

# AVERAGE RESULTS ON THE ORDER OF $a$ MODULO $p$

KIM, SUNGJIN  
DEPARTMENT OF MATHEMATICS  
UNIVERSITY OF CALIFORNIA, LOS ANGELES  
MATH SCIENCE BUILDING 6617A  
E-MAIL: 707107@GMAIL.COM

ABSTRACT. Let  $a > 1$  be an integer. Denote by  $l_a(p)$  the multiplicative order of  $a$  modulo primes  $p$ . We prove that if  $\frac{x}{\log x \log \log x} = o(y)$ , then

$$\frac{1}{y} \sum_{a \leq y} \sum_{p \leq x} \frac{1}{l_a(p)} = \log x + C \log \log x + O(1) + O\left(\frac{x}{y \log \log x}\right),$$

which is an improvement over a theorem by Felix [Fe]. Additionally, we also prove two other related average results.

## 1. INTRODUCTION

Throughout this paper, we use the letter  $p$  to denote prime numbers. Let  $a > 1$  be an integer. If  $p$  does not divide  $a$ , we denote the multiplicative order of  $a$  modulo  $p$  by  $l_a(p)$ . Artin's Conjecture on Primitive Roots (AC) states that  $l_a(p) = p - 1$  for a positive proportion of primes  $p$  where the proportion is a rational multiple of the Artin's constant  $A = \prod_p \left(1 - \frac{1}{p(p-1)}\right)$ . Assuming the Generalized Riemann Hypothesis (GRH), Hooley [Ho] proved that  $l_a(p) = p - 1$  for positive proportion of primes  $p \leq x$ . It is expected that  $l_a(p)$  is large for majority of primes  $p \leq x$ . In [EM], Erdos and Murty showed that  $l_a(p) \geq p^{1/2+\epsilon(p)}$  for all but  $o(\pi(x))$  primes  $p \leq x$  where  $\epsilon(p) \rightarrow 0$ . With much simpler method, they showed a weaker result  $l_a(p) > \frac{\sqrt{p}}{\log p}$  for all but  $O(x/\log^3 x)$  primes  $p \leq x$ . Pappalardi [P] showed that there exist  $\alpha, \delta > 0$  such that  $l_a(p) \geq p^{1/2} \exp(\log^\delta p)$  for all but  $O(x/\log^{1+\alpha} x)$ . Kurlberg and Pomerance [KP2] applied Fouvry [Fo] to show that there is  $\gamma > 0$  such that  $l_a(p) > p^{1/2+\gamma}$  for positive proportion of primes  $p \leq x$ .

Therefore, it is natural to expect that the average reciprocal of  $l_a(p)$  is quite small. Murty and Srinivasan [MS] showed that  $\sum_{p < x} \frac{1}{l_a(p)} = O(\sqrt{x})$  and that  $\sum_{p < x} \frac{1}{l_a(p)} = O(x^{1/4})$  implies AC for  $a$ . Pappalardi [P] proved that for some positive constant  $\gamma$ ,

$$\sum_{p < x} \frac{1}{l_a(p)} = O\left(\frac{\sqrt{x}}{\log^{1+\gamma} x}\right).$$

For fixed  $a$ , it seems that it is very difficult to reduce  $\sqrt{x}$  with current knowledge. However, we expect that averaging over  $a$  would give some information. The following result by Felix [Fe] supports that  $l_a(p)$  is mostly large:

If  $\frac{x}{\log x} = o(y)$ , then

$$\frac{1}{y} \sum_{a \leq y} \sum_{p \leq x} \frac{1}{l_a(p)} = \log x + O(\log \log x) + O\left(\frac{x}{y}\right).$$

Felix remarked that the first error term  $O(\log \log x)$  can be  $C \log \log x + O(1)$  by applying Fiorilli's method [Fi], but did not explicitly find  $C$ . We find the  $C$  in Theorem 1.1. This detailed estimate takes effect when  $\frac{x}{(\log \log x)^2} = o(y)$ . We apply a deep result on exponential sums by Bourgain [B] to obtain Corollary 2.2 which will be the key for all average results in this paper.

**Theorem 1.1.** If  $\frac{x}{\log x \log \log x} = o(y)$ , then

$$\frac{1}{y} \sum_{a \leq y} \sum_{p \leq x} \frac{1}{l_a(p)} = \log x + C \log \log x + O(1) + O\left(\frac{x}{y \log \log x}\right),$$

where

$$C = 2\gamma - 2 \sum_p \frac{\log p}{p^2 - p + 1} + \frac{\zeta(2)\zeta(3)}{\zeta(6)} \sum_{k=1}^{\infty} \frac{\mu(k)}{k^2} \left( -2 \sum_{p|k} \frac{(p-1)p \log p}{p^2 - p + 1} + \log k \right) \prod_{p|k} \left( 1 + \frac{p-1}{p^2 - p + 1} \right).$$

Assuming GRH for Kummer extensions  $\mathbb{Q}(\zeta_d, a^{1/d})$ , Pappalardi [P, Theorem 4.1] proved that for increasing function  $\psi(x)$  tending to infinity,  $l_a(p) \geq \frac{p}{\psi(p)}$  for all but  $O\left(\frac{\pi(x) \log \psi(x)}{\psi(\sqrt{x})}\right)$  primes  $p \leq x$ . We prove an unconditional average version of Pappalardi's result proven within  $\log^2 x = o(\psi(x))$ .

