Basic Number Theory Supplementary Material

Timo Chang
timo65537@protonmail.com

Last edited: August 30, 2026

Abstract

These lecture notes are supplementary material written by the author and delivered in his recitation classes as a teaching assistant for the *Basic Number Theory* course in Spring 2024. The main goal of these sessions is to cover topics outside standard introductory courses and explore as many areas of number theory as possible. We assume the students are beginners in mathematics with a minimal background in calculus and linear algebra, but no prior knowledge of abstract algebra or analysis.

Contents

1 Liouville's Theorem and Liouville Number

1.1 Algebraic and Transcendental Numbers

Definition 1.1. A number $\alpha \in \mathbb{C}$ is called *algebraic* if it is a root of some polynomial $0 \neq f(x) \in \mathbb{Q}[x]$[1]. Otherwise, it is called *transcendental*.

We let $\overline{\mathbb{Q}}$ be the set of all algebraic numbers.

Example 1.2. It is not hard to come up with some examples of algebraic numbers. For instance, all rational numbers, $\sqrt{2}$, $i = \sqrt{-1}$, $(-1+\sqrt{-3})/2$, or more generally, the n-th root of unity $e^{2\pi i/n}$, are all algebraic.

In fact, using "field theory", one can show that if $\alpha, \beta \in \overline{\mathbb{Q}}$, then so are $\alpha \pm \beta, \alpha\beta, \alpha/\beta$ (when $\beta \neq 0$). In other words, $\overline{\mathbb{Q}}$ is a "field". (A constructive proof of this fact uses the notion of "resultant", which requires only a basic knowledge of linear algebra. See my note [Chaa] for further discussion.)

For example, take $\alpha = \sqrt{2}, \beta = \sqrt{3}$, and let $x = \sqrt{2} + \sqrt{3}$. Then we have $x^2 = 5 + 2\sqrt{6}$, and so

$$x^4 - 10x^2 + 1 = 0.$$

This shows that x is an algebraic number as well.

Example 1.3. For transcendental numbers, one might wonder whether the constants e and π serve as examples. This is indeed the case: both of them are proven to be transcendental, meaning that they are not roots of any non-zero rational polynomials. (See my note [Chab] for an elementary proof of the transcendence of e using only calculus.)

Some important "transcendental" functions also provide examples of transcendental numbers. For instance,

$$\log 2, \quad \zeta(2) := \sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}, \quad \text{and} \quad \Gamma\left(\frac{1}{2}\right) := \int_0^{\infty} t^{-\frac{1}{2}} e^{-t} dt = \sqrt{\pi}$$

are all transcendental. (We will prove the second identity in Lecture 2. See Theorem 2.3 and Example 2.4.)

While the constants e and π are known to be transcendental, they were actually not the first proven examples in history. Rather, it was constructed by Liouville in 1844 after he discovered a certain approximation property of irrational algebraic numbers, which is now called *Liouville constant*. In this week's lecture, we aim to present his theorem and construction.

[1] The notation $\mathbb{Q}[x]$ means the set of all polynomials with rational coefficients.

1.2 Liouville's Theorem on Diophantine Approximation

Our first goal is to prove the following theorem.

Theorem 1.4 (Liouville's Theorem on Diophantine Approximation)**.** *Let α be an irrational algebraic number. Suppose α satisfies an irreducible polynomial $f(x) = \sum_{i=0}^{n} c_i x^i \in \mathbb{Z}[x]$ with degree n. Then there exists an $A > 0$ such that for any $p, q \in \mathbb{Z}$ and $q > 0$, we have*

$$\left|\alpha - \frac{p}{q}\right| > \frac{A}{q^n}.$$

Roughly speaking, Theorem 1.4 says that every irrational algebraic number α can not be approximated very "nicely" by rational numbers, because there is always a gap A/q^n between α and any rational number p/q. In other words, if an irrational number α can be approximated nicely (to certain degree) by rational numbers, then it must be transcendental.

Proof. Since f is a polynomial, we know f' is a continuous function. So by extreme value theorem, $M := \max_{[\alpha-1,\alpha+1]} |f'(x)|$ is finite. Let $\{\alpha_1, \ldots, \alpha_m\}$ be the set of distinct roots of f other than α, and set

$$0 < A < \min\{1, 1/M, |\alpha - \alpha_1|, \ldots, |\alpha - \alpha_m|\}.$$

We claim that this A satisfies our desired property.

Suppose on the contrary, there exist $p, q \in \mathbb{Z}$ and $q > 0$ such that

$$\left|\alpha - \frac{p}{q}\right| \leq \frac{A}{q^n}. \tag{1.1}$$

It follows that

$$\left|\alpha - \frac{p}{q}\right| \leq \frac{A}{q^n} \leq A < 1,$$

and so $p/q \in [\alpha - 1, \alpha + 1]$. By mean value theorem, we have

$$f(\alpha) - f\left(\frac{p}{q}\right) = f'(\xi)\left(\alpha - \frac{p}{q}\right) \tag{1.2}$$

for some ξ between α and p/q. Note that $\xi \in [\alpha - 1, \alpha + 1]$ because p/q also lies in this interval. So by the definition of M, we have

$$M \geq |f'(\xi)|.$$

Note that $f(\alpha) = 0$ because α is a root of f. And from $|\alpha - p/q| \leq A < |\alpha - \alpha_i|$ for all $i = 1, \ldots, m$, we know p/q cannot be any of $\alpha_1, \ldots, \alpha_m$. In other words, $f(p/q) \neq 0$. These two together with (1.2) imply that $f'(\xi) \neq 0$. So we may write

$$\left|\alpha - \frac{p}{q}\right| = \frac{|f(\alpha) - f(p/q)|}{|f'(\xi)|} = \frac{|f(p/q)|}{|f'(\xi)|}.$$

We see that the numerator

$$\begin{aligned}\left|f\left(\frac{p}{q}\right)\right| &= \left|c_0 + c_1\frac{p}{q} + \cdots + c_n\left(\frac{p}{q}\right)^n\right| \\ &= \frac{1}{q^n}\left|c_0q^n + c_1pq^{n-1} + \cdots + c_np^n\right| \\ &\geq \frac{1}{q^n}.\end{aligned}$$

The last equality is because $c_0, \ldots, c_n, p, q$ are all integers and their combination in the last absolute value is non-zero (as $f(p/q) \neq 0$). We can now combine all inequalities together and see that

$$\left|\alpha - \frac{p}{q}\right| = \frac{|f(p/q)|}{|f'(\xi)|} \geq \frac{1}{|f'(\xi)|q^n} \geq \frac{1}{Mq^n} > \frac{A}{q^n}.$$

This contradicts to (1.1). □

1.3 Liouville Number

In the last section, we proved an important theorem describing an approximation property of irrational algebraic numbers due to Liouville. This enabled him to construct the very first example of transcendental numbers. Now, we're in position to see how this was done.

Definition 1.5 (Liouville Number)**.** A number $\alpha \in \mathbb{R}$ is called a *Liouville number* if for any $n \in \mathbb{N}$, there exist $a, b \in \mathbb{Z}$ and $b > 1$ such that

$$0 < \left|\alpha - \frac{a}{b}\right| < \frac{1}{b^n}. \tag{1.3}$$

By definition, every Liouville number is approximated nicely by rational numbers to some extent. In view of Theorem 1.4, this suggests that every Liouville number is transcendental. And this is indeed the case. We will prove this in two steps.

Proposition 1.6. *All Liouville numbers are irrational.*

Proof. We suppose on the contrary that there exists a Liouville number α which is rational. Then we may write $\alpha = p/q$ where $p, q \in \mathbb{Z}$ and $q \geq 1$. Take $n \in \mathbb{N}$ large enough so that $2^{n-1} > q$. Then given this particular n, we claim that for any $a, b \in \mathbb{Z}$ and $b > 1$, one of the inequalities in (1.3) fails. This will imply that α is not a Liouville number, which leads to a contradiction.

Note that $|\alpha - a/b| = |p/q - a/b| = |(pb - aq)/qb|$.

- Case 1: $|pb - aq| = 0$. Then $|\alpha - a/b| = 0$. This contradicts to the first inequality in (1.3).

