ARTIN TWIN PRIMES

MAGDALÉNA TINKOVÁ, EZRA WAXMAN, MIKULÁŠ ZINDULKA

ABSTRACT. We say that a prime number p is an *Artin prime* for g if g mod p generates the group $(\mathbb{Z}/p\mathbb{Z})^\times$. For appropriately chosen integers d and g, we present a conjecture for the asymptotic number $\pi_{d,g}(x)$ of primes $p \leq x$ such that both p and $p+d$ are Artin primes for g. In particular, we identify a class of pairs (d, g) for which $\pi_{d,g}(x) = 0$. Our results suggest that the distribution of Artin prime pairs, amongst the ordinary prime pairs, is largely governed by a Poisson binomial distribution.

1. INTRODUCTION

1.1. **Background.** Fix an odd prime $p \in \mathbb{N}$, and let $(\mathbb{Z}/p\mathbb{Z})^\times$ denote the finite cyclic group of order $p - 1$. We say that $g \in \mathbb{Z}$ is a *primitive root* mod p (denoted $ord_p(g) = p - 1$) if g mod p generates the group $(\mathbb{Z}/p\mathbb{Z})^\times$. Let $\mathcal{P}$ denote the set of odd primes, and let $\mathcal{P}_g \subseteq \mathcal{P}$ denote the subset of primes p for which $ord_p(g) = p - 1$. If $p \in \mathcal{P}_g$, we say that p is an *Artin prime* (for the root g). *Artin's primitive root conjecture* states that if g is neither a perfect square nor -1, then $\mathcal{P}_g$ contains infinitely many primes.

Fix an integer $g \neq -1$ such that g is not a perfect square. Let $h \in \mathbb{N}$ denote the largest integer such that $g = b^h$ for some $b \in \mathbb{Z}$, and let Δ denote the discriminant of $\mathbb{Q}(\sqrt{g})$. Assuming the *Generalized Riemann Hypothesis* (GRH) for Dedekind zeta functions, Hooley [12] demonstrated that the *natural density* of $\mathcal{P}_g$ in $\mathcal{P}$ is equal to $c_g \cdot A(h)$, i.e. that

$$(1.1) \qquad \lim_{x \to \infty} \frac{\#\{p \in \mathcal{P}_g : p \leq x\}}{\#\{p \in \mathcal{P} : p \leq x\}} = c_g \cdot A(h),$$

where

$$A(h) := \prod_{\substack{p \nmid h \\ p \text{ prime}}} \left(1 - \frac{1}{p(p-1)}\right) \prod_{\substack{p \mid h \\ p \text{ prime}}} \left(1 - \frac{1}{p-1}\right)$$

and

$$(1.2) \qquad c_g := 1 - \mu(|\Delta|) \prod_{\substack{q \mid \Delta \\ q \mid h}} \frac{1}{q - 2} \prod_{\substack{q \mid \Delta \\ q \nmid h}} \frac{1}{q^2 - q - 1}$$

is a correction factor. In particular, $c_g = 1$ whenever $\Delta \not\equiv 1 (4)$.

Date: January 13, 2023.

The correction factor c_g has an interesting history [25]. Initially, E. Artin (1927) actually claimed that the natural density of $\mathcal{P}_g \subseteq \mathcal{P}$ is equal to $A(h)$ for *every* integer $g \neq -1$ that is not a square. Such a model is based on the assumption that any two field extensions of the form $K_q = \mathbb{Q}(\zeta_q, \sqrt[q]{g})$, q prime, are linearly disjoint over $\mathbb{Q}$. In 1957, Derrick and Emma Lehmer brought to Artin's attention the need for an additional correction factor, by noting numerical deviations from Artin's predicted asymptotic in the case $g = 5$. In a reply letter to the Lehmers, Artin suggested a modified heuristic that could correct for the 'entanglement' of the corresponding splitting fields. The adjusted heuristic could better explain the newly observed numerical data, and Artin concluded his letter by noting that he had been "careless but the machine caught up with me." In 1965, Lang and Tate eventually listed the full correction factor, c_g, in their preface to Artin's collected works, and this corrected conjecture was then proved by Hooley two years later, under GRH. For a derivation of c_g (as well as more general Artin-type correction factors) in terms of character sums, see [15].

1.2. **Modeling Artin Primes.** The group $(\mathbb{Z}/p\mathbb{Z})^\times$ has $p-1$ elements and $\varphi(p-1)$ generators, where φ denotes Euler's totient function. The fraction $\varphi(p-1)/p-1$ may thus be viewed as the "probability" that g is a primitive root mod p, and the quantity

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

naively modeled by the sum

$$W_g(x) = \sum_{\substack{p \leq x \\ p \text{ prime}}} \frac{\varphi(p-1)}{p-1}.$$

Indeed,

$$W_g(x) \sim A(1)\mathrm{Li}(x),$$

i.e. the above heuristic yields the correct asymptotic whenever $h = 1$ and $c_g = 1$.

To improve upon this model, we make two simple observations. First, a necessary condition for p to lie in $\mathcal{P}_g$ is for $(p-1, h) = 1$. Indeed, if $(p-1, h) > 1$, then

(1.3) $$g^{\frac{p-1}{(p-1,h)}} \equiv \left(b^{\frac{h}{(p-1,h)}}\right)^{p-1} \equiv 1 \ (\mathrm{mod}\ p),$$

and therefore $p \notin \mathcal{P}_g$. This corresponds to the necessary adjustment of the main term, $A(h)$. Second, a further necessary condition for p to lie in $\mathcal{P}_g$ is for g to be a non-square mod p, i.e. for $(g/p) = -1$, where (g/p) is the Legendre symbol. This corresponds to the additional correction factor, c_g.

If the above two conditions are met, then g lies in one of $(p-1)/2$ specified

residue classes in $(\mathbb{Z}/p\mathbb{Z})^\times$, of which $\varphi(p-1)$ are primitive roots. Such considerations motivated Moree [16, 17] to propose the following general model to mimic the properties of Artin primes. Fix an integer $g \neq -1$ such that g is not a perfect square. For $p \in \mathcal{P}$, define

$$(1.4) \qquad w_g(p) := \begin{cases} 2\frac{\varphi(p-1)}{p-1} & \text{if } \left(\frac{g}{p}\right) = -1 \text{ and } (p-1, h) = 1 \\ 0 & \text{otherwise.} \end{cases}$$

Suppose $S \subseteq \mathcal{P}$ has a non-zero natural density, and let

$$N_S(x) := \{p \leq x : p \in S\}$$

denote the set of primes $p \leq x$ contained in S. To each $p_i \in S$, we assign a Bernoulli distributed random variable E_{p_i}, such that $E_{p_i} = 1$ occurs with probability $w_g(p_i)$, and $E_{p_i} = 0$ occurs with probability $1 - w_g(p_i)$. For fixed x, we then consider the random variable

$$X_{g,N} = \sum_{\substack{p_i \leq x \\ p_i \in S}} E_{p_i}$$

representing the number of successes amongst a sequence of $N := |N_S(x)|$ trials, $\{E_{p_i}\}_{i=1}^N$. As $X_{g,N}$ is a finite sum of independent Bernoulli distributed random variables, it has a *Poisson binomial distribution*, with mean value

$$\mu_{g,N} = \sum_{\substack{p \leq x \\ p \in S}} w_g(p)$$

and variance

$$\sigma_{g,N}^2 = \sum_{\substack{p \leq x \\ p \in S}} (1 - w_g(p)) w_g(p).$$

If the elements in $\mathcal{P}_g$ are well distributed across admissible residue classes, one expects

$$\#\{p : p \leq x, p \in \mathcal{P}_g \cap S\}$$

to look like a "typical" instance of $X_{g,N}$.

