PMATH 440 Course Notes

Analytic Number Theory

Wentang Kuo • Fall 2021 • University of Waterloo

Table of Contents

1 Introduction to Prime Numbers and Their Counting Function 3
1.1 Primes . . . 3
1.2 Elementary Approximations of $\pi(x)$ . . . 5
1.3 Bertrand's Postulate . . . 8
1.4 Gaps Between Twin Primes . . . 9

2 Asymptotic Analysis for $\pi(x)$ 11
2.1 The Möbius Function . . . 11
2.2 The von Mangoldt Function . . . 13
2.3 Abel's Summation Formula . . . 15

3 Riemann's Zeta Function and the Prime Number Theorem 19
3.1 The Riemann Zeta Function . . . 19
3.2 Newman's Theorem . . . 21
3.3 Revisiting the Möbius Function . . . 23
3.4 Divisor Function . . . 24
3.5 The Prime Number Theorem . . . 26

4 Divisor Counting Functions 30
4.1 Asymptotic Formulas for Divisor Counting Functions . . . 30
4.2 Summatory Functions for $\omega(n)$ and $\Omega(n)$ . . . 33
4.3 Asymptotic Density and Normal Order . . . 34
4.4 Normal Order of $\omega(n)$ and $\Omega(n)$ . . . 36

5 Quadratic Reciprocity 41
5.1 Euler's Totient Function . . . 41
5.2 Quadratic Residues . . . 42
5.3 Special Cases of Quadratic Reciprocity . . . 44
5.4 The Law of Quadratic Reciprocity . . . 47
5.5 The Jacobi Symbol . . . 48

6 Primitive Roots 51
6.1 Cyclicity of $(\mathbb{Z}/p\mathbb{Z})^*$ . . . 51
6.2 Primitive Roots: The Prime Power Case . . . 53
6.3 Primitive Roots: The General Case . . . 55

7 L-functions and Dirichlet's Theorem 58
7.1 Some Results on Primes in Arithmetic Progressions . . . 58
7.2 Characters . . . 59
7.3 Dirichlet Characters . . . 60
7.4 Dirichlet L-functions . . . 61
7.5 Dirichlet Series . . . 63
7.6 Dirichlet's Theorem . . . 65
7.7 Distribution of Primes in Arithmetic Progressions . . . 67

8 Waring's Problem 69
8.1 History of Waring's Problem 69
8.2 Special Cases of Waring's Problem 69

1 Introduction to Prime Numbers and Their Counting Function

1.1 Primes

Definition 1.1

A **prime number** is a positive integer greater than 1 such that its only factors are 1 and itself. We denote by $\mathcal{P}$ the set of all prime numbers. For a positive real number x, we define the **prime counting function** by

$$\pi(x) = \#\{p \leq x : p \in \mathcal{P}\},$$

where $\#S$ denotes the cardinality of the set S.

We would like to know how the primes are distributed among the integers. Let p_n denote the n-th prime. Is there a formula to obtain p_n? Is there a polynomial $f(x) \in \mathbb{Z}[x]$ such that $f(n) = p_n$ for all $n \in \mathbb{N}$? The answer to the latter question is no, due to the following result.

Proposition 1.2

There is no non-constant polynomial $f(x) \in \mathbb{Z}[x]$ such that $f(n)$ is prime for all $n \in \mathbb{N}$.

Proof. Suppose such a polynomial $f(x) \in \mathbb{Z}[x]$ existed, and write

$$f(x) = a_n x^n + \cdots + a_1 x + a_0.$$

Let q be a prime with $f(n) = q$ for some $n \in \mathbb{N}$. Then $q \mid f(n + kq)$ for each $k \in \mathbb{N}$. In particular, notice that if $f(m)$ is prime for every positive integer m, then $f(x)$ must be constant with $f(x) = q$ for some prime q. $\square$

Remark 1.3

(1) There are examples of polynomials whose initial values are surprisingly often prime. For example, the polynomial $n^2 + n + 41$ is prime for all $0 \leq n \leq 39$, and the polynomial $(n-40)^2 + (n-40) + 41$ is prime for all $0 \leq n \leq 79$.

(2) In the 1970s, Matijasevic proved Hilbert's tenth problem, and in the process, he was able to show that there is a polynomial $f \in \mathbb{Z}[a, b, \ldots, z]$ such that the set of positive values in $f(\mathbb{N}^{26})$ is exactly the set of primes. In 1977, he showed that only 10 variables are needed.

Let us instead ask a weaker question. Can we find a non-constant polynomial $f(x) \in \mathbb{Z}[x]$ such that $f(n)$ yields a prime for infinitely many $n \in \mathbb{N}$? Trivially, we see that $f(x) = x + k$ works for any $k \in \mathbb{Z}$. When the coefficient of x is not equal to 1, we have the following result, which we will prove at the end of this course.

Theorem 1.4: Dirichlet

Let k and ℓ be coprime positive integers. Then $kn + \ell$ is prime for infinitely many positive integers n.

Remark 1.5

(1) At the moment, there is no known polynomial of degree greater than 1 in one variable known to take prime values infinitely often. The best result known to date is that n^2+1 is a product of two primes for infinitely many n.

(2) If we instead consider polynomials of two variables, we can go further. It is known that an odd prime p is the sum of two squares if and only if $p \equiv 1 \pmod 4$. In 1998, Friedlander and Iwaniec proved that there are infinitely many primes of the form n^2+m^4. In 2001, Heath-Brown showed that there are infinitely many primes of the form n^3+2m^3.

Theorem 1.6: Euclid

There are infinitely many prime numbers.

Proof. Assume that there are only finitely many primes, say $p_1, \ldots, p_n$, and consider

$$m = p_1 \cdots p_n + 1.$$

Then m can be written as a product of primes by unique factorization, and $p_k \mid m$ for some $1 \leq k \leq n$. Hence, we see that $p_k \mid m - p_1 \cdots p_n$ and $p_k \mid 1$, which is a contradiction. $\square$

We would like to estimate the prime counting function $\pi(x)$.

Proposition 1.7

For all $n \in \mathbb{N}$, we have $p_n \leq 2^{2^n}$.

Proof. We proceed by induction. For $n = 1$, we have $2 = p_1 \leq 2^{2^1} = 4$. Suppose the result holds for all $1 \leq k \leq n$. By Euclid's argument, we obtain $p_{n+1} \leq p_1 \cdots p_n + 1$. It follows from induction that

$$p_{n+1} \leq 2^{2^1} 2^{2^2} \cdots 2^{2^n} + 1 \leq 2^{2^{n+1}-2} + 1 \leq 2^{2^{n+1}},$$

which completes the proof. $\square$

Corollary 1.8

For all $x \geq 2$, we have $\pi(x) > \log\log x$. (In this course, log denotes the natural logarithm.)

Proof. Let $x \geq 2$, and let s be the integer satisfying

$$2^{2^s} \leq x < 2^{2^{s+1}}.$$

By Proposition 1.7, we have $\pi(x) \geq s$. On the other hand, since $x < 2^{2^{s+1}}$, taking logarithms yields $\log_2(\log_2 x) < s+1$, and hence

$$\frac{\log(\frac{\log x}{\log 2})}{\log 2} < s+1.$$

It follows that

$$\pi(x) \geq s > \frac{\log(\frac{\log x}{\log 2})}{\log 2} - 1 \geq \log\log x. \qquad \square$$

There is an alternative way to prove Euclid's theorem, due to Euler, which is left as part of the homework. Using the same idea, we can derive a slightly better lower bound for $\pi(x)$.

Proposition 1.9

For all $x \geq 2$, we have
$$\pi(x) \geq \frac{\log\log x}{\log 2}.$$

Proof. Suppose that $x \geq 2$. Then we have
$$2^{\pi(x)} \geq \prod_{p \leq x}\left(1-\frac{1}{p}\right)^{-1} = \prod_{p\leq x}\left(1+\frac{1}{p}+\frac{1}{p^2}+\cdots\right) \geq \sum_{n\leq x}\frac{1}{n} \geq \int_1^{\lfloor x\rfloor+1}\frac{1}{u}\,\mathrm{d}u \geq \log x,$$
where the product $\prod_{p\leq x}$ means that p runs through all primes at most x, and $\lfloor y\rfloor$ is the greatest integer less than or equal to y. We will will use this notation for the rest of the course. Taking logarithms yields the desired inequality. □

Fermat had conjectured that the numbers of the form $2^{2^n}+1$ are prime for $n \in \mathbb{N}$. He had checked it for the values $0 \leq n \leq 4$. These are known as the **Fermat numbers** and are denoted by
$$F_n = 2^{2^n}+1.$$
In 1732, Euler showed that $641 \mid F_5$. It is also known that $F_6, \ldots, F_{21}$ are composite. It is quite likely that only finitely many Fermat numbers are prime.

Theorem 1.10: Polyá

If n and m are positive integers with $1 \leq n < m$, then $(F_n, F_m) = 1$.

Proof. Write $m = n+k$ with $k \geq 1$. First, we will show that $F_n \mid F_m - 2$. Observe that
$$F_m - 2 = (2^{2^{n+k}}+1)-2 = 2^{2^{n+k}}-1.$$
The polynomial $x^{2^k}-1$ is divisible by $x+1$ in $\mathbb{Z}[x]$. Now, letting $x = 2^{2^n}$, we get
$$\frac{F_m-2}{F_n} = \frac{x^{2^k}-1}{x+1} = x^{2^k-1} - x^{2^k-2} + \cdots - 1 \in \mathbb{Z}.$$
Hence, we have $F_n \mid F_m - 2$. Suppose now that $d \mid F_n$ and $d \mid F_m$. Then $d \mid 2$ and $2 \nmid F_n$, which implies that $d = \pm 1$. The result follows. □

This gives yet another proof of Euclid's theorem, as well as the bound $p_n \leq 2^{2^n}+1$.

1.2 Elementary Approximations of $\pi(x)$

In 1896, Hadamard and de la Vallée Poussin each proved the Prime Number Theorem independently.

Theorem 1.11: Prime Number Theorem

We have
$$\lim_{x\to\infty}\frac{\pi(x)}{x/\log x} = 1.$$

This was initially conjectured by Gauss. We will prove this theorem later in the course; for now, we will see how to approach this problem using elementary methods.

Theorem 1.12

For all $x \geq 2$, we have
$$\pi(x) \geq \frac{\log x}{2\log 2}.$$
Moreover, for all $n \geq 1$, we have $p_n \leq 4^n$.

Proof. Let $x \geq 2$ be an integer. Let $p_1, \ldots, p_j$ be the primes less than or equal to x. Note that we have $j = \pi(x)$ here. For every integer n with $n \leq x$, we can write $n = n_1^2 m$ where n_1 is a positive integer and m is squarefree. Then m is of the form
$$m = p_1^{\varepsilon_1} \cdots p_j^{\varepsilon_j},$$
where $\varepsilon_i \in \{0, 1\}$ for each $1 \leq i \leq j$. We see that there are at most 2^j possible values for m. Moreover, there are at most $\sqrt{x}$ possible values for n_1. Hence, we have $2^j\sqrt{x} \geq x$, which implies that $2^j \geq \sqrt{x}$. Denote this inequality by $(\star)$. Since $j = \pi(x)$, we see that
$$\pi(x)\log 2 \geq \frac{\log x}{2},$$
so the first equality follows. For the second equality, take $x = p_n$ so that $\pi(p_n) = n$. By $(\star)$, we obtain $2^n \geq \sqrt{p_n}$ and hence $4^n \geq p_n$. $\square$

Let n be a positive integer and let p be a prime. Recall that the exact power of p dividing $n!$ is
$$\sum_{n=1}^{\infty} \left\lfloor \frac{n}{p^k} \right\rfloor = \sum_{n=1}^{\left\lfloor \frac{\log n}{\log p} \right\rfloor} \left\lfloor \frac{n}{p^k} \right\rfloor.$$

Theorem 1.13

For all $x \geq 2$, we have
$$\left(\frac{3\log 2}{8}\right)\frac{x}{\log x} < \pi(x) < (6\log 2)\frac{x}{\log x}.$$

Proof. This argument was given by Erdős. First, we will prove the lower bound. Note that $\binom{2n}{n}$ is an integer, and
$$\binom{2n}{n} = \frac{(2n)!}{(n!)^2} \;\Big|\; \prod_{p \leq 2n} p^{r_p},$$
where r_p is an integer satisfying $p^{r_p} \leq 2n < p^{r_p+1}$. Indeed, note that the exact power of p dividing $(2n)!$ is
$$\sum_{k=1}^{r_p} \left\lfloor \frac{2n}{p^k} \right\rfloor,$$
and the exact power of p dividing $n!$ is
$$\sum_{k=1}^{r_p} \left\lfloor \frac{n}{p^k} \right\rfloor.$$
Thus, the exact power of p dividing $\binom{2n}{n}$ is
$$\sum_{k=1}^{r_p} \left(\left\lfloor \frac{2n}{p^k} \right\rfloor - \left\lfloor \frac{n}{p^k} \right\rfloor \right) \leq r_p,$$

since $\lfloor 2a \rfloor - 2\lfloor a \rfloor \leq 1$ for all $a \in \mathbb{R}$. In particular, we have

$$\binom{2n}{n} \leq \prod_{p \leq 2n} p^{r_p} \leq (2n)^{\pi(2n)}.$$

Notice that

$$\binom{2n}{n} = \frac{2n \cdot (2n-1) \cdots (n+1)}{n \cdot (n-1) \cdots 1} = \frac{2n}{n} \cdots \frac{n+1}{1} \geq 2^n.$$

Hence, we get $2^n \leq (2n)^{\pi(2n)}$. Now, we have

$$\pi(2n) \geq \left(\frac{\log 2}{2}\right) \frac{2n}{\log(2n)}.$$

Recall that $\frac{x}{\log x}$ is increasing for $x > e$. If $x \geq 6$, choose $n \in \mathbb{N}$ such that $3x/4 \leq 2n \leq x$. We see that

$$\pi(x) \geq \pi(2n) \geq \left(\frac{\log 2}{2}\right) \frac{2n}{\log(2n)} \geq \left(\frac{\log 2}{2}\right) \frac{\frac{3}{4}x}{\log(\frac{3}{4}x)} > \frac{3\log 2}{8} \frac{x}{\log x}.$$

One can manually check that the result holds for $2 \leq x \leq 6$, which finishes the proof of the lower bound.

We now turn to the upper bound. Observe that

$$\prod_{n < p \leq 2n} p \,\Bigg|\, \binom{2n}{n},$$

so by the binomial theorem, we have

$$\prod_{n < p \leq 2n} p \leq \binom{2n}{n} \leq (1+1)^{2n} = 2^{2n}.$$

On the other hand, notice that

$$\prod_{n < p \leq 2n} p \geq n^{\pi(2n) - \pi(n)},$$

so it follows that

$$\pi(2n)\log n - \pi(n)\log(n/2) < (\log 2)2n + (\log 2)\pi(n) < (3\log 2)n.$$

By taking $n = 2^k, 2^{k-1}, \ldots, 4$, we obtain a telescoping collection of inequalities, given by

$$\begin{aligned}
\pi(2^{k+1})\log 2^k - \pi(2^k)\log 2^{k-1} &< (3\log 2)2^k, \\
\pi(2^k)\log 2^{k-1} - \pi(2^{k-1})\log 2^{k-2} &< (3\log 2)2^{k-1}, \\
&\vdots \\
\pi(8)\log 4 - \pi(4)\log 2 &< (3\log 2)4.
\end{aligned}$$

Putting these inequalities together, we have

$$\pi(2^{k+1})\log 2^k < (3\log 2)(2^k + 2^{k+1} + \cdots + 4) + \pi(4)\log 2 < (3\log 2)2^{k+1},$$

and hence

$$\pi(2^{k+1}) < (3\log 2)\left(\frac{2^{k+1}}{\log(2^k)}\right).$$

If $x > e$, choose k such that $2^k \leq x \leq 2^{k+1}$. Then $\pi(x) \leq \pi(2^{k+1})$, and so

$$\pi(x) \leq (3\log 2)\left(\frac{2^{k+1}}{\log(2^k)}\right) \leq (6\log 2)\left(\frac{2^k}{\log(2^k)}\right) \leq (6\log 2)\left(\frac{x}{\log x}\right),$$

where in the last equality, we use the fact that $\frac{x}{\log x}$ is increasing for $x > e$. The values $2 \leq x \leq e$ can be checked manually, proving the lower bound. $\square$

We should note that $\frac{3\log 2}{8}$ is in some sense arbitrary. In the proof, we could have picked $n \in \mathbb{N}$ such that $1-\varepsilon \leq 2n \leq x$ instead of $3x/4 \leq 2n \leq x$ for ε arbitrarily small. However, this comes at the cost that the bound may potentially fail for small x, and there is little purpose in a better lower bound for large x as it is overshadowed by the Prime Number Theorem.

1.3 Bertrand's Postulate

In 1845, Bertrand showed that there is always a prime p in the interval $[n, 2n]$ for $n \in \mathbb{Z}^+$ provided that $n < 6 \cdot 10^6$, and he had conjectured that this holds for all $n \in \mathbb{Z}^+$. Chebyshev proved that this was indeed the case in 1950. Note that this is not a trivial result; it doesn't occur for free just because $\pi(x) \sim x/\log x$.

Proposition 1.14

For all $n \in \mathbb{Z}^+$, we have

$$\prod_{p \leq n} p < 4^n.$$

Proof. The result is clearly true for $n = 1$ and $n = 2$. Suppose that it holds for all $1 \leq n \leq k-1$. Note that we can restrict our attention to the case where n is odd, because if n is even and $n > 2$, then

$$\prod_{p \leq n} p = \prod_{p \leq n-1} p,$$

and the result will follow by induction. Write $n = 2m+1$ for some $m \in \mathbb{Z}^+$, and consider $\binom{2m+1}{m}$. In particular, we have

$$\prod_{m+1 < p \leq 2m+1} p \,\bigg|\, \binom{2m+1}{m}.$$

Since $\binom{2m+1}{m}$ and $\binom{2m+1}{m+1}$ both appear in the binomial expansion of $(1+1)^{2m+1}$ with $\binom{2m+1}{m} = \binom{2m+1}{m+1}$, we obtain

$$\binom{2m+1}{m} \leq \frac{1}{2}(2^{2m+1}) = 4^m.$$

By our inductive hypothesis and the previous inequality, it follows that

$$\prod_{p \leq 2m+1} p = \left(\prod_{p \leq m+1} p\right)\left(\prod_{m+1 < p \leq 2m+1} p\right) \leq 4^{m+1}4^m = 4^{2m+1}.$$

□

Lemma 1.15

If $n \geq 3$ and p is a prime with $\frac{2}{3}n < p \leq n$, then $p \nmid \binom{2n}{n}$.

Proof. Since $n \geq 3$, we see that if p is in the range $\frac{2}{3}n < p \leq n$, then $p > 2$. Then p and $2p$ are the only multiples of p at most $2n$, and so

$$p^2 \parallel (2n)!,$$

where we write $p^k \parallel b$ to mean that $p^{k+1} \nmid b$ and $p^k \mid b$. Furthermore, since $\frac{2}{3}n < p \leq n$, we have $p \parallel n!$ and hence $p^2 \parallel (n!)^2$. Using the identity

$$\binom{2n}{n} = \frac{(2n)!}{(n!)^2},$$

we see that $p \nmid \binom{2n}{n}$.

□

Theorem 1.16: Chebyshev

For every $n \in \mathbb{Z}^+$, there exists a prime satisfying $n < p \le 2n$.

Proof. This argument was given by Erdős. Note that the result holds for $1 \le n \le 3$. Assume that the result is false for some integer $n \ge 4$. By Lemma 1.15, every prime dividing $\binom{2n}{n}$ is at most $\frac{2}{3}n$.

Let p be a prime divisor of $\binom{2n}{n}$ where we have $p \le \frac{2}{3}n$. Suppose that $p^{\alpha_p} \parallel \binom{2n}{n}$ for some integer α_p. Recall that in the proof of Theorem 1.13, we defined r_p to be the integer satisfying $p^{r_p} \le 2n < p^{r_p+1}$. Then we have $\alpha_p \le r_p$, and hence $p^{\alpha_p} \le p^{r_p} \le 2n$.

If $\alpha_p \ge 2$, then $p^2 \le p^{\alpha_p} \le 2n$ so that $p \le \sqrt{2n}$. By Proposition 1.14, we have

$$\binom{2n}{n} \le \left(\prod_{\substack{p \le \frac{2}{3}n \\ \alpha_p \le 1}} p\right)\left(\prod_{\substack{p \le \frac{2}{3}n \\ \alpha_p \ge 2}} p\right) \le 4^{2n/3}(2n)^{\pi(\sqrt{2n})} \le 4^{2n/3}(2n)^{\sqrt{2n}}.$$

Note that $\binom{2n}{n}$ is the largest of the $2n+1$ terms in the binomial expansion of

$$(1+1)^{2n} = \binom{2n}{0} + \binom{2n}{1} + \cdots + \binom{2n}{2n},$$

so we get

$$\binom{2n}{n} \ge \frac{2^{2n}}{2n+1}.$$

Combining the above inequalities gives

$$\frac{4^n}{2n+1} \le \binom{2n}{n} \le 4^{2n/3}(2n)^{\sqrt{2n}},$$

which implies that

$$4^{n/3} \le (2n)^{\sqrt{2n}}(2n+1) < (2n)^{\sqrt{2n}+2}.$$

One can check manually that the result holds for $4 \le n \le 16$, so assume that $n > 16$. Taking logarithms, we find that

$$\frac{n}{3}\log 4 < (\sqrt{2n}+2)\log(2n) < 2\sqrt{n}\log(2n) < 2\sqrt{n}\log(n^{5/4}) < \frac{5}{2}\sqrt{n}\log n.$$

Notice that $\frac{\sqrt{n}}{\log n}$ is increasing for $n > e^2$. Putting this together with the fact that

$$\frac{\sqrt{1600}}{\log 1600} \approx 5.421 > 5.410 \approx \frac{15}{2\log 4},$$

we have $n \le 1600$. Finally, we know that $\{2, 3, 5, 7, 13, 23, 43, 83, 163, 317, 557, 1109, 2207\}$ are all primes, where each number in the set is the largest prime less than twice the previous one. Thus, no counterexample exists, and the result holds for all $n \ge 4$. $\square$

1.4 Gaps Between Twin Primes

By Theorem 1.16, we have

$$p_{n+1} - p_n \le p_n$$

as there is a prime between p_n and $2p_n$. What more can we say about differences of consecutive primes?

By the Prime Number Theorem, there are about $x/\log x$ primes p at most x. Therefore, the "average gap" between primes p at most x is $\log x$. However, the value of $p_{n+1} - p_n$ can vary widely.

Notice that for any $n \geq 2$, the numbers $n! + k$ for $2 \leq k \leq n$ are all composite. This implies that

$$\limsup_{n\to\infty} (p_{n+1} - p_n) = \infty.$$

In 1931, Weszynthius showed that

$$\limsup_{n\to\infty} \left(\frac{p_{n+1} - p_n}{\log p_n} \right) = \infty.$$

By probabilistic reasoning, Cramer had conjectured in 1936 that

$$\limsup_{n\to\infty} \left(\frac{p_{n+1} - p_n}{(\log p_n)^2} \right) \leq 1.$$

In the 1930s, Erdős proved that for infinitely many integers n, we have

$$p_{n+1} - p_n > c \log p_n \frac{\log\log p_n}{(\log\log\log p_n)^2}$$

for some positive constant c. In 1938, Rankin added a factor of $\log\log\log\log p_n$.

What about small gaps between consecutive primes? The famous Twin Prime Conjecture states that there are infinitely many $n \in \mathbb{Z}^+$ such that $p_{n+1} - p_n = 2$. Equivalently, it can be stated that

$$\liminf_{n\to\infty} (p_{n+1} - p_n) = 2.$$

If we assume that the primes are randomly distributed and an integer is prime with probability $1/\log x$, then we might expect x and $x + 2$ to both be prime with probability $1/(\log x)^2$.

Therefore, we expect about $x/(\log x)^2$ primes p such that $p + 2$ is also prime and $p \leq x$. A more careful heuristic suggests that there are about $Cx/(\log x)^2$ such primes p where $C > 0$ and $C \neq 1$. In the 1960s, Chen proved that there are more than $0.6x/(\log x)^2$ primes p with $p \leq x$ such that $p + 2$ is a product of at most two primes (called a P_2), provided that x is sufficiently large.

In 2005, Goldston, Pintz, and Yildirim showed that

$$\liminf_{n\to\infty} \left(\frac{p_{n+1} - p_n}{\log p_n} \right) = 0.$$

However, this is still quite far from the Twin Prime Conjecture; the bound between consecutive primes can still go to infinity.

Astoundingly, Zhang made a breakthrough in 2013 and showed that

$$\liminf_{n\to\infty} (p_{n+1} - p_n) \leq 7 \cdot 10^7.$$

This was independently improved by Tao and Maynard (via the Polymath Project) in the same year to get

$$\liminf_{n\to\infty} (p_{n+1} - p_n) \leq 246.$$

2 Asymptotic Analysis for $\pi(x)$

2.1 The Möbius Function

Definition 2.1

Let f and g be functions from $\mathbb{N}$ or $\mathbb{R}^+$ to $\mathbb{R}$, and suppose that g maps to $\mathbb{R}^+$.

(1) We write $f = O(g)$ if there exist constants $c_1, c_2 > 0$ such that for all $x > c_1$, we have $|f(x)| \leq c_2 g(x)$.

(2) We write $f = o(g)$ if $\lim_{n\to\infty} f(n)/g(n) = 0$.

(3) We write $f \sim g$ if $\lim_{n\to\infty} f(n)/g(n) = 1$, and we say that f is **asymptotic** to g.