**Theorem 1.2.** Let  $\psi(x)$  be an increasing function such that  $\log^2 x = o(\psi(x))$ . Let  $\delta$  be the positive constant in Corollary 2.8. If  $x^{1-\delta} \log^3 x = o(y)$ , then

$$\frac{1}{y} \sum_{a < y} \sum_{\substack{p < x \\ l_a(p) > \frac{x}{\psi(x)}}} 1 = \pi(x) + O\left(\frac{x \log x}{\psi(x)}\right) + O\left(\frac{x^{2-\delta} \log^2 x}{y}\right).$$

Assuming the GRH for Kummer extensions  $\mathbb{Q}(\zeta_d, a^{1/d})$ , Kurlberg and Pomerance [KP] showed that

$$\frac{1}{\pi(x)} \sum_{p < x} l_a(p) = c_a x + O\left(\frac{x}{(\log x)^{1-4/\log \log \log x}}\right)$$

where  $c_a$  is a rational multiple of  $c = \prod_p \left(1 - \frac{p}{p^3-1}\right)$ . In fact, they proved this for rational  $a = c/d$  with  $(c, d) = 1$ ,  $|c| \leq x$ , and  $|d| \leq x$ . For integral  $a$ , Stephens [S, Theorem 1] proved assuming the GRH that

$$\frac{1}{\pi(x)} \sum_{p \leq x} \frac{l_a(p)}{p-1} = c_a \frac{x}{\log x} + O\left(\frac{x \log \log x}{\log^2 x}\right).$$

An average result over all possible nonzero residue classes is obtained by Luca [L]: For any constant  $A > 0$ ,

$$\frac{1}{\pi(x)} \sum_{p < x} \frac{1}{(p-1)^2} \sum_{a=1}^{p-1} l_a(p) = c + O\left(\frac{1}{\log^A x}\right).$$

By partial summation, this gives the following statistics on average order:

$$\frac{1}{\pi(x)} \sum_{p < x} \frac{1}{p-1} \sum_{a=1}^{p-1} l_a(p) = \frac{1}{2} c x + O\left(\frac{x}{\log x}\right).$$

We prove that unconditionally on average, a similar result holds with average order  $cp$ .

**Theorem 1.3.** Let  $A > 0$  be any constant, and  $\delta > 0$  be the constant in Corollary 2.8. If  $x^{1-\delta} \log^3 x = o(y)$ , then

$$\frac{1}{y} \sum_{a < y} \sum_{\substack{p < x \\ p \nmid a}} l_a(p) = c Li(x^2) + O\left(\frac{x^2}{\log^A x}\right) + O\left(\frac{x^{3-\delta} \log^2 x}{y}\right)$$

where

$$c = \prod_p \left(1 - \frac{p}{p^3-1}\right).$$

An immediate corollary of Theorem 1.3 is the following average version of Kurlberg and Pomerance:

**Corollary 1.4.** *Let  $\delta > 0$  be the constant in Corollary 2.4. If  $x^{1-\delta} \log^3 x = o(y)$ , then*

$$\frac{1}{y} \sum_{a < y} \frac{1}{\pi(x)} \sum_{\substack{p < x \\ p \nmid a}} l_a(p) = \frac{1}{2} cx + O\left(\frac{x}{\log x}\right) + O\left(\frac{x^{2-\delta} \log^3 x}{y}\right).$$

Theorem 1.3 and Corollary 1.4 are indeed weaker than what we can achieve from [S, Theorem 1]. We state and prove these to exhibit the use of Erdős-Turán inequality and Bourgain's exponential sum result.

## 2. BACKGROUND

**2.1. Equidistribution.** A sequence  $\{a_n\}$  of real numbers with  $a_n \in [0, 1]$  are said to be *equidistributed modulo 1* if the following is satisfied:

**Definition 2.1.** *Let  $0 \leq a < b \leq 1$ . Suppose that*

$$\lim_{N \rightarrow \infty} \frac{1}{N} \#\{n \leq N : a_n \in (a, b)\} = b - a.$$

*Then we say that  $\{a_n\}$  is equidistributed modulo 1.*

We have the following well-known criterion by Weyl [W]:

**Theorem 2.2.** *For any integer  $k \neq 0$ , suppose that*

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n \leq N} e^{2\pi i k a_n} = 0.$$

*Then the sequence  $\{a_n\}$  is equidistributed modulo 1.*

There was an effort to obtain a quantitative form of the equidistribution theorem. Erdős and Turán [ET] succeed in obtaining such form:

**Theorem 2.3** (Erdős-Turán Inequality). *Let  $\{a_n\}$  be a sequence of real numbers in  $[0, 1]$ . Then*

$$\sup_{0 \leq a < b \leq 1} \left| \#\{n \leq N : a_n \in (a, b)\} - (b - a)N \right| \leq c_1 \frac{N}{M + 1} + c_2 \sum_{m=1}^M \frac{1}{m} \left| \sum_{n \leq N} e^{2\pi i m a_n} \right|.$$

Montgomery [M] obtained  $c_1 = 1$ ,  $c_2 = 3$ . Mauduit, Rivat, Sárközy [MRS] obtained  $c_1 = c_2 = 1$ . Thus, we have a quantitative upper bound of discrepancy when we have good upper bounds for exponential sums.

**2.2. Exponential Sums in Prime Fields.** Bourgain [B] obtained the following equidistribution result for the subgroup  $H < \mathbb{F}_p^*$  when  $|H| > p^{\frac{C}{\log \log p}}$  for some absolute constant  $C > 1$  by sum-product method. See also [BG].