- Case 2: $|pb - aq| \geq 1$. Then

$$\left|\alpha - \frac{a}{b}\right| \geq \frac{1}{bq} > \frac{1}{b2^{n-1}} \geq \frac{1}{b \cdot b^{n-1}} = \frac{1}{b^n}.$$

The second inequality is due to our choice of n, and the third inequality is because $b \geq 2$. But this contradicts to the second inequality in (1.3).

□

Proposition 1.7. *All Liouville numbers are transcendental.*

Proof. We again suppose on the contrary that there exists a Liouville number α which is algebraic. By Proposition 1.6, we know α is irrational. So it satisfies the assumption of Theorem 1.4. Hence, there exists $n \in \mathbb{N}$ and $A > 0$ such that for any $p, q \in \mathbb{Z}$ and $q > 0$, we have

$$\left|\alpha - \frac{p}{q}\right| > \frac{A}{q^n}. \tag{1.4}$$

We take $r \in \mathbb{N}$ large enough so that $2^r > 1/A$. Then since α is a Liouville number, by Definition 1.5, we can find $a, b \in \mathbb{Z}$ and $b > 1$ such that $0 < |\alpha - a/b| < 1/b^{n+r}$. But this implies

$$0 < \left|\alpha - \frac{a}{b}\right| < \frac{1}{b^{n+r}} = \frac{1}{b^n b^r} \leq \frac{1}{b^n 2^r} < \frac{A}{b^n},$$

which contradicts to (1.4). □

We have shown that all Liouville numbers are transcendental. An example of such numbers was given by Liouville himself, called *Liouville's constant*.

Example 1.8 (Liouville's Constant)**.** Define the infinite summation

$$L := \sum_{m=1}^{\infty} \frac{1}{10^{m!}} = 0.110\ 001\ 000\ 00\cdots.$$

Observe that its decimal representation has 1 at every $m!$-th digit and 0 otherwise. In particular, L is irrational.[2] We claim that L is a Liouville number, and hence, by Proposition 1.7, L is transcendental.

Given any $n \in \mathbb{N}$, write

$$L = \sum_{m=1}^{n} \frac{1}{10^{m!}} + \sum_{m=n+1}^{\infty} \frac{1}{10^{m!}} = \frac{a'}{10^{n!}} + \sum_{m=n+1}^{\infty} \frac{1}{10^{m!}}$$

[2]Recall that a real number is rational if and only if its decimal representation is either terminating or repeating.

where a' is whatever it takes to make the first sum become a single fraction. Now, we take $a := a'$ and $b := 10^{n!}$. It follows that

$$0 < \left| L - \frac{a}{b} \right| = \sum_{m=n+1}^{\infty} \frac{1}{10^{m!}} = \frac{1}{10^{(n+1)!}} + \frac{1}{10^{(n+1)!}} \sum_{m=n+2}^{\infty} \frac{1}{10^{m!-(n+1)!}}$$
$$\leq \frac{1}{10^{(n+1)!}} + \frac{1}{10^{(n+1)!}} \sum_{m=1}^{\infty} \frac{1}{2^m} = \frac{2}{10^{(n+1)!}} < \frac{1}{10^{(n+1)!-1}} \leq \left(\frac{1}{10^{n!}} \right)^n$$

where the last inequality is due to the fact that $(n+1)! - 1 \geq n(n!)$ for all $n \in \mathbb{N}$. So by Definition 1.5, L is a Liouville number.

Exercise 1.9. Following the idea of Example 1.8, give another example of a Liouville number.

2 Riemann Zeta Function

The Riemann zeta function, undoubtedly one of the most important functions in number theory, is known to possess rich arithmetic and analytic properties. For example, the famous Riemann hypothesis concerns the locations of its "zeros", which are closely related to the distribution of prime numbers. In this week's lecture, we aim to prove some basic properties of this function, including the Euler product, its special values at positive even integers, and Riemann's "analytic continuation" and functional equation. (We will need some tools from complex analysis. However, all technical terms and proofs will be kept at a minimum.)

2.1 Euler's Formula

Definition 2.1. The *Riemann zeta function* is the complex function defined by the series

$$\zeta(s) := \sum_{n=1}^{\infty} \frac{1}{n^s}, \quad \mathrm{Re}(s) > 1. \tag{2.1}$$

The first step to connect the Riemann zeta function with prime numbers is through the following Euler product.

Proposition 2.2 (Euler)**.** *For* $\mathrm{Re}(s) > 1$*, we have*

$$\zeta(s) = \prod_{p} (1 - p^{-s})^{-1}$$

where the product runs through all prime numbers.

Proof. Since $|p^{-s}| < 1$ for all prime number p and $s \in \mathbb{C}$ with $\mathrm{Re}(s) > 1$, by the geometric series, the right-hand side is seen to be

$$\begin{aligned}&\prod_p (1-p^{-s})^{-1} \\ &= \left(\frac{1}{1} + \frac{1}{2^s} + \frac{1}{2^{2s}} + \cdots\right)\left(\frac{1}{1} + \frac{1}{3^s} + \frac{1}{3^{2s}} + \cdots\right)\left(\frac{1}{1} + \frac{1}{5^s} + \frac{1}{5^{2s}} + \cdots\right)\cdots.\end{aligned}$$

Now, expanding the product on the right, we see that this coincides with the series expression of the Riemann zeta function (2.1) by the fundamental theorem of arithmetic.[3] □

Our main goal in this section is to prove the following Euler's formula of the Riemann zeta function at positive even integers.

Theorem 2.3 (Euler)**.** *For a positive even integer $n \in \mathbb{N}$, we have*

$$\zeta(n) = -\frac{B_n}{2 \cdot n!}(2\pi i)^n$$

where $B_k \in \mathbb{Q}$ is the Bernoulli number defined by

$$\frac{t}{e^t - 1} = \sum_{k=0}^{\infty} B_k \frac{t^k}{k!}.$$

Note that due to the transcendence of π (recall Example 1.3), this formula implies that the Riemann zeta function is transcendental at positive even integers. However, the transcendence at positive odd integers is still widely open.

Example 2.4. A simple calculation of the Taylor expansion shows that

$$\frac{t}{e^t - 1} = 1 - \frac{t}{2} + \frac{t^2}{12} - \frac{t^4}{720} + \cdots.$$

Thus, we have

$$B_0 = 1, \quad B_1 = -\frac{1}{2}, \quad B_2 = \frac{1}{6}, \quad B_3 = 0, \quad B_4 = -\frac{1}{30}.$$

Then by Theorem 2.3, we have (recall Example 1.3)

$$\zeta(2) = \frac{\pi^2}{6}, \quad \zeta(4) = \frac{\pi^4}{90},$$

and so on.

[3] For example, the term $1/12^s$ in (2.1) comes from

$$\frac{1}{12^s} = \frac{1}{(2^2 \cdot 3)^s} = \frac{1}{2^{2s}} \cdot \frac{1}{3^s} \cdot \frac{1}{1} \cdots.$$

Proof. It is a fact from complex analysis that the sine function has the following infinite product formula

$$\sin(\pi z) = \pi z \prod_{n=1}^{\infty} \left(1 - \frac{z^2}{n^2}\right).$$

From here, we take the log derivative and get

$$\pi \cot(\pi z) = \frac{\pi \cos(\pi z)}{\sin(\pi z)} = \frac{1}{z} + \sum_{n=1}^{\infty} \frac{-\frac{2z}{n^2}}{1 - \frac{z^2}{n^2}} = \frac{1}{z} + \sum_{n=1}^{\infty} \left(\frac{1}{n+z} - \frac{1}{n-z}\right). \tag{2.2}$$

On the other hand, by complex analysis, we have the following exponential expressions of trigonometric functions

$$\cos(z) = \frac{e^{iz} + e^{-iz}}{2} \quad \text{and} \quad \sin(z) = \frac{e^{iz} - e^{-iz}}{2i}.$$

These two imply

$$\cot(z) = i \cdot \frac{e^{iz} + e^{-iz}}{e^{iz} - e^{-iz}} = i + \frac{2i}{e^{2iz} - 1},$$

and so

$$\pi \cot(\pi z) = \pi i + \frac{2\pi i}{e^{2\pi i z} - 1}. \tag{2.3}$$

Equating (2.2) and (2.3), we get

$$\pi i + \frac{2\pi i}{e^{2\pi i z} - 1} = \frac{1}{z} + \sum_{n=1}^{\infty} \left(\frac{1}{n+z} - \frac{1}{n-z}\right).$$

