, although the smallest number $x > 10$ such that $\pi(x) < \text{Li}(x)$ (known as Skewes's number) is probably very huge.

4.2 Counting irreducible polynomials with zeta function

How about polynomials? First of all, there are infinitely many irreducible polynomials:

Theorem 4.7. For each prime p , there are infinitely many irreducible polynomials over \mathbb{F}_p .

Exercise 4.2.1. Prove Theorem 4.7. You can follow the argument of Theorem 4.1 above, or give more interesting proof.

Now, we can ask more interesting question, i.e. number of irreducible polynomials of certain degree.

For a given prime p and integer n , how many irreducible monic polynomials of degree n in $A = \mathbb{F}_p[T]$?

We will fix p , and denote the number of such polynomials as a_n . When n is small, we can get nice formulas.

Proposition 4.8.

$$a_2 = \frac{p^2 - p}{2}. \quad (18)$$

Proof. One can count number of *reducible* monic polynomials, and subtract from the total number of monic quadratic polynomials, which is simply p^2 (you have p choices for each c_0 and c_1 in $f(T) = T^2 + c_1T + c_0$). If $f(T)$ is not irreducible, then it should factor as $f(T) = (T - \alpha)(T - \beta)$ for some $\alpha, \beta \in \mathbb{F}_p$. Now, there are p cases where $\alpha = \beta$, and $\binom{p}{2}$ cases where $\alpha \neq \beta$. This gives

$$a_2 = p^2 - p - \binom{p}{2} = \frac{p^2 - p}{2}.$$

□

Exercise 4.2.2. Prove

$$a_3 = \frac{p^3 - p}{3}. \quad (19)$$

You may need to exclude the polynomials $f(T)$ which factor as 1) products of three linear polynomials, or 2) products of linear and irreducible quadratic polynomial.

It seems that we may even able to find a nice formula for a_n . To do this, we will define a *zeta function* for polynomials, which is similar to the Riemann zeta function but much simpler.

Definition 4.9 (Zeta function for A). We define *zeta function* as

$$\zeta_A(s) = \sum_{0 \neq f \text{ monic}} \frac{1}{|f|^s} \quad (20)$$

which converges for $\Re(s) > 1$.

We have a simple formula for ζ_A :

Proposition 4.10. We have

$$\zeta_A(s) = \frac{1}{1 - p^{1-s}}. \quad (21)$$

Proof. We can group the sum in (20) by degree, since $|f|$ only depends on degree of f . Since there are p^d -many monic polynomials of degree d , we have

$$\zeta_A(s) = \sum_{d \geq 1} \frac{p^d}{p^{ds}} = \sum_{d \geq 0} p^{d(1-s)} = \frac{1}{1 - p^{1-s}}.$$

□

Especially, it has a simple pole at $s = 1$ with residue $1/\log p$, and analytic continuation is clear from the formula. Also, there are no zeros of $\zeta_A(s)$, so Riemann Hypothesis for polynomials is trivially true. Note that the notation ζ_A somehow emphasizes A , since we will eventually study more general “zeta functions” for other polynomial rings (“extensions” of A), where the Riemann Hypothesis is more interesting (but still true).

What about Euler factorization? Since A is also UFD, the same argument as in Theorem 4.6 gives us

Theorem 4.11 (Euler factorization for polynomials). We have

$$\zeta_A(s) = \prod_{P \text{ monic irred}} (1 - |P|^{-s})^{-1}. \quad (22)$$

As in Proposition 4.10, we can group the product in (22) by degree, which gives

$$\zeta_A(s) = \prod_{d \geq 1} (1 - p^{-ds})^{-a_d}. \quad (23)$$

By comparing (21) and (23), we can deduce the following formula for a_n :

Theorem 4.12. We have

$$p^n = \sum_{d|n} d a_d. \quad (24)$$

Proof. Let $u = p^{-s}$, so

$$\frac{1}{1 - pu} = \prod_{d \geq 1} (1 - u^d)^{-a_d}. \quad (25)$$

By taking logarithmic derivative, we have

$$\frac{p}{1 - pu} = \sum_{d \geq 1} a_d \frac{du^{d-1}}{1 - u^d} \Leftrightarrow \frac{pu}{1 - pu} = \sum_{d \geq 1} a_d \frac{du^d}{1 - u^d}. \quad (26)$$

Using geometric series, we can rewrite the equations as

$$\sum_{n \geq 1} p^n u^n = \sum_{d \geq 1} d a_d \sum_{k \geq 1} u^{dk} = \sum_{n \geq 1} \left(\sum_{d|n} d a_d \right) u^n. \quad (27)$$

Now, we can compare the coefficients of u^n on both sides and get (24). □

We can *invert* this formula to get a formula for a_n . Such a process is called *Möbius inversion* - before we state the general formula, we will first give small examples that may leads to the general formula (in other words, *you can invent Möbius inversion formula yourself*). Let $b_n := n a_n$. For $n = 1, 2, 3, 4, 6, 12$, (24) gives us the following equations:

$$p = b_1,$$

$$\begin{aligned}
p^2 &= b_1 + b_2, \\
p^3 &= b_1 + b_3, \\
p^4 &= b_1 + b_2 + b_4, \\
p^6 &= b_1 + b_2 + b_3 + b_6, \\
p^{12} &= b_1 + b_2 + b_3 + b_4 + b_6 + b_{12}.
\end{aligned}$$

From these equations, we can compute b_n recursively:

$$\begin{aligned}
b_1 &= p, \\
b_2 &= p^2 - p, \\
b_3 &= p^3 - p, \\
b_4 &= p^4 - p^2, \\
b_6 &= p^6 - p^3 - p^2 + p, \\
b_{12} &= p^{12} - p^6 - p^4 + p^2.
\end{aligned}$$

Let's see how we actually computed b_6 and b_{12} . b_6 can be written as

$$\begin{aligned}
b_6 &= p^6 - (b_1 + b_2 + b_3) \\
&= p^6 - (b_1 + b_3) - (b_1 + b_2) + b_1 \\
&= p^6 - p^3 - p^2 + p.
\end{aligned}$$

Similarly, we can compute b_{12} as

$$\begin{aligned}
b_{12} &= p^{12} - (b_1 + b_2 + b_3 + b_4 + b_6) \\
&= p^{12} - (b_1 + b_2 + b_3 + b_6) - (b_1 + b_2 + b_4) - (b_1 + b_2) \\
&= p^{12} - p^6 - p^4 + p^2.
\end{aligned}$$

This may remind you the principle of inclusion-exclusion, and indeed, we can write the general formula as follows: for $n = p_1^{e_1} p_2^{e_2} \cdots p_k^{e_k}$,

$$\begin{aligned}
b_n &= p^n - \sum_{d|n, d < n} b_d \\
&= p^n - (b_1 + \cdots + b_{n/p_1}) - (b_1 + \cdots + b_{n/p_2}) - \cdots - (b_1 + \cdots + b_{n/p_k}) \\
&\quad + (b_1 + \cdots + b_{n/p_1 p_2}) + (b_1 + \cdots + b_{n/p_1 p_3}) + \cdots \\
&\quad \dots \\
&= p^n - p^{n/p_1} - p^{n/p_2} - \cdots - p^{n/p_k} + p^{n/(p_1 p_2)} + p^{n/(p_1 p_3)} + \cdots
\end{aligned}$$

This gives us the following formula for a_n :

Corollary 4.13. For a given prime p and integer n , we have

$$a_n = \frac{1}{n} \sum_{d|n} \mu(d) p^{n/d}, \quad (28)$$

where μ is the Möbius function, defined as

$$\mu(n) = \begin{cases} 1 & \text{if } n = 1, \\ (-1)^k & \text{if } n = p_1 p_2 \cdots p_k \text{ with } k \text{ distinct primes} \\ 0 & \text{otherwise.} \end{cases} \quad (29)$$

Exercise 4.2.3. Complete the proof of Corollary 4.13.


Note that (28) gives an *exact* formula for a_n . This also gives an analogue of the prime number theorem for polynomials: the dominating term of a_n is

$$\frac{p^n}{n} = \frac{p^n}{\log_p p^n},$$

where p^n is equal to the number of monic polynomials of degree n over \mathbb{F}_p . Moreover, the error term is bounded by $O\left(\frac{p^{n/2}}{n}\right)$, so we get

$$a_n = \frac{p^n}{\log_p p^n} + O\left(\frac{\sqrt{p^n}}{\log_p p^n}\right). \quad (30)$$

which has a similar form as Koch's bound (16) (or even better!).

Exercise 4.2.4.  Let $p = 3$. For each $1 \leq n \leq 10$, find the number of irreducible monic polynomials of degree n over \mathbb{F}_3 , by using the factorization of $T^{3^n} - T$. Check that the results match with the formula (28).

Exercise 4.2.5. Prove the following non-asymptotic version of (30):


$$\left| a_n - \frac{p^n}{n} \right| \leq \frac{p^{n/2}}{n}$$

Exercise 4.2.6. For each p and d , prove that the polynomial

$$\left(\prod_{0 \leq \deg(f) < d} f \right) + 1$$

is equal to a constant multiple of the product of all monic irreducible polynomials of degree d over \mathbb{F}_p . What is the constant?

Exercise 4.2.7. Prove that the polynomial $T^p - T - 1$ is irreducible over \mathbb{F}_p for all prime p .¹⁷

Exercise 4.2.8.  For a given prime p and an integer n , design an algorithm to find an irreducible polynomial of degree n over \mathbb{F}_p ? (This is not an easy question.)