**Theorem 2.4.** *Let  $p$  be a prime. There exist absolute constants  $C > 1$  and  $C_1 > 0$  such that for any subgroup  $H$  of  $\mathbb{F}_p^*$  with  $|H| > p^{\frac{C}{\log \log p}}$ ,*

$$\max_{(k,p)=1} \left| \sum_{a \in H} e^{2\pi i k \frac{a}{p}} \right| < e^{-\log^{C_1} p |H|}.$$

Since any subgroup  $H$  of  $\mathbb{F}_p^*$  is cyclic, we consider  $|H| = d|p - 1|$ . Then  $H$  consists of all  $d$ -th roots of unity in  $\mathbb{F}_p$ . This yields

**Corollary 2.5.** *Let  $1 \leq d|p - 1$ . Suppose that  $d > p^{\frac{C}{\log \log p}}$ . Then we have*

$$\max_{(m,p)=1} \left| \sum_{a \in \mathbb{F}_p, a^d=1} e^{2\pi i m \frac{a}{p}} \right| < d e^{-\log^{C_1} p}.$$

Combining this with Erdős-Turán inequality, we obtain the following

**Corollary 2.6.** *Let  $y \geq 1$ . Assume that  $d|p-1$  and  $d > p^{\frac{C}{\log \log p}}$ . Then for any constant  $C_2 \in (0, C_1)$ , where  $C_1$  is given in Corollary 2.5, we have*

$$\sum_{a < y, a^d \equiv 1(p)} 1 = \frac{y}{p}d + O(de^{-\log^{C_2} p}).$$

*Proof.* Since  $d|p-1$ , the congruence  $a^d \equiv 1$  yields  $d$  roots in  $\mathbb{F}_p$ . Thus, we need to count  $a < y$  satisfying those  $d$  congruences modulo  $p$ . Considering  $\frac{y}{p} = \lfloor \frac{y}{p} \rfloor + \frac{y}{p} - \lfloor \frac{y}{p} \rfloor$ , it is enough to prove the result for  $y < p$ . We apply Erdős-Turán inequality to the set  $\{\frac{a}{p} : 0 \leq a \leq p-1, a^d \equiv 1(p)\} = \{a_1, \dots, a_d\}$ . Then

$$\begin{aligned} \left| \#\{1 \leq i \leq d : a_i \in (0, \frac{y}{p}) \bmod 1\} - \frac{y}{p}d \right| &\leq \frac{d}{p} + \sum_{m=1}^{p-1} \frac{1}{m} \left| \sum_{a \leq p-1, a^d \equiv 1(p)} e^{2\pi i m \frac{a}{p}} \right| \\ &\leq \frac{d}{p} + (2 \log p)de^{-\log^{C_1} p} \\ &\leq de^{-\log^{C_2} p}. \end{aligned}$$

This completes the proof. □

For the Theorem 1.2 and 1.3, we use a weaker form of Theorem 2.4.

**Theorem 2.7.** *For any fixed  $\epsilon > 0$ , There exist a constant  $\delta = \delta(\epsilon) > 0$  such that for any subgroup  $H$  of  $\mathbb{F}_p^*$  with  $|H| > p^\epsilon$ ,*

$$\max_{(k,p)=1} \left| \sum_{a \in H} e^{2\pi i k \frac{a}{p}} \right| < p^{-\delta} |H|.$$

Similarly, we have the following corollary:

**Corollary 2.8.** *Let  $1 \leq d|p-1$ . Let  $\epsilon > 0$  be fixed. Suppose that  $d > p^\epsilon$ . Then there exists a constant  $\delta = \delta(\epsilon) > 0$  such that*

$$\max_{(m,p)=1} \left| \sum_{a \in \mathbb{F}_p, a^d=1} e^{2\pi i m \frac{a}{p}} \right| < dp^{-\delta}.$$

We omit the proof of the following corollary because it is similar to that of Corollary 2.6.

**Corollary 2.9.** *Let  $y \geq 1$ . Let  $\epsilon > 0$  be fixed. Assume that  $d|p-1$  and  $d > p^\epsilon$ . Then there exists  $\delta = \delta(\epsilon) > 0$  such that*

$$\sum_{a < y, a^d \equiv 1(p)} 1 = \frac{y}{p}d + O(dp^{-\delta}).$$

Corollary 2.6 and 2.8 play key roles in proving Theorem 1.1, 1.2, and 1.3. Note that this is significantly better than the trivial bound when  $p$  is large:

$$\sum_{a < y, a^d \equiv 1(p)} 1 = \frac{y}{p}d + O(d).$$

### 3. PROOF OF THEOREMS

**3.1. Proof of Theorem 1.1.** Let  $\epsilon = \frac{4C}{\log \log x}$  and consider the summation change:

$$\begin{aligned} \sum_{a \leq y} \sum_{p \leq x} \frac{1}{l_a(p)} &= \sum_{d < x} \frac{1}{d} \sum_{\substack{p \leq x \\ p \equiv 1(d)}} \sum_{\substack{a \leq y \\ l_a(p)=d}} 1 \\ &= \sum_{d < x^\epsilon} + \sum_{x^\epsilon \leq d < x} \\ &= \Sigma_1 + \Sigma_2. \end{aligned}$$

First, we treat  $\Sigma_1$  by trivial bound and Brun-Titchmarsh inequality:

$$\begin{aligned}
\Sigma_1 &= \sum_{d < x^\epsilon} \frac{1}{d} \sum_{\substack{p \leq x \\ p \equiv 1(d)}} \sum_{\substack{a \leq y \\ l_a(p) = d}} 1 \\
&= \sum_{d < x^\epsilon} \frac{1}{d} \sum_{\substack{p \leq x \\ p \equiv 1(d)}} \left( \phi(d) \frac{y}{p} + O(\phi(d)) \right) \\
&= \sum_{d < x^\epsilon} \frac{1}{d} \sum_{\substack{p \leq x \\ p \equiv 1(d)}} \phi(d) \frac{y}{p} + O(E_1),
\end{aligned}$$

where

$$\begin{aligned}
E_1 &= \sum_{d < x^\epsilon} \frac{1}{d} \sum_{\substack{p \leq x \\ p \equiv 1(d)}} \phi(d) = \sum_{d < x^\epsilon} \frac{\phi(d)}{d} \sum_{\substack{p \leq x \\ p \equiv 1(d)}} 1 \\
&= \sum_{d < x^\epsilon} \frac{\phi(d)}{d} \pi(x; d, 1) \\
&\ll \sum_{d < x^\epsilon} \frac{\phi(d)}{d} \frac{x}{\phi(d) \log x} \ll \epsilon x.
\end{aligned}$$