And so

$$\pi i z + \frac{2\pi i z}{e^{2\pi i z} - 1} = 1 + \sum_{n=1}^{\infty} \left(\frac{2\pi i z}{2\pi i n + 2\pi i z} - \frac{2\pi i z}{2\pi i n - 2\pi i z}\right).$$

Let $t = 2\pi i z$, then we get

$$\frac{t}{2} + \frac{t}{e^t - 1} = 1 + \sum_{n=1}^{\infty} \left(\frac{t}{2\pi i n + t} - \frac{t}{2\pi i n - t}\right).$$

The left-hand side is by definition,

$$\text{LHS} = \frac{t}{2} + \sum_{k=0}^{\infty} B_k \frac{t^k}{k!}.$$

And the right-hand side is

$$\begin{aligned}
\text{RHS} &= 1 + \sum_{n=1}^{\infty} \frac{t}{2\pi i n} \left(\frac{1}{1+\frac{t}{2\pi i n}} - \frac{1}{1-\frac{t}{2\pi i n}} \right) \\
&= 1 + \sum_{n=1}^{\infty} \frac{t}{2\pi i n} \left(\sum_{k=0}^{\infty} \left(-\frac{t}{2\pi i n} \right)^k - \sum_{k=0}^{\infty} \left(\frac{t}{2\pi i n} \right)^k \right) \\
&= 1 + \sum_{n=1}^{\infty} \frac{-2t}{2\pi i n} \left(\sum_{k=0}^{\infty} \left(\frac{t}{2\pi i n} \right)^{2k+1} \right) \\
&= 1 - 2 \sum_{n=1}^{\infty} \sum_{k=0}^{\infty} \left(\frac{t}{2\pi i n} \right)^{2k+2} \\
&= 1 - 2 \sum_{k=0}^{\infty} \left(\sum_{n=1}^{\infty} \frac{1}{(2\pi i n)^{2k+2}} \right) t^{2k+2}.
\end{aligned}$$

(In the last equality, we should be careful when changing the order of summations. But here, we will omit such a detail.) So we have the equality

$$\frac{t}{2} + \sum_{k=0}^{\infty} B_k \frac{t^k}{k!} = 1 - 2 \sum_{k=0}^{\infty} \left(\sum_{n=1}^{\infty} \frac{1}{(2\pi i n)^{2k+2}} \right) t^{2k+2}. \tag{2.4}$$

Now, comparing the coefficients of t^m for positive even integer m, we have

$$\frac{B_m}{m!} = -2 \sum_{n=1}^{\infty} \frac{1}{(2\pi i n)^m} = \frac{-2}{(2\pi i)^m} \sum_{n=1}^{\infty} \frac{1}{n^m} = \frac{-2}{(2\pi i)^m} \cdot \zeta(m).$$

So finally,

$$\zeta(m) = -\frac{B_m}{2 \cdot m!} (2\pi i)^m.$$

□

Corollary 2.5. *As $i \in \mathbb{C} \setminus \mathbb{R}$ and $B_k \in \mathbb{Q}$, we see that for $n \geq 2$,*

$$\frac{\zeta(n)}{(2\pi i)^n} \in \mathbb{Q} \iff n \text{ is even.}$$

Corollary 2.6. *By comparing the coefficients of other terms in (2.4), we have (i) $B_0 = 1$ (ii) $B_1 = -1/2$ (iii) $B_k = 0$ for odd $k > 1$. (That is, the k-th Bernoulli number vanishes when $k > 1$ is odd.) This matches our calculation in Example 2.4.*

2.2 Riemann's Functional Equation and Riemann Hypothesis

In complex analysis, there is a concept called "analytic[4] continuation", which means to extend the domain of an analytic function to a larger domain. For example, consider the geometric series $1+z+z^2+\cdots$, which defines an analytic function on $D := \{z \in \mathbb{C} \mid |z| < 1\}$, and we have

$$1 + z + z^2 + \cdots = \frac{1}{1-z}, \quad z \in D.$$

However, the function on the right is not only analytic on D but $\mathbb{C} \setminus \{1\}$. Thus, we can use this expression to extend the domain of the geometric series, and call the right-hand side its analytic continuation.

With some intuition of analytic continuation, in this section, we show that the Riemann zeta function $\zeta(s)$ admits such a continuation and derive its functional equation. We begin by defining the *Euler gamma function* as the integral

$$\Gamma(s) := \int_0^\infty t^s e^{-t} \frac{dt}{t}, \quad \mathrm{Re}(s) > 0. \tag{2.5}$$

Theorem 2.7 (Riemann)**.** *Let*

$$\xi(s) := \pi^{-\frac{s}{2}} \Gamma\left(\frac{s}{2}\right) \zeta(s), \quad \mathrm{Re}(s) > 1.$$

Then $\xi(s)$ has an analytic continuation to $\mathbb{C} \setminus \{0, 1\}$ and satisfies the functional equation

$$\xi(s) = \xi(1-s).$$

Proof. From (2.5), we substitute $s \mapsto s/2$, multiply the factor $\pi^{-s/2} n^{-s}$ on both sides, do some change of variables, and get

$$\begin{aligned}
\pi^{-\frac{s}{2}} n^{-s} \Gamma\left(\frac{s}{2}\right) &= \int_0^\infty \pi^{-\frac{s}{2}} n^{-s} t^{\frac{s}{2}} e^{-t} \frac{dt}{t} \\
&= \int_0^\infty \left(\frac{t}{\pi n^2}\right)^{\frac{s}{2}} e^{-t} \frac{dt}{t} \\
&= \int_0^\infty t^{\frac{s}{2}} e^{-\pi n^2 t} \frac{dt}{t} \qquad \left(\frac{t}{\pi n^2} \mapsto t\right).
\end{aligned}$$

Summing both sides from 1 to ∞ with respect to n, we get

$$\begin{aligned}
\sum_{n=1}^\infty \pi^{-\frac{s}{2}} n^{-s} \Gamma\left(\frac{s}{2}\right) &= \sum_{n=1}^\infty \int_0^\infty t^{\frac{s}{2}} e^{-\pi n^2 t} \frac{dt}{t} \\
&= \frac{1}{2} \sum_{\substack{n=-\infty \\ n \neq 0}}^\infty \int_0^\infty t^{\frac{s}{2}} e^{-\pi n^2 t} \frac{dt}{t} \\
&= \int_0^\infty \left(\frac{\theta(it) - 1}{2}\right) t^{\frac{s}{2}} \frac{dt}{t}.
\end{aligned} \tag{2.6}$$

[4]Roughly speaking, an analytic function is a function which "locally" admits a power series expansion.

Here,

$$\theta(z) := \sum_{n\in\mathbb{Z}} e^{\pi i n^2 z}, \quad \mathrm{Im}(z) > 0,$$

is the *Jacobi theta series*. (Be careful that in the last equality, we have changed the order of summation and integration without justification.)

Note that by the definition of $\xi(s)$, the left-hand side of (2.6) is

$$\sum_{n=1}^{\infty} \pi^{-\frac{s}{2}} n^{-s} \Gamma\left(\frac{s}{2}\right) = \pi^{-\frac{s}{2}} \Gamma\left(\frac{s}{2}\right) \zeta(s) = \xi(s).$$

So we get

$$\xi(s) = \int_0^{\infty} \left(\frac{\theta(it) - 1}{2}\right) t^{\frac{s}{2}} \frac{dt}{t}. \tag{2.7}$$

In order to calculate the right-hand side of (2.7), we need the following functional equation of the theta series, which can be proved by the "Poisson summation formula".