By the Prime Number Theorem, we have $\pi(x) \sim x/\log x$, or equivalently,

$$\pi(x) = \frac{x}{\log x} + o\left(\frac{x}{\log x}\right). \tag{2.1}$$

Remark 2.2

Let $\varepsilon > 0$. Then the number of primes in the interval $[x, (1+\varepsilon)x]$ is

$$\pi((1+\varepsilon)x) - \pi(x) = \frac{(1+\varepsilon)x}{\log((1+\varepsilon)x)} - \frac{x}{\log x} + o\left(\frac{x}{\log x}\right).$$

Notice that

$$\frac{(1+\varepsilon)x}{\log((1+\varepsilon)x)} = \frac{(1+\varepsilon)x}{\log x + \log(1+\varepsilon)} = \frac{(1+\varepsilon)x}{(\log x)(1+\log(1+\varepsilon)/\log x)} = \frac{(1+\varepsilon)x}{\log x} + o\left(\frac{x}{\log x}\right).$$

Therefore, it follows that

$$\pi((1+\varepsilon)x) - \pi(x) = \frac{(1+\varepsilon)x}{\log x} - \frac{x}{\log x} + o\left(\frac{x}{\log x}\right) = \frac{\varepsilon x}{\log x} + o\left(\frac{x}{\log x}\right).$$

By taking $\varepsilon = 1$, we have

$$\pi(2x) - \pi(x) = \frac{x}{\log x} + o\left(\frac{x}{\log x}\right). \tag{2.2}$$

Equation (2.2) might look odd together with equation (2.1). Nonetheless, the result is correct; it's just that the bounds in the notation o are different.

Definition 2.3

We define the **Möbius function** on $\mathbb{N}$ by

$$\mu(n) = \begin{cases} 1 & \text{if } n = 1, \\ 0 & \text{if } n \text{ is not squarefree}, \\ (-1)^r & \text{if } n \text{ is a product of } r \text{ distinct primes.} \end{cases}$$

For example, we have $\mu(48) = \mu(2^4 \cdot 3) = 0$ and $\mu(30) = \mu(2 \cdot 3 \cdot 5) = (-1)^3 = -1$.

Proposition 2.4

We have

$$\sum_{d \mid n} \mu(d) = \begin{cases} 1 & \text{if } n = 1, \\ 0 & \text{otherwise,} \end{cases}$$

where $\sum_{d \mid n}$ means that the summation runs through the positive divisors d of n.

Proof. The result is true for $n = 1$. For $n > 1$, let $n = p_1^{a_1} \cdots p_r^{a_r}$ be the unique factorization of n into distinct prime numbers. Set $N = p_1 \cdots p_r$ (which is called the **radical** of n). Since $\mu(d) = 0$ when d is not squarefree, we have

$$\sum_{d \mid n} \mu(d) = \sum_{d \mid N} \mu(d).$$

Note that the divisors of N are in bijective correspondence with the subsets of $\{p_1, \ldots, p_r\}$. Since the number of k element subsets is $\binom{r}{k}$ and the corresponding divisor d of such a set satisfies $\mu(d) = (-1)^k$, we have

$$\sum_{d \mid n} \mu(d) = \sum_{d \mid N} \mu(d) = \sum_{k=0}^{r} \binom{r}{k} (-1)^k = (1-1)^r = 0. \qquad \square$$

Proposition 2.5: Möbius Inversion Formula

(1) For two functions $f, g : \mathbb{R}^+ \to \mathbb{C}$, we have

$$g(x) = \sum_{1 \le n \le x} f(x/n)$$

if and only if

$$f(x) = \sum_{1 \le n \le x} \mu(n) g(x/n).$$

(2) For two functions $f, g : \mathbb{N} \to \mathbb{C}$, we have

$$f(n) = \sum_{d \mid n} g(d)$$

if and only if

$$g(n) = \sum_{d \mid n} \mu(d) f(n/d).$$

Proof. This is on Homework 1. $\square$

2.2 The von Mangoldt Function

DEFINITION 2.6

We define the **von Mangoldt function** on $\mathbb{N}$ by

$$\Lambda(n) = \begin{cases} \log p & \text{if } n = p^k \text{ for } p \text{ prime and } k \in \mathbb{N}, \\ 0 & \text{otherwise.} \end{cases}$$

Moreover, for all $x \in \mathbb{R}$, we define the functions

$$\theta(x) = \sum_{p \le x} \log p = \log \prod_{p \le x} p,$$
$$\psi(x) = \sum_{p^k \le x} \log p = \sum_{n \le x} \Lambda(n).$$

Notice that

$$\psi(x) = \sum_{p \le x} \left\lfloor \frac{\log x}{\log p} \right\rfloor \log p.$$

Since $p^2 \le x$ is equivalent to $p \le x^{1/2}$ and $p^3 \le x$ if and only if $p \le x^{1/3}$, we see that

$$\psi(x) = \theta(x) + \theta(x^{1/2}) + \theta(x^{1/3}) + \cdots.$$

Note that $\theta(x^{1/m}) = 0$ when $m > \frac{\log x}{\log 2}$. Therefore, we get

$$\psi(x) = \sum_{k=1}^{\lfloor \frac{\log x}{\log 2} \rfloor} \theta(x^{1/k}).$$

Observe that we have the inequality

$$\theta(x) = \sum_{p \le x} \log p \le x \log x,$$

so it follows that

$$\sum_{k \ge 2} \theta(x^{1/k}) = O\left(x^{1/2} (\log x)^2\right).$$

Therefore, we obtain

$$\psi(x) = \theta(x) + O\left(x^{1/2} (\log x)^2\right)$$

and so by Theorem 1.12, we get

$$\theta(x) = \sum_{p \le x} \log p \le \pi(x) \log x < c_1 x$$

for $x \ge 2$ and a constant $c_1 > 0$. Similarly, one finds that $\psi(x) < c_2 x$ for $x \ge 2$ and a positive constant c_2. Furthermore, from the proof of Theorem 1.12, we have $2^n \le \binom{2n}{n}$ and $\binom{2n}{n} \mid \prod_{p \le 2n} p^{r_p}$, where r_p is the integer satisfying $p^{r_p} \le 2n < p^{r_p+1}$. It follows that

$$n \log 2 = \log(2^n) \le \log \binom{2n}{n} \le \sum_{p \le 2n} r_p \log p \le \sum_{p \le 2n} \left\lfloor \frac{\log(2n)}{\log p} \right\rfloor \log p \le \psi(2n).$$

For $x \geq 2$, choosing n such that $2n \leq x < 2n+2$ gives

$$\psi(x) \geq \psi(2n) \geq n \log 2 > \frac{x-2}{2} \log 2.$$

Hence, we have $\psi(x) > c_3 x$ and $\theta(x) > c_4 x$ for positive constants c_3 and c_4.

What is the relationship between $\theta(x)$, $\psi(x)$, and $\pi(x)$? We note that

$$\theta(x) = \sum_{p \leq x} \log p \leq x \log p \leq \pi(x) \log x,$$

so it follows that

$$\pi(x) \geq \frac{\theta(x)}{\log x} > c_4 \frac{x}{\log x}.$$

Theorem 2.7

We have

$$\pi(x) \sim \frac{\theta(x)}{\log x} \sim \frac{\psi(x)}{\log x}.$$

Proof. Since $\psi(x) = \theta(x) + O(x^{1/2}(\log x)^2)$ and $\theta(x) > c_4 x$, we see that $\theta(x) \sim \psi(x)$. In particular, we have $\theta(x)/\log x \sim \psi(x)/\log x$, so it only remains to show that $\pi(x) \sim \theta(x)/\log x$.

We have already shown that $\pi(x) \geq \theta(x) \geq \log x$, so

$$\liminf_{n \to \infty} \frac{\pi(x) \log x}{\theta(x)} \geq 1.$$

We need an upper bound for $\pi(x)$ in terms of $\theta(x)$. Note that for any $\delta > 0$, we have

$$\theta(x) = \sum_{p \leq x} \log p \geq \log(x^{1-\delta}) \sum_{x^{1-\delta} \leq p \leq x} 1 \geq (1-\delta)(\log x)\left(\pi(x) - \pi(x^{1-\delta})\right).$$

Since $\pi(y) \leq y$ for all real numbers $y > 0$, we get

$$\theta(x) + (1-\delta)x^{1-\delta} \log x \geq (1-\delta)(\log x)\pi(x).$$

Rearranging the above gives

$$\frac{\theta(x)}{(1-\delta)\log x} + x^{1-\delta} \geq \pi(x),$$

and therefore

$$\frac{1}{1-\delta} + \frac{x^{1-\delta}\log x}{\theta(x)} \geq \frac{\pi(x)\log x}{\theta(x)}.$$

Given $\varepsilon > 0$, we can choose $\delta > 0$ such that $\frac{1}{1-\delta} < 1 + \frac{\varepsilon}{2}$, and then pick x_0 such that if $x > x_0$, then

$$\frac{x^{1-\delta}\log x}{\theta(x)} < \frac{\varepsilon}{2}$$

since $\theta(x) > c_1 x$ for $x \geq 2$. Then for all $x > x_0$, we have

$$1 \leq \frac{\pi(x)\log x}{\theta(x)} < 1 + \varepsilon,$$

which completes the proof. $\square$

2.3 Abel's Summation Formula

We will prove Abel's summation formula and give some of its applications.

Lemma 2.8: Abel's summation formula

Let $\{a_n\}_{n=1}^{\infty}$ be a sequence of complex numbers. Let $f : \{x \in \mathbb{R} : x \geq 1\} \to \mathbb{C}$ be a function. For all $x \geq 1$, we define

$$A(x) := \sum_{n \leq x} a_n,$$

where the summation runs through all positive integers up to x. If f' is continuous at every $x \geq 1$, then

$$\sum_{n \leq x} a_n f(n) = A(x)f(x) - \int_1^x A(u)f'(u)\,\mathrm{d}u.$$

Proof. Set $N = \lfloor x \rfloor$. Note that $a_n = A(n) - A(n-1)$ for all $n \geq 2$, so we can write

$$\begin{aligned}\sum_{n \leq N} a_n f(n) &= A(1)f(1) + (A(2) - A(1))f(2) + \cdots + (A(N) - A(N-1))f(N) \\ &= A(1)(f(1) - f(2)) + \cdots + A(N-1)(f(N-1) - f(N)) + A(N)f(N).\end{aligned}$$

Observe that if $i \in \mathbb{Z}^+$ and $t \in \mathbb{R}$ with $i \leq t < i+1$, then $A(t) = A(i)$. It follows that

$$A(i)(f(i) - f(i+1)) = -\int_i^{i+1} A(u)f'(u)\,\mathrm{d}u.$$

Therefore, we have

$$\sum_{n \leq N} a_n f(n) = -\int_1^N A(u)f'(u)\,\mathrm{d}u + A(N)f(N),$$

so the result holds when x is an integer. Now, notice that $A(t) = A(N)$ for all $x \geq t \geq N$, so we obtain

$$\int_N^x A(u)f'(u)\,\mathrm{d}u = A(x)(f(x) - f(N)) = A(x)f(x) - A(N)f(N).$$

Thus, the result holds for all $x \geq 1$. □

Definition 2.9

Given $x \in \mathbb{R}$, we denote the **fractional part** of x by $\{x\}$; that is,

$$\{x\} := x - \lfloor x \rfloor.$$

We define **Euler's constant** by

$$\gamma := 1 - \int_1^\infty \frac{\{t\}}{t^2}\,\mathrm{d}t = 1 - \int_1^\infty \frac{t - \lfloor t \rfloor}{t^2}\,\mathrm{d}t.$$

Note that $\gamma \approx 0.55721$.

This has not been proven, but it has been conjectured that γ is irrational and transcendental.

THEOREM 2.10

We have
$$\sum_{n\le x}\frac{1}{n}=\log x+\gamma+O\left(\frac{1}{x}\right).$$

PROOF. Taking $a_n=1$ and $f(t)=1/t$ in Abel's summation formula, we have
$$A(x)=\sum_{n\le x}a_n=\sum_{n\le x}1=\lfloor x\rfloor$$
so that
$$\begin{aligned}\sum_{n\le x}\frac{1}{n}&=\frac{\lfloor x\rfloor}{x}+\int_1^x\frac{\lfloor u\rfloor}{u^2}\,\mathrm{d}u\\&=\frac{x-(x-\lfloor x\rfloor)}{x}+\int_1^x\frac{u-(u-\lfloor u\rfloor)}{u^2}\,\mathrm{d}u\\&=1+O\left(\frac{1}{x}\right)+\int_1^x\frac{\mathrm{d}u}{u}-\int_1^x\frac{u-\lfloor u\rfloor}{u^2}\,\mathrm{d}u\\&=1+O\left(\frac{1}{x}\right)+\log x-\left(\int_1^\infty\frac{u-\lfloor u\rfloor}{u^2}\,\mathrm{d}u-\int_x^\infty\frac{u-\lfloor u\rfloor}{u^2}\,\mathrm{d}u\right)\\&=\log x+\gamma+O\left(\frac{1}{x}\right)+\int_x^\infty\frac{u-\lfloor u\rfloor}{u^2}\,\mathrm{d}u\\&=\log x+\gamma+O\left(\frac{1}{x}\right)+O\left(\int_x^\infty\frac{1}{u^2}\,\mathrm{d}u\right)\\&=\log x+\gamma+O\left(\frac{1}{x}\right).\end{aligned}$$
□

THEOREM 2.11

We have
$$\sum_{n\le x}\frac{\Lambda(n)}{n}=\log x+O(1).$$

PROOF. First, we apply Abel's summation formula with $a_n=1$ and $f(n)=\log n$ to get
$$\begin{aligned}\sum_{n\le x}\log n&=\lfloor x\rfloor\log x-\int_1^x\frac{\lfloor u\rfloor}{u}\,\mathrm{d}u\\&=(x-(x-\lfloor x\rfloor))\log x-\int_1^x\frac{u-(u-\lfloor u\rfloor)}{u}\,\mathrm{d}u\\&=x\log x+O(\log x)-(x-1)+\int_1^x\frac{u-\lfloor u\rfloor}{u}\,\mathrm{d}u\\&=x\log x-x+O(\log x).\end{aligned}$$

On the other hand, we have

$$\begin{aligned}
\sum_{n\le x}\log n = \log(\lfloor x\rfloor !) &= \sum_{p\le x}\left(\sum_{k=1}^{\infty}\left\lfloor \frac{x}{p^k}\right\rfloor\right)\log p \\
&= \sum_{p^m\le x}\left\lfloor \frac{x}{p^m}\right\rfloor \log p \\
&= \sum_{n\le x}\left\lfloor \frac{x}{n}\right\rfloor \Lambda(n) \\
&= \sum_{n\le x}\frac{x}{n}\Lambda(n) - \sum_{n\le x}\left(\frac{x}{n}-\left\lfloor\frac{x}{n}\right\rfloor\right)\Lambda(n) \\
&= x\sum_{n\le x}\frac{\Lambda(n)}{n} - O\left(\sum_{n\le x}\Lambda(n)\right).
\end{aligned}$$

Since $\sum_{n\le x}\Lambda(n)=\psi(x)=O(x)$, we have

$$\sum_{n\le x}\log x = x\sum_{n\le x}\frac{\Lambda(n)}{n} - O(n).$$

By the asymptotic formula of $\sum_{n\le x}\log n$ above, we see that

$$x\log x - x + O(\log x) = x\sum_{n\le x}\frac{\Lambda(n)}{n} - O(x).$$

Rearranging and tucking some terms under $O(x)$ gives

$$x\sum_{n\le x}\frac{\Lambda(n)}{n} = x\log x + O(x).$$

Finally, dividing through by x gives

$$\sum_{n\le x}\frac{\Lambda(n)}{n} = \log x + O(1).$$

□

Theorem 2.12

We have

$$\sum_{p\le x}\frac{\log p}{p} = \log x + O(1).$$

Proof. Note that

$$\sum_{p\le x}\frac{\log p}{p} = \sum_{n\le x}\frac{\Lambda(n)}{n} - \sum_{m\ge 2}\sum_{p^m\le x}\frac{\log p}{p^m} = \log x + O(1) - \sum_{m\ge 2}\sum_{p^m\le x}\frac{\log p}{p^m}.$$

Moreover, we see that

$$\sum_{m\ge 2}\sum_{p^m\le x}\frac{\log p}{p^m} \le \sum_{p}\left(\frac{1}{p^2}+\frac{1}{p^3}+\cdots\right)\log p \le \sum_{p}\frac{\log p}{p(p-1)} \le \sum_{n=2}^{\infty}\frac{\log n}{n(n-1)} = O(1),$$

which completes the proof. □

Theorem 2.13: Merten

There exists a real number β such that

$$\sum_{p\le x}\frac{1}{p}=\log\log x+\beta+O\left(\frac{1}{\log x}\right).$$

Proof. We apply Abel's summation formula with

$$a_n=\begin{cases}\frac{\log p}{p} & \text{if } n=p \text{ for a prime } p,\\ 0 & \text{otherwise,}\end{cases}$$

and $f(n)=1/\log n$. Setting $A(x)=\sum_{n\le x}a_n$, we have

$$\sum_{p\le x}\frac{1}{p}=\frac{A(x)}{\log x}+\int_1^x\frac{A(u)}{u(\log u)^2}\,\mathrm{d}u.$$

By Theorem 2.12, we have

$$A(x)=\sum_{p\le x}\frac{\log p}{p}=\log x+O(1),$$

so we see that

$$\sum_{p\le x}\frac{1}{p}=1+O\left(\frac{1}{\log x}\right)+\int_2^x\frac{\log u+\tau(u)}{u(\log u)^2}\,\mathrm{d}u,$$

where $\tau(u)=A(u)-\log u=O(1)$. Therefore, we have

$$\begin{aligned}\sum_{p\le x}\frac{1}{p}&=1+O\left(\frac{1}{\log x}\right)+\log\log x-\log\log 2+\int_2^x\frac{\tau(u)}{u(\log u)^2}\,\mathrm{d}u\\&=\log\log x+1-\log\log 2+\int_2^\infty\frac{\tau(u)}{u(\log u)^2}\,\mathrm{d}u-\int_x^\infty\frac{\tau(u)}{u(\log u)^2}\,\mathrm{d}u+O\left(\frac{1}{\log x}\right).\end{aligned}$$

By setting β to the middle terms above, we are done. ☐

In fact, we have

$$\beta=\gamma+\sum_p\left[\log\left(1-\frac{1}{p}\right)+\frac{1}{p}\right]\approx 0.261497,$$

and β is called **Merten's constant**.

3 Riemann's Zeta Function and the Prime Number Theorem

3.1 The Riemann Zeta Function

In order to prove the Prime Number Theorem, we need to first introduce the Riemann zeta function.

Definition 3.1

For $s \in \mathbb{C}$ with $\mathrm{Re}(s) > 1$, we define the **Riemann zeta function** by

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

We will denote $s = \sigma + it$ where $\sigma, t \in \mathbb{R}$.

Note that the series $\sum_{n=1}^{\infty} n^{-s}$ converges absolutely when $\mathrm{Re}(s) > 1$.

Recall that the infinite product $\prod_n (1 + a_n)$ converges absolutely (that is, it is finite and non-zero) if and only if $\sum_n |a_n|$ converges. We have the **Euler product representation** of $\zeta(s)$ given in the following lemma.

Lemma 3.2: Euler product

For $s \in \mathbb{C}$ with $\mathrm{Re}(s) > 1$, we have

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

Proof. Note that

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

A typical term in the sum is of the form

$$\frac{1}{p_1^{\alpha_1 s} \cdots p_k^{\alpha_k s}} = \frac{1}{(p_1^{\alpha_1} \cdots p_k^{\alpha_k})^s}.$$

By the Fundamental Theorem of Arithmetic, every positive integer can be expressed uniquely as a product of primes, so the identity holds. $\square$

Theorem 3.3

$\zeta(s)$ can be analytically continued to $s \in \mathbb{C}$ with $\mathrm{Re}(s) > 0$ and $s \neq 1$. It is analytic except at the point $s = 1$ where it has a simple pole with residue 1.

Proof. For $s \in \mathbb{C}$ with $\mathrm{Re}(s) > 1$, we have $\zeta(s) = \sum_{n=1}^{\infty} n^{-s}$. By Abel's summation formula with $a_n = 1$ and $f(x) = x^{-s}$, we find that

$$\sum_{n \le x} \frac{1}{n^s} = \frac{\lfloor x \rfloor}{x^s} + s \int_1^x \frac{\lfloor u \rfloor}{u^{s+1}} \, du.$$

Letting $x \to \infty$, we obtain

$$\begin{aligned}
\zeta(s) &= 0 + s\int_1^\infty \frac{\lfloor u \rfloor}{u^{s+1}}\,\mathrm{d}u \\
&= s\int_1^\infty \frac{u-(u-\lfloor u \rfloor)}{u^{s+1}}\,\mathrm{d}u \\
&= s\int_1^\infty \frac{u}{u^{s+1}}\,\mathrm{d}u - s\int_1^\infty \frac{u-\lfloor u \rfloor}{u^{s+1}}\,\mathrm{d}u \\
&= s\left(\frac{u^{1-s}}{1-s}\bigg|_1^\infty\right) - s\int_1^\infty \frac{u-\lfloor u \rfloor}{u^{s+1}}\,\mathrm{d}u \\
&= \frac{s}{s-1} - s\int_1^\infty \frac{u-\lfloor u \rfloor}{u^{s+1}}\,\mathrm{d}u
\end{aligned}$$

for $\mathrm{Re}(s) > 1$. Note that

$$\int_1^\infty \frac{u-\lfloor u \rfloor}{u^{s+1}}\,\mathrm{d}u$$

converges for $\mathrm{Re}(s) > 0$ and represents an analytic function. Therefore, we see that

$$\frac{s}{s-1} - s\int_1^\infty \frac{u-\lfloor u \rfloor}{u^{s+1}}\,\mathrm{d}u$$

is an analytic function for $\mathrm{Re}(s) > 0$ with $s \neq 1$. This gives a meromorphic continuation of $\zeta(s)$ to the region $\{s \in \mathbb{C} : \mathrm{Re}(s) > 0\}$. Finally, note that $\frac{s}{s-1}$ has a simple pole with residue 1 at $s = 1$. □

Theorem 3.4

$\zeta(s)$ has no zeroes in the region $\{s \in \mathbb{C} : \mathrm{Re}(s) \geq 1\}$.

Proof. If $\mathrm{Re}(s) > 1$, then $\prod_p (1 - \frac{1}{p^s})^{-1}$ converges, so $\zeta(s) \neq 0$.

It only remains to consider the case where $\mathrm{Re}(s) = 1$. We will first do some preliminary work.

Recall that we denote $s = \sigma + it$ where $\sigma, t \in \mathbb{R}$. Let $\sigma > 1$. Then for all $t \in \mathbb{R}$, we have

$$\log^*(\zeta(\sigma + it)) = \log\left(\prod_p \left(1 + \frac{1}{p^s}\right)^{-1}\right) = \sum_p \sum_{n=1}^\infty \frac{1}{n}\left(\frac{1}{p^{ns}}\right),$$

where log denotes the principal branch and $\log^*$ denotes some branch of the logarithm (we have to be careful here as we are considering complex numbers). Comparing the real parts of the above equality, we have

$$\log|\zeta(\sigma + it)| = \sum_p \sum_{n=1}^\infty \frac{p^{-\sigma n}\cos(nt\log p)}{n},$$

since we can write

$$p^{-int} = e^{-int\log p} = \cos(-nt\log p) + i\sin(-nt\log p) = \cos(nt\log p) - i\sin(nt\log p)$$

and therefore $\mathrm{Re}(p^{-int}) = \cos(nt\log p)$. Moreover, observe that we have the inequality

$$\begin{aligned}
0 \leq 2(1+\cos\theta)^2 &= 2(1 + 2\cos\theta + \cos^2\theta) \\
&= 2 + 4\cos\theta + 2\cos^2\theta \\
&= 3 + 4\cos\theta + (2\cos^2\theta - 1) \\
&= 3 + 4\cos\theta + \cos(2\theta).
\end{aligned}$$

From this, we can deduce that

$$\sum_p \sum_{n=1}^{\infty} \frac{p^{-\sigma n}}{n}(3 + 4\cos(nt\log p) + \cos(2nt\log p)) \geq 0.$$

Therefore, we have

$$\log|\zeta(\sigma)|^3 + \log|\zeta(\sigma+it)|^4 + \log|\zeta(\sigma+2it)| \geq 0.$$

In particular, we see that

$$|\zeta(\sigma)|^3 \cdot |\zeta(\sigma+it)|^4 \cdot |\zeta(\sigma+2it)| \geq 1 \tag{3.1}$$

for $\sigma > 1$ and $t \in \mathbb{R}$.

Suppose now that $1 + it_0$ is a zero of $\zeta(s)$, and note that $t_0 \neq 0$ as $\zeta(s)$ has a pole at $s = 1$. By taking $t \to 1^+$ (that is, from the right), we observe tht

$$|\zeta(s)| = O((\sigma-1)^{-1})$$

since 1 is a simple pole of $\zeta(s)$. Moreover, since $1 + it_0$ is a zero of $\zeta(s)$, we have $|\zeta(\sigma+it_0)| = O(\sigma-1)$ as $\sigma \to 1^+$. Finally, we have $|\zeta(\sigma+2it_0)| = O(1)$ as $\sigma \to 1^+$ since $1 + 2it_0$ is not a simple pole of $\zeta(s)$. It follows that

$$|\zeta(\sigma)|^3 \cdot |\zeta(\sigma+it)|^4 \cdot |\zeta(\sigma+2it)| = O((\sigma-1)^{-3}) \cdot O((\sigma-1)^4) \cdot O(1) = O(\sigma-1).$$

Thus, $|\zeta(s)|^3 \cdot |\zeta(\sigma+it)|^4 \cdot |\zeta(\sigma+2it)|$ tends to 0 as $\sigma \to 1^+$. But this contradicts that the lower bound we found in (3.1), so we conclude that $\zeta(s)$ cannot have a zero when $\operatorname{Re}(s) = 1$. ☐

3.2 Newman's Theorem

Theorem 3.5: Newman

Let $\{a_n\}_{n=1}^{\infty}$ be a sequence of complex numbers with $|a_n| \leq 1$ for all $n \geq 1$. Consider the series $\sum_{n=1}^{\infty} a_n/n^s$, which converges to an analytic function $F(s)$ for $\operatorname{Re}(s) > 1$. If $F(s)$ can be analytically continued to $\operatorname{Re}(s) \geq 1$, then $\sum_{n=1}^{\infty} a_n/n^s$ converges to $F(s)$ for $\operatorname{Re}(s) \geq 1$.