Thus,

$$\Sigma_1 = \sum_{d < x^\epsilon} \frac{1}{d} \sum_{\substack{p \leq x \\ p \equiv 1(d)}} \phi(d) \frac{y}{p} + O(\epsilon x).$$

Now, we treat  $\Sigma_2$  by Möbius inversion and Corollary 2.6:

$$\begin{aligned}
\Sigma_2 &= \sum_{x^\epsilon \leq d < x} \frac{1}{d} \sum_{\substack{p \leq x \\ p \equiv 1(d)}} \sum_{\substack{a \leq y \\ l_a(p) = d}} 1 \\
&= \sum_{x^\epsilon \leq d < x} \frac{1}{d} \sum_{\substack{p \leq x \\ p \equiv 1(d)}} \sum_{d' | d} \mu\left(\frac{d}{d'}\right) \sum_{\substack{a \leq y \\ a^{d'} \equiv 1(p)}} 1 \\
&= \sum_{x^\epsilon \leq d < x} \frac{1}{d} \sum_{\substack{p \leq x \\ p \equiv 1(d)}} \sum_{\substack{d' | d \\ d' < p^{\frac{C}{\log \log p}}}} \mu\left(\frac{d}{d'}\right) \sum_{\substack{a \leq y \\ a^{d'} \equiv 1(p)}} 1 + \sum_{x^\epsilon \leq d < x} \frac{1}{d} \sum_{\substack{p \leq x \\ p \equiv 1(d)}} \sum_{\substack{d' | d \\ d' \geq p^{\frac{C}{\log \log p}}}} \mu\left(\frac{d}{d'}\right) \sum_{\substack{a \leq y \\ a^{d'} \equiv 1(p)}} 1 \\
&= \sum_{x^\epsilon \leq d < x} \frac{1}{d} \sum_{\substack{p \leq x \\ p \equiv 1(d)}} \sum_{\substack{d' | d \\ d' < p^{\frac{C}{\log \log p}}}} \mu\left(\frac{d}{d'}\right) \left( \frac{y}{p} d' + O(d') \right) \\
&\quad + \sum_{x^\epsilon \leq d < x} \frac{1}{d} \sum_{\substack{p \leq x \\ p \equiv 1(d)}} \sum_{\substack{d' | d \\ d' \geq p^{\frac{C}{\log \log p}}}} \mu\left(\frac{d}{d'}\right) \left( \frac{y}{p} d' + O(d' e^{-\log^{C_2} p}) \right).
\end{aligned}$$

Then we have

$$\begin{aligned}
& \Sigma_1 + \Sigma_2 \\
&= \sum_{d < x^\epsilon} \frac{1}{d} \sum_{\substack{p \leq x \\ p \equiv 1(d)}} \phi(d) \frac{y}{p} + \sum_{x^\epsilon \leq d < x} \frac{1}{d} \sum_{\substack{p \leq x \\ p \equiv 1(d)}} \sum_{\substack{d' | d \\ d' < p^{\frac{C}{\log \log p}}}} \mu\left(\frac{d}{d'}\right) \frac{y}{p} d' + \sum_{x^\epsilon \leq d < x} \frac{1}{d} \sum_{\substack{p \leq x \\ p \equiv 1(d)}} \sum_{\substack{d' | d \\ d' \geq p^{\frac{C}{\log \log p}}}} \mu\left(\frac{d}{d'}\right) \frac{y}{p} d' \\
&+ O(E_1) + O(E_2) + O(E_3) \\
&= \sum_{d < x} \frac{\phi(d)}{d} \sum_{\substack{p \leq x \\ p \equiv 1(d)}} \frac{y}{p} + O(E_1) + O(E_2) + O(E_3),
\end{aligned}$$

where

$$E_2 = \sum_{x^\epsilon \leq d < x} \frac{1}{d} \sum_{\substack{p \leq x \\ p \equiv 1(d)}} \sum_{\substack{d' | d \\ d' < p^{\frac{C}{\log \log p}}}} \left| \mu\left(\frac{d}{d'}\right) \right| d'$$

and

$$E_3 = \sum_{x^\epsilon \leq d < x} \frac{1}{d} \sum_{\substack{p \leq x \\ p \equiv 1(d)}} \sum_{\substack{d' | d \\ d' \geq p^{\frac{C}{\log \log p}}}} \left| \mu\left(\frac{d}{d'}\right) \right| d' e^{-\log^{C_2} p}.$$

Here, the term

$$\sum_{d < x} \frac{\phi(d)}{d} \sum_{\substack{p \leq x \\ p \equiv 1(d)}} \frac{y}{p}$$

is the main term in [Fe, Theorem 1.4]. It is proven to be  $y \log x + O(y \log \log x)$  in [Fe, Theorem 1.4] which will be shown to be  $y \log x + Cy \log \log x + O(1)$  later.