Lemma 2.8. *The theta series $\theta(z)$ satisfies the functional equation*

$$\theta\left(-\frac{1}{z}\right) = \sqrt{\frac{z}{i}}\theta(z), \quad \mathrm{Im}(z) > 0.$$

In particular,

$$\theta\left(\frac{i}{t}\right) = t^{\frac{1}{2}}\theta(it), \quad t > 0.$$

Returning to the right-hand side of (2.7), we see that

$$\begin{aligned}
&\int_0^{\infty} \left(\frac{\theta(it) - 1}{2}\right) t^{\frac{s}{2}} \frac{dt}{t} \\
&= \frac{1}{2}\int_0^1 (\theta(it) - 1) t^{\frac{s}{2}} \frac{dt}{t} + \boxed{\frac{1}{2}\int_1^{\infty} (\theta(it) - 1) t^{\frac{s}{2}} \frac{dt}{t}} \\
&= \frac{1}{2}\int_1^{\infty} \left(\theta\left(\frac{i}{t}\right) - 1\right) t^{-\frac{s}{2}} \frac{dt}{t} + \boxed{} && \left(t \mapsto \frac{1}{t}\right) \\
&= \frac{1}{2}\int_1^{\infty} \left(t^{\frac{1}{2}}\theta(it) - 1\right) t^{-\frac{s}{2}} \frac{dt}{t} + \boxed{} && \text{(by Lemma 2.8)} \\
&= \frac{1}{2}\int_1^{\infty} \left((\theta(it) - 1) t^{\frac{1}{2}(1-s)} + t^{\frac{1}{2}(1-s)} - t^{-\frac{s}{2}}\right) \frac{dt}{t} + \boxed{} \\
&= \frac{1}{2}\int_1^{\infty} (\theta(it) - 1) t^{\frac{1}{2}(1-s)} \frac{dt}{t} - \frac{1}{1-s} - \frac{1}{s} + \boxed{\frac{1}{2}\int_1^{\infty} (\theta(it) - 1) t^{\frac{s}{2}} \frac{dt}{t}}.
\end{aligned}$$

So we get

$$\xi(s) = \frac{1}{2}\int_1^{\infty} (\theta(it) - 1) t^{\frac{1}{2}(1-s)} \frac{dt}{t} + \frac{1}{2}\int_1^{\infty} (\theta(it) - 1) t^{\frac{s}{2}} \frac{dt}{t} - \frac{1}{1-s} - \frac{1}{s}. \tag{2.8}$$

It can be shown that the integrals in (2.8) are analytic for all $s \in \mathbb{C}$. Thus, using this expression, $\xi(s)$ is now analytically continued to $\mathbb{C}\setminus\{0,1\}$. (Note the extra terms $-1/(1-s)$ and $-1/s$, which make $\xi(s)$ undefined at $s = 0, 1$.) Moreover, note that the expression on the right is unchanged under the substitution $s \mapsto 1 - s$. Thus, the same is true for the left-hand side. In other words, we have

$$\xi(s) = \xi(1 - s).$$

This completes the proof. □

Remark 2.9. Using the language of complex analysis, (2.8) actually shows that $\xi(s)$ has a "meromorphic continuation" to the whole complex plane. It is "holomorphic" on $\mathbb{C} \setminus \{0, 1\}$ and has "simple pole" at $s = 0, 1$ with "residue" $-1, 1$, respectively.[5]

Corollary 2.10. *The Riemann zeta function $\zeta(s)$ has an analytic continuation to $\mathbb{C} \setminus \{1\}$ and satisfies the functional equation*

$$\zeta(s) = 2^s \pi^{s-1} \sin\left(\frac{\pi s}{2}\right) \Gamma(1 - s)\zeta(1 - s).$$

Proof. We only prove the second assertion.[6] From Theorem 2.7, we multiply $\Gamma(1 - s/2)$ on both sides and get

$$\pi^{-\frac{s}{2}}\Gamma\left(\frac{s}{2}\right)\Gamma\left(1 - \frac{s}{2}\right)\zeta(s) = \pi^{-\frac{1-s}{2}}\Gamma\left(\frac{1-s}{2}\right)\Gamma\left(1 - \frac{s}{2}\right)\zeta(1 - s). \tag{2.9}$$

Using the reflection formula of the gamma function:

$$\Gamma(z)\Gamma(1 - z) = \frac{\pi}{\sin(\pi z)}, \quad z \notin \mathbb{Z},$$

the left-hand side of (2.9) becomes

$$\pi^{-\frac{s}{2}}\Gamma\left(\frac{s}{2}\right)\Gamma\left(1 - \frac{s}{2}\right)\zeta(s) = \pi^{-\frac{s}{2}}\frac{\pi}{\sin\left(\frac{\pi s}{2}\right)}\zeta(s). \tag{2.10}$$

On the other hand, using Legendre's duplication formula of the gamma function:

$$\Gamma(z)\Gamma\left(z + \frac{1}{2}\right) = 2^{1-2z}\pi^{\frac{1}{2}}\Gamma(2z),$$

[5] Very roughly, it means the function $\xi(s)$ can be extended "nicely" (holomorphic) to the complex plane, except at $s = 0, 1$, it "behaves like" $-1/s, -1/(1 - s)$ (meromorphic), respectively.

[6] The idea of the first assertion goes as follows. By definition, we have

$$\zeta(s) = \frac{\pi^{s/2}}{\Gamma(s/2)}\xi(s).$$

From Remark 2.9, we know that $\xi(s)$ behaves like $-1/s$ at $s = 0$. On the other hand, a fact from complex analysis is that $\Gamma(s)$ behaves like $1/s$ at $s = 0$. So the "bad behaviors" in the fraction cancel out with each other, showing that $\zeta(s)$ is actually defined at 0.

the right-hand side of (2.9) becomes

$$\pi^{-\frac{1-s}{2}}\Gamma\left(\frac{1-s}{2}\right)\Gamma\left(1-\frac{s}{2}\right)\zeta(1-s)=\pi^{-\frac{1-s}{2}}2^{1-2\left(\frac{1-s}{2}\right)}\pi^{\frac{1}{2}}\Gamma(1-s)\zeta(1-s). \tag{2.11}$$

We may now equate (2.10) and (2.11) and get

$$\pi^{-\frac{s}{2}}\frac{\pi}{\sin\left(\frac{\pi s}{2}\right)}\zeta(s)=\pi^{-\frac{1-s}{2}}2^{1-2\left(\frac{1-s}{2}\right)}\pi^{\frac{1}{2}}\Gamma(1-s)\zeta(1-s).$$

This simplifies to the desired functional equation. □

Let s be a negative even integer. In this case, the sine factor in Corollary 2.10 is zero while the others are non-zero. Thus, we have $\zeta(s)=0$ for all $s\in\{-2n\mid n\in\mathbb{N}\}$. Such s is called a *trivial zero* of $\zeta(s)$. The famous million-dollar problem, Riemann hypothesis, concerns the locations of all non-trivial zeros of $\zeta(s)$.

Conjecture 2.11 (Riemann Hypothesis)**.** All non-trivial zeros of $\zeta(s)$ lie on the line $\mathrm{Re}(s)=1/2$.

Next, we express the Riemann zeta function at negative integers in terms of Bernoulli numbers.

Corollary 2.12. *For $k\in\mathbb{N}$ with $k>1$, we have*

$$\zeta(1-k)=-\frac{B_k}{k}$$

where B_k is the k-th Bernoulli number.

Proof. By Corollary 2.10 and Theorem 2.3, when $k>1$ is even, we have

$$\begin{aligned}\zeta(1-k)&=2^{1-k}\pi^{-k}\sin\left(\frac{\pi(1-k)}{2}\right)\Gamma(k)\zeta(k)\\&=2^{1-k}\pi^{-k}(-1)^{\frac{k}{2}}(k-1)!\left(-\frac{B_k}{2\cdot k!}(2\pi i)^k\right)\\&=-\frac{B_k}{k}.\end{aligned}$$

When $k>1$ is odd, we have $\zeta(1-k)=0$ as $1-k$ is a trivial zero. Also, recall Corollary 2.6 that $B_k=0$ for all odd $k>1$. This completes the proof. □

Remark 2.13. From Corollary 2.12, we take $k=2$ and see that (recall Example 2.4)

$$\zeta(-1)=-\frac{B_2}{2}=-\frac{1}{6}\cdot\frac{1}{2}=-\frac{1}{12}.$$

Some people would interpret this identity as

$$\zeta(-1)=\sum_{n=1}^{\infty}\frac{1}{n}=1+2+3+4+\cdots=-\frac{1}{12}$$

and claim that the sum of all natural numbers is $-1/12$. This is incorrect because the series expression (2.1) for $\zeta(s)$ is only valid for $\mathrm{Re}(s)>1$.