Proof. Let $w \in \mathbb{C}$ with $\operatorname{Re}(w) \geq 1$. Then $F(z+w)$ is analytic for $\operatorname{Re}(z) \geq 0$. Choose $R \geq 1$ and let $\delta = \delta(R) > 0$ so that $F(z+w)$ is analytic on the region

$$\tilde{\Gamma} := \{z \in \mathbb{C} : \operatorname{Re}(z) \geq -\delta \text{ and } |z| \leq R\}.$$

To see why such a $\delta > 0$ exists, first note that $F(z+w)$ is analytic for $\operatorname{Re}(z) \geq 0$. Consider the line $L = \{z = iy : |y| \leq R\}$. Every point in L has an open cover such that $F(z+w)$ is analytic on that cover; call the union of these covers U. Since L is compact[1], there exists a finite open subcover $\tilde{U}$ of U such that $L \subseteq \tilde{U} \subseteq U$. Since the number of open sets in $\tilde{U}$ is finite, it follows that such a $\delta > 0$ exists.

Let M denote the maximum of $|F(z+w)|$ on $\tilde{\Gamma}$, and let Γ denote the contour obtained by following the outside of $\tilde{\Gamma}$ in a counterclockwise path. Let A be the part of Γ in $\operatorname{Re}(z) > 0$, and let $B = \Gamma \setminus A$. For $N \in \mathbb{N}$, consider the function

$$F(z+w)N^z\left(\frac{1}{z} + \frac{z}{R^2}\right),$$

which is analytic on $\tilde{\Gamma}$ except at $z = 0$ where there is a simple pole with residue $F(0+w)N^0 = F(w)$. By Cauchy's residue theorem, we obtain

$$\begin{aligned} 2\pi i F(w) &= \int_\Gamma F(z+w)N^z\left(\frac{1}{z} + \frac{z}{R^2}\right) \mathrm{d}z \\ &= \int_A F(z+w)N^z\left(\frac{1}{z} + \frac{z}{R^2}\right) \mathrm{d}z + \int_B F(z+w)N^z\left(\frac{1}{z} + \frac{z}{R^2}\right) \mathrm{d}z. \end{aligned} \tag{3.2}$$

[1]Recall that a set X is compact if every open cover of X has a finite subcover.

Observe that $F(z+w)$ is equal to its series on A. We split the series as

$$S_N(z+w) = \sum_{n=1}^{N} \frac{a_n}{n^{z+w}}$$

and $R_N(z+w) = F(z+w) - S_N(z+w)$. Note that $S_N(z+w)$ is analytic for all $z \in \mathbb{C}$. Let C be the contour given by the path $|z| = R$ taken in the counterclockwise direction. By Cauchy's residue theorem, we obtain

$$2\pi i S_N(w) = \int_C S_N(z+w) N^z \left(\frac{1}{z} + \frac{z}{R^2}\right) \mathrm{d}z$$

since the integrand has a simple pole at $z = 0$ with residue $S_N(0+w)N^0 = S_N(w)$. Note that

$$C = A \cup (-A) \cup \{iR, -iR\}.$$

Therefore, we see that

$$2\pi i S_N(w) = \int_A S_N(z+w) N^z \left(\frac{1}{z} + \frac{z}{R^2}\right) \mathrm{d}z + \int_{-A} S_N(z+w) N^z \left(\frac{1}{z} + \frac{z}{R^2}\right) \mathrm{d}z.$$

Consider the second integral above. Using the change of variables $z \to -z$, we find that

$$\int_{-A} S_N(z+w) N^z \left(\frac{1}{z} + \frac{z}{R^2}\right) \mathrm{d}z = \int_A S_N(-z+w) N^{-z} \left(\frac{1}{z} + \frac{z}{R^2}\right) \mathrm{d}z.$$

Thus, we obtain

$$2\pi i S_N(w) = \int_A \left(S_N(z+w)N^z + S_N(-z+w)N^{-z}\right) \left(\frac{1}{z} + \frac{z}{R^2}\right) \mathrm{d}z.$$

Combining the above equality with (3.2), we have

$$\begin{aligned} 2\pi i(F(w) - S_N(w)) &= \int_A \left(R_N(z+w)N^z - S_N(-z+w)N^{-z}\right) \left(\frac{1}{z} + \frac{z}{R^2}\right) \mathrm{d}z \\ &\quad + \int_B F(z+w) N^z \left(\frac{1}{z} + \frac{z}{R^2}\right) \mathrm{d}z. \end{aligned} \tag{3.3}$$

Our goal is to show that $S_N(w)$ converges to $F(w)$ as $N \to \infty$. Write $z = x + iy$ where $x, y \in \mathbb{R}$. Then for $z \in A$, we have $x > 0$ and $|z| = R$, so

$$\frac{1}{z} + \frac{z}{R^2} = \frac{x - iy}{R^2} + \frac{x + iy}{R^2} = \frac{2x}{R^2}.$$

Since $|n^z| = n^x$, we have

$$|R_N(z+w)| \leq \sum_{n=N+1}^{\infty} \frac{1}{n^{\operatorname{Re}(z+w)}} \leq \sum_{n=N+1}^{\infty} \frac{1}{n^{x+1}} \leq \int_N^{\infty} \frac{1}{u^{x+1}} \,\mathrm{d}u = \frac{1}{xN^x}.$$

Also, we have

$$|S_N(-z+w)| \leq \sum_{n=1}^{N} \frac{1}{n^{-x+1}} \leq N^{x-1} + \int_1^N u^{x-1} \,\mathrm{d}u \leq N^{x-1} + \frac{N^x}{x} = N^x \left(\frac{1}{N} + \frac{1}{x}\right).$$

Putting the above estimates together, we get

$$\begin{aligned}\left|\int_A \left(R_N(z+w)N^z - S_N(-z+w)N^{-z}\right)\left(\frac{1}{z}+\frac{z}{R^2}\right)\mathrm{d}z\right| &\le \int_A \left(\frac{1}{xN^x}N^x + N^x\left(\frac{1}{N}+\frac{1}{x}\right)N^{-x}\right)\frac{2x}{R^2}\,\mathrm{d}z \\ &= \int_A \left(\frac{2}{x}+\frac{1}{N}\right)\frac{2x}{R^2}\,\mathrm{d}z \\ &= \int_A \left(\frac{4}{R^2}+\frac{2x}{NR^2}\right)\mathrm{d}z \\ &\le \pi R\left(\frac{4}{R^2}+\frac{2}{NR}\right) \quad \text{(since } x \le R\text{)} \\ &\le \frac{4\pi}{R}+\frac{2\pi}{N}.\end{aligned}$$

We now estimate the integral along B. We can divide B into two parts; one part with $\mathrm{Re}(z) = -\delta$, and the other with $-\delta < \mathrm{Re}(z) \le 0$. For $z \in B$ with $\mathrm{Re}(z) = -\delta$, we use the fact that $|z| \le R$ to find that

$$\left|\frac{1}{z}+\frac{z}{R^2}\right| = \left|\frac{1}{z}\right|\left|\frac{\bar{z}}{z}+\frac{z\bar{z}}{R^2}\right| \le \frac{1}{\delta}\left(1+\frac{|z|^2}{R^2}\right) \le \frac{2}{\delta}.$$

Since $|F(z+w)| \le M$ for $z \in B$, we have

$$\begin{aligned}\left|\int_B F(z+w)N^z\left(\frac{1}{z}+\frac{z}{R^2}\right)\mathrm{d}z\right| &\le \int_{-R}^{R} MN^{-\delta}\frac{2}{\delta}\,\mathrm{d}z + 2\left|\int_{-\delta}^{0} MN^x\frac{2x}{R^2}\,\mathrm{d}x\right| \\ &= \frac{4MR}{\delta N^\delta} + \frac{4M}{R^2}\left|\int_{-\delta}^{0} xN^x\,\mathrm{d}x\right| \\ &\le \frac{4MR}{\delta N^\delta} + \frac{4M\delta}{R^2}\left(\frac{1}{(\log N)^2} - \frac{\delta+1}{N^\delta \log N}\right) \\ &\le \frac{4MR}{\delta N^\delta} + \frac{4M\delta}{R^2(\log N)^2}.\end{aligned}$$

Combining this estimate with (3.2) and (3.3) yields

$$|2\pi i(F(w) - S_N(w))| \le \frac{4\pi}{R}+\frac{2\pi}{N}+\frac{4MR}{\delta N^\delta}+\frac{4M\delta}{R^2(\log N)^2}.$$

That is, we have

$$|F(w) - S_N(w)| \le \frac{2}{R}+\frac{1}{N}+\frac{MR}{\delta N^\delta}+\frac{M\delta}{R^2(\log N)^2}.$$

Given $\varepsilon > 0$, choose $R = 3/\varepsilon$. Then for sufficiently large N, we have

$$|F(w) - S_N(w)| < \varepsilon.$$

This implies that $S_N(w) \to F(w)$ as $N \to \infty$, which completes the proof. $\square$

3.3 Revisiting the Möbius Function

Recall that we defined the Möbius function $\mu : \mathbb{N} \to \{-1, 0, 1\}$ by

$$\mu(n) = \begin{cases} 1 & \text{if } n = 1, \\ 0 & \text{if } n \text{ is not squarefree}, \\ (-1)^r & \text{if } n \text{ is the product of } r \text{ distinct primes.} \end{cases}$$

We will show on Homework 2 that for $\mathrm{Re}(s) > 1$, we have

$$\frac{1}{\zeta(s)} = \prod_p \left(1 - \frac{1}{p^s}\right) = \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s}.$$

Theorem 3.6

We have
$$\sum_{n=1}^{\infty} \frac{\mu(n)}{n} = 0.$$

Proof. For all $\mathrm{Re}(s) > 1$, equation (3.4) holds. Moreover, we have shown that $(s-1)\zeta(s)$ is analytic and non-zero in $\mathrm{Re}(s) \geq 1$, so $1/\zeta(s)$ is analytic on $\mathrm{Re}(s) \geq 1$. Now, $\zeta(s)$ can be analytically continued up to $\mathrm{Re}(s) > 0$ and it is nonzero for $\mathrm{Re}(s) \geq 1$, so we see that the series
$$\sum_{n=1}^{\infty} \frac{\mu(n)}{n^s}$$
converges to $1/\zeta(s)$ for $\mathrm{Re}(s) \geq 1$. In particular, it converges at $s = 1$. But $\zeta(s)$ has a simple pole at $s = 1$, so $1/\zeta(1) = 0$. $\square$

Theorem 3.7

We have
$$\sum_{n \leq x} \mu(n) = o(x).$$

Proof. Applying Abel's summation formula with $a_n = \mu(n)/n$ and $f(x) = x$, we obtain
$$\sum_{n \leq x} \mu(n) = A(x)x - \int_1^x A(u)\,\mathrm{d}u,$$
where we hae
$$A(t) = \sum_{n \leq t} \frac{\mu(n)}{n}.$$
By Theorem 3.5, we know that $A(t) = o(1)$. It follows that $A(x)x = o(x)$ and
$$\int_1^x A(u)\,\mathrm{d}u = o(x),$$
so the result holds. $\square$

3.4 Divisor Function

Definition 3.8

For a positive integer $n \in \mathbb{N}$, let $d(n)$ be the number of positive integers that divide n.

For example, we have $d(1) = 1$, $d(4) = 3$, and $d(p) = 2$ for all primes p.

Theorem 3.9

We have
$$\sum_{m=1}^{n} d(m) = \sum_{m=1}^{n} \left\lfloor \frac{n}{m} \right\rfloor = n\log n + (2\gamma - 1)n + O(n^{1/2}).$$
where γ denotes Euler's constant.

Proof. Let D_n be the region in the upper right-hand quadrant not containing the x or y axes, which is under and includes the hyperbola $xy = n$. That is,
$$D_n := \{(x, y) \in \mathbb{R}^2 : x > 0,\, y > 0,\, xy \leq n\}.$$
Define a **lattice point** to be a point in the plane with integer coordinates; that is, a point $(x, y) \in \mathbb{R}^2$ with $x, y \in \mathbb{Z}$. Notice that every lattice point in D_n is contained in some hyperbola $xy = s$ where s is an integer with $1 \leq s \leq n$.

Therefore, $\sum_{s=1}^{n} d(s)$ is the number of lattice points in D_n; that is,
$$\sum_{s=1}^{n} d(s) = \#\{(x, y) \in \mathbb{R}^2 : x, y \in \mathbb{N},\, xy \leq n\}.$$
We now count the number of lattice points in a different way. Given $x \in \mathbb{N}$ with $1 \leq x \leq n$, there are exactly $\lfloor \frac{x}{n} \rfloor$ many integers y such that $xy \leq n$. Thus, we see that
$$\#\{(x, y) \in \mathbb{R}^2 : x, y \in \mathbb{N},\, xy \leq n\} = \sum_{x=1}^{n} \left\lfloor \frac{n}{x} \right\rfloor.$$
Observe that the number of lattice points above the line $x = y$ inside D_n is equal to the number of lattice points below it. Divide the lattice points in D_n into three disjoint regions given by
$$\begin{aligned} D_{n,1} &= \{(x, y) \in \mathbb{N}^2 : xy \leq n,\, x < y\}, \\ D_{n,2} &= \{(x, y) \in \mathbb{N}^2 : xy \leq n,\, x > y\}, \\ D_{n,3} &= \{(x, y) \in \mathbb{N}^2 : xy \leq n,\, x = y\}. \end{aligned}$$
Our observation above shows that $|D_{n,1}| = |D_{n,2}|$. Suppose that $(x, y) \in D_{n,1}$. Then $x^2 < xy \leq n$, which implies that $x < \sqrt{n}$. Moreover, for a fixed integer x, the number of integers y satisfying $xy \leq n$ and $y > x$ is $\lfloor \frac{n}{x} \rfloor - \lfloor x \rfloor$. We also see that $|D_{n,3}| = \lfloor \sqrt{n} \rfloor$, so we obtain
$$\begin{aligned} \sum_{x=1}^{n} \left\lfloor \frac{n}{x} \right\rfloor &= |D_{n,1}| + |D_{n,2}| + |D_{n,3}| \\ &= 2\sum_{x=1}^{\lfloor \sqrt{n} \rfloor} \left(\left\lfloor \frac{n}{x} \right\rfloor - \lfloor x \rfloor \right) + \lfloor \sqrt{n} \rfloor \\ &= 2\sum_{x=1}^{\lfloor \sqrt{n} \rfloor} \left(\frac{n}{x} - x + O(1) \right) + \lfloor \sqrt{n} \rfloor. \end{aligned}$$
By Theorem 2.10, we see that
$$\sum_{x=1}^{n} \left\lfloor \frac{n}{x} \right\rfloor = 2n\left(\log\lfloor \sqrt{n} \rfloor + \gamma + O\left(\frac{1}{\sqrt{n}} \right) \right) - \left(n + O(\sqrt{n}) \right) + O(\sqrt{n}).$$

Note that if we use the fact that $\log\lfloor\sqrt{n}\rfloor = \log\sqrt{n} + O(1)$, then the resulting error term $O(n)$ will be too large. Therefore, we need a finer estimate. Indeed, since $\lfloor\sqrt{n}\rfloor = \sqrt{n} - \{\sqrt{n}\}$ where $\{t\}$ denotes the fractional part of t for $t \in \mathbb{R}$, we have

$$\begin{aligned}\log\lfloor\sqrt{n}\rfloor = \log\left(\sqrt{n} - \{\sqrt{n}\}\right) &= \log\left(\sqrt{n}\left(1 - \frac{\{\sqrt{n}\}}{\sqrt{n}}\right)\right) \\ &= \log\sqrt{n} + \log\left(1 - \frac{\{\sqrt{n}\}}{\sqrt{n}}\right) \\ &= \log\sqrt{n} + O\left(\frac{1}{\sqrt{n}}\right).\end{aligned}$$

Combining this with the previous equality gives

$$\sum_{x=1}^{n}\left\lfloor\frac{n}{x}\right\rfloor = n\log n + (2\gamma - 1)n + O(\sqrt{n}).$$ □

3.5 The Prime Number Theorem

We now have everything we need to prove the Prime Number Theorem.

Theorem 3.10: Prime Number Theorem

We have

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

Proof. In Theorem 2.7, we showed that

$$\pi(x) \sim \frac{\psi(x)}{\log x}.$$

Therefore, it suffices to show that $\psi(x) \sim x$. Define the function

$$F(x) = \sum_{n\le x}\left(\psi\left(\frac{x}{n}\right) - \left\lfloor\frac{x}{n}\right\rfloor + 2\gamma\right),$$

where γ denotes Euler's constant. By the Möbius inversion formula (Proposition 2.5), we have

$$\psi(x) - \lfloor x\rfloor + 2\gamma = \sum_{n\le x}\mu(n)F\left(\frac{x}{n}\right).$$

In particular, we get

$$\psi(x) = x + O(1) + \sum_{n\le x}\mu(n)F\left(\frac{x}{n}\right).$$

Now, it is enough to show that $\sum_{n\le x}\mu(n)F(x/n) = o(x)$. First, we will estimate $F(x)$. Observe that

$$F(x) = \sum_{n\le x}\psi\left(\frac{x}{n}\right) - \sum_{n\le x}\left\lfloor\frac{x}{n}\right\rfloor + 2\gamma\lfloor x\rfloor. \tag{3.4}$$

Looking at the first sum in (3.4), we have

$$\begin{aligned}\sum_{n\le x}\psi\left(\frac{x}{n}\right) = \sum_{n\le x}\sum_{m\le \frac{x}{n}}\Lambda(m) &= \sum_{n\le x}\Lambda(n)\sum_{m\le\frac{x}{n}}1\\ &= \sum_{n\le x}\Lambda(n)\left\lfloor\frac{x}{n}\right\rfloor\\ &= \sum_{p^k\le x}\log p\left\lfloor\frac{x}{p^k}\right\rfloor\\ &= \sum_{p\le x}\left(\left\lfloor\frac{x}{p}\right\rfloor+\left\lfloor\frac{x}{p^2}\right\rfloor+\cdots+\left\lfloor\frac{x}{p^k}\right\rfloor\right) \quad (\text{where } p^k \parallel \lfloor x\rfloor)\\ &= \log(\lfloor x\rfloor!) = \sum_{n\le x}\log n.\end{aligned}$$

In the proof of Theorem 2.11, we showed that

$$\sum_{n\le x}\log n = x\log x - x + O(\log x).$$

Hence, we obtain

$$\sum_{n\le x}\psi\left(\frac{x}{n}\right) = x\log x - x + O(\log x). \tag{3.5}$$

Moreover, by Theorem 3.9, we have

$$\sum_{n=1}^{\lfloor x\rfloor}\left\lfloor\frac{\lfloor x\rfloor}{n}\right\rfloor = \lfloor x\rfloor\log\lfloor x\rfloor + (2\gamma-1)\lfloor x\rfloor + O(x^{1/2}).$$

For all $y\in\mathbb{R}$, notice that $\lfloor y\rfloor \le y \le \lfloor y\rfloor + 1$. In particular, we obtain the inequalities

$$\sum_{n=1}^{\lfloor x\rfloor}\left\lfloor\frac{\lfloor x\rfloor}{n}\right\rfloor \le \sum_{n=1}^{\lfloor x\rfloor}\left\lfloor\frac{x}{n}\right\rfloor \le \sum_{n=1}^{\lfloor x\rfloor+1}\left\lfloor\frac{\lfloor x\rfloor+1}{n}\right\rfloor,$$

and it follows that

$$\sum_{n=1}^{\lfloor x\rfloor}\left\lfloor\frac{x}{n}\right\rfloor = x\log x + (2\gamma-1) + O(x^{1/2}). \tag{3.6}$$

Combining equations (3.4), (3.5), and (3.6) gives

$$F(x) = (x\log x - x + O(\log x)) - (x\log x + (2\gamma-1)x + O(x^{1/2})) + (2\gamma x + O(1)) = O(x^{1/2}).$$

Hence, there exists a positive constant $c>0$ such that $|F(x)|\le cx^{1/2}$ for all $x\ge 1$. If $t>1$ is an integer, then

$$\begin{aligned}\left|\sum_{n\le\frac{x}{t}}\mu(n)F\left(\frac{x}{n}\right)\right| &\le \sum_{n\le\frac{x}{t}}\left|F\left(\frac{x}{n}\right)\right|\\ &\le \sum_{n\le\frac{x}{t}}c\left(\frac{x}{n}\right)^{1/2}\\ &\le cx^{1/2}\left(1+\int_1^{x/t}\frac{1}{u^{1/2}}\,\mathrm{d}u\right)\\ &= cx^{1/2}\left(1+2\left(\frac{x}{t}\right)^{1/2}-2\right)\\ &\le 2\cdot\frac{cx}{t^{1/2}}.\end{aligned} \tag{3.7}$$

Observe that F is a step function. That is, if a is an integer and $a \le x < a+1$, then $F(x) = F(a)$. Therefore, we have

$$\sum_{\frac{x}{t}<n\le x} \mu(n)F\left(\frac{x}{n}\right) = F(1)\sum_{\frac{x}{2}<n\le x}\mu(n) + F(2)\sum_{\frac{x}{3}<n\le\frac{x}{2}}\mu(n) + \cdots + F(t-1)\sum_{\frac{x}{t}<n\le\frac{x}{t-1}}\mu(n).$$

We see that

$$\begin{aligned}
\left|\sum_{\frac{x}{t}<n\le x}\mu(n)F\left(\frac{x}{n}\right)\right| &\le |F(1)|\left|\sum_{\frac{x}{2}<n\le x}\mu(n)\right| + |F(2)|\left|\sum_{\frac{x}{3}<n\le\frac{x}{2}}\mu(n)\right| + \cdots + |F(t-1)|\left|\sum_{\frac{x}{t}<n\le\frac{x}{t-1}}\mu(n)\right| \\
&\le (|F(1)| + \cdots + |F(t-1)|)\max_{2\le i\le t}\left|\sum_{\frac{x}{i}<n\le\frac{x}{i-1}}\mu(n)\right| \\
&\le \left(\sum_{i=1}^{t} ci^{1/2}\right)\max_{2\le i\le t}\left|\sum_{\frac{x}{i}<n\le\frac{x}{i-1}}\mu(n)\right|.
\end{aligned}$$

Notice that

$$\sum_{\frac{x}{i}<n\le\frac{x}{i-1}}\mu(n) = \sum_{n\le\frac{x}{i-1}}\mu(n) - \sum_{\frac{x}{i}<n}\mu(n) = o(x),$$

so we obtain

$$\left|\sum_{\frac{x}{t}<n\le x}\mu(n)F\left(\frac{x}{n}\right)\right| = o(t^{3/2}x).$$

By Theorem 3.7, we have $\sum_{n\le x}\mu(n) = o(x)$. Hence, for any $\varepsilon > 0$, we can find sufficiently large x such that

$$-\varepsilon x \le \sum_{n\le x}\mu(n) \le \varepsilon x.$$

In particular, when x is sufficiently large, we get

$$-\frac{\varepsilon x}{i-1} - \frac{\varepsilon x}{i} \le \sum_{\frac{x}{i}<n\le\frac{x}{i-1}}\mu(n) \le \frac{\varepsilon x}{i-1} + \frac{\varepsilon x}{i}.$$

For any given $\varepsilon > 0$, choose $t = t(\varepsilon)$ such that

$$\frac{2c}{t^{1/2}} < \frac{\varepsilon}{2}.$$

By equation (3.7), we have

$$\left|\sum_{n\le\frac{x}{t}}\mu(n)F\left(\frac{x}{n}\right)\right| \le 2\cdot\frac{cx}{t^{1/2}} < \frac{\varepsilon}{2}x. \tag{3.8}$$

For fixed $\varepsilon > 0$ and t as above, we can choose x sufficiently large so that $o(xt^{3/2}) \le \varepsilon x/2$. Indeed, we have $2c/t^{1/2} < \varepsilon/2$ if and only if $t > (4c)^2/\varepsilon^2$. In particular, we have $t = A^2\varepsilon^{-2}$ for some $A > 4c$, and we can pick x large enough so that

$$o(x) \le \frac{\varepsilon^4}{2A^3}x.$$

Then we get

$$o(xt^{3/2}) \le \frac{\varepsilon^4}{2A^3}x\cdot A^3\varepsilon^{-3} = \frac{\varepsilon}{2}x.$$

It follows that

$$\left| \sum_{\frac{x}{t} < n \leq x} \mu(n) F\left(\frac{x}{n}\right) \right| < \frac{\varepsilon}{2}. \tag{3.9}$$

Combining inequalities (3.8) and (3.9) yields

$$\left| \sum_{n \leq x} \mu(n) F\left(\frac{x}{n}\right) \right| = o(x),$$

which completes the proof. □

Remark 3.11

(1) In 1896, Hadamard and de la Vallée Poussin proved the Prime Number Theorem independently. Consider the logarithmic integral

$$\mathrm{Li}(x) = \int_2^x \frac{1}{\log t} \, \mathrm{d}t \sim \frac{x}{\log x} \sum_{k=0}^{\infty} \frac{k!}{(\log x)^k}.$$

In 1899, de la Vallée Poussin proved that as $x \to \infty$, there exists some $a > 0$ such that

$$\pi(x) = \mathrm{Li}(x) + O(xe^{-a\sqrt{\log x}}).$$

(2) The main ingredient of our proof of the Prime Number Theorem is the fact that $\sum_{n \leq x} \mu(n) = o(x)$, which is a consequence of the analytic continuation and non-vanishing of $\zeta(s)$ at $\mathrm{Re}(s) = 1$. The **Riemann hypothesis**, proposed by Riemann in 1859, states that the non-trivial zeros of $\zeta(s)$ all have real part $1/2$. (The trivial zeros of $\zeta(s)$ are of the form $2n$ for $n \in \mathbb{Z}$ and $n < 0$; these can be obtained by functional equations.) In 1901, Helge von Koch proved that the Riemann hypothesis is true if and only if

$$\pi(x) = \mathrm{Li}(x) + O(\sqrt{x} \log x).$$

4 Divisor Counting Functions

4.1 Asymptotic Formulas for Divisor Counting Functions

Definition 4.1

For a positive integer $n \in \mathbb{N}$, we denote by $\Omega(n)$ the number of prime factors of n counted with multiplicity, and $\omega(n)$ the number of distinct prime factors of n.