We treat  $E_2$ . Since  $\pi(x; d, 1) \ll \frac{x}{d}$ , we have:

$$\begin{aligned}
E_2 &= \sum_{x^\epsilon \leq d < x} \frac{1}{d} \sum_{\substack{p \leq x \\ p \equiv 1(d)}} \sum_{\substack{d' | d \\ d' < p^{\frac{C}{\log \log p}}}} \left| \mu\left(\frac{d}{d'}\right) \right| d' \\
&\ll \sum_{x^\epsilon \leq d < x} \frac{1}{d} \sum_{\substack{p \leq x \\ p \equiv 1(d)}} \sum_{d' < p^{\frac{C}{\log \log p}}} d' \\
&\ll \sum_{x^\epsilon \leq d < x} \frac{1}{d} \sum_{\substack{p \leq x \\ p \equiv 1(d)}} p^{\frac{2C}{\log \log p}} \\
&\ll x^{\frac{2C}{\log \log x}} \sum_{x^\epsilon \leq d < x} \frac{1}{d} \pi(x; d, 1) \\
&\ll x^{\frac{2C}{\log \log x}} \sum_{x^\epsilon \leq d < x} \frac{x}{d^2}.
\end{aligned}$$

Since  $\sum_{d \geq x} \frac{1}{d^2} \ll \frac{1}{x}$ , we have

$$E_2 \ll x^{1 + \frac{2C}{\log \log x} - \epsilon} \ll x^{1 - \frac{\epsilon}{2}}.$$

We are left with  $E_3$ . First, we have the following:

$$\begin{aligned}
\sum_{\substack{d'|d \\ d' \geq p^{\frac{C}{\log \log p}}}} \left| \mu \left( \frac{d}{d'} \right) \right| d' &\leq \sum_{d'|d} \left| \mu \left( \frac{d}{d'} \right) \right| d' \\
&\leq d \prod_{p|d} \left( 1 + \frac{1}{p} \right) \\
&= d \prod_{p|d} \left( 1 - \frac{1}{p} \right) \frac{1 + \frac{1}{p}}{1 - \frac{1}{p}} \\
&\leq \phi(d) \mathfrak{Z}^{\omega(d)},
\end{aligned}$$

where  $\omega(d)$  is the number of distinct prime factors of  $d$ .

Again by  $\pi(x; d, 1) \ll \frac{x}{d}$ , we have

$$\begin{aligned}
E_3 &\ll \sum_{x^\epsilon \leq d < x} \frac{1}{d} \sum_{\substack{p \leq x \\ p \equiv 1(d)}} \phi(d) \mathfrak{Z}^{\omega(d)} e^{-\log^{C_3} x} \\
&\ll \sum_{x^\epsilon \leq d < x} \frac{1}{d} \phi(d) \mathfrak{Z}^{\omega(d)} \frac{x}{d} e^{-\log^{C_3} x}.
\end{aligned}$$

By partial summation with  $\sum_{d \leq t} \mathfrak{Z}^{\omega(d)} \ll t \log^2 t$ ,

$$E_3 \ll \sum_{x^\epsilon \leq d < x} \frac{\mathfrak{Z}^{\omega(d)}}{d} x e^{-\log^{C_3} x} \ll x (\log^3 x) e^{-\log^{C_3} x} \ll x e^{-\log^{C_4} x}.$$

Combining these estimates, we have

$$\sum_{a \leq y} \sum_{p \leq x} \frac{1}{l_a(p)} = \sum_{d < x} \frac{\phi(d)}{d} \sum_{\substack{p \leq x \\ p \equiv 1(d)}} \frac{y}{p} + O(\epsilon x) + O(x^{1-\frac{\epsilon}{2}}) + O(x e^{-\log^{C_4} x}),$$

with the first error term dominating the other two. Hence,

$$\sum_{a \leq y} \sum_{p \leq x} \frac{1}{l_a(p)} = \sum_{d < x} \frac{\phi(d)}{d} \sum_{\substack{p \leq x \\ p \equiv 1(d)}} \frac{y}{p} + O\left(\frac{x}{\log \log x}\right).$$

Following the proof of [Fe, Theorem 1.4], we have

$$\begin{aligned}
\sum_{d < x} \frac{\phi(d)}{d} \pi(x; d, 1) &= \sum_{k < x} \frac{\mu(k)}{k} \sum_{\substack{p \leq x \\ p \equiv 1(k)}} \tau\left(\frac{p-1}{k}\right) \\
&= \sum_{k \leq \log^{A+2} x} + \sum_{\log^{A+2} x < k < x} \\
&= \sum_{k \leq \log^{A+2} x} + O\left(\frac{x}{\log^A x}\right).
\end{aligned}$$

As Fiorilli and Felix pointed out, we apply

$$\sum_{\substack{p \leq x \\ p \equiv 1(k)}} \tau\left(\frac{p-1}{k}\right) = \frac{x}{k} C_1(k) + \frac{1}{k} \left( 2C_2(k) + C_1(k) \log\left(\frac{(k')^2}{k}\right) \right) \text{li}(x) + O\left(\frac{x}{\log^A x}\right),$$

where

$$C_1(k) = \frac{\zeta(2)\zeta(3)}{\zeta(6)} \prod_{p|k} \left(1 + \frac{p-1}{p^2-p+1}\right),$$

$$C_2(k) = C_1(k) \left( \gamma - \sum_p \frac{\log p}{p^2-p+1} - \sum_{p|k} \frac{(p-1)p \log p}{p^2-p+1} \right),$$

and  $k' = \prod_{p|k} p$ .

As in [Fe, Theorem 1.4], all the sums over  $k$  are absolutely convergent and  $\sum \frac{\mu(k)C_1(k)}{k^2} = 1$ , so we have

$$\begin{aligned} \sum_{k \leq \log^{A+2} x} \frac{\mu(k)}{k} \sum_{\substack{p \leq x \\ p \equiv 1(k)}} \tau\left(\frac{p-1}{k}\right) &= \sum_{k \leq \log^{A+2} x} \frac{\mu(k)}{k^2} \left( x C_1(k) + \left( 2C_2(k) + C_1(k) \log\left(\frac{(k')^2}{k}\right) \right) \text{li}(x) \right) \\ &\quad + O\left(\frac{x}{\log^A x}\right) \\ &= x + \left( 2\gamma - 2 \sum_p \frac{\log p}{p^2-p+1} \right) \text{li}(x) \\ &\quad + \left( \sum_{k=1}^{\infty} \frac{\mu(k)C_1(k)}{k^2} \left( -2 \sum_{p|k} \frac{(p-1)p \log p}{p^2-p+1} + \log\left(\frac{(k')^2}{k}\right) \right) \right) \text{li}(x) \\ &\quad + O\left(\frac{x}{\log^A x}\right). \end{aligned}$$