Exercise 4.2.9. What would be a good analogue of the logarithmic integral $\text{Li}(x)$ for polynomials?

4.3 Counting other things

The power of the above method is that we can count other objects by replacing the zeta function with *Dirichlet series*

$$L(s, a) = \sum_{n \geq 1} \frac{a_n}{n^s} \quad (31)$$

or

$$L(s, a) = \sum_{\substack{0 \neq f \in A \\ f \text{ monic}}} \frac{a_f}{|f|^s}, \quad (32)$$

¹⁷Such a polynomial is called *Artin-Schreier polynomial*.

for given sequence $(a_n)_{n \geq 1}$ or sequence of polynomials $(a_f)_{f \in A}$. When these sequences are “interesting”, we call the Dirichlet series *L-function*. One condition for being “interesting” is that the Dirichlet series has an Euler factorization, i.e. it can be written as a product over primes (which might be more complicated than (12) or (22)).

For example, we will count the number of *square-free* monic polynomials of given degree n over \mathbb{F}_p . Note that $f \in A$ is square-free if and only if f is not divisible by P^2 for any monic irreducible polynomial P . In other words, it has no repeated factors in its factorization. To count such polynomials, one can consider the indicator function

$$\delta(f) = \begin{cases} 1 & \text{if } f \text{ is square-free,} \\ 0 & \text{otherwise} \end{cases} \quad (33)$$

and the corresponding Dirichlet series

$$L(s, \delta) = \sum_{\substack{0 \neq f \in A \\ f \text{ monic}}} \frac{\delta(f)}{|f|^s} = \sum_{n \geq 0} \frac{c_n}{p^{ns}}, \quad (34)$$

where $c_n = \sum_{\deg f = n} \delta(f)$ is the number of square-free monic polynomials of degree n over \mathbb{F}_p (here $c_0 = 1$). Now, only the polynomials with no repeated factors contribute to the sum, so they are product of distinct monic irreducible polynomials. This gives us the Euler factorization

$$L(s, \delta) = \prod_{P \text{ monic irred}} (1 + |P|^{-s}) \quad (35)$$

From $1 + |P|^{-s} = (1 - |P|^{-2s}) / (1 - |P|^{-s})$, comparing with the Euler factorization (22) we can rewrite (35) as

$$L(s, \delta) = \frac{\zeta_A(s)}{\zeta_A(2s)} = \frac{1 - p^{1-2s}}{1 - p^{1-s}}. \quad (36)$$

Substituting with $u = p^{-s}$ gives us

$$\frac{1 - pu^2}{1 - pu} = \sum_{n \geq 0} c_n u^n.$$

and expanding the left hand side as a geometric series gives us

Theorem 4.14. Let c_n be the number of square-free monic polynomials of degree n over \mathbb{F}_p . Then $c_1 = p$ and $c_n = p^n(1 - p^{-1})$ for $n \geq 2$.

Since there are p^n monic polynomials of degree n , the Theorem shows that the proportion of square-free monic polynomials of degree n is $(1 - p^{-1})$ for $n \geq 2$.

We can ask the same question for integers. The analogous *L-series* would be

$$L(s, \delta) = \sum_{n \geq 1} \frac{\delta(n)}{n^s} = \prod_{p \text{ prime}} (1 + p^{-s}) = \frac{\zeta(s)}{\zeta(2s)}, \quad (37)$$

where $\delta(n)$ is the indicator function for square-free integers, i.e.

$$\delta(n) = \begin{cases} 1 & \text{if } n \text{ is square-free,} \\ 0 & \text{otherwise.} \end{cases} \quad (38)$$

Then we are interested in the number of square-free integers $n \leq x$, i.e. the sum

$$\sum_{n \leq x} \delta(n). \quad (39)$$

The main idea is that the above sum (39) can be approximated by the limit of $L(s, \delta)$ as $s \rightarrow 0$ (so you need to know analytic continuation of it), and use analytic properties of $L(s, \delta)$ to get the asymptotic formula (such an argument is often called as *Tauberian arguments or theorems*). Here we only give the result.

Theorem 4.15. The number of square-free integers $n \leq x$ is asymptotically equal to

$$\sum_{n \leq x} \delta(n) \sim \frac{1}{\zeta(2)} x = \frac{6}{\pi^2} x. \quad (40)$$

Note that Theorem 4.15 and 4.14 are analogous to each other; the proportion of square-free integers (resp. polynomials) is $\frac{1}{\zeta(2)}$ (resp. $\frac{1}{\zeta_A(2)} = 1 - p^{-1}$).

Exercise 4.3.1. Prove that $\delta(n) = \mu(n)^2$ for all integers $n \geq 1$, where μ is the Möbius function defined in (29).

Exercise 4.3.2. Compute the number of *cube-free* polynomials of degree n over \mathbb{F}_p , i.e. the number of monic polynomials $f \in A$ such that f is not divisible by P^3 for any monic irreducible polynomial P . Can you generalize this to higher powers?

4.4 Sum of arithmetic functions

More generally, we can adapt the above method to compute the sum of *arithmetic functions*, i.e. functions defined on integers or polynomials.

For example, we consider the divisor counting function

$$d(n) := \sum_{d|n} 1 \quad (41)$$

for integer n , or

$$d(f) := \sum_{\substack{h|f \\ h \text{ monic}}} 1 \quad (42)$$

for (monic) polynomial $f \in A$. You may found that the function $d(n)$ behaves quite irregularly, e.g. $d(p) = 2$ for all prime p , while $d(n)$ can be large for composite n . But the sum or average of $d(n)$ (or arithmetic functions in general) behaves much better. For example, we have

Theorem 4.16 (Dirichlet).

$$\sum_{n \leq x} d(n) = x \log x + (2\gamma - 1)x + O(\sqrt{x} \log x), \quad (43)$$

where $\gamma = 0.577 \dots$ is the *Euler-Mascheroni constant*.

Proof. See [1, Theorem 3.3]. □

As you expect, it is much easier to compute the sum of $d(f)$ for (monic) polynomials $f \in A$ of given degree n . In fact, we can obtain an exact formula for the sum of $d(f)$:

Theorem 4.17. For a given prime p and integer n , we have

$$\sum_{\deg(f)=n} d(f) = (n+1)p^n. \quad (44)$$

Proof. Proof also uses L -series. In fact, we will prove that the L -series of $d(f)$ is simply given by

$$\sum_{\substack{0 \neq f \in A \\ f \text{ monic}}} \frac{d(f)}{|f|^s} = \zeta_A(s)^2 \quad (45)$$

Let's assume that the above formula is true for a moment. Denote the sum of $d(f)$ for monic polynomials of degree n as D_n . Then we have ($u = p^{-s}$)

$$\sum_{n \geq 0} D_n u^n = \sum_{\substack{0 \neq f \in A \\ f \text{ monic}}} \frac{d(f)}{|f|^s} = \zeta_A(s)^2 = \left(\frac{1}{1 - p^{1-s}} \right)^2 = \frac{1}{(1 - pu)^2}.$$

One can compute the right hand side by differentiating the geometric series $\frac{1}{1-pu} = 1 + pu + p^2u^2 + \dots$, which gives the desired formula.

To prove (45), we can simply take square of $\zeta_A(s)$:

$$\zeta_A(s)^2 = \left(\sum_f \frac{1}{|f|^s} \right)^2 = \left(\sum_{f_1} \frac{1}{|f_1|^s} \right) \left(\sum_{f_2} \frac{1}{|f_2|^s} \right) = \sum_f \left(\sum_{h_1 h_2 = f} 1 \right) \frac{1}{|f|^s} = \sum_f \frac{d(f)}{|f|^s}.$$

□

Note that one can write (44) as

$$\sum_{\deg(f)=n} d(f) = np^n + p^n = p^n \log_p p^n + p^n,$$

which is analogous to the main terms $x \log x + (2\gamma - 1)x$ in (43).

Why do we care about such sums? One nice example is the following theorem by Littlewood.

Theorem 4.18 (Littlewood). Define the *Mertens function* as

$$M(x) = \sum_{n \leq x} \mu(n). \quad (46)$$

Then the Riemann Hypothesis is equivalent to the following bound:

$$|M(x)| = O(x^{\frac{1}{2} + \epsilon}) \quad (47)$$

for any $\epsilon > 0$.

Exercise 4.4.1. For integer n , *Dedekind's ψ -function* is defined as

$$\psi(n) = n \prod_{p|n} \left(1 + \frac{1}{p} \right). \quad (48)$$

1. Prove that $\psi(n)$ is multiplicative, i.e. if $\gcd(m, n) = 1$, then $\psi(mn) = \psi(m)\psi(n)$.

2. Prove that $\psi(n) = n \sum_{d|n} \frac{|\mu(d)|}{d}$, where μ is the Möbius function defined in (29).
3. Prove that $\psi(n)$ is equal to the number of cyclic subgroups of order n in $(\mathbb{Z}/n\mathbb{Z})^2$.
4. Prove

$$\sum_{n \geq 1} \frac{\psi(n)}{n^s} = \frac{\zeta(s)\zeta(s-1)}{\zeta(2s)}$$

for $\Re(s) > 2$.¹⁸

It is known that the sum of $\psi(n)$ has an asymptotic formula:

$$\sum_{n \leq x} \psi(n) \sim \frac{\zeta(2)}{2\zeta(4)} x^2 + O(x \log x). \quad (49)$$

(See Apostol [1, p. 71, Exercise 11].)