1.3. **Artin Primes in Arithmetic Progressions.** As an example, set

$$S = \{p : p \equiv a \pmod{f}\},$$

where $(a, f) = 1$, and let

$$\mathcal{P}_g(x; f, a) := \#\{p : p \leq x, p \equiv a \pmod{f}, \text{ and } p \in \mathcal{P}_g\}.$$

For simplicity sake, we moreover assume that $h = 1$, i.e. that $g \in \mathcal{H}_1$, where

$$\mathcal{H}_1 := \{g \in \mathbb{Z} : g \neq -1, g \text{ is not a perfect power}\} \subset \mathbb{Z}.$$

Under GRH,

$$(1.5) \qquad \mathcal{P}_g(x; f, a) = \mathbf{A}(a, f) \mathbf{c}_g(a, f) \frac{x}{\log x} + O\left(\frac{x \log \log x}{\log 2x}\right)$$

where

$$(1.6) \qquad \mathbf{A}(a,f) := \frac{1}{\varphi(f)} \prod_{p|(a-1,f)} \left(1 - \frac{1}{p}\right) \prod_{p \nmid f} \left(1 - \frac{1}{p(p-1)}\right)$$

and $\mathbf{c}_g(a,f)$ is a correction factor that takes into account quadratic obstructions [14, 25]. In particular, $\mathbf{c}_g(a,f) = 0$ if and only if $\Delta|f$ and $(\Delta/a) = 1$, where (a/b) denotes the Kronecker symbol.

The probabilistic model presented above suggests that in the limit as $x \to \infty$,

$$(1.7) \qquad \mathcal{P}_g(x; f, a) \sim W_{g,S}(x) := \sum_{\substack{p \leq x \\ p \equiv a (\mathrm{mod}\ f)}} w_g(p).$$

Indeed, Moree [16, 17] showed that

$$W_{g,S}(x) \sim \mathbf{A}(a,f)\mathbf{c}_g(a,f)\frac{x}{\log x},$$

i.e. that the model provides an exact asymptotic heuristic for $\mathcal{P}_g(x; f, a)$, under GRH. Variants of Moree's model moreover provide exact asymptotic heuristics for a wide class of problems related to Artin's conjecture, such as an asymptotic count for the number of primes $p \leq x$ for which $ord_p(g)$ is even [6, 18] (unconditional) or for which g is a near primitive root [19, 20] (conditional on GRH). Such models are related to observations suggesting that the distribution of primitive roots across admissible residue classes modulo a large fixed prime $p \in \mathcal{P}$ is largely governed by Poisson processes [7, 23].

1.4. A Hardy-Littlewood Conjecture for Artin Primes.

Let $\mathbf{n} = (n_1, \ldots, n_k)$ denote an integer k-tuple, i.e. an ordered set of k integers, and write $\mathbf{n} \leq x$ if $n_i \leq x$ for all $1 \leq i \leq k$.

Fix $\mathbf{d} = (d_1, \ldots, d_k) \in (\mathbb{Z}_{\geq 0})^k$ to be a k-tuple of even integers. We say that $\mathbf{d}$ is *admissible* if it does not include the complete residue system of any prime p. Let

$$\mathbf{D} := \{\mathbf{n} \in \mathbb{N}^k : n_i = n + d_i \text{ for all } 1 \leq i \leq k, \text{ for some odd } n \in \mathbb{N}\}$$

denote the set of (odd) integer k-tuples with spacing $\mathbf{d}$. Define the *prime k-tuples* (of spacing $\mathbf{d}$) to be the collection

$$\pi_{\mathbf{d}} := \{\mathbf{p} \in \mathbf{D} : p_i = n + d_i \in \mathcal{P} \text{ for all } 1 \leq i \leq k\}.$$

In 1922, Hardy and Littlewood [10] conjectured that in the limit as $x \to \infty$,

$$\#\{\mathbf{p} : \mathbf{p} \leq x \text{ and } \mathbf{p} \in \pi_{\mathbf{d}}\} = \mathfrak{S}(\mathbf{d}) \cdot \frac{x}{(\log x)^k} (1 + o(1)),$$

where $\mathfrak{S}(\mathbf{d})$ is the *singular series*

$$\mathfrak{S}(\mathbf{d}) := \prod_p \frac{1 - \omega_{\mathbf{d}}(p)/p}{(1 - 1/p)^k},$$

and where $\omega_{\boldsymbol{d}}(p)$ counts the number of distinct residue classes amongst $d_1, \ldots, d_k$ modulo p. Note that $\mathfrak{S}(\boldsymbol{d}) > 0$ for any admissible $\boldsymbol{d}$.

We wish to study subsets of $\pi_{\boldsymbol{d}}$ with fixed primitive roots. Fix $\boldsymbol{d} = (d_1, \ldots, d_k)$ to be an admissible k-tuple, and $\boldsymbol{g} = (g_1, \ldots, g_k) \in (\mathbb{Z} \setminus \{-1\})^k$ to be an integer k-tuple such that g_i is not a perfect square for any $1 \le i \le k$. We define the set of *Artin prime k-tuples* (of spacing $\boldsymbol{d}$ with respect to $\boldsymbol{g}$) to be

$$\pi_{\boldsymbol{d},\boldsymbol{g}} := \{\boldsymbol{p} \in \boldsymbol{D} : p_i = n + d_i \in \mathcal{P}_{g_i} \text{ for all } 1 \le i \le k\} \subseteq \pi_{\boldsymbol{d}}.$$

To model this set quantitatively, we employ a variation of the heuristic function $w_g(p)$ defined above. Specifically, for $\boldsymbol{p} \in \pi_{\boldsymbol{d}}$, define

$$w_{\boldsymbol{g}}(\boldsymbol{p}) := w_{g_1}(p_1) \cdots w_{g_k}(p_k).$$

To each $\boldsymbol{p}_i \in \pi_{\boldsymbol{d}}$, we assign a Bernoulli distributed random variable $E_{\boldsymbol{p}_i}$, such that $E_{\boldsymbol{p}_i} = 1$ with probability $w_{\boldsymbol{g}}(\boldsymbol{p}_i)$. We then model $\pi_{\boldsymbol{d},\boldsymbol{g}}$ by a Poisson binomial distribution indexed by the events $\{E_{\boldsymbol{p}_i}\}_{\boldsymbol{p}_i \in \pi_{\boldsymbol{d}}}$. Specifically, define

$$\pi_{\boldsymbol{d},\boldsymbol{g}}(x) := \#\{\boldsymbol{p} : \boldsymbol{p} \le x \text{ and } \boldsymbol{p} \in \pi_{\boldsymbol{d},\boldsymbol{g}}\}.$$

If $\pi_{\boldsymbol{d},\boldsymbol{g}}(x)$ is unbounded, we expect that in the limit as $x \to \infty$,

$$(1.8) \qquad \pi_{\boldsymbol{d},\boldsymbol{g}}(x) \sim W_{\boldsymbol{g},\boldsymbol{d}}(x) := \sum_{\substack{\boldsymbol{p} \le x \\ \boldsymbol{p} \in \pi_{\boldsymbol{d}}}} w_{\boldsymbol{g}}(\boldsymbol{p}).$$

1.5. **Artin Prime Pairs.** In this paper, we use the heuristic model suggested in (1.8) to study the asymptotic growth of *Artin Prime Pairs*, i.e. $\pi_{\boldsymbol{d},\boldsymbol{g}}$ when $k = 2$, $\boldsymbol{d} = (0, d)$ and $\boldsymbol{g} = (g, g)$. Specializing our notation, we let

$$\pi_d := \{p : p, p + d \in \mathcal{P}\}$$

denote the set of prime gaps of size $d \in 2\mathbb{N}$, and

$$\pi_{d,g} := \{p : p, p + d \in \mathcal{P}_g\}$$

denote the set of $p \in \pi_d$ such that g is a primitive root modulo both p and $p + d$. Under the assumption of GRH, Pollack [4, 21] showed the existence of some $C > 0$ such that $\pi_{d,g}$ is an infinite set for some $d \le C$.