3 Dirichlet Density Theorem

The nature of prime numbers has long been a mystery to number theorists. On the one hand, they show randomness that makes them hard to predict. For example, the distance between consecutive prime numbers (i.e., the problem of prime gaps) varies wildly and erratically. But on the other hand, mathematicians have discovered beautiful patterns among primes, such as prime number theorem. In this week's lecture, we prove a remarkable theorem concerning the distribution of prime numbers, called the *Dirichlet density theorem*.

3.1 Dirichlet Density

Definition 3.1. Let $\mathcal{P}$ be a set of prime numbers. We define the *Dirichlet (analytic) density* of $\mathcal{P}$ by

$$d(\mathcal{P}) := \lim_{s \to 1^+} \frac{\sum_{p \in \mathcal{P}} \frac{1}{p^s}}{\sum_p \frac{1}{p^s}}$$

(if the limit exists), where the denominator runs through all prime numbers.

Note that if $\mathcal{P}$ is finite, then its Dirichlet density is 0 as the denominator goes to ∞ (why?). On the other hand, the Dirichlet density of all prime numbers is 1. In general, if the density exists, it must be a number between 0 and 1. It is also easy to see that if $\mathcal{P}_1, \mathcal{P}_2$ are disjoint and both have Dirichlet density, then $d(\mathcal{P}_1 \cup \mathcal{P}_2) = d(\mathcal{P}_1) + d(\mathcal{P}_2)$.

The main objective of this week's lecture is to prove the following theorem due to Dirichlet.

Theorem 3.2 (Dirichlet)**.** *Let $a, m \in \mathbb{N}$ with $\gcd(a, m) = 1$. Define*

$$\mathcal{P}(a; m) := \{p\text{: prime} \mid p \equiv a \pmod{m}\}.$$

Then we have

$$d(\mathcal{P}(a; m)) = \frac{1}{\phi(m)},$$

where ϕ is the Euler phi-function.

Note that this theorem immediately implies that the set $\mathcal{P}(a; m)$ must be infinite, i.e., there are infinitely many primes which are congruent to a modulo m. Yet, the result asserts something more profound. Note that for each fixed $m \in \mathbb{N}$, if $p \nmid m$ is a prime number, then $p \in \mathcal{P}(a; m)$ for some $a \in \mathbb{N}$ with $\gcd(a, m) = 1$. In other words, using these sets, we can partition such prime numbers by their residue modulo m. Then since there are exactly $\phi(m)$ of them, this theorem also says that the prime numbers (with only finitely many exceptions, which have Dirichlet density zero) are equally distributed among $\mathcal{P}(a; m)$.

For example, say $m = 4$. Then Theorem 3.2 says that

$$d(\mathcal{P}(1; 4)) = d(\mathcal{P}(3; 4)) = \frac{1}{2}.$$

In other words, half of the odd prime numbers are congruent to 1 modulo 4, and the other half are congruent to 3 modulo 4.

3.2 An Equivalent Definition of Dirichlet Density

In order to prove Theorem 3.2, we first give an equivalent definition of the Dirichlet density. This will be done by investigating more analytic properties of the Riemann zeta function (recall Definition 2.1):

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}, \quad \mathrm{Re}(s) > 1.$$

Proposition 3.3. *We have*

$$\lim_{s \to 1^+} (s-1)\zeta(s) = 1.$$

Proof. Let $s > 1$, then since t^{-s} decreases as $t \to \infty$, we have

$$\frac{1}{(n+1)^s} \leq \int_n^{n+1} \frac{1}{t^s} dt \leq \frac{1}{n^s}.$$

Summing from $n = 1$ to ∞ and calculating the improper integral in the middle, we obtain

$$\zeta(s) - 1 \leq \frac{1}{s-1} \leq \zeta(s).$$

Thus, we have

$$1 \leq (s-1)\zeta(s) \leq s - 1 \leq s.$$

Letting $s \to 1^+$, the result now follows from the squeeze theorem. □

Corollary 3.4. *We have*

$$\lim_{s \to 1^+} \frac{\log \zeta(s)}{\log(s-1)^{-1}} = 1.$$

Proof. From

$$\log \zeta(s) = \log \zeta(s)(s-1) - \log(s-1),$$

we have

$$\frac{\log \zeta(s)}{\log(s-1)^{-1}} = -\frac{\log \zeta(s)(s-1)}{\log(s-1)} + 1.$$

Letting $s \to 1^+$, the result now follows from Proposition 3.3. □

Proposition 3.5. *We have*

$$\log \zeta(s) = \sum_p \frac{1}{p^s} + R(s),$$

where $R(s)$ remains bounded as $s \to 1^+$.

Proof. Recall that we have the Taylor expansion

$$-\log(1-x) = \sum_{m=1}^{\infty} \frac{x^m}{m}, \quad -1 \le x < 1.$$

Then by Proposition 2.2, for $s > 1$, we have

$$\log \zeta(s) = \sum_p -\log(1-p^{-s}) = \sum_p \sum_{m=1}^{\infty} \frac{p^{-sm}}{m} = \sum_p \frac{1}{p^s} + \sum_p \sum_{m=2}^{\infty} \frac{1}{mp^{sm}}.$$

A simple estimate of the second term shows that

$$\sum_p \sum_{m=2}^{\infty} \frac{1}{mp^{sm}} \le \sum_p \sum_{m=2}^{\infty} \frac{1}{p^{sm}} = \sum_p \frac{p^{-2s}}{1-p^{-s}} \le \frac{1}{1-2^{-s}} \sum_p \frac{1}{p^{2s}} \le 2\zeta(2s) < \infty$$

as $s \to 1^+$. This completes the proof. □

Corollary 3.6. *We have*

$$\lim_{s \to 1^+} \frac{\sum_p \frac{1}{p^s}}{\log(s-1)^{-1}} = 1.$$

Proof. From Proposition 3.5, we have

$$\frac{\log \zeta(s)}{\log(s-1)^{-1}} = \frac{\sum_p \frac{1}{p^s}}{\log(s-1)^{-1}} + \frac{R(s)}{\log(s-1)^{-1}}.$$

Letting $s \to 1^+$, the result follows from Corollary 3.4 and Proposition 3.5. □

Recall Definition 3.1 that the Dirichlet density for a set $\mathcal{P}$ of prime numbers is defined through the limit

$$d(\mathcal{P}) = \lim_{s \to 1^+} \frac{\sum_{p \in \mathcal{P}} \frac{1}{p^s}}{\sum_p \frac{1}{p^s}}.$$

Now, by Corollary 3.6, we see that it can also be defined as

$$d(\mathcal{P}) = \lim_{s \to 1^+} \frac{\sum_{p \in \mathcal{P}} \frac{1}{p^s}}{\log(s-1)^{-1}}. \tag{3.1}$$

Remark 3.7. Note that Corollary 3.6 also implies the infinitude of prime numbers (because otherwise, the limit would be 0).

3.3 Dirichlet Characters

In this section, we fix $m \in \mathbb{N}$. We let $G := (\mathbb{Z}/m\mathbb{Z})^\times$ be the set consisting of all integers modulo m which are relatively prime to m. Our next ingredient of proving Theorem 3.2 is the notion of *Dirichlet characters*.

Definition 3.8. A *character* χ of G is a function $\chi : G \to \mathbb{C}^\times := \mathbb{C} \setminus \{0\}$ such that

$$\chi(ab) = \chi(a)\chi(b)$$

for all $a, b \in G$. We let $\widehat{G}$ be the set of all characters of G.

An easy example of a character is the *trivial character* $\chi_0 \in \widehat{G}$, which is defined by $\chi_0(a) := 1$ for every $a \in G$. In general, observe that for any $\chi \in \widehat{G}$, we have $\chi(1) = \chi(1^2) = \chi(1)^2$. And since by definition, $\chi(1) \neq 0$, it follows that $\chi(1) = 1$.

Example 3.9. Let's take $m = 4$. Then $G = (\mathbb{Z}/4\mathbb{Z})^\times = \{1, 3\}$. Since $3^2 \equiv 1 \pmod 4$, we have $\chi(3)^2 = \chi(3^2) = \chi(1) = 1$. This shows that $\chi(3) = \pm 1$. Thus, there are two characters for G. We can write down the *character table* of G as follows.

