Artin's Primitive Root Conjecture and its Extension to Composite Moduli

by

Patrice Camiré

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Mathematics
in
Pure Mathematics

Waterloo, Ontario, Canada, 2008

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

If we fix an integer $a \neq -1$, which is not a perfect square, we are interested in estimating the quantity $N_a(x)$ representing the number of prime integers p up to x such that a is a generator of the cyclic group $(\mathbb{Z}/p\mathbb{Z})^*$. We will first show how to obtain an asymptotic formula for $N_a(x)$ under the assumption of the generalized Riemann hypothesis. We then investigate the average behaviour of $N_a(x)$. More precisely, we study the quantity

$$\frac{1}{N} \sum_{1<a\leq N} N_a(x)\,.$$

Finally, we discuss how to generalize the problem over $(\mathbb{Z}/m\mathbb{Z})^*$, where $m > 0$ is not necessarily prime. We present an average result in this setting and prove the existence of oscillation.

Acknowledgements

I would like to thank my advisors, Professor Yu-Ru Liu and Professor Wentang Kuo, for suggesting that I work on Artin's primitive root conjecture and its extension to composite moduli. I am also grateful for their constant help and suggestions while making the final corrections to this thesis. I would also like to thank the University of Waterloo and the National Sciences and Engineering Research Council of Canada for their financial support. Last but not least, I wish to thank the readers, Professor Kevin Hare and Professor Cameron Stewart, for taking the time to read my thesis and make valuable suggestions.

Dedication

This thesis is dedicated to my parents, Daniel and Huguette.

Contents

Chapter 1

Introduction

Given a positive prime integer p, we can consider the set of invertible residue classes modulo p, denoted by $(\mathbb{Z}/p\mathbb{Z})^*$. It is straightforward to show that under multiplication, this set is a cyclic group of order $p-1$ with $\phi(p-1)$ generators, where ϕ is the Euleur totient function. A generator of $(\mathbb{Z}/p\mathbb{Z})^*$ is called a *primitive root* modulo p. Given an integer a, we are interested in determining whether or not a generates $(\mathbb{Z}/p\mathbb{Z})^*$ for infinitely many primes. A necessary condition for this to hold is that a does not equal -1 or a perfect square, but are these conditions sufficient? In 1927 Emil Artin [2, p.viii-x] conjectured that these conditions were indeed sufficient. Furthermore, letting $N_a(x)$ denote the number of primes p up to x for which a is a primitive root, he conjectured that

$$N_a(x) \sim A(a)\frac{x}{\log x}$$

as $x \to \infty$ and where $A(a) > 0$ is a constant depending on a. Let us now discuss briefly and informally how one may arrive to such a conclusion. Since $(a,p) = 1$ for all but a finite number of primes p, we may assume that $(a,p) = 1$. The first step is to classify the primes p for which a is not a primitive root modulo p. Notice that a is not a primitive root modulo p if and only if there exists a prime q such that the equation $\nu^q \equiv a (\mathrm{mod}\, p)$ has exactly q distinct roots modulo p. By a famous theorem of Dedekind, assuming that the polynomial $u^q - a$ is irreducible over $\mathbb{Q}[u]$, the previous condition is equivalent to the fact that $p \nmid q$ and that p splits completely over the Galois extension $\mathbb{Q}(\sqrt[q]{a}, \sqrt[q]{1})$ as a product of distinct linear prime ideals. With the above assumption, the degree of $\mathbb{Q}(\sqrt[q]{a}, \sqrt[q]{1})$ over $\mathbb{Q}$ can be shown to be $q(q-1)$. From the Chebotarev density theorem, the density of primes splitting completely over $\mathbb{Q}(\sqrt[q]{a}, \sqrt[q]{1})$ is $1/q(q-1)$. Therefore the probability that a

given prime p does not split completely over $\mathbb{Q}(\sqrt[q]{a}, \sqrt[q]{1})$ is equal to $1 - 1/q(q-1)$. It thus follows that for a to be a primitive root modulo p, we need that p does not split completely over $\mathbb{Q}(\sqrt[q]{a}, \sqrt[q]{1})$ for any prime q such that $p \nmid q$. Hence, the probability that a is a primitive root modulo p should be

$$A \;=\; \prod_{q\,:\,\text{prime}} \left(1 - \frac{1}{q(q-1)}\right),$$

leading one to believe that

$$N_a(x) \;\sim\; \prod_{q\,:\,\text{prime}} \left(1 - \frac{1}{q(q-1)}\right) \frac{x}{\log x}$$

as $x \to \infty$. Observe that the above argument may fail since the polynomial $u^q - a$ may be reducible over $\mathbb{Q}[u]$ for certain values of a and q. This leads one to suspect that the asymptotic constant $A(a)$ should depend on a in possibly different subtle ways. Concerns about the value of $A(a)$ were first brought up by D. H. Lehmer whose work revealed that the original formula did not appear to predict values for $N_a(x)$ that were in accord with the numerical evidence. In the light of this knowledge Heilbronn then suggested a revised form of the formula (see [27]). In 1967 Christopher Hooley [10] showed that Artin's primive root conjecture holds true under the assumption of the generalized Riemann hypothesis for certain Galois extensions. He also provided a complete description of the asymptotic constant $A(a)$. This is the subject of study of the second chapter of this thesis. Coming back to the polynomial $u^q - a$, it is straightforward to show that it is irreducible over $\mathbb{Q}[u]$ for all primes q for almost all integers a. This leads one to believe that

$$\prod_{q\,:\,\text{prime}} \left(1 - \frac{1}{q(q-1)}\right)$$

should be the right asymptotic constant on average. This is the subject of study of the third chapter where the work of P. J. Stephens [26] is presented. In the final chapter, a portion of the work by Shuguang Li [14] is presented on the extension of Artin's conjecture to composite moduli. An average result is demonstrated and the presence of oscillation is exhibited.

Chapter 2

Artin's Primitive Root Conjecture and the Generalized Riemann Hypothesis

Let

$$N_a(x) = \#\{\, p \leq x \mid \langle a \rangle \,=\, (\mathbb{Z}/p\mathbb{Z})^* \},$$

where $\langle a \rangle$ denotes the subgroup of $(\mathbb{Z}/p\mathbb{Z})^*$ generated by a. In this chapter, our goal is to prove both Artin's primitive root conjecture and provide an asymptotic formula for $N_a(x)$ subject to the assumption that the generalized Riemann hypothesis holds over a certain class of Galois extensions. The value of the constant $A(a)$ we obtain in the asymptotic formula for $N_a(x)$ agrees with the one conjectured by Heilbronn. Finally, we wish to mention that throughout this section the condition $p \nmid a$ remains always implicit, the reason being that any fixed integer a posesses only finitely many distinct prime divisors.

2.1 Notation

The letter a is a given non-zero integer that is not equal to $1, -1$, or a perfect square. p and q are positive prime numbers. l, m and r are positive integers. ν is an integer and k is a square-free integer.

x is a continuous real variable to be regarded as tending to infinity. All the inequalities given are valid for sufficiently large values of x. The function $\omega(l)$ is the

number of distinct prime factors of l. We let (h,k) denote the greatest common divisor of h and k. $\left(\frac{b}{c}\right)$ denotes the standard Jacobi symbol.

2.2 Formulation of the Method

We first observe the following equivalent statement for what it means for a to be a primitive root modulo p:

a is a primitive root modulo p if, and only if, $p \nmid a$ and there is no prime divisor q of $p-1$ for which there exists an integer ν such that $\nu^q \equiv a \pmod{p}$.

For a prime q, we let $R(q,p)$ denote the simultaneous conditions:

$$q \,|\, p-1 \text{ and there exists an integer } \nu \text{ such that } \nu^q \equiv a \pmod{p}.$$

We also denote by V the logical valuation operator. Hence given a sentence, say $\mathcal{S}$, we have that $V(\mathcal{S}) = \top$ if the sentence $\mathcal{S}$ is true and $V(\mathcal{S}) = \bot$ if the sentence $\mathcal{S}$ is false.

In order to study the sum $N_a(x)$ and to isolate the main contribution, we have to partition the interval $[1,x]$ into subintervals and we need to introduce several auxiliary sums. Let us first do the later. Although it may not be completely clear at first why we need to define such auxiliary sums, it will become apparent once we observe their properties and the different relationships that exist among them. Let us first define, given a set of conditions $\mathcal{C}$, the index function of $\mathcal{C}$ in the following way

$$\mathbb{1}\cdot\mathcal{C} := \begin{cases} 1\,, & \text{if every condition in } \mathcal{C} \text{ is satisfied} \\ 0\,, & \text{otherwise.} \end{cases}$$

Using the above definition, we define the following three auxiliary sums:

$$N_a(x,\eta) := \sum_{p\le x} \mathbb{1}\cdot\{\, V(R(q,p)) = \bot \ \forall \text{ prime } q \le \eta \,\},$$

$$M_a(x,\eta_1,\eta_2) := \sum_{p\le x} \mathbb{1}\cdot\{\, V(R(q,p)) = \top \text{ for at least one prime } q \text{ such that } \eta_1 < q \le \eta_2 \,\},$$

$$P_a(x,k) := \sum_{p\le x} \mathbb{1}\cdot\{\, V(R(q,p)) = \top \ \ \forall \,\text{prime}\, q \,|\, k \,\}\,,\ \text{for any square-free integer } k.$$

It should be clear that $N_a(x) = N_a(x, x-1)$ since $V(R(q,p)) = \bot$ for any prime number $q > p-1$ and that if $k = 1$, then $P_a(x,k) = \pi(x)$ where $\pi(x)$ is the prime counting function.

The partition of the interval $[1,x]$ is given by the following subintervals: $[1,\xi_1]$, $[\xi_1,\xi_2]$, $[\xi_2,\xi_3]$ and $[\xi_3,x]$, where

$$\xi_1 := \frac{1}{6}\log x\,,\ \xi_2 := \sqrt{x}\log^{-2} x \ \text{ and } \ \xi_3 := \sqrt{x}\log x.$$

We now present a series of inequalities and equalities between $N_a(x)$ and the auxiliary sums defined above. Each such relation can be deduced from definition and the equivalence mentioned at the beginning of this section. On the one hand, we have that

$$N_a(x) \le N_a(x,\xi_1)\,,$$

while on the other hand

$$N_a(x) \ge N_a(x,\xi_1) - M_a(x,\xi_1,x-1)\,.$$

Combining these two inequalities implies that

$$N_a(x) = N_a(x,\xi_1) + O\big(M_a(x,\xi_1,x-1)\big).$$

Moreover

$$M_a(x,\xi_1,x-1) \le M_a(x,\xi_1,\xi_2) + M_a(x,\xi_2,\xi_3) + M_a(x,\xi_3,x-1)$$

and we therefore obtain the fundamental equation

$$\begin{aligned} N_a(x) = N_a(x,\xi_1) + O\big(M_a(x,\xi_1,\xi_2)\big) + O\big(M_a(x,\xi_2,\xi_3)\big) \\ + O\big(M_a(x,\xi_3,x-1)\big)\,. \end{aligned} \tag{2.1}$$

2.3 Applications of Algebraic Number Theory

Let h be the largest positive integer such that a is a perfect h-th power. Since a is not a perfect square, h is odd. Also, by the unique factorization in $\mathbb{Z}$, a is also a perfect r-th power if and only if $r \mid h$. For any square-free integer k, we define

$$k_1 := \frac{k}{(h,k)}\,.$$

Since h is odd, then k and k_1 are either both even or both odd. Furthermore, the primes contributing to $P_a(x,k)$ are those primes p, relatively prime to a, for which the simultaneous conditions

$$\nu^q \equiv a \,(\mathrm{mod}\, p) \;\text{ has a solution } \nu \in \mathbb{Z}, \text{ and } p \equiv 1 \,(\mathrm{mod}\, q)$$

hold for every prime divisor q of k. Since we can always find a solution ν to the congruence $\nu^q \equiv a \,(\mathrm{mod}\, p)$ when $q \mid h$, we obtain the equivalent simultaneous conditions

$$\nu^{k_1} \equiv a \,(\mathrm{mod}\, p) \;\text{ has a solution } \nu \in \mathbb{Z}, \text{ and } p \equiv 1 \,(\mathrm{mod}\, k). \tag{2.2}$$

The proof of this equivalence only requires the knowledge that the group $(\mathbb{Z}/p\mathbb{Z})^*$ is a cyclic group of order $p-1$ and if $a \equiv g^n \,(\mathrm{mod}\, p)$ where g is a generator of $(\mathbb{Z}/p\mathbb{Z})^*$, then the order of a is $\dfrac{p-1}{(n,p-1)}$.

We now let $\mathbb{Q}$ denote the field of rational numbers and for any algebraic extension M over a field L, we indicate the degree of M over L by $[M:L]$. Consider the polynomial

$$u^{k_1} - a\,.$$

Over $\overline{\mathbb{Q}}$, $u^{k_1} - a$ factorizes as

$$\prod_{j=0}^{k_1-1} \left(u - \zeta_{k_1}^j a^{1/k_1}\right)\,, \tag{2.3}$$

where $\zeta_{k_1} = e^{2\pi i/k_1}$. Since k is square-free and $(k_1, h) = 1$, we see that the constant term in any combination of linear factors from (2.3) cannot be a rational number, thus $u^{k_1} - a$ is irreducible over $\mathbb{Q}$. This shows that the field

$$F_k = \mathbb{Q}(\sqrt[k_1]{a})$$

has degree k_1 over $\mathbb{Q}$. Moreover, the prime factors of the discriminant of F_k divide either a or k_1 [19, p.45-47]. Similarly, if we let $\sqrt[k]{1}$ denote a primitive k-th root of unity, then the cyclotomic extension

$$Z_k = \mathbb{Q}(\sqrt[k]{1})$$

has degree $\phi(k)$ and its discriminant is composed entirely of prime divisors of k [19, p.52]. We now want to state a lemma having as a goal to help us provide an equivalent, but more useful formulation of condition (2.2).

Lemma 2.3.1. *If there exists an integer $\nu \in \mathbb{Z}$ such that $\nu^{k_1} \equiv a \,(mod\, p)$ and $p \equiv 1 \,(mod\, k_1)$, then the congruence $y^{k_1} \equiv a \,(mod\, p)$ has exactly k_1 distinct solutions in $(\mathbb{Z}/p\mathbb{Z})^*$.*

Proof. Since $p \equiv 1 \,(\mathrm{mod}\, k_1)$ and $(\mathbb{Z}/p\mathbb{Z})^*$ is cyclic, there exists a unique subgroup $H_{k_1} \leq (\mathbb{Z}/p\mathbb{Z})^*$ of size k_1. Then the set νH_{k_1} provides k_1 distinct solutions to the congruence $y^{k_1} \equiv a \,(\mathrm{mod}\, p)$. Since $\mathbb{Z}/p\mathbb{Z}$ is field, νH_{k_1} is the complete set of solutions. □

Then, by a famous principle due to Dedekind, the condition

$$\nu^{k_1} \equiv a \,(\mathrm{mod}\, p) \text{ having exactly } k_1 \text{ distinct roots}$$

is equivalent to the assertion that both $p \nmid k_1$ and p factorizes in F_k as a product of k_1 distinct linear prime ideals. Similarly, the statement

$$p \equiv 1 \,(\mathrm{mod}\, k)$$

is equivalent to the condition that $p \nmid k$ and p factorizes in Z_k as a product of $\phi(k)$ distinct linear prime ideals. From the above and Lemma 2.3.1, we obtain the following theorem.

Theorem 2.3.1. *(Dedekind) Let p be a positive prime integer. Then p satisfies condition (2.2) if and only if $p \nmid k$ and p factorizes totally in the Galois extension*

$$L_k = \mathbb{Q}(\sqrt[k_1]{a}, \sqrt[k]{1})$$

as a product of distinct linear prime ideals. Observe that L_k is Galois over $\mathbb{Q}$ since it is the splitting field of the polynomial $(u^{k_1} - a)(u^k - 1) \in \mathbb{Q}[u]$ and $\mathbb{Q}$ is a field of characteristic zero.

We now wish to prove Theorem 2.3.1 by presenting the main steps of the proof of Dedekind's principle quoted above. We refer to *Appendix A* and to [19] for any omitted details in the following discussion.

Preliminary Results

Theorem 2.3.2. *Let K be an algebraic number field of degree n over $\mathbb{Q}$. Suppose that there exists an element $\theta \in K$ such that $\mathcal{O}_K = \mathbb{Z}[\theta]$. Let $f(x)$ be the minimal polynomial of θ over $\mathbb{Z}[x]$. Let p be a rational prime, and suppose*

$$f(x) \equiv f_1(x)^{e_1} \cdots f_g(x)^{e_g} \ (mod p)\,,$$

where each $f_i(x)$ is irreducible in $\mathbb{F}_p[x]$. Then $p\mathcal{O}_K = \wp_1^{e_1} \cdots \wp_g^{e_g}$ where $\wp_i = (p, f_i(\theta))$ are prime ideals, with norm $|\wp_i| = p^{deg f_i}$.

Proof. See *Appendix A*. □

Theorem 2.3.3. *If in the previous theorem, given $\theta \in K$, we do not assume that $\mathcal{O}_K = \mathbb{Z}[\theta]$ but instead that $p \nmid [\mathcal{O}_K : \mathbb{Z}[\theta]]$ and $[\mathbb{Q}(\theta) : \mathbb{Q}] = [K : \mathbb{Q}] = n$, then the same result holds.*

Proof. See *Appendix A*. □

Theorem 2.3.4. *Let K be Galois extension of degree n over $\mathbb{Q}$ with Galois group $Gal(K/\mathbb{Q})$. Let p be a rational prime. Then $Gal(K/\mathbb{Q})$ acts transitively on the set of prime ideals of $\mathcal{O}_K$ lying above p. As a corollary, any two prime ideals of $\mathcal{O}_K$ lying above p have the same ramification index and inertial degree.*

Proof. This is a straightforward application of the Chinese remainder theorem. For complete details, see [21, p.54]. □

Proof of Theorem 2.3.1

We are now in a position to prove Theorem 2.3.1. Recall that $Z_k = \mathbb{Q}(\sqrt[k]{1})$ and assuming $p \equiv 1 \,(\mathrm{mod}\, k)$ implies that the polynomial $x^k - 1$ splits completely modulo p into a product of distinct linear factors. Moreover, it can be shown that the ring of integers of Z_k, denoted by $\mathcal{O}_{Z_k}$, is equal to $\mathbb{Z}[\sqrt[k]{1}]$. Thus applying Theorem 2.3.2 yields the desired result in this case. In the other case, we consider $F_k = \mathbb{Q}(\sqrt[k_1]{a})$ and from (2.2) and Lemma 2.3.1 we have that $f_k(x) := x^{k_1} - a$ splits completely modulo p into a product of distinct linear factors. Before proceeding any further, we need the following definitions.

Definition 2.3.1. *Let K be an algebraic number field of degree n over $\mathbb{Q}$ and $\sigma_1, \sigma_2, ..., \sigma_n$ be the embeddings of K into $\mathbb{C}$. We define the discriminant of K over $\mathbb{Q}$ as*

$$d_{K/\mathbb{Q}} := det(\sigma_i(\omega_j))^2 \,,$$

where $\omega_1, \omega_2, ..., \omega_n$ is an integral basis for K over $\mathbb{Q}$. We can generalize the notion of a discriminant for arbitrary elements of K. If we let $\alpha \in K$, we define

$$d_{K/\mathbb{Q}}(\alpha) := det(\sigma_i(\alpha^{j-1}))^2 \,.$$

For more details, see [19, p.45].

If we let $m_k = [\mathcal{O}_{F_k} : \mathbb{Z}[\sqrt[k_1]{a}\,]]$ and $\zeta_i\,'s$ be the distinct roots of $f_k(x)$, then it can be shown that (see [19, p.210])

$$d_{F_k/\mathbb{Q}}(\sqrt[k_1]{a}) = m_k^2 d_{F_k/\mathbb{Q}} \,,$$

but

$$\begin{aligned} d_{F_k/\mathbb{Q}}(\sqrt[k_1]{a}) &= (-1)^{\frac{k_1(k_1-1)}{2}} \prod_{i=1}^{k_1} f_k'(\zeta_i) \\ &= (-1)^{\frac{k_1(k_1-1)}{2}} \prod_{i=1}^{k_1} k_1(\zeta_i)^{k_1-1} \\ &= (-1)^{\frac{k_1(k_1-1)}{2}} k_1^{k_1}(\pm a)^{k_1-1} \\ &= \pm k_1^{k_1} a^{k_1-1} \,. \end{aligned}$$

Since p does not divide k_1 nor a by assumption, this implies that $p \nmid m_k$. Applying Theorem 2.3.3 yields the desired result.

We can now show that p splits completely over L_k. Let $\wp$ be one of the prime ideals of Z_k lying above p. Since $x^{k_1} - a$ factors into k_1 distinct linear polynomials modulo p, it follows that the same holds true modulo $\wp$. It is then possible to show that p does not divide $d_{L_k/Z_k}(\sqrt[k_1]{a})$ and hence that the same is true for $\wp$. We are now in a similar situation as the one where we proved that p splits completely over F_k. From an analogous argument, we conclude that $\wp$ splits completely in L_k. Therefore p possesses at least one linear prime factor in L_k, but L_k is a Galois extension and so from Theorem 2.3.4, p splits completely over L_k.

Conversely, we assume that $p \nmid k$ and that p splits completely in the Galois enten-

sion L_k. Our goal is to show that this implies the conditions in (2.2) are satisfied. Using the fact that the ramification index and the inertial degree are both multiplicative in towers of field extensions, we can see that p splits completely over both Z_k and F_k. Since p splits completely over F_k, it follows from Theorem 2.3.3 that the polynomial $x^{k_1} - a$ factors as a product of distinct linear polynomials modulo p. This proves that the congruence $\nu^{k_1} \equiv a \pmod{p}$ is indeed solvable. It remains to show that $p \equiv 1 \pmod{k}$. To do this we need to consider the number field Z_k. We first recall that

$$x^k - 1 = \prod_{d \,|\, k} \Phi_d(x)$$

where $\Phi_d(x)$ is the d-th cyclotomic polynomial. If we let ζ_d be a primitive d-th root of unity for any divisor d of k, then $Z_k = \mathbb{Q}(\zeta_k)$ and $\mathbb{Q}(\zeta_d)$ is a subfield of Z_k for any $d \,|\, k$. Knowing that p splits completely over $\mathbb{Q}(\zeta_d)$ for any $d \,|\, k$, Theorem 2.3.3 implies that $\Phi_d(x)$ factors into a product of distinct linear polynomials modulo p for any $d \,|\, k$. We thus conclude that the congruence $x^k \equiv 1 \pmod{p}$ has exactly k solutions modulo p. This we know implies that $p \equiv 1 \pmod{k}$, hence completing the proof.

□

In order to make use of the Theorem 2.3.1, we need to determine the degree

$$n_k = [L_k : \mathbb{Q}] \tag{2.4}$$

of L_k over $\mathbb{Q}$. To accomlish this task, it is enough to compute $[L_k : Z_k]$ because

$$[L_k : \mathbb{Q}] = [L_k : Z_k][Z_k : \mathbb{Q}] = [L_k : Z_k]\,\phi(k)\,. \tag{2.5}$$

The following theorem is crucial in determining $[L_k : Z_k]$.

Theorem 2.3.5. *Keeping the above setup and notation, we have that* $[L_k : Z_k] \,|\, k_1$.

Proof. See *Appendix A*. □

We now let

$$k_1 = m_k[L_k : Z_k]\,. \tag{2.6}$$

Then, if q is a prime factor of m_k, we have that $[Z_k(\sqrt[q]{a}) : Z_k] \in \{1, q\}$ and that

$$[Z_k(\sqrt[q]{a}) : Z_k] \,|\, \frac{k_1}{m_k} = [L_k : Z_k]$$

since $Z_k \subseteq Z_k(\sqrt[q]{a}) \subseteq L_k$. As $(k_1/m_k, q) = 1$, because k_1 is square-free, it implies that $[Z_k(\sqrt[q]{a}) : Z_k] = 1$ therefore $\sqrt[q]{a} \in Z_k$. Our next goal is to show that $m_k \in \{1, 2\}$, but we need the following two lemmas first.

Lemma 2.3.2. *Let L and K be Galois extensions of finite degree over $\mathbb{Q}$ such that $K \subseteq L$. If $L/\mathbb{Q}$ is an abelian extension, so is $K/\mathbb{Q}$.*

Proof. See [3, p.558-559]. □

Lemma 2.3.3. *Let q be an odd prime, and let a be an integer which is not a q-th power. Let K be a splitting field of the polynomial $u^q - a$ over $\mathbb{Q}$. Let α denote any q-th root of a and ζ be a primitive q-th root of unity. Then $K = \mathbb{Q}(\alpha, \zeta)$ and $[K : \mathbb{Q}] = q(q-1)$. Furthermore, we have that*

$$Gal(K/\mathbb{Q}) \simeq \left\{ \begin{bmatrix} b & c \\ 0 & 1 \end{bmatrix} \ : \ b \in (\mathbb{Z}/p\mathbb{Z})^*, \ c \in \mathbb{Z}/p\mathbb{Z} \right\}.$$

This isomorphism implies that $Gal(K/\mathbb{Q})$ is not an abelian group.

Proof. See [3, p.565-568, 582]. □

Let us now assume that there exists and odd prime q such that $q \,|\, m_k$. Then as we proved above, this implies that $\sqrt[q]{a} \in Z_k$. Thus we have $\mathbb{Q} \subset \mathbb{Q}(\sqrt[q]{a}, \sqrt[q]{1}) \subseteq Z_k$. Since $Z_k/\mathbb{Q}$ is a Galois and abelian extension, applying Lemma 2.3.2 shows that $\mathbb{Q}(\sqrt[q]{a}, \sqrt[q]{1})$, being a Galois extension over $\mathbb{Q}$, is also abelian. On the other hand, Lemma 2.3.3 implies that $\mathbb{Q}(\sqrt[q]{a}, \sqrt[q]{1})$ is not an abelian extension over $\mathbb{Q}$. This is a contradiction. Hence, m_k does not possess any odd prime divisors and since it is square-free it must be true that $m_k \in \{1, 2\}$. Observe also that m_k may equal 2 only when k_1 and k are both even. Furthermore, the fact that

$$L_k = Z_k(\sqrt[k_1/q]{a}, \sqrt[q]{a})$$

is a direct consequence of the Euclidean algorithm. This is because k_1 is square-free and thus implies that q and k_1/q are coprime. It thus follows from this observation and the above discussion that $m_k = 2$ if and only if $\sqrt{a} \in Z_k$. To reformulate this condition in a more appropriate fashion, let

$$a = \tilde{a} a_2^2$$

where $\tilde{a}$ is the square-free part of a and possibly negative. Let also D be a positive odd divisor of k other than 1. Then, from the theory of cyclotomic fields (see [3, p.567]), the only quadratic subfields of Z_k are of the form

$$\mathbb{Q}\left(\sqrt{\left(\frac{-1}{D}\right)D}\right),$$

we obtain that $m_k = 2$ if and only if $\tilde{a}|k$, $|\tilde{a}| > 1$, and $\tilde{a}$ is an odd integer with the same sign as the Legendre symbol

$$\left(\frac{-1}{|\tilde{a}|}\right) = \begin{cases} 1\,, & \text{if } |\tilde{a}| \equiv 1 \,(\mathrm{mod}\, 4) \\ -1\,, & \text{if } |\tilde{a}| \equiv 3 \,(\mathrm{mod}\, 4). \end{cases}$$

Moreover, $\tilde{a} \neq 1$ since a is not a square. Thus the above conditions are equivalent to $\tilde{a}|k$, $\tilde{a} \equiv 1 \,(\mathrm{mod}\, 4)$. We therefore reach the conclusion from (2.4), (2.5) and (2.6) that

$$n_k = \frac{k_1 \phi(k)}{\varepsilon(k)}\,, \tag{2.7}$$

where $\varepsilon(k)$ is given by

$$\varepsilon(k) = \begin{cases} 2\,, & \text{if } 2\tilde{a}|k \ \text{ and } \ \tilde{a} \equiv 1 \,(\mathrm{mod}\, 4) \\ 1\,, & \text{otherwise.} \end{cases} \tag{2.8}$$

2.4 Estimation of the Remainder Terms

In this section, we wish to estimate the remainder terms in equation (2.1). We first obtain upper bounds for the last two terms. To estimate the first one we observe that

$$M_a(x, \xi_2, \xi_3) \ \leq \sum_{\xi_2 < q \leq \xi_3} P_a(x, q)\,,$$

which can be seen by interchanging the order of summation from the sum on the right-hand side of the inequality.