Fix $d \in 2\mathbb{N}$. Henceforth, we will moreover assume $g \in \mathcal{H}_1$. For $p \in \pi_d$, define

$$\Delta_{d,g}(p) := \begin{cases} 1 & \text{if } \left(\frac{g}{p}\right) = \left(\frac{g}{p+d}\right) = -1, \\ 0 & \text{otherwise.} \end{cases}$$

Note that since $p \in \mathcal{P}_g \Rightarrow \left(\frac{g}{p}\right) = -1$, it follows that $\Delta_{d,g}(p) = 0 \Rightarrow p \notin \pi_{d,g}$. In particular, if $\Delta_{g,d}(p) = 0$ for all $p \in \pi_d$, then $\pi_{d,g}$ is an empty set. Upon identifying a set of pairs (g, d) such that $\Delta_{g,d}(p) = 0$ for all $p \in \pi_d$, we obtain the following theorem:

Theorem 1.1. *Fix $d \in 2\mathbb{N}$ and $g \in \mathcal{H}_1$, and write $g = g_0 g_s^2$, where g_0 is square-free. Then $\pi_{d,g}$ is empty if the pair (d, g) satisfies any of the following conditions:*

(1) $g_0 = 5$ *and* $d \equiv 2, 3 \pmod 5$,
(2) $g_0 | d$, *and either*
    (a) $d \equiv 2 \pmod 4$ *and* $g_0 \equiv 3 \pmod 4$,
    (b) $d \equiv 4 \pmod 8$ *and* $g_0 \equiv 2 \pmod 4$,
(3) $g_0 \nmid d$, $g_0 | 3d$, *and either*
    (a) $g_0 \equiv 1 \pmod 4$
    (b) $g_0 \equiv 3 \pmod 4$ *and* $d \equiv 0 \pmod 4$,
    (c) $g_0 \equiv 2 \pmod 4$ *and* $d \equiv 0 \pmod 8$.

Conversely, if $d \in 2\mathbb{N}$ and $g \in \mathcal{H}_1$ do not satisfy any of the conditions identified in Theorem 1.1, we conjecture $\pi_{d,g}$ to be an infinite set (and moreover $\pi_{d,g} \subseteq \pi_d$ to have a non-zero natural density amongst the twin-prime pairs). Let

$$\pi_d(x) := \#\{p \in \pi_d : p, p + d \leq x\}$$

and

$$\pi_{d,g}(x) := \#\{p \in \pi_{d,g} : p, p + d \leq x\}.$$

The main aim of this paper is to suggest − and provide support for − the following conjecture:

Conjecture 1.2. (***Artin Two-Tuple Conjecture***): *Fix* $d \in 2\mathbb{N}$ *and* $g \in \mathcal{H}_1$. *Then*

$$\lim_{x \to \infty} \frac{\pi_{d,g}(x)}{\pi_d(x)} = \boldsymbol{c}_g(d)\mathbf{A}(d),$$

where
(1.9)
$$\boldsymbol{A}(d) := \prod_{q|d} \left(1 + \frac{1}{q^2(q-1)} - \frac{2}{q(q-1)}\right) \prod_{q|d\pm1} \left(1 - \frac{1}{q(q-2)}\right) \prod_{q\nmid d(d\pm1)} \left(1 - \frac{2}{q(q-2)}\right),$$

and $\boldsymbol{c}_g(d)$ *is as in* (1.10), (4.20), *or* (4.21). *In particular,* $\boldsymbol{c}_g(d) = 0$ *if and only if the pair* (d, g) *satisfies any of the conditions identified in Theorem 1.1.*

$\mathbf{A}(d)$ may be viewed as a 'linear' constant which emerges upon ignoring the contribution of all non-principal characters in the computation of $\pi_{d,g}$. $\boldsymbol{c}_g(d)$ is a correction factor that includes the contribution from quadratic characters. Conjecture 1.2 is derived from the heuristic assumption that the combined contribution of all additional higher-order characters is asymptotically negligible (Conjecture 2.1).

When $|g| = 2$ the correction factor is particularly simple, namely

(1.10)
$$\boldsymbol{c}_g(d) := \begin{cases} 2 & \text{if } d \equiv 0(8) \\ 1 & \text{if } d \equiv 2, 6(8) \\ 0 & \text{if } d \equiv 4(8). \end{cases}$$

Pollack [21] states a conjecture for $\pi_{2,2}(x)$, i.e. for the particular case $g = 2$ and $d = 2$, though he does not provide any theoretical justification. Zoeteman [26] extends this to a conjecture for $\pi_{d,2}(x)$ ($d \not\equiv 4(8)$), and outlines a

theoretical approach based upon the circle method of Ramanujan, Hardy, and Littlewood. Our work provides independent corroboration, further generalization, and additional theoretical justification, for these earlier conjectures. In particular, our results are obtained by quite different methods than those outlined by Zoeteman, who employs an alternative heuristic for counting Artin prime solutions to general Diophantine equations, implicit in the work of Frei, Koymans, and Sofos [9].

A strong form of the Hardy-Littlewood conjecture asserts that

$$\lim_{x \to \infty} \pi_d(x) = \mathfrak{S}(d) \cdot \mathrm{Li}_2(x) + O(x^{1/2+\epsilon}),$$

where

$$(1.11) \qquad \mathfrak{S}(d) := 2 \prod_{\substack{p \mid d \\ p > 2}} \left(\frac{p-1}{p-2} \right) \prod_{p > 2} \left(1 - \frac{1}{(p-1)^2} \right)$$

and

$$(1.12) \qquad \mathrm{Li}_2(x) := \int_2^x \frac{1}{(\log t)^2} dt.$$

We expect a similar square-root cancellation to hold for Artin twin primes, namely that

$$(1.13) \qquad \pi_{d,g}(x) = \mathbf{c}_g(d)\mathbf{A}(d)\mathfrak{S}(d)\mathrm{Li}_2(x) + O(x^{1/2+\epsilon}).$$

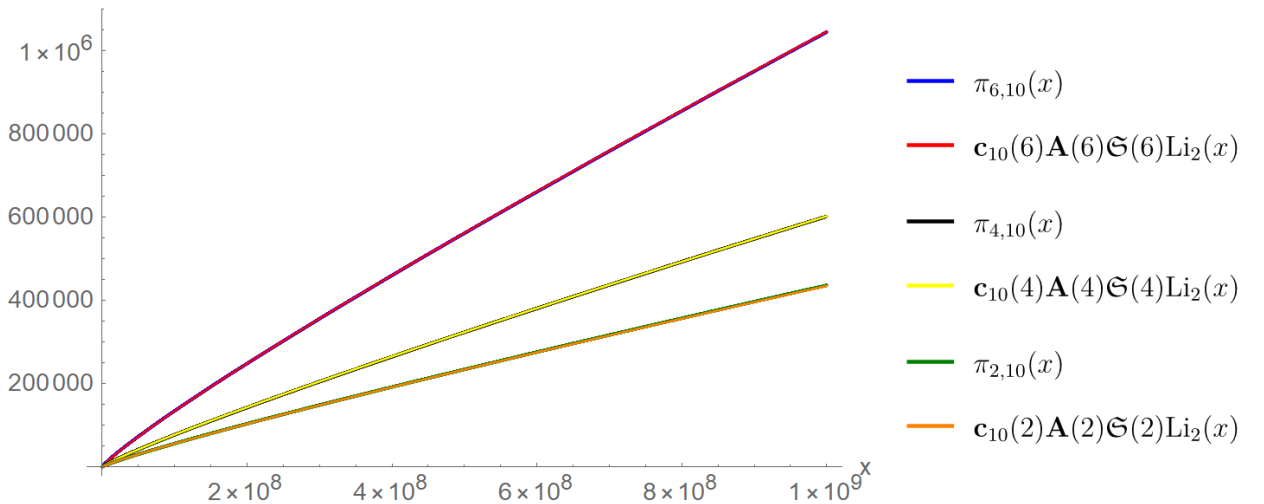Numerical data for Conjecture 1.2 is provided in Figure 1 and Tables $1-3$.

FIGURE 1.
Plots of $\pi_{2,10}(x)$, $\pi_{4,10}(x)$, and $\pi_{6,10}(x)$ for $x \leq 10^9$.