Exercise 4.4.2. For $f \in A$, define $\psi(f)$ as

$$\psi(f) := |f| \prod_{P|f} (1 + |P|^{-1}), \quad (50)$$

where the product is over all monic irreducible polynomials P dividing f .

1. Prove all the analogous statements in the previous exercise.
2. Compute the sum

$$\sum_{f \in A, \deg(f)=n} \psi(f).$$

Compare the result with (49).

¹⁸Hint: use Euler factorization.

5 How many primes in arithmetic progressions?

In Section 3, we saw that the polynomial $T^2 + 1$ has a root in \mathbb{F}_p if and only if $p \equiv 1 \pmod{4}$ ¹⁹. Especially, we also saw that almost half of primes $p < 100$ satisfy $p \equiv 1 \pmod{4}$, and the other half satisfy $p \equiv 3 \pmod{4}$. One might ask if the “density” of primes actually converges to $1/2$ as we consider more and more primes.

In this section, we will show that this is indeed true. More generally, we will study Dirichlet’s theorem on primes in arithmetic progressions, which states that this is a general phenomenon for any arithmetic progression, and the densities are uniform across all congruence classes. Of course, we will also study the polynomial analogue of Dirichlet’s theorem on irreducible polynomials.

5.1 Dirichlet’s theorem on primes

Before we talk about densities of primes in arithmetic progressions $4k + 1$ and $4k + 3$, let’s first prove that there are infinitely many primes in each of these two congruence classes, by modifying the proof of Euclid’s theorem on the infinitude of primes.

Theorem 5.1. There are infinitely many primes of the form $4k + 1$.


Proof. Assume that there are only finitely many primes of the form $4k + 1$, and let p_1, p_2, \dots, p_r be the set of such primes. Consider $N = 4(p_1 p_2 \cdots p_r)^2 + 1$, and let p be a prime dividing N . Since N is odd, we have $p > 2$. Also, since p divides N , p has to be different from p_1, \dots, p_r . However, $(2p_1 p_2 \cdots p_r)^2 \equiv -1 \pmod{p}$ implies that -1 is a quadratic residue modulo p , hence $p \equiv 1 \pmod{4}$ by Theorem 3.1 and we get a contradiction. \square

Exercise 5.1.1. Prove that there are infinitely many primes of the form $4k + 3$.²⁰

One may ask if such an argument can be used to prove that there are infinitely many primes in any arithmetic progression $ak + b$ for $a > 0$ and $\gcd(a, b) = 1$, by choosing the polynomial defining N carefully. Unfortunately, this is not the case, and it turns out that such *Euclidean proof* only exists in very special cases.

Theorem 5.2 (Schur [30], Murty [22]). *Euclidean proof* exists for the arithmetic progression $ak + b$ if and only if $b^2 \equiv 1 \pmod{a}$.

However, one can find that the primes of the form $ak + b$ for different a, b (with $\gcd(a, b) = 1$) seems to be infinitely many, and also seems to be evenly distributed across all congruence classes.

Exercise 5.1.2.  Count the number of primes $p < 10^6$ in the arithmetic progressions $8k + b$ for $b = 1, 3, 5, 7$.

Dirichlet [13] proved that this is indeed true for any arithmetic progression $ak + b$ with $\gcd(a, b) = 1$. He actually proved that they are evenly distributed across all congruence classes

¹⁹Or $p = 2$.

²⁰Hint: Assume there are finitely many, p_1, \dots, p_r . Now consider $N = 4p_1 p_2 \cdots p_r + 3$. Show that there is at least one prime p dividing N such that $p \equiv 3 \pmod{4}$.

modulo a in the following sense: for a subset S of positive integers, the Dirichlet density of S is defined as the limit (if exists)

$$\delta(S) := \lim_{s \rightarrow 1^+} \frac{\sum_{p \in S} p^{-s}}{\sum_p p^{-s}}. \quad (51)$$

Note that the denominator $\sum_p p^{-s}$ goes to ∞ as $s \rightarrow 1^+$, which can be shown by Euler factorization:

$$\log \zeta(s) = \log \prod_p (1 - p^{-s})^{-1} = \sum_p \log(1 - p^{-s})^{-1} = \sum_p \frac{1}{p^s} + \sum_{p, k \geq 2} \frac{1}{kp^{ks}} = \sum_p \frac{1}{p^s} + R(s)$$

where $R(s)$ is bounded as $s \rightarrow 1^+$, and the claim follows from $\lim_{s \rightarrow 1^+} \zeta(s) = \infty$. This might be different from what you'd expect as a density, which is usually called *natural density*:

$$\delta_{\text{nat}}(S) := \lim_{n \rightarrow \infty} \frac{\#\{p \in S : p \leq n\}}{\#\{p : p \leq n\}}. \quad (52)$$

Dirichlet proved the following theorem:

Theorem 5.3 (Dirichlet [13]). For each a, b with $\gcd(a, b) = 1$, there are infinitely many primes of the form $ak + b$. Moreover, the Dirichlet density of such primes is $\frac{1}{\phi(a)}$, where ϕ is the Euler ϕ -function.²¹

The proof is often considered as a starting point of modern analytic number theory. It can be found in many places, and we will only give a sketch. For the historical point of view, see the article by Avigad and Morris [2].

The main idea is to use *Dirichlet characters* and associated *Dirichlet L-functions*. As a motivating example, let's consider the case when $a = 4$ and $b = 1$, which is already proved in Theorem 5.1. In this case, the Dirichlet's theorem states that

$$\delta(\{p : p \equiv 1 \pmod{4}\}) = \lim_{s \rightarrow 1^+} \frac{\sum_{p \equiv 1 \pmod{4}} p^{-s}}{\sum_p p^{-s}} = \frac{1}{2}. \quad (53)$$

From the previous discussion, one may expect to use the Dirichlet series

$$L_1(s) = \sum_{n \equiv 1 \pmod{4}} \frac{1}{n^s} = \frac{1}{1^s} + \frac{1}{5^s} + \frac{1}{9^s} + \frac{1}{13^s} + \dots$$

However, this series is not good enough, since it does not admit Euler factorization. Similarly, the series

$$L_3(s) = \sum_{n \equiv 3 \pmod{4}} \frac{1}{n^s} = \frac{1}{3^s} + \frac{1}{7^s} + \frac{1}{11^s} + \frac{1}{15^s} + \dots$$

is not the right choice. However, their sum and differences are nice: we have

$$L_1(s) + L_3(s) = \sum_{n \equiv 1 \pmod{4}} \frac{1}{n^s} = \prod_{p \neq 2} (1 - p^{-s})^{-1} = (1 - 2^{-s})\zeta(s) \quad (54)$$

$$L_1(s) - L_3(s) = \frac{1}{1^s} - \frac{1}{3^s} + \frac{1}{5^s} - \frac{1}{7^s} + \dots = \prod_{p \equiv 1 \pmod{4}} (1 - p^{-s})^{-1} \prod_{p \equiv 3 \pmod{4}} (1 + p^{-s})^{-1} \quad (55)$$

²¹In fact, natural density also exists and is equal to $\frac{1}{\phi(a)}$.

Especially, taking logarithm of (55) gives

$$\log(L_1(s) - L_3(s)) = \sum_{p \equiv 1 \pmod{4}} \frac{1}{p^s} - \sum_{p \equiv 3 \pmod{4}} \frac{1}{p^s} + R(s)$$

where $R(s)$ is bounded as $s \rightarrow 1^+$. Assuming $\lim_{s \rightarrow 1^+} L_1(s) - L_3(s) \neq 0$, we can conclude that the Dirichlet density of primes $p \equiv 1 \pmod{4}$ and primes $p \equiv 3 \pmod{4}$ are the same as $1/2$.

The series $L_1(s) - L_3(s)$ is a special case of *Dirichlet L-functions*, associated to a *Dirichlet character*: for the function $\chi : \mathbb{Z} \rightarrow \mathbb{C}$ defined by

$$\chi(n) = \chi_{-4}(n) := \begin{cases} 1 & n \equiv 1 \pmod{4} \\ -1 & n \equiv 3 \pmod{4} \\ 0 & \text{otherwise} \end{cases} \quad (56)$$

we have $L_1(s) - L_3(s) = \sum_{n \geq 1} \chi(n)n^{-s}$. More generally,

Definition 5.4. A *Dirichlet character modulo a* is a function $\chi : \mathbb{Z} \rightarrow \mathbb{C}$ such that:

1. $\chi(n + a) = \chi(n)$ for all $n \in \mathbb{Z}$.
2. $\chi(mn) = \chi(m)\chi(n)$ for all $m, n \in \mathbb{Z}$.
3. $\chi(n) = 0$ if and only if $\gcd(n, a) > 1$.

Note that this induces a group homomorphism $(\mathbb{Z}/a\mathbb{Z})^\times \rightarrow \mathbb{C}^\times$, and conversely, any group homomorphism $\chi : (\mathbb{Z}/a\mathbb{Z})^\times \rightarrow \mathbb{C}^\times$ can be extended to a Dirichlet character modulo a .