Keeping only the condition $q \,|\, p - 1$ in $R(q, p)$, we obtain directly the following inequality

$$P_a(x, q) \ \leq \sum_{\substack{p \leq x \\ p \equiv 1 \,(\mathrm{mod}\, q)}} 1 \quad .$$

To bound the right-hand side above we need the following important theorem and elementary lemma.

Theorem 2.4.1. (Brun-Titchmarsh Theorem) *Let a and k be positive coprime integers and let x be a positive real number such that $k \leq x^\theta$ for some $\theta < 1$. Then, for any $\epsilon > 0$, there exists $x_0 = x_0(\epsilon) > 0$ such that*

$$\sum_{\substack{p \leq x \\ p \equiv a \,(mod\, k)}} 1 \quad \leq \frac{(2+\epsilon)x}{\phi(k)\log(2x/k)}$$

for all $x > x_0$ where ϕ is the Euler totient function.

Proof. See [5, p.125]. □

Lemma 2.4.1.

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

Proof. See [5, p.9]. □

From Theorem 2.4.1 we have

$$M_a(x, \xi_2, \xi_3) \;\ll\; \sum_{\xi_2 < q \leq \xi_3} \sum_{\substack{p \leq x \\ p \equiv 1 \,(\mathrm{mod}\ q)}} 1 \;\ll\; \sum_{\xi_2 < q \leq \xi_3} \frac{x}{(q-1)\log(x/q)}\,.$$

Then, since $\xi_2 < q \leq \xi_3$, we have

$$\sum_{\xi_2 < q \leq \xi_3} \frac{x}{(q-1)\log(x/q)} \;\ll\; \frac{x}{\log x} \sum_{\xi_2 < q \leq \xi_3} \frac{1}{q} \;\ll\; \frac{x}{\log^2 x} \sum_{\xi_2 < q \leq \xi_3} \frac{\log q}{q}\,.$$

Finally Lemma 2.4.1 yields that

$$\frac{x}{\log^2 x} \sum_{\xi_2 < q \leq \xi_3} \frac{\log q}{q} \;\ll\; \frac{x}{\log^2 x}\left(\log \frac{\xi_3}{\xi_2} + 1\right) \;\ll\; \frac{x \log\log x}{\log^2 x}\,,$$

the last inequality following from our choice of ξ_2 and ξ_3.

Therefore we can see that

$$M_a(x, \xi_2, \xi_3) = O\left(\frac{x \log\log x}{\log^2 x}\right). \tag{2.9}$$

Our next task is to provide an upper bound for $M_a(x, \xi_3, x-1)$. We first observe that the condition $R(q,p)$ implies that

$$a^{\frac{p-1}{q}} \equiv 1 \,(\mathrm{mod}\, p),$$

thus

$$a^{\frac{2(p-1)}{q}} \equiv 1 \,(\mathrm{mod}\, p).$$

Hence, since $q > \xi_3 = \sqrt{x}\log x$ and $p \leq x$, the prime numbers p counted by $M_a(x, \xi_3, x-1)$ must divide the positive product

$$\prod_{m<\frac{\sqrt{x}}{\log x}} (a^{2m}-1)\,.$$

Since $p \geq 2$, we have

$$2^{M_a(x,\xi_3,x-1)} \;<\; \prod_{m<\frac{\sqrt{x}}{\log x}} a^{2m}\,,$$

so

$$M_a(x,\xi_3,x-1) \;<\; \frac{2\log|a|}{\log 2} \sum_{m<\frac{\sqrt{x}}{\log x}} m \;=\; O\left(\frac{x}{\log^2 x}\right). \tag{2.10}$$

It now remains to evaluate $M_a(x,\xi_1,\xi_2)$. As in the derivation of the inequality for $M_a(x,\xi_2,\xi_3)$, we have

$$M_a(x,\xi_1,\xi_2) \leq \sum_{\xi_1<q\leq\xi_2} P_a(x,q)\,.$$

The sum $P_a(x,k)$ can now be expressed in terms of the familiar prime ideals counting function

$$\pi(x,k) \;:=\; \sum_{\substack{\mathfrak{p}\subseteq\mathcal{O}_{L_k}\\ \mathfrak{p}:\text{prime ideal}\\ N\mathfrak{p}\leq x}} 1\,.$$

We can write

$$\pi(x,k) = W(x,k) + W'(x,k)\,, \tag{2.11}$$

where $W(x,k)$ is the contribution to $\pi(x,k)$ coming from linear prime ideals that do not divide ak, and $W'(x,k)$ is the remaining contribution. Since L_k is a Galois extension over $\mathbb{Q}$, each rational prime p relatively prime to ak either has n_k linear prime ideal factors or has no such factor in L_k. In the latter case, for a prime ideal

$\mathfrak{p}$ lying above p, we have that $N\mathfrak{p} \geq p^2$. Thus

$$W(x,k) = n_k P_a(x,k) \tag{2.12}$$

and

$$W'(x,k) \leq n_k \omega(ak) + n_k \sum_{p^2 \leq x} 1 \, . \tag{2.13}$$

By (2.11), (2.12) and (2.13), we have

$$n_k P_a(x,k) = \pi(x,k) + O\left(n_k \omega(k)\right) + O\left(n_k \sqrt{x}\right) \, . \tag{2.14}$$

Theorem 2.4.2. *Assuming the generalized Riemann hypothesis for the field extensions $L_k/\mathbb{Q}$ and defining $li(x) := \displaystyle\int_2^\infty \frac{dt}{\log t}$, we have that*

$$\pi(x,k) = li(x) + O\left(n_k \sqrt{x} \log kx\right) \tag{2.15}$$

and combining this with (2.14) yields that

$$P_a(x,k) = \frac{1}{n_k} li(x) + O\left(\sqrt{x} \log kx\right) \, . \tag{2.16}$$

Proof. See [11]. □

It thus follows from Theorem 2.4.2 and our choice of ξ_2 that

$$\begin{aligned} M_a(x,\xi_1,\xi_2) &\ll \sum_{\xi_1 < q \leq \xi_2} \left(\frac{li(x)}{n_q} + \sqrt{x} \log x \right) \\ &\ll \sum_{\xi_1 < q \leq \xi_2} \left(\frac{li(x)}{q(q-1)} + \sqrt{x} \log x \right) \\ &\ll \frac{x}{\log x} \sum_{q > \xi_1} \frac{1}{q^2} \; + \; \sqrt{x} \log x \sum_{q \leq \xi_2} 1 \\ &\ll \frac{x}{\xi_1 \log x} \; + \; \sqrt{x} \log x \frac{\xi_2}{\log \xi_2} \\ &\ll \frac{x}{\log^2 x} \, . \end{aligned} \tag{2.17}$$

We gather from (2.1), (2.9), (2.10) and (2.17) that

$$N_a(x) = N_a(x,\xi_1) + O\left(\frac{x \log\log x}{\log^2 x} \right) . \tag{2.18}$$

2.5 GRH Implies Artin's Primitive Root Conjecture

In this section, we estimate the main term in (2.1). This leads us to conclude that assuming the generalized Riemann hypothesis allows us to solve completely Artin's primitive root conjecture. We begin by expressing $N_a(x, \xi_1)$ in terms of $P_a(x, k)$. The reason is that it is possible to characterize the prime integers p counted in the sum $P_a(x, k)$ in terms of conditions formulated in the language of algebraic number theory. Now, we can see from a direct application of the inclusion-exclusion principle that

$$N_a(x, \xi_1) \;=\; \sum_{\substack{k \geq 1 \text{ , square-free} \\ \forall \text{ prime } q,\, q \,|\, k \Rightarrow q \leq \xi_1}} \mu(k) P_a(x, k) \tag{2.19}$$

where μ denotes the standard Möbius function. We now wish to give an upper bound for k, but in order to do so we need the following lemma.

Lemma 2.5.1. *If we let*

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

for any real number $x \geq 2$, then $\theta(x) \leq 2x \log 2$.

Proof. See [19, p.248]. □

Then, from Lemma 2.5.1, we have that

$$k \leq \prod_{q \leq \xi_1} q = e^{\theta(\xi_1)} \leq e^{(2 \log 2)\xi_1} \leq e^{2\xi_1} = x^{\frac{1}{3}} . \tag{2.20}$$

It then follows from (2.19) and (2.16) that

$$N_a(x, \xi_1) \quad = \sum_{\substack{k \geq 1 \text{ , square-free} \\ \forall \text{ prime } q,\, q \,|\, l' \Rightarrow q \leq \xi_1}} \mu(k) \left(\frac{1}{n_k} li(x) + O\left(\sqrt{x} \log (xk)\right) \right) . \tag{2.21}$$

Moreover, from (2.20), we have

$$\sum_{\substack{k \geq 1, \text{ square-free} \\ \forall \text{ prime } q,\, q \,|\, k \Rightarrow q \leq \xi_1}} \sqrt{x} \log x \;\ll\; \sum_{k \leq x^{\frac{1}{3}}} \sqrt{x} \log x \;\ll\; \frac{x}{\log^2 x} . \tag{2.22}$$

Combining (2.21) and (2.22) yields

$$N_a(x,\xi_1) \;=\; li(x) \sum_{\substack{k\geq 1\,,\ \text{square-free} \\ \forall\ \text{prime } q,\ q\,|\,k \Rightarrow q \leq \xi_1}} \frac{\mu(k)}{n_k} \;+\; O\left(\frac{x}{\log^2 x}\right). \tag{2.23}$$

Lemma 2.5.2. *Let ϕ be the Euler totient function. Then there exist positive constants A, B such that*

$$\sum_{1\leq n\leq x} \frac{1}{\phi(n)} \;=\; A\log x + B + O\left(\frac{\log x}{x}\right).$$

Proof. See [5, p.109]. □

Corollary 2.5.1. *From partial summation, we have*

$$\sum_{n>y} \frac{1}{n\phi(n)} \;\ll\; \frac{\log y}{y}.$$

Now, since all square-free integers $k \leq \xi_1$ satisfy the condition:

$$\forall\ \text{prime } q,\ q\,|\,k \;\Rightarrow\; q \;\leq\; \xi_1\,,$$

we obtain from (2.7) and Corollary 2.5.1 that

$$\begin{aligned} N_a(x,\xi_1) &= li(x)\sum_{k=1}^{\infty}\frac{\varepsilon(k)\mu(k)}{k_1\phi(k)} + O\left(li(x)\sum_{k>\xi_1}\frac{1}{k\phi(k)}\right) + O\left(\frac{x}{\log^2 x}\right) \\ &= li(x)\sum_{k=1}^{\infty}\frac{\varepsilon(k)\mu(k)}{k_1\phi(k)} + O\left(\frac{x\log\xi_1}{\xi_1\log x}\right) + O\left(\frac{x}{\log^2 x}\right) \\ &= \frac{x}{\log x}\sum_{k=1}^{\infty}\frac{\varepsilon(k)\mu(k)}{k_1\phi(k)} + O\left(\frac{x\log\xi_1}{\xi_1\log x}\right) + O\left(\frac{x}{\log^2 x}\right), \end{aligned}$$

since $li(x) = x/\log x + O\left(x/\log^2 x\right)$.

Since $\xi_1 = \dfrac{1}{6}\log x$,

$$N_a(x,\xi_1) \;=\; A(a)\frac{x}{\log x} + O\left(\frac{x\log\log x}{\log^2 x}\right), \tag{2.24}$$

where $A(a) = \displaystyle\sum_{k=1}^{\infty}\frac{\varepsilon(k)\mu(k)}{k_1\phi(k)}$.

We now verify the positivity of $A(a)$ and we break our analysis into two cases.

Case 1: $\tilde{a} \not\equiv 1 \pmod 4$.

It follows from (2.8) that $\varepsilon(k)$ is always equal to 1. Expanding $A(a)$ into its Euler product form, we have that

$$\begin{aligned} A(a) &= \sum_{k=1}^{\infty} \frac{\mu(k)(h,k)}{k\phi(k)} = \prod_{q|h}\left(1-\frac{1}{q-1}\right)\prod_{q\nmid h}\left(1-\frac{1}{q(q-1)}\right) \qquad (2.25)\\ &= C(h)\,,\ \text{say}\,. \end{aligned}$$

It thus follows that $A(a) > 0$ when $\tilde{a} \not\equiv 1 \pmod 4$.

Case 2: $\tilde{a} \equiv 1 \pmod 4$.

We firts obtain from (2.8) that

$$\begin{aligned} A(a) &= \sum_{k\not\equiv 0 \,(\mathrm{mod}\, 2|\tilde{a}|)} \frac{\mu(k)}{k_1\phi(k)} \;+\; 2\sum_{k\equiv 0 \,(\mathrm{mod}\, 2|\tilde{a}|)} \frac{\mu(k)}{k_1\phi(k)}\\ &= \sum_{k=1}^{\infty} \frac{\mu(k)}{k_1\phi(k)} \;+\; \sum_{k\equiv 0 \,(\mathrm{mod}\, 2|\tilde{a}|)} \frac{\mu(k)(h,k)}{k\phi(k)}\\ &= C(h) \;+\; \sum_{k\equiv 0 \,(\mathrm{mod}\, 2|\tilde{a}|)} \frac{\mu(k)(h,k)}{k\phi(k)}\,. \qquad (2.26) \end{aligned}$$

For a square-free integer k as in (2.26), letting $k = 2|\tilde{a}|k'$ with $(k', 2|\tilde{a}|) = 1$, the remaining sum in (2.26) is equal to

$$\sum_{(k',2|\tilde{a}|)=1} \frac{\mu(2|\tilde{a}|k')(h,2|\tilde{a}|k')}{2|\tilde{a}|k'\phi(2|\tilde{a}|k')} = \frac{\mu(2|\tilde{a}|)(h,2|\tilde{a}|)}{2|\tilde{a}|\phi(2|\tilde{a}|)} \sum_{(k',2|\tilde{a}|)=1} \frac{\mu(k')(h,k')}{k'\phi(k')}\,.$$

Observe that the right-hand sum over k' can be written as a product that is very similar to $C(h)$ except that it is lacking the factors corresponding to the prime divisors of $2|\tilde{a}|$.

Thus, since $1/\phi(2|\tilde{a}|) = \prod_{q|2\tilde{a}} 1/(q-1)$,

$$\begin{aligned}
\frac{A(a)}{C(h)} &= 1 + \frac{\mu(2|\tilde{a}|)(h,2|\tilde{a}|)}{2|\tilde{a}|\phi(2|\tilde{a}|)} \prod_{\substack{q|h \\ q|2\tilde{a}}} \left(1 - \frac{1}{q-1}\right)^{-1} \prod_{\substack{q\nmid h \\ q|2\tilde{a}}} \left(1 - \frac{1}{q(q-1)}\right)^{-1} \\
&= 1 + \frac{\mu(2|\tilde{a}|)(h,2|\tilde{a}|)}{2|\tilde{a}|\phi(2|\tilde{a}|)} \prod_{\substack{q|h \\ q|2\tilde{a}}} \left(\frac{q-1}{q-2}\right) \prod_{\substack{q\nmid h \\ q|2\tilde{a}}} \left(\frac{q(q-1)}{q^2-q-1}\right) \\
&= 1 + \frac{\mu(2|\tilde{a}|)(h,2|\tilde{a}|)}{2|\tilde{a}|} \prod_{\substack{q|h \\ q|2\tilde{a}}} \left(\frac{1}{q-2}\right) \prod_{\substack{q\nmid h \\ q|2\tilde{a}}} \left(\frac{q}{q^2-q-1}\right).
\end{aligned}$$

Now, since h and $|\tilde{a}|$ are odd and $\displaystyle\prod_{\substack{q|h \\ q|\tilde{a}}} q \cdot \prod_{q|\tilde{a}} \frac{1}{q} \cdot \prod_{\substack{q\nmid h \\ q|\tilde{a}}} q = 1$, we have

$$\begin{aligned}
\frac{A(a)}{C(h)} &= 1 + \mu(2)\mu(|\tilde{a}|) \prod_{\substack{q|h \\ q|\tilde{a}}} q \cdot \prod_{q|\tilde{a}} \frac{1}{q} \cdot \prod_{\substack{q|h \\ q|2\tilde{a}}} \frac{1}{q-2} \cdot \prod_{\substack{q\nmid h \\ q|\tilde{a}}} \frac{1}{q^2-q-1} \cdot \prod_{\substack{q\nmid h \\ q|\tilde{a}}} q \\
&= 1 - \mu(|\tilde{a}|) \prod_{\substack{q|h \\ q|\tilde{a}}} \frac{1}{q-2} \cdot \prod_{\substack{q\nmid h \\ q|\tilde{a}}} \frac{1}{q^2-q-1}. \qquad (2.27)
\end{aligned}$$

Observing that

$$\prod_{\substack{q|h \\ q|\tilde{a}}} \frac{1}{q-2} \cdot \prod_{\substack{q\nmid h \\ q|\tilde{a}}} \frac{1}{q^2-q-1} \leq 1$$

and that this product can possibly be equal to 1 only if $|\tilde{a}| = 3$, in which case $\mu(|\tilde{a}|) = -1$, shows that $A(a) > 0$ when $\tilde{a} \equiv 1 \pmod 4$. Hence, the value of $A(a)$ is always strictly positive.

Combining (2.18), (2.24), (2.25) and (2.27), we obtain our main theorem.

Theorem 2.5.1. *If the generalized Riemann hypothesis holds for the class of Dedekind zeta functions over Galois extensions of the type $\mathbb{Q}(\sqrt[k_1]{b}, \sqrt[k]{1})$, where b is an integer, k is a square-free integer and $k_1 \mid k$, then we have:*

Let a be a non-zero integer that is not equal to 1, -1, or to a perfect square. Let $N_a(x)$ be the number of primes p up to x for which a is a primitive root modulo p. We denote by $\tilde{a}$ the square-free part of a and h the largest integer

such that a is a perfect h-th power. Let

$$C(h) := \prod_{q|h}\left(1-\frac{1}{q-1}\right)\prod_{q\nmid h}\left(1-\frac{1}{q(q-1)}\right).$$

Then, if $\tilde{a} \not\equiv 1\,(mod\ 4)$, we have

$$N_a(x) = C(h)\frac{x}{\log x} + O\left(\frac{x\log\log x}{\log^2 x}\right),$$

while if $\tilde{a} \equiv 1\,(mod\ 4)$, we have

$$N_a(x) = C(h)\left(1-\mu(|\tilde{a}|)\prod_{\substack{q|h\\ q|\tilde{a}}}\left(\frac{1}{q-2}\right)\prod_{\substack{q\nmid h\\ q|\tilde{a}}}\left(\frac{1}{q^2-q-1}\right)\right)\frac{x}{\log x} + O\left(\frac{x\log\log x}{\log^2 x}\right).$$

Corollary 2.5.2. *If a is a non-zero integer not equal to 1, -1, or to a perfect square, then there are infinitely many prime integers p such that a is a primitive root modulo p.*

2.6 Numerical Evidence

The following table, where we set x to be the 50 000$^{\text{th}}$ prime, provides numerical support for Theorem 2.5.1.

Value of a	$\mathbf{N_a(x)}$	$\mathbf{A(a)\cdot li(x)}$	|Error|
2	18 701	18 724	23
3	18 761	18 724	37
5	19 699	19 709	10
7	18 687	18 724	37
8	11 225	11 235	10
11	18 772	18 724	48
13	18 863	18 845	18
17	18 796	18 793	3
$5^3\cdot 2^6$	11 844	11 826	18

Chapter 3

An Average Result for Artin's Primitive Root Conjecture

3.1 Introduction

In this chapter, we present the work of P. J. Stephens on average results for Artin's primitive root conjecture. It is to be noted that in this case, the results obtained are unconditional. As before, we denote by $N_a(x)$ the number of primes $p \leq x$ for which a is a primitive root modulo p and we let

$$A = \prod_{p\,:\,\text{prime}} \left(1 - \frac{1}{p(p-1)}\right).$$

Our goal is to prove, in order, the following two theorems and corollary.

Theorem 3.1.1. *If*

$$N > \exp\left(4(\log x \log\log x)^{\frac{1}{2}}\right), \tag{3.1}$$

then

$$\frac{1}{N}\sum_{1<a\leq N} N_a(x) = A\,li(x) + O\left(\frac{x}{(\log x)^D}\right), \tag{3.2}$$

where D is an arbitrary constant greater than 1.

Theorem 3.1.2. *If*

$$N > \exp\left(6(\log x \log\log x)^{\frac{1}{2}}\right), \tag{3.3}$$

then

$$\frac{1}{N}\sum_{1<a\leq N} \Big(N_a(x) - A\,li(x)\Big)^2 \ll \frac{x^2}{(\log x)^E}, \tag{3.4}$$

where E is an arbitrary constant greater than 2.

As a consequence of Theorem 3.1.2, we have

Corollary 3.1.1. *Let* $\mathbf{E}$ *be the set of integers* $a \leq N$ *for which*

$$|N_a(x) - A\,li(x)| > \varepsilon\, li(x)\,,$$

for a given $\varepsilon > 0$. *Assuming that* $N > \exp\big(6(\log x \log\log x)^{\frac{1}{2}}\big)$, *then*

$$\#\mathbf{E} \;=\; O\left(\frac{N}{\varepsilon^2(\log x)^F}\right) \;=\; o(N)\,,$$

where F is an arbitrary positive constant.

3.2 Preliminary Lemmas

If we define $M_p(N)$ in the following way:

$$M_p(N) := \#\{1 < a \leq N \,|\, \langle a\rangle \,= (\mathbb{Z}/p\mathbb{Z})^*\}\,,$$

then, by reordering any of the two given summations, we see that

$$\frac{1}{N}\sum_{1<a\leq N} N_a(x) = \frac{1}{N}\sum_{p\leq x} M_p(N)\,. \tag{3.5}$$

Furthermore, since there are precisely $\phi(p-1)$ integers which are primitive roots mod p in any interval of length p, we see that

$$M_p(N) = \phi(p-1)\left(\frac{N}{p} + O(1)\right)\,.$$

Hence, it follows from the above equality, (3.5) and the prime number theorem that

$$\begin{aligned}
\frac{1}{N}\sum_{1<a\leq N} N_a(x) \;&=\; \frac{1}{N}\sum_{p\leq x}\left(\phi(p-1)\left(\frac{N}{p}+O(1)\right)\right)\\
&=\; \sum_{p\leq x}\frac{\phi(p-1)}{p} + O\left(\frac{1}{N}\sum_{p\leq x}\phi(p-1)\right)\\
&=\; \sum_{p\leq x}\frac{\phi(p-1)}{p} + O\left(\frac{x^2}{N\log x}\right)\,. \qquad (3.6)
\end{aligned}$$

To evaluate the main term of (3.6), we need the following theorems and lemma.

Theorem 3.2.1.

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

Proof. See [5, p.10]. □

Theorem 3.2.2. *(Merten's Theorem)*

$$\prod_{p\le x}\left(1-\frac{1}{p}\right) \;=\; \frac{e^{-\gamma}}{\log x}\left(1+O\left(\frac{1}{\log x}\right)\right)$$

as $x\to\infty$ and where γ is Euler's constant.

Proof. See [5, p.65]. □

Theorem 3.2.3. *(Siegel-Walfisz) Letting $\pi(x,l,d) := \#\{p\le x\,|\,p\equiv l\,(mod\,d)\}$, we have*

$$\pi(x,l,d)=\frac{li(x)}{\phi(d)}+O\left(\frac{x}{(\log x)^C}\right), \tag{3.7}$$

provided that $(l,d)=1$ and $d\le(\log x)^B$ where B and C are arbitrary positive constants.

Proof. See [24]. □

Lemma 3.2.1.

$$\sum_{p\le x}\frac{\phi(p-1)}{p}=A\,li(x)+O\left(\frac{x}{(\log x)^D}\right) \tag{3.8}$$

where D is an arbitrary constant greater than 1.

Proof. We first write

$$\begin{aligned}
S_1 \;&=\; \sum_{p\le x}\frac{\phi(p-1)}{p} \;=\; \sum_{p\le x}\phi(p-1)\left(\frac{1}{p-1}-\frac{1}{p(p-1)}\right)\\
&=\; \sum_{p\le x}\frac{\phi(p-1)}{p-1} \;+\; \sum_{p\le x}\frac{1}{p}\left(\frac{\phi(p-1)}{p-1}\right)\\
&=\; \sum_{p\le x}\frac{\phi(p-1)}{p-1}+O\left(\sum_{p\le x}\frac{1}{p}\right). \qquad (3.9)
\end{aligned}$$

Now, by reordering the following summation, we obtain

$$\begin{aligned}\sum_{p\le x}\frac{\phi(p-1)}{p-1} &= \sum_{p\le x}\sum_{d\,|\,p-1}\frac{\mu(d)}{d} = \sum_{d\le x}\frac{\mu(d)}{d}\pi(x,1,d)\\ &= \sum_{d\le(\log x)^B}\frac{\mu(d)}{d}\pi(x,1,d) + O\left(\sum_{(\log x)^B<d\le x}\frac{\pi(x,1,d)}{d}\right) \quad (3.10)\end{aligned}$$

where $B>0$ is an arbitrary constant.

We first estimate

$$\begin{aligned}\sum_{(\log x)^B<d\le x}\frac{\pi(x,1,d)}{d} &\le \sum_{d>(\log x)^B}\frac{1}{d}\sum_{\substack{1<n\le x\\ n\equiv 1\,(\mathrm{mod}\,d)}}1 \le \sum_{d>(\log x)^B}\frac{1}{d}\left(\frac{x}{d}\right)\\ &\le \frac{x}{(\log x)^B}\,. \quad (3.11)\end{aligned}$$

Setting $l=1$ in Theorem 3.2.3, we see from Theorem 3.2.1, (3.9), (3.10) and (3.11) that

$$\begin{aligned}S_1 &= \sum_{d\le(\log x)^B}\frac{\mu(d)}{d}\left(\frac{li(x)}{\phi(d)} + O\left(\frac{x}{(\log x)^C}\right)\right) + O\left(\frac{x}{(\log x)^B}\right)\\ &= li(x)\sum_{d\le(\log x)^B}\frac{\mu(d)}{d\phi(d)} + O\left(\frac{x}{(\log x)^C}\sum_{d\le(\log x)^B}\frac{1}{d}\right) + O\left(\frac{x}{(\log x)^B}\right)\\ &= li(x)\sum_{d=1}^{\infty}\frac{\mu(d)}{d\phi(d)} + O\left(li(x)\sum_{d>(\log x)^B}\frac{1}{d\phi(d)}\right) + O\left(\frac{x}{(\log x)^{C-1}}\right)\\ &\quad + O\left(\frac{x}{(\log x)^B}\right)\,. \quad (3.12)\end{aligned}$$

Moreover,

$$\sum_{d=1}^{\infty}\frac{\mu(d)}{d\phi(d)} = \prod_p\left(1-\frac{1}{p(p-1)}\right) = A\,.$$

To estimate the second sum in (3.12), we use Theorem 3.2.2 to get

$$\frac{d}{\phi(d)} = \prod_{p\,|\,d}\left(1-\frac{1}{p}\right)^{-1} \le \prod_{p\le d}\left(1-\frac{1}{p}\right)^{-1} = \frac{\log d}{e^{-\gamma}}\left(1+O\left(\frac{1}{\log d}\right)\right)^{-1} \ll \log d\,.$$

Then

$$\sum_{d>(\log x)^B} \frac{1}{d\phi(d)} \ll \sum_{d>(\log x)^B} \frac{\log d}{d^2} \ll \frac{\log\log x}{(\log x)^B}.$$

Hence

$$li(x) \sum_{d>(\log x)^B} \frac{1}{d\phi(d)} \ll \frac{x}{(\log x)^B}.$$

Finally, choosing B and C sufficiently large we have that

$$S_1 = Ali(x) + O\left(\frac{x}{(\log x)^D}\right),$$

where D is an arbitrary constant greater than 1. This completes the proof. □

We now wish to observe that if we were to take $N \geq x^{1+\varepsilon}$ for any given $\varepsilon > 0$, then combining (3.6) and Lemma 3.2.1 would give us a proof of equation (3.2). This is unsatisfactory since we wish to average over 1 to N for as small an N as possible. A finer analysis is therefore required.