TABLE 1.
Artin Twin Primes (i.e. $p \in \pi_{2,g}$) up to $x = 10^9$ ($2 \leq g \leq 19$)

d	g	$\pi_{d,g}(10^9)$	$\pi_d(10^9)$	$\pi_{d,g}/\pi_d$	$c_g(d)\mathbf{A}(d)$
2	2	434292	3424506	0.126818	0.126943
2	3	869877	3424506	0.254015	0.253886
2	6	433653	3424506	0.126632	0.126943
2	7	501029	3424506	0.146306	0.146176
2	10	436385	3424506	0.127430	0.126943
2	11	485267	3424506	0.141704	0.141338
2	12	870116	3424506	0.254085	0.253886
2	13	320350	3424506	0.093546	0.093631
2	14	434533	3424506	0.126889	0.126943
2	15	267775	3424506	0.078193	0.078118
2	17	467112	3424506	0.136402	0.136476
2	18	435362	3424506	0.127131	0.126943
2	19	457300	3424506	0.133537	0.133665

TABLE 2.
Artin Cousin Primes (i.e. $p \in \pi_{4,g}$) up to $x = 10^9$ ($5 \leq g \leq 20$)

d	g	$\pi_{d,g}(10^9)$	$\pi_d(10^9)$	$\pi_{d,g}/\pi_d$	$c_g(d)\mathbf{A}(d)$
4	5	668522	3424680	0.195207	0.195296
4	6	936167	3424680	0.273358	0.273415
4	7	369095	3424680	0.107775	0.107709
4	10	601828	3424680	0.175732	0.175767
4	11	415533	3424680	0.121334	0.121204
4	13	518948	3424680	0.151531	0.151251
4	14	566971	3424680	0.165554	0.165706
4	15	602032	3424680	0.175792	0.175767
4	17	504006	3424680	0.147168	0.146974
4	19	438500	3424680	0.128041	0.127764
4	20	668812	3424680	0.195291	0.195296

TABLE 3.
Artin Sexy Primes (i.e. $p \in \pi_{6,g}$) up to $x = 10^9$ ($2 \leq g \leq 20$)

d	g	$\pi_{d,g}(10^9)$	$\pi_d(10^9)$	$\pi_{d,g}/\pi_d$	$c_g(d)\mathbf{A}(d)$
6	2	1045109	6849047	0.152591	0.152588
6	5	1492420	6849047	0.217901	0.217982
6	6	1045093	6849047	0.152589	0.152588
6	7	1290264	6849047	0.188385	0.188491
6	10	1044298	6849047	0.152473	0.152588
6	11	1142428	6849047	0.166801	0.166745
6	13	770054	6849047	0.112432	0.112547
6	14	1044000	6849047	0.152429	0.152588
6	15	1344686	6849047	0.196331	0.196184
6	17	842745	6849047	0.123045	0.123035
6	18	1045470	6849047	0.152644	0.152588
6	19	1107167	6849047	0.161652	0.161619
6	20	1492389	6849047	0.217897	0.217982

The remainder of this paper is structured as follows. In Section 2, we write

$$(1.14) \qquad \pi_{d,g}(x) = W_{d,g}(x) + E_{d,g}(x),$$

where

$$(1.15) \qquad W_{d,g}(x) := \sum_{\substack{p \leq x \\ p \in \pi_d}} w_g(p) w_g(p+d)$$

is viewed as a weighted summation over principal and quadratic Dirichlet characters, and $E_{d,g}(x)$ (defined in (2.2)) is a weighted summation over higher-order characters. Similarly flavored identities have previously been employed to study Artin's conjecture over function fields [13, 22], and serve as the theoretical underpinnings for various Poisson binomial distribution models referred to above [16, 18, 19, 20]. Guided by these case studies, we expect the contribution from $E_{d,g}(x)$ to be asymptotically negligible (Conjecture 2.1). In particular, this suggests that the asymptotic behavior of $\pi_{d,g}(x)$ may be modeled by $W_{d,g}(x)$, in the limit as $x \to \infty$.

In Sections 3 we then employ a conjecture concerning twin primes in arithmetic progressions (Conjecture 3.1) to evaluate the sum

$$(1.16) \qquad S_d(x; a, f) := \sum_{\substack{p \leq x \\ p \in \pi_d \\ p \equiv a(f)}} \frac{\varphi(p-1)}{p-1} \frac{\varphi(p+d-1)}{p+d-1}.$$

This enables us to directly compute $W_{d,2}(x)$ (Corollary 3.5), yielding Conjecture 1.2 for the specific the case $|g| = 2$. The cases in which $|g| \neq 2$ are then dealt with in Section 4, and Section 5 explicitly computes those instances in which the correction factor $\mathbf{c}_g(d)$ is equal to zero.

2. Modeling $\pi_{d,g}(x)$ by $W_{d,g}(x)$

Let $(\mathbb{Z}/p\mathbb{Z})^\times$ denote the finite cyclic group of order $p - 1$. For any $g \in (\mathbb{Z}/p\mathbb{Z})^\times$, we define

$$f_p(g) := \begin{cases} 1 & \text{if } ord_p(g) = p - 1 \\ 0 & \text{otherwise.} \end{cases}$$

From [13], it follows that $f_p(g)$ may be written as a *sifting function*

$$f_p(g) = \sum_{k|p-1} \mathcal{S}_{k,p}(g),$$

where

$$\mathcal{S}_{k,p}(g) := \frac{\varphi(p-1)}{p-1} \frac{\mu(k)}{\varphi(k)} \sum_{\text{ord}\chi=k} \chi(g).$$

Here μ is the Möbius function, and the inner sum runs over the multiplicative characters of $(\mathbb{Z}/p\mathbb{Z})^\times$ with order precisely k. In particular, the *linear* contribution to the above sifting function is given by

$$\mathcal{S}_{1,p}(g) = \frac{\varphi(p-1)}{p-1}\chi_0(g),$$

and the *quadratic* contribution to the sifting function is equal to

$$\mathcal{S}_{2,p}(g) = -\frac{\varphi(p-1)}{p-1}\left(\frac{g}{p}\right),$$

and thus we find that

$$\mathcal{S}_{1,p}(g) + \mathcal{S}_{2,p}(g) = w_g(p),$$

where $w_g(p)$ is as in (1.4).

Fix $g \in \mathcal{H}_1$, and write

$$\pi_{d,g}(x) = \sum_{p \in \pi_d(x)} f_p(g) f_{p+d}(g)$$

$$= \sum_{k,l \leq x} H_{k,l}(x; d, g),$$

where

$$H_{k,l}(x; d, g) := \sum_{\substack{p \in \pi_d(x) \\ p \equiv 1(k) \\ p+d \equiv 1(l)}} \mathcal{S}_k(p) \mathcal{S}_l(p+d)$$

$$= \sum_{\substack{p \in \pi_d(x) \\ p \equiv 1(k) \\ p+d \equiv 1(l)}} \frac{\mu(k)}{\varphi(k)} \frac{\mu(l)}{\varphi(l)} \frac{\varphi(p-1)}{p-1} \frac{\varphi(p+d-1)}{p+d-1} \sum_{\mathrm{ord}\chi=k} \chi(g) \sum_{\mathrm{ord}\rho=l} \rho(g).$$

Here χ is a character on $(\mathbb{Z}/p\mathbb{Z})^\times$ while ρ is a character on $(\mathbb{Z}/(p+d)\mathbb{Z})^\times$. In particular,

$$H_{1,1}(x; d, g) = \sum_{\substack{p \in \pi_d(x) \\ p \nmid g}} \frac{\varphi(p-1)}{p-1} \frac{\varphi(p+d-1)}{p+d-1}$$

$$H_{2,1}(x; d, g) = -\sum_{p \in \pi_d(x)} \left(\frac{g}{p}\right) \frac{\varphi(p-1)}{p-1} \frac{\varphi(p+d-1)}{p+d-1}$$

$$H_{1,2}(x; d, g) = -\sum_{p \in \pi_d(x)} \left(\frac{g}{p+d}\right) \frac{\varphi(p-1)}{p-1} \frac{\varphi(p+d-1)}{p+d-1}$$

and

$$H_{2,2}(x; d, g) = \sum_{p \in \pi_d(x)} \left(\frac{g}{p}\right) \left(\frac{g}{p+d}\right) \frac{\varphi(p-1)}{p-1} \frac{\varphi(p+d-1)}{p+d-1}.$$