$(\mathbb{Z}/4\mathbb{Z})^\times$	1	3
χ_0	1	1
χ_1	1	-1

Example 3.10. Next, let's take $m = 10$. Then $G = (\mathbb{Z}/10\mathbb{Z})^\times = \{1, 3, 7, 9\}$. Note that in G, we have

$$\begin{aligned} 3^1 &\equiv 3 \pmod{10}, \\ 3^2 &\equiv 9 \pmod{10}, \\ 3^3 &\equiv 7 \pmod{10}, \\ 3^4 &\equiv 1 \pmod{10}. \end{aligned}$$

In other words, the element $3 \in G$ "generates" all the other elements. Thus, to find all characters of G, it suffices to determine their values at 3. Further, note that $\chi(3)^4 = \chi(3^4) = \chi(1) = 1$. Hence, we have $\chi(3) = \pm 1, \pm i$. So there are four characters of G.

Exercise 3.11. Write down the character table of $G = (\mathbb{Z}/10\mathbb{Z})^\times$.

Definition 3.12. We extend the domain of $\chi \in \widehat{G}$ to $\mathbb{Z}$ as follows:

- $\chi(a) := 0$ if $\gcd(a, m) > 1$.
- $\chi(a) := \chi(a \bmod m)$ for all $a \in \mathbb{Z}$.

The resulting character is called a *Dirichlet character modulo* m.

Exercise 3.13. Prove that every Dirichlet character χ is completely multiplicative. That is, it satisfies $\chi(ab) = \chi(a)\chi(b)$ for all $a, b \in \mathbb{Z}$.

3.4 Proof of the Dirichlet Density Theorem: a Special Case

We are now ready to prove Theorem 3.2. In fact, we will only prove the case where $m = 4$, as it would be enough to grasp the main idea of the proof. We will also explain how the general situation follows from exactly the same process (with more technical issues, of course).

Proof. ($m = 4$) We recall from Example 3.9 and Definition 3.12 that there are two Dirichlet characters modulo 4:

$$\chi_0(n) = \begin{cases} 1, & n\text{: odd}, \\ 0, & n\text{: even}, \end{cases} \quad \text{and} \quad \chi_1(n) = \begin{cases} 1, & n \equiv 1 \pmod 4, \\ -1, & n \equiv 3 \pmod 4, \\ 0, & n\text{: even}. \end{cases} \tag{3.2}$$

Next, for $s > 1$, we define the *Dirichlet L-functions* with respect to χ_0, χ_1 as

$$\begin{cases} L(s, \chi_0) := \displaystyle\sum_{n=1}^{\infty} \frac{\chi_0(n)}{n^s} = 1 + \frac{1}{3^s} + \frac{1}{5^s} + \frac{1}{7^s} + \cdots, \\ L(s, \chi_1) := \displaystyle\sum_{n=1}^{\infty} \frac{\chi_1(n)}{n^s} = 1 - \frac{1}{3^s} + \frac{1}{5^s} - \frac{1}{7^s} + \cdots. \end{cases}$$

By Exercise 3.13, the characters χ_0, χ_1 are completely multiplicative. Thus, a similar proof as in Proposition 2.2 shows that

$$\begin{cases} L(s, \chi_0) = \displaystyle\prod_{p\text{: odd}} (1 - p^{-s})^{-1} = (1 - 2^{-s})\zeta(s), \\ L(s, \chi_1) = \displaystyle\prod_{p\text{: odd}} (1 - \chi_1(p)p^{-s})^{-1}. \end{cases}$$

With these Euler products in hand, a similar proof as in Proposition 3.5 shows that

$$\begin{cases} \log L(s, \chi_0) = \displaystyle\sum_{p\text{: odd}} \frac{1}{p^s} + R_1(s), \\ \log L(s, \chi_1) = \displaystyle\sum_{p\text{: odd}} \frac{\chi_1(p)}{p^s} + R_2(s), \end{cases}$$

where $R_1(s), R_2(s)$ remain bounded as $s \to 1^+$. By (3.2), this implies

$$\begin{cases} \log L(s, \chi_0) + \log L(s, \chi_1) = 2 \displaystyle\sum_{p \equiv 1 \pmod 4} \frac{1}{p^s} + R_3(s), \\ \log L(s, \chi_0) - \log L(s, \chi_1) = 2 \displaystyle\sum_{p \equiv 3 \pmod 4} \frac{1}{p^s} + R_4(s), \end{cases} \tag{3.3}$$

where $R_3(s), R_4(s)$ remain bounded as $s \to 1^+$. Dividing both sides by $\log(s-1)^{-1}$, we obtain

$$\frac{\log L(s,\chi_0)}{\log(s-1)^{-1}} + \frac{\log L(s,\chi_1)}{\log(s-1)^{-1}} = \frac{2\sum_{p\equiv 1 \pmod 4} \frac{1}{p^s}}{\log(s-1)^{-1}} + \frac{R_3(s)}{\log(s-1)^{-1}} \tag{3.4}$$

and

$$\frac{\log L(s,\chi_0)}{\log(s-1)^{-1}} - \frac{\log L(s,\chi_1)}{\log(s-1)^{-1}} = \frac{2\sum_{p\equiv 3 \pmod 4} \frac{1}{p^s}}{\log(s-1)^{-1}} + \frac{R_4(s)}{\log(s-1)^{-1}}. \tag{3.5}$$

Now, by (3.1), the right-hand side of (3.4), (3.5) goes to $2d(\mathcal{P}(1;4)), 2d(\mathcal{P}(3;4))$, respectively, as $s \to 1^+$. For the left-hand side, note that by Corollary 3.4, we have

$$\frac{\log L(s,\chi_0)}{\log(s-1)^{-1}} = \frac{\log(1-2^{-s})}{\log(s-1)^{-1}} + \frac{\log \zeta(s)}{\log(s-1)^{-1}} \longrightarrow 1$$

as $s \to 1^+$. On the other hand, note that for $s > 1$, we have

$$L(s,\chi_1) = 1 - \frac{1}{3^s} + \frac{1}{5^s} - \frac{1}{7^s} + \cdots = \left(1 - \frac{1}{3^s}\right) + \left(\frac{1}{5^s} - \frac{1}{7^s}\right) + \cdots \geq \frac{2}{3}$$

and

$$L(s,\chi_1) = 1 - \frac{1}{3^s} + \frac{1}{5^s} - \frac{1}{7^s} + \cdots = 1 - \left(\frac{1}{3^s} - \frac{1}{5^s}\right) - \left(\frac{1}{7^s} - \frac{1}{9^s}\right) + \cdots \leq 1.$$

So by the squeeze theorem, we have

$$\frac{\log L(s,\chi_1)}{\log(s-1)^{-1}} \longrightarrow 0$$

as $s \to 1^+$. Combining every information we got, this implies that

$$d(\mathcal{P}(1;4)) = d(\mathcal{P}(3;4)) = \frac{1}{2},$$

which is what we want. □

3.5 General Case

We now finish this week's lecture by reexamining all the steps in the above proof, and explain which parts can be generalized easily, and which need more consideration. We let $a, m \in \mathbb{N}$ with $\gcd(a,m) = 1$.

In the first step, we listed all Dirichlet characters modulo 4. For general m, they can be classified by "group theory". It is worth mentioning that the number of characters of $(\mathbb{Z}/m\mathbb{Z})^\times$ is exactly $\phi(m)$, which is the same as the cardinality of $(\mathbb{Z}/m\mathbb{Z})^\times$ (recall Examples 3.9 and 3.10).

Next, for each Dirichlet character χ modulo m, we can define the Dirichlet L-function $L(s,\chi)$ in the same way. That is, for $s > 1$, we define

$$L(s,\chi) := \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s}.$$

Since by Exercise 3.13, each χ is completely multiplicative, we can easily obtain the Euler product

$$L(s,\chi) = \prod_p (1-\chi(p)p^{-s})^{-1}$$

as in Proposition 2.2. Using this, we can also establish an estimate of $\log L(s,\chi)$ similar to Proposition 3.5. That is, we have[7]

$$\log L(s,\chi) \text{“}=\text{”} \sum_{p\nmid m} \frac{\chi(p)}{p^s} + R_\chi(s)$$

where $R_\chi(s)$ remains bounded as $s \to 1^+$.