Theorem 3.2.4. *(The Large Sieve Inequality) For each character χ modulo k, we let*

$$S(\chi) = \sum_{n=M+1}^{M+N} a_n \chi(n),$$

where a_n is any complex number, $M \in \mathbb{Z}$ and $N \in \mathbb{Z}^+$. Then

$$\sum_{k\leq K} \sideset{}{'}\sum_{\chi \bmod k} |S(\chi)|^2 \ll (K^2 + N) \sum_{n=M+1}^{M+N} |a_n|^2,$$

where $\sum'$ denotes summation over primitive characters, $K \in \mathbb{Z}^+$ and $k \geq 1$.

Proof. See *Appendix B.* □

Let us now define $\tau'_r(a)$ to be the number of r-ordered tuples such that the product over all entries is equal to a and each entry does not exceed N and may possibly be equal to 1. We now wish to prove the following two lemmas, which provide upper bounds for the first and second moment of $\tau'_r(a)$.

Lemma 3.2.2. *With the above definition, we have*

$$\sum_{a\le N^r} \tau_r'(a) \;\le\; N^r\big(\log(eN^{r-1})\big)^{r-1}. \tag{3.13}$$

Proof. We proceed to prove the lemma by induction on r. If $r=1$, then $\tau_r'(a)=1$ for all a, and the result follows. We now suppose the lemma holds for $r=k\ge 1$. Then

$$\begin{aligned}\sum_{a\le N^{k+1}} \tau_{k+1}'(a) \;&\le\; \sum_{a\le N^{k+1}} \sum_{\substack{d\,|\,a\\ d\le N^k}} \tau_k'(d) \;=\; \sum_{d\le N^k} \tau_k'(d) \sum_{\substack{a\le N^{k+1}\\ d\,|\,a}} 1\\ &\le\; N^{k+1}\sum_{d\le N^k}\frac{\tau_k'(d)}{d}. \end{aligned}\tag{3.14}$$

Observe that (3.14) implies the result when $r=2$ since $\tau_1'(d)=1$ for all d. We may therefore assume that $k\ge 2$. Our goal is now to find an upper bound for (3.14). From (3.14), we may write

$$\frac{1}{N^{k+1}}\sum_{a\le N^{k+1}}\tau_{k+1}'(a) \;\le\; \sum_{d\le N^k}\frac{\tau_k'(d)}{d} \tag{3.15}$$

for $k\ge 1$. Then, from partial summation and (3.15), we have that

$$\begin{aligned}\sum_{d\le N^k}\frac{\tau_k'(d)}{d} \;&=\; \left(\sum_{d\le N^k}\tau_k'(d)\right)\frac{1}{N^k} + \sum_{d\le N^k-1}\left(\int_d^{d+1}\frac{dt}{t^2}\right)\sum_{\delta\le d}\tau_k'(\delta)\\ &=\; \frac{1}{N^k}\sum_{d\le N^k}\tau_k'(d) + \sum_{d\le N^k-1}\left(\frac{1}{d(d+1)}\right)\sum_{\delta\le d}\tau_k'(\delta)\\ &\le\; \sum_{t\le N^{k-1}}\frac{\tau_{k-1}'(t)}{t} + \sum_{d\le N^k-1}\left(\frac{1}{d(d+1)}\right)\sum_{\delta\le d}\sum_{\substack{t\,|\,\delta\\ t\le N^{k-1}}}\tau_{k-1}'(t).\end{aligned}$$

Since

$$\sum_{\delta\le d}\sum_{\substack{t\,|\,\delta\\ t\le N^{k-1}}}\tau_{k-1}'(t) \;=\; \sum_{t\le N^{k-1}}\tau_{k-1}'(t)\sum_{\substack{\delta\le d\\ t\,|\,d}}1 \;\le\; \sum_{t\le N^{k-1}}\tau_{k-1}'(t)\frac{d}{t},$$

then

$$\sum_{d\le N^k}\frac{\tau_k'(d)}{d} \;\le\; \sum_{t\le N^{k-1}}\frac{\tau_{k-1}'(t)}{t}\left(1+\sum_{d\le N^k-1}\frac{1}{d+1}\right) \;\le\; \log\left(eN^k\right)\sum_{t\le N^{k-1}}\frac{\tau_{k-1}'(t)}{t}\,.$$

Thus

$$\sum_{d\le N^k}\frac{\tau_k'(d)}{d} \;\le\; \log\left(eN^k\right)\sum_{t\le N^{k-1}}\frac{\tau_{k-1}'(t)}{t}$$

holds for $k\ge 2$. Applying this inequality recursively shows that

$$\sum_{d\le N^k}\frac{\tau_k'(d)}{d} \;\le\; \big(\log\left(eN^k\right)\big)^k\,.$$

By (3.15), it follows that

$$\sum_{a\le N^{k+1}}\tau_{k+1}'(a) \;\le\; N^{k+1}\big(\log\left(eN^k\right)\big)^k,$$

which concludes the proof of the lemma. □

Lemma 3.2.3. *We have that*

$$\sum_{a\le N^r}(\tau_r'(a))^2 \;\le\; N^r\left(\log(eN^{r-1})\right)^{r^2-1}\,. \tag{3.16}$$

Proof. We prove the result by induction on r. If $r=1$, the lemma follows since both sides of (3.16) are equal to N. We now suppose the lemma is true for $r=k\ge 1$.

First,

$$\begin{aligned}
\sum_{a\le N^{k+1}}(\tau_{k+1}'(a))^2 &\le \sum_{a\le N^{k+1}}\left(\sum_{\substack{d\,|\,a\\ d\le N^k}}\tau_k'(d)\right)^2\\
&\le \sum_{d\le N^k}\tau_k'(d)\sum_{l\le N^k}\tau_k'(l)\sum_{\substack{a\le N^{k+1}\\ d\,|\,a,\,l\,|\,a}}1\\
&\le \sum_{d\le N^k}\tau_k'(d)\sum_{l\le N^k}\tau_k'(l)\frac{N^{k+1}}{[d,l]}\\
&= \sum_{d\le N^k}\tau_k'(d)\sum_{l\le N^k}\tau_k'(l)N^{k+1}\frac{(d,l)}{dl}\\
&= N^{k+1}\sum_{d\le N^k}\frac{\tau_k'(d)}{d}\sum_{l\le N^k}\frac{\tau_k'(l)}{l}(d,l)\,. \qquad (3.17)
\end{aligned}$$

Since $n = \sum_{d\,|\,n} \phi(d)$ and $\tau'_k(mt) \le \tau'_k(m)\tau'_k(t)$, we have

$$\begin{aligned}
\sum_{l\le N^k} \frac{\tau'_k(l)}{l}(d,l) &= \sum_{l\le N^k} \frac{\tau'_k(l)}{l} \sum_{\substack{t\,|\,d\\ t\,|\,l}} \phi(t) \\
&= \sum_{\substack{t\le N^k\\ t\,|\,d}} \phi(t) \sum_{\substack{l\le N^k\\ t\,|\,l}} \frac{\tau'_k(l)}{l} \\
&= \sum_{\substack{t\le N^k\\ t\,|\,d}} \phi(t) \sum_{m\le \frac{N^k}{t}} \frac{\tau'_k(mt)}{mt} \\
&\le \sum_{\substack{t\le N^k\\ t\,|\,d}} \phi(t) \sum_{m\le \frac{N^k}{t}} \frac{\tau'_k(m)}{m}\frac{\tau'_k(t)}{t}\,. \qquad (3.18)
\end{aligned}$$

By (3.17) and (3.18), we have

$$\begin{aligned}
\sum_{a\le N^{k+1}} (\tau'_{k+1}(a))^2 &\le N^{k+1} \sum_{d\le N^k} \frac{\tau'_k(d)}{d} \sum_{\substack{t\le N^k\\ t\,|\,d}} \frac{\tau'_k(t)\phi(t)}{t} \sum_{m\le \frac{N^k}{t}} \frac{\tau'_k(m)}{m} \\
&\le N^{k+1} \sum_{d\le N^k} \frac{\tau'_k(d)}{d} \sum_{\substack{t\le N^k\\ t\,|\,d}} \tau'_k(t) \sum_{m\le N^k} \frac{\tau'_k(m)}{m} \\
&= \left(N^{k+1} \sum_{m\le N^k} \frac{\tau'_k(m)}{m}\right) \sum_{d\le N^k} \frac{\tau'_k(d)}{d} \sum_{\substack{t\le N^k\\ t\,|\,d}} \tau'_k(t)\,.
\end{aligned}$$

However,

$$\begin{aligned}
\sum_{d\le N^k} \frac{\tau'_k(d)}{d} \sum_{\substack{t\le N^k\\ t\,|\,d}} \tau'_k(t) &= \sum_{t\le N^k} \frac{\tau'_k(t)}{t} \sum_{n\le \frac{N^k}{t}} \frac{\tau'_k(nt)}{n} \\
&\le \sum_{t\le N^k} \frac{\tau'_k(t)}{t} \sum_{n\le \frac{N^k}{t}} \frac{\tau'_k(n)\tau'_k(t)}{n} \\
&\le \sum_{t\le N^k} \frac{(\tau'_k(t))^2}{t} \sum_{n\le N^k} \frac{\tau'_k(n)}{n}\,.
\end{aligned}$$

Therefore

$$\sum_{a\leq N^{k+1}}(\tau'_{k+1}(a))^2 \;\leq\; N_{k+1}\left(\sum_{u\leq N^k}\frac{\tau'_k(u)}{u}\right)^2\sum_{t\leq N^k}\frac{(\tau'_k(t))^2}{t}\,.$$

From the proof of Lemma 3.2.2, we have that

$$\left(\sum_{u\leq N^k}\frac{\tau'_k(u)}{u}\right)^2 \;\leq\; \big(\log(eN^k)\big)^{2k}$$

and combining the induction hypothesis with partial summation yields that

$$\sum_{t\leq N^k}\frac{(\tau'_k(t))^2}{t} \;\leq\; \big(\log(eN^k)\big)^{k^2}\,.$$

Hence

$$\sum_{a\leq N^{k+1}}(\tau'_{k+1}(a))^2 \;\leq\; N^{k+1}\big(\log(eN^k)\big)^{(k+1)^2-1}\,,$$

which completes the proof. □

Lemma 3.2.4.

$$\sum_{k\leq K}\sideset{}{'}\sum_{\chi \bmod k}\left|\sum_{a\leq N}\chi(a)\right|^{2r} \;\ll\; (K^2+N^r)N^r\big(\log(eN^{r-1})\big)^{r^2-1}\,, \tag{3.19}$$

where $\sum'$ *denotes summation over primitive characters,* $K\in\mathbb{Z}^+$ *and* $k\geq 1$.

Proof.

$$\begin{aligned}\sum_{k\leq K}\sideset{}{'}\sum_{\chi \bmod k}\left|\sum_{a\leq N}\chi(a)\right|^{2r} &= \sum_{k\leq K}\sideset{}{'}\sum_{\chi \bmod k}\left|\left(\sum_{a\leq N}\chi(a)\right)^r\right|^2\\ &= \sum_{k\leq K}\sideset{}{'}\sum_{\chi \bmod k}\left|\sum_{b\leq N^r}\tau'_r(b)\chi(b)\right|^2\,.\end{aligned}$$

Applying Lemma 3.2.1 and 3.2.3, we ontain

$$\begin{aligned}\sum_{k\leq K}\sideset{}{'}\sum_{\chi \bmod k}\left|\sum_{a\leq N^r}\tau'_r(a)\chi(a)\right|^2 &\ll (K^2+N^r)\sum_{a\leq N^r}|\tau'_r(a)|^2\\ &\leq (K^2+N^r)N^r\big(\log(eN^{r-1})\big)^{r^2-1}\,.\end{aligned}$$

This completes the proof. □

3.3 Proof of Theorem 3.1.1.

In this section, we prove an asymptotic formula for the quantity

$$\frac{1}{N}\sum_{1<a\leq N} N_a(x)\,.$$

For a prime p, we define the following function:

$$t_p(a) := \begin{cases} 1\,, & \text{if } a \text{ is a primitive root mod } p \\ 0\,, & \text{otherwise.} \end{cases}$$

Then we can see that

$$M_p(N) = \sum_{a\leq N} t_p(a)\,. \tag{3.20}$$

The following lemma enables us to rewrite $t_p(a)$ in a more effective way.

Lemma 3.3.1. *For any character χ modulo p, if we define $c(\chi)$ by*

$$c(\chi) := \frac{1}{p-1}\sum_{1<b<p}{}'' \chi(b) \tag{3.21}$$

where $\sum''$ means that we are summing over primitive roots b modulo p, then we have that

$$t_p(a) = \sum_{\chi \bmod p} c(\chi)\chi(a)\,. \tag{3.22}$$

Proof.

$$\sum_{\chi \bmod p} c(\chi)\chi(a) \;=\; \sum_{\chi \bmod p}\left(\frac{1}{p-1}\sum_{1<b<p}{}''\chi(b)\right)\chi(a) \;=\; \frac{1}{p-1}\sum_{1<b<p}{}''\sum_{\chi \bmod p}\chi(ba)\,,$$

but from the orthogonality relation for characters, we know that

$$\sum_{\chi \bmod p}\chi(ba) \;=\; \begin{cases} p-1\,, & \text{if } ba \equiv 1 \,(\mathrm{mod}\, p) \\ 0\,, & \text{otherwise.} \end{cases}$$

Since b is by assumption a primitive root modulo p, if $a \equiv b^{-1} \,(\mathrm{mod}\, p)$, then a is also a primitive root modulo p.

Hence

$$\frac{1}{p-1}\sideset{}{''}\sum_{1<b<p}\sum_{\chi \bmod p}\chi(ba)=\begin{cases}1\,, & \text{if } a \text{ is a primitive root mod } p\\ 0\,, & \text{otherwise.}\end{cases}$$

This completes the proof. □

It is also clear that if χ_0 denotes the principal character modulo p, then

$$c(\chi_0)=\frac{\phi(p-1)}{p-1}\,. \tag{3.23}$$

If $\chi \neq \chi_0$ we may express the sum in (3.21) in terms of Ramanujan sums (see [20, p.6-7]) to obtain

$$|c(\chi)| \leq \frac{1}{ord(\chi)}\,, \tag{3.24}$$

where $ord(\chi)$ is the smallest positive integer d such that $\chi^d=\chi_0$. As a final remark, note that if χ is a non-principal character modulo p, then it is automatically a primitive character modulo p. Thus, from (3.5), (3.20) and (3.22), we have

$$\begin{aligned}\frac{1}{N}\sum_{a\leq N}N_a(x) &= \frac{1}{N}\sum_{p\leq x}\sum_{a\leq N}\sum_{\chi \bmod p}c(\chi)\chi(a) \;=\; \frac{1}{N}\sum_{p\leq x}\sum_{\chi \bmod p}c(\chi)\sum_{a\leq N}\chi(a)\\ &= \frac{1}{N}\sum_{p\leq x}\sum_{\chi=\chi_0}c(\chi)\sum_{a\leq N}\chi(a) \;+\; \frac{1}{N}\sum_{p\leq x}\sum_{\substack{\chi \bmod p\\ \chi\neq\chi_0}}c(\chi)\sum_{a\leq N}\chi(a)\,.\end{aligned}$$

Then, by (3.23), (3.24) and since $1/(p-1)=1/p+1/p(p-1)$,

$$\begin{aligned}\frac{1}{N}\sum_{a\leq N}N_a(x) &= \frac{1}{N}\sum_{p\leq x}\frac{\phi(p-1)}{p-1}\left(N-\left\lfloor\frac{N}{p}\right\rfloor\right)+O\left(\frac{1}{N}\sum_{p\leq x}\sideset{}{'}\sum_{\chi \bmod p}\frac{1}{ord(\chi)}\left|\sum_{a\leq N}\chi(a)\right|\right)\\ &= \sum_{p\leq x}\frac{\phi(p-1)}{p-1}-\frac{1}{N}\sum_{p\leq x}\frac{\phi(p-1)}{p-1}\left\lfloor\frac{N}{p}\right\rfloor+O\left(\frac{S_2}{N}\right)\\ &= \sum_{p\leq x}\frac{\phi(p-1)}{p}+\sum_{p\leq x}\frac{\phi(p-1)}{p-1}\left(\frac{1}{p}-\frac{1}{N}\left\lfloor\frac{N}{p}\right\rfloor\right)+O\left(\frac{S_2}{N}\right)\,,\end{aligned}$$

where

$$S_2=\sum_{p\leq x}\sideset{}{'}\sum_{\chi \bmod p}\frac{1}{ord(\chi)}\left|\sum_{a\leq N}\chi(a)\right|\,.$$

Then,

$$\begin{aligned}\sum_{p\le x}\frac{\phi(p-1)}{p-1}\left(\frac{1}{p}-\frac{1}{N}\left\lfloor\frac{N}{p}\right\rfloor\right) &= \sum_{p\le x}\frac{\phi(p-1)}{p-1}\left(\frac{1}{p}-\frac{1}{N}\left(\frac{N}{p}-\left\{\frac{N}{p}\right\}\right)\right)\\ &= \sum_{p\le x}\frac{\phi(p-1)}{p-1}\left(\frac{1}{N}\left\{\frac{N}{p}\right\}\right)\\ &= O\left(\frac{1}{N}\sum_{p\le x}1\right) = O\left(\frac{x}{N\log x}\right).\end{aligned}$$

Therefore

$$\frac{1}{N}\sum_{a\le N}N_a(x) = \sum_{p\le x}\frac{\phi(p-1)}{p}+O\left(\frac{x}{N\log x}\right)+O\left(\frac{S_2}{N}\right). \tag{3.25}$$

Lemma 3.3.2. *Let m be a positive integer. The group of multiplicative characters $\chi:(\mathbb{Z}/m\mathbb{Z})^*\mapsto\mathbb{C}^*$ is isomorphic to $(\mathbb{Z}/m\mathbb{Z})^*$. As a consequence, if $m=p$ is a prime number, then for any divisor d of $p-1$, there are exactly $\phi(d)$ characters of order d.*

Proof. See [6, p.29]. □

We now wish to evaluate S_2. In order to do so, we apply Hölder's inequality, Lemma 3.2.4 and Lemma 3.3.2 to obtain

$$\begin{aligned}S_2^{2r} &\le \left(\sum_{p\le x}{\sum_{\chi \bmod p}}'\left(\frac{1}{ord(\chi)}\right)^{2r/(2r-1)}\right)^{2r-1}\sum_{p\le x}{\sum_{\chi \bmod p}}'\left|\sum_{a\le N}\chi(a)\right|^{2r}\\ &\ll \left(\sum_{p\le x}{\sum_{\chi \bmod p}}'\left(\frac{1}{ord(\chi)}\right)^{2r/(2r-1)}\right)^{2r-1}(x^2+N^r)N^r\left(\log\left(eN^{r-1}\right)\right)^{r^2-1}\\ &\le \left(\sum_{p\le x}\sum_{d\,|\,p-1}\frac{\phi(d)}{d^{2r/(2r-1)}}\right)^{2r-1}(x^2+N^r)N^r\left(\log\left(eN^{r-1}\right)\right)^{r^2-1}\\ &\le \left(\sum_{p\le x}\sum_{d\,|\,p-1}\frac{\phi(d)}{d}\right)^{2r-1}(x^2+N^r)N^r\left(\log\left(eN^{r-1}\right)\right)^{r^2-1}.\end{aligned} \tag{3.26}$$

Before we can proceed with our evaluation of S_2, we need to state the following result.

Lemma 3.3.3. *(Titchmarsh)*

If we let $\tau(n)$ denote the number of positive divisors of $n \in \mathbb{Z}^+$, then we have that

$$\sum_{p \leq x} \tau(p-1) \leq c_3\, x\,,$$

where c_3 is some positive real constant.

Proof. See [20, p.413]. □

Corollary 3.3.1.

$$\sum_{p \leq x} \sum_{d \,|\, p-1} \frac{\phi(d)}{d} \;\leq\; \sum_{p \leq x} \sum_{d \,|\, p-1} 1 \;=\; \sum_{p \leq x} \tau(p-1) \;\leq\; c_3\, x$$

We see from Corollary 3.3.1 that

$$S_2{}^{2r} \ll (c_3\, x)^{2r-1}(x^2 + N^r)N^r \left(\log\left(eN^{r-1}\right)\right)^{r^2-1}. \tag{3.27}$$

Dividing the above inequality by N^{2r} and then taking the $2r$-th root yields

$$\frac{S_2}{N} \ll x^{1-1/2r} \left(\frac{x^2}{N^r} + 1\right)^{1/2r} \left(\log\left(eN^{r-1}\right)\right)^{(r^2-1)/2r}. \tag{3.28}$$

Our next objective is to choose a value of r that minimizes the right-hand side of the above inequality. If N is very large with respect to x, then the term

$$\left(\log\left(eN^{r-1}\right)\right)^{(r^2-1)/2r}$$

becomes problematic and so we must require r to be equal to 1 in this case. On the other hand, if N is relatively small compared to x, then the term

$$\left(\frac{x^2}{N^r} + 1\right)^{1/2r}$$

is now problematic and so we must require r to be larger and at least greater than 2. From this argument, we now let

$$r = \left\lfloor \frac{2\log x}{\log N} \right\rfloor + 1\,.$$

It follows directly that $N^{r-1} \leq x^2 < N^r$. With this choice of r, we obtain from

(3.28) that

$$\frac{S_2}{N} \ll x^{1-1/2r} \left(\log\left(ex^2\right)\right)^{(r^2-1)/2r} . \tag{3.29}$$

If $N > x^2$, then $r = 1$ and by (3.29) we obtain

$$\frac{S_2}{N} \ll x^{\frac{1}{2}} . \tag{3.30}$$

We now assume that $N \leq x^2$, hence

$$\frac{2\log x}{\log N} \geq 1$$

and $r \geq 2$. We wish to show that

$$\frac{r^2-1}{2r} \leq \frac{3}{2}\frac{\log x}{\log N} .$$

First, we notice that

$$\frac{r^2-1}{2r} \;\leq\; \frac{r}{2} \;=\; \frac{1}{2}\left\lfloor \frac{2\log x}{\log N} \right\rfloor + \frac{1}{2} \;\leq\; \frac{\log x}{\log N} + \frac{1}{2} \;\leq\; \frac{3}{2}\frac{\log x}{\log N} ,$$

assuming that $\log x/\log N \geq 1$.

If $\dfrac{\log x}{\log N} < 1$, then we must have that

$$1 \leq \frac{2\log x}{\log N} < 2 ,$$

since $r \geq 2$. This actually implies that $r = 2$ and

$$\frac{3}{4} \leq \frac{3}{2}\frac{\log x}{\log N} \leq \frac{3}{2} .$$

Moreover, when $r = 2$, we have that

$$\frac{r^2-1}{2r} = \frac{3}{4}$$

and so

$$\frac{r^2-1}{2r} \leq \frac{3}{2}\frac{\log x}{\log N}$$

in this case as well.

Assuming that (3.1) holds, this implies that

$$\begin{aligned}\frac{S_2}{N} &\ll x^{1-1/2r}\left(\log\left(ex^2\right)\right)^{(3\log x)/(2\log N)} \\ &\ll x^{1-1/2r}\left(\log\left(ex^2\right)\right)^{(3(\log x)^{\frac{1}{2}})/(8(\log\log x)^{\frac{1}{2}})}\,.\end{aligned}$$

Now, since

$$-\frac{1}{2r}\log x+\frac{3}{8}\left(\frac{\log x}{\log\log x}\right)^{\frac{1}{2}}\log\log\left(ex^2\right)<-\frac{1}{4}\log N+\frac{3}{4}(\log x\log\log x)^{\frac{1}{2}}\,,$$

we obtain that

$$\begin{aligned}\frac{S_2}{N} &\ll xe^{-\frac{1}{4}\log N}e^{\frac{3}{4}(\log x\log\log x)^{\frac{1}{2}}} \\ &\leq xN^{-\frac{1}{4}}N^{\frac{3}{16}}\,,\ \text{since } e^{4(\log x\log\log x)^{\frac{1}{2}}}<N \\ &= xN^{-\frac{1}{16}}\,. \qquad (3.31)\end{aligned}$$

Finally, combining (3.30) and (3.31) yields

$$\frac{S_2}{N}\ll\frac{x}{(\log x)^D}\,, \qquad (3.32)$$

for any arbitrary $D>1$ provided (3.1) holds. Substituting (3.32) and the result of Lemma 3.2.1 into (3.25) completes the proof of Theorem 3.1.1.

3.4 Proof of Theorem 3.1.2.

In this section, we provide an upper bound for the quantity

$$\frac{1}{N}\sum_{1<a\leq N}\Big(N_a(x)-A\,li(x)\Big)^2\,.$$

We let p and q denote positive prime integers and we define

$$M_{p,q}(N):=\#\{a\leq N\mid <a>=(\mathbb{Z}/p\mathbb{Z})^*\ \text{and}\ <a>=(\mathbb{Z}/q\mathbb{Z})^*\}\,.$$

With the above definition and by Theorem 3.1.1, we can write

$$\begin{aligned} T &= \frac{1}{N}\sum_{a\le N}\left(N_a(x)-A\,li(x)\right)^2 \\ &= \frac{1}{N}\left(\sum_{p\le x}\sum_{q\le x}M_{p,q}(N)-2A\,li(x)\sum_{p\le x}M_p(N)+NA^2\left(li(x)\right)^2\right) \\ &= \frac{1}{N}\sum_{p\le x}\sum_{q\le x}M_{p,q}(N)-2A\,li(x)\left(A\,li(x)+O\left(\frac{x}{(\log x)^D}\right)\right)+A^2\left(li(x)\right)^2 \\ &= \frac{1}{N}\sum_{p\le x}\sum_{q\le x}M_{p,q}(N)-A^2\left(li(x)\right)^2+O\left(\frac{x^2}{(\log x)^E}\right), \end{aligned} \tag{3.33}$$

where $E=D+1$.