The $H_{1,1}(x; d, g)$ term corresponds to the 'linear' contribution, i.e. to a naive model constructed without taking into consideration any quadratic obstructions. In particular, it is independent of g. The necessary correction factor is then obtained upon including the contribution from quadratic characters, namely by considering the sum

$$(2.1) \quad H_{1,1}(x; d, g) + H_{1,2}(x; d, g) + H_{2,1}(x; d, g) + H_{2,2}(x; d, g) = W_{d,g}(x).$$

It follows that

$$\pi_{d,g}(x) = W_{d,g}(x) + E_{d,g}(x),$$

where

$$(2.2) \qquad E_{d,g}(x) := \sum_{\substack{k,l\in\mathbb{N} \\ (k,l)\neq(1,1) \\ (k,l)\neq(2,1) \\ (k,l)\neq(1,2) \\ (k,l)\neq(2,2)}} H_{k,l}(x;d,g).$$

Drawing guidance from similarly flavored identities previously studied [6, 13, 22, 16, 18, 19, 20], we expect $W_{d,g}(x)$ to contribute the main term, and for the contribution coming from $E_{d,g}(x)$ to be of lower order:

Conjecture 2.1. (*Moree's Model*) *Fix $d \in 2\mathbb{N}$ and $g \in \mathcal{H}_1$. The contribution of higher-order characters to $\pi_{d,g}(x)$ is asymptotically negligible, i.e.*

$$E_{d,g}(x) = o\left(\frac{x}{(\log x)^2}\right).$$

Conjecture 2.1 amounts to the heuristic assumption that the distribution of Artin prime pairs amongst the ordinary primes pairs is largely governed by a Poisson binomial distribution. In practice, we expect $E_{d,g}(x) = O(x^{1-\delta})$ for some $0 < \delta < 1$.

3. COMPUTING $S_d(x; a, f)$

Let $g \in \mathbb{N}$, and write $g = q_1^{n_1} q_2^{n_2} \ldots q_r^{n_r} \in \mathbb{N}$, where the q_i's are pairwise distinct prime numbers. Fix $d \in 2\mathbb{N}$, and consider the set

$$\varphi_d(g) := \{1 \leq \alpha \leq g : (\alpha, g) = (\alpha + d, g) = 1\}.$$

Note that for prime q,

$$(3.1) \qquad |\varphi_d(q^n)| = \begin{cases} q^{n-1}(q-2) & \text{if } q \nmid d, \\ q^{n-1}(q-1) & \text{if } q|d. \end{cases}$$

Since $|\varphi_d|$ is multiplicative, it follows that

$$(3.2) \qquad |\varphi_d(g)| = g \prod_{i=1}^{s} \frac{(q_i - 1)}{q_i} \prod_{i=s+1}^{r} \frac{(q_i - 2)}{q_i},$$

where $q_1 \ldots, q_s | (g, d)$ and $q_{s+1} \ldots, q_r \nmid (g, d)$.

Let

$$\pi_d(x; a, f) := \#\{p \in \pi_d : p \leq x, \ p \equiv a \pmod{f}\}.$$

The following generalization of the twin prime conjecture (alternatively: a special case of the *Bateman Horn conjecture* [1, 5]) provides an estimate for the size of $\pi_d(x; a, f)$. We require some uniformity in the error term [3], analogous to how the Siegal Walfisz theorem is employed to count Artin primes in previous variants of Moree's model [16, 18, 19, 20]:

Conjecture 3.1. (*Prime Gaps in Arithmetic Progressions*) *Fix* $B > 0$, *and* $a, f \in \mathbb{N}$. *Then for all* $f \leq (\log x)^B$ *with* $(a, f) = (a + d, f) = 1$,

$$\pi_d(x; a, f) = \frac{\mathfrak{S}(d)}{|\varphi_d(f)|} \cdot \mathrm{Li}_2(x) + O_B\left(\frac{x}{(\log x)^C}\right)$$

for any $C > 2$, *where* $\mathfrak{S}(d)$ *is as in* (1.11) (*and where the implied constant depends on the choice of* B).

Lemma 3.2. *Fix* $a, f \in \mathbb{N}$ *such that* $(a, f) = (a + d, f) = 1$, *and moreover assume Conjecture* 3.1. *Then for any* $B > 2$,

$$S_d(x; a, f) = \frac{\mathfrak{S}(d)}{|\varphi_d(f)|} K_{d,f}(a) \boldsymbol{A}_f(d) \mathrm{Li}_2(x) + O\left(\frac{x}{(\log x)^B}\right),$$

where

$$(3.3) \qquad K_{d,f}(a) := \prod_{p \mid (f, a-1)} \left(1 - \frac{1}{p}\right) \prod_{p \mid (f, a+d-1)} \left(1 - \frac{1}{p}\right),$$

and

$$\boldsymbol{A}_f(d) := \sum_{\substack{m_1^1, m_2^2 \\ (m_1^1 m_2^2, f)=1 \\ (m_1^1, m_2^2)=1 \\ (m_1^1, 1+d)=1 \\ (m_2^2, 1-d)=1}} \frac{\mu(m_1^1)}{m_1^1 |\varphi_d(m_1^1)|} \frac{\varphi((m_1^1, d))}{(m_1^1, d)} \frac{\mu(m_2^2)}{m_2^2 |\varphi_d(m_2^2)|}.$$

Proof. By Möbius inversion,

$$(3.4) \qquad \frac{\varphi(p-1)}{p-1} = \sum_{k \mid p-1} \frac{\mu(k)}{k},$$

and therefore

$$S_d(x; a, f) = \sum_{\substack{p \leq x \\ p \in \pi_d \\ p \equiv a(f)}} \left(\sum_{M_1 \mid p-1} \frac{\mu(M_1)}{M_1}\right) \left(\sum_{M_2 \mid p+d-1} \frac{\mu(M_2)}{M_2}\right)$$

$$(3.5) \qquad = \sum_{\substack{M_1, M_2 \\ M_1, M_2 \; \square\text{-free}}} \frac{\mu(M_1)\mu(M_2)}{M_1 M_2} \sum_{\substack{p \leq x \\ p \in \pi_d \\ p \equiv a(f) \\ p \equiv 1(M_1) \\ p \equiv 1-d(M_2)}} 1.$$

Write

$$(3.6) \qquad M_1 = m_1^0 \cdot m_1^1, \qquad \text{and} \qquad M_2 = m_2^0 \cdot m_2^1 \cdot m_2^2,$$

where

$$m_1^0 := (f, M_1), \qquad m_1^1 := \frac{M_1}{(f, M_1)},$$

and

$$m_2^0 := (f, M_2), \quad m_2^1 := \left(\frac{M_1}{(f, M_1)}, M_2 \right) \ m_2^2 := \frac{M_2}{(f, M_2) \cdot \left(\frac{M_1}{(f, M_1)}, M_2 \right)}.$$

In other words, we group the primes p dividing M_2 as follows: if $p \mid f$ then $p|m_2^0$; if $p \nmid f$ but $p|M_1$ then $p|m_2^1$; and if $p \nmid f \cdot M_1$ then $p|m_2^2$.

By the Chinese Remainder Theorem (CRT), the congruence conditions

$$
\begin{aligned}
p &\equiv a \pmod{f} \\
p &\equiv 1 \pmod{M_1} \\
p &\equiv 1 - d \pmod{M_2}
\end{aligned}
$$

(3.7)

are simultaneously solvable if and only if

(3.8) $$a \equiv 1 \pmod{m_1^0}$$

(3.9) $$a \equiv 1 - d \pmod{m_2^0}$$

(3.10) $$1 \equiv 1 - d \pmod{m_2^1},$$

where we further employ the fact that

$$1 \equiv 1 - d \pmod{m_2^1} \Leftrightarrow 1 \equiv 1 - d \pmod{(M_1, M_2)}$$

whenever (3.8) and (3.9) are both satisfied. Such a solution is moreover unique modulo

$$LCM(f, M_1, M_2) = f m_1^1 m_2^2,$$

and the explicit solution is given by

(3.11) $$p \equiv a \pmod{f}$$

(3.12) $$p \equiv 1 \pmod{m_1^1}$$

(3.13) $$p \equiv 1 - d \pmod{m_2^2},$$

where $\{f, m_1^1, m_2^2\}$ now form a pairwise coprime set.