For example, if $n = 2^{10} \cdot 3^2 \cdot 7$, then $\Omega(n) = 10 + 2 + 1 = 13$ and $\omega(n) = 3$.

Definition 4.2

Let $k \in \mathbb{N}$. For each real number $x \in \mathbb{R}$, we define $\tau_k(x)$ to be the number of positive integer with $n \leq x$ and $\Omega(n) = k$. That is,

$$\tau_k(x) = \#\{n \leq x : \Omega(n) = k\}.$$

Furthermore, we let $\pi_k(x)$ be the number of positive integers n with $n \leq x$ and $\omega(n) = \Omega(n) = k$. That is,

$$\pi_k(x) = \#\{n \leq x : \omega(n) = \Omega(n) = k\}.$$

In particular, $\pi_k(x)$ counts the positive integers n up to x which are squarefree and have k prime factors. Note that $\pi(x) = \pi_1(x) = \tau_1(x)$.

Theorem 4.3: Landau, 1900

Let $k \in \mathbb{N}$ be a positive integer. Then

$$\pi_k(x) \sim \tau_k(x) \sim \frac{1}{(k-1)!} \frac{x}{\log x} (\log \log x)^{k-1}.$$

Proof. We first introduce the functions

$$L_k(x) = \sum_{p_1 \cdots p_k \leq x}^{*} \frac{1}{p_1 \cdots p_k}, \qquad \Pi_k(x) = \sum_{p_1 \cdots p_k \leq x}^{*} 1, \qquad \Theta_k(x) = \sum_{p_1 \cdots p_k \leq x}^{*} \log(p_1 \ldots p_k),$$

where the $*$ means that the sum is taken over all k-tuples of primes $(p_1, \ldots, p_k)$ with $p_1 \cdots p_k \leq x$. Note that different k-tuples can correspond to the same product $p_1 \cdots p_k$.

For each positive integer $n \geq 1$, we let $c_n = c_n(k)$ denote the number of k-tuples $(p_1, \ldots, p_k)$ such that $p_1 \cdots p_k = n$. Observe that

$$\Pi_k(x) = \sum_{n \leq x} c_n,$$
$$\Theta_k(x) = \sum_{n \leq x} c_n \log n.$$

Moreover, we have

$$c_n = \begin{cases} 0 & \text{if } n \text{ is not a product of } k \text{ primes,} \\ k! & \text{if } n \text{ is squarefree and } \omega(n) = \Omega(n) = k. \end{cases}$$

We also see that $0 < c_n < k!$ if $\Omega(n) = k$ but n is not squarefree. Therefore, we obtain the inequalities

$$k!\pi_k(x) \leq \Pi_k(x) \leq k!\tau_k(x). \tag{4.1}$$

For $k \geq 2$, note that the number of positive integers up to x with k prime factors and divisible by the square of some prime is $\tau_k(x) - \pi_k(x)$. Therefore, we have

$$\tau_k(x) - \pi_k(x) = \sideset{}{^*}\sum_{\substack{p_1 \cdots p_k \leq x \\ p_i = p_j \text{ for some } i \neq j}} 1 \leq \binom{k}{2} \sideset{}{^*}\sum_{p_1 \cdots p_k \leq x} 1 = \binom{k}{2} \Pi_{k-1}(x).$$

Claim. We have

$$\Pi_k(x) \sim k \frac{x (\log \log x)^{k-1}}{\log x}.$$

Proof of Claim. Applying Abel's summation formula with $a_n = c_n$ and $f(u) = \log u$, we have

$$\Theta_k(x) = \sum_{n \leq x} c_n \log n = \Pi_k(x) \log x - \int_1^x \frac{\Pi_k(u)}{u} \, \mathrm{d}u.$$

Observe that

$$\Pi_k(x) \leq k!\tau_k(x) \leq k!x,$$

so $\Pi_k(u) = O(u)$, and hence

$$\Theta_k(x) = \Pi_k(x) \log x + O(x).$$

Thus, it suffices to show that for all $k \in \mathbb{N}$, we have

$$\Theta_k(x) \sim kx (\log \log x)^{k-1}. \tag{4.2}$$

We'll proceed by induction on k. This will be somewhat similar to the proof of the Prime Number Theorem, but with the weighting function $\log(p_1 \cdots p_k)$ on the k-tuple $(p_1, \ldots, p_k)$.

For $k = 1$, we have $\Theta_1(x) = \theta(x) \sim x$ by Theorem 2.7 and the Prime Number Theorem. Assume now that $\Theta_k(x) \sim kx(\log \log x)^{k-1}$ for some $k \geq 1$. We'll prove the result for $\Theta_{k+1}(x)$. First, note that

$$\left(\sum_{p \leq x^{1/k}} \frac{1}{p} \right)^k \leq L_k(x) \leq \left(\sum_{p \leq x} \frac{1}{p} \right)^k$$

for all $k \geq 1$. By Theorem 2.13, we have

$$\left(\sum_{p \leq x^{1/k}} \frac{1}{p} \right)^k \sim \left(\log \log (x^{1/k}) \right)^k,$$

$$\left(\sum_{p \leq x} \frac{1}{p} \right)^k \sim (\log \log x)^k.$$

Notice that

$$\left(\log \log (x^{1/k}) \right)^k = (\log \log x - \log k)^k \sim (\log \log x)^k,$$

so $L_k \sim (\log \log x)^k$. Therefore, we have

$$\Theta_{k+1}(x) - (k+1)(\log \log x)^k = \Theta_{k+1}(x) - (k+1) x L_k(x) + o\left(x (\log \log x)^k \right).$$

Note that

$$\begin{aligned}
k\Theta_{k+1}(x) &= \sideset{}{^*}\sum_{p_1\cdots p_{k+1}\le x} k\cdot\log(p_1\cdots p_{k+1}) \\
&= \sideset{}{^*}\sum_{p_1\cdots p_{k+1}\le x} (\log(p_2\cdots p_{k+1}) + \log(p_1p_3\cdots p_{k+1}) + \cdots + \log(p_1\cdots p_k)) \\
&= (k+1)\sum_{p_1\le x}\ \sideset{}{^*}\sum_{p_2\cdots p_{k+1}\le x/p_1} \log(p_2\cdots p_{k+1}) \\
&= (k+1)\sum_{p_1\le x}\Theta_k\left(\frac{x}{p_1}\right).
\end{aligned}$$

Since $L_0(x)=1$ and

$$L_k(x) = \sideset{}{^*}\sum_{p_1\cdots p_k\le x}\frac{1}{p_1\cdots p_k} = \sum_{p_1\le x}\frac{1}{p_1}L_{k-1}\left(\frac{x}{p_1}\right),$$

it follows that

$$\begin{aligned}
\Theta_{k+1}(x) - (k+1)xL_k(x) &= (k+1)\sum_{p_1\le x}\left(\frac{1}{k}\Theta_k\left(\frac{x}{p_1}\right) - \frac{x}{p_1}L_{k-1}\left(\frac{x}{p_1}\right)\right) \\
&= \frac{k+1}{k}\sum_{p_1\le x}\left(\Theta_k\left(\frac{x}{p_1}\right) - k\frac{x}{p_1}L_{k-1}\left(\frac{x}{p_1}\right)\right).
\end{aligned}$$

By the induction hypothesis, we have

$$\Theta_k(y) - kyL_{k-1}(y) = o(y(\log\log y)^{k-1}).$$

Given $\varepsilon>0$, there exists $x_0 = x_0(\varepsilon,k)$ such that for all $y > x_0$, we have

$$|\Theta_k(y) - kyL_{k-1}(y)| \le \varepsilon y(\log\log y)^{k-1}.$$

Furthermore, there exists a positive constant $c = c(\varepsilon,k) > 0$ such that for all $y\le x_0$, we have

$$|\Theta_k(y) - kyL_{k-1}(y)| \le c.$$

Note that $x/p_1 > x_0$ implies that $p_1 < x/x_0$, so for sufficiently large x, we obtain

$$\begin{aligned}
|\Theta_{k+1}(x) - (k+1)xL_k(x)| &\le \frac{k+1}{k}\left(\sum_{\frac{x}{x_0}<p_1\le x} c + \sum_{p_1\le\frac{x}{x_0}}\varepsilon\frac{x}{p_1}\left(\log\log\frac{x}{p_1}\right)^{k-1}\right) \\
&\le 2cx + 2\varepsilon x(\log\log x)^{k-1}\sum_{p_1\le\frac{x}{x_0}}\frac{1}{p_1} \\
&\le 2cx + 4\varepsilon x(\log\log x)^k < 5\varepsilon x(\log\log x)^k,
\end{aligned}$$

where the second last inequality comes from choosing x large enough so that

$$\sum_{p\le x}\frac{1}{p} \le 2\log\log x.$$

Therefore, we see that

$$\Theta_{k+1}(x) - (k+1)xL_k(x) = o\left(x(\log\log x)^k\right).$$

We conclude that

$$\Theta_{k+1}(x) \sim (k+1)x(\log\log x)^k,$$

which proves the claim. ■

From equation (4.1) and the claim, we have

$$\pi_k(x) \le \frac{1}{k!}\Pi_k(x) \sim \frac{1}{(k-1)!}\frac{x}{\log x}(\log\log x)^{k-1}.$$

Moreover, combining equations (4.1) and (4.2) with the claim yields

$$\pi_k(x) = \tau_k(x) + O(\Pi_{k-1}(x)) \ge \frac{1}{k!}\Pi_k(x) + O(\Pi_{k-1}(x)) \sim \frac{1}{(k-1)!}\frac{x}{\log x}(\log\log x)^{k-1}.$$

In particular, we get

$$\pi_k(x) \sim \tau_k(x) \sim \frac{1}{(k-1)!}\frac{x}{\log x}(\log\log x)^{k-1},$$

which finishes the proof of the theorem. □

4.2 Summatory Functions for $\omega(n)$ and $\Omega(n)$

Let's now consider the averages of $\omega(n)$ and $\Omega(n)$.

Theorem 4.4

We have

$$\sum_{n\le x}\omega(n) = x\log\log x + \beta x + o(x),$$
$$\sum_{n\le x}\Omega(n) = x\log\log x + \tilde{\beta} x + o(x),$$

where β is Merten's constant as in Theorem 2.13 and

$$\tilde{\beta} = \beta + \sum_p \frac{1}{p(p-1)}.$$

Proof. Set $S_1 = S_1(x) = \sum_{n\le x}\omega(n)$. Then we have

$$S_1 = \sum_{n\le x}\sum_{p\mid n} 1 = \sum_{p\le x}\left\lfloor \frac{x}{p}\right\rfloor.$$

By Theorem 2.13, we obtain

$$\begin{aligned} S_1 &= \sum_{p\le x}\left\lfloor \frac{x}{p}\right\rfloor \\ &= x\sum_{p\le x}\frac{1}{p} + O(\pi(x)) \\ &= x(\log\log x + \beta + o(1)) + O(\pi(x)) \\ &= x\log\log x + x\beta + o(x), \end{aligned}$$

where the last equality follows from the Prime Number Theorem.

On the other hand, if we set $S_2 = S_2(x) = \sum_{n\le x}\Omega(n)$, then

$$S_2 - S_1 = \sum_{p^m\le x,\, m\ge 2}\left\lfloor \frac{x}{p^m}\right\rfloor = \sum_{p^m\le x,\, m\ge 2}\frac{x}{p^m} + O\left(\sum_{p^m\le x,\, m\ge 2} 1\right).$$

Note that $2^m \le p^m \le x$, so $m \le \frac{\log x}{\log 2}$. Moreover, $p^2 \le p^m \le x$ implies that $p \le x^{1/2}$. Therefore, we have

$$S_2 - S_1 = \sum_{p^m\le x,\, m\ge 2}\frac{x}{p^m} + O(x^{1/2}\log x) = x\left(\sum_p\left(\frac{1}{p^2}+\frac{1}{p^3}+\cdots\right) - \sum_{p^m\ge x}\frac{1}{p^m}\right) + O(x^{1/2}\log x).$$

Observe that

$$\sum_{\substack{p^m>x\\ m\ge x}}\frac{1}{p^m} \le \sum_{\substack{p^m>x\\ m\ge 2\\ 2\mid n}}\frac{1}{p^m} + \sum_{\substack{p^m>x\\ m\ge 2\\ 2\nmid n}}\frac{1}{p^m} \le \sum_{n^2>x}\frac{1}{n^2} + \sum_{\substack{p^m>x\\ m\ge 2\\ 2\mid m\\ p\le\sqrt{x}}}\frac{1}{p^m} + \sum_{\substack{p^m>x\\ m\ge 2\\ 2\mid m\\ p>\sqrt{x}}}\frac{1}{p^m}.$$

Notice that if $p\le\sqrt{x}$, then since $p^m > x$, we get $p^{m-1} > x/p > \sqrt{x}$. On the other hand, if $p > \sqrt{x}$, then $p^{m-1} > \sqrt{x}$. Hence, we get

$$\sum_{\substack{p^m>x\\ m\ge 2}}\frac{1}{p^m} \le \sum_{n^2>x}\frac{1}{n^2} + 2\sum_{\substack{p^{m-1}>\sqrt{x}\\ m\ge 2\\ 2\mid m}}\frac{1}{p^{m-1}} \le \sum_{n^2>x}\frac{1}{n^2} + 2\sum_{m^2>\sqrt{x}}\frac{1}{m^2} \le 3\sum_{k>\sqrt[4]{x}}\frac{1}{k^2} = O\left(\frac{1}{\sqrt[4]{x}}\right).$$

Therefore, we have

$$S_2 - S_1 = x\left(\sum_p \frac{1}{p(p-1)} + o(1)\right) + O(x^{1/2}\log x) = x\sum_p\frac{1}{p(p-1)} + o(x).$$

Together with our estimate of S_1, we see that

$$S_2 = x\log\log x + x\left(\beta + \sum_p\frac{1}{p(p-1)}\right) + o(x).$$

□

4.3 Asymptotic Density and Normal Order

Definition 4.5

Let A be a subset of $\mathbb{N}$. For any $n\in\mathbb{N}$, we set $A(n) = \{1,\ldots,n\}\cap A$. We define the **upper asymptotic density** of A by

$$\bar{d}(A) := \limsup_{n\to\infty}\frac{|A(n)|}{n}.$$

Similarly, we define the **lower asymptotic density** of A to be

$$\underline{d}(A) := \liminf_{n\to\infty}\frac{|A(n)|}{n}.$$

We say that A has **asymptotic density** $d(A)$ when $\bar{d}(A) = \underline{d}(A)$, in which case we set $d(A)$ to be this common value.

Now, let's look at some simple examples of asymptotic density of subsets $A\subseteq\mathbb{N}$.

Example 4.6

(1) When A is the set of all primes, we have $d(A) = \bar{d}(A) = \underline{d}(A) = 0$.

(2) For $A = \{n \in \mathbb{N} : n \equiv 0 \pmod 5\}$, we have $d(A) = \bar{d}(A) = \underline{d}(A) = 1/5$.

(3) For $A = \{n \in \mathbb{N} : n \neq k^2 + 1 \text{ for any } k \in \mathbb{Z}\}$, we have $d(A) = \bar{d}(A) = \underline{d}(A) = 1$.

(4) Let $A = \{a \in \mathbb{N} : (2k)! < a < (2k+1)! \text{ for some } k \in \mathbb{Z}\}$. Notice that for $n = (2k+1)!$, any $a \in \mathbb{N}$ satisfying $(2k)! < a < (2k+1)!$ is included in $A(n)$. Therefore, we have

$$1 \geq \frac{|A((2k+1)!)|}{(2k+1)!} \geq \frac{(2k+1)! - (2k)!}{(2k+1)!} = \frac{2k}{2k+1}.$$

By taking $k \to \infty$, we see that

$$\frac{|A((2k+1)!)|}{(2k+1)!} \to 1,$$

and hence $\bar{d}(A) = 1$. On the other hand, when $n = (2k)!$, then only $a \in \mathbb{N}$ such that $a < (2k-1)!$ are included in $A(n)$. Thus, we have

$$0 \leq \frac{|A((2k)!)|}{(2k)!} \leq \frac{(2k-1)!}{(2k)!} = \frac{1}{2k}.$$

As $k \to \infty$, we have

$$\frac{|A((2k)!)|}{(2k)!} \to 0,$$

and hence $\underline{d}(A) = 0$.

From asymptotic density, we can define normal order. Moreover, we will define average order.

Definition 4.7

Let $f(n)$ and $F(n)$ be functions from $\mathbb{N}$ to $\mathbb{R}$.

- We say that $f(n)$ has **normal order** $F(n)$ if for every $\varepsilon > 0$, the set

$$A(\varepsilon) = \{n \in \mathbb{N} : (1-\varepsilon)F(n) < f(n) < (1+\varepsilon)F(n)\}$$

  has the property that $d(A(\varepsilon)) = 1$. Equivalently, if $B(\varepsilon) = \mathbb{N} \setminus A(\varepsilon)$, then $d(B(\varepsilon)) = 0$.

- We say that $f(n)$ has **average order** $F(n)$ if

$$\sum_{j=1}^{n} f(j) \sim \sum_{j=1}^{n} F(j).$$

These definitions seem rather abstract, so let's look at some examples of normal and average order. It's not too difficult to check the details.

Example 4.8

(1) If we define
$$f(n) = \begin{cases} 1 & \text{if } n \neq k! \text{ for any } k \in \mathbb{N}, \\ n & \text{if } n = k! \text{ for some } k \in \mathbb{N}, \end{cases}$$
then f has normal order 1 but not average order 1.

(2) If we define
$$f(n) = \begin{cases} 2 & \text{if } n \equiv 1 \pmod 2, \\ 0 & \text{if } n \equiv 0 \pmod 2, \end{cases}$$
then f has average order 1 but not normal order 1.

(3) If we define
$$f(n) = \begin{cases} \log n + (\log n)^{1/2} & \text{if } n \equiv 1 \pmod 2, \\ \log n - (\log n)^{1/2} & \text{if } n \equiv 0 \pmod 2, \end{cases}$$
then f has both normal and average order $\log n$.

Theorem 4.9

Both $\omega(n)$ and $\Omega(n)$ have average order $\log \log n$.

Proof. First, note that
$$\begin{aligned} \sum_{n \leq x} \log \log n &= \sum_{x^{1/2} < n \leq x} \log \log n + \sum_{n \leq x^{1/2}} \log \log n \\ &= \sum_{x^{1/2} < n \leq x} \log \log n + O(x^{1/2} \log \log x). \end{aligned}$$

Moreover, we have
$$\sum_{x^{1/2} < n \leq x} \log \log n \leq \log \log x \sum_{x^{1/2} < n \leq x} 1 = x \log \log x + O(x^{1/2} \log \log x).$$

Also, we have the lower bound
$$\sum_{x^{1/2} < n \leq x} \log \log n \geq (\log \log x - \log 2) \sum_{x^{1/2} < n \leq x} 1 = x \log \log x + O(x^{1/2} \log \log x).$$

It follows that
$$\sum_{n \leq x} \log \log n = x \log \log x + O(x^{1/2} \log \log x).$$

Combining this estimate with Theorem 4.4 shows that $\omega(n)$ and $\Omega(n)$ both have average order $\log \log n$. $\square$

4.4 Normal Order of $\omega(n)$ and $\Omega(n)$

We have shown that $\omega(n)$ and $\Omega(n)$ have average order $\log \log n$. In this section, we'll work towards proving that they have normal order $\log \log n$.

THEOREM 4.10

Let $\delta > 0$. The number of positive integers $n \leq x$ satisfying

$$|f(n) - \log\log n| > (\log\log n)^{\frac{1}{2}+\delta}$$

is $o(x)$, where $f(n) = \omega(n)$ or $f(n) = \Omega(n)$. In particular, both $\omega(n)$ and $\Omega(n)$ have normal order $\log\log n$.

PROOF. It is enough to prove that the number of positive integers $n \leq x$ with

$$|f(n) - \log\log x| > (\log\log x)^{\frac{1}{2}+\delta}$$

is $o(x)$, because for $x^{1/e} \leq n \leq x$, we have

$$\log\log x \geq \log\log n \geq \log\left(\frac{\log x}{e}\right) = \log\log x - 1.$$

In other words, we can replace $\log\log n$ in the statement of the theorem with $\log\log x$.

Moreover, we can restrict our attention to the case where $f(n) = \omega(n)$, because by Theorem 4.4, we have

$$\sum_{n\leq x}(\Omega(n) - \omega(n)) = O(x).$$

Thus, the number of integers $n \leq x$ for which $\Omega(n) - \omega(n) > (\log\log n)^{1/2}$ is $o(x)$.

CLAIM. We have

$$\begin{aligned}\sum_{n\leq x}\omega(n)^2 &= x(\log\log x)^2 + O(x\log\log x),\\ \sum_{n\leq x}(\omega(n) - \log\log x)^2 &= O(x\log\log x).\end{aligned}$$

PROOF OF CLAIM. For each $n \leq x$, consider the ordered pairs (p, q) where p and q are distinct prime factors of n. There are $\omega(n)$ choices for p and $\omega(n) - 1$ choices for q, which gives

$$\omega(n)(\omega(n) - 1) = \sum_{\substack{pq\,|\,n\\ p\neq q}} 1 = \sum_{pq\,|\,n} 1 - \sum_{p^2\,|\,n} 1.$$

Therefore, we have

$$\begin{aligned}\sum_{n\leq x}\omega(n)^2 - \sum_{n\leq x}\omega(n) &= \sum_{n\leq x}\omega(n)(\omega(n) - 1)\\ &= \sum_{n\leq x}\left(\sum_{pq\,|\,n} 1 - \sum_{p^2\,|\,n} 1\right)\\ &= \sum_{pq\leq x}\left\lfloor\frac{x}{pq}\right\rfloor - \sum_{p^2\leq x}\left\lfloor\frac{x}{p^2}\right\rfloor.\end{aligned}$$

Observe that

$$\begin{aligned}\sum_{p^2\leq x}\left\lfloor\frac{x}{p^2}\right\rfloor &\leq x\sum_{p^2\leq x}\frac{1}{p^2} = O(x),\\ \sum_{pq\leq x}\left\lfloor\frac{x}{pq}\right\rfloor &= \sum_{pq\leq x}\frac{x}{pq} + O(x),\end{aligned}$$

which implies that
$$\sum_{n\le x}\omega(n)^2-\sum_{n\le x}\omega(n)=\sum_{pq\le x}\frac{x}{pq}+O(x). \tag{4.3}$$
Next, note that
$$\left(\sum_{p\le x^{1/2}}\frac{1}{p}\right)^2-\left(\sum_{p\le x}\frac{1}{p^2}\right)\le\sum_{pq\le x}\frac{1}{pq}\le\left(\sum_{p\le x}\frac{1}{p}\right)^2.$$
Furthermore, Merten's theorem (Theorem 2.13) tells us that
$$\left(\sum_{p\le x}\frac{1}{p}\right)^2=(\log\log x)^2+O(\log\log x),$$
so it follows that
$$\left(\sum_{p\le x^{1/2}}\frac{1}{p}\right)^2=\left(\log\log x^{1/2}+O(1)\right)^2=(\log\log x-\log 2+O(1))^2=(\log\log x)^2+O(\log\log x).$$
Thus, we obtain
$$\sum_{pq\le x}\frac{1}{pq}=(\log\log x)^2+O(\log\log x). \tag{4.4}$$
By Theorem 4.4, we get
$$\sum_{n\le x}\omega(n)=O(x\log\log x). \tag{4.5}$$
Combining equations (4.3), (4.4), and (4.5) together yields
$$\sum_{n\le x}\omega(n)^2=x(\log\log x)^2+O(x\log\log x),$$
which proves the first equality. Now, we have
$$\begin{aligned}\sum_{n\le x}(\omega(n)-\log\log x)^2&=\sum_{n\le x}\omega(n)^2-2\sum_{n\le x}\omega(n)\log\log x+\sum_{n\le x}(\log\log x)^2\\&=x(\log\log x)^2+O(x\log\log x)-2\log\log x\sum_{n\le x}\omega(n)+\lfloor x\rfloor(\log\log x)^2\\&=x(\log\log x)^2+O(x\log\log x)-2x(\log\log x)^2+O(\log\log x)\\&\qquad+x(\log\log x)^2+O((\log\log x)^2)\\&=O(x\log\log x),\end{aligned}$$
where the second last equality follows from Theorem 4.4. This finishes the proof of the claim. ■

Finally, as we stated in the beginning of the proof, it suffices to show that
$$E(x):=\#\{n\le x:|\omega(n)-\log\log x|>(\log\log x)^{\frac{1}{2}+\delta}\}$$
is $o(x)$. By the claim, we have
$$E(x)\cdot(\log\log x)^{1+2\delta}\le\sum_{n\le x}(\omega(n)-\log\log x)^2=O(x\log\log x).$$
It follows that
$$E(x)=O\left(\frac{x\log\log x}{(\log\log x)^{1+2\delta}}\right)=o(x). \qquad\square$$

REMARK 4.11

Since the average order of $\omega(n)$ is $\log\log n$, which is asymptotic to $\log\log x$ for "almost all" n (namely, all except $o(x)$ many $n \le x$), we can view the sum

$$\frac{1}{x}\sum_{n\le x}(\omega(n) - \log\log x)^2$$

as the variance of $\omega(n)$; that is, the squares of the standard deviation. In Homework 3, we will show that

$$\sum_{n\le x}(\omega(n) - \log\log x)^2 \sim x\log\log x,$$

which implies that the standard deviation of $\omega(n)$ is about $\sqrt{\log\log n}$. Now, consider the term

$$\frac{\omega(n) - \log\log n}{\sqrt{\log\log n}}.$$

In 1934, Erdős and Kac proved (without knowing probability theory) that

$$\lim_{x\to\infty}\frac{1}{x}\#\left\{n \le x : \frac{\omega(n) - \log\log n}{\sqrt{\log\log n}} \le \gamma\right\} = G(\gamma),$$

where we define

$$G(\gamma) := \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\gamma} e^{-t^2/2}\,\mathrm{d}t$$

to be the Gaussian normal distribution. This result forms a foundation of probabilistic number theory.