Write

$$\sum_{p\le x}\sum_{q\le x}M_{p,q}(N)=\sum_{p\le x}M_p(N)+\sum_{p\le x}\sum_{\substack{q\le x\\ q\ne p}}M_{p,q}(N)\,. \tag{3.34}$$

Since (3.3) holds, we can apply Theorem 3.1.1 to obtain

$$\sum_{p\le x}M_p(N)\ll N\,li(x)\,. \tag{3.35}$$

Also, by (3.22),

$$\begin{aligned} \sum_{p\le x}\sum_{\substack{q\le x\\ q\ne p}}M_{p,q}(N) &= \sum_{p\le x}\sum_{\substack{q\le x\\ q\ne p}}\sum_{a\le N}t_p(a)t_q(a) \\ &= \sum_{p\le x}\sum_{\substack{q\le x\\ q\ne p}}\sum_{\chi_1 \bmod p}\sum_{\chi_2 \bmod q}c(\chi_1)c(\chi_2)\sum_{a\le N}\chi_1(a)\chi_2(a) \\ &= T_1+2T_2+T_3\,, \end{aligned}$$

where

$$T_1=\sum_{p\le x}\sum_{\substack{q\le x\\ q\ne p}}\sum_{\substack{\chi_1 \bmod p\\ \chi_1=\chi_0}}\sum_{\substack{\chi_2 \bmod q\\ \chi_2=\chi_0}}c(\chi_1)c(\chi_2)\sum_{a\le N}\chi_1(a)\chi_2(a)\,,$$

$$T_2=\sum_{p\le x}\sum_{\substack{q\le x\\ q\ne p}}\sum_{\substack{\chi_1 \bmod p\\ \chi_1=\chi_0}}\sum_{\substack{\chi_2 \bmod q\\ \chi_2\ne\chi_0}}c(\chi_1)c(\chi_2)\sum_{a\le N}\chi_1(a)\chi_2(a)\,,$$

and

$$T_3=\sum_{p\le x}\sum_{\substack{q\le x\\ q\ne p}}\sum_{\substack{\chi_1 \bmod p\\ \chi_1\ne\chi_0}}\sum_{\substack{\chi_2 \bmod q\\ \chi_2\ne\chi_0}}c(\chi_1)c(\chi_2)\sum_{a\le N}\chi_1(a)\chi_2(a)\,.$$

Therefore

$$\sum_{p\le x}\sum_{q\le x} M_{p,q}(N) \ll N\, li(x) + T_1 + 2T_2 + T_3\,. \tag{3.36}$$

Applying the result of Lemma 3.2.1, we obtain

$$\begin{aligned}
T_1 &= \sum_{\substack{p,q\le x\\ p\neq q}} \frac{\phi(p-1)}{p-1}\frac{\phi(q-1)}{q-1}\left(\lfloor N\rfloor - \left\lfloor\frac{N}{p}\right\rfloor - \left\lfloor\frac{N}{q}\right\rfloor + \left\lfloor\frac{N}{pq}\right\rfloor\right)\\
&= \sum_{\substack{p,q\le x\\ p\neq q}} \frac{\phi(p-1)}{p-1}\frac{\phi(q-1)}{q-1}\left(N\frac{(p-1)(q-1)}{pq} + O(1)\right)\\
&= N\sum_{\substack{p,q\le x\\ p\neq q}} \frac{\phi(p-1)}{p}\frac{\phi(q-1)}{q} + O\left(\frac{x^2}{(\log x)^2}\right)\\
&= N\left(\sum_{p\le x}\frac{\phi(p-1)}{p}\right)^2 + O(N\, li(x)) + O\left(\frac{x^2}{(\log x)^2}\right)\\
&= NA^2\,(li(x))^2 + O\left(\frac{Nx^2}{(\log x)^E}\right) + O\left(\frac{x^2}{(\log x)^2}\right)\,.
\end{aligned} \tag{3.37}$$

Also, since

$$\chi_0\chi_2(a) = \begin{cases} \chi_2(a)\,, & \text{if } p\nmid a\\ 0\,, & \text{if } p\mid a,\end{cases}$$

we see that

$$T_2 = \sum_{p\le x}\frac{\phi(p-1)}{p-1}\sum_{\substack{q\le x\\ q\neq p}}\sum_{\chi_2\neq\chi_0} c(\chi_2)\sum_{\substack{a\le N\\ p\nmid a}}\chi_2(a)\,.$$

Let us recall that

$$S_2 = \sum_{p\le x}{\sum_{\chi \bmod p}}' \frac{1}{ord(\chi)}\left|\sum_{a\le N}\chi(a)\right|\,.$$

Then

$$\begin{aligned}
T_2 &\ll \sum_{p\le x}\sum_{q\le x}\sum_{\chi_2\neq\chi_0}\frac{1}{ord(\chi_2)}\left|\sum_{a\le N}\chi_2(a)\right| + \sum_{p\le x}\sum_{q\le x}\sum_{\chi_2\neq\chi_0}\frac{1}{ord(\chi_2)}\sum_{\substack{a\le N\\ p\mid a}} 1\\
&= S_2\sum_{p\le x} 1 + \sum_{p\le x}\sum_{q\le x}\sum_{\chi_2\neq\chi_0}\frac{1}{ord(\chi_2)}\frac{N}{p}\,.
\end{aligned} \tag{3.38}$$

Now, from Corollary 3.3.1 and Lemma 3.3.3, we have that

$$\begin{aligned}\sum_{p\le x}\sum_{q\le x}\sum_{\chi_2\neq\chi_0}\frac{1}{ord(\chi_2)}\frac{N}{p} &\le N\sum_{p\le x}\frac{1}{p}\sum_{q\le x}\sum_{d\,|\,q-1}\frac{\phi(d)}{d}\\ &\ll Nx\sum_{p\le x}\frac{1}{p}\,. \end{aligned}\tag{3.39}$$

Finally, by (3.32), (3.38), (3.39) and Theorem 3.2.1, we obtain that

$$T_2 \ll \frac{Nx^2}{(\log x)^E}+Nx\log\log x\,.\tag{3.40}$$

Before we can proceed to estimate T_3, we need to make the following remarks. First, if $p\neq q$ and χ_1, χ_2 are non-principal characters modulo p and q respectively, then $\chi_1\chi_2$ is a primitive character modulo pq. Furthermore,

$$ord(\chi_1\chi_2)=[ord(\chi_1),ord(\chi_2)]\le ord(\chi_1)ord(\chi_2)\,,$$

where $[m,n]$ denotes the least common multiple of m and n. Thus

$$T_3 \le \sum_{\substack{p,q\le x\\ p\neq q}}\sideset{}{'}\sum_{\chi \bmod pq}\frac{1}{ord(\chi)}\left|\sum_{a\le N}\chi(a)\right|.$$

Applying Hölder's inequality to T_3 and using Lemma 3.3.2 yields

$$\begin{aligned}T_3^{2r} &\le \left(\sum_{\substack{p,q\le x\\ p\neq q}}\sideset{}{'}\sum_{\chi \bmod pq}\left(\frac{1}{ord(\chi)}\right)^{2r/(2r-1)}\right)^{2r-1}\sum_{\substack{p,q\le x\\ p\neq q}}\sideset{}{'}\sum_{\chi \bmod pq}\left|\sum_{a\le N}\chi(a)\right|^{2r}\\ &\le \left(\sum_{\substack{p,q\le x\\ p\neq q}}\sum_{d\,|\,\phi(pq)}\frac{\phi(d)}{d}\right)^{2r-1}\sum_{\substack{p,q\le x\\ p\neq q}}\sideset{}{'}\sum_{\chi \bmod pq}\left|\sum_{a\le N}\chi(a)\right|^{2r}\\ &\le \left(\sum_{\substack{p,q\le x\\ p\neq q}}\tau(\phi(pq))\right)^{2r-1}\sum_{k\le x^2}\sideset{}{'}\sum_{\chi \bmod k}\left|\sum_{a\le N}\chi(a)\right|^{2r}. \end{aligned}\tag{3.41}$$

Applying Lemma 3.2.4 to (3.41) yields

$$\begin{aligned} T_3^{2r} &\ll \left(\sum_{\substack{p,q \le x \\ p \neq q}} \tau(\phi(pq)) \right)^{2r-1} (x^4 + N^r) N^r \left(\log \left(eN^{r-1} \right) \right)^{r^2-1} \\ &\le \left(\sum_{\substack{p,q \le x \\ p \neq q}} \tau(p-1)\tau(q-1) \right)^{2r-1} (x^4 + N^r) N^r \left(\log \left(eN^{r-1} \right) \right)^{r^2-1} , \end{aligned}$$

where we have also used the fact that $\phi(pq) = (p-1)(q-1)$ and $\tau(mn) \le \tau(m)\tau(n)$.

Since

$$\sum_{\substack{p,q \le x \\ p \neq q}} \tau(p-1)\tau(q-1) \ \le \ \left(\sum_{p \le x} \tau(p-1) \right)^2 ,$$

we see from Lemma 3.3.3 that

$$\begin{aligned} T_3^{2r} &\ll \left(\sum_{p \le x} \tau(p-1) \right)^{2(2r-1)} (x^4 + N^r) N^r \left(\log \left(eN^{r-1} \right) \right)^{r^2-1} \\ &\ll (c_3\, x)^{2(2r-1)} (x^4 + N^r) N^r \left(\log \left(eN^{r-1} \right) \right)^{r^2-1} . \end{aligned}$$

Following the same steps as in the proof of Theorem 3.1.1, we let

$$r = \left\lfloor \frac{4 \log x}{\log N} \right\rfloor + 1 ,$$

hence $N^{r-1} \le x^4 < N^r$. With this choice of r, we obtain directly that

$$\frac{T_3}{N} \ll x^{2-1/r} \left(\log \left(ex^4 \right) \right)^{(r^2-1)/(2r)} . \tag{3.42}$$

If $N > x^4$, then $r = 1$ and we have from (3.42) that

$$\frac{T_3}{N} \ll x . \tag{3.43}$$

If $N \leq x^4$, then $r \geq 2$ and again from (3.42)

$$\begin{aligned} \frac{T_3}{N} &\ll x^{2-1/r} \left(\log\left(ex^4\right)\right)^{(3\log x)/\log N} \\ &\ll x^2\, N^{-1/32} \\ &\ll \frac{x^2}{(\log x)^E}\,, \end{aligned} \tag{3.44}$$

for an arbitrary positive constant E provided (3.3) holds. The proof of the two inequalities preceeding (3.44) is essentially the same as the one which was provided in the proof of Theorem 3.1.1 while determining a proper upper bound for S_2/N.

Assuming 3.3 holds, we conclude from (3.27), (3.34), (3.35), (3.36), (3.37), (3.43) and (3.44) that

$$T \ll \frac{x^2}{(\log x)^E}\,,$$

for any arbitrary constant E greater than 2. This completes the proof of Theorem 3.1.2.

3.5 Proof of Corollary 3.1.1.

In this section, assuming $N > \exp\big(6(\log x \log\log x)^{\frac{1}{2}}\big)$, we provide an upper bound for the set

$$\mathbf{E} = \Big\{1 < a \leq N \,:\, |N_a(x) - A\, li(x)| > \varepsilon\, li(x)\Big\}\,,$$

for a given $\varepsilon > 0$. First note that for $a \in \mathbf{E}$, we have

$$\big(N_a(x) - A li(x)\big)^2 > \varepsilon^2 \left(li(x)\right)^2\,.$$

Thus

$$\frac{1}{N}\sum_{a\leq N}\big(N_a(x) - A\, li(x)\big)^2 > \frac{\#\mathbf{E}}{N}\varepsilon^2 (li(x))^2\,.$$

On the other hand, Theorem 3.1.2 yields

$$\frac{1}{N}\sum_{a\leq N}\big(N_a(x) - A\, li(x)\big)^2 \ll \frac{x^2}{(\log x)^E}\,.$$

Hence

$$\frac{\#\mathbf{E}}{N}\varepsilon^2 (li(x))^2 \ll \frac{x^2}{(\log x)^E}\,.$$

Therefore

$$\#\mathbf{E} \ll \frac{N}{\varepsilon^2 (\log x)^F} ,$$

where $F\,(= E-2)$ is an arbitrary positive constant.

As an example of Corollary 3.1.1, if we take

$$\varepsilon = \frac{1}{(\log x)^{D_1}} ,$$

where D_1 is an arbitrary positive constant, then we obtain that

$$|N_a(x) - A\,li(x)| < \frac{x}{(\log x)^{D_1+1}}$$

for all positive integers $a \leq N$ with at most

$$O\left(\frac{N}{(\log x)^{D_2}}\right)$$

exceptions, where D_2 is a positive constant depending on D_1, provided (3.3) holds.

Chapter 4

An Average Result for Composite Moduli

4.1 Introduction

The concept of a primitive root modulo a prime can be generalized. This was done by R. D. Carmichael [4]. Since the multiplicative group $(\mathbb{Z}/n\mathbb{Z})^*$ is not necessarily cyclic for a given positive integer n, he defined a *primitive λ-root modulo n* as any integer coprime to n having maximal multiplicative order. Therefore a primitive root for a prime p is a primitive λ-root modulo p. In analogy with the previous two chapters, we denote by $N_a(x)$ the number of positive integers up to x for which a is a primitive λ-root. Our goal in this chapter will be to demonstrate that the average of $N_a(x)$ oscillates. More precisely, we will prove the following theorem.

Theorem 4.1.1. *[14, Li] If we let*

$$N_a(x) := \{1 < n \leq x \,|\, a \text{ is a primitive } \lambda\text{-root modulo } n\}\,,$$

then

$$\limsup_{x\to\infty} \frac{1}{x^2} \sum_{1\leq a\leq x} N_a(x) \;>\; 0 \quad \text{and} \quad \liminf_{x\to\infty} \frac{1}{x^2} \sum_{1\leq a\leq x} N_a(x) \;=\; 0\,.$$

Our strategy will be to obtain information about the behaviour of the distribution function of $r(n)$, the density of primitive λ-roots modulo n. It does not seem possible to do so directly hence we will introduce an auxiliary function $\tilde{f}(n)$ and study this function instead. The reason $\tilde{f}(n)$ is useful is because we are able to find

the correct order of magnitude for its average order. Moreover, we can relate the first and second moment of $\tilde{f}(n)$ in a very nice way. This will allow us to extract information about the distribution function of $r(n)$. We will also use the very nice properties of the distribution function of $\phi(n)/n$, where ϕ is the Euler phi function. Finally, we will combine our knowledge of $r(n)$, $\tilde{f}(n)$ and $\phi(n)/n$ to show that the average of $N_a(x)$ exhibits extreme behaviour.

4.2 Preliminary Results and Definitions

Definition 4.2.1. *Let k be a positive integer and x a positive real number. We let $\log_k x$ denote the k-fold iteration of the natural logarithm of x whenever this makes sense, and zero otherwise.*

Definition 4.2.2. *Given $x \in \mathbb{R}$, we define the floor and ceiling of x by*

$$\lfloor x \rfloor := max\left\{n \in \mathbb{Z} \,|\, n \leq x\right\}$$

and

$$\lceil x \rceil := min\left\{n \in \mathbb{Z} \,|\, n \geq x\right\}.$$

Proposition 4.2.1. *Let $l_a(n)$ denote the multiplicative order of a modulo n when $gcd(a,n) = 1$. We define the Carmichael function in the following way:*

$$\lambda(n) := max\left\{l_a(n) \,|\, gcd(a,n) = 1 \text{ and } a \in \mathbb{Z}\right\}.$$

Moreover, if we use the notation $p^e \| n$ to mean that $p^e | n$ while $p^{e+1} \nmid n$, then we have

$$\lambda(n) = \operatorname*{lcm}_{p^e \| n}\left\{\lambda(p^e)\right\}$$

where $\lambda(p^e) = \phi(p^e)$ for all prime powers p^e, with the exception $\lambda(2^e) = \frac{1}{2}\phi(2^e)$ for $e \geq 3$.

Proof. See [20, p.23-24]. □

Definition 4.2.3. *Let a and n be coprime integers. If $l_a(n) = \lambda(n)$, we say that a is a primitive λ-root modulo n.*

Theorem 4.2.1. *From the structure theorem for finitely generated abelian groups, we can write*

$$(\mathbb{Z}/n\mathbb{Z})^* = \bigoplus_{q \,|\, \lambda(n)} \left(\left(\bigoplus_{j=1}^{\Delta_q(n)} C_{q^{e_q}} \right) \bigoplus H_q \right)$$

where $C_{q^{e_q}}$ is a cyclic group of order q^{e_q} and H_q is a direct sum of cyclic groups having order some power of q strictly less than e_q. Hence $\Delta_q(n)$ represents the number of direct summands whose order is q^{e_q}. Define

$$r(n) \;:=\; \frac{1}{\phi(n)}\#\{1 \le m \le n \,|\, m \text{ is a primitive } \lambda\text{-root modulo } n\}\,.$$

Then we have

$$r(n) \;=\; \prod_{q \,|\, \lambda(n)} \left(1 - \frac{1}{q^{\Delta_q(n)}}\right)\,. \tag{4.1}$$

Proof. Let us first consider a single prime q so that $q \,|\, \lambda(n)$. Since $C_{q^{e_q}}$ is cyclic, it possesses $\phi(q^{e_q})$ generators and hence $q^{e_q} - \phi(q^{e_q})$ elements not having maximal order. This implies that

$$\left(\bigoplus_{j=1}^{\Delta_q(n)} C_{q^{e_q}}\right) \bigoplus H_q$$

contains exactly $(q^{e_q} - \phi(q^{e_q}))^{\Delta_q(n)}|H_q|$ elements not having maximal order. This implies that

$$\begin{aligned}
r(n) &= \frac{1}{\phi(n)} \prod_{q \,|\, \lambda(n)} \left(q^{e_q\Delta_q(n)}|H_q| - (q^{e_q} - \phi(q^{e_q}))^{\Delta_q(n)}|H_q|\right) \\
&= \frac{1}{\phi(n)} \prod_{q \,|\, \lambda(n)} |H_q| q^{e_q\Delta_q(n)} \left(1 - \frac{1}{q^{\Delta_q(n)}}\right)\,, \text{ since } \phi(q^{e_q}) = q^{e_q} - q^{e_q - 1} \\
&= \frac{1}{\phi(n)} \prod_{q \,|\, \lambda(n)} |H_q| q^{e_q\Delta_q(n)} \prod_{q \,|\, \lambda(n)} \left(1 - \frac{1}{q^{\Delta_q(n)}}\right)\,.
\end{aligned}$$

Since

$$\frac{1}{\phi(n)} \prod_{q \,|\, \lambda(n)} |H_q| q^{e_q\Delta_q(n)} \;=\; 1\,,$$

this completes the proof. □

Theorem 4.2.2. *(Linnik's Theorem) There exists an absolute constant C such that, if gcd(a, q)=1, there is always a prime $p \equiv a (mod q)$ satisfying $p < q^C$.*

Proof. See [16]. □

Theorem 4.2.3. *(Merten's Theorem)*

$$\prod_{p<z} \left(1 - \frac{1}{p}\right) = \frac{e^{-\gamma}}{\log z} \left(1 + O\left(\frac{1}{\log z}\right)\right)$$

as $z \to \infty$ and where γ is Euler's constant.

Proof. See [5, p.67]. □

Theorem 4.2.4. *(Prime Number Theorem) If we denote by $\pi(x)$ the number of positive prime integers less than or equal to x, then*

$$\pi(x) = \frac{x}{\log x} + O\left(\frac{x}{\log^2 x}\right).$$

Proof. See [20, p.35-62] □

Lemma 4.2.1. *For any $t \in (0,1)$ the limit*

$$\lim_{x\to\infty} \frac{1}{x} \# \left\{ n \leq x \;\middle|\; \frac{\phi(n)}{n} \leq t \right\} := w(t)$$

exists. Moreover, the function $w(t)$ is continuous and strictly increasing in the interval $(0,1)$ with

$$\lim_{t\to 0^+} w(t) = 0 \quad \textit{and} \quad \lim_{t\to 1^-} w(t) = 1\,.$$

Proof. See [25, Theorem 1]. □

4.3 Applications of Sieve Theory

We now wish to present a few results that will be essential for our applications in the following sections. The reader may refer to [5] or [9] for more details. Let us first introduce some notation. Let $\mathcal{A}$ be the set of positive integers up to $x \geq 1$. Throughout the remaining of this chapter the set $\mathcal{A}$ will always be of this type. Let $\mathcal{P}$ be a set of primes. Define

$$P(z) \;:=\; \prod_{\substack{p\leq z \\ p\in\mathcal{P}}} p$$

and

$$W(z) \;:=\; \prod_{\substack{p\leq z \\ p\in\mathcal{P}}} \left(1 - \frac{1}{p}\right)$$

for any $z \leq x$. We are interested in estimating the counting function defined by

$$S(\mathcal{A},\mathcal{P},z) \;:=\; \sum_{\substack{n\in\mathcal{A} \\ \gcd(n,P(z))=1}} 1\,.$$

Lemma 4.3.1. *(See [14, p.105]) For any integer $k \geq 2$ and any $x \geq 2$, we have*

$$\sum_{\substack{p \leq x \\ p \equiv 1 \,(mod\, k)}} \frac{1}{p} \;=\; \frac{\log_2 x}{\phi(k)} + O\left(\frac{\log k}{\phi(k)}\right),$$

where the implied constant is uniform and effectively computable.

Theorem 4.3.1. *(See [9, Theorem 7.2]) Let $\mathcal{P}$ be a set of primes and assume that $2 \leq z \leq t$. Let $u = \log t / \log z$. Then we have*

$$\sum_{\substack{n \leq t \\ gcd(n, P(z))=1}} 1 \;=\; tW(z)\left(1 + O\Big(\exp\big(-1/2 \cdot u \log u\big)\Big) + O\Big(\exp\big(-\sqrt{\log t}\big)\Big)\right),$$

where $W(z)$ is defined as above and the implied constants are absolute.

Lemma 4.3.2. *Let $\mathcal{P}$ be a set of primes and assume that $\varepsilon > 0$ is a number depending on $\mathcal{P}$ such that*

$$\sum_{\substack{w < p \leq ew \\ p \in \mathcal{P}}} \frac{1}{p} \;\leq\; \frac{\varepsilon}{\log w}$$

for all $w \geq w_0$, w_0 depending on $\mathcal{P}$. Then if $z = \exp(\log t / \log_2^2 t) \geq w_0$ and $\varepsilon > \exp(-\log_2^2 t \log_3 t)$, we have

$$\sum_{\substack{m \leq t \\ gcd(m, P(t))=1}} \;=\; tW(z) + O\Big(\varepsilon t \log_3(t) W(z)\Big) + O\left(\frac{\varepsilon t}{\log_2^2 t}\right),$$

where the implied constants are absolute and $W(z)$ is defined as above.

Proof. First, we have the equality

$$\sum_{\substack{m \leq t \\ \gcd(m, P(t))=1}} 1 \;=\; \sum_{\substack{m \leq t \\ \gcd(m, P(z))=1}} 1 \;+\; O\left(\sum_{\substack{m \leq t \\ \gcd(m, P(z))=1 \\ \gcd(m, P(t))>1}} 1\right). \tag{4.2}$$

Recalling our choice of z, we see that the error term in (4.2) is bounded above by

$$\sum_{\substack{z < p \leq t/z \\ p \in \mathcal{P}}} \sum_{\substack{m \leq t/p \\ \gcd(m, P(z))=1}} 1 \;+\; \sum_{\substack{t/z < p \leq t \\ p \in \mathcal{P}}} \sum_{\substack{m \leq t/p \\ \gcd(m, P(z))=1}} 1\,.$$

Since $t/p < z$, the above sum is bounded by $E_1 + E_2$ where

$$E_1 = \sum_{\substack{z<p\le t/z\\ p\in\mathcal{P}}} \sum_{\substack{m\le t/p\\ \gcd(m,P(z))=1}} 1 \quad \text{and} \quad E_2 = \sum_{\substack{t/z<p\le t\\ p\in\mathcal{P}}} \sum_{\substack{m\le t/p\\ \gcd(m,P(t/p))=1}} 1\,.$$

Note that in E_1 we have $t/p \ge z$. The inner sum of E_1 is $S(\mathcal{A},\mathcal{P},z)$ where $\mathcal{A} = \{\, n \,|\, n \le t/p \,\}$. Applying Theorem 4.3.1 to the inner sum of E_1 and using a trivial estimate for the inner sum of E_2 , we can rewrite (4.2) as

$$\sum_{\substack{m\le t\\ \gcd(m,P(t))=1}} 1 \;=\; \sum_{\substack{m\le t\\ \gcd(m,P(z))=1}} 1 \;+\; O\left(W(z)\sum_{\substack{z<p\le t/z\\ p\in\mathcal{P}}}\frac{t}{p}\right) \;+\; O\left(\sum_{\substack{t/z<p\le t\\ p\in\mathcal{P}}}\frac{t}{p}\right),$$

where the implied constants are absolute. We now wish to show that the two error terms are equal to the ones appearing in the statement of the lemma. Because $z \ge w_0$, by the conditions of the lemma, and choosing $k \in \mathbb{Z}^+$ such that $e^k z \le t/z < e^{k+1}z$ yields

$$\begin{aligned}
\sum_{\substack{z<p\le t/z\\ p\in\mathcal{P}}}\frac{1}{p} &\le \frac{\varepsilon}{\log z} + \frac{\varepsilon}{\log z+1} + \cdots + \frac{\varepsilon}{\log z + k}\\
&\le \varepsilon\int_{\log z-1}^{\log t/z}\frac{1}{u}du\\
&= \varepsilon\log\left((\log_2^2 t - 1)\left(1-\frac{\log_2^2 t}{\log t}\right)^{-1}\right) \;\ll\; \varepsilon\log_3 t\,.
\end{aligned}$$

The same method applies to the other error term, which yields

$$\sum_{\substack{m\le t\\ \gcd(m,P(t))=1}} 1 \;=\; \sum_{\substack{m\le t\\ \gcd(m,P(z))=1}} 1 \;+\; O\Big(\varepsilon t\log_3(t)W(z)\Big) + O\left(\frac{\varepsilon t}{\log_2^2 t}\right). \qquad (4.3)$$

By Theorem 4.3.1, the main term in (4.3) is equal to

$$tW(z)\Big(1 + O\big(\exp(-\log_2^2 t\log_3 t)\big)\Big),$$

where the implied constant is absolute. Since we assumed that $\varepsilon > \exp(-\log_2^2 t\log_3 t)$, the above error can be taken in the first error in (4.3). This completes the proof of Lemma 4.3.2. □

For our purpose, we are interested in investigating the set $\mathcal{P}$ of primes in an arithmetic progression since this is where our applications of Lemma 4.3.2 lie.

Lemma 4.3.3. *Let $m \geq 2$ be an integer. For all $w \geq m^{12}$, we have*

$$\sum_{\substack{w<p\leq ew \\ p\equiv 1 \,(mod\, m)}} \frac{1}{p} \;\leq\; \frac{6}{\phi(m)\log w}\,.$$

Proof. By the Montgomery-Vaughan version of the Brun-Titchmarsh inequality in [18] and since $w \geq m^{12}$, we have

$$\begin{aligned}
\sum_{\substack{w<p\leq ew \\ p\equiv 1 \,(\mathrm{mod}\, m)}} \frac{1}{p} \;&\leq\; \frac{1}{w} \sum_{\substack{w<p\leq ew \\ p\equiv 1 \,(\mathrm{mod}\, m)}} 1 \\
&\leq\; \frac{1}{w}\frac{2ew}{\phi(m)\log\left(ew/m\right)} \\
&<\; \frac{12}{11}\frac{2e}{\phi(m)\log w} \\
&<\; \frac{6}{\phi(m)\log w}\,.
\end{aligned}$$

□

4.4 Prime Factorization of $\lambda(n)$

Let q be a prime integer and m be a positive integer. Then $v_q(m)$ denotes the exponent of q in the prime factorization of m, that is $q^{v_q(m)} \| m$.

Theorem 4.4.1. *Let ε be a number in the interval $(0,1)$ and let $q \leq \log_2^{\varepsilon/2} x$ be a prime integer. Then for $x \geq 16$, we have*

$$\#\left\{ n \leq x \,:\, \left| v_q(\lambda(n)) - \frac{\log_3 x}{\log q} \right| > \varepsilon \frac{\log_3 x}{\log q} \right\} \;=\; O\left(\frac{x}{\log_2^{\varepsilon} x}\right)\,,$$

where the O-constant depends only on ε.

Proof. We prove the theorem in two parts. Letting $K = \varepsilon \frac{\log_3 x}{\log q}$, we first wish to show that

$$\#\left\{ n \leq x \,:\, v_q(\lambda(n)) < \frac{\log_3 x}{\log q} - K \right\} \;=\; O\Big(x \exp(-\log_2^{\varepsilon/2} x)\Big)\,.$$

Let $K_1 = \left\lceil \frac{\log_3 x}{\log q} - K \right\rceil$. Since $q \le \log_2^{\varepsilon/2} x$, we have that $K \ge 2$ and $K_1 \ge 1$ when x is sufficiently large. Then, since $v_q(\lambda(n)) < K_1$ implies that $q^{K_1} \nmid p-1$ for any $p|n$, we have

$$\begin{aligned}
\# \left\{ n \le x \,:\, v_q(\lambda(n)) < \frac{\log_3 x}{\log q} - K \right\} &\le \# \left\{ n \le x \,:\, v_q(\lambda(n)) < K_1 \right\} \\
&\le \# \left\{ n \le x \,:\, p \not\equiv 1 \,(\mathrm{mod}\, q^{K_1}) \text{ for all } p \,|\, n \right\} \\
&= S(\mathcal{A}, \mathcal{P}_{q^{K_1}}, x)\,,
\end{aligned}$$

where $\mathcal{A} = \{ n \le x \}$ and $\mathcal{P}_{q^{K_1}} = \{ p \,:\, q^{K_1} | p-1 \}$.