It follows that if A_M denotes the unique residue class modulo $f m_1^1 m_2^2$ for which the conditions in (3.11)−(3.13) are satisfied, then

$$S_d(x; a, f) = \sum_{\substack{m_1^1, m_2^2 \\ (m_1^1 m_2^2, f) = 1 \\ (m_1^1, m_2^2) = 1}} \frac{\mu(m_1^1)\mu(m_2^2)}{m_1^1 m_2^2} \sum_{\substack{m_1^0, m_2^0 \\ m_1^0 | (f, a-1) \\ m_2^0 | (f, a+d-1)}} \frac{\mu(m_1^0)\mu(m_2^0)}{m_1^0 m_2^0} \sum_{m_2^1 | (d, m_1^1)} \frac{\mu(m_2^1)}{m_2^1}$$

(3.14) $$\times \pi_d(x; A_M, f m_1^1 m_2^2).$$

Next, suppose $A_M \in \varphi_d \left(f m_1^1 m_2^2 \right)$, i.e.,

$$\left(A_M, f m_1^1 m_2^2 \right) = \left(A_M + d, f m_1^1 m_2^2 \right) = 1.$$

By the multiplicative property of the greatest common divisor function,

$$\left(A_M, fm_1^1 m_2^2\right) = 1 \Leftrightarrow (A_M, f) = \left(A_M, m_1^1\right) = \left(A_M, m_2^2\right) = 1$$
$$\Leftrightarrow (1-d, m_2^2) = 1,$$

where in the last step we note that $(A_M, m_1^1) = 1$ already follows from (3.12), and that $(A_M, f) = 1$ already follows from the assumption of Lemma 3.2. Similarly, we find that

$$\left(A_M + d, fm_1^1 m_2^2\right) = 1 \Leftrightarrow (1+d, m_1^1) = 1.$$

In other words, given the assumptions of Lemma 3.2, $A_M \in \varphi_d(fm_1^1 m_2^2)$ implies the additional constraints

$$(3.15) \qquad (m_2^2, 1-d) = 1 \quad \text{and} \quad (m_1^1, 1+d) = 1.$$

Next, note that if $A_M \notin \varphi_d(fm_1^1 m_2^2)$, then

$$\pi_d(x; A_M, fm_1^1 m_2^2) \leq 1$$

uniformly in m_1^1 and m_2^2. Moreover, since the inner sums in (3.14) are finite, and since $\pi_d(x; A_M, fm_1^1 m_2^2) = 0$ whenever either $m_1^1, m_2^2 > x$, we may bound the contribution from terms for which $A_M \notin \varphi_d(fm_1^1 m_2^2)$ by

$$\ll \sum_{m_1^1, m_2^2 < x} \frac{\mu(m_1^1)\mu(m_2^2)}{m_1^1 m_2^2} \pi_d\left(x; A_M, fm_1^1 m_2^2\right) \ll \sum_{m_1^1 < x} \frac{1}{m_1^1} \sum_{m_2^2 < x} \frac{1}{m_2^2} \ll (\log x)^2.$$

It then follows that

$$(3.16)$$
$$S_d(x; a, f) = \sum_{\substack{m_1^1, m_2^2 \\ (m_1^1 m_2^2, f)=1 \\ (m_1^1, m_2^2)=1 \\ (m_1^1, 1+d)=1 \\ (m_2^2, 1-d)=1}} \frac{\mu(m_1^1)\mu(m_2^2)}{m_1^1 m_2^2} \sum_{\substack{m_1^0, m_2^0 \\ m_1^0 | (f, a-1) \\ m_2^0 | (f, a+d-1)}} \frac{\mu(m_1^0)\mu(m_2^0)}{m_1^0 m_2^0} \sum_{m_2^1 | (d, m_1^1)} \frac{\mu(m_2^1)}{m_2^1}$$

$$\times \pi_d(x; A_M, fm_1^1 m_2^2) + O\left((\log x)^2\right).$$

Next, suppose $A_M \in \varphi_d(fm_1^1 m_2^2)$. Since $\pi_d(x; A_M, fm_1^1 m_2^2) \leq \frac{x}{m_1^1 m_2^2}$, we find that for any fixed $B > 2$,

$$\sum_{\substack{m_1^1, m_2^2 \\ m_1^1 > (\log x)^B}} \frac{\mu(m_1^1)\mu(m_2^2)}{m_1^1 m_2^2} \pi_d(x; A_M, fm_1^1 m_2^2) \leq \sum_{m_1^1 > (\log x)^B} \frac{x}{(m_1^1)^2} \sum_{m_2^2 = 1}^{\infty} \frac{1}{(m_2^2)^2}$$

$$= O\left(\frac{x}{(\log x)^B}\right),$$

and similarly for the sum over $m_2^2 > (\log x)^B$. Upon applying Conjecture 3.1, we conclude that

(3.17)

$$S_d(x; a, f) = \frac{\mathfrak{S}(d)}{|\varphi_d(f)|} \cdot \mathrm{Li}_2(x) \times$$

$$\sum_{\substack{m_1^1, m_2^2 < (\log x)^B \\ (m_1^1 m_2^2, f) = 1 \\ (m_1^1, m_2^2) = 1 \\ (m_1^1, 1+d) = 1 \\ (m_2^2, 1-d) = 1}} \sum_{\substack{m_1^0, m_2^0 \\ m_1^0 | (f, a-1) \\ m_2^0 | (f, a+d-1)}} \sum_{m_2^1 | (d, m_1^1)} \frac{\mu(m_1^0) \mu(m_1^1) \mu(m_2^0) \mu(m_2^1) \mu(m_2^2)}{m_1^0 m_1^1 m_2^0 m_2^1 m_2^2 \left| \varphi_d(m_1^1 m_2^2) \right|}$$

$$+ O\left(\frac{x}{(\log x)^B} \right).$$

Here we note that all the error terms coming from each application of Conjecture 3.1 may together be incorporated into $O\left(x(\log x)^{-B}\right)$, since upon fixing $\epsilon > 0$ and setting $C = B + \epsilon$ in each application of Conjecture 3.1, we bound the error by

$$\ll \sum_{\substack{m_1^1 < (\log x)^B \\ m_2^2 < (\log x)^B}} \frac{\mu(m_1^1) \mu(m_2^2)}{m_1^1 m_2^2} \cdot \frac{x}{(\log x)^C} \ll (\log \log x)^2 \cdot \frac{x}{(\log x)^{B+\epsilon}} \ll \frac{x}{(\log x)^B}.$$

Finally, by Mertens' third theorem,

$$\lim_{n \to \infty} \log n \prod_{p \leq n} \left(1 - \frac{1}{p} \right) = e^{-\gamma},$$

where γ is the Euler-Mascheroni constant. It thus follows from (3.2) that for square-free $n \in \mathbb{N}$,

$$\frac{1}{|\varphi_d(n)|} \leq \frac{2}{n} \prod_{\substack{p \leq n \\ p \neq 2}} \frac{p}{p-2} = \frac{2}{n} \prod_{\substack{p \leq n \\ p \neq 2}} \left(\frac{p}{p-1} \right)^2 \left(1 - \frac{1}{(p-1)^2} \right)^{-1} \sim \frac{e^{2\gamma}}{\mathfrak{S}(2)} \frac{(\log n)^2}{n},$$

in the limit as $n \to \infty$. Thus

$$\sum_{m_1^1 > (\log x)^B} \sum_{m_2^2 = 1}^{\infty} \frac{\mu(m_1^1) \mu(m_2^2)}{m_1^1 m_2^2 |\varphi_d(m_1^1 m_2^2)|} \ll \sum_{m_1^1 > (\log x)^B} \sum_{m_2^2 = 1}^{\infty} \frac{\log(m_1^1 m_2^2)^2}{(m_1^1 m_2^2)^2} \ll (\log x)^{-(B-\epsilon')},$$

for any $0 < \epsilon' < 2$, i.e.