Recall that for all $n \in \mathbb{N}$, the divisor function $d(n)$ gives the number of positive divisors of n. In particular, if we have $n = p_1^{a_1}\cdots p_r^{a_r}$ where $a_1, \ldots, a_r \in \mathbb{N}$ and $p_1, \ldots, p_r$ are distinct primes, then

$$\begin{aligned}\omega(n) &= r,\\ \Omega(n) &= a_1 + \cdots + a_r,\\ d(n) &= (a_1 + 1)\cdots(a_r + 1).\end{aligned}$$

THEOREM 4.12

Given $\varepsilon > 0$, define the set

$$S(\varepsilon) = \{n \in \mathbb{N} : 2^{(1-\varepsilon)\log\log n} < d(n) < 2^{(1+\varepsilon)\log\log n}\}.$$

Then $S(\varepsilon)$ has asymptotic density 1.

PROOF. Note that for any $a \in \mathbb{N}$, we have

$$2 \le a + 1 \le 2^a.$$

In particular, we get

$$2^{\omega(n)} \le d(n) \le 2^{\Omega(n)},$$

and the result follows from Theorem 4.10. $\square$

Remark 4.13

We saw in Theorem 3.9 that

$$\sum_{n\leq x} d(n) \sim x\log x \sim \sum_{n\leq x} \log n.$$

Therefore, the average order of $d(n)$ is $\log n$. However, using Theorem 4.12, one can show that for almost all $n \in \mathbb{N}$, the divisor function $d(n)$ satisfies

$$(\log n)^{\log 2-\varepsilon} < d(n) < (\log n)^{\log 2+\varepsilon}$$

for any $\varepsilon > 0$.

5 Quadratic Reciprocity

5.1 Euler's Totient Function

Definition 5.1

For $n \in \mathbb{N}$, we define **Euler's totient function** $\phi(n)$ to be the number of integers m such that $1 \leq m \leq n$ and $\gcd(m,n) = 1$. That is, we have

$$\phi(n) = \#\{1 \leq m \leq n : \gcd(m,n) = 1\}.$$

A **reduced residue system modulo** n is a subset $R \subseteq \mathbb{Z}$ such that

(i) $\gcd(r,n) = 1$ for each $r \in R$;

(ii) R contains $\phi(n)$ elements; and

(iii) no two elements of R are congruent modulo n.

Theorem 5.2: Euler

Let $a, n \in \mathbb{N}$ with $\gcd(a,n) = 1$. Then

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

Proof. Let $\{c_1, \ldots, c_{\phi(n)}\}$ be a reduced residue system modulo n. Since $\gcd(a,n) = 1$, $\{ac_1, \ldots, ac_{\phi(n)}\}$ is also a reduced residue system modulo n. Hence, we have

$$c_1 \cdots c_{\phi(n)} \equiv ac_1 \cdots ac_{\phi(n)} \pmod{n}.$$

In particular, we see that

$$c_1 \cdots c_{\phi(n)} \equiv a^{\phi(n)} c_1 \cdots c_{\phi(n)} \pmod{n}.$$

Since $c_1, \ldots, c_{\phi(n)}$ are all coprime with n, it follows that

$$a^{\phi(n)} \equiv 1 \pmod{n}.$$ □

Notice that when p is prime, we have $\phi(p) = p - 1$, so we immediately obtain the following corollary.

Corollary 5.3: Fermat's little theorem

Let p be a prime. For any $a \in \mathbb{Z}$ with $p \nmid a$, we have

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

Theorem 5.4: Wilson's theorem

If p is a prime, then $(p-1)! \equiv -1 \pmod{p}$.

PROOF. Consider the element $x^{p-1} - 1 \in (\mathbb{Z}/p\mathbb{Z})[x]$. By Fermat's little theorem and using the fact that $\mathbb{Z}/p\mathbb{Z}$ is a field, this factors as

$$x^{p-1} - 1 \equiv (x-1)(x-2)\cdots(x-(p-1)) \pmod p$$

in $(\mathbb{Z}/p\mathbb{Z})[x]$, as $1, 2, \ldots, p-1$ are all roots. Looking at the constant coefficient, we find that

$$-1 \equiv (-1)(-2)\cdots(-(p-1)) \pmod p.$$

Therefore, we have $-1 \equiv (-1)^{p-1}(p-1)! \pmod p$. When $p = 2$, the result holds since $-1 \equiv 1 \pmod 2$; otherwise, p is odd, so $-1 \equiv (p-1)! \pmod p$ as required. □

5.2 Quadratic Residues

DEFINITION 5.5

Let p be a prime. A nonzero integer a coprime to p that is congruent to a square modulo p is called a **quadratic residue modulo** p (or QR for short). If not, then a is said to be a **quadratic nonresidue modulo** p (or NR for short). Moreover, we define the Legendre symbol $(\frac{a}{p})$ by

$$\left(\frac{a}{p}\right) = \begin{cases} 1 & \text{if } x^2 \equiv a \pmod p \text{ has a solution,} \\ -1 & \text{if } x^2 \equiv a \pmod p \text{ has no solution.} \end{cases}$$

In other words, if $(\frac{a}{p}) = 1$, then a is a quadratic residue; otherwise, it is a quadratic nonresidue.

REMARK 5.6

Let p be an odd prime. Then there are exactly $(p-1)/2$ quadratic residues modulo p, and exactly $(p-1)/2$ quadratic nonresidues modulo p.

PROOF. The quadratic residues modulo p are given by

$$1^2, 2^2, \ldots, \left(\frac{p-1}{2}\right)^2 \pmod p.$$

To see that there are exactly $(p-1)/2$ of them, it suffices to show that these are all different modulo p. Indeed, suppose that $1 \leq b_2 \leq b_1 \leq (p-1)/2$ with $b_1^2 \equiv b_2^2 \pmod p$. Then we have

$$(b_1 - b_2)(b_1 + b_2) \equiv 0 \pmod p,$$

or equivalently, $p \mid (b_1 - b_2)(b_1 + b_2)$. Since p is prime, at least one of $(b_1 - b_2)$ or $(b_1 + b_2)$ must be divisible by p. Note that

$$2 = 1 + 1 \leq b_1 + b_2 \leq \frac{p-1}{2} + \frac{p-1}{2} = p - 1,$$

so $p \nmid (b_1 + b_2)$, and hence $p \mid (b_1 - b_2)$. But we know that $0 \leq b_1 - b_2 \leq (p-1)/2 < p$, so $b_1 = b_2$. □

We now consider the products between quadratic residues and nonresidues modulo p, and derive a nice property about the product of Legendre symbols.

Lemma 5.7

If a_1 and a_2 are quadratic residues modulo p, then so is a_1a_2.

Proof. Suppose that $b_1^2 \equiv a_1 \pmod{p}$ and $b_2^2 \equiv a_2 \pmod{p}$. Then we obtain

$$(b_1b_2)^2 \equiv b_1^2b_2^2 \equiv a_1a_2 \pmod{p}.$$

□

Lemma 5.8

If a_1 is a quadratic residue and a_2 is a quadratic nonresidue modulo p, then a_1a_2 is a quadratic nonresidue modulo p.

Proof. Suppose that $b_1^2 \equiv a_1 \pmod{p}$. Taking the inverse, we have

$$(b_1^{-1})^2 \equiv a_1^{-1} \pmod{p}.$$

Now, if a_1a_2 were a quadratic residue modulo p, then there would exist an integer b such that

$$b^2 \equiv a_1a_2 \pmod{p}.$$

Multiplying these equations together gives

$$(bb_1^{-1})^2 \equiv a_1a_2 \cdot a_1^{-1} \equiv a_2 \pmod{p},$$

contradicting our assumption that a_2 is a quadratic nonresidue modulo p. □

Theorem 5.9: Quadratic residue multiplication rules

Let p be an odd prime. Then

(i) $\mathrm{QR} \times \mathrm{QR} = \mathrm{QR}$;

(ii) $\mathrm{QR} \times \mathrm{NR} = \mathrm{NR}$; and

(iii) $\mathrm{NR} \times \mathrm{NR} = \mathrm{QR}$.

In particular, if p is an odd prime and $a, b \in (\mathbb{Z}/p\mathbb{Z})^*$, then

$$\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right).$$

Proof. We have already proved (i) and (ii) from Lemma 5.7 and Lemma 5.8. Let A be the subset of $(\mathbb{Z}/p\mathbb{Z})^*$ of quadratic residues, and let B be the subset of $(\mathbb{Z}/p\mathbb{Z})^*$ of quadratic nonresidues. We know that $A \cap B = \varnothing$ and $A \cup B = (\mathbb{Z}/p\mathbb{Z})^*$. Moreover, from Remark 5.6, we have $|A| = |B| = (p-1)/2$. Notice that if c is a quadratic nonresidue, then

$$c \cdot (\mathbb{Z}/p\mathbb{Z})^* = (\mathbb{Z}/p\mathbb{Z})^*.$$

Since $\mathrm{QR} \times \mathrm{NR} = \mathrm{NR}$, we see that multiplication by c sends A to B, which is a surjective map by counting numbers. In particular, we find that

$$c \cdot B = c \cdot ((\mathbb{Z}/p\mathbb{Z})^* \setminus A) = (c \cdot (\mathbb{Z}/p\mathbb{Z})^*) \setminus (c \cdot A) = (\mathbb{Z}/p\mathbb{Z})^* \setminus B = A,$$

and hence $\mathrm{NR} \times \mathrm{NR} = \mathrm{QR}$, proving (iii). □

5.3 Special Cases of Quadratic Reciprocity

Theorem 5.10: Euler's criterion

Let p be an odd prime, and let $a \in (\mathbb{Z}/p\mathbb{Z})^*$. Then

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

Proof. If a is a quadratic residue modulo p, say $a \equiv b^2 \pmod{p}$, then Fermat's little theorem tells us that

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

Next, consider the congruence

$$X^{(p-1)/2} - 1 \equiv 0 \pmod{p}.$$

We just showed that every quadratic residue modulo p is a solution to this congruence, and we know that there are exactly $(p-1)/2$ of them by Remark 5.6. Moreover, $X^{(p-1)/2} - 1 \equiv 0 \pmod{p}$ has at most $(p-1)/2$ distinct solutions since $(\mathbb{Z}/p\mathbb{Z})^*$ is a field. Therefore, there is a correspondence between the solutions of $X^{(p-1)/2} - 1 \equiv 0 \pmod{p}$ and the quadratic residues modulo p.

Now, let a be a quadratic nonresidue modulo p. Then Fermat's little theorem tells us that $a^{p-1} \equiv 1 \pmod{p}$, and hence

$$0 \equiv a^{p-1} - 1 \equiv \left(a^{(p-1)/2} - 1\right)\left(a^{(p-1)/2} + 1\right) \pmod{p}.$$

The first factor is not 0 modulo p because the solutions of $X^{(p-1)/2} - 1 \equiv 0 \pmod{p}$ are precisely the quadratic residues modulo p. Therefore, we have $a^{(p-1)/2} + 1 \equiv 0 \pmod{p}$, and so

$$a^{(p-1)/2} \equiv -1 \equiv \left(\frac{a}{p}\right) \pmod{p}. \qquad \square$$

We can use Euler's criterion to compute some Legendre symbols.

Theorem 5.11: Quadratic Reciprocity I

Let p be an odd prime. Then

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

Proof. By Euler's criterion, we have

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

Suppose that $p \equiv 1 \pmod{4}$. Then $p = 4k+1$ for some integer k, and we see that

$$(-1)^{(p-1)/2} = (-1)^{(4k+1-1)/2} = (-1)^{2k} = 1,$$

which gives us

$$\left(\frac{-1}{p}\right) = 1.$$

On the other hand, when $p \equiv 3 \pmod 4$ so that $p = 4k+3$ for some integer k, then

$$(-1)^{(p-1)/2} = (-1)^{(4k+3-1)/2} = (-1)^{2k+1} = -1,$$

and it follows that

$$\left(\frac{-1}{p}\right) = -1. \qquad \square$$

To go further, we first require another tool.

Theorem 5.12: Gauss' lemma

Let p be an odd prime, and let a be an integer coprime to p. Take the numbers $a, 2a, \ldots, [(p-1)/2]a$ and reduce each of them modulo p to get numbers lying between $-(p-1)/2$ and $(p-1)/2$. If s is the number of resulting residues less than 0, then

$$\left(\frac{a}{p}\right) = (-1)^s.$$

Proof. For each $1 \le i \le (p-1)/2$, let u_i be an integer such that $ia \equiv u_i \pmod p$ and $-(p-1)/2 \le u_i \le (p-1)/2$. Note that s is the number of elements in $u_1, \ldots, u_{(p-1)/2}$ less than 0.

We claim that

$$\{|u_1|, |u_2|, \ldots, |u_{(p-1)/2}|\} = \{1, 2, \ldots, p-1\}.$$

It is sufficient to show that no two of the integers in the first set are congruent modulo p, as there are exactly $(p-1)/2$ elements in the set, and they are all positive integers not exceeding $(p-1)/2$. Suppose that $|u_i| = |u_j|$. If $u_i = u_j$, then $ia \equiv ja \pmod p$, which implies that $i = j$ since $\gcd(a,p) = 1$ and $1 \le i, j \le (p-1)/2$. On the other hand, if $u_i = -u_j$, then $ia \equiv -ja \pmod p$. This implies that $(i+j)a \equiv 0 \pmod p$ and hence

$$i + j \equiv 0 \pmod p$$

since $\gcd(a,p) = 1$. Since $1 \le i, j \le (p-1)/2$, we have $2 \le i + j \le p - 1$. But there is no number congruent to 0 modulo p in this range, so this scenario is impossible. This proves our claim.

Finally, we find that

$$a^{(p-1)/2} \prod_{i=1}^{(p-1)/2} i \equiv \prod_{i=1}^{(p-1)/2} ia \equiv \prod_{i=1}^{(p-1)/2} u_i \equiv (-1)^s \prod_{i=1}^{(p-1)/2} i \pmod p,$$

where the third congruence follows from the claim. This implies that

$$\left(\frac{a}{p}\right) \equiv a^{(p-1)/2} \equiv (-1)^s \pmod p. \qquad \square$$

Now, we can compute $(\frac{2}{p})$.

Theorem 5.13: Quadratic Reciprocity II

Let p be an odd prime. Then

$$\left(\frac{2}{p}\right) = (-1)^{(p^2-1)/8} = \begin{cases} 1 & \text{if } p \equiv \pm 1 \pmod 8, \\ -1 & \text{if } p \equiv \pm 3 \pmod 8. \end{cases}$$

Proof. From Gauss' lemma (Theorem 5.12), we only need to find the amount of numbers s from the list

$$1\cdot 2,\ 2\cdot 2,\ 3\cdot 2, \ldots, \left(\frac{p-1}{2}\right)\cdot 2$$

which are greater than $p/2$. Note that for $1 \leq j \leq (p-1)/2$, the integer $2j$ is less than $p/2$ when $j \leq p/4$. Hence, there are $\lfloor p/4 \rfloor$ integers in the set less than $p/2$. Consequently, there are

$$s = \frac{p-1}{2} - \left\lfloor \frac{p}{4} \right\rfloor$$

of them greater than $p/2$. In particular, we find that

$$\left(\frac{2}{p}\right) = (-1)^{\frac{p-1}{2} - \lfloor p/4 \rfloor}.$$

To finish the proof, it remains to show that

$$\frac{p-1}{2} - \left\lfloor \frac{p}{4} \right\rfloor \equiv \frac{p^2-1}{8} \pmod{2}.$$

We can verify this case by case.

- If $p \equiv \pm 1 \pmod 8$, then $p = 8k \pm 1$ for some integer k. We have

  $$\frac{p-1}{2} - \left\lfloor \frac{p}{4} \right\rfloor = \frac{(8k\pm 1)-1}{2} - \left\lfloor \frac{8k\pm 1}{4} \right\rfloor = \frac{-1\pm 1}{2} + 4k + \left\lfloor 2k \pm \frac{1}{4} \right\rfloor.$$

  If we choose the $+$ sign, we obtain

  $$\frac{p-1}{2} - \left\lfloor \frac{p}{4} \right\rfloor \equiv \frac{-1+1}{2} + \left\lfloor 2k + \frac{1}{4} \right\rfloor \equiv 0 + 2k \equiv 0 \pmod{2},$$

  whereas if we choose the $-$ sign, we get

  $$\frac{p-1}{2} - \left\lfloor \frac{p}{4} \right\rfloor \equiv \frac{-1-1}{2} + \left\lfloor 2k - \frac{1}{4} \right\rfloor \equiv -1 + 2k - 1 \equiv 0 \pmod{2}.$$

  On the other hand, we have

  $$\frac{p^2-1}{8} \equiv \frac{(8k\pm 1)^2 - 1}{8} \equiv \frac{64k^2 \pm 16k + 1 - 1}{8} \equiv 8k^2 \pm 2k \equiv 0 \pmod{2},$$

  which proves this case.

- If $p \equiv \pm 3 \pmod 8$, then $p = 8k \pm 3$ for some integer k. Then, we see that

  $$\frac{p-1}{2} - \left\lfloor \frac{p}{4} \right\rfloor = \frac{(8k\pm 3)-1}{2} - \left\lfloor \frac{8k\pm 3}{4} \right\rfloor = \frac{-1\pm 3}{2} + 4k + \left\lfloor 2k \pm \frac{3}{4} \right\rfloor.$$

  Choosing $+$ gives

  $$\frac{p-1}{2} - \left\lfloor \frac{p}{4} \right\rfloor \equiv \frac{-1+3}{2} + \left\lfloor 2k + \frac{3}{4} \right\rfloor \equiv -1 + 2k \equiv 1 \pmod{2},$$

  while choosing $-$ gives

  $$\frac{p-1}{2} - \left\lfloor \frac{p}{4} \right\rfloor \equiv \frac{-1-3}{2} + \left\lfloor 2k - \frac{3}{4} \right\rfloor \equiv -2 + 2k - 1 \equiv 1 \pmod{2}.$$

  Moreover, we have

  $$\frac{p^2-1}{8} \equiv \frac{(8k\pm 3)^2 - 1}{8} \equiv \frac{64k^2 \pm 48k + 9 - 1}{8} \equiv 8k^2 \pm 6k + 1 \equiv 1 \pmod{2},$$

  which completes the proof of the theorem. $\square$

5.4 The Law of Quadratic Reciprocity

Theorem 5.14: Law of Quadratic Reciprocity

Let p and q be odd primes. Then

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

In particular, we have

$$\left(\frac{p}{q}\right) = \begin{cases} \left(\frac{q}{p}\right) & \text{if } p \equiv 1 \pmod 4 \text{ or } q \equiv 1 \pmod 4, \\ -\left(\frac{q}{p}\right) & \text{if } p \equiv q \equiv 3 \pmod 4. \end{cases}$$

There are many ways to prove this famous theorem. We will take an approach that is not the fastest, but is easy to understand. First, we require the following lemma.

Lemma 5.15

Let p be an odd prime, and let a be an odd integer such that $a \nmid p$. Then we have

$$\left(\frac{a}{p}\right) = (-1)^{T(a,p)},$$

where we define $T(a,p)$ to be

$$T(a,p) = \sum_{j=1}^{(p-1)/2} \left\lfloor \frac{ja}{p} \right\rfloor.$$

Proof. Consider the reduced residues of $a, 2a, \ldots, [(p-1)/2]a$ lying between $-(p-1)/2$ and $(p-1)/2$; let $u_1, \ldots, u_s$ be those less than 0, and $v_1, \ldots, v_t$ be those greater than 0. The division algorithm tells us that

$$ja = p\left\lfloor \frac{ja}{p} \right\rfloor + r,$$

where the remainder r is either in the form $p + u_j$ or v_j. By adding these $(p-1)/2$ equations, we obtain

$$\sum_{j=1}^{(p-1)/2} ja = \sum_{j=1}^{(p-1)/2} p\left\lfloor \frac{ja}{p} \right\rfloor + \sum_{j=1}^{s}(p + u_j) + \sum_{j=1}^{t} v_j. \tag{5.1}$$

In the proof of Gauss' lemma (Theorem 5.12), we saw that the integers

$$-u_1, -u_2, \ldots, -u_s, v_1, \ldots, v_t$$

are precisely the integers $1, 2, \ldots, (p-1)/2$. Therefore, we have

$$\sum_{j=1}^{(p-1)/2} j = -\sum_{j=1}^{s} u_j + \sum_{j=1}^{t} v_j. \tag{5.2}$$

Subtracting (5.2) from (5.1), we find that

$$\sum_{j=1}^{(p-1)/2} ja - \sum_{j=1}^{(p-1)/2} j = \sum_{j=1}^{(p-1)/2} p\left\lfloor \frac{ja}{p} \right\rfloor + ps + 2\sum_{j=1}^{s} u_j.$$

Using the definition of $T(a,p)$, we get

$$(a-1)\sum_{j=1}^{(p-1)/2} j = pT(a,p) + ps + 2\sum_{j=1}^{s} u_j. \tag{5.3}$$

Reducing equation (5.3) modulo 2 yields $0 \equiv T(a,p) + s \pmod 2$ since a and p are assumed to be odd, and hence $T(a,p) \equiv s \pmod 2$ since s and $-s$ have the same parity. We conclude that

$$\left(\frac{a}{p}\right) = (-1)^s = (-1)^{T(a,p)}.$$

□

Now, we are ready to prove our main theorem.

Proof of Theorem 5.14. Consider the pairs of integers (x,y) with $1 \le x \le (p-1)/2$ and $1 \le y \le (q-1)/2$. There are $(p-1)/2 \cdot (q-1)/2$ such pairs. Note that none of these pairs satisfy $qx = py$ since this would imply that $p \mid x$ and $y \mid q$, which is absurd. We divide these $(p-1)/2 \cdot (q-1)/2$ pairs into two groups, depending on the relative sizes of qx and py.

The pairs of integers (x,y) satisfying $qx > py$ are precisely those with $1 \le x \le (p-1)/2$ and $1 \le y \le qx/p$. Hence, for fixed $1 \le x \le (p-1)/2$, there are $\lfloor qx/p \rfloor$ possible values of y, so the number of pairs (x,y) satisfying $qx > py$ is

$$\sum_{j=1}^{(p-1)/2} \left\lfloor \frac{qj}{p} \right\rfloor.$$

Similarly, the pairs of integers (x,y) satisfying $qx < py$ are precisely those with $1 \le y \le (q-1)/2$ and $1 \le x \le py/q$. For fixed $1 \le y \le (q-1)/2$, there are $\lfloor py/q \rfloor$ possible values of x, so the number of pairs (x,y) satisfying $qx < py$ is

$$\sum_{j=1}^{(q-1)/2} \left\lfloor \frac{pj}{q} \right\rfloor.$$

Consequently, we find that

$$\frac{p-1}{2}\frac{q-1}{2} = \sum_{j=1}^{(p-1)/2} \left\lfloor \frac{qj}{p} \right\rfloor + \sum_{j=1}^{(q-1)/2} \left\lfloor \frac{pj}{q} \right\rfloor = T(q,p) + T(p,q).$$

It follows from Lemma 5.15 that

$$(-1)^{\frac{p-1}{2}\frac{q-1}{2}} = (-1)^{T(q,p)+T(p,q)} = (-1)^{T(q,p)} \cdot (-1)^{T(p,q)} = \left(\frac{q}{p}\right)\left(\frac{p}{q}\right),$$

which completes the proof. □

5.5 The Jacobi Symbol

We now give a generalization of the Legendre symbol to all odd positive integers instead of only odd primes, called the Jacobi symbol. The Jacobi symbol also shares many of the properties of the Legendre symbol, and we will leave the proof of these as homework.

Definition 5.16

Let $a \in \mathbb{Z}$, and let $n \in \mathbb{N}$ be odd. Let $n = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_k^{\alpha_k}$ be the prime factorization of n. We define the **Jacobi symbol** $(\frac{a}{n})$ to be

$$\left(\frac{a}{n}\right) := \left(\frac{a}{p_1}\right)^{\alpha_1}\left(\frac{a}{p_2}\right)^{\alpha_2} \cdots \left(\frac{a}{p_k}\right)^{\alpha_k}.$$

Theorem 5.17: Generalized Law of Quadratic Reciprocity

Let $a, b \in \mathbb{N}$ be odd with $\gcd(a,b) = 1$. Then we have

(1) $\left(\dfrac{-1}{b}\right) = \begin{cases} 1 & \text{if } b \equiv 1 \pmod 4, \\ -1 & \text{if } b \equiv 3 \pmod 4; \end{cases}$

(2) $\left(\dfrac{2}{b}\right) = \begin{cases} 1 & \text{if } b \equiv \pm 1 \pmod 8, \\ -1 & \text{if } b \equiv \pm 3 \pmod 8; \end{cases}$

(3) $\left(\frac{a}{b}\right) = \begin{cases} \left(\dfrac{b}{a}\right) & \text{if } a \equiv 1 \pmod 4 \text{ or } b \equiv 1 \pmod 4, \\ -\left(\dfrac{b}{a}\right) & \text{if } a \equiv b \equiv 3 \pmod 4. \end{cases}$

Proof. This is Question 1 of Homework 4. $\square$

To end off the chapter, we give some examples of computing Legendre symbols, as well as some applications.