Since  $\text{li}(u) = \frac{u}{\log u} + O\left(\frac{u}{\log^2 u}\right)$ , we finally obtain

$$\begin{aligned} \sum_{d < x} \frac{\phi(d)}{d} \sum_{\substack{p \leq x \\ p \equiv 1(d)}} \frac{1}{p} &= \int_2^x \frac{1}{u^2} \sum_{k \leq u} \frac{\phi(k)}{k} \pi(x; k, 1) du + O(1) \\ &= \int_2^x \frac{1}{u^2} \left( u + C \frac{u}{\log u} + O\left(\frac{u}{\log^2 u}\right) \right) du + O(1) \\ &= \log x + C \log \log x + O(1), \end{aligned}$$

where

$$\begin{aligned} C &= 2\gamma - 2 \sum_p \frac{\log p}{p^2-p+1} \\ &\quad + \sum_{k=1}^{\infty} \frac{\mu(k)C_1(k)}{k^2} \left( -2 \sum_{p|k} \frac{(p-1)p \log p}{p^2-p+1} + \log\left(\frac{(k')^2}{k}\right) \right). \end{aligned}$$

Since the terms in the second sum over  $k$  only appears when  $k$  is square free, we have  $k' = k$ . Thus,

$$\begin{aligned} C &= 2\gamma - 2 \sum_p \frac{\log p}{p^2-p+1} \\ &\quad + \sum_{k=1}^{\infty} \frac{\mu(k)C_1(k)}{k^2} \left( -2 \sum_{p|k} \frac{(p-1)p \log p}{p^2-p+1} + \log k \right). \end{aligned}$$

This completes the proof of Theorem 1.1.



**3.2. Proof of Theorem 1.2.** Throughout this and the next sections,  $\epsilon > 0$  is a fixed constant. Let  $\psi(x)$  be an increasing function which tends to infinity as  $x \rightarrow \infty$ . The rate of increase of  $\psi(x)$  is to be determined. We start with the change of order in summation. By Corollary 2.8,

$$\begin{aligned}
\sum_{a < y} \sum_{\substack{p < x \\ l_a(p) > \frac{x}{\psi(x)}}} 1 &= \sum_{\substack{\frac{x}{\psi(x)} < d < x \\ p \equiv 1(d)}} \sum_{\substack{p < x \\ p \equiv 1(d)}} \sum_{\substack{a < y \\ l_a(p) = d}} 1 \\
&= \sum_{\substack{\frac{x}{\psi(x)} < d < x \\ p \equiv 1(d)}} \sum_{\substack{p < x \\ p \equiv 1(d)}} \sum_{\substack{d' | d \\ d' < p^\epsilon}} \mu\left(\frac{d}{d'}\right) \sum_{\substack{a < y \\ a d' \equiv 1(p)}} 1 + \sum_{\substack{\frac{x}{\psi(x)} < d < x \\ p \equiv 1(d)}} \sum_{\substack{p < x \\ p \equiv 1(d)}} \sum_{\substack{d' | d \\ d' \geq p^\epsilon}} \mu\left(\frac{d}{d'}\right) \sum_{\substack{a < y \\ a d' \equiv 1(p)}} 1 \\
&= \sum_{\substack{\frac{x}{\psi(x)} < d < x \\ p \equiv 1(d)}} \sum_{\substack{p < x \\ p \equiv 1(d)}} \sum_{\substack{d' | d \\ d' < p^\epsilon}} \mu\left(\frac{d}{d'}\right) \left(\frac{y}{p} d' + O(d')\right) \\
&\quad + \sum_{\substack{\frac{x}{\psi(x)} < d < x \\ p \equiv 1(d)}} \sum_{\substack{p < x \\ p \equiv 1(d)}} \sum_{\substack{d' | d \\ d' \geq p^\epsilon}} \mu\left(\frac{d}{d'}\right) \left(\frac{y}{p} d' + O(d' p^{-\delta})\right) \\
&= \sum_{\substack{\frac{x}{\psi(x)} < d < x \\ p \equiv 1(d)}} \sum_{\substack{p < x \\ p \equiv 1(d)}} \frac{y}{p} \phi(d) + O(E_1) + O(E_2),
\end{aligned}$$

where

$$\begin{aligned}
E_1 &= \sum_{d < x} \sum_{\substack{p < x \\ p \equiv 1(d)}} \sum_{\substack{d' | d \\ d' < p^\epsilon}} \left| \mu\left(\frac{d}{d'}\right) \right| d' \\
&\ll \sum_{d < x} \sum_{\substack{p < x \\ p \equiv 1(d)}} \sum_{d' < p^\epsilon} d' \\
&\ll x^{2\epsilon} \sum_{d < x} \pi(x; d, 1) \\
&\ll x^{1+2\epsilon} \log x,
\end{aligned}$$

and

$$\begin{aligned}
E_2 &= \sum_{d < x} \sum_{\substack{p < x \\ p \equiv 1(d)}} \sum_{\substack{d' | d \\ d' \geq p^\epsilon}} \left| \mu\left(\frac{d}{d'}\right) \right| d' p^{-\delta} \\
&\ll \sum_{d < x} \sum_{\substack{p < x \\ p \equiv 1(d)}} \phi(d) 3^{\omega(d)} p^{-\delta} \\
&\ll \sum_{d < x} \phi(d) 3^{\omega(d)} \frac{x^{1-\delta}}{d} \\
&\ll x^{2-\delta} \log^2 x.
\end{aligned}$$