(3.18)

$$\sum_{m_1^1 > (\log x)^B} \sum_{m_2^2 = 1}^{\infty} \frac{\mathrm{Li}_2(x)}{m_1^1 m_2^2 |\varphi_d(m_1^1 m_2^2)|} = \sum_{m_2^2 > (\log x)^B} \sum_{m_1^1 = 1}^{\infty} \frac{\mathrm{Li}_2(x)}{m_1^1 m_2^2 |\varphi_d(m_1^1 m_2^2)|}$$

$$= O\left(\frac{x}{(\log x)^{2+B-\epsilon'}} \right) = O\left(\frac{x}{(\log x)^{B+\epsilon}} \right).$$

Lemma 3.2 then follows from (3.17) and (3.18), where by (3.4) we note that

$$\sum_{m_2^1 \mid (d, m_1^1)} \frac{\mu(m_2^1)}{m_2^1} = \frac{\varphi((m_1^1, d))}{(m_1^1, d)},$$

and that

$$\sum_{\substack{m_1^0, m_2^0 \\ m_1^0 \mid (f, a-1) \\ m_2^0 \mid (f, a+d-1)}} \frac{\mu(m_1^0)\mu(m_2^0)}{m_1^0 m_2^0} = \frac{\varphi((f, a-1))}{(f, a-1)} \frac{\varphi((f, a+d-1))}{(f, a+d-1)}$$

$$= \prod_{p \mid (f, a-1)} \left(1 - \frac{1}{p}\right) \prod_{p \mid (f, a+d-1)} \left(1 - \frac{1}{p}\right).$$

□

Remark 3.3. In fact, a slightly weaker form of Conjecture 3.1 is sufficient to derive Conjecture 1.2. Specifically, to compute the asymptotic growth of $S_d(x; a, f)$ in Lemma 3.2 (and subsequently to compute the asymptotic growth of $W_{d,g}(x)$ in Theorem 4.4), one need only assume Conjecture 3.1 for some fixed $B, C > 2$.

3.1. **A Product Formula for** $S_d(x; a, f)$**.** Let

$$(3.19) \qquad \mathbf{A}(d) = \prod_{p \text{ prime}} \rho_d(p),$$

where

$$\rho_d(p) := \begin{cases} 1 + \frac{1}{p^2(p-1)} - \frac{2}{p(p-1)} & \text{if } p \mid d \\ 1 - \frac{1}{p(p-2)} & \text{if } p \mid d \pm 1 \\ 1 - \frac{2}{p(p-2)} & \text{if } p \nmid d(d^2 - 1). \end{cases}$$

Corollary 3.4. *Fix* $a, f \in \mathbb{N}$ *such that* $(a, f) = (a+d, f) = 1$, *and moreover assume Conjecture 3.1. Then for any* $B > 2$,

$$S_d(x; a, f) = \mathfrak{S}(d) \boldsymbol{e}_f(d) \boldsymbol{A}(d) K_{d,f}(a) \mathrm{Li}_2(x) + O\left(\frac{x}{(\log x)^B}\right),$$

where

$$(3.20) \qquad \boldsymbol{e}_f(d) := \frac{1}{|\varphi_d(f)|} \prod_{p \mid f} \rho_d(p)^{-1}.$$

In particular,

$$S_d(x; 1, 1) = H_{1,1}(x; d, g) = \boldsymbol{A}(d)\mathfrak{S}(d)\mathrm{Li}_2(x) + O\left(\frac{x}{(\log x)^B}\right).$$

Proof. Let η be a multiplicative function, and note that

$$\sum_{\substack{n=1 \\ n \text{ square-free}}}^{\infty} \eta(n) = \prod_p (1 + \eta(p)),$$

whenever the series $\sum_{n=1}^{\infty} \eta(n)$ is absolutely convergent (see [2, Theorem 11.6]). Upon defining the multiplicative functions

$$\eta_1(n) := \begin{cases} \frac{\mu(n)}{n|\varphi_d(n)|} \frac{\varphi((n,d))}{(n,d)} & \text{if } (n, f) = 1 \text{ and } (n, 1 + d) = 1 \\ 0 & \text{otherwise,} \end{cases}$$

$$\eta_2(n) := \begin{cases} \frac{\mu(n)}{n|\varphi_d(n)|} & \text{if } (n, f) = 1 \text{ and } (n, 1 - d) = 1 \\ 0 & \text{otherwise,} \end{cases}$$

it then follows that

$$\mathbf{A}_f(d) = \sum_{\substack{n,m \\ n,m \text{ square-free} \\ (n,m)=1}} \eta_1(n)\eta_2(m) = \prod_p \left(1 + \eta_1(p) + \eta_2(p)\right).$$

Upon noting that for any $p \nmid f$,

$$(3.21) \qquad \eta_1(p) = \begin{cases} 0 & \text{if } p|d + 1 \\ -\frac{1}{p^2} & \text{if } p|d \\ -\frac{1}{p(p-2)} & \text{if } p \nmid d(d + 1), \end{cases}$$

and

$$(3.22) \qquad \eta_2(p) = \begin{cases} 0 & \text{if } p|d - 1 \\ -\frac{1}{p(p-1)} & \text{if } p|d \\ -\frac{1}{p(p-2)} & \text{if } p \nmid d(d - 1), \end{cases}$$

we conclude that

$$\mathbf{A}_f(d) = \prod_{p \nmid f} \rho_d(p) = \mathbf{A}(d)\mathbf{e}_f(d),$$

from which Corollary 3.4 now follows. $\square$

3.2. The Correction Factor for $g = 2$. We can now compute $W_{d,g}(x)$ in the particular case that $g = 2$ (the case $g = -2$ is very similar):

Corollary 3.5. *Fix $g = 2$ and assume Conjecture 3.1. Then*

$$(3.23) \qquad W_{d,2}(x) = \boldsymbol{c}_2(d)\boldsymbol{A}(d)\mathfrak{S}(d)\mathrm{Li}_2(x) + O\left(\frac{x}{(\log x)^C}\right)$$

where $\boldsymbol{c}_2(d)$ is as in (1.10).

Proof. Since

$$(3.24) \qquad \left(\frac{2}{p}\right) = \begin{cases} 1 & \text{if } p \equiv 1, 7(8) \\ -1 & \text{if } p \equiv 3, 5(8), \end{cases}$$

we find that

$$(3.25) \qquad W_{d,2}(x) = 4 \times \begin{cases} S_d(x; 3, 8) + S_d(x; 5, 8) & \text{if } d \equiv 0(8) \\ S_d(x; 3, 8) & \text{if } d \equiv 2(8) \\ 0 & \text{if } d \equiv 4(8) \\ S_d(x; 5, 8) & \text{if } d \equiv 6(8). \end{cases}$$

Equation (3.23) now follows from Corollary 3.4 upon noting that

$$S_d(x; 3, 8) = S_d(x; 5, 8) = \frac{1}{4} \mathbf{A}(d) \mathfrak{S}(d) \mathrm{Li}_2(x) + O\left(\frac{x}{(\log x)^C}\right).$$

$\square$

4. The Correction Factor for $g \neq 2$

4.1. Quadratic Reciprocity.

To compute $\mathbf{c}_g(d)$ when $g \neq 2$, we make extensive use of quadratic reciprocity. For $g \in \mathcal{H}_1 \setminus \{\pm 2\}$, write $g = g_0 g_s^2$, where g_0 is square-free, and define

$$(4.1) \qquad \tilde{g} := \begin{cases} |g_0| & \text{if } g_0 \text{ is odd,} \\ \left|\frac{g_0}{2}\right| & \text{if } g_0 \text{ is even.} \end{cases}$$

Fix $d \in 2\mathbb{N}$ and $\varepsilon_1, \varepsilon_2 \in \{0, \pm 1\}$. For odd $n \in \mathbb{N}$, define

$$\delta_{\varepsilon_1, \varepsilon_2}(\tilde{g}, d)(n) := \begin{cases} 1 & \text{if } \left(\frac{n}{\tilde{g}}\right) = \varepsilon_1 \text{ and } \left(\frac{n+d}{\tilde{g}}\right) = \varepsilon_2 \\ 0 & \text{otherwise,} \end{cases}$$

where $(n/\tilde{g})$ now denotes the *Jacobi* symbol.