For general modulus, we consider all Dirichlet characters modulo a and the associated Dirichlet L-functions, and consider the linear combination

$$\sum_{\chi} \overline{\chi(b)} \log L(s, \chi) = \phi(a) \sum_{p \equiv b \pmod{a}} \frac{1}{p^s} + R(s)$$

where $R(s)$ is bounded as $s \rightarrow 1^+$. Equality follows from the *orthogonality* of Dirichlet characters (see Theorem 5.8 for the polynomial case). Then Theorem 5.3 follows once we prove $L(1, \chi) \neq 0$ for all nontrivial χ , which is the main step of Dirichlet's proof.

Exercise 5.1.3. In this exercise, we will prove

$$L(1, \chi_{-4}) = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots = \frac{\pi}{4} \neq 0. \quad (57)$$

1. Without computing the actual value, show that the series converges to a positive number.
2. Consider the polynomial

$$f_n(x) = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \cdots + (-1)^n \frac{x^{2n+1}}{2n+1}.$$

Show that $\lim_{n \rightarrow \infty} f_n(1) = L(1, \chi_{-4})$ and

$$f'_n(x) = \frac{1 - (-x^2)^{n+1}}{1 + x^2}.$$

3. By integrating $f'_n(x)$ from 0 to 1, show that

$$f_n(1) = \int_0^1 f'_n(t) dt = \int_0^1 \frac{dt}{1+t^2} - \int_0^1 \frac{(-t^2)^{n+1}}{1+t^2} dt = \frac{\pi}{4} - \int_0^1 \frac{(-t^2)^{n+1}}{1+t^2} dt.$$

Now use dominated convergence theorem to conclude the proof.

5.2 Dirichlet's theorem on irreducible polynomials

We will follow the same idea as in the integer case to prove the polynomial analogue of Dirichlet's theorem. The notable difference is that proving $L(1, \chi) \neq 0$ for nontrivial χ is much easier, since the L -functions $L(s, \chi)$ are simply polynomials in p^{-s} . We can define Dirichlet character as you expect.

Definition 5.5. Let $h \in A$ be a (monic) polynomial of degree $d \geq 1$. A *Dirichlet character modulo h* is a function $\chi : A \rightarrow \mathbb{C}$ such that:

1. $\chi(a + bh) = \chi(a)$ for all $a \in A$ and $b \in A$.
2. $\chi(ab) = \chi(a)\chi(b)$ for all $a, b \in A$.
3. $\chi(a) \neq 0$ if and only if $\gcd(a, h) = 1$.

As you expect, this induces a group homomorphism $(A/hA)^\times \rightarrow \mathbb{C}^\times$, and conversely, any group homomorphism $\chi : (A/hA)^\times \rightarrow \mathbb{C}^\times$ can be extended to a Dirichlet character modulo h . The trivial Dirichlet character modulo h is defined by $\chi(a) = 1$ for all $a \in A$ with $\gcd(a, h) = 1$. We will denote the set of all Dirichlet characters modulo h by X_h .

Example 5.6. Let $p = 3$ and $h(T) = T$. Then we have the following nontrivial Dirichlet characters modulo h :

$$\chi(f) := \begin{cases} 1 & \text{if } f \equiv 1 \pmod{T} \\ -1 & \text{if } f \equiv 2 \pmod{T} \\ 0 & \text{otherwise} \end{cases}$$

Example 5.7. Let $p = 2$ and $h(T) = T^2 + T + 1$. Then the following defines a nontrivial Dirichlet character modulo h :

$$\chi(f) = \begin{cases} 1 & \text{if } f \equiv 1 \pmod{h} \\ \zeta_3 & \text{if } f \equiv T \pmod{h} \\ \zeta_3^2 & \text{if } f \equiv T + 1 \pmod{h} \\ 0 & \text{otherwise} \end{cases}$$

where $\zeta_3 = e^{2\pi i/3}$ is a third root of unity.

Exercise 5.2.1. For the above example, how many Dirichlet characters modulo h are there?

Theorem 5.8 (Orthogonality). Let χ, ψ be Dirichlet characters modulo h and $a, b \in A$ be such that $\gcd(a, h) = \gcd(b, h) = 1$. Then we have

1.
$$\sum_{a \in A/hA} \chi(a) \overline{\psi(a)} = \begin{cases} \phi(h) & \text{if } \chi = \psi \\ 0 & \text{if } \chi \neq \psi \end{cases} \quad (58)$$

2.
$$\sum_{\chi \in X_h} \chi(a) \overline{\chi(b)} = \begin{cases} \phi(h) & \text{if } a \equiv b \pmod{h} \\ 0 & \text{otherwise} \end{cases} \quad (59)$$

Here $\phi(h) = \#(A/hA)^\times$ is the Euler ϕ -function defined previously.

Proof. We will only prove (58), and leave the proof of (59) as an exercise. Note that we only need to consider $a \in (A/hA)^\times$ since $\chi(a) = 0$ and $\psi(a) = 0$ if $\gcd(a, h) \neq 1$. If $\chi = \psi$, then we have $\chi(a)\overline{\chi(a)} = 1$ for all $a \in (A/hA)^\times$, and hence the sum becomes $\#(A/hA)^\times = \phi(h)$. If $\chi \neq \psi$, there exists $a_0 \in (A/hA)^\times$ such that $\chi(a_0) \neq \psi(a_0)$, i.e. $\chi(a_0)\overline{\psi(a_0)} = \chi(a_0)\psi(a_0)^{-1} \neq 1$. Since the map $a \mapsto aa_0$ is a bijection on $(A/hA)^\times$, we have

$$\sum_{a \in (A/hA)^\times} \chi(a)\overline{\psi(a)} = \sum_{a \in (A/hA)^\times} \chi(aa_0)\overline{\psi(aa_0)} = \chi(a_0)\overline{\psi(a_0)} \sum_{a \in (A/hA)^\times} \chi(a)\overline{\psi(a)}$$

and $\chi(a_0)\overline{\psi(a_0)} \neq 1$ implies that the sum is zero. \square

Exercise 5.2.2. Prove (59).²²

Definition 5.9. For a Dirichlet character χ modulo h , we define the *Dirichlet L-function* of χ as

$$L(s, \chi) = \sum_{\substack{0 \neq f \in A \\ f \text{ monic}}} \frac{\chi(f)}{|f|^s} \quad (60)$$

Using orthogonality, we can show that Dirichlet L-functions for nontrivial characters are simply polynomials in p^{-s} .

Proposition 5.10. Let χ be a nontrivial Dirichlet character modulo h . Then the Dirichlet L-function $L(s, \chi)$ is a polynomial in p^{-s} of degree at most $\deg(h) - 1$.

Proof. Define

$$A(n, \chi) = \sum_{\substack{\deg f = n \\ f \text{ monic}}} \chi(f) \quad (61)$$

so that

$$L(s, \chi) = \sum_{n=0}^{\infty} A(n, \chi) p^{-ns} \quad (62)$$

Then our goal reduces to showing that $A(n, \chi) = 0$ for all $n \geq \deg(h)$. If $\deg(f) = n$, it can be uniquely written as $f = gh + r$ where $\deg(r) < \deg(h)$ or $r = 0$, and $\deg(g) = n - \deg(h)$. For each r , we have $p^{n-\deg(h)}$ possibilities for g , and $\chi(f) = \chi(r)$. Therefore, we have

$$A(n, \chi) = p^{n-\deg(h)} \sum_{\deg(r) < \deg(h)} \chi(r) = p^{n-\deg(h)} \sum_{r \in A/hA} \chi(r) = 0$$

where the summation is zero by the orthogonality of Dirichlet characters (58) (χ is nontrivial). \square

Exercise 5.2.3. Is it always true that $A(\deg(h) - 1, \chi) \neq 0$, i.e. the degree of $L(s, \chi)$ is exactly $\deg(h) - 1$? If not, give a counterexample.

Since χ is multiplicative, Dirichlet L-function admits Euler product formula:

Proposition 5.11. Let χ be a Dirichlet character modulo h . Then we have

$$L(s, \chi) = \prod_{P \nmid h} (1 - \chi(P)|P|^{-s})^{-1} \quad (63)$$

where P runs over all irreducible monic polynomials not dividing h .

²²Hint: When $a \not\equiv b \pmod{h}$, there exists χ such that $\chi(a) \neq \chi(b)$.

Exercise 5.2.4. Prove Proposition 5.11.

Example 5.12. Consider the Example 5.6 with $p = 3$ and $h(T) = T$. By Proposition 5.10, $L(s, \chi)$ is a polynomial of degree at most 0, i.e. a constant. One can directly check that $A(0, \chi) = 1^{23}$, hence $L(s, \chi) = 1$.

Example 5.13. Consider the Example 5.7 with $p = 2$ and $h(T) = T^2 + T + 1$. By Proposition 5.10, $L(s, \chi)$ is a polynomial of degree at most 1. One can directly check that $A(0, \chi) = 1$ and $A(1, \chi) = \zeta_3 + \zeta_3^2 = -1$, hence $L(s, \chi) = 1 - 2^{-s}$.

When $\chi = \chi_0$ is the trivial Dirichlet character modulo h , the Euler factorization of $L(s, \chi_0)$ is almost the same as (22) except that we exclude the factor corresponding to P with $P \mid h$, hence

$$L(s, \chi_0) = \prod_{P \nmid h} (1 - |P|^{-s}) \cdot \zeta_A(s) = \frac{\prod_{P \mid h} (1 - |P|^{-s})}{1 - p^{1-s}}. \quad (64)$$

Our main goal is to prove $L(1, \chi) \neq 0$ for all nontrivial Dirichlet characters χ , as in the integer case. We need the following lemma.