Now we treat the main term. Since we have  $\sum_{p \equiv 1(d)} \frac{1}{p} = \frac{\log \log x + O(\log d)}{\phi(d)}$  by [EP, Lemma 2.5],

$$\begin{aligned}
\sum_{\frac{x}{\psi(x)} < d < x} \phi(d) \sum_{\substack{p < x \\ p \equiv 1(d)}} \frac{1}{p} &= \sum_{d < x} \phi(d) \sum_{\substack{p < x \\ p \equiv 1(d)}} \frac{1}{p} - \sum_{d \leq \frac{x}{\psi(x)}} \phi(d) \sum_{\substack{p < x \\ p \equiv 1(d)}} \frac{1}{p} \\
&= \sum_{p < x} \frac{1}{p} \sum_{d|p-1} \phi(d) - \sum_{d \leq \frac{x}{\psi(x)}} \phi(d) \sum_{\substack{p < x \\ p \equiv 1(d)}} \frac{1}{p} \\
&= \sum_{p < x} \frac{p-1}{p} + O\left(\sum_{d \leq \frac{x}{\psi(x)}} \phi(d) \frac{\log \log x + \log d}{\phi(d)}\right) \\
&= \pi(x) + O(\log \log x) + O\left(\frac{x \log x}{\psi(x)}\right).
\end{aligned}$$

Combining all the estimates, we have

$$\sum_{a < y} \sum_{\substack{p < x \\ l_a(p) > \frac{x}{\psi(x)}}} 1 = y\pi(x) + O(y \log \log x) + O\left(\frac{xy \log x}{\psi(x)}\right) + O(x^{2-\delta} \log^2 x).$$

Since we have  $y < x$ , the error term  $O(y \log \log x)$  is dominated by  $O(x^{2-\delta} \log^2 x)$ . This completes the proof of Theorem 1.2.

**3.3. Proof of Theorem 1.3.** We begin with an application of Mobius inversion and Corollary 2.8:

$$\begin{aligned}
\sum_{a < y} \sum_{d < x} \sum_{\substack{p < x \\ l_a(p) = d}} d &= \sum_{d < x} d \sum_{\substack{p < x \\ p \equiv 1(d)}} \sum_{\substack{a < y \\ l_a(p) = d}} 1 \\
&= \sum_{d < x} d \sum_{\substack{p < x \\ p \equiv 1(d)}} \sum_{\substack{d' | d \\ d' < p^\epsilon}} \mu\left(\frac{d}{d'}\right) \sum_{\substack{a < y \\ a d' \equiv 1(p)}} 1 + \sum_{d < x} d \sum_{\substack{p < x \\ p \equiv 1(d)}} \sum_{\substack{d' | d \\ d' \geq p^\epsilon}} \mu\left(\frac{d}{d'}\right) \sum_{\substack{a < y \\ a d' \equiv 1(p)}} 1 \\
&= \sum_{d < x} d \sum_{\substack{p < x \\ p \equiv 1(d)}} \sum_{\substack{d' | d \\ d' < p^\epsilon}} \mu\left(\frac{d}{d'}\right) \left(\frac{y}{p} d' + O(d')\right) \\
&\quad + \sum_{d < x} d \sum_{\substack{p < x \\ p \equiv 1(d)}} \sum_{\substack{d' | d \\ d' \geq p^\epsilon}} \mu\left(\frac{d}{d'}\right) \left(\frac{y}{p} d' + O(d' p^{-\delta})\right) \\
&= \sum_{d < x} d \phi(d) \sum_{\substack{p < x \\ p \equiv 1(d)}} \frac{y}{p} + O(E_1) + O(E_2),
\end{aligned}$$

where

$$\begin{aligned}
E_1 &= \sum_{d < x} d \sum_{\substack{p < x \\ p \equiv 1(d)}} \sum_{\substack{d' | d \\ d' < p^\epsilon}} \left| \mu\left(\frac{d}{d'}\right) \right| d' \\
&\ll \sum_{d < x} d \sum_{\substack{p < x \\ p \equiv 1(d)}} \sum_{d' < p^\epsilon} d' \\
&\ll x^{2\epsilon} \sum_{d < x} d \pi(x; d, 1) \\
&\ll x^{2+2\epsilon},
\end{aligned}$$

and

$$\begin{aligned}
E_2 &= \sum_{d < x} d \sum_{\substack{p < x \\ p \equiv 1(d)}} \sum_{\substack{d'|d \\ d' \geq p^\epsilon}} \left| \mu \left( \frac{d}{d'} \right) \right| d' p^{-\delta} \\
&\ll \sum_{d < x} d \sum_{\substack{p < x \\ p \equiv 1(d)}} \phi(d) 3^{\omega(d)} p^{-\delta} \\
&\ll \sum_{d < x} d \phi(d) 3^{\omega(d)} \frac{x^{1-\delta}}{d} \\
&\ll x^{3-\delta} \log^2 x.
\end{aligned}$$

Now we treat the main term:

$$\begin{aligned}
\sum_{d < x} d \phi(d) \sum_{\substack{p < x \\ p \equiv 1(d)}} \frac{y}{p} &= y \sum_{p < x} \frac{1}{p} \sum_{d|p-1} d \phi(d) \\
&= y \sum_{p < x} \frac{(p-1)\alpha(p-1)}{p} \\
&= y \left( \sum_{p < x} \alpha(p-1) - \sum_{p < x} \frac{\alpha(p-1)}{p} \right).
\end{aligned}$$

Here,  $\alpha(n) = \frac{1}{n} \sum_{d|n} d \phi(d)$  is the average order of  $\mathbb{Z}/n\mathbb{Z}$ . We use the following theorem by Luca [L, Theorem 1]:

**Theorem 3.1.** *For any constant  $A > 0$ ,*

$$\frac{1}{\pi(x)} \sum_{p < x} \frac{\alpha(p-1)}{p-1} = c + O\left(\frac{1}{\log^A x}\right)$$

where

$$c = \prod_p \left( 1 - \frac{p}{p^3 - 1} \right).$$

Applying this theorem with partial summation, we obtain

$$\sum_{p < x} \alpha(p-1) - \sum_{p < x} \frac{\alpha(p-1)}{p} = c \operatorname{Li}(x^2) + O\left(\frac{x^2}{\log^A x}\right).$$

Therefore,

$$\sum_{a < y} \sum_{d < x} d \sum_{\substack{p < x \\ l_a(p)=d}} 1 = cy \operatorname{Li}(x^2) + O\left(\frac{yx^2}{\log^A x}\right) + O(x^{3-\delta} \log^2 x).$$

This completes the proof of Theorem 1.3.

For the proof of Corollary 1.1, we use  $\operatorname{Li}(x^2) = \frac{1}{2}x\pi(x) + O\left(\frac{x^2}{\log^2 x}\right)$ .

#### 4. REMARKS

The theorems in this paper have resemblance. If we change order of summation to put  $\sum_d$  first, Theorem 1.1 is essentially  $\sum_d d^{-1} \sum_p \sum_a$ . Theorem 1.2 is  $\sum_d d^0 \sum_p \sum_a$ , and Theorem 1.3 is  $\sum_d d^1 \sum_p \sum_a$ . There is a difference in the method of Theorem 1.1, and the other two. In Theorem 1.1, we split the sum into four parts, while we split into three parts in Theorem 1.2 and Theorem 1.3. This is because  $d^{-1}$  is large for small  $d$ 's. We do not have a better information than  $O(\epsilon x)$  for the error term  $O(E_1)$  in Theorem 1.1, unless we have better exponential sum results. However, the method presented in this paper has wide variety of applications. For various conditional results, we could obtain the corresponding unconditional

average results, and this method of exponential sums is powerful in shortening the range of averaging. In the upcoming paper, we will consider problems on the order of  $a$  modulo  $n$ , for general modulus  $n$ .

#### REFERENCES

- [B] J. Bourgain, *Multilinear Exponential Sums in Prime Fields Under Optimal Entropy Condition on the Sources*, Geometric and Functional Analysis, February 2009, Volume 18, Issue 5, pp 1477-1502
- [BG] J. Bourgain, A. A. Glibichuk, *Exponential Sum Estimates over a Subgroup in an Arbitrary Finite Field*, J. d'Analyse Math, 115 (2011), pp. 51-70
- [EM] P. Erdős, R. Murty, *On the Order of  $a \pmod p$* , CRM Proceedings and Lecture Notes, Volume 19, (1999) pp. 87-97.
- [EP] P. Erdős, C. Pomerance, *On the Normal Number of Prime Factors of  $\phi(n)$* , Rocky Mountain Journal of Mathematics, Volume 15, Number 2, Spring 1985.
- [ET] P. Erdős, P. Turán, *On a Problem in the Theory of Uniform Distribution I, II*, Nederll. Akad. Wetensch, 51 (1948), pp. 1146-1154, 1262-1269.
- [Fe] A. T. Felix, *Generalizing the Titchmarsh divisor problem*, Int. J. Number Theory 8 (2012), pp. 613-629
- [Fi] A. Fiorilli, *On a Theorem of Bombieri, Friedlander and Iwaniec*, 15 pages. Canad. J. Math. 64 (2012), 1019-1035.
- [Fo] E. Fouvry, *Théorème de Brun-Titchmarsh; application au théorème de Fermat*, Invent. Math., 79, (1985) pp. 383-407
- [Ho] C. Hooley, *On Artin's Conjecture*, Journal für die Reine und Angewandte Mathematik, Volume 225, (1967) pp. 209-220.
- [KR] P. Kurlberg, Z. Rudnick, *On Quantum Ergodicity for Linear Maps of the Torus*, Comm. Math. Phys., 222(1):201-227, 2001
- [KP] P. Kurlberg, C. Pomerance, *On a Problem of Arnold: the average multiplicative order of a given integer*, Algebra and Number Theory, 7 (2013), pp. 981-999.
- [KP2] P. Kurlberg, C. Pomerance, *On the periods of linear congruential and power generators*, Acta Arith. 119 (2005), pp. 149-169.
- [L] F. Luca, *Some Mean Values Related to Average Multiplicative Orders of Elements in Finite Fields*, Ramanujan Journal, March 2005, Volume 9, Issue 1-2, pp. 33-44
- [M] H. Montgomery, *Ten Lectures on the Interface Between Analytic Number Theory and Harmonic Analysis*, CBMS Regional Conference Series in Mathematics, Number 84, AMS.
- [MRS] C. Mauduit, J. Rivat, A. Sárközy, *On the Pseudo-Random Properties of  $n^c$* , Illinois Journal of Mathematics, Volume 46, Number 1, Spring 2002, pp. 185-197.
- [MS] M. R. Murty, S. Srinivasan, *Some remarks on Artin's conjecture*, Canad. Math. Bull. Vol. 30(1), 1987.
- [P] F. Pappalardi, *On the Order of Finitely Generated Subgroups of  $\mathbb{Q}^*$  (mod  $p$ ) and Divisors of  $p - 1$* , Journal of Number Theory 57 (1996), pp. 207-222
- [S] P. J. Stephens, *Prime Divisors of second-order linear recurrences. I*, Journal of Number Theory, Volume 8, August 1976, pp. 313-332.
- [W] H. Weyl, *Über ein Problem aus dem Gebiete der diophantischen*, Ges. Abh. I (Springer: Berlin 1968), 487-497. Approximationen