By Theorem 4.3.1 and Lemma 4.3.1, we see that

$$\begin{aligned}
S(\mathcal{A}, \mathcal{P}_{q^{K_1}}, x) &\ll x \prod_{\substack{p \le x \\ q^{K_1} | p-1}} \left(1 - \frac{1}{p}\right) \\
&= x \exp\left(- \sum_{\substack{p \le x \\ q^{K_1} | p-1}} \frac{1}{p} - \sum_{\substack{p \le x \\ q^{K_1} | p-1}} \sum_{j=2}^{\infty} \frac{1}{jp^j} \right) \\
&= x \exp\left(- \sum_{\substack{p \le x \\ q^{K_1} | p-1}} \frac{1}{p} + O\left(\frac{1}{q^{2K_1}}\right) \right) \\
&= x \exp\left(-\frac{\log_2 x}{\phi(q^{K_1})} + O\left(\frac{K_1 \log q}{\phi(q^{K_1})}\right) \right) \\
&\ll x \exp\left(-\frac{\log_2 x}{q^{K_1}(1-1/q)} \right),
\end{aligned}$$

where the last inequality follows from the fact that

$$\frac{K_1 \log q}{\phi(q^{K_1})} \ll 1\,.$$

Moreover, since $K_1 - 1 \le \log_3 x / \log q - K$ and $K \ge 2$, we have

$$\begin{aligned}
x \exp\left(-\frac{\log_2 x}{q^{K_1}(1-1/q)} \right) &\le x \exp\left(-\frac{\log_2 x}{q^{K_1}} \right) \\
&\le x \exp\left(-q^{K-1} \right) \\
&\le x \exp\left(-\log_2^{\varepsilon/2} x \right) \\
&= O\left(\frac{x}{\log_2^{\varepsilon} x} \right).
\end{aligned}$$

This concludes the first part of the proof.

We now wish to show that

$$\#\left\{ n \le x \,:\, v_q(\lambda(n)) > \frac{\log_3 x}{\log q} + K \right\} \;=\; O\left(\frac{x}{\log_2^{\varepsilon} x}\right).$$

Let $K_2 = \left\lfloor \frac{\log_3 x}{\log q} + K \right\rfloor$. Since $\nu_q(\lambda(n)) > K_2$ implies that $q^{K_2+2}|n$ or that $q^{K_2+1}|p-1$ for some $p|n$, we have

$$\begin{aligned}
\#\left\{ n \le x \,:\, v_q(\lambda(n)) > \frac{\log_3 x}{\log q} + K \right\} \;&\le\; \#\left\{ n \le x \,:\, v_q(\lambda(n)) > K_2 \right\} \\
&\le\; \#\left\{ n \le x \,:\, q^{K_2+2}|n \right\} \\
&\quad +\; \#\left\{ n \le x \,:\, q^{K_2+1}|p-1 \text{ for some } p|n \right\}.
\end{aligned}$$

From our choice of K_2 the first term above is bounded by

$$\frac{x}{q^{K_2+2}} \;\le\; \frac{x}{q^{K+1}\log_2 x} \;=\; O\left(\frac{x}{\log_2 x}\right).$$

By Lemma 4.3.1,the second term above is bounded by

$$\sum_{\substack{p\le x \\ q^{K_2+1}|p-1}} \frac{x}{p} \;=\; x\left(\frac{\log_2 x}{q^{K_2+1}(1-1/q)} + O\left(\frac{K_2 \log q}{q^{K_2+1}}\right)\right).$$

Moreover, since $K_2 + 1 \ge (1+\varepsilon)\frac{\log_3 x}{\log q}$, we see that

$$\frac{\log_2 x}{q^{K_2+1}(1-1/q)} \;\ll\; \frac{\log_2 x}{q^{\frac{\log_3 x}{\log q}(1+\varepsilon)}} \;=\; \frac{1}{\log_2^{\varepsilon} x}.$$

Finally, provided that x is sufficiently large, our choice of K and K_2 yields

$$\frac{K_2 \log q}{q^{K_2+1}} \;\le\; \frac{(1+\varepsilon)\log_3 x}{\log_2^{1+\varepsilon} x} \;\le\; \frac{1}{\log_2^{\varepsilon} x}.$$

The above two remarks show that

$$\sum_{\substack{p\le x \\ q^{K_2+1}|p-1}} \frac{x}{p} \;\ll\; \frac{x}{\log_2^{\varepsilon} x},$$

which concludes the proof. □

4.5 First Moment of $\tilde{f}$

In this section, we consider the first moment of $\tilde{f}$ where given $n \in \mathbb{Z}^+$, we let

$$\tilde{f}(n) := \sum_{\substack{q \le \log_4 n \\ q \mid \lambda(n) \\ \Delta_q(n)=1}} \frac{1}{q}$$

when $\log_4 n$ is defined and $\tilde{f}(n) := 0$ otherwise. Our goal is to prove the following theorem.

Lemma 4.5.1. *Let $\varepsilon \in (0, 1/4]$. There exists an x_0 such that for $x \ge x_0$,*

$$\sum_{n \le x} \tilde{f}(n) \;=\; x \sum_{q \le \log_4 x} \frac{1}{q} \sum_{\substack{k \ge 1 \\ \left|k - \frac{\log_3 x}{\log q}\right| \le \varepsilon \frac{\log_3 x}{\log q}}} \frac{\log_2 x}{q^k} \exp\left(-\frac{\log_2 x}{\phi(q^k)}\right) + O_\varepsilon\left(\frac{x}{\log_4 x}\right).$$

Proof. It follows from the definition of $\tilde{f}(n)$ that

$$\sum_{n \le x} \tilde{f}(n) \;=\; \sum_{n \le x} \sum_{\substack{q \le \log_4 n \\ \Delta_q(n)=1}} \frac{1}{q} \;=\; \sum_{q \le \log_4 x} \frac{1}{q} \sum_{\substack{n \le x \\ \Delta_q(n)=1}} 1 \;+ O\left(\frac{x}{\log_4 x}\right) \tag{4.4}$$

as the difference between the two sums in the second equality is bounded above by

$$\begin{aligned}
\sum_{n \le x} \sum_{\log_4 n < q \le \log_4 x} \frac{1}{q} \;&=\; \sum_{n \le x^{1/2}} \sum_{\log_4 n < q \le \log_4 x} \frac{1}{q} \;+ \sum_{x^{1/2} < n \le x} \sum_{\log_4 n < q \le \log_4 x} \frac{1}{q} \\
&\ll \sum_{n \le x^{1/2}} \log_6 x \;+ \sum_{x^{1/2} < n \le x} \frac{1}{\log_4 x} \\
&\le\; x^{1/2} \log_6 x + \frac{x}{\log_4 x} \;\ll\; \frac{x}{\log_4 x}.
\end{aligned}$$

Applying Theorem 4.4.1 to the inner sum on the right-hand side of (4.4) yields

$$\sum_{n \le x} \tilde{f}(n) \;=\; \sum_{q \le \log_4 x} \frac{1}{q} \sum_{\substack{n \le x \\ \Delta_q(n)=1 \\ \left|v_q(\lambda(n)) - \frac{\log_3 x}{\log q}\right| \le \varepsilon \frac{\log_3 x}{\log q}}} 1 \;+\; O\left(\frac{x \log_6 x}{\log_2^{\varepsilon} x}\right) + O\left(\frac{x}{\log_4 x}\right), \tag{4.5}$$

where $\varepsilon \in (0, 1)$ will be determined later.

We now wish to show that the following equality holds for the inner sum of (4.5):

$$\sum_{\substack{n\le x\\ \Delta_q(n)=1\\ \left|v_q(\lambda(n))-\frac{\log_3 x}{\log q}\right|\le\varepsilon\frac{\log_3 x}{\log q}}} 1 \;=\; \sum_{\left|k-\frac{\log_3 x}{\log q}\right|\le\varepsilon\frac{\log_3 x}{\log q}} \sum_{\substack{p\le x^{1/4}\\ q^k\,\|\,p-1}} \sum_{\substack{m\le x/p\\ \gcd\left(m,P_{q^k}(x/p)\right)=1}} 1 \;+\; O\left(\frac{x\log_3 x}{\log_2^{1-\varepsilon} x}\right),$$

where

$$P_{q^k}(x/p) \;:=\; \prod_{\substack{\tilde p\le x/p\\ q^k\,|\,\tilde p-1}} \tilde p\,.$$

To accomplish this, we begin by rearranging the left-hand side. Observe that the condition $q^k\|\lambda(n)$ implies that $q^k\|p-1$ for some prime $p|n$ or $q^{k+1}|n$ and these two conditions are exclusive since $\Delta_q(n)=1$. This implies that

$$\begin{aligned}\sum_{\substack{n\le x\\ \Delta_q(n)=1\\ \left|v_q(\lambda(n))-\frac{\log_3 x}{\log q}\right|\le\varepsilon\frac{\log_3 x}{\log q}}} 1 \;&=\; \sum_{\left|k-\frac{\log_3 x}{\log q}\right|\le\varepsilon\frac{\log_3 x}{\log q}} \sum_{\substack{n\le x\\ q^k\,\|\,\lambda(n)\\ \Delta_q(n)=1}} 1\\ &=\; \sum_{\left|k-\frac{\log_3 x}{\log q}\right|\le\varepsilon\frac{\log_3 x}{\log q}} \sum_{\substack{n\le x\\ q^k\,\|\,p-1\text{ for some }p|n\\ \Delta_q(n)=1,\,q^k\,\|\,\lambda(n)}} 1 \;+\; \sum_{\left|k-\frac{\log_3 x}{\log q}\right|\le\varepsilon\frac{\log_3 x}{\log q}} \sum_{\substack{n\le x\\ q^{k+1}\,|\,n\\ \Delta_q(n)=1,\,q^k\,\|\,\lambda(n)}} 1\,. \qquad (4.6)\end{aligned}$$

Let us show that the second sum in (4.6) falls into the error term:

$$\begin{aligned}\sum_{\left|k-\frac{\log_3 x}{\log q}\right|\le\varepsilon\frac{\log_3 x}{\log q}} \sum_{\substack{n\le x\\ q^{k+1}\,|\,n\\ \Delta_q(n)=1,\,q^k\,\|\,\lambda(n)}} 1 \;&\le\; \sum_{\left|k-\frac{\log_3 x}{\log q}\right|\le\varepsilon\frac{\log_3 x}{\log q}} \frac{x}{q^{k+1}}\\ &\le\; xq^{(\varepsilon-1)\frac{\log_3 x}{\log q}} \sum_{\left|k-\frac{\log_3 x}{\log q}\right|\le\varepsilon\frac{\log_3 x}{\log q}} 1\\ &=\; \frac{x}{\log_2^{1-\varepsilon} x} \sum_{\left|k-\frac{\log_3 x}{\log q}\right|\le\varepsilon\frac{\log_3 x}{\log q}} 1\\ &\le\; \frac{x}{\log_2^{1-\varepsilon} x}\left(2\varepsilon\frac{\log_3 x}{\log q}\right)\\ &\ll\; \frac{x\log_3 x}{\log_2^{1-\varepsilon} x}\,.\end{aligned}$$

We now rewrite the first sum in (4.6) in a more appropriate form. First, instead of summing over $n\le x$, if we sum over primes $p\le x$ satisfying the required

conditions, we obtain

$$\sum_{\substack{n\le x\\ q^k\,\|\,p-1 \text{ for some } p\mid n\\ \Delta_q(n)=1,\, q^k\,\|\,\lambda(n)}} 1 \;=\; \sum_{\substack{p\le x\\ q^k\,\|\,p-1}}\sum_{r\ge 1}\sum_{\substack{m\le x/p^r\\ \gcd\left(m,P_{q^k}(x/p^r)\right)=1}} 1 \;+\; O\left(\frac{x}{q^{k+1}}\right),$$

where the error term on the right-hand side appears because the triple sum counts the number of $n \le x$ taking into account every required condition, except that it allows for the possibility that $q^{k+1}|n$, which violates the condition $\Delta_q(n) = 1$. Since there are at most x/q^{k+1} such n's, the error term is justified. This implies

$$\begin{aligned}
&\sum_{\left|k-\frac{\log_3 x}{\log q}\right|\le \varepsilon\frac{\log_3 x}{\log q}}\ \sum_{\substack{n\le x\\ q^k\,\|\,p-1 \text{ for some } p\mid n\\ \Delta_q(n)=1,\, q^k\,\|\,\lambda(n)}} 1\\
&= \sum_{\left|k-\frac{\log_3 x}{\log q}\right|\le \varepsilon\frac{\log_3 x}{\log q}}\ \sum_{\substack{p\le x\\ q^k\,\|\,p-1}}\sum_{r\ge 1}\sum_{\substack{m\le x/p^r\\ \gcd\left(m,P_{q^k}(x/p^r)\right)=1}} 1 \;+\; O\left(\frac{x\log_3 x}{\log_2^{1-\varepsilon} x}\right)\\
&= S_1+S_2+S_3 \;+\; O\left(\frac{x\log_3 x}{\log_2^{1-\varepsilon} x}\right),
\end{aligned}$$

where S_1, S_2 and S_3 represent the contributions of the quadruple sum corresponding to the conditions $p \le x^{1/4}$ and $r = 1$, $p > x^{1/4}$ and $r = 1$, and $r \ge 2$ respectively.

Note that

$$\begin{aligned}
S_3 &= \sum_{\left|k-\frac{\log_3 x}{\log q}\right|\le \varepsilon\frac{\log_3 x}{\log q}}\ \sum_{\substack{p\le x\\ q^k\,\|\,p-1}}\sum_{r\ge 2}\sum_{\substack{m\le x/p^r\\ \gcd\left(m,P_{q^k}(x/p^r)\right)=1}} 1\\
&\le \sum_{\left|k-\frac{\log_3 x}{\log q}\right|\le \varepsilon\frac{\log_3 x}{\log q}}\ \sum_{\substack{p\le x\\ q^k\,\|\,p-1}}\sum_{r\ge 2}\frac{x}{p^r}\\
&\le 2x\sum_{\left|k-\frac{\log_3 x}{\log q}\right|\le \varepsilon\frac{\log_3 x}{\log q}}\ \sum_{\substack{p\le x\\ q^k\,\|\,p-1}}\frac{1}{p^2}\\
&\ll x\sum_{\left|k-\frac{\log_3 x}{\log q}\right|\le \varepsilon\frac{\log_3 x}{\log q}}\frac{1}{q^{2k}}\\
&\ll x q^{2(\varepsilon-1)\frac{\log_3 x}{\log q}}\left(2\varepsilon\frac{\log_3 x}{\log q}\right)\\
&\le \frac{x\log_3 x}{\log_2^{2(1-\varepsilon)} x} \;\ll\; \frac{x\log_3 x}{\log_2^{1-\varepsilon} x},
\end{aligned}$$

which again falls into the error term.

From Lemma 4.3.1,

$$\begin{aligned}
S_2 &= \sum_{\left|k-\frac{\log_3 x}{\log q}\right|\le\varepsilon\frac{\log_3 x}{\log q}} \sum_{\substack{x^{1/4}<p\le x\\ q^k \,\|\, p-1}} \sum_{\substack{m\le x/p\\ \gcd\left(m,P_{q^k}(x/p)\right)=1}} 1\\
&\le \sum_{\left|k-\frac{\log_3 x}{\log q}\right|\le\varepsilon\frac{\log_3 x}{\log q}} \sum_{\substack{x^{1/4}<p\le x\\ q^k \,\|\, p-1}} \frac{x}{p}\\
&= x \sum_{\left|k-\frac{\log_3 x}{\log q}\right|\le\varepsilon\frac{\log_3 x}{\log q}} \left(\frac{\log_2 x-\log_2 x^{1/4}}{\phi(q^k)}+O\left(\frac{\log q^k}{q^k}\right)\right)\\
&\ll x \sum_{\left|k-\frac{\log_3 x}{\log q}\right|\le\varepsilon\frac{\log_3 x}{\log q}} \left(\frac{1}{q^k}+\frac{\log q^k}{q^k}\right)\\
&\ll \frac{x\log_3 x}{\log_2^{1-\varepsilon} x} + x \sum_{\left|k-\frac{\log_3 x}{\log q}\right|\le\varepsilon\frac{\log_3 x}{\log q}} \frac{\log q^k}{q^k}\,.
\end{aligned}$$

Moreover, we have

$$\begin{aligned}
x \sum_{\left|k-\frac{\log_3 x}{\log q}\right|\le\varepsilon\frac{\log_3 x}{\log q}} \frac{\log q^k}{q^k} &= x\log q \sum_{\left|k-\frac{\log_3 x}{\log q}\right|\le\varepsilon\frac{\log_3 x}{\log q}} k\left(\frac{1}{q}\right)^k\\
&\le x\log q(1+\varepsilon)\frac{\log_3 x}{\log q} \sum_{\left|k-\frac{\log_3 x}{\log q}\right|\le\varepsilon\frac{\log_3 x}{\log q}} \left(\frac{1}{q}\right)^k\\
&\ll x\log_3 x\left(q^{(\varepsilon-1)\frac{\log_3 x}{\log q}}\right)\\
&= \frac{x\log_3 x}{\log_2^{1-\varepsilon} x}\,.
\end{aligned}$$

Therefore we conclude that

$$\sum_{\substack{n\le x\\ \Delta_q(n)=1\\ \left|v_q(\lambda(n))-\frac{\log_3 x}{\log q}\right|\le\varepsilon\frac{\log_3 x}{\log q}}} 1 = \sum_{\left|k-\frac{\log_3 x}{\log q}\right|\le\varepsilon\frac{\log_3 x}{\log q}} \sum_{\substack{p\le x^{1/4}\\ q^k \,\|\, p-1}} \sum_{\substack{m\le x/p\\ \gcd\left(m,P_{q^k}(x/p)\right)=1}} 1 + O\left(\frac{x\log_3 x}{\log_2^{1-\varepsilon} x}\right).$$

Substituting this into (4.5) yields

$$\sum_{n\le x}\tilde{f}(n) = \sum_{q\le\log_4 x}\frac{1}{q} \sum_{\left|k-\frac{\log_3 x}{\log q}\right|\le\varepsilon\frac{\log_3 x}{\log q}} \sum_{\substack{p\le x^{1/4}\\ q^k \,\|\, p-1}} \sum_{\substack{m\le x/p\\ \gcd\left(m,P_{q^k}(x/p)\right)=1}} 1 + E(x)\,, \qquad (4.7)$$

where

$$\begin{aligned} E(x) &= O\left(\frac{x\log_6 x}{\log_2^{\varepsilon} x}\right) + O\left(\frac{x\log_3 x\log_6 x}{\log_2^{1-\varepsilon} x}\right) + O\left(\frac{x}{\log_4 x}\right) \\ &\ll \frac{x}{\log_4 x} \end{aligned}$$

when $x \geq x_\varepsilon$ for some x_ε.

Let $t = x/p$ with $p \leq x^{1/4}$. Then $x^{3/4} \leq t \leq x$. If we let $\varepsilon \leq 1/2$, we can see that q and k satisfy $\log_2^{1/2} x \leq q^k \leq \log_2^{3/2} x$. Thus we have $\log_2^{1/2} t \leq q^k \leq \log_2^2 t$ when x is sufficiently large. If we let $\mathcal{P} = \{\, p : p \equiv 1 \,(\mathrm{mod}\, q^k)\,\}$ and ask that $\varepsilon \geq 6/\phi(q^k)$, we see that with the above conditions and Lemma 4.3.3, all the conditions of Lemma 4.3.2 are satisfied, which enables us to conclude that

$$\sum_{\substack{m\leq t\\ \gcd(m,P(t))=1}} 1 = t\left(W(z) + O\Big(\frac{\log_3 t}{q^k}W(z) + \frac{1}{q^k\log_2^2 t}\Big)\right), \tag{4.8}$$

where $z = \exp(\log t/\log_2^2 t)$. Using Lemma 4.3.1 and the fact that $\log_2^{1/2} x \leq q^k \leq \log_2^{3/2} x$ and $\log_2 t = \log_2 x + O(1)$, we may write $W(z)$ in a more effective way:

$$\begin{aligned} \log W(z) &= \sum_{\substack{p<z\\ p\in\mathcal{P}}} \log\left(1-\frac{1}{p}\right) = \sum_{\substack{p<z\\ p\equiv 1\,(\mathrm{mod}\, q^k)}} \log\left(1-\frac{1}{p}\right) \\ &= -\sum_{\substack{p<z\\ p\equiv 1\,(\mathrm{mod}\, q^k)}} \frac{1}{p} - \sum_{\substack{p<z\\ p\equiv 1\,(\mathrm{mod}\, q^k)}} \sum_{j=2}^{\infty} \frac{1}{jp^j} \\ &= -\frac{\log_2 z}{\phi(q^k)} + O\left(\frac{\log q^k}{q^k}\right) - \sum_{\substack{p<z\\ p\equiv 1\,(\mathrm{mod}\, q^k)}} \sum_{j=2}^{\infty} \frac{1}{jp^j} \end{aligned}$$

and

$$\sum_{\substack{p<z\\ p\equiv 1\,(\mathrm{mod}\, q^k)}} \sum_{j=2}^{\infty} \frac{1}{jp^j} = O\left(\frac{1}{q^{2k}}\right) = O\left(\frac{\log q^k}{q^k}\right).$$

Thus, since $z = \exp(\log t/\log_2^2 t)$ and $\log_2 t = \log_2 x + O(1)$,

$$\begin{aligned} W(z) &= \exp\left(-\frac{\log_2 z}{\phi(q^k)} + O\left(\frac{\log q^k}{q^k}\right)\right) \\ &= \exp\left(-\frac{\log_2 x}{\phi(q^k)} + O\left(\frac{\log_3 x}{q^k}\right) + O\left(\frac{\log q^k}{q^k}\right)\right). \end{aligned}$$

Finally, since $\log_3 x \ll \log q^k$ and using the series expansion of the exponential function, we have

$$\begin{aligned} W(z) &= \exp\left(-\frac{\log_2 x}{\phi(q^k)} + O\left(\frac{\log q^k}{q^k}\right)\right) \\ &= \left(1 + O\left(\frac{\log q^k}{q^k}\right)\right) \exp\left(-\frac{\log_2 x}{\phi(q^k)}\right). \end{aligned}$$

The first application of this formula is to simplify the expression within the big-O term in (4.8) as follows:

Since $\log_2 t = \log_2 x + O(1)$ and

$$\exp\left(-\frac{\log_2 x}{\phi(q^k)}\right) \ll \frac{q^{2k}}{\log_2^2 x}, \tag{4.9}$$

then

$$\begin{aligned} \frac{\log_3 t}{q^k} W(z) + \frac{1}{q^k \log_2^2 t} &\ll \frac{\log_3 x}{q^k} W(z) + \frac{1}{q^k \log_2^2 x} \\ &\ll \frac{\log_3 x}{q^k} \exp\left(-\frac{\log_2 x}{\phi(q^k)}\right) + \frac{1}{q^k \log_2^2 x} \\ &\ll \frac{\log_3 x}{q^k} \frac{q^{2k}}{\log_2^2 x} + \frac{1}{q^k \log_2^2 x} \\ &\ll \frac{q^k \log_3 x}{\log_2^2 x}. \end{aligned}$$

Secondly, from (4.9) and since $\log q^k \ll \log_3 x$, we can write $W(z)$ as

$$W(z) = \exp\left(-\frac{\log_2 x}{\phi(q^k)}\right) + O\left(\frac{q^k \log_3 x}{\log_2^2 x}\right).$$

Thus, by (4.8), we have

$$\sum_{\substack{m \le x/p \\ \gcd\left(m, P_{q^k}(x/p)\right)=1}} 1 = \frac{x}{p}\left(\exp\left(-\frac{\log_2 x}{\phi(q^k)}\right) + O\left(\frac{q^k \log_3 x}{\log_2^2 x}\right)\right).$$

When we put this formula in (4.7), by Lemma 4.3.1 and since $\log q^k \ll \log_3 x$, the

error generated is bounded by

$$\begin{aligned}
& \sum_{q\le \log_4 x} \frac{1}{q} \sum_{\left|k-\frac{\log_3 x}{\log q}\right|\le \varepsilon \frac{\log_3 x}{\log q}} \sum_{\substack{p\le x^{1/4}\\ q^k \,\|\, p-1}} \frac{x}{p} \frac{q^k \log_3 x}{\log_2^2 x} \\
\ll\ & x\left(\frac{\log_3 x}{\log_2^2 x}\right) \sum_{q\le \log_4 x} \frac{1}{q} \sum_{\left|k-\frac{\log_3 x}{\log q}\right|\le \varepsilon \frac{\log_3 x}{\log q}} q^k \sum_{\substack{p\le x^{1/4}\\ q^k \,\|\, p-1}} \frac{1}{p} \\
\ll\ & x\left(\frac{\log_3 x}{\log_2^2 x}\right) \sum_{q\le \log_4 x} \frac{1}{q} \sum_{\left|k-\frac{\log_3 x}{\log q}\right|\le \varepsilon \frac{\log_3 x}{\log q}} q^k \left(\frac{\log_2 x}{q^k}\right) \\
\ll\ & x\left(\frac{\log_3 x}{\log_2 x}\right) \sum_{q\le \log_4 x} \frac{1}{q} \sum_{\left|k-\frac{\log_3 x}{\log q}\right|\le \varepsilon \frac{\log_3 x}{\log q}} 1 \\
\ll\ & x\left(\frac{\log_3^2 x}{\log_2 x}\right) \sum_{q\le \log_4 x} \frac{1}{q\log q}\,.
\end{aligned}$$

However, since

$$\sum_{q\le \log_4 x} \frac{1}{q\log q} \ll 1\,,$$

the error term is bounded by

$$\frac{x\log_3^2 x}{\log_2 x}\,.$$

Therefore we can rewrite (4.7) as

$$\sum_{n\le x} \tilde{f}(n) \;=\; x \sum_{q\le \log_4 x} \frac{1}{q} \sum_{\left|k-\frac{\log_3 x}{\log q}\right|\le \varepsilon \frac{\log_3 x}{\log q}} \sum_{\substack{p\le x^{1/4}\\ q^k \,\|\, p-1}} \frac{1}{p} \exp\left(-\frac{\log_2 x}{\phi(q^k)}\right) \;+\; O\left(\frac{x}{\log_4 x}\right).$$

Applying Lemma 4.3.1 to the sum

$$\sum_{\substack{p\le x^{1/4}\\ q^k \,\|\, p-1}} \frac{1}{p}\,,$$

we obtain the main term exactly as in Lemma 4.5.1 while we can put the generated error term in $O(x/\log_4 x)$. This completes the proof of Lemma 4.5.1.

□

4.6 Second Moment of $\tilde{f}$

In this section, we consider the second moment of $\tilde{f}$ and we will prove the following theorem.

Lemma 4.6.1. *Let $\varepsilon \in (0, 1/5]$. There exists an x_0 such that for $x \geq x_0$,*

$$\sum_{n\leq x} \tilde{f}^2(n) \;=\; x \sum_{q_1,q_2\leq \log_4 x} {\sum_{k_1,k_2\geq 1}}' \frac{\log_2^2 x}{q_1^{k_1+1}q_2^{k_2+1}} \exp\left(-\sum_{j=1,2}\frac{\log_2 x}{\phi(q_j^{k_j})}\right) + O(x)\,,$$

where $\sum'$ means that the sum is taken over k_1 and k_2 with $|k_i - \log_3 x/\log q_i| \leq \varepsilon \log_3 x/\log q_i$ for $i = 1, 2$.