Lemma 5.14. Let χ vary over all Dirichlet characters modulo h . Then, for each irreducible polynomial P not dividing h , there exist positive integers f_P and g_P such that $f_P g_P = \phi(h)$ and

$$\prod_{\chi \in X_h} L(s, \chi) = \prod_{P \nmid h} (1 - |P|^{-f_P s})^{-g_P}. \quad (65)$$

Proof. For each $P \nmid h$, the map $\chi \mapsto \chi(P)$ gives a group homomorphism $X_h \rightarrow \mathbb{C}^\times$. Its image is a finite order subgroup of \mathbb{C}^\times , and hence it is cyclic. Let f_P be the order of the image, and g_P be the order of the kernel, so that $f_P g_P = \phi(h)$. Then the P -th Euler factor of $\prod_{\chi} L(s, \chi)$ is given by

$$\prod_{j=0}^{f_P-1} (1 - \zeta_{f_P}^j |P|^{-s})^{-g_P} = (1 - |P|^{-f_P s})^{-g_P}.$$

Here ζ_{f_P} is a primitive f_P -th root of unity. Multiplying over all irreducible polynomials $P \nmid h$ gives us the desired formula. \square

Using this, we can prove the nonvanishing result.

Lemma 5.15. Suppose χ is a complex Dirichlet character modulo h , i.e. $\bar{\chi} \neq \chi$. Then $L(1, \chi) \neq 0$.

Proof. Assume $L(1, \chi) = 0$. Consider (65) for h . Among the factors of the LHS of (65), $L(s, \chi_0)$ has a simple pole at $s = 1$ and $L(s, \chi)$ has a zero at $s = 1$, so is $L(s, \bar{\chi})$. Since $L(s, \chi)$ is regular (i.e. no pole) at $s = 1$ for all nontrivial character χ , the LHS has a zero at $s = 1$. However, the RHS is positive for all $s > 1$; each Euler factor has positive coefficients as

$$(1 - |P|^{-f_P s})^{-g_P} = \sum_{k \geq 0} \binom{g_P + k - 1}{g_P - 1} |P|^{-f_P k s}.$$

By taking the limit as $s \rightarrow 1$, we get a contradiction. \square

²³2 is not monic!

We need to be more careful about the case when χ is a real Dirichlet character modulo h , i.e. $\bar{\chi} = \chi$ (a quadratic character).

Lemma 5.16. Suppose χ is a nontrivial real Dirichlet character modulo h . Then $L(1, \chi) \neq 0$.

Proof. The main idea is to consider

$$G(s) = \frac{L(s, \chi_0)L(s, \chi)}{L(2s, \chi_0)} = \sum_f \frac{a(f)}{|f|^s} \quad (66)$$

where χ_0 is the trivial Dirichlet character modulo h . The P -th Euler factor of $G(s)$ for $P \nmid h$ is

$$\begin{cases} \frac{1+|P|^{-s}}{1-|P|^{-s}} = 1 + 2 \sum_{k \geq 1} |P|^{-ks} & \text{if } \chi(P) = 1 \\ 1 & \text{if } \chi(P) = -1 \end{cases}$$

so $G(s)$ has nonnegative coefficients, i.e. $a(f) \geq 0$ for all f . One can rewrite (66) as

$$\frac{(1 - pu^2)L^*(u, \chi)}{1 - pu} = \sum_{n \geq 0} A(n)u^n, \quad (67)$$

where $u = p^{-s}$, $L^*(u, \chi) = L(s, \chi)$, and $A(n) = \sum_{\deg f=n} a(f)$. By Proposition 5.10, $L^*(u, \chi)$ is a polynomial in u (of degree at most 1). Now, assume that $L(1, \chi) = 0$, i.e. $L^*(p^{-1}, \chi) = 0$. Then $1 - pu$ divides $L^*(u, \chi)$ and the LHS of (67) is a polynomial. Thus RHS is also a polynomial, where all the coefficients are nonnegative since $a(f)$'s are. However, the LHS has a zero $u = p^{1/2}$, which is impossible for polynomials with nonnegative coefficients. \square

Exercise 5.2.5. Let χ be a nontrivial real Dirichlet character modulo $h \in A$. Prove that there exists $c \in \mathbb{F}_p$ such that $\chi(T + c) = 1$.²⁴

Theorem 5.17 (Dirichlet's theorem for polynomials). Let $h \in A$ be a monic polynomial of degree $d \geq 1$. Let $a \in A$ be a polynomial such that $\gcd(a, h) = 1$. Then the set of irreducible monic polynomials $f \in A$ such that $f \equiv a \pmod{h}$ has Dirichlet density $\frac{1}{\phi(h)} > 0$. Especially, there are infinitely many such polynomials.

Proof. From Proposition 5.11 and $-\log(1 - x) = \sum_{k \geq 1} \frac{x^k}{k}$, we have

$$\log L(s, \chi) = - \sum_P \log \left(1 - \frac{\chi(P)}{|P|^s} \right) = \sum_P \frac{\chi(P)}{|P|^s} + \sum_{P, k \geq 2} \frac{\chi(P)^k}{k|P|^{ks}} = \sum_P \frac{\chi(P)}{|P|^s} + R(s, \chi) \quad (68)$$

where $R(s, \chi)$ is bounded as $s \rightarrow 1$. Now, multiplying $\overline{\chi(a)}$ and summing over χ gives

$$\sum_{\chi} \overline{\chi(a)} \log L(s, \chi) = \phi(h) \sum_{P \equiv a \pmod{h}} \frac{1}{|P|^s} + R(s) \quad (69)$$

where $R(s) = \sum_{\chi} \overline{\chi(a)} R(s, \chi)$ is still remain bounded as $s \rightarrow 1$. Dividing by $\sum_P 1/|P|^s$ gives

$$\frac{\log L(s, \chi_0)}{\sum_P 1/|P|^s} + \sum_{\chi \neq \chi_0} \overline{\chi(a)} \cdot \frac{\log L(s, \chi)}{\sum_P 1/|P|^s} = \phi(h) \frac{\sum_{P \equiv a \pmod{h}} 1/|P|^s}{\sum_P 1/|P|^s} + \frac{R(s)}{\sum_P 1/|P|^s} \quad (70)$$

Using (64) and $L(1, \chi) \neq 0$ for $\chi \neq \chi_0$, one can show that the first term of the LHS of the above equation converges to 1, while the other terms correspond to the nontrivial characters converges to 0, as $s \rightarrow 1$. This gives the desired result. \square

²⁴Hint: This is equivalent to $L(1, \chi) \neq 0$. In other words, proving this without using the nonvanishing result would give a new proof of Lemma 5.16.

Corollary 5.18. Let $h \in A$ be a monic polynomial of degree $d \geq 1$. For each $a \in A$ such that $\gcd(a, h) = 1$ and $n \geq 0$, consider the set of irreducible monic polynomials

$$S_n(a, h) := \{P \in A : P \equiv a \pmod{h}, \deg(P) = n\}. \quad (71)$$

Then we have

$$S_n(a, h) = \frac{1}{\phi(h)} \frac{p^n}{n} + O\left(\frac{p^{n/2}}{n}\right). \quad (72)$$

If we denote as S_n for the set of irreducible monic polynomials of degree n , then (72) implies

$$\lim_{n \rightarrow \infty} \frac{\#S_n(a, h)}{\#S_n} = \frac{1}{\phi(h)},$$

i.e. the *natural density* of $S_n(a, h)$ is $\frac{1}{\phi(h)}$.

Proof. Proof can be found in [26, p. 40, Theorem 4.8]. Note that it uses Riemann hypothesis for $L(s, \chi)$, which will be covered later. \square

Exercise 5.2.6 ([26, p. 43, Exercise 6]). Let $f_0(T) \in A$ be a polynomial of degree d with a non-zero constant term. Show that there are infinitely many irreducible monic polynomials $f(T) \in A$ whose first $m + 1$ terms coincide with $f_0(T)$. What is the Dirichlet density of such irreducible polynomials?

5.3 More precise formula and Chebyshev's bias

Now we know that Dirichlet's theorem is true for polynomials, i.e. the irreducible polynomials are equidistributed in the congruence classes modulo h . In the previous section, we were able to prove an exact formula for the number of irreducible polynomials (28). Then it is natural to ask if we can also prove an exact formula for the number of irreducible polynomials in a given congruence class modulo h . We mainly work with the case when $p = 3$ and $h(T) = T$, but the same argument may generalize to any prime p and any polynomial h .

For each n , let $a_{n,1}$ and $a_{n,2}$ be the number of irreducible monic polynomials of degree n in $\mathbb{F}_3[T]$ that are congruent to 1 and 2 modulo T , respectively. Our main goal is express the Dirichlet L -functions $L(s, \chi)$ and $L(s, \chi_0)$ in terms of $a_{n,1}$ and $a_{n,2}$, which eventually gives us the exact formulas for $a_{n,1}$ and $a_{n,2}$. Here χ_0 is the trivial Dirichlet character modulo T , which is defined by $\chi_0(f) = 1$ for all $f \in A$ with $\gcd(f, T) = 1$.