The next key step in our proof was taking some linear combination of $\log L(s,\chi)$ on the left and “generate” the equality

$$\sum_{\chi\in\widehat{G}} c_\chi \log L(s,\chi) = \phi(m) \sum_{p\equiv a \pmod m} \frac{1}{p^s} + R_a(s) \tag{3.6}$$

where $R_a(s)$ remains bounded as $s \to 1^+$. The existence of such linear combination in general is in fact, guaranteed by linear algebra. More specifically, it can be shown that all characters of $(\mathbb{Z}/m\mathbb{Z})^\times$ form an “orthogonal basis” of the $\mathbb{C}$-vector space of all $\mathbb{C}$-valued functions on $(\mathbb{Z}/m\mathbb{Z})^\times$.[8] We will discuss this in Lecture 4.

For the last step, we can once again divide both sides of (3.6) by $\log(s-1)^{-1}$. Letting $s \to 1^+$, the right-hand side will similarly converge to $\phi(m)d(\mathcal{P}(a;m))$ by (3.1). For the left-hand side, it is still not hard to show by Corollary 3.4 that

$$\lim_{s\to 1^+} \frac{\log L(s,\chi_0)}{\log(s-1)^{-1}} = 1,$$

where recall χ_0 is the trivial character modulo m. The trickest part of the whole proof lies in proving the identity

$$\lim_{s\to 1^+} \frac{\log L(s,\chi)}{\log(s-1)^{-1}} = 0$$

for non-trivial χ. Equivalently, we are required to show that $L(1,\chi)$ is non-vanishing for all such χ. However, this is a deep result in number theory whose proof lies far beyond what we can cover in this lecture.

4 Orthogonal Relations of Characters

In the last week's lecture, we proved a special case of the Dirichlet density theorem (Theorem 3.2), and explained how the strategy can be generalized to arbitrary situation. It was also

[7] One technical issue arises here. As characters are complex-valued in general, so are the Dirichlet L-functions $L(s,\chi)$. Thus, one needs to be careful about taking logarithm, as it is in fact, a multivalued function in complex analysis.

[8] This is related to the “Fourier analysis” on $(\mathbb{Z}/m\mathbb{Z})^\times$.

mentioned that a key step involves the orthogonal property of characters of $(\mathbb{Z}/m\mathbb{Z})^\times$. In this week's lecture, we aim to explain this in detail.

4.1 The Character Group of $(\mathbb{Z}/m\mathbb{Z})^\times$

Throughout this lecture, we fix $m \in \mathbb{N}$. We let $G := (\mathbb{Z}/m\mathbb{Z})^\times$ be the set consisting of all integers modulo m which are relatively prime to m. Recall Definition 3.8 that a character χ of G is a function $\chi : G \to \mathbb{C}^\times = \mathbb{C} \setminus \{0\}$ satisfying

$$\chi(ab) = \chi(a)\chi(b)$$

for all $a, b \in G$. We also let $\widehat{G}$ be the set of all characters of G.

We now endow a natural binary operation on $\widehat{G}$. Given $\chi, \lambda \in \widehat{G}$, we define $\chi\lambda : G \to \mathbb{C}^\times$ by $\chi\lambda(a) := \chi(a)\lambda(a)$ for all $a \in G$. Then we have the following simple observations.

- (Closure) As χ, λ are characters, we have

  $$\chi\lambda(ab) = \chi(ab)\lambda(ab) = \chi(a)\chi(b)\lambda(a)\lambda(b) = \chi(a)\lambda(a)\chi(b)\lambda(b) = \chi\lambda(a)\chi\lambda(b)$$

  for all $a, b \in G$. In other words, $\chi\lambda$ is also a character.

- (Commutativity) We have

  $$\chi\lambda(a) = \chi(a)\lambda(a) = \lambda(a)\chi(a) = \lambda\chi(a)$$

  for all $a \in G$. That is, $\chi\lambda = \lambda\chi$.

- (Identity) Recall §3.3 that there is a trivial character $\chi_0 \in \widehat{G}$ given by $\chi_0(a) := 1$ for all $a \in G$. It satisfies

  $$\chi\chi_0(a) = \chi(a)\chi_0(a) = \chi(a)$$

  for all $\chi \in \widehat{G}$ and $a \in G$. That is, we have $\chi\chi_0 = \chi$.

- (Inverse) Given $\chi \in \widehat{G}$, we define $\chi^{-1} : G \to \mathbb{C}^\times$ by $\chi^{-1}(a) := \chi(a)^{-1}$ for all $a \in G$. Then χ^{-1} is also a character of G (see Exercise 4.1). Moreover, it satisfies

  $$\chi\chi^{-1}(a) = \chi(a)\chi^{-1}(a) = \chi(a)\frac{1}{\chi(a)} = 1 = \chi_0(a)$$

  for all $a \in G$. In other words, we have $\chi\chi^{-1} = \chi_0$.

- (Associativity) For any $\chi, \lambda, \psi \in \widehat{G}$, we have (see Exercise 4.1)

  $$(\chi\lambda)\psi = \chi(\lambda\psi).$$

Exercise 4.1. Fill out the missing details in the above discussion.

To summarize, this shows that the set $\widehat{G}$ together with the given binary operation forms an "abelian group". We call it the *character group* of G.

We have already observed in §3.3 that $\chi(1) = 1$ for every $\chi \in \widehat{G}$. We now make more observations as follows. Let ϕ be the Euler phi-function. By the Euler's theorem, for every $a \in G$, we have

$$a^{\phi(m)} \equiv 1 \pmod{m}.$$

Thus, for every $\chi \in \widehat{G}$,

$$\chi(a)^{\phi(m)} = \chi(a^{\phi(m)}) = \chi(1) = 1.$$

This shows that $\chi(a)$ is a $\phi(m)$-th root of unity. In particular, it is an algebraic number (recall Example 1.2). Moreover, if we let $a^{-1} \in G$ be the element such that $aa^{-1} \equiv 1 \pmod{m}$. Then we have

$$\chi(a)\chi(a^{-1}) = \chi(aa^{-1}) = \chi(1) = 1.$$

Thus,

$$\chi(a^{-1}) = \chi(a)^{-1} = \chi^{-1}(a).$$

This identity suggests that the inverse of $a \in G$, the inverse of $\chi(a) \in \mathbb{C}^\times$, and the inverse of $\chi \in \widehat{G}$, are in fact compatible with each other. Also, since $\chi(a)$ is a root of unity, it lies on the unit circle of the complex plane. Thus, it satisfies $\chi(a)^{-1} = \overline{\chi(a)}$, where $\overline{\cdot}$ denotes the complex conjugation.

4.2 The Orthogonal Relations

We mentioned in §3.5 without proof that the character group $\widehat{G}$ consists of $\phi(m)$ elements. Thus, we have

$$\#(G) = \#(\widehat{G}) = \phi(m).$$

In this section, we establish two orthogonal relations of characters and explain their linear algebra point of view.

Proposition 4.2 (The Orthogonal Relation of Characters)**.** *For all $\chi, \lambda \in \widehat{G}$, we have*

$$\sum_{a \in G} \chi(a)\overline{\lambda(a)} = \begin{cases} \phi(m), & \chi = \lambda, \\ 0, & \chi \neq \lambda. \end{cases}$$

Proof. Note that

$$\sum_{a \in G} \chi(a)\overline{\lambda(a)} = \sum_{a \in G} \chi(a)\lambda^{-1}(a) = \sum_{a \in G} \chi\lambda^{-1}(a).$$

Thus, it suffices to show that

$$\sum_{a \in G} \chi(a) = \begin{cases} \phi(m), & \chi = \chi_0, \\ 0, & \chi \neq \chi_0. \end{cases} \tag{4.1}$$

When $\chi = \chi_0$ is the trivial character, it is easily seen that

$$\sum_{a \in G} \chi(a) = \sum_{a \in G} 1 = \#(G) = \phi(m).$$

When $\chi \neq \chi_0$, there must exist $t \in G$ such that $\chi(t) \neq 1$. Observe that

$$\chi(t) \sum_{a \in G} \chi(a) = \sum_{a \in G} \chi(t)\chi(a) = \sum_{a \in G} \chi(ta) = \sum_{a \in G} \chi(a). \tag{4.2}$$

The last equality can be justified as follows. Suppose we have $a, b \in G$ with $at = bt$. This means $at \equiv bt \pmod{m}$. Then since $\gcd(t, m) = 1$, this implies that $a \equiv b \pmod{m}$, or equivalently, $a = b$ in G. Thus, the multiplication-by-t map on G is an injection, and hence, a bijection.