Proof. By definition of $\tilde{f}$ and the fact that $\Delta_q(n) = 1$ implies $q \,|\, \lambda(n)$, we have

$$\begin{aligned}
\sum_{n\leq x} \tilde{f}^2(n) &= \sum_{n\leq x} \sum_{\substack{q_i\leq \log_4 n\\ \Delta_{q_i}(n)=1}} \frac{1}{q_1q_2} \\
&= \sum_{q_i\leq \log_4 x} \frac{1}{q_1q_2} \sum_{\substack{n\leq x\\ \Delta_{q_i}(n)=1}} 1 + O\left(\sum_{n\leq x} \sum_{\substack{q_1\leq \log_4 x\\ \log_4 n<q_2\leq \log_4 x}} \frac{1}{q_1q_2}\right) \\
&= \sum_{q_i\leq \log_4 x} \frac{1}{q_1q_2} \sum_{\substack{n\leq x\\ \Delta_{q_i}(n)=1}} 1 + O\left(\frac{x\log_6 x}{\log_4 x}\right).
\end{aligned}$$

Observe that the last equality follows by partial summation in the following way:

$$\begin{aligned}
&\sum_{n\leq x} \sum_{\substack{q_1\leq \log_4 x\\ \log_4 n<q_2\leq \log_4 x}} \frac{1}{q_1q_2} = \sum_{n\leq x} \sum_{q_1\leq \log_4 x} \frac{1}{q_1}\left(\sum_{\log_4 n<q_2\leq \log_4 x} \frac{1}{q_2}\right) \\
\ll\;& \sum_{n\leq x} \log_6 x\,(\log_6 x - \log_6 n) \\
=\;& (\log_6 x)^2 \sum_{n\leq x} 1 \;-\; \log_6 x \sum_{\alpha\leq n\leq x} \log_6 n \\
\ll\;& (\log_6 x)^2 \sum_{n\leq x} 1 \;- \log_6 x \left(\left(\sum_{n\leq x} 1\right)\log_6 x - \int_\alpha^x \frac{\lfloor t\rfloor}{\log_5 t \log_4 t\cdots\log t}\frac{dt}{t}\right) \\
\ll\;& \log_6 x \int_\alpha^x \frac{dt}{\log_5 t\log_4 t\cdots\log t} \\
\ll\;& \log_6 x \int_\alpha^x \frac{d}{dt}\left(\frac{t}{\log_4 t}\right)dt \;\ll\; \frac{x\log_6 x}{\log_4 x}\,,
\end{aligned}$$

where $\alpha > 0$ is some sufficiently large positive constant so that the above integrals are well-defined.

When $q_1 = q_2$, we have

$$\sum_{q \le \log_4 x} \frac{1}{q^2} \sum_{n \le x} 1 \;=\; O(x)\,,$$

hence

$$\sum_{n \le x} \tilde{f}^2(n) \;=\; \sum_{\substack{q_i \le \log_4 x \\ q_1 \ne q_2}} \frac{1}{q_1 q_2} \sum_{\substack{n \le x \\ \Delta_{q_i}(n)=1}} 1 \;+\; O(x)\,. \tag{4.10}$$

In the remaining part of this section, our goal will be to show that for $x \ge x_0 > 0$ the inner sum of (4.10) satisfies the equality

$$\sum_{\substack{n \le x \\ \Delta_{q_i}(n)=1}} 1 \;=\; x \sum_{k_1,k_2}{}' \frac{\log_2^2 x}{q_1^{k_1} q_2^{k_2}} \exp\left(- \sum_{i=1,2} \frac{\log_2 x}{\phi(q_i^{k_i})} \right) \;+\; O\left(\frac{x}{\log_2^{\varepsilon} x} \right), \tag{4.11}$$

where $\varepsilon \in (0, 1/4]$ is fixed and $\sum_{k_1,k_2}{}'$ indicates that k_1 and k_2 are subject to the condition

$$\left| k_i - \frac{\log_3 x}{\log q_i} \right| \le \varepsilon \frac{\log_3 x}{\log q_i}\,, \text{ for } i = 1, 2.$$

Lemma 4.6.1 will then follow immediately from (4.10) and (4.11).

Before we proceed to prove (4.11) in a series of steps, let us define

$$P_i(t) := \prod_{\substack{p \le t \\ p \equiv 1 \,(\mathrm{mod}\, q_i^{k_i})}} p\,.$$

Step 1. An explicit formula for $\sum_{\substack{n \le x \\ \Delta_{q_i}(n)=1}} 1\,.$

Let us denote by $H(q_1, q_2)$ the following condition:

$$\left| \nu_{q_i}(\lambda(n)) - \frac{\log_3 x}{\log q_i} \right| \le \varepsilon \frac{\log_3 x}{\log q_i}\,, \text{ for } i = 1, 2.$$

By Theorem 4.4.1, since $\log_4 x < \log_2^{\varepsilon/2} x$ for x sufficiently large and $q_i \le \log_4 x$,

$$\begin{aligned}\sum_{\substack{n\le x\\ \Delta_{q_i}(n)=1}} 1 \;&=\; \sum_{\substack{n\le x\\ \Delta_{q_i}(n)=1\\ H(q_1,q_2)}} 1 \;+\; \sum_{\substack{n\le x\\ \Delta_{q_i}(n)=1\\ \neg H(q_1,q_2)}} 1 \\ &=\; \#\,\{\,n\le x\,:\,\Delta_{q_i}(n)=1 \;\text{ and }\; H(q_1,q_2)\,\} + O\left(\frac{x}{\log_2^{\varepsilon} x}\right), \qquad (4.12)\end{aligned}$$

where the number $\varepsilon \in (0, 1/2)$ will be chosen later. By analyzing the main term in (4.12), we will prove how

$$\begin{aligned}\sum_{\substack{n\le x\\ \Delta_{q_i}(n)=1}} 1 \;&=\; \sum_{k_1,k_2}{}' \sum_{\substack{p_1,p_2\le x\\ q_i^{k_i}\,\|\,p_i-1}} \sum_{r_1,r_2\ge 1} \sum_{\substack{m\le x/p_1^{r_1}p_2^{r_2}\\ gcd\left(m,P_1(x/p_1^{r_1}p_2^{r_2})P_2(x/p_1^{r_1}p_2^{r_2})\right)=1}} 1 \\ &\quad + \sum_{k_1,k_2}{}' \sum_{\substack{p\le x\\ q_i^{k_i}\,\|\,p-1}} \sum_{\substack{m\le x/p\\ gcd(m,P_1(x/p)P_2(x/p))=1}} 1 \;+\; O\left(\frac{x}{\log_2^{\varepsilon} x}\right). \qquad (4.13)\end{aligned}$$

The second triple sum on the right of (4.13) counts the numbers of $n \le x$, subject to $H(q_1, q_2)$, such that $p_1 = p_2 = p$ where $p \,\|\, n$. Let q be a prime and let $k = v_q(\lambda(n))$. By definition of $\Delta_q(n)$, if $\Delta_q(n) = 1$, then either $q^{k+1} \,|\, n$ or n contains only one prime factor p such that $q^k \,|\, p - 1$. Conversely, if n contains only one prime factor p such that $q^k \,|\, p-1$, then either $\Delta_q(n) = 1$ or $\Delta_q(n) \ge 2$, in which case $q^{k+1} \,|\, n$.

For q_1 fixed, since we assumed that $\varepsilon \in (0, 1/2)$, then

$$\begin{aligned}&\#\{n \le x \,:\, q_1^{k_1+1} | n \;\text{ and }\; |k_1 - \log_3 x/\log q_1| \le \varepsilon \log_3 x/\log q_1\} \\ =\;& O(x/(q_1 \log_2^{1-\varepsilon} x)) \;=\; O(x/\log_2^{\varepsilon} x)\,.\end{aligned}$$

The same bound holds if we reverse the role of q_1 and q_2. The number of $n \le x$ such that n has only one prime p_1 with $v_{q_1}(p_1 - 1) = v_{q_1}(\lambda(n))$ and only one prime p_2 with $v_{q_2}(p_2 - 1) = v_{q_2}(\lambda(n))$, where both $v_{q_i}(\lambda(n))$ are subject to the condition $|v_{q_i}(\lambda(n)) - \log_3 x/\log q_i| \le \varepsilon \log_3 x/\log q_i$, is counted by the four-fold sum and the three-fold sum in (4.13). The numbers n counted by the above two cases exhaust all the numbers n counted by the main term of (4.12). This shows that equation (4.13) is indeed valid.

We are going to simplify the four-fold sum and the three-fold sum in a few steps while always keeping track of the error terms arising in the process. Recall that the

integers k_i fall in the range $[(1-\varepsilon)\log_3 x/\log q_i, (1+\varepsilon)\log_3 x/\log q_i]$.

Step 2. The three-fold sum in (4.13) is equal to $O\left(\dfrac{x}{\log_2^{1-2\varepsilon} x}\right)$.

By Lemma 4.3.1, the three-fold sum is

$$\begin{aligned}
\sum_{k_1,k_2}{}' \sum_{\substack{p\le x \\ q_i^{k_i}\,\|\,p-1}} \sum_{\substack{m\le x/p \\ gcd(m,P_1(x/p)P_2(x/p))=1}} 1 &\le \sum_{k_1,k_2}{}' \sum_{\substack{p\le x \\ q_i^{k_i}\,\|\,p-1}} \frac{x}{p} \\
&= x\sum_{k_1,k_2}{}' \frac{\log_2 x + O(k_1\log q_1 + k_2 \log q_2)}{\phi(q_1^{k_1}q_2^{k_2})} \\
&\ll x\log_2 x \sum_{k_1,k_2}{}' \frac{1}{q_1^{k_1}q_2^{k_2}} \\
&\ll x\log_2 x \left(\sum_{k_1\ge(1-\varepsilon)\frac{\log_3 x}{\log q_1}} \frac{1}{q_1^{k_1}}\right)\left(\sum_{k_2\ge(1-\varepsilon)\frac{\log_3 x}{\log q_2}} \frac{1}{q_2^{k_2}}\right) \\
&\ll x\log_2 x \left(q_1^{(\varepsilon-1)\frac{\log_3 x}{\log q_1}}\right)\left(q_2^{(\varepsilon-1)\frac{\log_3 x}{\log q_2}}\right) \\
&\ll \frac{x}{\log_2^{1-2\varepsilon} x}\,.
\end{aligned}$$

Step 3. The contribution from the terms with $r_1 > 1$ or $r_2 > 1$ in (4.13) is $O\left(\dfrac{x}{\log_2^{2-3\varepsilon} x}\right)$.

To simplify what follows, let us use $\sum_m''$ to denote the sum over m subject to the condition $gcd\,(m, P_1(x/p_1^{r_1}p_2^{r_2})P_2(x/p_1^{r_1}p_2^{r_2})) = 1$. Write

$$\sum_{r_1,r_2\ge 1} \sum_{m\le x/p_1^{r_1}p_2^{r_2}}{}'' 1 = \sum_{m\le x/p_1p_2}{}'' 1 + \sum_{r_1+r_2\ge 3} \sum_{m\le x/p_1^{r_1}p_2^{r_2}}{}'' 1\,.$$

If we let E denote the second sum on the right-hand side, then

$$E \le \sum_{r_1+r_2\ge 3} \frac{x}{p_1^{r_1}p_2^{r_2}} \ll \frac{x}{p_1p_2^2} + \frac{x}{p_1^2p_2}\,.$$

Applying Lemma 4.3.1 yields

$$\begin{aligned}\sideset{}{'}\sum_{k_1,k_2}\sum_{\substack{p_1,p_2\le x\\ q_i^{k_i}\parallel p_i-1}}\frac{x}{p_1p_2^2} &\le x\sideset{}{'}\sum_{k_1,k_2}\sum_{\substack{p_1\le x\\ p_1\equiv 1\,(\mathrm{mod}\,q_1^{k_1})}}\frac{1}{p_1}\sum_{\substack{p_2\le x\\ p_2\equiv 1\,(\mathrm{mod}\,q_2^{k_2})}}\frac{1}{p_2^2}\\ &\ll x\sideset{}{'}\sum_{k_1,k_2}\left(\frac{\log_2 x}{q_1^{k_1}}\right)\left(\frac{1}{q_2^{2k_2}}\right) = x\sideset{}{'}\sum_{k_1,k_2}\frac{\log_2 x}{q_1^{k_1}q_2^{2k_2}}\\ &\ll x\log_2 x\left(\frac{1}{\log_2^{1-\varepsilon}x}\right)\left(\frac{1}{\log_2^{2-2\varepsilon}x}\right) = \frac{x}{\log_2^{2-3\varepsilon}x}\,.\end{aligned}$$

This shows that

$$\sideset{}{'}\sum_{k_1,k_2}\sum_{\substack{p_1,p_2\le x\\ q_i^{k_i}\parallel p_i-1}}\sum_{r_1+r_2\ge 3}\sideset{}{''}\sum_{m\le x/p_1^{r_1}p_2^{r_2}}1 = O\left(\frac{x}{\log_2^{2-3\varepsilon}x}\right).$$

By *Steps* 2 and 3 and choosing $\varepsilon < 1/3$, (4.13) can be written as

$$\sum_{\substack{n\le x\\ \Delta_{q_i}(n)=1}}1 = \sideset{}{'}\sum_{k_i}\sum_{\substack{p_i\le x\\ q_i^{k_i}\parallel p_i-1}}\sideset{}{''}\sum_{m\le x/p_1p_2}1 + O\left(\frac{x}{\log_2^{\varepsilon}x}\right). \tag{4.14}$$

Step 4. The contribution from the terms in (4.14) with $p_1 > x^{1/4}$ or $p_2 > x^{1/4}$ is bounded above by $O\left(\dfrac{x}{\log_2^{1-2\varepsilon}x}\right)$.

By Lemma 4.3.1,

$$\begin{aligned}\sideset{}{'}\sum_{k_i}\sum_{\substack{x^{1/4}<p_1\le x\\ p_2\le x\\ q_i^{k_i}\parallel p_i-1}}\sideset{}{''}\sum_{m\le x/p_1p_2}1 &\le \sideset{}{'}\sum_{k_i}\sum_{\substack{x^{1/4}<p_1\le x\\ p_2\le x\\ q_i^{k_i}\parallel p_i-1}}\frac{x}{p_1p_2}\\ &\le x\sideset{}{'}\sum_{k_i}\left(\sum_{\substack{x^{1/4}<p_1\le x\\ p_1\equiv 1\,(\mathrm{mod}\,q_1^{k_1})}}\frac{1}{p_1}\right)\left(\sum_{\substack{p_2\le x\\ p_2\equiv 1\,(\mathrm{mod}\,q_2^{k_2})}}\frac{1}{p_2}\right)\\ &\ll x\sideset{}{'}\sum_{k_i}\left(\sum_{\substack{x^{1/4}<p_1\le x\\ p_1\equiv 1\,(\mathrm{mod}\,q_1^{k_1})}}\frac{1}{p_1}\right)\left(\frac{\log_2 x}{q_2^{k_2}}\right).\end{aligned}$$

Applying Lemma 4.3.3 to the other sum and choosing $\alpha > 0$ satisfying $e^{\alpha}x^{1/4} < x \le e^{\alpha+1}x^{1/4}$, we obtain

$$\begin{aligned}
\sum_{\substack{x^{1/4}<p_1\le x\\ p_1\equiv 1 \,(\mathrm{mod}\, q_1^{k_1})}} \frac{1}{p_1} &\le \sum_{\substack{x^{1/4}<p_1\le ex^{1/4}\\ p_1\equiv 1 \,(\mathrm{mod}\, q_1^{k_1})}} \frac{1}{p_1} + \sum_{\substack{ex^{1/4}<p_1\le e^2x^{1/4}\\ p_1\equiv 1 \,(\mathrm{mod}\, q_1^{k_1})}} \frac{1}{p_1} + \cdots + \sum_{\substack{e^{\alpha}x^{1/4}<p_1\le e^{\alpha+1}x^{1/4}\\ p_1\equiv 1 \,(\mathrm{mod}\, q_1^{k_1})}} \frac{1}{p_1} \\
&\le \frac{6}{\phi(q_1^{k_1})} \sum_{j=0}^{\alpha} \frac{1}{\frac{1}{4}\log x + j} \\
&\le \frac{24}{\phi(q_1^{k_1})} \left(\frac{1}{\log x}\right) \sum_{j=0}^{\alpha} 1 \\
&\ll \frac{1}{q_1^{k_1}}\,.
\end{aligned}$$

Thus, it follows that

$$\begin{aligned}
\sideset{}{'}\sum_{k_i} \sum_{\substack{x^{1/4}<p_1\le x\\ p_2\le x\\ q_i^{k_i}\,\|\, p_i-1}} \sideset{}{''}\sum_{m\le x/p_1p_2} 1 &\ll (x\log_2 x) \sideset{}{'}\sum_{k_i} \frac{1}{q_1^{k_1}q_2^{k_2}} \\
&\ll (x\log_2 x)\,\frac{1}{\log_2^{2-2\varepsilon} x} \;=\; \frac{x}{\log_2^{1-2\varepsilon} x}\,.
\end{aligned}$$

Observe that the same bound holds if we reverse the role of p_1 and p_2. Substituting the estimate in (4.14) yields

$$\sum_{\substack{n\le x\\ \Delta_{q_i}(n)=1}} 1 \;=\; \sideset{}{'}\sum_{k_i} \sum_{\substack{p_i\le x^{1/4}\\ q_i^{k_i}\,\|\, p_i-1}} \sideset{}{''}\sum_{m\le x/p_1p_2} 1 \;+\; O\left(\frac{x}{\log_2^{\varepsilon} x}\right). \tag{4.15}$$

Step 5. Simplification of the main term in (4.15).

Let $t = x/p_1p_2$ with $p_i \le x^{1/4}$, so we have $x^{1/2} \le t \le x$. Let us choose $\varepsilon \in (0, 1/5]$ so that $\log_2^{4/5} x \le q_i^{k_i} \le \log_2^{6/5} x$ and so $\log_2^{4/5} t \le q_i^{k_i} \le 2\log_2^{6/5} t$ if x is sufficiently large. Moreover, if we take $\varepsilon \ll \sum_{i=1,2} \frac{1}{q_i^{k_i}}$, then we may use Lemma 4.3.3 to show

that the conditions of Lemma 4.3.2 are satisfied, which implies that

$$\begin{aligned}\sum_{m\le x/p_1p_2}{}'' 1 &= tW(z)+O\big(\varepsilon t(\log_3 t)W(z)\big)+O\left(\frac{\varepsilon t}{\log_2^2 t}\right)\\ &= tW(z)+O\left(t(\log_3 t)W(z)\sum_{i=1,2}\frac{1}{q_i^{k_i}}\right)+O\left(\frac{t}{\log_2^2 t}\sum_{i=1,2}\frac{1}{q_i^{k_i}}\right),\end{aligned}$$

where $z=\exp(\log t/\log_2^2 t)$ and $W(z) = \displaystyle\prod_{\substack{p\le z\\ q_1^{k_1}\mid p-1 \text{ or } q_2^{k_2}\mid p-1}}\left(1-\frac{1}{p}\right)$.

From the above setup, we can see that

$$\sum_{m\le x/p_1p_2}{}'' 1 = tW(z)+O\left(\frac{t\log_3 x}{\log_2^{1-\varepsilon}x}W(z)\right)+O\left(\frac{t}{\log_2^{3-\varepsilon}x}\right). \tag{4.16}$$

We now wish to evaluate $W(z)$.

$$\begin{aligned}\log W(z) &= \sum_{\substack{p\le z\\ q_1^{k_1}\mid p-1 \text{ or } q_2^{k_2}\mid p-1}}\log\left(1-\frac{1}{p}\right)\\ &= -\sum_{\substack{p\le z\\ q_1^{k_1}\mid p-1 \text{ or } q_2^{k_2}\mid p-1}}\frac{1}{p}-\sum_{\substack{p\le z\\ q_1^{k_1}\mid p-1 \text{ or } q_2^{k_2}\mid p-1}}\sum_{j=2}^{\infty}\frac{1}{j\,p^j}.\end{aligned}$$

By Lemma 4.3.1,

$$\begin{aligned}-\sum_{\substack{p\le z\\ q_1^{k_1}\mid p-1 \text{ or } q_2^{k_2}\mid p-1}}\frac{1}{p} &= -\sum_{\substack{p\le z\\ q_1^{k_1}\mid p-1}}\frac{1}{p}-\sum_{\substack{p\le z\\ q_2^{k_2}\mid p-1}}\frac{1}{p}+O\left(\sum_{\substack{p\le z\\ q_1^{k_1}q_2^{k_2}\mid p-1}}\frac{1}{p}\right)\\ &= -\sum_{i=1,2}\frac{\log_2 z+O(\log q_i^{k_i})}{\phi(q_i^{k_i})}+O\left(\frac{\log_2 z+O(\log q_1^{k_1}q_2^{k_2})}{\phi(q_1^{k_1}q_2^{k_2})}\right).\end{aligned}$$

We now wish to collect all the error terms into a single term. First, since

$$\log_2^{1-\varepsilon}x\le q_i^{k_i},$$

then

$$\sum_{\substack{p\le z\\ q_1^{k_1}|p-1 \text{ or } q_2^{k_2}|p-1}} \sum_{j=2}^{\infty} \frac{1}{j\,p^j} \;\ll\; \sum_{\substack{p\le z\\ q_1^{k_1}\,|\,p-1}} \frac{1}{p^2} + \sum_{\substack{p\le z\\ q_2^{k_2}\,|\,p-1}} \frac{1}{p^2}$$
$$\ll\; \frac{1}{q_1^{2k_1}} + \frac{1}{q_2^{2k_2}} \;\ll\; \frac{1}{\log_2^{2-2\varepsilon} x}\,.$$

Second, since $\log_2 z = \log_2 x + O(\log_3 x)$ and $\log q_i^{k_i} = O(\log_3 x)$, then

$$\sum_{i=1,2} \frac{\log_2 z + O(\log q_i^{k_i})}{\phi(q_i^{k_i})} \;=\; \sum_{i=1,2} \frac{\log_2 x}{\phi(q_i^{k_i})} + O\left(\frac{\log_3 x}{\log_2^{1-\varepsilon} x}\right)$$

and

$$\frac{\log_2 z + O(\log q_1^{k_1} q_2^{k_2})}{\phi(q_1^{k_1} q_2^{k_2})} \;\ll\; \frac{\log_2 x}{\log_2^{2-2\varepsilon} x} \;\ll\; \frac{1}{\log_2^{1-\varepsilon} x}\,.$$

The above estimates allow us to conclude that

$$\begin{aligned} W(z) \;&=\; \exp\left(-\sum_{i=1,2} \frac{\log_2 x}{\phi(q_i^{k_i})} + O\left(\frac{1}{\log_2^{1-2\varepsilon} x}\right)\right) \\ &=\; \exp\left(-\sum_{i=1,2} \frac{\log_2 x}{\phi(q_i^{k_i})}\right)\left(1 + O\left(\frac{1}{\log_2^{1-2\varepsilon} x}\right)\right). \end{aligned} \tag{4.17}$$

Combining (4.16) and (4.17) yields

$$\sum_{m\le x/p_1p_2}'' 1 \;=\; t\,\exp\left(-\sum_{i=1,2} \frac{\log_2 x}{\phi(q_i^{k_i})}\right) + O\left(\frac{t}{\log_2^{1-2\varepsilon} x}\right). \tag{4.18}$$

Put (4.18) into (4.15). The first term in (4.18) provides the main term for $\sum_{\substack{n\le x\\ \Delta_{q_i}(n)=1}} 1$, which is

$$x \sum_{k_i}{}' \sum_{\substack{p_i\le x^{1/4}\\ q_i^{k_i}\,\|\,p_i-1}} \frac{1}{p_1p_2} \exp\left(-\sum_{i=1,2} \frac{\log_2 x}{\phi(q_i^{k_i})}\right). \tag{4.19}$$

Then, we have

$$\begin{aligned}\sum_{\substack{p_i\le x^{1/4}\\ q_i^{k_i}\,\|\,p_i-1}}\frac{1}{p_1p_2} &= \left(\sum_{\substack{p_1\le x^{1/4}\\ q_1^{k_1}\,\|\,p_1-1}}\frac{1}{p_1}\right)\left(\sum_{\substack{p_2\le x^{1/4}\\ q_2^{k_2}\,\|\,p_2-1}}\frac{1}{p_2}\right)\\ &= \left(\sum_{\substack{p_1\le x^{1/4}\\ q_1^{k_1}\,|\,p_1-1}}\frac{1}{p_1}-\sum_{\substack{p_1\le x^{1/4}\\ q_1^{k_1+1}\,|\,p_1-1}}\frac{1}{p_1}\right)\left(\sum_{\substack{p_2\le x^{1/4}\\ q_2^{k_2}\,|\,p_2-1}}\frac{1}{p_2}-\sum_{\substack{p_2\le x^{1/4}\\ q_2^{k_2+1}\,|\,p_2-1}}\frac{1}{p_2}\right).\end{aligned}$$

From Lemma 4.3.1, it follows that

$$\begin{aligned}\sum_{\substack{p_i\le x^{1/4}\\ q_i^{k_i}\,|\,p_i-1}}\frac{1}{p_i}-\sum_{\substack{p_i\le x^{1/4}\\ q_i^{k_i+1}\,|\,p_i-1}}\frac{1}{p_i} &= \frac{\log_2 x}{\phi(q_i^{k_i})}+O\left(\frac{\log q_i^{k_i}}{q_i^{k_i}}\right)-\frac{\log_2 x}{\phi(q_i^{k_i+1})}+O\left(\frac{\log q_i^{k_i+1}}{q_i^{k_i+1}}\right)\\ &= \log_2 x\left(\frac{1}{q_i^{k_i-1}(q_i-1)}-\frac{1}{q_i^{k_i}(q_i-1)}\right)+O\left(\frac{\log q_i^{k_i}}{q_i^{k_i}}\right)\\ &= \frac{\log_2 x}{q_i^{k_i}}+O\left(\frac{\log q_i^{k_i}}{q_i^{k_i}}\right).\end{aligned}$$

Thus

$$\sum_{\substack{p_i\le x^{1/4}\\ q_i^{k_i}\,\|\,p_i-1}}\frac{1}{p_1p_2} = \prod_{i=1,2}\frac{\log_2 x+O(\log q_i^{k_i})}{q_i^{k_i}} = \frac{\log_2^2 x+O(\log_2 x\log_3 x)}{q_1^{k_1}q_2^{k_2}}.$$

Therefore we have that

$$\begin{aligned}& x\sum_{k_i}{}'\sum_{\substack{p_i\le x^{1/4}\\ q_i^{k_i}\,\|\,p_i-1}}\frac{1}{p_1p_2}\exp\left(-\sum_{i=1,2}\frac{\log_2 x}{\phi(q_i^{k_i})}\right)\\ =\;& x\sum_{k_i}{}'\frac{\log_2^2 x}{q_1^{k_1}q_2^{k_2}}\exp\left(-\sum_{i=1,2}\frac{\log_2 x}{\phi(q_i^{k_i})}\right)+O\left(x\sum_{k_i}{}'\frac{\log_2 x\log_3 x}{q_1^{k_1}q_2^{k_2}}\right)\\ =\;& x\sum_{k_i}{}'\frac{\log_2^2 x}{q_1^{k_1}q_2^{k_2}}\exp\left(-\sum_{i=1,2}\frac{\log_2 x}{\phi(q_i^{k_i})}\right)+O\left(\frac{x\log_3 x}{\log_2^{1-2\varepsilon}x}\right). \qquad (4.20)\end{aligned}$$

What is left to estimate is the accumulation of the error of (4.18) into (4.15). From

our previous calculations, we see that it is bounded above by

$$\left(\frac{x}{\log_2^{1-2\varepsilon} x}\right) \sideset{}{'}\sum_{k_i} \sum_{\substack{p_i \le x^{1/4} \\ q_i^{k_i} \,\|\, p_i - 1}} \frac{1}{p_1 p_2} \;\ll\; \left(\frac{x}{\log_2^{1-2\varepsilon} x}\right) \sideset{}{'}\sum_{k_i} \frac{\log_2^2 x}{q_1^{k_1} q_2^{k_2}} \;\ll\; \frac{x}{\log_2^{1-4\varepsilon} x}\,.$$

Assuming $\varepsilon \in (0, 1/5]$, the above estimate and (4.20) yields that

$$\sum_{\substack{n \le x \\ \Delta_{q_i}(n)=1}} 1 \;=\; x \sideset{}{'}\sum_{k_i} \frac{\log_2^2 x}{q_1^{k_1} q_2^{k_2}} \exp\left(-\sum_{i=1,2} \frac{\log_2 x}{\phi(q_i^{k_i})}\right) + O\left(\frac{x}{\log_2^{\varepsilon} x}\right). \tag{4.21}$$

Lemma 4.6.1 then follows by substituting (4.21) into (4.10).