Lemma 5.19. Let χ be the Dirichlet character modulo T defined in Example 5.6. Then we have

$$1 = L(s, \chi) = \prod_{n \geq 1} (1 - 3^{-ns})^{-a_{n,1}} (1 + 3^{-ns})^{-a_{n,2}} \quad (73)$$

Proof. The first equality is shown in Example 5.12. The second equality follows from the Euler factorization of $L(s, \chi)$. \square

Lemma 5.20. Let χ_0 be the trivial Dirichlet character modulo T . Then we have

$$\frac{1 - 3^{-s}}{1 - 3^{1-s}} = L(s, \chi_0) = \prod_{n \geq 1} (1 - 3^{-ns})^{-a_{n,1} - a_{n,2}}. \quad (74)$$

Proof. Both of the equalities follow from Euler factorization of $L(s, \chi_0)$. On the one hand, the Euler factorization of $L(s, \chi_0)$ is the same as (22) except that we exclude the factor corresponding to T , which is $(1 - 3^{-s})^{-1}$. This gives us the first equality. On the other hand, grouping the factors by degree gives us the second equality. \square

Now, we can divide (74) by (73) to get

$$\frac{1-u}{1-3u} = \prod_{n \geq 1} \left(\frac{1+u^n}{1-u^n} \right)^{a_{n,2}}.$$

for $u = 3^{-s}$. By taking logarithmic derivative, we have

$$\begin{aligned} -\frac{1}{1-u} + \frac{3}{1-3u} &= -(1+u+u^2+\dots) + (3+3^2u+3^3u^2+\dots) \\ &= \sum_{n \geq 1} a_{n,2} \left(\frac{nu^{n-1}}{1+u^n} + \frac{nu^{n-1}}{1-u^n} \right) \\ &= \sum_{n \geq 1} 2na_{n,2}u^{n-1}(1+u^{2n}+u^{4n}+\dots) \end{aligned}$$

and hence

$$\sum_{n \geq 1} (3^n - 1)u^n = \sum_{n \geq 1} 2na_{n,2}(u^n + u^{3n} + u^{5n} + \dots) = \sum_{n \geq 1} \left(\sum_{\substack{d|n \\ n/d \text{ odd}}} 2da_{d,2} \right) u^n.$$

By comparing the coefficients, we have

$$\frac{3^n - 1}{2} = \sum_{\substack{d|n \\ n/d \text{ odd}}} da_{d,2}. \quad (75)$$

By applying the Möbius inversion formula, we can express $a_{n,2}$ in terms of n :

Theorem 5.21. We have

$$a_{n,1} = \frac{1}{2n} \left(\sum_{2 \nmid d|n} \mu(d) \cdot 3^{\frac{n}{d}} + 2 \sum_{2|d|n} \mu(d) \cdot 3^{\frac{n}{d}} + \epsilon_n \right) \quad (76)$$

$$a_{n,2} = \frac{1}{2n} \left(\sum_{2 \nmid d|n} \mu(d) \cdot 3^{\frac{n}{d}} - \epsilon_n \right) \quad (77)$$

where

$$\epsilon_n = \begin{cases} 1 & n = 2^k \text{ for some } k \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad (78)$$

Proof. Equation 77 follows from the above discussion, and (76) follows from the fact that $a_{n,1} + a_{n,2} = a_n$ for $n \neq 1$ where the formula for a_n is given in (28). \square

Not only this explains the exact number of irreducible polynomials in each congruence class modulo T , but also it shows that one of them is larger than the other, even if they are asymptotically equal as $n \rightarrow \infty$.

Corollary 5.22. We have $a_{n,2} \geq a_{n,1}$ for all $n \geq 1$, and the equality holds if and only if n is odd.

Proof. It is clear that $a_{n,1} = a_{n,2}$ when n is odd. For n even, we have

$$\begin{aligned} a_{n,2} - a_{n,1} &= \frac{1}{n} \left(- \sum_{2|d|n} \mu(d) \cdot 3^{\frac{n}{d}} - \epsilon_n \right) = \frac{1}{n} \left(- \sum_{d|\frac{n}{2}} \mu(2d) \cdot 3^{\frac{n}{2d}} - \epsilon_n \right) \\ &\geq \frac{1}{n} \left(3^{\frac{n}{2}} - 3^{\frac{n}{2}-1} - \dots - 3 - 1 \right) > 0. \end{aligned}$$

□

In other words, although $a_{n,1}$ and $a_{n,2}$ are asymptotically same as $n \rightarrow \infty$, one of them is always larger than or equal to the other. Such a *bias* is observed by Chebyshev for the primes congruent to 1 and 3 modulo 4 [6]. Let $\pi(x; m, a)$ be the number of prime numbers $p \leq x$ such that $p \equiv a \pmod{m}$. Chebyshev found that $\pi(x; 4, 3) > \pi(x; 4, 1)$ occurs much more often than $\pi(x; 4, 1) > \pi(x; 4, 3)$. It might be even plausible to conjecture that the inequality is true for sufficiently large x . However, this turned out to be false:

Theorem 5.23 (Littlewood [20]). There are infinitely many x such that $\pi(x; 4, 1) > \pi(x; 4, 3)$. More precisely, we have

$$\pi(x; 4, 1) - \pi(x; 4, 3) > \frac{1}{2} \frac{\sqrt{x}}{\log x} \log \log \log x$$


for infinitely many x .

The second thing we can guess is that the inequality $\pi(x; 4, 3) > \pi(x; 4, 1)$ might be true for almost all x , i.e. the density of the set $\{x : \pi(x; 4, 3) > \pi(x; 4, 1)\}$ is 1. However, this is also false; in fact, Rubinstein and Sarnak [28] showed that the *natural density* of the set $\{x : \pi(x; 4, 3) > \pi(x; 4, 1)\}$ does not exist. Instead, they showed that *logarithmic density* exists and very close to 1.

Theorem 5.24 (Rubinstein–Sarnak [28]).

$$\lim_{X \rightarrow \infty} \frac{1}{\log X} \sum_{\substack{x \leq X \\ \pi(x; 4, 3) > \pi(x; 4, 1)}} \frac{1}{x} = 0.9959 \dots$$

In the case of polynomials, Cha [5] proved that Chebyshev's bias still holds for odd characters, under certain assumption on the zeros of L -functions (*Grand Simplicity Hypothesis*, GSH). For more story about Chebyshev's bias, see the excellent article of Granville and Martin [16].

Exercise 5.3.1.  Reproduce Chebyshev's observation by plotting the graphs of $\pi(x; 4, 1)$ and $\pi(x; 4, 3)$.

Exercise 5.3.2. Adapt the same argument to the Example 5.7, where $p = 2$ and $h(T) = T^2 + T + 1$. For each $a = 1, T, T + 1$, find formulas for the number of irreducible polynomials of degree n in $\mathbb{F}_2[T]$ that are congruent to a modulo h . Which is the largest and which is the smallest?²⁵

²⁵Hint: You may need to consider three L -functions $L(s, \chi)$, $L(s, \chi^2)$, and $L(s, \chi^3)$ where χ^3 is the trivial Dirichlet character modulo h .

6 It is as easy as ABC

Infamous Fermat's Last Theorem (FLT) states that there are no positive integer solutions to the equation

$$x^n + y^n = z^n$$

for any integer $n \geq 2$. The theorem was famously proven by Andrew Wiles in 1994, using sophisticated techniques from algebraic geometry and number theory. See the great expository book by Cornell, Silverman, and Stevens [9] for a detailed account of the proof.

As you expect, the goal of this section is to prove a polynomial analogue of FLT. Indeed, we will show a stronger result, named *Mason–Stothers theorem* or *Polynomial ABC* [21, 35]. The proof given in this section is based on the elementary argument by Snyder [33], but we will also give a *geometric* proof in the next section.

6.1 Polynomial FLT

Let k be a field and $k[T]$ the polynomial ring over k . We have the following version of FLT:

Theorem 6.1 (Polynomial FLT). Let k be a field and $n \geq 3$ an integer. Let $f, g, h \in k[T]$ be mutually coprime polynomials such that not all derivatives f', g', h' are zero. Then $f^n + g^n \neq h^n$.

Note that we have solutions for $n \leq 2$, e.g. $(1 - T^2)^2 + (2T)^2 = (1 + T^2)^2$. Also, when k is of characteristic 0, $f' = 0$ is equivalent to f being a constant polynomial. However, there is a nonconstant solution when k is of characteristic p , e.g. $1^p + T^p = (1 + T)^p$ for $n = p$.

The main idea of a proof is to use “derivatives” of polynomials.

Proof. Let's assume that $f(T)^n + g(T)^n = h(T)^n$. Differentiating both sides gives

$$f(T)^{n-1}f'(T) + g(T)^{n-1}g'(T) - h(T)^{n-1}h'(T) = 0.$$

By multiplying $h(T)$ on the both sides, we can rewrite this as

$$\begin{aligned} f^{n-1}f'h + g^{n-1}g'h - h^n h' &= f^{n-1}f'h + g^{n-1}g'h - (f^n + g^n)h' = 0 \\ \Leftrightarrow f^{n-1}(f'h - fh') &= -g^{n-1}(g'h - gh'). \end{aligned}$$

Since f and g are coprime, this implies that f^{n-1} divides $g'h - gh'$. Similarly, we can show that g^{n-1} divides $h'f - hf'$ and h^{n-1} divides $f'g - fg'$. Also, we have $gh' - g'h \neq 0$; otherwise, $gh' = g'h$ and coprimality of g, h implies that $g \mid g'$, so $g' = 0$ from $\deg(g') < \deg g$. Without loss of generality, assume that f has the largest degree among the three polynomials. Then

$$2 \deg f \leq (n-1) \deg f = \deg f^{n-1} \leq \deg(g'h - gh') \leq \deg g' + \deg h - 1 \leq 2 \deg f - 1,$$

which is a contradiction. \square

Compared to the Wiles and Taylor's proof of original FLT [38, 36], the proof of Theorem 6.1 is very elementary and short. However, the above argument cannot be applied to the original FLT, since we don't have a good notion of derivatives for integers.