Now, from (4.2), we obtain

$$(\chi(t) - 1) \sum_{a \in G} \chi(a) = 0.$$

Since $\chi(t) \neq 1$ by the assumption, it follows that the sum vanishes. □

Example 4.3. Recall Example 3.9 that the character table of $(\mathbb{Z}/4\mathbb{Z})^\times$ is

$(\mathbb{Z}/4\mathbb{Z})^\times$	1	3
χ_0	1	1
χ_1	1	-1

Then Proposition 4.2 (or rather, (4.1)) indicates the fact that the first row for χ_0 sum to $\phi(4) = 2$, and the second row for χ_1 sum to 0.

Example 4.4. In Exercise 3.11, you were asked to find the character table of $(\mathbb{Z}/10\mathbb{Z})^\times$. The answer is as follows.

$(\mathbb{Z}/10\mathbb{Z})^\times$	1	3	7	9
χ_0	1	1	1	1
χ_1	1	-1	-1	1
χ_2	1	i	$-i$	-1
χ_3	1	$-i$	i	-1

Then we see that the first row for χ_0 sum to $\phi(10) = 4$, and the others sum to 0.

We now interpret Proposition 4.2 through linear algebra. Consider the $\mathbb{C}$-vector space $\mathcal{F}(G, \mathbb{C})$ of all $\mathbb{C}$-valued functions on G.

Definition 4.5. We define the inner product on $\mathcal{F}(G, \mathbb{C})$ by

$$(f_1 \mid f_2) := \frac{1}{\phi(m)} \sum_{a \in G} f_1(a)\overline{f_2(a)}$$

for all $f_1, f_2 \in \mathcal{F}(G, \mathbb{C})$.

Exercise 4.6. Check that $(\cdot \mid \cdot)$ is indeed an inner product on $\mathcal{F}(G, \mathbb{C})$.

With this definition, we can rewrite Proposition 4.2 as

$$(\chi \mid \lambda) = \frac{1}{\phi(m)} \sum_{a \in G} \chi(a)\overline{\lambda(a)} = \begin{cases} 1, & \chi = \lambda, \\ 0, & \chi \neq \lambda, \end{cases}$$

for all $\chi, \lambda \in \widehat{G}$. That is, the group $\widehat{G}$ of all characters of G is orthogonal (in fact, orthonormal, as we normalized the inner product by $\#(G)^{-1} = \phi(m)^{-1}$) in $\mathcal{F}(G, \mathbb{C})$. In particular, they are linearly independent, and so

$$\phi(m) \leq \dim_{\mathbb{C}} \mathcal{F}(G, \mathbb{C}).$$

On the other hand, suppose $G = \{a_1, \ldots, a_{\phi(m)}\}$. For each $1 \leq i \leq \phi(m)$, we define $f_i \in \mathcal{F}(G, \mathbb{C})$ by $f_i(a) := 1$ if $a = a_i$ and 0 otherwise. That is, f_i is the characteristic function of $\{a_i\} \subseteq G$. Then it is easy to see that $\{f_1, \ldots, f_{\phi(m)}\}$ generates the $\mathbb{C}$-vector space $\mathcal{F}(G, \mathbb{C})$. In particular, we have

$$\phi(m) \geq \dim_{\mathbb{C}} \mathcal{F}(G, \mathbb{C}).$$

To summarize, we have shown that the dimension of $\mathcal{F}(G, \mathbb{C})$ is precisely $\phi(m)$ and the character group $\widehat{G}$ forms an orthogonal (orthonormal) basis of $\mathcal{F}(G, \mathbb{C})$.

Proposition 4.7 (Vertical Orthogonal Relation)**.** *For all $a, b \in G$, we have*

$$\sum_{\chi \in \widehat{G}} \chi(a)\overline{\chi(b)} = \begin{cases} \phi(m), & a = b, \\ 0, & a \neq b. \end{cases}$$

Proof. As $\widehat{G}$ is a group, one can give a proof similar to Proposition 4.2. Here, we use linear algebra.

Fix $b \in G$, we define $f_b \in \mathcal{F}(G, \mathbb{C})$ to be the characteristic function of the singleton $\{b\}$ normalized by $\phi(m)$. That is,

$$f_b(a) := \begin{cases} \phi(m), & a = b, \\ 0, & a \neq b. \end{cases}$$

Since the characters form an orthonormal basis of $\mathcal{F}(G, \mathbb{C})$, we have

$$f_b = \sum_{\chi \in \widehat{G}} (f_b \mid \chi)\chi.$$

For the inner product $(f_b \mid \chi)$, by Definition 4.5, one sees that

$$(f_b \mid \chi) = \frac{1}{\phi(m)} \sum_{a \in G} f_b(a)\overline{\chi(a)} = \frac{1}{\phi(m)} \cdot f_b(b)\overline{\chi(b)} = \overline{\chi(b)}.$$

This shows that

$$f_b(a) = \sum_{\chi \in \widehat{G}} \overline{\chi(b)}\chi(a) = \begin{cases} \phi(m), & a = b, \\ 0, & a \neq b, \end{cases}$$

which is the desired identity. □

Remark 4.8. Note that Proposition 4.7 is equivalent to

$$\sum_{\chi \in \widehat{G}} \chi(a) = \begin{cases} \phi(m), & a = 1, \\ 0, & a \neq 1. \end{cases}$$

This formulation indicates the fact that the first column (for $a = 1$) in Example 4.3 sum to $\phi(4) = 2$, and the second column (for $a = 3$) sum to 0. The same observation applies to Example 4.4.

To end this lecture, let us revisit the proof of the Dirichlet density theorem for the general case (recall §3.5). Suppose during the proof, we have found all Dirichlet characters modulo m and established the estimate

$$\log L(s, \chi) \text{“=”} \sum_{p \nmid m} \frac{\chi(p)}{p^s} + R_\chi(s).$$

We now want to take $c_\chi \in \mathbb{C}$ for each $\chi \in \widehat{G}$ so that (3.6) holds.

With Proposition 4.7, one sees that if we take $c_\chi := \overline{\chi(a)}$ and sum over $\chi \in \widehat{G}$, we will obtain

$$\begin{aligned} \sum_{\chi \in \widehat{G}} \overline{\chi(a)} \log L(s, \chi) &= \sum_{\chi \in \widehat{G}} \sum_{p \nmid m} \frac{\chi(p)\overline{\chi(a)}}{p^s} + \sum_{\chi \in \widehat{G}} \overline{\chi(a)} R_\chi(s) \\ &= \sum_{p \nmid m} \frac{1}{p^s} \sum_{\chi \in \widehat{G}} \chi(p)\overline{\chi(a)} + R_a(s) \\ &= \phi(m) \sum_{p \equiv a \pmod{m}} \frac{1}{p^s} + R_a(s). \end{aligned}$$

And this is exactly what we want in (3.6).

Indeed, consider the special case $m = 4$ we proved in §3.4, let us take $a = 3 \in (\mathbb{Z}/4\mathbb{Z})^\times$. Then from the character table of $(\mathbb{Z}/4\mathbb{Z})^\times$ (recall Example 3.9), we should take

$$c_{\chi_0} = \overline{\chi_0(3)} = 1 \quad \text{and} \quad c_{\chi_1} = \overline{\chi_1(3)} = -1.$$

This agrees the second equality in (3.3), where we took

$$\sum_{\chi \in \widehat{G}} c_\chi \log L(s, \chi) = \log L(s, \chi_0) - \log L(s, \chi_1).$$

References

[Chaa] Timo Chang. *Resultant and its Applications*. URL: https://timomath.com/resultant_and_its_applications.pdf.

[Chab] Timo Chang. *The Transcendence of e*. URL: https://timomath.com/the_transcendence_of_e.pdf.