Example 5.18

Consider the Legendre symbol $(\frac{13}{17})$. By the Law of Quadratic Reciprocity (Theorem 5.14), we have

$$\left(\frac{13}{17}\right) = \left(\frac{17}{13}\right)(-1)^{(17-1)(13-1)/4} = \left(\frac{17}{13}\right) = \left(\frac{4}{13}\right) = 1.$$

Example 5.19

Let's compute $(\frac{713}{1009})$. Note that 1009 is prime and $713 = 23 \cdot 31$, so Theorem 5.9 tells us that

$$\left(\frac{713}{1009}\right) = \left(\frac{23}{1009}\right)\left(\frac{31}{1009}\right).$$

By the Law of Quadratic Reciprocity (Theorem 5.14), we find that $(\frac{23}{1009}) = (\frac{1009}{23}) = (\frac{20}{23})$ and $(\frac{31}{1009}) = (\frac{1009}{31}) = (\frac{17}{31})$. Then, we see that

$$\left(\frac{20}{23}\right) = \left(\frac{4}{23}\right)\left(\frac{5}{23}\right) = \left(\frac{5}{23}\right) = \left(\frac{23}{5}\right) = \left(\frac{3}{5}\right) = -1.$$

On the other hand, we have

$$\left(\frac{17}{31}\right) = \left(\frac{31}{17}\right) = \left(\frac{14}{17}\right) = \left(\frac{2}{17}\right)\left(\frac{7}{17}\right) = \left(\frac{7}{17}\right) = \left(\frac{17}{7}\right) = \left(\frac{3}{7}\right) = -1,$$

where the third last equality above comes from Quadratic Reciprocity II (Theorem 5.13). Putting these together, we obtain

$$\left(\frac{713}{1009}\right) = \left(\frac{23}{1009}\right)\left(\frac{31}{1009}\right) = (-1)(-1) = 1.$$

Example 5.20

Note that 5 is a quadratic residue for all primes p of the form $10k \pm 1$, whereas it is a quadratic nonresidue for all primes p of the form $10k \pm 3$. This is because we have

$$\left(\frac{5}{p}\right) = \left(\frac{p}{5}\right)$$

by the Law of Quadratic Reciprocity (Theorem 5.14); we see that ± 1 are quadratic residues modulo 5, while ± 3 are not.

Example 5.21

The equation $x^4 - 17y^4 = 2\omega^2$ has no solution in the integers.

Suppose towards a contradiction that an integer solution exists. We may assume without loss of generality that x is coprime with y, and hence x is coprime with ω. If p is an odd prime that divides ω, then

$$\left(\frac{17}{p}\right) = 1$$

since $x^4 \equiv 17y^4 \pmod{p}$. Alternatively, by the Law of Quadratic Reciprocity (Theorem 5.14), we obtain

$$\left(\frac{17}{p}\right) = \left(\frac{p}{17}\right)(-1)^{(17-1)(p-1)/4} = 1,$$

so p is a quadratic residue of 17. Furthermore, we have $(\frac{2}{17}) = (-1)^{(17^2-1)/8} = 1$. In particular, every prime factor p of ω is a quadratic residue modulo 17, so $\omega \equiv t^2 \pmod{17}$ for some integer t. Then $x^4 \equiv 2t^4 \pmod{17}$, and it follows that there exists an integer r such that $r^4 \equiv 2 \pmod{17}$. But the order of 2 modulo 17 is 8, so no such r exists. This is a contradiction, so the equation has no solution in the integers.

Example 5.22

We can check if the congruence $3x^2 - 7x - 42 \equiv 0 \pmod{391}$ has a solution.

Note that $391 = 17 \cdot 23$. Multiplying the above equation by 12, we find that

$$36x^2 - 84x - 516 \equiv 0 \pmod{391}.$$

By completing the square, we see that

$$(6x-7)^2 \equiv 516 + 49 \equiv 565 \pmod{391}.$$

This is now equivalent to solving $x^2 \equiv 174 \pmod{391}$. Notice that $x^2 \equiv 174 \equiv 4 \pmod{17}$, which has the solution $x = 2$. On the other hand, we have $x^2 \equiv 174 \equiv 13 \pmod{23}$. We compute

$$\left(\frac{13}{23}\right) = \left(\frac{23}{13}\right) = \left(\frac{10}{13}\right) = \left(\frac{2}{13}\right)\left(\frac{5}{13}\right) = -\left(\frac{5}{13}\right) = -\left(\frac{13}{5}\right) = -\left(\frac{3}{5}\right) = 1.$$

Then $x^2 \equiv 13 \pmod{23}$ has a solution, so it follows from the Chinese remainder theorem that the original congruence has a solution.

6 Primitive Roots

6.1 Cyclicity of $(\mathbb{Z}/p\mathbb{Z})^*$

Recall that for $a, b \in \mathbb{Z}$, we can find $x, y \in \mathbb{Z}$ such that $ax + by = \gcd(a, b)$ using the Euclidean algorithm.

Theorem 6.1: Chinese remainder theorem

Let $m_1, \dots, m_t \in \mathbb{N}$ with $\gcd(m_i, m_j) = 1$ whenever $i \neq j$, and set $m = m_1 \cdots m_t$. Let $b_1, \dots, b_t \in \mathbb{Z}$. Then the simultaneous congruences

$$\begin{aligned} x &\equiv b_1 \pmod{m_1}, \\ x &\equiv b_2 \pmod{m_2}, \\ &\vdots \\ x &\equiv b_t \pmod{m_t} \end{aligned}$$

has a unique solution modulo m.

Theorem 6.2

Let $m_1, \dots, m_t \in \mathbb{N}$ with $\gcd(m_i, m_j) = 1$ whenever $i \neq j$, and set $m = m_1 \cdots m_t$. Then we have the ring isomorphism

$$\mathbb{Z}/m\mathbb{Z} \cong \mathbb{Z}/m_1\mathbb{Z} \times \cdots \times \mathbb{Z}/m_t\mathbb{Z},$$

as well as the group isomorphism

$$(\mathbb{Z}/m\mathbb{Z})^* \cong (\mathbb{Z}/m_1\mathbb{Z})^* \times \cdots \times (\mathbb{Z}/m_t\mathbb{Z})^*.$$

Proof. Let $\psi : \mathbb{Z} \to \mathbb{Z}/m_1\mathbb{Z} \times \cdots \times \mathbb{Z}/m_t\mathbb{Z}$ be defined by

$$\psi(n) = (n + m_1\mathbb{Z}, \dots, n + m_t\mathbb{Z}).$$

It is readily checked that ψ is a ring homomorphism. By the Chinese remainder theorem, ψ is surjective and $\ker \psi = m\mathbb{Z}$. It follows from the first isomorphism theorem that

$$\mathbb{Z}/m\mathbb{Z} \cong \mathbb{Z}/m_1\mathbb{Z} \times \cdots \times \mathbb{Z}/m_t\mathbb{Z}.$$

On the other hand, if we define $\lambda : (\mathbb{Z}/m\mathbb{Z})^* \to (\mathbb{Z}/m_1\mathbb{Z})^* \times \cdots \times (\mathbb{Z}/m_t\mathbb{Z})^*$ by

$$\lambda(n + m\mathbb{Z}) = (n + m_1\mathbb{Z}, \dots, n + m_t\mathbb{Z}),$$

then λ is a group homomorphism, and it is bijective by the Chinese remainder theorem. $\square$

Corollary 6.3

Let $m_1, \dots, m_t$ be pairwise coprime positive integers. Set $m = m_1 \cdots m_t$. Then

$$\phi(m) = \phi(m_1) \cdots \phi(m_t).$$

Proof. Recall that $\phi(m) = |(\mathbb{Z}/m\mathbb{Z})^*|$, and

$$\phi(m_1) \cdots \phi(m_t) = |(\mathbb{Z}/m_1\mathbb{Z})^*| \cdots |(\mathbb{Z}/m_t\mathbb{Z})^*| = |(\mathbb{Z}/m_1\mathbb{Z})^* \times \cdots \times (\mathbb{Z}/m_t\mathbb{Z})^*|.$$

The result follows from Theorem 6.2. $\square$

Corollary 6.4

Let $m = p_1^{a_1} \cdots p_t^{a_t}$, where $p_1, \ldots, p_t$ are distinct primes and $a_1, \ldots, a_t$ are positive integers. Then

$$\phi(m) = m \cdot \prod_{i=1}^{t} \left(1 - \frac{1}{p_i}\right).$$

Proof. Select $m_i = p_i^{a_i}$ for $i = 1, \ldots, t$ in Corollary 6.3. Observe that

$$\phi(p_i^{a_i}) = p_i^{a_i} - p_i^{a_i - 1} = p_i^{a_i} \left(1 - \frac{1}{p_i}\right).$$

It follows that

$$\phi(m) = \phi(p_1^{a_1}) \cdots \phi(p_t^{a_t}) = p_1^{a_1} \cdots p_t^{a_t} \left(1 - \frac{1}{p_1}\right) \cdots \left(1 - \frac{1}{p_t}\right) = m \cdot \prod_{i=1}^{t} \left(1 - \frac{1}{p_i}\right). \qquad \square$$

Proposition 6.5

Let p be prime. If $d \mid (p-1)$, then $x^d \equiv 1 \pmod{p}$ has exactly d solutions modulo p.

Proof. Write $p - 1 = dk$ for some integer k. Then we have

$$\frac{x^{p-1} - 1}{x^d - 1} = \frac{(x^d)^k - 1}{x^d - 1} = (x^d)^{k-1} + \cdots + x^d + 1 = g(x) \in (\mathbb{Z}/p\mathbb{Z})[x].$$

By Fermat's little theorem (Corollary 5.3), $x^{p-1} - 1$ has $p - 1$ distinct roots in $\mathbb{Z}/p\mathbb{Z}$. Since $\mathbb{Z}/p\mathbb{Z}$ is a field, any polynomial of degree n in $(\mathbb{Z}/p\mathbb{Z})[x]$ has at most n roots. In particular, $(x^d - 1)g(x)$ factors into linear polynomials in $(\mathbb{Z}/p\mathbb{Z})[x]$, and the result follows. $\square$

Theorem 6.6

If p is a prime, then $(\mathbb{Z}/p\mathbb{Z})^*$ is a cyclic group.

Proof. For each divisor d of $p - 1$, let $\lambda(d)$ denote the number of elements in $(\mathbb{Z}/p\mathbb{Z})^*$ of order d. By Proposition 6.5, there are exactly d elements of $(\mathbb{Z}/p\mathbb{Z})^*$ whose order divides d, so we obtain

$$d = \sum_{c \mid d} \lambda(c).$$

By the Möbius inversion formula (Proposition 2.5), we have

$$\lambda(d) = \sum_{c \mid d} \mu(c) \frac{d}{c} = d \cdot \sum_{c \mid d} \frac{\mu(c)}{c} = d \cdot \prod_{p \mid d} \left(1 - \frac{1}{p}\right) = \phi(d),$$

where the final equality follows from Corollary 6.4. Hence, there are $\phi(p-1)$ elements of $(\mathbb{Z}/p\mathbb{Z})^*$ of order $p - 1$, so $(\mathbb{Z}/p\mathbb{Z})^*$ is cyclic. $\square$

6.2 Primitive Roots: The Prime Power Case

DEFINITION 6.7

Let $n \in \mathbb{Z}^+$, and let $a \in \mathbb{Z}$. We say that a is a **primitive root** modulo n if $a + n\mathbb{Z}$ generates $(\mathbb{Z}/n\mathbb{Z})^*$.

REMARK 6.8

(1) For any prime p, we saw that $(\mathbb{Z}/p\mathbb{Z})^*$ is cyclic by Theorem 6.6, so there exists a primitive root modulo p; in fact, there are $\phi(p-1)$ of them.

Artin conjectured that if a is a positive integer that is not a perfect square, then a is a primitive root modulo p for infinitely many primes p. This conjecture is still open, but it can be deduced from the generalized Riemann hypothesis by the work of Hooley.

(2) Why do we require that a is not a perfect square? Note that if p is an odd prime, then $p-1$ is even. We want a to have order $p-1$. Assume that $a = k^2$ for some integer k, and that a is a primitive root modulo p. Then there exists an integer i such that $a^i \equiv k \pmod{p}$. We see that $a^{2i} \equiv a \pmod{p}$, and hence $a^{2i-1} \equiv 1 \pmod{p}$. Since the order of a is $p-1$, we have $(p-1) \mid (2i-1)$. But $p-1$ is even and $2i-1$ is odd, which is a contradiction.

(3) Observe that 2 is a primitive root modulo 5, but 2 is not a primitive root modulo 7 since $2^3 \equiv 1 \pmod{7}$. In general $(\mathbb{Z}/n\mathbb{Z})^*$ is not cyclic, as primitive roots might not exist modulo n. For example, $(\mathbb{Z}/8\mathbb{Z})^* = \{[1],[3],[5],[7]\}$ with $1^2 \equiv 3^2 \equiv 5^2 \equiv 7^2 \equiv 1 \pmod{8}$, so this group is not cyclic.

PROPOSITION 6.9

Let p be prime, and let ℓ be a positive integer. If $a \equiv b \pmod{p^\ell}$, then

$$a^p \equiv b^p \pmod{p^{\ell+1}}.$$

PROOF. Write $a = b + cp^\ell$ for some $c \in \mathbb{Z}$. Then we have

$$a^p = (b+cp^\ell)^p = b^p + \binom{p}{1} b^{p-1} cp^\ell + \binom{p}{2} b^{p-2}(cp^\ell)^2 + \cdots + \binom{p}{p}(cp^\ell)^p,$$

so $a^p \equiv b^p \pmod{p^{\ell+1}}$, since $2\ell \geq \ell + 1$. $\square$

PROPOSITION 6.10

If $\ell \geq 2$ is an integer and p is an odd prime, then for any $a \in \mathbb{Z}$, we have

$$(1+ap)^{p^{\ell-2}} \equiv 1 + ap^{\ell-1} \pmod{p^\ell}.$$

PROOF. We proceed by induction on ℓ. The result is clear for $\ell = 2$. Suppose the result holds for some integer $\ell \geq 2$. We prove it for $\ell + 1$.

By Proposition 6.9 and our inductive hypothesis, we have

$$(1+ap)^{p^{\ell-1}} \equiv (1+ap^{\ell-1})^p \equiv 1 + \binom{p}{1} ap^{\ell-1} + \binom{p}{2}(ap^{\ell-1})^2 + \cdots + \binom{p}{p}(ap^{\ell-1})^p \pmod{p^{\ell+1}}.$$

Since $\ell \geq 2$ implies $2(\ell-1)+1 \leq 3(\ell-1) \leq k(\ell-1)$, we see that $p^{2(\ell-1)+1}$ divides $(ap^{\ell-1})^k$ for $k = 3, \ldots, p$. Furthermore, $p^{2(\ell-1)+1}$ divides $\binom{p}{2}(ap^{\ell-1})^2$ since

$$\binom{p}{2}(ap^{\ell-1})^2 = \frac{p(p-1)}{2}(ap^{\ell-1})^2 = \frac{p-1}{2}a^2p^{2\ell-1}.$$

Note that $(p-1)/2$ is an integer since p is odd. Hence, $p^{2(\ell-1)+1}$ divides the sum

$$\binom{p}{2}(ap^{\ell-1})^2 + \cdots + \binom{p}{p}(ap^{\ell-1})^p \pmod{p^{\ell+1}}.$$

Now, since $\ell \geq 2$ implies $2(\ell-1)+1 \geq \ell+1$ and p is odd, we have

$$1 + \binom{p}{1}ap^{\ell-1} + \binom{p}{2}(ap^{\ell-1})^2 + \cdots + \binom{p}{p}(ap^{\ell-1})^p \equiv 1 + \binom{p}{1}ap^{\ell-1} \equiv 1 + ap^{\ell} \pmod{p^{\ell+1}}.$$

The result holds for all integers $\ell \geq 2$ by induction. □

Proposition 6.11

Let p be an odd prime, ℓ be a positive integer, and a be an integer coprime with p. Then $1 + ap$ has order $p^{\ell-1}$ in $(\mathbb{Z}/p^{\ell}\mathbb{Z})^*$.

Proof. By Proposition 6.10, we have

$$(1+ap)^{p^{\ell-2}} \equiv 1 + ap^{\ell-1} \pmod{p^{\ell}}.$$

Since a is coprime with p, we see that

$$(1+ap)^{p^{\ell-2}} \not\equiv 1 \pmod{p^{\ell}}.$$

Applying Proposition 6.10 again, we have

$$(1+ap)^{p^{\ell-1}} \equiv 1 + ap^{\ell} \pmod{p^{\ell+1}},$$

which implies that

$$(1+ap)^{p^{\ell-1}} \equiv 1 \pmod{p^{\ell}}.$$

Thus, $1 + ap$ has order $p^{\ell-1}$ in $(\mathbb{Z}/p^{\ell}\mathbb{Z})^*$. □

Theorem 6.12

Let p be an odd prime, and ℓ be a positive integer. Then $(\mathbb{Z}/p^{\ell}\mathbb{Z})^*$ is cyclic group.

Proof. Since $(\mathbb{Z}/p\mathbb{Z})^*$ is cyclic by Theorem 6.6, there is a primitive root g modulo p. If $g^{p-1} \equiv 1 \pmod{p^2}$, then

$$(g+p)^{p-1} \equiv g^{p-1} + \binom{p-1}{1}g^{p-2}p + \binom{p-1}{2}g^{p-3}p^2 + \cdots + \binom{p}{p}p^{p-1} \equiv 1 + \binom{p-1}{1}g^{p-2}p \pmod{p^2},$$

so $(g+p)^{p-1} \not\equiv 1 \pmod{p^2}$. In particular, at least one of g^{p-1} and $(g+p)^{p-1}$ is not congruent to 1 modulo p^2. Without loss of generality, we assume that $g^{p-1} \not\equiv 1 \pmod{p^2}$. We claim that g is a primitive root modulo p^{ℓ}, and it will follow that $(\mathbb{Z}/p^{\ell}\mathbb{Z})^*$ is cyclic.

Suppose that g has order m in $(\mathbb{Z}/p^\ell\mathbb{Z})^*$. By Euler's theorem (Theorem 5.2), we have

$$g^{\phi(p^\ell)} \equiv 1 \pmod{p^\ell},$$

and hence $m \mid (p^\ell - p^{\ell-1}) = (p-1)p^{\ell-1}$. Write $m = dp^s$ where $d \mid (p-1)$ and $0 \le s \le \ell - 1$. By Fermat's little theorem (Corollary 5.3), we have $g^p \equiv g \pmod p$. Provided that $s \neq 0$, we have

$$g^{p^s} \equiv g \pmod p.$$

However, we have $g^m \equiv 1 \pmod{p^\ell}$, so $g^m \equiv 1 \pmod p$, which implies that $g^d \equiv 1 \pmod p$ as well. Since g is a primitive root modulo p, we see that $(p-1) \mid d$. Then $d = p-1$, so $m = (p-1)p^s$. Since $g^{p-1} \not\equiv 1 \pmod{p^2}$ and $g^{p-1} \equiv 1 \pmod p$, there exists an integer a coprime with p such that $g^{p-1} \equiv 1 + ap \pmod{p^2}$. By Proposition 6.11, $1 + ap$ has order $p^{\ell-1}$ in $(\mathbb{Z}/p\mathbb{Z})^*$. Then g has order $(p-1)p^\ell$ in $(\mathbb{Z}/p^\ell\mathbb{Z})^*$, so g is a primitive root of p^ℓ. $\square$

6.3 Primitive Roots: The General Case

Theorem 6.13

If $\ell = 1, 2$, then $(\mathbb{Z}/2^\ell\mathbb{Z})^*$ is cyclic. For $\ell \ge 3$, we have the group isomorphism

$$(\mathbb{Z}/2^\ell\mathbb{Z})^* \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2^{\ell-2}\mathbb{Z}.$$

In particular, we can write

$$(\mathbb{Z}/2^\ell\mathbb{Z})^* = \{(-1)^a 5^b + 2^\ell\mathbb{Z} : a \in \{0,1\},\, b \in \{0, \ldots, 2^{\ell-2} - 1\}\}.$$

Proof. It is clear that $(\mathbb{Z}/2\mathbb{Z})^*$ and $(\mathbb{Z}/4\mathbb{Z})^*$ are cyclic. Suppose that $\ell \ge 3$. We claim that

$$5^{2^{\ell-3}} \equiv 1 + 2^{\ell-1} \pmod{2^\ell}. \tag{6.1}$$

We proceed by induction on ℓ. For $\ell = 3$, we have $5 \equiv 1 + 2^2 \pmod{2^3}$. Assume that (6.1) holds for some $\ell \ge 3$. Note that $(1 + 2^{\ell-1})^2 = 1 + 2^\ell + 2^{2(\ell-1)}$ and $2(\ell-1) \ge \ell + 1$ for $\ell \ge 3$. By the induction hypothesis, we know that $5^{2^{\ell-3}} = 1 + 2^{\ell-1} + k2^\ell$ for some $k \in \mathbb{Z}$. It follows that

$$\begin{aligned} 5^{2^{\ell-2}} &= (1 + 2^{\ell-1} + k2^\ell)^2 \\ &= 1 + (2^{\ell-1})^2 + (k2^\ell)^2 + 2 \cdot 2^{\ell-1} + 2 \cdot k2^\ell + 2 \cdot 2^{\ell-1} \cdot k2^\ell \\ &= 1 + 2^\ell + k2^{\ell+1} + 2^{2\ell-2} + k2^{2\ell} + k^2 2^{2\ell}. \end{aligned}$$

Since $2\ell - 2 \ge \ell + 1$ for $\ell \ge 3$, we see that

$$5^{2^{\ell-2}} \equiv 1 + 2^\ell \pmod{2^{\ell+1}},$$

which completes the induction. In particular, we have $5^{2^{\ell-3}} \not\equiv 1 \pmod{2^\ell}$ and $5^{2^{\ell-2}} \equiv 1 \pmod{2^\ell}$, so 5 has order $2^{\ell-2}$ in $(\mathbb{Z}/2^\ell\mathbb{Z})^*$.

We now show that the numbers $(-1)^a 5^b$ with $a \in \{0,1\}$ and $b \in \{0, \ldots, 2^{\ell-2} - 1\}$ are distinct modulo 2^ℓ for $\ell \ge 3$. Suppose that

$$(-1)^{a_1} 5^{b_1} \equiv (-1)^{a_2} 5^{b_2} \pmod{2^\ell}$$

for some $a_1, a_2 \in \{0,1\}$ and $b_1, b_2 \in \{0, \ldots, 2^{\ell-2} - 1\}$. Then, we obtain

$$(-1)^{a_1} 5^{b_1} \equiv (-1)^{a_2} 5^{b_2} \pmod 4.$$

Since $5 \equiv 1 \pmod 4$, we have

$$(-1)^{a_1} \equiv (-1)^{a_2} \pmod 4,$$

which implies that $a_1 = a_2$. On the other hand, we get

$$5^{b_1} \equiv 5^{b_2} \pmod{2^\ell},$$

and since 5 has order $2^{\ell-2}$ with $b_1, b_2 \in \{0, \ldots, 2^{\ell-2}-1\}$, we have $b_1 = b_2$. □

Theorem 6.14

The only positive integers that have primitive roots are 1, 2, 4, p^a, or $2p^a$, where p is an odd prime and a is a positive integer.

Proof. Let $n = 2^{\ell_0} p_1^{\ell_1} \cdots p_r^{\ell_r}$, where $p_1, \ldots, p_r$ are distinct odd primes and $\ell_0, \ldots, \ell_r$ are non-negative integers. We have shown in Theorem 6.2 that

$$(\mathbb{Z}/n\mathbb{Z})^* \cong (\mathbb{Z}/2^{\ell_0}\mathbb{Z})^* \times (\mathbb{Z}/p_1^{\ell_1}\mathbb{Z})^* \times \cdots \times (\mathbb{Z}/p_r^{\ell_r}\mathbb{Z})^*.$$

By Theorem 6.12, $(\mathbb{Z}/p_i^{\ell_i}\mathbb{Z})^*$ is cyclic for all $i = 1, \ldots, r$. Moreover, by Theorem 6.13, $(\mathbb{Z}/2^{\ell_0}\mathbb{Z})^*$ is cyclic for $0 \leq \ell_0 \leq 2$ and is isomorphic to $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2^{\ell_0-2}\mathbb{Z}$ for $\ell \geq 3$. Hence, the order of any element of $\mathbb{Z}/n\mathbb{Z}$ is a divisor of $\lambda(n) = \operatorname{lcm}(b, \phi(p_1^{\ell_1}), \ldots, \phi(p_r^{\ell_r}))$, where we define

$$b = \begin{cases} \phi(2^{\ell_0}) & \text{if } 0 \leq \ell_0 \leq 2, \\ \phi(2^{\ell_0})/2 & \text{if } \ell_0 \geq 3. \end{cases}$$

It is clear that $\lambda(n) < \phi(2^{\ell_0})\phi(p_1^{\ell_1}) \cdots \phi(p_r^{\ell_r})$ except in the cases where n is of the form 1, 2, 4, p^a, or $2p^a$ where p is a prime and a is a positive integer. □

Definition 6.15

Write $n = 2^{\ell_0} p_1^{\ell_1} \cdots p_r^{\ell_r}$, where $p_1, \ldots, p_r$ are distinct odd primes and $\ell_0, \ell_1, \ldots, \ell_r$ are non-negative integers. If we define

$$b = \begin{cases} \phi(2^{\ell_0}) & \text{if } 0 \leq \ell_0 \leq 2, \\ \phi(2^{\ell_0})/2 & \text{if } \ell_0 \geq 3, \end{cases}$$

then $\lambda(n) = \operatorname{lcm}(b, \phi(p_1^{\ell_1}), \ldots, \phi(p_r^{\ell_r}))$ is called the **universal exponent** of n.