Exercise 6.1.1. 1. Show that the only map $D : \mathbb{Z} \rightarrow \mathbb{Z}$ such that $D(m+n) = D(m) + D(n)$ (additive) and $D(mn) = mD(n) + nD(m)$ (Leibniz rule) is the zero map.

2. If we give up the additivity condition, then show that there is a nonzero map $D : \mathbb{Z} \rightarrow \mathbb{Z}$ such that $D(mn) = mD(n) + nD(m)$. Such a map is called an *arithmetic derivative*, which is used in [24] to given an alternative proof of infinitude of primes.

6.2 Polynomial ABC

In fact, we can prove more general result called *Mason–Stothers theorem* or *Polynomial ABC* [21, 35, 33]. Before stating the theorem, we introduce the original ABC conjecture for integers first. For $n \in \mathbb{Z}$, we define the *radical* of n as the product of distinct prime factors of n : $\text{rad}(n) = \prod_{p|n} p$. Oesterlé and Masser conjectured the following:

Conjecture 6.2 (ABC conjecture). For any $\epsilon > 0$, there exists a constant $K = K_\epsilon > 0$ such that

$$c < K \cdot \text{rad}(abc)^{1+\epsilon} \quad (79)$$

for any coprime integers a, b, c satisfying $a + b = c$.

Compared to n , $\text{rad}(n)$ is small when n can be divisible by high powers of primes. In other words, to violate (79), we need to have a, b, c to be divisible by many distinct primes so that $\text{rad}(abc)$ is small. It means that the conjecture is saying that the sum of two coprime integers that are divisible by high powers of primes is not divisible by high powers anymore. If we define the *quality* of numbers a, b, c as

$$q(a, b, c) := \frac{\log c}{\log \text{rad}(abc)},$$

then the current record for the highest quality is achieved by

$$(a, b, c) = (2, 3^{10} \cdot 109, 23^5), \quad q(a, b, c) \approx 1.6299.$$

There are numerous applications of the conjecture, which can be found in the [Wikipedia page](#). Here we introduce few of them, where some of them are proven (by different methods) or follows from the conjecture. The simplest example is the following weak version of FLT:

Theorem 6.3 (Weak FLT). Assume Conjecture 6.2 holds. If $n \geq 4$, then there are finitely many coprime integers x, y, z such that $x^n + y^n = z^n$.

Proof. Take $\epsilon = 1/4$ and $(a, b, c) = (x^n, y^n, z^n)$ in Conjecture 6.2. Then there exists a constant $K > 0$ such that

$$z^n < K \cdot \text{rad}(x^n y^n z^n)^{5/4} = K \cdot \text{rad}(xyz)^{5/4} \leq K \cdot (xyz)^{5/4} < K \cdot z^{15/4}$$

and there are only finitely many z satisfying this inequality. □

Another example is the Roth's theorem on Diophantine approximation.

Theorem 6.4 (Roth [27]). Let α be an irrational algebraic number. Then the inequality

$$\left| \alpha - \frac{p}{q} \right| < \frac{1}{q^{2+\epsilon}}$$

can have only finitely many solutions in coprime integers p, q for any $\epsilon > 0$.

Proof. Bombieri [4] gave a proof of Roth's theorem assuming Conjecture 6.2. In fact, he proved a stronger version of a Roth's theorem, which can be stated as an inequality on heights of algebraic numbers. See Frankenhuysen's note [15] for a detailed account of the proof. \square

The last example is the weak version of Hall's conjecture, which is about the difference of a cube and a square.

Theorem 6.5 (Weak Hall's conjecture [18]). Assume Conjecture 6.2 holds. For any $\epsilon > 0$, there exists a constant $K = K_\epsilon > 0$ such that $|y^2 - x^3| > K \cdot x^{1/2-\epsilon}$ for any integers x, y such that $y^2 - x^3 \neq 0$.

Exercise 6.2.1. Prove Theorem 6.5. Answer can be found in [29].

See Granville and Tucker's survey [17] for more applications of the conjecture.

Exercise 6.2.2. For any $K > 0$, prove that there exist infinitely many coprime integers a, b, c such that $a + b = c$ and $c > K \cdot \text{rad}(abc)$. In other words, the condition $\epsilon > 0$ in Conjecture 6.2 is necessary.²⁶

Let's move on to the polynomial analogue of Conjecture 6.2. As you expect, the radical of a polynomial $f \in k[T]$ is defined as the product of distinct irreducible monic factors of f . The polynomial version of the ABC conjecture is now a theorem by Mason and Stothers [21, 35].

Theorem 6.6 (Mason–Stothers theorem). Let k be a field and $f, g, h \in k[T]$ be mutually coprime polynomials such that $f + g + h = 0$ and not all of f', g', h' are zero. Then

$$\max\{\deg f, \deg g, \deg h\} \leq \deg \text{rad}(fgh) - 1. \quad (80)$$

We follow the proof of Snyder [33], which is similar to the proof of Theorem 6.1. The main step of the proof is the following lemma:

Lemma 6.7. For any nonzero $f \in k[T]$, $f/\text{rad}(f)$ divides f' .

Proof. This follows from factorization of f into irreducible monic factors. If we write $f = \prod_{i=1}^m f_i^{e_i}$ where f_i are distinct irreducible monic polynomials and $e_i \geq 1$, then

$$f' = \sum_{i=1}^m e_i f_i^{e_i-1} f'_i \prod_{j \neq i} f_j^{e_j}$$

$$\frac{f}{\text{rad}(f)} = \prod_{i=1}^m f_i^{e_i-1}$$

so $f/\text{rad}(f)$ divides f' . \square

Proof of Mason–Stothers theorem. For two polynomials $f, g \in k[T]$, we define the Wronskian $W(f, g)$ as $W(f, g) = fg' - gf'$. Then $f + g + h = 0$ implies that all three Wronskians $W(f, g), W(g, h), W(h, f)$ are the same:

$$W(f, g) = fg' - gf' = (-g - h)g' - g(-g' - h') = gh' - g'h = W(g, h).$$

²⁶Hint: Consider $(a, b, c) = (1, 2^{p(p-1)n} - 1, 2^{p(p-1)n})$ for large prime p and integer n . Prove that b is divisible by p^2 .

Let W be the common Wronskian of f, g, h . If $W = 0$, then $fg' = f'g$ and $\gcd(f, g) = 1$ gives $f \mid f'$ and $f' = 0$, and similarly $g' = 0$ and $h' = 0$, which contradicts the assumption that not all of f', g', h' are zero. Hence $W \neq 0$. By the previous Lemma, $f/\text{rad}(f), g/\text{rad}(g), h/\text{rad}(h)$ divide W , so does $fgh/\text{rad}(fgh)$ since f, g, h are mutually coprime. Thus

$$\begin{aligned} \deg f + \deg g + \deg h - \deg \text{rad}(fgh) &= \deg \left(\frac{fgh}{\text{rad}(fgh)} \right) \\ &\leq \deg W \\ &= \deg(fg' - gf') \\ &\leq \deg f + \deg g - 1 \end{aligned}$$

which implies $\deg h \leq \deg \text{rad}(fgh) - 1$. Similar inequality holds for $\deg f$ and $\deg g$ and we get the desired inequality. \square

This theorem has numerous applications. For example, one can prove non-solvability of the Fermat–Catalan equations, which is a generalization of the Fermat equation.

Theorem 6.8. Let k be a field and $p, q, r \geq 0$ be integers satisfying

$$\frac{1}{p} + \frac{1}{q} + \frac{1}{r} \leq 1 \quad (81)$$

and not divisible by the characteristic of k . Then there are no nonconstant and mutually coprime polynomials $f, g, h \in k[T]$ such that $f^p + g^q + h^r = 0$.

Proof. Apply the Mason–Stothers theorem to the polynomials f^p, g^q, h^r to show that $f' = g' = h' = 0$. When k has a positive characteristic, use the fact that $f' = 0$ implies that $f(T) = f_0(T^\ell)$ for some $f_0 \in k[T]$ and $\ell = \text{char}(k)$ and apply infinite descent. \square

Note that the finiteness result of the previous theorem for integers and fixed p, q, r satisfying $1/p + 1/q + 1/r < 1$ is a consequence of the original ABC conjecture, and also proved by Darmon and Granville [10] using Falting’s theorem.

We also have a version of the weak Hall’s conjecture for polynomials, which is a theorem of Davenport [11].

Theorem 6.9 (Davenport [11]). Let k be a field and $f, g \in k[T]$ be mutually coprime polynomials with nonzero derivatives such that $f^3 - g^2 \neq 0$. Then $\deg f + 2 \leq 2 \deg(f^3 - g^2)$.

Exercise 6.2.3. Prove Davenport’s theorem using the Mason–Stothers theorem.