□

4.7 Extreme Behavior of $D(x, u)$

In this section, with the help of Lemma 4.5.1 and Lemma 4.6.1, we will prove the following theorem where given $x \in \mathbb{R}^+$ and $0 < u < 1$, we define

$$D(x, u) := \frac{1}{x} \sum_{\substack{n \le x \\ r(n) \le u}} 1\,.$$

Theorem 4.7.1.

$$\limsup_{x \to \infty} D(x, \log_5^{-c_1} x) = 1$$

for some constant $c_1 > 0$.

We first observe that Theorem 4.7.1 is equivalent to

$$\liminf_{x \to \infty} \; \#\{\, n \le x \mid r(n) > \log_5^{-c_1} \,\} \;=\; 0\,.$$

Moreover, since the function $1/x \;+\; \log(1 - 1/x)$ is negative for $x \in [2, \infty)$, combining it with (4.1) yields

$$r(n) \;=\; \exp\left(\sum_{q \,|\, \lambda(n)} \log\left(1 - \frac{1}{q^{\Delta_q(n)}}\right)\right) \;\le\; \exp\left(-\tilde{f}(n)\right).$$

An important fact about the function $\tilde{f}$ is that we can get control of the upper

order of its first and second moments as we shall see below. This allows us to extract information related to $r(n)$.

Lemma 4.7.1. *For all $x \geq 1$, we have*

$$\sum_{n \leq x} \tilde{f}^2(n) \;=\; \frac{1}{x}\left(\sum_{n \leq x} \tilde{f}(n)\right)^2 + O(x)\,.$$

Proof. Taking $\varepsilon = 1/5$ in Lemma 4.5.1 we have

$$\sum_{n \leq x} \tilde{f}(n) \;=\; x \sum_{q \leq \log_4 x} \frac{1}{q} {\sum_{k \geq 1}}' \frac{\log_2 x}{q^k} \exp\left(-\frac{\log_2 x}{\phi(q^k)}\right) + O\left(\frac{x}{\log_4 x}\right), \tag{4.22}$$

where $\sum'$ means that the sum is taken over all $k \geq 1$ in the interval $\left[\frac{4\log_3 x}{5\log q}, \frac{6\log_3 x}{5\log q}\right]$. Now, observe that

$$\begin{aligned}
&\sum_{k \geq 1} \frac{\log_2 x}{q^k} \exp\left(-\frac{\log_2 x}{\phi(q^k)}\right) \;\leq\; \sum_{k \geq 1} \frac{\log_2 x}{q^k} \exp\left(-\frac{\log_2 x}{q^k}\right) \\
&\quad= \sum_{k \leq \log_3 x/\log q} \frac{\log_2 x}{q^k} \exp\left(-\frac{\log_2 x}{q^k}\right) + \sum_{k > \log_3 x/\log q} \frac{\log_2 x}{q^k} \exp\left(-\frac{\log_2 x}{q^k}\right).
\end{aligned}$$

The choice of partition is justified by the fact that

$$\frac{\log_2 x}{q^k} < 1 \iff k > \frac{\log_3 x}{\log q}\,.$$

Considering the second term, since $q \geq 2$, we see that

$$\begin{aligned}
\sum_{k > \log_3 x/\log q} \frac{\log_2 x}{q^k} \exp\left(-\frac{\log_2 x}{q^k}\right) \;&\leq\; \sum_{k > \log_3 x/\log q} \frac{\log_2 x}{q^k} \\
&\leq\; (\log_2 x)\left(\frac{q^{-\log_3 x/\log q}}{1-q^{-1}}\right) \\
&=\; \frac{1}{1-q^{-1}} \;\leq\; 2\,.
\end{aligned}$$

Before we can provide a bound for the first term, we need to make the following two observations. First,

$$k \leq \frac{\log_3 x}{\log q} \implies \frac{\log_2 x}{q^k} - \frac{\log_2 x}{q^{k+1}} \;=\; \frac{\log_2 x}{q^k}\left(1 - \frac{1}{q}\right) \;\geq\; 1 - \frac{1}{q} \;\geq\; \frac{1}{2}\,.$$

Second, the function te^{-t} is strictly decreasing in the interval $[1, \infty)$ and is bounded

by 1. Interpreting the first term as a Riemann sum, we can see that

$$\begin{aligned}\sum_{k\le \log_3 x/\log q} \frac{\log_2 x}{q^k}\exp\left(-\frac{\log_2 x}{q^k}\right) &\le 2\left(\int_1^\infty te^{-t}dt\right)+1\\ &= 2\left(\frac{2}{e}\right)+1 \;=\; \frac{4}{e}+1\,.\end{aligned}$$

Therefore we conclude that

$$\sum_{k\ge 1}\frac{\log_2 x}{q^k}\exp\left(-\frac{\log_2 x}{\phi(q^k)}\right) \;\le\; 3+\frac{4}{e}\,,$$

which holds independently of x and q for any allowed values of x and q. This implies that

$$\sum_{q\le \log_4 x}\frac{1}{q}\sum_{k\ge 1}\frac{\log_2 x}{q^k}\exp\left(-\frac{\log_2 x}{\phi(q^k)}\right) \;=\; O\left(\sum_{q\le \log_4 x}\frac{1}{q}\right) \;=\; O(\log_6 x)\,.$$

Using this fact and squaring (4.22) yields

$$\frac{1}{x}\left(\sum_{n\le x}\tilde{f}(n)\right)^2 \;=\; x\sum_{q_1,q_2\le \log_4 x}{\sum_{k_1,k_2\ge 1}}'\frac{\log_2^2 x}{q_1^{k_1+1}q_2^{k_2+1}}\exp\left(-\sum_{j=1,2}\frac{\log_2 x}{\phi(q_j^{k_j})}\right)+O\left(\frac{x\log_6 x}{\log_4 x}\right)\,,$$

where $\sum'$ carries similar meaning as above. The main term above is the same as that of $\sum_{n\le x}\tilde{f}^2(n)$ given in Lemma 4.6.1 with the corresponding $\varepsilon = 1/5$. This completes the proof of Lemma 4.7.1. □

Corollary 4.7.1. *There is an unbounded set of numbers x for which we have*

$$\sum_{n\le x}\left(\tilde{f}(n)-c_5\log_6 x\right)^2 \;=\; o(x\log_6^2 x)$$

for some constant $c_5>0$, depending on the unbounded set of x's.

Proof. By [12, Theorem 5.1] there exists an unbounded set S of numbers x such that on the set S,

$$\sum_{n\le x}\tilde{f}(n) \;\ge\; bx\log_6 x$$

for some constant $2 \geq b > 0$. On the other hand, for $n \leq x$ and x sufficiently large,

$$\tilde{f}(n) \;\leq\; \sum_{q \leq \log_2 x} \frac{1}{q} \;\leq\; \log_6 x + O(1) \;\leq\; 2\log_6 x\,.$$

Thus

$$\sum_{n \leq x} \tilde{f}(n) \;\leq\; 2x \log_6 x\,.$$

It therefore follows that for any given sufficiently large $x \in S$,

$$\sum_{n \leq x} \tilde{f}(n) \;=\; b_x(x \log_6 x)$$

for some $b_x \in [b, 2]$. By compactness of the interval $[b, 2]$ the sequence $\{\, b_x \,|\, x \in S \,\}$ has an accumulation point c_5 in the interval. Without loss of generality we can assume that

$$\lim_{\substack{x \in S \\ x \to \infty}} b_x \;=\; c_5\,.$$

Then when $x \in S$, by Lemma 4.7.1, we have

$$\begin{aligned}
\sum_{n \leq x} \Big(\tilde{f}(n) - c_5 \log_6 x\Big)^2 \;&=\; \sum_{n \leq x} \tilde{f}^2(n) \;-\; (c_5 \log_6 x) 2 \sum_{n \leq x} \tilde{f}(n) \;+\; x(c_5 \log_6 x)^2 \\
&=\; \frac{1}{x}\big(b_x(x \log_6 x)\big)^2 + O(x) - 2(c_5 \log_6 x)\big(b_x(x \log_6 x)\big) + x(c_5 \log_6 x)^2 \\
&=\; (b_x - c_5)^2 x \log_6^2 x + O(x) = o(x \log_6^2 x)\,.
\end{aligned}$$

This completes the proof. □

Proof. (*of Theorem 4.7.1*) As we noted at the beginning of the section, Theorem 4.7.1 is equivalent to the following statement:

$$\#\{\, n \leq x \,|\, r(n) > \log_5^{-c_5/2} x \,\} \;=\; o(x)$$

on an unbounded set of numbers x. Let S be the unbounded set of numbers x in Corollary 4.7.1. Then by Corollary 4.7.1 we have

$$\lim_{\substack{x \to \infty \\ x \in S}} \frac{1}{x} \# \left\{\, n \leq x \,|\, \tilde{f}(n) \leq \frac{c_5}{2} \log_6 x \,\right\} \;=\; 0\,.$$

Since $r(n) \le \exp(-\tilde{f}(n))$,

$$\{\, n \le x \,|\, r(n) > \log_5^{-c_5/2} x \,\} \subseteq \left\{ n \le x \,|\, \tilde{f}(n) \le \frac{c_5}{2} \log_6 x \right\}.$$

Theorem 4.7.1 follows from the above argument and by letting $c_1 = c_5/2$. □

4.8 Average Order of $N_a(x)$

If we let $R(n)$ be the number of primitive λ-roots modulo n in the interval $[1, n]$, then we notice that $R(n) = r(n)\phi(n)$ and it follows from the definition of $N_a(x)$ that

$$\begin{aligned} \sum_{a \le y} N_a(x) &= \sum_{a \le y} \sum_{\substack{n \le x \\ l_a(n) = \lambda(n)}} 1 = \sum_{n \le x} \sum_{\substack{a \le y \\ l_a(n) = \lambda(n)}} 1 \\ &= \sum_{n \le x} \left\lfloor \frac{y}{n} \right\rfloor R(n) + \sum_{n \le x} \sum_{\substack{1 \le a \le \{y/n\}n \\ l_a(n) = \lambda(n)}} 1, \end{aligned} \tag{4.23}$$

where $\{y\}$ denotes the fractional part of y. It is then easy to see that for any $y \ge x \ge 1$, we have

$$\sum_{a \le y} N_a(x) \ge \sum_{n \le x} \left\lfloor \frac{y}{n} \right\rfloor R(n) \ge \frac{y}{2} \sum_{n \le x} \frac{R(n)}{n}$$

and

$$\sum_{a \le y} N_a(x) \le \sum_{n \le x} \left\lceil \frac{y}{n} \right\rceil R(n) \le 2y \sum_{n \le x} \frac{R(n)}{n}.$$

We therefore have just proved the following lemma.

Lemma 4.8.1. *For any $y \ge x$, we have*

$$\frac{1}{2} \sum_{n \le x} \frac{R(n)}{n} \le \frac{1}{y} \sum_{a \le y} N_a(x) \le 2 \sum_{n \le x} \frac{R(n)}{n}.$$

Observe that from (4.23) we have

$$\sum_{a \le y} N_a(x) = y \sum_{n \le x} \frac{R(n)}{n} + O\left(\sum_{n \le x} R(n) \right).$$

Since

$$\sum_{n\le x} R(n) \le x \sum_{n\le x} \frac{R(n)}{n}\,,$$

we have

$$\frac{1}{y}\sum_{a\le y} N_a(x) = \left(1 + O\left(\frac{x}{y}\right)\right)\sum_{n\le x}\frac{R(n)}{n}\,.$$

Hence, if

$$\lim_{x\to\infty}\frac{x}{y} = 0\,,$$

then

$$\frac{1}{y}\sum_{a\le y} N_a(x) \sim \sum_{n\le x}\frac{R(n)}{n}\,.$$

It would be preferable to let a run over a smaller interval, say $y \le \sqrt{x}$, in analogy with P. J. Stephen's average result discussed in the previous chapter.

Theorem 4.8.1. *[14, Li](Extreme orders of $R(n)/n$.)*

(i)

$$\limsup_{n\to\infty}\ \frac{R(n)}{n} = 1$$

(ii)

$$\liminf_{n\to\infty}\ \frac{R(n)}{n}(\log_2 n)^2 = e^{-2\gamma}$$

where γ is Euler's constant.

Proof. (i) Since $R(n)/n \le 1$, we only need to find a sequence of positive integers n_x such that $\lim_{x\to\infty} n_x = \infty$ and $\lim_{x\to\infty} R(n_x)/n_x = 1$. Let

$$\mathcal{B} := \{\ p \le x \,|\, p \equiv 3 (\mathrm{mod}\, 4) \text{ and } \gcd\left(p-1, P(x^{1/5})\right) = 1\ \}$$

where

$$P(z) := \prod_{2<p\le z} p\,.$$

Then, applying [9, Theorem 7.4] to sieve the set $\mathcal{A} := \{\ p-1 \,|\, p \le x \text{ and } p \equiv 3(\mathrm{mod}\, 4)\ \}$ with the set of primes $\mathcal{P} := \{\ p \,|\, 2 < p \le x^{1/5}\ \}$, taking $\kappa = 1$ and $\alpha = 1/2$ yields

$$\#\mathcal{B} \ge \delta\frac{x}{\log^2 x}$$

for some constant $\delta > 0$ and for all $x \ge 3$. Let $p \in \mathcal{B}$. If q is an odd prime factor of $p-1$ then $q > x^{1/5}$. Since $p \le x$, it follows that $p-1$ has at most 5 odd

prime factors, counting multiplicity. Choose $\lfloor \log x \rfloor$ such primes $\{p_i\}_{i=1}^{\lfloor \log x \rfloor} \subset \mathcal{B}$, $p_1 < p_2 < ...$, and let

$$n_x := \prod_{i=1}^{\lfloor \log x \rfloor} p_i \, .$$

Then we have

$$\lambda(n_x) = lcm_{p_i}\{\lambda(p_i)\} = lcm_{p_i}\{p_i - 1\}$$

and

$$(\mathbb{Z}/n_x\mathbb{Z})^* \simeq \bigoplus_{i=1}^{\lfloor \log x \rfloor} (\mathbb{Z}/p_i\mathbb{Z})^* \simeq \bigoplus_{i=1}^{\lfloor \log x \rfloor} C_{p_i-1} \, .$$

Now, since $p \equiv 3 (\mathrm{mod}\, 4)$, then $C_{p_i-1} \simeq C_2 \bigoplus H_{p_i}$ where H_{p_i} is a direct sum of cyclic groups with odd orders. This implies that $\Delta_2(n_x) = \lfloor \log x \rfloor$. Moreover, since $p_i - 1$ has at most 5 odd prime factors, we see that $\lambda(n_x)$ has at most $5\lfloor \log x \rfloor$ distinct odd prime factors. Finally, if q is an odd prime factor of $\lambda(n_x)$, then $q > x^{1/5}$ and so $q^{\Delta_q(n_x)} > x^{1/5}$. Thus, by the definition of n_x and (4.1), we have

$$r(n_x) = \prod_{q \,|\, \lambda(n_x)} \left(1 - \frac{1}{q^{\Delta_q(n_x)}}\right) \geq \left(1 - \frac{1}{2^{\lfloor \log x \rfloor}}\right) \left(1 - \frac{1}{x^{1/5}}\right)^{5\lfloor \log x \rfloor} ,$$

while

$$\frac{\phi(n_x)}{n_x} = \prod_{i=1}^{\lfloor \log x \rfloor} \left(1 - \frac{1}{p_i}\right) \geq \left(1 - \frac{1}{x^{1/5}}\right)^{\lfloor \log x \rfloor} .$$

Note that

$$\lim_{x \to \infty} \left(1 - \frac{1}{2^{\lfloor \log x \rfloor}}\right) \left(1 - \frac{1}{x^{1/5}}\right)^{5\lfloor \log x \rfloor} = 1$$

and

$$\lim_{x \to \infty} \left(1 - \frac{1}{x^{1/5}}\right)^{\lfloor \log x \rfloor} = 1 \, .$$

Since $r(n) \leq 1$ and $\phi(n)/n \leq 1$ for all $n \geq 1$, then

$$\lim_{x \to \infty} r(n_x) = 1 \quad \text{and} \quad \lim_{x \to \infty} \frac{\phi(n_x)}{n_x} = 1 \, .$$

Since $R(n)/n = r(n)\phi(n)/n$ for all $n \geq 1$, we conclude that

$$\lim_{x \to \infty} \frac{R(n_x)}{n_x} = 1 \, .$$

This concludes the proof of *(i)*.

(ii) If we let

$$N(n) := \min_{m \geq 1} \left\{ \prod_{p \leq m} p \geq \lambda(n) \right\},$$

then

$$r(n) \geq \prod_{q \,|\, \lambda(n)} \left(1 - \frac{1}{q}\right) \geq \prod_{q \leq N(n)} \left(1 - \frac{1}{q}\right).$$

From Theorem 4.2.4, we get

$$N(n) = \big(1 + o(1)\big) \log \lambda(n) \leq \big(1 + o(1)\big) \log n \,.$$

From Theorem 4.2.3, it follows that

$$r(n) \geq \frac{e^{-\gamma} + o(1)}{\log_2 n} \,.$$

Moreover, Theorems 4.2.3 and 4.2.4 yield

$$\frac{\phi(n)}{n} \geq \frac{e^{-\gamma} + o(1)}{\log_2 n} \,,$$

thus

$$\frac{R(n)}{n} = r(n) \frac{\phi(n)}{n} \geq \left(\frac{e^{-\gamma} + o(1)}{\log_2 n} \right)^2 = \frac{e^{-2\gamma}\big(1 + o(1)\big)}{(\log_2 n)^2} \,.$$

Hence

$$\liminf_{n \to \infty} \ \frac{R(n)}{n} (\log_2 n)^2 \geq e^{-2\gamma} \,.$$

We now need to show that this is best possible.

For each prime $q < \log x$, let a_q be the least integer such that $q^{a_q} > \log x$ and let

$$m = \prod_{q < \log x} q^{a_q} \,.$$

Then, for x sufficiently large, we have from the definition of m and the prime number theorem that

$$x^{1/2} \leq (\log x)^{\pi(\log x)} \leq m \leq \left(\log^2 x\right)^{\pi(\log x)} \leq x^3 \,.$$

Moreover, we know from Theorem 4.2.2 that there exists a prime p_0 such that

$m < p_0 \le m^{c_3}$ and $p_0 \equiv 1 (\mathrm{mod}\, m)$ where c_3 is an absolute constant. If we let

$$n'_x := p_0 \prod_{p \le \log x} p\,,$$

then, since

$$x^{1/2} < \prod_{p \le \log x} p < x^2 \quad \text{and} \quad x^{1/2} < p_0 \le x^{3c_3}\,,$$

we have

$$x < n'_x \le x^{3c_3+2}.$$

Let $q \le \log x$. If q is a prime factor of $p-1$ and $p \le \log x$, then its maximal power in $p-1$ is less than that in $p_0 - 1$ (by the definitions of m and p_0). Thus, it follows that $\Delta_q(n'_x) = 1$ for all $q \le \log x$. Observing that $q \le \log x \Rightarrow q \,|\, \lambda(n'_x)$, we see that

$$r(n'_x) = \prod_{q \,|\, \lambda(n'_x)} \left(1 - \frac{1}{q^{\Delta_q(n'_x)}}\right) \le \prod_{q \le \log x} \left(1 - \frac{1}{q}\right)$$

and

$$\frac{\phi(n'_x)}{n'_x} = \prod_{p \,|\, n'_x} \left(1 - \frac{1}{p}\right) \le \prod_{p \le \log x} \left(1 - \frac{1}{p}\right).$$

Finally, noticing that $\log_2 x = \log_2 n'_x + O(1)$ yields

$$\prod_{p \le \log x} \left(1 - \frac{1}{p}\right) = \frac{e^{-\gamma}\big(1 + o(1)\big)}{\log_2 x} = \frac{e^{-\gamma}\big(1 + o(1)\big)}{\log_2 n'_x}$$

which implies that

$$\liminf_{x \to \infty} \frac{R(n'_x)}{n'_x} (\log_2 n'_x)^2 \le e^{-2\gamma}\,.$$

This completes the proof. □

Using Theorem 4.8.1, one can obtain a lower bound of the form

$$\frac{1}{x} \sum_{1 \le a \le x} N_a(x) \gg \frac{x}{(\log_2 x)^2}\,.$$

However, this can be improved as in the next theorem.

Theorem 4.8.2. *For a positive constant x_0, we have*

$$\frac{1}{x} \sum_{1 \le a \le x} N_a(x) \gg \frac{x}{\log_3 x}$$

for all $x \geq x_0$.

Proof. By Lemma 4.8.1,

$$\frac{1}{x}\sum_{1\leq a\leq x} N_a(x) \geq \frac{1}{2}\sum_{n\leq x}\frac{R(n)}{n} = \frac{1}{2}\sum_{n\leq x} r(n)\frac{\phi(n)}{n}$$

and by (4.1),

$$r(n) \geq \prod_{q\,|\,\lambda(n)}\left(1-\frac{1}{q}\right) = \prod_{q\,|\,\phi(n)}\left(1-\frac{1}{q}\right).$$

Let $\mathcal{S}$ be the set of integers $n \geq 1$ such that $\phi(n)$ has at most $(\log_2 n)^2$ distinct prime factors. By [7, 22], $\mathcal{S}$ has density 1 and by Theorem 4.2.3, we have

$$r(n) \gg \frac{1}{\log_3 n}$$

uniformly for all $n \in \mathcal{S}$ and $n \geq n_0$ for some $n_0 > 0$. On the other hand, if we let

$$\mathcal{S}' = \left\{ n \;\middle|\; \frac{\phi(n)}{n} \geq \frac{1}{2} \right\},$$

then $\mathcal{S}'$ has density greater than zero by Lemma 4.2.1. Therefore we have

$$\frac{1}{x}\sum_{1\leq a\leq x} N_a(x) \gg \sum_{\substack{n_0\leq n\leq x\\ n\in\mathcal{S}\cap\mathcal{S}'}}\frac{1}{\log_3 n} \gg \frac{x}{\log_3 x},$$

since the density of $\mathcal{S}\cap\mathcal{S}'$ is equal to $1-w(1/2) > 0$. This completes the proof. □

Lemma 4.8.2. *There exists a positive constant* c_4 *and an unbounded set of numbers* $\tilde{\mathcal{S}}$ *such that if* $x \in \tilde{\mathcal{S}}$, *then*

$$D(x,u) \leq \frac{c_4}{|\log u|}$$

for all u *with* $0 < u < 1$.

Proof. See [12, Corollary 5.2]. □

4.9 Proof of Theorem 4.1.1

We now have everything in place to prove our main result.

Let $u, t \in (0,1)$ be constants to be fixed later. We have from Lemma 4.2.1 and 4.8.2 that, for $x \in \tilde{\mathcal{S}}$ sufficiently large,

$$\begin{aligned}\sum_{n\le x}\frac{R(n)}{n} &= \sum_{n\le x} r(n)\frac{\phi(n)}{n} \ge u\sum_{\substack{n\le x\\ r(n)\ge u}}\frac{\phi(n)}{n} \ge ut\sum_{\substack{n\le x\\ r(n)\ge u\\ \phi(n)/n\ge t}} 1\\ &\ge ut\left(\lfloor x\rfloor - \sum_{\substack{n\le x\\ r(n)\le u}} 1 - \sum_{\substack{n\le x\\ \phi(n)/n<t}} 1\right)\\ &\ge ut\left(x - \frac{c_4 x}{|\log u|} - 2w(t)x\right).\end{aligned}$$

By Lemma 4.2.1, we know that

$$\lim_{t\to 0^+} w(t) = 0\,,$$

thus we can choose u and t small enough to ensure that

$$c_2 := \frac{ut}{2}\left(1 - \frac{c_4}{|\log u|} - 2w(t)\right) > 0\,.$$

Lemma 4.8.1 implies that for $x \in \tilde{\mathcal{S}}$ sufficiently large,

$$\frac{1}{x}\sum_{1\le a\le x} N_a(x) \ge c_2 x\,.$$

Therefore

$$\limsup_{x\to\infty}\ \frac{1}{x^2}\sum_{1\le a\le x} N_a(x) \ > \ 0\,.$$

By Theorem 4.7.1, on an unbounded set of numbers x, we have

$$\begin{aligned}
\sum_{n\le x}\frac{R(n)}{n} &= \sum_{n\le x} r(n)\frac{\phi(n)}{n} \quad \le \quad \sum_{n\le x} r(n) \\
&= \sum_{\substack{n\le x \\ r(n)\le \log_5^{-c_1} x}} r(n) \quad + \sum_{\substack{n\le x \\ r(n)> \log_5^{-c_1} x}} r(n) \\
&\le \sum_{\substack{n\le x \\ r(n)\le \log_5^{-c_1} x}} \frac{1}{\log_5^{c_1} x} \quad + \sum_{\substack{n\le x \\ r(n)> \log_5^{-c_1} x}} 1 \\
&\le \frac{x}{\log_5^{c_1} x} \quad + \sum_{\substack{n\le x \\ r(n)> \log_5^{-c_1} x}} 1 \\
&\le \frac{x}{\log_5^{c_1} x} \quad + \quad x\left(1 - D(x, \log_5^{-c_1} x)\right) \\
&= o(x)\,.
\end{aligned}$$

We thus conclude from Lemma 4.8.1 that

$$\liminf_{x\to\infty} \frac{1}{x^2} \sum_{1\le a\le x} N_a(x) = 0\,.$$

This completes the proof.

4.10 Related Results and Recent Developments

In this chaper, we averaged $N_a(x)$ over a in the interval $[1, x]$. Our main result was the proof of Theorem 4.1.1, but can we do better? Namely, can we average $N_a(x)$ over a in the interval $[1, y]$ where y is less than x in order of magnitude? The best result up to date was obtained by S. Li [15], who was inspired by P. J. Stephens [26]. The result states that if $y > \exp\big((\log x)^{3/4}\big)$, then

$$\limsup_{x\to\infty} \frac{1}{xy} \sum_{1\le a\le y} N_a(x) \;>\; 0 \quad \text{ and } \quad \liminf_{x\to\infty} \frac{1}{xy} \sum_{1\le a\le y} N_a(x) \;=\; 0\,.$$

Interesting results have also been obtained concerning individual $N_a(x)$'s. In analogy to Artin's primitive root conjecture, one is tempted to guess that if a is not in some exceptional set, then there exists a positive constant $B(a)$ such that

$$N_a(x) \sim B(a)x\,.$$

As we will now see, this is not the case. In [13], S. Li showed that for any integer a,

$$\liminf_{x\to\infty} N_a(x)/x = 0\,.$$

Conversely, if we denote by $\mathcal{E}$ the set of integers which are a power with an exponent larger than 1, or a square times either -1 or ± 2, then S. Li and C. Pomerance [23] were able to demonstrate the following theorem.

Theorem 4.10.1. *On assumption of the GRH, there is a positive number C such that if a is an integer with $a \notin \mathcal{E}$, then*

$$\limsup_{x\to\infty} N_a(x)/x \geq C\phi(|a|)/|a|\,.$$

Moreover, there is an unbounded set $\mathcal{D}$ of positive real numbers such that for any $a \notin \mathcal{E}$,

$$\liminf_{\substack{x\to\infty \\ x\in\mathcal{D}}} N_a(x)/x \geq C\phi(|a|)/|a|\,.$$

The set $\mathcal{E}$ is the analogue of the exceptional set in Artin's primitive root conjecture.