Theorem 6.16

Let n be a positive integer, and let $\lambda(n)$ be the universal exponent of n. Then for any integer a coprime with n, we have

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

Proof. This follows from Theorem 6.14. □

This theorem gives us a strengthening of Euler's theorem (Theorem 5.2). Given a prime p, one can ask what an upper bound is for the smallest positive integer a which is a primitive root modulo p. Hua proved that

$$a < 2^{\omega(p-1)+1}\sqrt{p}.$$

Theorem 6.17

If p is a prime of the form $4q+1$ where q is an odd prime, then 2 is a primitive root modulo p.

Proof. Let t be the order of 2 modulo p. By Fermat's little theorem (Corollary 5.3), we have $t \mid (p-1)$ and hence $t \mid 4q$. By Theorem 6.14, t is one of 1, 2, 4, $2q$, or $4q$. Note that $p = 13$ or $p > 20$, so t cannot be 1, 2, or 4. Furthermore, by Euler's criterion (Theorem 5.10), we have

$$2^{(p-1)/2} \equiv 2^{2q} \equiv \left(\frac{2}{p}\right) \pmod{p}.$$

But we have

$$\left(\frac{2}{p}\right) = (-1)^{(p^2-1)/8} = (-1)^{[(4q)^2+8q]/8} = (-1)^q = -1.$$

Then t cannot be q or $2q$, so we must have $t = 4q = p - 1$, as required. $\square$

7 L-functions and Dirichlet's Theorem

7.1 Some Results on Primes in Arithmetic Progressions

Let k and ℓ be coprime positive integers. Recall that Dirichlet's theorem asserts that $kn + \ell$ is prime for infinitely many integers n. For many pairs (k, ℓ), this result can be proved using elementary means. However, this is hard to prove generally, and we'll require more tools to do so.

For example, consider the pair $(k, \ell) = (4, 3)$. Suppose that there are only finitely many primes $p_1, \ldots, p_k$ of the form $4n + 3$. Then $4p_1 \cdots p_k + 3$ must be divisible by a prime of the form $4n + 3$, since a product of primes congruent to 1 modulo 4 can only yield numbers congruent to 1 modulo 4. Notice that such a prime cannot be any of $p_1, \ldots, p_k$, which is a contradiction.

The following result is Dirichlet's theorem in the case where k is an arbitrary positive integer and $\ell = 1$.

Theorem 7.1

Let $n \in \mathbb{Z}^+$. There are infinitely many primes congruent to 1 modulo n.

Proof. This proof is due to Birkhoff and Vandiver (1904). Let $a > 2$ be an integer, and consider the n-th cyclotomic polynomial

$$\Phi_n(x) = \prod_{\substack{1 \le j \le n \\ \gcd(j,n)=1}} (x - \zeta_n^j),$$

where $\zeta_n = e^{2\pi i/n}$. We know that $\Phi_n(x) \in \mathbb{Z}[x]$ is irreducible, and $x^n - 1 = \prod_{d|n} \Phi_d(x)$. We now consider $\Phi_n(x)$ evaluated at a.

Claim. If p is a prime dividing $\Phi_n(a)$, then $p \mid n$ or $p \equiv 1 \pmod{n}$.

Proof of Claim. Since $x^n - 1 = \prod_{d|n} \Phi_d(x)$, we have $p \mid a^n - 1$. We consider two cases.

First, if $p \nmid a^d - 1$ for all proper divisors d of n, then the order of a modulo p must be n. By Fermat's little theorem (Corollary 5.3), we have $n \mid (p - 1)$, so $p \equiv 1 \pmod{n}$.

Suppose there is a proper divisor d of n such that $p \mid a^d - 1$. Since $p \mid \Phi_n(a)$, we obtain $p \mid (a^n - 1)/(a^d - 1)$. Notice that

$$a^n = \big(1 + (a^d - 1)\big)^{n/d} = 1 + \frac{n}{d}(a^d - 1) + \binom{n/d}{2}(a^d - 1)^2 + \binom{n/d}{3}(a^d - 1)^3 + \cdots,$$

so it follows that

$$\frac{a^n - 1}{a^d - 1} = \frac{n}{d} + \binom{n/d}{2}(a^d - 1) + \binom{n/d}{3}(a^d - 1)^2 + \cdots.$$

Since $p \mid (a^n - 1)/(a^d - 1)$ and $p \mid (a^d - 1)$, we get $p \mid n/d$ as well. Therefore, we have $p \mid n$ as required. ■

Now, assume that there are only finitely many primes $p_1, \ldots, p_k$ congruent to 1 modulo n. The n-th cyclotomic polynomial is of the form

$$\Phi_n(x) = x^{\phi(n)} + \cdots \pm 1.$$

Let m be an integer. We see that $\Phi_n(np_1 \cdots p_k m)$ is not divisible by p_i for $i = 1, \ldots, k$ and is coprime with n. Notice that for sufficiently large m, we have $\Phi_n(np_1 \cdots p_k m) \ge 2$. In particular, $\Phi_n(np_1 \cdots p_k m)$ has a prime divisor congruent to 1 modulo n which is not in the set $\{p_1, \ldots, p_k\}$, which is a contradiction. □

7.2 Characters

Definition 7.2

Let G be a finite abelian group. A **character** of G is a homomorphism $\chi : G \to \mathbb{C}^*$. Note that the set of characters of G forms a group under the operation $(\chi_1 \cdot \chi_2)(g) = \chi_1(g)\chi_2(g)$. We call this group the **dual group** of G, and denote it by $\hat{G}$. The identity of $\hat{G}$ is the character χ_0 with $\chi_0(g) = 1$ for all $g \in G$. We call χ_0 the **principal character**.

Notice that if $|G| = n$, then $g^n = e$ for all $g \in G$. In particular, we see that $(\chi(g))^n = \chi(g^n) = \chi(e) = 1$, so $\chi(g)$ is an n-th root of unity for all $g \in G$.

Theorem 7.3

Let G be a finite abelian group.

(1) The order of $\hat{G}$ is equal to the order of G.

(2) The dual group $\hat{G}$ is isomorphic to G.

(3) We have the formulas

$$\sum_{\chi \in \hat{G}} \chi(g) = \begin{cases} |G| & \text{if } g = e, \\ 0 & \text{otherwise;} \end{cases}$$

$$\sum_{g \in G} \chi(g) = \begin{cases} |G| & \text{if } \chi = \chi_0, \\ 0 & \text{otherwise.} \end{cases}$$

Proof.

(1) Recall that a finite abelian group is the direct product of cyclic groups. Hence, there exist elements $g_1, \ldots, g_r \in G$ and $h_1 \cdots h_r \in \mathbb{N}$ with $h_1 \cdots h_r = |G|$ such that every element $g \in G$ has a unique representation $g = g_1^{a_1} \cdots g_r^{a_r}$ with $1 \le a_i \le h_i$, and $g_i^{h_i} = e$ for $i = 1, \ldots, r$.

Any character χ is uniquely determined by its action on $g_1, \ldots, g_r$. Since $g_i^{h_i} = e$, we have $(\chi(g_i))^{h_1} = 1$, which shows that $\chi(g_i)$ is an h_i-th root of unity. Hence, there are at most $h_1 \cdots h_r$ characters.

On the other hand, there are at least $h_1 \cdots h_r$ characters because if ω_i is an h_i-th root of unity, then we may define $\chi(g_i) = \omega_i$ for $i = 1, \ldots, r$ and extend multiplicatively to G. We conclude that $|\hat{G}| = |G|$.

(2) For each $i = 1, \ldots, r$, let χ_i be the character which sends g_i to $e^{2\pi i/h_i}$ and g_j to 1 when $j \neq i$. Define $\phi : G \to \hat{G}$ by

$$\phi(g_1^{a_1} \cdots g_r^{a_r}) = \chi_1^{a_1} \cdots \chi_r^{a_r}.$$

Notice that ϕ is a homomorphism. We see that ϕ is injective because for $j = 1, \ldots, r$, we have

$$(\chi_1^{a_1} \cdots \chi_r^{a_r})(g_j) = e^{2\pi i a_j/h_j}$$

Since G is finite and $|\hat{G}| = |G|$ by (1), we see that ϕ is also surjective. Therefore, we have $\hat{G} \cong G$.

(3) Let $S(g) = \sum_{\chi \in \hat{G}} \chi(g)$. Notice that $\chi(e) = 1$ for all $\chi \in \hat{G}$, so we obtain $S(g) = |\hat{G}| = |G|$.

Assume now that $g \neq e$. Then there exists a character $\chi_1 \in \hat{G}$ such that $\chi_1(g) \neq 1$. Now, we have

$$S(g) = \sum_{\chi \in \hat{G}} \chi(g) = \sum_{\chi \in \hat{G}} (\chi_1\chi)(g) = \chi_1(g) \sum_{\chi \in \hat{G}} \chi(g) = \chi_1(g)S(g).$$

Since $\chi_1(g) \neq 1$, we must have $S(g) = 0$.

On the other hand, let $T(\chi) = \sum_{g \in G} \chi(g)$. Notice that $\chi_0(g) = 1$ for all $g \in G$, so $T(\chi_0) = |G|$. If $\chi \neq \chi_0$, then there exists $g_1 \in G$ such that $\chi(g_1) \neq 1$. We have

$$T(\chi) = \sum_{g \in G} \chi(g) = \sum_{g \in G} \chi(g_1 g) = \chi(g_1) \sum_{g \in G} \chi(g) = \chi(g_1) T(\chi).$$

Since $\chi(g_1) \neq 1$, it follows that $T(\chi) = 0$. $\square$

7.3 Dirichlet Characters

Definition 7.4

Let $k \in \mathbb{Z}^+$, and denote $(\mathbb{Z}/k\mathbb{Z})^*$ by $G(k)$. Let χ be a character on $G(k)$. We can associate χ with a map $\mathbb{Z} \to \mathbb{C}^*$, which we also call χ, by setting

$$\chi(a) = \begin{cases} \chi([a]) & \text{if } [a] \in G(k), \\ 0 & \text{if } [a] \notin G(k), \end{cases}$$

where $[a]$ denotes the conjugacy class of a. We call χ a **character modulo** k.

Theorem 7.5

Let $k \in \mathbb{Z}^+$, and let χ be a character modulo k.

(1) If $\gcd(n, k) = 1$, then $\chi(n)$ is a $\phi(k)$-th root of unity.

(2) We have $\chi(mn) = \chi(m)\chi(n)$; that is, χ is completely multiplicative.

(3) We have $\chi(n + k) = n$ for all $n \in \mathbb{Z}$; that is, χ is periodic with period k.

(4) We have the formulas

$$\sum_{n=1}^{k} \chi(n) = \begin{cases} \phi(k) & \text{if } \chi = \chi_0, \\ 0 & \text{otherwise;} \end{cases}$$

$$\sum_{\chi} \chi(n) = \begin{cases} \phi(k) & \text{if } n \equiv 1 \pmod{k}, \\ 0 & \text{otherwise,} \end{cases}$$

where the second sum runs through all characters modulo k.

(5) Let $\overline{\chi}$ be the conjugate character of χ with $\overline{\chi}(n) = \overline{\chi(n)}$ for all $n \in \mathbb{Z}$. Let χ' be a character modulo k. Then we have the formulas

$$\sum_{\chi \in \hat{G}(k)} \chi(n)\overline{\chi}(m) = \begin{cases} \phi(k) & \text{if } n \equiv m \pmod{k} \text{ and } \gcd(n, k) = 1, \\ 0 & \text{otherwise;} \end{cases}$$

$$\sum_{n=1}^{k} \chi(n)\chi'(n) = \begin{cases} \phi(k) & \text{if } \chi' = \overline{\chi}, \\ 0 & \text{otherwise.} \end{cases}$$

PROOF. Properties (1) to (4) either follow from definitions or Theorem 7.3. Note that $\overline{\chi}(m)\chi(m) = 1 = \chi(m^{-1})\chi(m)$, where m^{-1} denotes the multiplicative inverse of m in $G(k)$. Therefore, we have $\overline{\chi}(m) = \chi(m^{-1})$. It follows that

$$\sum_{\chi\in\hat{G}(k)} \chi(n)\overline{\chi}(m) = \sum_{\chi\in\hat{G}(k)} \chi(n)\chi(m^{-1}) = \sum_{\chi\in\hat{G}(k)} \chi(nm^{-1}).$$

By (4), the last sum is $\phi(k)$ if $nm^{-1} \equiv 1 \pmod{k}$, or equivalently $n \equiv m \pmod{k}$, and the sum is 0 otherwise. This gives us the first equation in (5). Moreover, note that if $\chi' = \overline{\chi}$, then $\chi\chi' = \chi_0$. Otherwise, $\chi\chi'$ is a non-principal character, so the second equation in (5) follows from (4). $\square$

We now describe the group of characters modulo k. By the multiplicative property of characters, it is enough to discuss the characters modulo p^a where p is prime and $a \in \mathbb{Z}^+$. First, assume that p is an odd prime. Let g be a primitive root modulo p^a. If n is coprime with p, then there is a unique integer $1 \le \nu \le \phi(p^a)$ such that $n \equiv g^\nu \pmod{p^a}$. For each integer $1 \le b \le \phi(p^a)$, we define the character χ^b by

$$\chi^b(a) = \exp\left(\frac{2\pi i b\nu}{\phi(p^a)}\right).$$

In this way, we get $\phi(p^a)$ different characters modulo p^a, so this is the complete list. Now, let $k = 2^a$. If $a = 1$, then we simply have the principal character. For $a = 2$, we have the principal character together with the character χ_4 given by

$$\chi_4(n) = \begin{cases} 1 & \text{if } n \equiv 1 \pmod 4, \\ -1 & \text{if } n \equiv -1 \pmod 4, \\ 0 & \text{otherwise.} \end{cases}$$

If $a \ge 3$, then $(\mathbb{Z}/2^a\mathbb{Z})^*$ is not cyclic by Theorem 6.13. Moreover, we saw in the proof of that theorem that for any odd integer n, there is a unique pair of integers (x, y) with $x \in \{0, 1\}$ and $y \in \{0, \ldots, 2^{a-2} - 1\}$ such that

$$n \equiv (-1)^x 5^y \pmod{2^a}.$$

For $c, d \in \mathbb{Z}$ with $c \in \{0, 1\}$ and $d \in \{0, \ldots, 2^{a-2} - 1\}$, we define

$$\chi_{2^a}^{c,d}(n) = \begin{cases} \exp(\frac{2\pi i c x}{2} + \frac{2\pi i d y}{2^{a-2}}) & \text{if } n \equiv 1 \pmod 2, \\ 0 & \text{otherwise.} \end{cases}$$

We obtain $\phi(2^a)$ different characters modulo 2^a, giving us the complete list.

7.4 Dirichlet L-functions

The Riemann zeta function was a powerful tool for studying the prime counting function $\pi(x)$. This suggests that it might be helpful to introduce complex functions in order to understand the primes in arithmetic progressions.

DEFINITION 7.6

Let $k \in \mathbb{Z}^+$, and let χ be a character modulo k. For $\mathrm{Re}(s) > 1$, we define the **Dirichlet L-function** by

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

As with the Riemann zeta function, we can establish the analytic continuation of $L(s, \chi)$ up to $\mathrm{Re}(s) > 0$.

Theorem 7.7

The function $L(s,\chi)$ can be analytically continued to $\mathrm{Re}(s) > 0$ except when χ is the principal character. If χ_0 is the principal character modulo k, then $L(s,\chi_0)$ can be analytically continued to $\mathrm{Re}(s) > 0$ except at the point $s = 1$, where we have a simple pole with residue $\phi(k)/k$.

Proof. Let $A(x) = \sum_{n \le x} \chi(n)$. By (4) of Theorem 7.5, we see that

$$A(x) = \begin{cases} \lfloor \frac{x}{k} \rfloor \phi(k) + T(x) & \text{if } \chi = \chi_0, \\ \lfloor \frac{x}{k} \rfloor 0 + T(x) & \text{if } \chi \neq \chi_0, \end{cases}$$

where $|T(x)| < \phi(k)$. It follows that

$$A(x) = E(\chi)\frac{\phi(k)x}{k} + R(x),$$

with $|R(x)| < 2\phi(k)$ and

$$E(\chi) = \begin{cases} 1 & \text{if } \chi = \chi_0, \\ 0 & \text{if } \chi \neq \chi_0. \end{cases}$$

Let $f(n) = 1/n^s$. By Abel's summation formula (Lemma 2.8), we have

$$\begin{aligned} \sum_{n \le x} \frac{\chi(n)}{n^s} &= \frac{A(x)}{x^s} + s\int_1^x \frac{A(u)}{u^{s+1}}\,\mathrm{d}u \\ &= E(\chi)\frac{\phi(k)}{k}\frac{1}{x^{s-1}} + \frac{R(x)}{x^s} + sE(\chi)\frac{\phi(k)}{k}\left(\frac{-u^{-s+1}}{s-1}\bigg|_1^x\right) + s\int_1^x \frac{R(u)}{u^{s+1}}\,\mathrm{d}u \\ &= E(\chi)\frac{\phi(k)}{k}\left(x^{1-s} + \frac{s}{1-s}(x^{1-s}-1)\right) + \frac{R(x)}{x^s} + s\int_1^x \frac{R(u)}{u^{s+1}}\,\mathrm{d}u. \end{aligned}$$

We now consider two cases.

- If $\chi \neq \chi_0$, then $E(\chi) = 0$. We see from above that

$$\sum_{n \le x} \frac{\chi(n)}{n^s} = \frac{R(x)}{x^s} + s\int_1^x \frac{R(u)}{u^{s+1}}\,\mathrm{d}u.$$

We have $|R(x)| < 2\phi(k)$, so by letting $x \to \infty$, we see that

$$L(s,\chi) = s\int_1^\infty \frac{R(u)}{u^{s+1}}.$$

This integral converges for $\mathrm{Re}(s) > 0$, so $L(s,\chi)$ has an analytic continuation to $\mathrm{Re}(s) > 0$.

- Note that $E(\chi_0) = 1$. By the above equation, we have

$$\sum_{n \le x} \frac{\chi_0(n)}{n^s} = \frac{\phi(k)}{k}\left(x^{1-s} + \frac{s}{1-s}(x^{1-s}-1)\right) + \frac{R(x)}{x^s} + s\int_1^x \frac{R(u)}{u^{s+1}}\,\mathrm{d}u.$$

Since $|R(x)| < 2\phi(k)$, letting $x \to \infty$ again gives

$$L(s,\chi_0) = \frac{\phi(k)}{k}\frac{s}{s-1} + s\int_1^\infty \frac{R(u)}{u^{s+1}}.$$

The integral converges for $\mathrm{Re}(s) > 0$, so $L(s,\chi_0)$ has an an analytic continuation to $\mathrm{Re}(s) > 0$, except at the simple pole $s = 1$ with residue $\phi(k)/k$. $\square$

7.5 Dirichlet Series

Definition 7.8

Let $\{\lambda_n\}_{n=1}^{\infty}$ be a strictly increasing sequence of positive real numbers. A **Dirichlet series** attached to $\{\lambda_n\}_{n=1}^{\infty}$ is a series of the form

$$\sum_{n=1}^{\infty} a_n e^{-\lambda_n z},$$

where $\{a_n\}_{n=1}^{\infty}$ is a sequence of complex numbers and $z \in \mathbb{C}$.

Theorem 7.9

If the Dirichlet series $\sum_{n=1}^{\infty} a_n e^{-\lambda_n z}$ converges at $z = z_0$, then it converges uniformly for $\mathrm{Re}(z - z_0) \geq 0$ and $|\arg(z - z_0)| \leq \alpha$ with $\alpha < \pi/2$.

Proof. Without loss of generality, we may assume that $z_0 = 0$. Note that $\sum_{n=1}^{\infty} a_n$ converges, so for any $\varepsilon > 0$, there exists $N = N(\varepsilon) \in \mathbb{N}$ such that if $\ell, m > N$, then

$$\left| \sum_{n=\ell}^{m} a_n \right| < \varepsilon.$$

Defining $A_{\ell,m} = \sum_{n=\ell}^{m} a_n$ and taking the convention that $A_{\ell,\ell-1} = 0$, we have

$$\begin{aligned}
\sum_{n=\ell}^{\infty} a_n e^{-\lambda_n z} &= \sum_{n=\ell}^{m} (A_{\ell,n} - A_{\ell,n-1}) e^{-\lambda_n z} \\
&= \sum_{n=\ell}^{m-1} A_{\ell,n}(e^{-\lambda_n z} - e^{-\lambda_{n+1} z}) + A_{\ell,m} e^{-\lambda_m z}.
\end{aligned}$$

For $\mathrm{Re}(z) \geq 0$, we see that

$$\left| \sum_{n=\ell}^{m} a_n e^{-\lambda_n z} \right| \leq \varepsilon \left(\sum_{n=\ell}^{m-1} |e^{-\lambda_n z} - e^{-\lambda_{n+1} z}| + 1 \right).$$

Note that

$$e^{-\lambda_n z} - e^{-\lambda_{n+1} z} = z \int_{\lambda_n}^{\lambda_{n+1}} e^{-tz}\, \mathrm{d}t.$$

Moreover, for $z = x + iy \in \mathbb{C}$ with $x, y \in \mathbb{R}$, we have $|e^{-tz}| = e^{-tx}$. Hence, we obtain

$$\begin{aligned}
|e^{-\lambda_n z} - e^{-\lambda_{n+1} z}| &\leq |z| \int_{\lambda_n}^{\lambda_{n+1}} e^{-tx}\, \mathrm{d}t \\
&\leq |z| \left(-\frac{e^{-tx}}{x} \Bigg|_{\lambda_n}^{\lambda_{n+1}} \right) \\
&= \frac{|z|}{x} (e^{-\lambda_n x} - e^{-\lambda_{n+1} x}).
\end{aligned}$$

It follows that

$$\left| \sum_{n=\ell}^{m} a_n e^{-\lambda_n z} \right| \leq \varepsilon \left(\frac{|z|}{x} \sum_{n=\ell}^{m-1} (e^{-\lambda_n x} - e^{-\lambda_{n+1} x}) + 1 \right) = \varepsilon \left(\frac{|z|}{x} (e^{-\lambda_\ell x} - e^{-\lambda_m x}) + 1 \right).$$

Now, for $|\arg(z)| \leq \alpha$ with $\alpha < \pi/2$, we have

$$\frac{|z|}{x} = \frac{1}{\cos(\arg z)} < c$$

for some constant $c = c(\alpha)$. Moreover, note that $|e^{-\lambda_\ell x} - e^{-\lambda_m x}| \leq 2$. Therefore, we have

$$\left|\sum_{n=\ell}^{m} a_n e^{-\lambda_n z}\right| < (2c+1)\varepsilon.$$

In particular, the Dirichlet series converges uniformly for $\mathrm{Re}(z) \geq 0$ and $|\arg(z)| \leq \alpha$. $\square$

We have proved that if the Dirichlet series converges at $z = z_0$, then it determines an analytic function for $\mathrm{Re}(z - z_0) \geq 0$ and $|\arg(z - z_0)| \leq \alpha$ with $\alpha < \pi/2$. Next, we'll show that if $\{a_n\}_{n=1}^{\infty}$ is in addition a sequence of non-negative real numbers, then the domain of convergence for the analytic function determined by the series is limited only by a singularity on the real axis.

Theorem 7.10

Let $f(z) = \sum_{n=1}^{\infty} a_n e^{-\lambda_n z}$ be a Dirichlet series where $\{a_n\}_{n=1}^{\infty}$ is a sequence of non-negative real numbers. Suppose that the series converges for $\mathrm{Re}(z) > \sigma_0$ where $\sigma_0 \in \mathbb{R}$, and f can be analytically continued in a neighbourhood of σ_0. Then there exists $\varepsilon > 0$ such that $\sum_{n=1}^{\infty} a_n e^{-\lambda_n z}$ converges for $\mathrm{Re}(z) > \sigma_0 - \varepsilon$.

Proof. Without loss of generality, we may assume that $\sigma_0 = 0$. Since this series converges on $\mathrm{Re}(z) > 0$, then for any $z \in \mathbb{C}$ with $\mathrm{Re}(z) > 0$, the series converges uniformly on a neighbourhood centered at z by Theorem 7.9 (we can find $w \in \mathbb{C}$ with $\mathrm{Re}(w) > 0$ such that the fan-like uniform convergence area covers a neighbourhood of z). Then f is analytic at z, and hence analytic for $\mathrm{Re}(z) > 0$. Since f is analytic for $\mathrm{Re}(z) > 0$ and f is also analytic in a neighbourhood of $\sigma_0 = 0$, there exists $\varepsilon > 0$ such that f is analytic in $|z - 1| \leq 1 + \varepsilon$. We now consider the Taylor series expansion of f around 1 in $|z - 1| \leq 1 + \varepsilon$. Note that for $\mathrm{Re}(z) > 0$, we have

$$f^{(m)}(z) = \sum_{n=1}^{\infty} a_n (-\lambda_n)^m e^{-\lambda_n z}.$$

This implies that

$$f^{(m)}(1) = \sum_{n=1}^{\infty} a_n (-\lambda_n)^m e^{-\lambda_n}.$$

Now, the Taylor series expansion of f about 1 in $|z - 1| \leq 1 + \varepsilon$ is of the form

$$\sum_{m=0}^{\infty} \frac{f^{(m)}(1)}{m!} (z-1)^m.$$

Consider f at the point $z = -\varepsilon$. We have

$$\begin{aligned} f(-\varepsilon) &= \sum_{m=0}^{\infty} \left(\sum_{n=1}^{\infty} a_n (-\lambda_n)^m e^{-\lambda_n} \right) \frac{(-1-\varepsilon)^m}{m!} \\ &= \sum_{m=0}^{\infty} \left(\sum_{n=1}^{\infty} a_n \lambda_n^m e^{-\lambda_n} \right) \frac{(1+\varepsilon)^m}{m!}. \end{aligned}$$

Since $a_n \geq 0$ and all the other terms above are positive, we can switch the order of summation to obtain

$$f(-\varepsilon) = \sum_{n=1}^{\infty} a_n e^{-\lambda_n} \left(\sum_{m=0}^{\infty} \frac{\lambda_n^m (1+\varepsilon)^m}{m!} \right) = \sum_{n=1}^{\infty} a_n e^{-\lambda_n} e^{\lambda_n (1+\varepsilon)} = \sum_{n=1}^{\infty} a_n e^{\lambda_n \varepsilon} = \sum_{n=1}^{\infty} a_n e^{(-\lambda_n)(-\varepsilon)}.$$

Hence, the series $\sum_{n=1}^{\infty} a_n e^{-\lambda_n z}$ converges to f at $z = -\varepsilon$. By Theorem 7.9, it converges to f for $\mathrm{Re}(z) > -\varepsilon$ because for any $z \in \mathbb{C}$ with $\mathrm{Re}(z) > -\varepsilon$, we can find some $\alpha < \pi/2$ such that $|\arg(z - (-\varepsilon))| < \alpha$. $\square$

7.6 Dirichlet's Theorem

Theorem 7.11

If χ is a character modulo k, then $L(s,\chi)$ is nonzero for $\text{Re}(s) > 1$. Furthermore, if χ is not principal, then $L(1,\chi)$ is nonzero.