Exercise 6.2.4. Search for the result that follows from ABC conjecture for integers, and prove polynomial analogue of the result (if exists) using Mason–Stothers theorem.

A Appendix

A.1 Dictionary between integers and polynomials

Here we summarize the dictionary between the integers and the polynomials over finite fields. Here all g, g_1, g_2, \dots on the A -side are all irreducible polynomials over \mathbb{F}_p . Also, we omit all the technical assumptions for each theorem, which can be found in the main text.

	\mathbb{Z}	$A = \mathbb{F}_p[T]$
indecomposable	prime	irreducible
number of units	$2 = \#(\mathbb{Z}^\times)$	$p - 1 = \#(\mathbb{F}_p[T]^\times) = \#(\mathbb{F}_p^\times)$
normalized	\mathbb{N}	monic polynomials
absolute value	$ n = \#(\mathbb{Z}/n\mathbb{Z})$	$ f = \#(A/fA) = p^{\deg(f)}$
Euler φ function	$\varphi(n) = \#(\mathbb{Z}/n\mathbb{Z})^\times$	$\varphi(f) = \#(A/fA)^\times$
Fermat's little theorem	$a^{p-1} \equiv 1 \pmod{p}$	$f^{ g -1} \equiv 1 \pmod{g}$
Euler's theorem	$a^{\varphi(n)} \equiv 1 \pmod{n}$	$f^{\varphi(f_0)} \equiv 1 \pmod{f_0}$
Wilson's theorem	$(p-1)! \equiv -1 \pmod{p}$	$\prod_{f \in (A/gA)^\times} f \equiv -1 \pmod{g}$
Quadratic reciprocity	$\left(\frac{p}{q}\right) \left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2} \frac{q-1}{2}}$	$\left(\frac{g_1}{g_2}\right) \left(\frac{g_2}{g_1}\right) = (-1)^{\frac{ g_1 -1}{2} \frac{ g_2 -1}{2}}$
Riemann Zeta function	$\zeta(s) = \sum_{n \geq 1} \frac{1}{n^s}$	$\zeta_A(s) = \sum_{f \neq 0, \text{ monic}} \frac{1}{ f ^s}$
Euler factorization	$\zeta(s) = \prod_p (1 - p^{-s})^{-1}$	$\zeta_A(s) = \prod_g (1 - g ^{-s})^{-1}$
Möbius function	$\mu(p_1 \cdots p_k) = (-1)^k$	$\mu(g_1 \cdots g_k) = (-1)^k$
Sum of arithmetic functions	$\sum_{n \leq x} a_n$	$\sum_{f, \deg(f)=n} a_f$
Dirichlet series	$L(s) = \sum_{n \geq 1} \frac{a_n}{n^s}$	$L(s) = \sum_{0 \neq f, \text{ monic}} \frac{a_f}{ f ^s}$

Table 1: Integers and Polynomials.

References

- [1] APOSTOL, T. M. *Introduction to analytic number theory*. Springer Science & Business Media, 2013.
- [2] AVIGAD, J., AND MORRIS, R. The concept of “character” in dirichlet’s theorem on primes in an arithmetic progression. *Archive for history of exact sciences* 68 (2014), 265–326.
- [3] BAUMGART, O. *The quadratic reciprocity law: A collection of classical proofs*. Birkhäuser, 2015.
- [4] BOMBIERI, E. Roth’s theorem and the abc-conjecture. *preprint ETH Zürich* (1994).
- [5] CHA, B. Chebyshev’s bias in function fields. *Compositio Mathematica* 144, 6 (2008), 1351–1374.
- [6] CHEBYSHEV, P. L. Lettre de m. le professeur tchébychev à m. fuss sur un nouveau théorème relatif aux nombres premiers contenus dans les formes $4n + 1$ et $4n + 3$. *Bull. Classe Phys. Acad. Imp. Sci. St. Petersburg* 11 (1853), 208.
- [7] CONRAD, K. Quadratic reciprocity in characteristic 2. <https://kconrad.math.uconn.edu/blurbs/ugradnumthy/QRchar2.pdf>. Accessed: 2025-06-05.
- [8] CONRAD, K. Quadratic reciprocity in odd characteristic. <https://kconrad.math.uconn.edu/blurbs/ugradnumthy/QRcharp.pdf>. Accessed: 2025-06-05.
- [9] CORNELL, G., SILVERMAN, J. H., AND STEVENS, G. *Modular forms and Fermat’s last theorem*. Springer Science & Business Media, 2013.
- [10] DARMON, H., AND GRANVILLE, A. On the equations $z^m = f(x, y)$ and $ax^p + by^q = cz^r$. *Bulletin of the London Mathematical Society* 27, 6 (1995), 513–543.
- [11] DAVENPORT, H. On $f^3(t) - g^2(t)$. *Norske Vid. Selsk. Forh.(Trondheim)* 38 (1965), 86–87.
- [12] DE LA VALLÉE POUSSIN, C. Sur la fonction $\zeta(s)$ de Riemann et le nombre des nombres premiers inférieurs à une limite donnée. *Mémoires de l’Académie royale de Belgique* 59, 1 (1899), 1–74.
- [13] DIRICHLET, P. L. Beweis des Satzes, dass jede unbegrenzte arithmetische Progression, deren erstes Glied und Differenz ganze Zahlen ohne gemeinschaftlichen Factor sind, unendlich viele Primzahlen enthält. *Abhandlungen der Königlich Preussischen Akademie der Wissenschaften* 45 (1837), 81.
- [14] ERDÖS, P. On a new method in elementary number theory which leads to an elementary proof of the prime number theorem. *Proceedings of the National Academy of Sciences* 35, 7 (1949), 374–384.
- [15] FRANKENHUYSEN, M. V. The abc conjecture implies roth’s theorem and mordell’s conjecture. <https://swc-math.github.io/notes/files/98Frankenhuysen.pdf>. Accessed: 2025-07-06.
- [16] GRANVILLE, A., AND MARTIN, G. Prime number races. *The American Mathematical Monthly* 113, 1 (2006), 1–33.
- [17] GRANVILLE, A., AND TUCKER, T. It’s as easy as abc. *Notices of the AMS* 49, 10 (2002), 1224–1231.
- [18] HALL JR, M. The diophantine equation $x^3 - y^2 = k$, 1971.

- [19] LEMMERMEYER, F. Proofs of the Quadratic Reciprocity Law. https://web.archive.org/web/20250106010310/https://www.mathi.uni-heidelberg.de/%7Eflemmermeyer/qrg_proofs.html. Accessed: 2025-06-05.
- [20] LITTLEWOOD, J. E. Sur la distribution des nombres premiers. *CR Acad. Sci. Paris* 158, 1869 (1914), 1872.
- [21] MASON, R. C. *Diophantine equations over function fields*, vol. 96. Cambridge University Press, 1984.
- [22] MURTY, M. R. Primes in Certain Arithmetic Progressions. *J. Madras Univ.* 51 (1988), 161–169.
- [23] NEWMAN, D. J. Simple analytic proof of the prime number theorem. *The American Mathematical Monthly* 87, 9 (1980), 693–696.
- [24] PASTEN, H. A Derivation of the Infinitude of Primes. *The American Mathematical Monthly* 131, 1 (2024), 66–73.
- [25] RIEMANN, B. Ueber die Anzahl der Primzahlen unter einer gegebenen Grosse. *Ges. Math. Werke und Wissenschaftlicher Nachlaß* 2, 145-155 (1859), 2.
- [26] ROSEN, M. *Number theory in function fields*, vol. 210. Springer Science & Business Media, 2013.
- [27] ROTH, K. F. Rational approximations to algebraic numbers. *Mathematika* 2, 1 (1955), 1–20.
- [28] RUBINSTEIN, M., AND SARNAK, P. Chebyshev’s bias. *Experimental Mathematics* 3, 3 (1994), 173–197.
- [29] SCHMIDT, W. M. *Diophantine approximations and Diophantine equations*. Springer, 2006.
- [30] SCHUR, I. *Über die Existenz unendlich vieler Primzahlen in einigen speziellen arithmetischen Progressionen*. 1912.
- [31] SELBERG, A. An elementary proof of the prime-number theorem. *Annals of Mathematics* 50, 2 (1949), 305–313.
- [32] SERRE, J.-P. *A course in arithmetic*, vol. 7. Springer Science & Business Media, 2012.
- [33] SNYDER, N. An alternate proof of Mason’s theorem. *Elemente der Mathematik* 55, 3 (2000), 93–94.
- [34] STEIN, E. M., AND SHAKARCHI, R. *Complex analysis*, vol. 2. Princeton University Press, 2010.
- [35] STOTHERS, W. W. Polynomial identities and Hauptmoduln. *The Quarterly Journal of Mathematics* 32, 3 (1981), 349–370.
- [36] TAYLOR, R., AND WILES, A. Ring-theoretic properties of certain hecke algebras. *Annals of Mathematics* 141, 3 (1995), 553–572.
- [37] VON KOCH, H. Sur la distribution des nombres premiers. *Acta Mathematica* 24, 1 (1901), 159.
- [38] WILES, A. Modular elliptic curves and fermat’s last theorem. *Annals of mathematics* 141, 3 (1995), 443–551.
- [39] ZAGIER, D. Newman’s short proof of the prime number theorem. *The American mathematical monthly* 104, 8 (1997), 705–708.