To conclude, we wish to mention that for some $a \in \mathbb{Z}$, it is possible to prove unconditionally that a is a primitive λ-root for infinitely many integers n. For example, if a is a primitive λ-root for p^2, where p is an odd prime, then a is also a primitive λ-root for p^j for every $j \geq 2$ (see [1, p.209]). This is an important distinction from Artin's primitive root conjecture since in this case, we still cannot prove unconditionally that any given a is a primitive root modulo p for infinitely many primes p.

Appendices

Appendix A

Results From Algebraic Number Theory

Proof. (*Theorem 2.3.2*) From the above congruence, we automatically have that $(p, f_1(\theta))^{e_1} \cdots (p, f_g(\theta))^{e_g} \subseteq p\mathcal{O}_K$. Now, since $f_i(x)$ is irreducible in $\mathbb{F}_p[x]$, then $\mathbb{F}_p[x]/(f_i(x))$ is a field. Moreover,

$$\mathbb{Z}[x]/(p) \simeq \mathbb{F}_p[x] \;\Rightarrow\; \mathbb{Z}[x]/(p, f_i(x)) \simeq \mathbb{F}_p[x]/(f_i(x))\,,$$

and so $\mathbb{Z}[x]/(p, f_i(x))$ is a field.

Let us now consider the map $\varphi_i : \mathbb{Z}[x] \mapsto \mathbb{Z}[\theta]/(p, f_i(\theta))$. Our goal now is to show that $\ker(\varphi_i) = (p, f_i(x))$. Clearly

$$(p, f_i(x)) \subseteq \ker(\varphi_i) = \{\, h(x) \,|\, h(\theta) \in (p, f_i(\theta)) \,\}\,.$$

If $h(x) \in \ker(\varphi_i)$, we can divide by $f_i(x)$ to yield

$$h(x) = q(x)f_i(x) + r_i(x)\,,\ \deg(r_i) < \deg(f_i)\,.$$

If $r_i(x)$ is the zero polynomial, then $h(x) \in (p, f_i(x))$ and we are done. Thus, we assume that $r_i(x)$ is not the zero polynomial. Since $h(\theta) \in (p, f_i(\theta))$, then $r_i(\theta) \in (p, f_i(\theta))$, so $r_i(\theta) = pa(\theta) + f_i(\theta)b(\theta)$. Here we have used the fact that

$\mathcal{O}_K = \mathbb{Z}[\theta]$. Now define the polynomial $H(x) := r_i(x) - pa(x) - f_i(x)b(x)$. Since $H(\theta) = 0$ and $f(x)$ is the minimal polynomial of θ, then $H(x) = G(x)f(x)$ for some polynomial $G(x) \in \mathbb{Z}[x]$. We conclude that $r_i(x) = p\tilde{a}(x) + f_i(x)\tilde{b}(x)$ for some $\tilde{a}(x), \tilde{b}(x) \in \mathbb{Z}[x]$. Therefore $r_i(x) \in (p, f_i(x))$ and so $\ker(\varphi_i) = (p, f_i(x))$. The first isomorphism theorem yields

$$\mathbb{Z}[\theta]/(p, f_i(\theta)) \simeq \mathbb{Z}[x]/(p, f_i(x)) \simeq \mathbb{F}_p[x]/(f_i(x))$$

and is therefore a field. This proves that $(p, f_i(\theta))$ is a maximal ideal of $\mathbb{Z}[\theta] = \mathcal{O}_K$, but a maximal ideal is necessarily prime, hence $(p, f_i(\theta))$ is a prime ideal of $\mathcal{O}_K$. We now let $\wp_i = (p, f_i(\theta))$ and as was previously observed at the beginning of the proof, we have that $\wp_1^{e_1} \cdots \wp_g^{e_g} \subseteq p\mathcal{O}_K$. This implies that $p\mathcal{O}_K = \wp_1^{e_1'} \cdots \wp_g^{e_g'}$ where $0 \leq e_i' \leq e_i$ is the ramification index of $\wp_i$, since for ideals to contain is to divide. Moreover, let d_i be the inertial degree of $\wp_i$. Then $d_i = [\mathcal{O}_K/\wp_i : \mathbb{Z}/(p)]$ and it is clear from the above isomorphisms that d_i is the degree of the polynomial $f_i(x)$. Furthermore, we know that

$$\sum_{i \leq g} e_i d_i = \deg(f) = n = [K : \mathbb{Q}] = \sum_{i \leq g} e_i' d_i$$

hence $e_i = e_i'$ for all $1 \leq i \leq g$. Therefore $p\mathcal{O}_K = \wp_1^{e_1} \cdots \wp_g^{e_g}$, which completes the proof. □

Proof. (*Theorem 2.3.3*) Given that $\mathbb{Z}[\theta]$ and $\mathcal{O}_K$ are both $\mathbb{Z}$-modules, we define $[\mathcal{O}_K : \mathbb{Z}[\theta]]$ to be the number of elements in the quotient module $\mathcal{O}_K/\mathbb{Z}[\theta]$. As we will show, this is finite since the degree of θ over $\mathbb{Q}$ is equal to n by assumption. Let $\omega_1, \ldots, \omega_n$ be an integral basis for $\mathcal{O}_K$ and observe that we can write

$$\mathbb{Z}[\theta] \;=\; \mathbb{Z} + \mathbb{Z}\theta + \cdots + \mathbb{Z}\theta^{n-1} \;=\; \mathbb{Z}\alpha_1 + \mathbb{Z}\alpha_2 + \cdots + \mathbb{Z}\alpha_n$$

where

$$\alpha_i = \sum_{j=1}^{n} a_{ij}\omega_j \quad \text{for } 1 \leq i \leq n \text{ and } a_{ij} \in \mathbb{Z} \text{ for } 1 \leq i, j \leq n\,.$$

This implies that

$$\begin{aligned}\mathbb{Z}[\theta] &= (a_{11}\mathbb{Z}\omega_1 + a_{21}\mathbb{Z}\omega_1 + \cdots + a_{n1}\mathbb{Z}\omega_1) + \cdots + (a_{1n}\mathbb{Z}\omega_n + a_{2n}\mathbb{Z}\omega_n + \cdots + a_{nn}\mathbb{Z}\omega_n) \\ &= \gcd(a_{11}, a_{21}, \ldots, a_{n1})\mathbb{Z}\omega_1 + \cdots + \gcd(a_{1n}, a_{2n}, \ldots, a_{nn})\mathbb{Z}\omega_n\end{aligned}$$

and since

$$\mathcal{O}_K = \mathbb{Z}\omega_1 + \mathbb{Z}\omega_2 + \cdots + \mathbb{Z}\omega_n\,,$$

we obtain that

$$m := [\mathcal{O}_K : \mathbb{Z}[\theta]] = \prod_{j=1}^{n} \gcd(a_{1j}, a_{2j}, ..., a_{nj}) \;<\; \infty\,.$$

We see from this that given any $\alpha \in \mathcal{O}_K$, $m\alpha \in \mathbb{Z}[\theta]$. More precisely, given any $\alpha \in \mathcal{O}_K$, we can write $m\alpha = b_0 + b_1\theta + \cdots + b_{n-1}\theta^{n-1}$. Consider this expression modulo p. Since m is coprime to p there exists an m' such that $mm' \equiv 1 \,(\mathrm{mod}\, p)$. Then

$$\alpha \equiv b_0m' + b_1m'\theta + \cdots + b_{n-1}m'\theta^{n-1} \,(\mathrm{mod}\, p)\,.$$

Since α was arbitrary, this implies that $\mathcal{O}_K \equiv \mathbb{Z}[\theta] \,(\mathrm{mod}\, p)$. Recall now that in the proof of the previous theorem, we only used the fact that $\mathcal{O}_K = \mathbb{Z}[\theta]$ at one point. This was when we wrote that $r_i(\theta) = pa(\theta) + f_i(\theta)b(\theta)$, but we simply need that $r_i(\theta) \equiv pa(\theta) + f_i(\theta)b(\theta) \,(\mathrm{mod}\, p)$ since

$$\mathcal{O}_K \equiv \mathbb{Z}[\theta] \,(\mathrm{mod}\, p) \;\;\Longrightarrow\;\; \mathcal{O}_K/(p) \simeq \mathbb{Z}[\theta]/(p)\,,$$

thus

$$\mathcal{O}_K/(p, f_i(\theta)) \simeq \mathbb{Z}[\theta]/(p, f_i(\theta))\,.$$

The rest of the argument is now identical to the one given in the demonstration of the previous theorem. This completes the proof. □

Proof. (*Theorem 2.3.5*) Let us first recall that given a finite group H acting on a set $\mathcal{B}$, for $b \in \mathcal{B}$, the orbit of b is defined by

$$\mathrm{O}_H(b) := \{\, h(b) \,|\, h \in H \,\}$$

and the stabilizer of b is defined by

$$S_H(b) := \{\, h \in H \,|\, h(b) = b \,\}\,.$$

It is a well-known fact from group theory that $|H| = |\mathrm{O}_H(b)|\,|S_H(b)|$ for any $b \in \mathcal{B}$. We first have that

$$\begin{aligned} n_k \;=\; |Gal(L_k/\mathbb{Q})| \;&=\; \left|\mathrm{O}_{Gal(L_k/\mathbb{Q})}(\sqrt[k_1]{a})\right| \left|S_{Gal(L_k/\mathbb{Q})}(\sqrt[k_1]{a})\right| \\ &=\; k_1 \left|S_{Gal(L_k/\mathbb{Q})}(\sqrt[k_1]{a})\right|\,, \end{aligned}$$

where the last equality follows since the element $\sqrt[k_1]{a}$ has exactly k_1 conjugates in L_k. Moreover, since Z_k is Galois over $\mathbb{Q}$, we have

$$n_k \;=\; |Gal(L_k/\mathbb{Q})| \;=\; |Gal(Z_k/\mathbb{Q})|\;|Gal(L_k/Z_k)|\;.$$

It follows that

$$\begin{aligned} n_k &= |Gal(Z_k/\mathbb{Q})|\;\left|\mathrm{O}_{Gal(L_k/Z_k)}(\sqrt[k_1]{a})\right|\;\left|S_{Gal(L_k/Z_k)}(\sqrt[k_1]{a})\right| \\ &= |Gal(Z_k/\mathbb{Q})|\;\left|\mathrm{O}_{Gal(L_k/Z_k)}(\sqrt[k_1]{a})\right|\;, \end{aligned}$$

since $\tau \in S_{Gal(L_k/Z_k)}(\sqrt[k_1]{a}) \implies \tau = 1_{L_k}$.

Thus,

$$|Gal(Z_k/\mathbb{Q})|\;\left|\mathrm{O}_{Gal(L_k/Z_k)}(\sqrt[k_1]{a})\right| \;=\; k_1\;\left|S_{Gal(L_k/\mathbb{Q})}(\sqrt[k_1]{a})\right|\;.$$

We now wish to show that $S_{Gal(L_k/\mathbb{Q})}(\sqrt[k_1]{a})$ is isomorphic to a subgroup $\mathcal{S}$ of $Gal(Z_k/\mathbb{Q})$. Consider the isomorphism given by

$$\Phi \;:\; S_{Gal(L_k/\mathbb{Q})}(\sqrt[k_1]{a}) \mapsto \mathcal{S} \leq Gal(Z_k/\mathbb{Q})$$

where

$$\Phi(\tau) = \tau|_{Z_k} \text{ is the restriction of } \tau \text{ to } Z_k \text{ for any } \tau \in S_{Gal(L_k/\mathbb{Q})}(\sqrt[k_1]{a})\,.$$

This implies that $\left|S_{Gal(L_k/\mathbb{Q})}(\sqrt[k_1]{a})\right|$ divides $|Gal(Z_k/\mathbb{Q})|$, thus we see that

$$\frac{|Gal(Z_k/\mathbb{Q})|}{\left|S_{Gal(L_k/\mathbb{Q})}(\sqrt[k_1]{a})\right|}\;\left|\mathrm{O}_{Gal(L_k/Z_k)}(\sqrt[k_1]{a})\right| \;=\; k_1 \;\implies\; \left|\mathrm{O}_{Gal(L_k/Z_k)}(\sqrt[k_1]{a})\right| \text{ divides } k_1\,.$$

Therefore $[L_k : Z_k]\,|\,k_1$ since $[L_k : Z_k] \;=\; \left|\mathrm{O}_{Gal(G/Z)}(\sqrt[k_1]{a})\right|$. This completes the proof. □

Appendix B

A Proof of the Large Sieve Inequality

Definition B.1. *Let χ be a character modulo k. We say that χ is primitive if there is no positive integer $m < k$ such that $m \mid k$ and $\chi(n) = \chi(n \bmod m)$.*

Definition B.2. *Let χ be a character modulo k. Then the Gauss sum $\tau(\chi)$ is defined by*

$$\tau(\chi) \;=\; \sum_{m=1}^{k} \chi(m) e\left(\frac{m}{k}\right) ,$$

where $e(t) = e^{2\pi i t}$.

Lemma B.1. *Let χ be a character modulo k. If $(n,k) = 1$, then*

$$\chi(n)\tau(\bar{\chi}) \;=\; \sum_{m=1}^{k} \bar{\chi}(m) e\left(\frac{mn}{k}\right) .$$

Proof. Letting $h \equiv mn^{-1} \pmod{k}$, which we can do since $(n,k) = 1$, we have

$$\chi(n)\tau(\bar{\chi}) \;=\; \sum_{m=1}^{k} \bar{\chi}(m)\chi(n) e\left(\frac{m}{k}\right) \;=\; \sum_{h=1}^{k} \bar{\chi}(h) e\left(\frac{nh}{k}\right) .$$

□

Lemma B.2. *If χ is a primitive, nonprincipal character modulo k and $(n,k) > 1$, then*

$$\chi(n)\tau(\bar{\chi}) \;=\; \sum_{m=1}^{k} \bar{\chi}(m) e\left(\frac{mn}{k}\right) .$$

Proof. Let us write

$$\frac{n}{k} \;=\; \frac{n_1}{k_1}\,,$$

where $(n_1, k_1) = 1$ and $k_1 \,|\, k$, $k_1 < k$. If n is a multiple of k, the left-hand side is zero, and so is the right-hand side, since

$$\sum_{m=1}^{k} \bar{\chi}(m) \;=\; 0\,.$$

So, we may assume that $1 < k_1 < k$. It now remains to show that

$$\sum_{m=1}^{k} \bar{\chi}(m) e\left(\frac{mn_1}{k_1}\right) \;=\; 0\,.$$

Write $k = k_1 k_2$ and put $m = ak_1 + b$, where $0 \le a < k_2$, $1 \le b \le k_1$. The above sum can then be rewritten as

$$\sum_{1 \le b \le k_1} e\left(\frac{bn_1}{k_1}\right) \sum_{0 \le a < k_2} \bar{\chi}(ak_1 + b)\,.$$

It therefore suffices to prove that the inner sum is zero. Let us write

$$S(b) \;=\; \sum_{0 \le a < k_2} \bar{\chi}(ak_1 + b)\,.$$

A straightforward reordering argument shows that $S(b + k_1) = S(b)$. Moreover, if c is any integer satisfying

$$(c, k) = 1 \;\text{ and }\; c \equiv 1 \,(\mathrm{mod}\, k_1)\,,$$

then, since $S(b + k_1) = S(b)$,

$$\chi(c) S(b) \;=\; \sum_{0 \le a < k_2} \bar{\chi}(cak_1 + cb) \;=\; \sum_{0 \le a < k_2} \bar{\chi}(ak_1 + cb)\,.$$

Now, dividing bc by k_1 yields $bc = qk_1 + r$ where $0 \le r < k_1$. Looking at this equation modulo k_1, we see that $r = 0$ and $b = k_1$, or $r = b$, since $c \equiv 1 \,(\mathrm{mod}\, k_1)$ and $1 \le b \le k_1$. This shows that

$$\sum_{0 \le a < k_2} \bar{\chi}(ak_1 + cb) \;=\; \sum_{0 \le a < k_2} \bar{\chi}(ak_1 + b) \;=\; S(b)\,.$$

Therefore,

$$\chi(c)S(b) \;=\; S(b)\,.$$

Since χ is a primitive character modulo k, there are integers c_1 and c_2 such that

$$(c_1,k) = (c_2,k) = 1 \;\text{ and }\; c_1 \equiv c_2 \,(\mathrm{mod}\, k_1)\,,$$

where $\chi(c_1) \neq \chi(c_2)$. Hence, there exists $c \equiv c_1 c_2^{-1} \,(\mathrm{mod}\, k_1)$ such that $(c,k) = 1$ and $\chi(c) \neq 1$. This in turns imply that $S(b) = 0$, which completes the proof. □

Theorem B.1. *If χ is a primitive character modulo k, then $|\tau(\chi)| = \sqrt{k}$.*

Proof. Take any integer n such that $(n,k) = 1$. Then, from Lemma B.2, we have that

$$\begin{aligned}
|\tau(\chi)| \;&=\; |\bar{\chi}(-n)\tau(\chi)| \;=\; \left|\sum_{m=1}^{k} \chi(m) e\left(\frac{-mn}{k}\right)\right| \\
&=\; \left|\sum_{m=1}^{k} \bar{\chi}(m) e\left(\frac{mn}{k}\right)\right| \\
&=\; |\chi(n)\tau(\bar{\chi})| \\
&=\; |\tau(\bar{\chi})|\,,
\end{aligned}$$

hence $|\tau(\bar{\chi})| \;=\; |\tau(\chi)|$. From this equality and Lemma B.2, we see that

$$|\chi(n)|^2 \,|\tau(\chi)|^2 \;=\; \sum_{m_1}^{k}\sum_{m_2}^{k} \bar{\chi}(m_1)\chi(m_2) e\left(\frac{n(m_1-m_2)}{k}\right)\,.$$

Summing over n for $1 \leq n \leq k$, we obtain from the left-hand side

$$\phi(k)\,|\tau(\chi)|^2\,.$$

The right-hand side yields

$$\begin{aligned}
&\sum_{n=1}^{k} \phi(k) \;+\; \sum_{n=1}^{k} \sum_{\substack{1\leq m_1,m_2\leq k \\ m_1\neq m_2}} \bar{\chi}(m_1)\chi(m_2) e\left(\frac{n(m_1-m_2)}{k}\right) \\
=\;& k\phi(k) \;+\; \sum_{\substack{1\leq m_1,m_2\leq k \\ m_1\neq m_2}} \bar{\chi}(m_1)\chi(m_2) \sum_{n=1}^{k} e\left(\frac{n(m_1-m_2)}{k}\right)\,,
\end{aligned}$$

but

$$\sum_{n=1}^{k} e\left(\frac{n(m_1 - m_2)}{k}\right) \;=\; 0\,,$$

whenever $m_1 \neq m_2$. Therefore, we have that

$$\phi(k)\,|\tau(\chi)|^2 \;=\; k\phi(k)\,,$$

which implies that $|\tau(\chi)| = \sqrt{k}$, as required. □

Theorem B.2. *(The Large Sieve Inequality) For each character χ modulo k, we let*

$$S(\chi) \;=\; \sum_{n=M+1}^{M+N} a_n \chi(n)\,,$$

where a_n is any complex number, $M \in \mathbb{Z}$ and $N \in \mathbb{Z}^+$. Then

$$\sum_{k \leq K} \sum_{\chi \bmod k}{}' |S(\chi)|^2 \;\ll\; (K^2 + N) \sum_{n=M+1}^{M+N} |a_n|^2\,,$$

where $\sum'$ denotes summation over primitive characters, $K \in \mathbb{Z}^+$ and $k \geq 1$.

Proof. We first let

$$S(t) = \sum_{n=M+1}^{M+N} a_n e(nt) \quad \text{and} \quad Z = \sum_{n=M+1}^{M+N} |a_n|^2\,.$$

To prove the theorem, we proceed in two steps.

Step 1.

From Lemma B.1 and lemma B.2, we know that for each integer n,

$$\chi(n)\tau(\bar{\chi}) \;=\; \sum_{m=1}^{k} \bar{\chi}(m) e\left(\frac{mn}{k}\right)\,,$$

for any primitive character χ modulo k. Multiplying this by a_n and summing over n from $M+1$ to $M+N$, we get

$$\tau(\bar{\chi}) S(\chi) \;=\; \sum_{m=1}^{k} \bar{\chi}(m) S\left(\frac{m}{k}\right)$$

for any primitive character χ modulo k.

Now, using the fact that $|\tau(\chi)| = \sqrt{k}$ for any primitive character χ modulo k, we see that

$$k \sum_{\chi \bmod k}{}' |S(\chi)|^2 \;\le\; \sum_{\chi \bmod k} \left| \sum_{m=1}^{k} \bar{\chi}(m) S\left(\frac{m}{k}\right) \right|^2 \;=\; \phi(k) \sum_{\substack{m=1 \\ (m,k)=1}}^{k} \left| S\left(\frac{m}{k}\right) \right|^2 .$$

We thus have

$$\frac{1}{\phi(k)} \sum_{\chi \bmod k}{}' |S(\chi)|^2 \;\le\; \frac{1}{k} \sum_{\substack{m=1 \\ (m,k)=1}}^{k} \left| S\left(\frac{m}{k}\right) \right|^2 . \tag{B.1}$$

Step 2.

Let $F : \mathbb{R} \mapsto \mathbb{C}$ be any complex-valued function with continuous derivative and period 1. Then

$$\int_{m/k}^{\alpha} dF(\beta) \;=\; F(\alpha) - F(m/k)$$

and

$$\begin{aligned} |F(\alpha)| &= |F(\alpha) - F(m/k) + F(m/k)| \\ &= |F(m/k) - (F(m/k) - F(\alpha))| \\ &\ge |F(m/k)| - |F(m/k) - F(\alpha)| \, . \end{aligned}$$

Thus

$$\begin{aligned} |F(m/k)| &\le |F(\alpha)| + |F(\alpha) - F(m/k)| \\ &= |F(\alpha)| + \left| \int_{m/k}^{\alpha} dF(\beta) \right| \\ &\le |F(\alpha)| + \int_{m/k}^{\alpha} |F'(\beta)| \, d\beta \, . \end{aligned}$$

Averaging the above inequality over the interval $I(m/k)$ of length $1/K^2$ centered at m/k yields that

$$\left| F\left(\frac{m}{k}\right) \right| \;\le\; K^2 \int_{I(m/k)} |F(\alpha)| d\alpha \;+\; \frac{1}{2} \int_{I(m/k)} |F'(\beta)| d\beta \, .$$

Observe that the intervals $I(m/k)$ with $1 \le m \le k$, $(m,k) = 1$, and $k \le K$ do not

overlap, modulo 1.

Hence

$$\sum_{k\le K}\sum_{\substack{m=1\\(m,k)=1}}^{k}\left|F\left(\frac{m}{k}\right)\right| \;\le\; K^2\int_0^1|F(\alpha)|d\alpha \;+\; \frac{1}{2}\int_0^1|F'(\beta)|d\beta\,. \tag{B.2}$$

Now let $F = S^2$. Then the first integral on the right of (B.2) is Z. Applying Hölder's inequality to the second integral on the right of (B.2) yields

$$\begin{aligned}\frac{1}{2}\int_0^1|F'(\beta)|d\beta &= \int_0^1|S(\beta)S'(\beta)|d\beta\\ &\le \left(\int_0^1|S(\beta)|^2d\beta\right)^{\frac{1}{2}}\left(\int_0^1|S'(\beta)|^2d\beta\right)^{\frac{1}{2}}.\end{aligned}$$

The first integral on the right is again equal to Z. Before estimating the second integral, we may first multiply $S(\alpha)$ by $e(-\hat{m}\alpha)$ for a suitalbe choice of $\hat{m}\in\mathbb{Z}$ so that the range of n becomes $|n|\le\frac{1}{2}N$. This leaves $|S(\alpha)|$ and Z unchanged and is therefore legitimate. The second integral can then be evaluated to be

$$\sum_{|n|\le\frac{1}{2}N}|2\pi i n a_n|^2 \;\le\; (\pi N)^2 Z\,.$$

Finally, combining the above implies that

$$\sum_{k\le K}\sum_{\substack{m=1\\(m,k)=1}}^{k}\left|S\left(\frac{m}{k}\right)\right|^2 \;\le\; (K^2+\pi N)Z\,. \tag{B.3}$$

Combining (B.1) and (B.3), we see that

$$\begin{aligned}\sum_{k\le K}{\sum_{\chi \bmod k}}' |S(\chi)|^2 &\le \sum_{k\le K}\sum_{\substack{m=1\\(m,k)=1}}^{k}\left|S\left(\frac{m}{k}\right)\right|^2\\ &\le (K^2+\pi N)Z\\ &\ll (K^2+N)Z\,.\end{aligned}$$

This completes the proof.

□

List of References

[1] T. M. Apostol. *Introduction to Analytic Number Theory.* Springer-Verlag, New York Inc., 1976. 79

[2] E. Artin. *Collected Papers.* Reading, MA: Addison-Wesley, 1965. 1

[3] M. Artin. *Algebra.* Prentice-Hall, Inc., 1991. 11, 12

[4] R. D. Carmichael. *The Theory of Numbers.* Wiley, New York, 1914. 42

[5] A. C. Cojocaru and M. R. Murty. *An Introduction to Sieve Methods and their Applications.* Cambridge University Press, 2005. 13, 17, 23, 45

[6] H. Davenport. *Multiplicative Number Theory, second ed.* Springer-Verlag, New York., 1980. 32

[7] P. Erdos and C. Pomerance. *On the Normal Number of Prime Factors of $\phi(n)$. Rocky Mountain Math. J.*, 15:343–352, 1985. 76

[8] P. X. Gallagher. *The Large Sieve.* 14:14–20, 1967.

[9] H. Halberstan and H. E. Richert. *Sieve Methods.* Academic Press, New York., 1974. 45, 46, 72

[10] C. Hooley. *On Artin's conjecture.* 225:209–220, 1967. 2

[11] E. Landau. *Uber ideale und primideale in idealklassen. Math. Zeito*, 2:52–154, 1918. 15

[12] S. Li. *On the Number of Elements with Maximal Order in the Multiplicative Group Modulo n. Acta Arithmetica*, 86:113–132, 1998. 69, 76

[13] S. Li. *On Extending Artin's Conjecture to Composite Moduli. Mathematika*, 46:373–390, 1999. 79

[14] S. Li. *Artin's Conjecture on Average for Composite Moduli. Journal of Number Theory*, 84:93–118, 2000. 2, 42, 46, 72

[15] S. Li. *An Improvement of Artin's Conjecture on Average for Composite Moduli. Mathematika*, 51:97–110, 2004. 78

[16] Yu. V. Linnik. *On the Least Prime in an Arithmetic Progression. I. The Basic Theorem. Mat. Sbornik N. S.*, 15 (57):139–178, 1944. 44

[17] D. A. Marcus. *Number Fields.* Springer-Verlag, New York Inc., 1977.

[18] H. L. Montgomery and R. C. Vaughan. *The Large Sieve. Mathematika*, 20:119–134, 1973. 48

[19] M. R. Murty. *Problems in Algebraic Number Theory.* Springer-Verlag, New York Inc., 2001. 7, 8, 9, 16

[20] M. R. Murty. *Problems in Analytic Number Theory.* Springer-Verlag, New York Inc., 2001. 31, 33, 43, 45

[21] J. Neukirch. *Algebraic Number Theory.* Springer-Verlag Berlin Heidelberg, 1999. 8

[22] G. Pomerance P. Erdos, A. Granville and C. Spiro. *On the Normal Behaviour of the Iterates of Some Arithmetic Functions. Progr. Math.*, 85:165–204, 1990. 76

[23] C. Pomerance and S. Li. *On Generalizing Artin's Conjecture on Primitive Roots to Composite Moduli. J. Reine Angew. Math.*, 556:205–224, 2003. 79

[24] K. Prachar. *Primzahlverteilung.* Berlin, 1957. 23

[25] I. J. Schoenberg. *On Asymptotic Distributions of Arithmetical Functions. Trans. Amer. Math. Soc.*, 39:315–330, 1936. 45

[26] P. J. Stephens. *An Average Result for Artin's Conjecture.* 16:178–188, 1969. 2, 78

[27] A. E. Western and J. C. P. Miller. *Tables of Indices and Primitive Roots. Royal Society Mathematical Tables*, 9:38, 1968. 2