Proof. Note that $L(s,\chi)$ converges absolutely for $\text{Re}(s) > 1$. Moreover, we showed in Theorem 7.5 that χ is completely multiplicative, so $L(s,\chi)$ has an Euler product representation for $\text{Re}(s) > 1$ given by

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

Recall that given a sequence of complex numbers $\{a_n\}_{n=1}^{\infty}$, the product $\prod_{n=1}^{\infty}(1+a_n)$ with $1 + a_n \neq 0$ converges absolutely (to a nonzero value) if and only if the series $\sum_{n=1}^{\infty} |a_n|$ converges. Since $\sum_p |\chi(p)/p^s|$ converges for $\text{Re}(s) > 1$, so does the Euler product representation above. Thus, $L(s,\chi) \neq 0$ for $\text{Re}(s) > 1$.

For the second assertion, we have two cases, depending on whether χ is a real or complex character. For $\text{Re}(s) > 1$, the Euler product representation of $L(s,\chi)$ gives

$$\log^* L(s,\chi) = \sum_p -\log\left(1 - \frac{\chi(p)}{p^s}\right) = \sum_p \sum_{a=1}^{\infty} \frac{\chi(p^a)}{ap^{as}},$$

where log denotes the principal branch and $\log^*$ indicates a branch of the logarithm.

Let $k \geq 2$ be an integer, and let ℓ be an integer coprime with k. Then we have

$$\sum_{\chi \in \hat{G}(k)} \overline{\chi}(\ell) \log^* L(s,\chi) = \sum_p \sum_{a=1}^{\infty} \frac{1}{ap^{as}} \sum_{\chi \in \hat{G}(k)} \overline{\chi}(\ell)\chi(p^a).$$

By (5) of Theorem 7.5, we obtain

$$\sum_{\chi \in \hat{G}(k)} \overline{\chi}(\ell) \log^* L(s,\chi) = \phi(k) \sum_{a=1}^{\infty} \sum_{p^a \equiv \ell \pmod{k}} \frac{1}{ap^{as}}. \tag{7.1}$$

Taking $\ell = 1$ in equation (7.1) and exponentiating both sides, we get

$$\prod_{\chi \in \hat{G}(k)} L(s,\chi) = \exp\left(\phi(k) \sum_{a=1}^{\infty} \sum_{p^a \equiv 1 \pmod{k}} \frac{1}{ap^{as}}\right).$$

Therefore, if s is real with $s > 1$, then

$$\prod_{\chi \in \hat{G}(k)} L(s,\chi) \geq 1. \tag{7.2}$$

First, suppose that $L(1,\chi) = 0$ where χ is not a real character. Then $\overline{\chi}$ is a character modulo k with $\chi \neq \chi_0$. Notice that when s is real with $s > 1$, we have $\overline{L(s,\chi)} = L(s,\overline{\chi})$, and hence

$$L(1,\overline{\chi}) = \overline{L(1,\chi)} = 0.$$

By Theorem 7.7, $L(s,\chi_0)$ has a simple pole at $s = 1$, and $L(s,\chi)$ does not have a pole at $s = 1$ when $\chi \neq \chi_0$. Hence, as $s \to 1$ on the real axis, we have

$$\prod_{\chi \in \hat{G}(k)} L(s,\chi) = O((s-1)^{-1}(s-1)^2) = O(s-1).$$

However, this contradicts equation (7.2), so $L(1,\chi) \neq 0$ when χ is not a real character.

Suppose now that $L(1,\chi) = 0$ where χ is a real character. For $\mathrm{Re}(s) > 1$, we set

$$g(s) = \frac{\zeta(s)L(s,\chi)}{\zeta(2s)}.$$

The Euler product representation of g for $\mathrm{Re}(s) > 1$ is

$$\begin{aligned}
g(s) &= \prod_p \left(\frac{1 - p^{-2s}}{(1-p^{-s})(1-\chi(p)/p^s)} \right) \\
&= \prod_p \left(\frac{1 + p^{-s}}{1-\chi(p)/p^s} \right) \\
&= \prod_p \left(1 + \frac{1}{p^s} \right) \sum_{a=0}^{\infty} \frac{\chi(p^a)}{p^{as}} \\
&= \prod_p \left(1 + \sum_{a=1}^{\infty} \frac{\chi(p^{a-1}) + \chi(p^a)}{p^{as}} \right) \\
&= \prod_p \left(1 + \sum_{a=1}^{\infty} \frac{b(p^a)}{p^{as}} \right),
\end{aligned}$$

where $b(p^a) = \chi(p^{a-1}) + \chi(p^a)$. Since χ is a real character, it takes on values from $\{-1,0,1\}$. Moreover, we know that χ is multiplicative by Theorem 7.5, so we have

$$b(p^a) = \chi(p^{a-1}) + \chi(p^a) = \begin{cases} 0 & \text{if } \chi(p) = 0, \\ 2 & \text{if } \chi(p) = 1, \\ 0 & \text{if } \chi(p) = -1. \end{cases}$$

In all cases, we have $b(p^a) \geq 0$ for $a \geq 1$. Therefore, we see that $g(s) = \sum_{n=1}^{\infty} a_n/n^s$ where $a_1 = 1$ and $a_n \geq 0$ for all $n \geq 2$. Moreover, we had set $g(s) = \zeta(s)L(s,\chi)/\zeta(2s)$ for $\mathrm{Re}(s) > 1$. Since $L(1,\chi) = 0$ eliminates the pole of $\zeta(s)$ at $s = 1$ and $\zeta(2s)$ is nonzero and analytic for $\mathrm{Re}(s) > 1/2$, it follows that $g(s)$ has an analytic continuation to $\mathrm{Re}(s) > 1/2$.

We now apply Theorem 7.10 to conclude that the series defining g converges to g for $\mathrm{Re}(s) > 1/2$. Letting $s \to 1/2$ from above on the real axis, we have

$$g(s) = O(s - 1/2) = o(1)$$

since $\zeta(2s)$ has a pole at $s = 1/2$. However, since

$$g(s) = 1 + \sum_{n=2}^{\infty} \frac{a_n}{n^s}$$

with $a_n \geq 0$ for $n \geq 2$, we obtain $g(s) \geq 1$ for $\mathrm{Re}(s) > 1/2$. This is a contradiction, so we must have $L(1,\chi) \neq 0$ when χ is a real character. $\square$

Theorem 7.12

If k and ℓ are coprime integers with $k \geq 2$, then the series

$$\sum_{p \equiv \ell \pmod{k}} \frac{1}{p}$$

diverges. Consequently, there are infinitely many primes in the arithmetic progression $kn + \ell$.

PROOF. From equation (7.1) from Theorem 7.11, we have

$$\frac{1}{\phi(k)} \sum_{\chi \in \hat{G}(k)} \overline{\chi}(\ell) \log L(s,\chi) = \sum_{a=1}^{\infty} \sum_{p^a \equiv \ell \pmod{k}} \frac{1}{ap^{as}}.$$

As $s \to 1$ from the right on the real axis, $(s-1)^{E(\chi)} L(s,\chi)$ tends to a finite nonzero limit, where $E(\chi) = 1$ if $\chi = \chi_0$ and $E(\chi) = 0$ otherwise. Then $E(\chi)\log(s-1) + \log L(s,\chi)$ also tends to a limit. It follows that as $s \to 1$ from the right on the real axis, we have

$$\log L(s,\chi) = -E(\chi)\log(s-1) + O(1).$$

Hence, we get

$$\begin{aligned} \frac{1}{\phi(k)} \sum_{\chi \in \hat{G}(k)} \overline{\chi}(\ell) \log L(s,\chi) &= \frac{1}{\phi(k)} \log L(s,\chi_0) + \frac{1}{\phi(k)} \sum_{\substack{\chi \in \hat{G}(k) \\ \chi \neq \chi_0}} \overline{\chi}(\ell) \log L(s,\chi) \\ &= -\frac{1}{\phi(k)} \log(s-1) + O(1). \end{aligned}$$

Combining this with (7.1) yields

$$\sum_{a=1}^{\infty} \sum_{p^a \equiv \ell \pmod{k}} \frac{1}{ap^{as}} = -\frac{1}{\phi(k)} \log(s-1) + O(1).$$

Thus, we have

$$\sum_{p \equiv \ell \pmod{k}} \frac{1}{p^s} + \sum_{a=2}^{\infty} \sum_{p^a \equiv \ell \pmod{k}} \frac{1}{ap^{as}} = -\frac{1}{\phi(k)} \log(s-1) + O(1).$$

On the other hand, for $s \in \mathbb{R}$ with $s \geq 1$, we have

$$\begin{aligned} \sum_{a=2}^{\infty} \sum_{p^a \equiv \ell \pmod{k}} \frac{1}{ap^{as}} &\leq \frac{1}{2} \sum_{a=2}^{\infty} \sum_{p^a \equiv \ell \pmod{k}} \frac{1}{p^{as}} \\ &\leq \frac{1}{2} \sum_{n=2}^{\infty} \left(\frac{1}{n^{2s}} + \frac{1}{n^{3s}} + \cdots \right) \\ &\leq \frac{1}{2} \sum_{n=2}^{\infty} \frac{1}{n^{2s}} \left(\frac{1}{1 - 1/n^s} \right) \\ &\leq \sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}. \end{aligned}$$

Therefore, as $s \to 1$ from the right on the real axis, we obtain

$$\sum_{p \equiv \ell \pmod{k}} \frac{1}{p^s} = -\frac{1}{\phi(k)} \log(s-1) + O(1).$$

Since the quantity $\log(s-1)$ blows up as $s \to 1$, the series diverges. $\square$

7.7 Distribution of Primes in Arithmetic Progressions

Let k and ℓ be coprime integers with $k \geq 2$. For each $x \in \mathbb{R}$, define $\pi(x,k,\ell)$ to be the number of primes p satisfying $p \leq x$ and $p \equiv \ell \pmod{k}$. Then it can be shown that

$$\pi(x,k,\ell) \sim \frac{1}{\phi(k)} \frac{x}{\log x} \sim \frac{\text{Li}(x)}{\phi(k)}.$$

For $t \in \mathbb{R}$ and $k \in \mathbb{Z}^+$, set $\tau(k,t) = \max\{|t|, k+2\}$. Let $c \in \mathbb{R}$ with $0 < c < 1$, and define the set $R_c(k)$ by

$$R_c(k) = \left\{ \sigma + it : 1 - \frac{c}{\log \tau(k,t)} < \sigma \right\}.$$

One can show that there exists a positive real number c_0 such that if χ is a non-real character modulo k for $k \geq 2$, then $L(s,\chi)$ is nonzero in $R_{c_0}(k)$.

When χ is a real non-principal character, this is not true in general. However, such a c_0 exists if we allow for the possibility that there is a point β on the real axis in $R_{c_0}(k)$ where $L(s,\chi)$ is zero.

Definition 7.13

If $L(s,\chi)$ vanishes at $\beta \in R_{c_0}(k)$, then β is a simple zero of $L(s,\chi)$ and is called a **Siegel zero**.

The extended Riemann hypothesis implies that $L(s,\chi)$ is nonzero for $\mathrm{Re}(s) > 1/2$, so no Siegel zero exists under this hypothesis.

Let k and ℓ be coprime integers with $k \geq 2$. Put $b = \beta(k)$ if there is a real non-principal character χ where β is a zero of $L(s,\chi)$ in $R_{c_0}(k)$, and set $b = 1$ otherwise. Then there exists $a > 0$ such that

$$\pi(x,k,\ell) = \frac{\mathrm{Li}(x)}{\phi(k)} - \frac{\lambda(b)}{b}\frac{x^b}{\phi(k)} + O(x\exp(-a\sqrt{\log x})),$$

where $\lambda(b) = 0$ if $b = 1$, and $\lambda(b) = \chi(\ell)$ if $b \neq 1$. We would like to know if the term

$$\frac{\lambda(b)}{b}\frac{x^b}{\phi(k)}$$

exists, or in other words, whether a Siegel zero exists. We haven't been able to do so yet, however. The best "effective" estimate for the size of a Siegel zero $\beta(k)$ associated to $L(s,\chi)$ where χ is a real character modulo k is due to Pintz, who proved that

$$\beta(k) < 1 - \frac{c}{\sqrt{k}},$$

where c is an effectively computable positive number. On the other hand, Siegel proved that for every $\varepsilon > 0$, there exists a positive number $c(\varepsilon)$ such that

$$\beta(k) < 1 - \frac{c(\varepsilon)}{k^\varepsilon}.$$

Unfortunately, there is no known algorithm for computing $c(\varepsilon)$ given $\varepsilon > 0$. Using Siegel's estimate, one can prove that if H is a positive number satisfying $k \leq (\log x)^H$, then

$$\pi(x,k,\ell) = \frac{\mathrm{Li}(x)}{\phi(k)} + O\left(\frac{x}{\exp(C\sqrt{\log x})}\right)$$

for some $C > 0$. However, notice that this big-O term is quite ineffective.

8 Waring's Problem

8.1 History of Waring's Problem

A famous problem in additive number theory is Waring's problem. In 1770, Edward Waring asserted without proof that every natural number is the sum of at most 4 squares, 9 cubes, and 19 biquadrates. In general, Waring's problem is stated as follows.

Proposition 8.1: Waring's problem

For every $k \in \mathbb{N}$ with $k \geq 2$, there exists an integer $s = s(k)$ such that every natural number n is the sum of at most s k-th powers of natural numbers. That is, we have

$$n = x_1^k + \cdots + x_s^k$$

where $x_i \in \mathbb{N} \cup \{0\}$ for $i = 1, \ldots, s$.

Let $g(k)$ denote the smallest integer s such that the above statement holds. Then Waring's problem asserts that $g(2) = 4$, $g(3) = 9$, $g(4) = 19$, and in general, $g(k) < \infty$ for all $k \geq 2$.

In 1770, Lagrange proved that $g(2) = 4$. By 1909, the only known cases were $k = 2, 3, 4, 5, 6, 7, 8, 10$. In 1909, Hilbert proved using a combinatorial method that $g(k) < \infty$ for every $k \geq 2$.

By the work of Vinogradov, we now have an almost complete solution to $g(k)$. Consider the integer

$$n = 2^k \lfloor (3/2)^k \rfloor - 1 < 3^k.$$

The most efficient representation for n is to use $\lfloor (3/2)^k \rfloor - 1$ copies of 2^k, and $2^k - 1$ copies of 1^k. In other words, we can write

$$n = 2^k \left(\lfloor (3/2)^k \rfloor - 1 \right) + 1^k (2^k - 1).$$

By a result of Euler, we obtain the inequality

$$g(k) \geq 2^k + \lfloor (3/2)^k \rfloor - 2,$$

with equality holding for all but finitely many k. Recall that the fractional part of $x \in \mathbb{R}$ is defined by $\{x\}$. We see that equality holds if

$$2^k \{(3/2)^k\} + \lfloor (3/2)^k \rfloor \leq 2^k. \tag{8.1}$$

On the other hand, if we have $2^k \{(3/2)^k\} + \lfloor (3/2)^k \rfloor < 2^k$, then we have one of the cases

$$g(k) = \begin{cases} 2^k + \lfloor (3/2)^k \rfloor + \lfloor (4/3)^k \rfloor - 2 & \text{if } \lfloor (4/3)^k \rfloor \lfloor (3/2)^k \rfloor + \lfloor (4/3)^k \rfloor + \lfloor (3/2)^k \rfloor = 2^k, \\ 2^k + \lfloor (3/2)^k \rfloor + \lfloor (4/3)^k \rfloor - 3 & \text{if } \lfloor (4/3)^k \rfloor \lfloor (3/2)^k \rfloor + \lfloor (4/3)^k \rfloor + \lfloor (3/2)^k \rfloor > 2^k. \end{cases}$$

In 1957, Mahler showed that (8.1) holds for all but finitely k, and no exception is known.

8.2 Special Cases of Waring's Problem

We'll now establish that $g(2) = 4$. Determining $g(4)$ is significantly harder, but we can prove an upper bound for it by using some elementary arguments.

First, for each $x \in \mathbb{Z}/8\mathbb{Z}$, we observe that x^2 is congruent to either 0, 1, or 4 modulo 8. Then $x_1^2 + x_2^2 + x_3^2 \not\equiv 7 \pmod 8$ for any $x_1, x_2, x_3 \in \mathbb{Z}$, so we must have $g(2) \geq 4$.

Proposition 8.2

If p is an odd prime, then there exist $x, y \in \mathbb{Z}$ such that

$$1 + x^2 + y^2 = mp,$$

where $m \in \mathbb{Z}$ satisfies $1 \leq m \leq p - 1$.

Proof. First, consider the sets

$$S_1 = \{x^2 + p\mathbb{Z} : x \in \mathbb{Z},\ 0 \leq x \leq (p-1)/2\},$$
$$S_2 = \{-1 - y^2 + p\mathbb{Z} : y \in \mathbb{Z},\ 0 \leq y \leq (p-1)/2\}.$$

Note that $x_1^2 \equiv x_2^2 \pmod p$ if and only if $x_1 \equiv \pm x_2 \pmod p$. Then the elements of S_1 are distinct since $0 \leq x \leq (p-1)/2$, and the elements of S_2 are also distinct. Since $|S_1| = |S_2| = (p+1)/2$, we have $S_1 \cap S_2 \neq \varnothing$. Thus, there exist $x, y \in \mathbb{Z}$ with $0 \leq x, y \leq (p-1)/2$ such that $x^2 \equiv -1 - y^2 \pmod p$, or equivalently, $1 + x^2 + y^2 \equiv 0 \pmod p$. In particular, we have $1 + x^2 + y^2 = mp$ for some $m \in \mathbb{Z}$. Moreover, notice that

$$0 < m = \frac{1 + x^2 + y^2}{p} \leq \frac{1 + [(p-1)/2]^2 + [(p-1)/2]^2}{p} < p,$$

so $1 \leq m \leq p - 1$, as desired. ☐

Theorem 8.3: Lagrange's theorem

We have $g(2) = 4$. In other words, every natural number can be expressed as a sum of 4 squares.

Proof. Observe that we have the Lagrange identity

$$\begin{aligned}(x_1^2 + x_2^2 + x_3^2 + x_4^2)(y_1^2 + y_2^2 + y_3^2 + y_4^2) &= (x_1y_1 + x_2y_2 + x_3y_3 + x_4y_4)^2 + (x_1y_2 - x_2y_1 + x_3y^4 - x_4y_3)^2 \\ &\quad + (x_1y_3 - x_3y_1 + x_4y_2 - x_2y_4)^2 + (x_1y_4 - x_4y_1 + x_2y_3 - x_3y_2)^2,\end{aligned}$$

so the product of two numbers that are representable as a sum of 4 squares is also representable as a sum of 4 squares. In order to prove the theorem, it suffices to show that all primes can be written as the sum of 4 squares. It is clear that

$$2 = 1^2 + 1^2 + 0^2 + 0^2,$$

so it remains to show that any odd prime p can be written as the sum of 4 squares. By Proposition 8.2, there exist $x_1, x_2, x_3, x_4 \in \mathbb{Z}$ such that

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 = mp$$

where $1 \leq m \leq p - 1$. Indeed, we can take $x_1 = 1$, $x_2 = x$, $x_3 = y$, and $x_4 = 0$ to get $x_1^2 + x_2^2 + x_3^2 + x_4^2 = 1 + x^2 + y^2$. Next, let m_0 be the smallest natural number such that m_0p is the sum of 4 squares. We claim that $m_0 = 1$, and thus $x_1^2 + x_2^2 + x_3^3 + x_4^2 = p$. First, suppose that m_0 is even. Note that

$$(x_1 + x_2 + x_3 + x_4)^2 = x_1^2 + x_2^2 + x_3^2 + x_4^2 + 2 \sum_{1 \leq i < j \leq n} x_ix_j.$$

Since $x_1^2 + x_2^2 + x_3^2 + x_4^2 = m_0p$ is even, it follows that $x_1 + x_2 + x_3 + x_4$ is even. They may be all even, all odd, or two even and two odd. In the last case, suppose without loss of generality that x_1 and x_2 are even, and x_3 and x_4 are odd. Then in all cases, the terms $x_1 + x_2$, $x_1 - x_2$, $x_3 + x_4$, and $x_3 - x_4$ are all even. Thus, we obtain

$$\left(\frac{x_1 + x_2}{2}\right)^2 + \left(\frac{x_1 - x_2}{2}\right)^2 \left(\frac{x_3 + x_4}{2}\right)^2 \left(\frac{x_3 - x_4}{2}\right)^2 = \frac{x_1^2 + x_2^2 + x_3^2 + x_4^2}{2} = \frac{m_0}{2}p.$$

But this contradicts the minimality of m_0, so m_0 must be odd.

Next, suppose that $m_0 > 1$. Since $x_1^2 + x_2^2 + x_3^2 + x_4^2 = m_0 p$ and $1 \leq m_0 \leq p - 1$, we see that not all of x_1, x_2, x_3, x_4 are divisible by m_0, for otherwise $m_0^2 \mid m_0 p$ and hence $m_0 \mid p$, a contradiction. Hence, there exist $b_1, b_2, b_3, b_4 \in \mathbb{Z}$ such that $y_i = x_i - b_i m_0$ and $|y_i| < m_0/2$ for $i = 1, 2, 3, 4$ (note that m_0 being odd gives us the condition $|y_i| < m_0/2$), where the y_i are not all 0. Then we have

$$0 < y_1^2 + y_2^2 + y_3^2 + y_4^2 < 4\left(\frac{m_0}{2}\right)^2 = m_0^2,$$

and moreover, we know that $y_1^2 + y_2^2 + y_3^2 + y_4^2 \equiv 0 \pmod{m_0}$. Hence, there exists $m_1 \in \mathbb{N}$ with $m_1 < m_0$ such that

$$y_1^2 + y_2^2 + y_3^2 + y_4^2 = m_0 m_1.$$

Recall that $x_1^2 + x_2^2 + x_3^2 + x_4^2 = m_0 p$, so by multiplying these equations together, the Lagrange identity tells us that there exist $z_1, z_2, z_3, z_4 \in \mathbb{Z}$ such that

$$z_1^2 + z_2^2 + z_3^2 + z_4^2 = m_0^2 m_1 p.$$

Notice that

$$z_1 = x_1 y_1 + x_2 y_2 + x_3 y_3 + x_4 y_4 = \sum_{i=1}^{4} x_i y_i = \sum_{i=1}^{4} x_i (x_i - b_i m_0) = \sum_{i=1}^{4} x_i^2 + m_0 K$$

for some $K \in \mathbb{Z}$. Since $x_1^2 + x_2^2 + x_3^2 + x_4^2 = m_0 p$, this tells us that $z_1 \equiv 0 \pmod{m_0}$. By a similar argument, we can verify that z_2, z_3, z_4 are all divisible by m_0. Now, define $t_i = z_i/m_0$ for $i = 1, 2, 3, 4$. It follows that

$$t_1^2 + t_2^2 + t_3^2 + t_4^2 = m_1 p$$

with $1 \leq m_1 < m_0$, contradicting the minimality of m_0. Thus, we have $m_0 = 1$, so we are done. $\square$

Theorem 8.4

We have $g(4) \leq 53$.

Proof. By Theorem 8.3, every non-negative integer x can be written in the form $x = a^2 + b^2 + c^2 + d^2$ for some $a, b, c, d \in \mathbb{N} \cup \{0\}$. Observe that we have the identity

$$\begin{aligned} 6(a^2 + b^2 + c^2 + d^2) &= (a+b)^4 + (a-b)^4 + (c+d)^4 + (c-d)^4 \\ &\quad + (a+c)^4 + (a-c)^4 + (b+d)^4 + (b-d)^4 \\ &\quad + (a+d)^4 + (a-d)^4 + (b+c)^4 + (b-c)^4. \end{aligned}$$

Hence, every integer of the form $6x^2$ can be expressed as the sum of 12 fourth powers. Now, every natural number can be written in the form $6k + r$ where $k \in \mathbb{N} \cup \{0\}$ and $0 \leq r \leq 5$. By Theorem 8.3, we can write k as a sum of 4 squares, say $k = x_1^2 + x_2^2 + x_3^2 + x_4^2$. Then we have

$$6k = 6x_1^2 + 6x_2^2 + 6x_3^2 + 6x_4^2.$$

Each term in the above sum is a sum of 12 fourth powers, so $6k$ can be expressed as a sum of 48 fourth powers. Finally, we can write $r = 1^4 + \cdots + 1^4$ (r times). Since $0 \leq r \leq 5$, we conclude that $6k + r$ is a sum of 53 fourth powers. $\square$