Lemma 4.1. *Fix $p \in \pi_d$ and suppose $(p, g_s) = (p + d, g_s) = 1$. Then*

$$(4.2) \qquad \Delta_{d,g}(p) = \begin{cases} \delta_{-1,-1}(\tilde{g}, d)(p) & \text{if } g_0 \equiv 1(4) \\ \delta_{T_1(p), T_1(p+d)}(\tilde{g}, d)(p) & \text{if } g_0 \equiv 3(4) \\ \delta_{T_2(p), T_2(p+d)}(\tilde{g}, d)(p) & \text{if } g_0 \equiv 2(8) \\ \delta_{T_3(p), T_3(p+d)}(\tilde{g}, d)(p) & \text{if } g_0 \equiv 6(8), \end{cases}$$

where T_1, T_2, and T_3 are as in (4.8), (4.11), (4.13), respectively.

Proof. Since $(p, g_s) = (p + d, g_s) = 1$, we have $\left(\frac{g}{p}\right) = \left(\frac{g_0}{p}\right)$ and $\left(\frac{g}{p+d}\right) = \left(\frac{g_0}{p+d}\right)$. The idea of the proof is to use quadratic reciprocity to write $\left(\frac{g}{p}\right)$ and $\left(\frac{g}{p+d}\right)$ in terms of $\left(\frac{p}{\tilde{g}}\right)$ and $\left(\frac{p+d}{\tilde{g}}\right)$. The result depends heavily on the particular residue classes of n, d, and g_0.

First, suppose $g_0 \equiv 1 \pmod 4$. If $g_0 > 0$, then

$$(4.3) \qquad \left(\frac{g_0}{p}\right) = \left(\frac{p}{g_0}\right) = \left(\frac{p}{\tilde{g}}\right),$$

and similarly if $g_0 < 0$, then

$$\left(\frac{g_0}{p}\right) = \left(\frac{-1}{p}\right)\left(\frac{\tilde{g}}{p}\right) = \left(\frac{-1}{p}\right)(-1)^{\frac{p-1}{2}}\left(\frac{p}{\tilde{g}}\right) = \left(\frac{p}{\tilde{g}}\right).$$

It follows that

$$(4.4) \qquad \Delta_{d,g}(p) = \delta_{-1,-1}(\tilde{g}, d)(p).$$

Next, suppose $g_0 \equiv 3 \pmod 4$. If $g_0 > 0$, then

$$(4.5) \qquad \left(\frac{g_0}{p}\right) = \begin{cases} -\left(\frac{p}{\tilde{g}}\right) & \text{if } p \equiv 3(4) \\ \left(\frac{p}{\tilde{g}}\right) & \text{if } p \equiv 1(4), \end{cases}$$

and similarly if $g_0 < 0$, then

$$(4.6) \qquad \left(\frac{g_0}{p}\right) = \left(\frac{-1}{p}\right)\left(\frac{\tilde{g}}{p}\right) = \left(\frac{-1}{p}\right)\left(\frac{p}{\tilde{g}}\right) = \begin{cases} -\left(\frac{p}{\tilde{g}}\right) & \text{if } p \equiv 3(4) \\ \left(\frac{p}{\tilde{g}}\right) & \text{if } p \equiv 1(4). \end{cases}$$

Upon defining

$$(4.7) \qquad T_1(n) := \begin{cases} 1 & \text{if } n \equiv 3(4) \\ -1 & \text{if } n \equiv 1(4) \end{cases}$$

for odd $n \in \mathbb{N}$, it follows that for $g_0 \equiv 3(4)$,

$$\left(\frac{g_0}{p}\right) = -1 \Leftrightarrow \left(\frac{p}{\tilde{g}}\right) = T_1(p),$$

and therefore that

$$(4.8) \qquad \Delta_{d,g}(p) = \delta_{T_1(p), T_1(p+d)}(\tilde{g}, d)(p).$$

Next, suppose g_0 is even. If $g_0 > 0$, then

$$(4.9) \qquad \left(\frac{g_0}{p}\right) = \left(\frac{2}{p}\right)\left(\frac{\tilde{g}}{p}\right) = \begin{cases} \left(\frac{\tilde{g}}{p}\right) & \text{if } p \equiv 1, 7(8) \\ -\left(\frac{\tilde{g}}{p}\right) & \text{if } p \equiv 3, 5(8), \end{cases}$$

while if $g_0 < 0$,

$$(4.10) \qquad \left(\frac{g_0}{p}\right) = \left(\frac{2}{p}\right)\left(\frac{-1}{p}\right)\left(\frac{\tilde{g}}{p}\right) = \begin{cases} \left(\frac{\tilde{g}}{p}\right) & \text{if } p \equiv 1, 3(8) \\ -\left(\frac{\tilde{g}}{p}\right) & \text{if } p \equiv 5, 7(8). \end{cases}$$

It follows that when $g_0 \equiv 2(8)$,

$$\left(\frac{g_0}{p}\right) = -1 \Leftrightarrow \left(\frac{p}{\tilde{g}}\right) = T_2(p),$$

where for odd $n \in \mathbb{N}$,

$$(4.11) \qquad T_2(n) := \begin{cases} 1 & \text{if } n \equiv 3, 5(8) \\ -1 & \text{if } n \equiv 1, 7(8). \end{cases}$$

Hence

$$(4.12) \qquad \Delta_{d,g}(p) = \delta_{T_2(p), T_2(p+d)}(\tilde{g}, d)(p).$$

Similarly, for $g_0 \equiv 6(8)$, we find that

$$\left(\frac{g_0}{p}\right) = -1 \Leftrightarrow \left(\frac{p}{\tilde{g}}\right) = T_3(p),$$

where

$$(4.13) \qquad T_3(n) := \begin{cases} 1 & \text{if } n \equiv 5, 7(8) \\ -1 & \text{if } n \equiv 1, 3(8). \end{cases}$$

Hence

(4.14) $$\Delta_{d,g}(p) = \delta_{T_3(p),T_3(p+d)}(\tilde{g},d)(p),$$

as desired. $\qquad\square$

4.2. Properties of $\Omega_{d,f}$.

Fix $d \in 2\mathbb{N}$, and $f \in \mathbb{N}$ to be odd and square-free. Define

(4.15) $$\Omega_{d,f}(\varepsilon_1,\varepsilon_2) := \mathbf{e}_f(d) \sum_{a \in \Phi_{\varepsilon_1,\varepsilon_2}(f,d)} K_{d,f}(a),$$

where

(4.16) $$\Phi_{\varepsilon_1,\varepsilon_2}(f,d) := \left\{ 1 \le \alpha \le f : \left(\frac{\alpha}{f}\right) = \varepsilon_1 \text{ and } \left(\frac{\alpha+d}{f}\right) = \varepsilon_2 \right\},$$

and $\varepsilon_1, \varepsilon_2 \in \{-1,1\}$.

Furthermore, if $1 \le a \le f$ is the unique solution to the congruence conditions

$$a \equiv a_1 \pmod{f_1}$$
(4.17) $$a \equiv a_2 \pmod{f_2},$$

where $(f_1, f_2) = 1$ and $f = f_1 f_2$, we moreover define

$$\Omega_{a_2}^{f_2}(\varepsilon_1,\varepsilon_2) := \mathbf{e}_f(d) \sum_{a_1 \in \Phi_{\varepsilon_1,\varepsilon_2}(f_1,d)} K_{d,f}(a).$$

Lemma 4.2.
$$\Omega_{a_2}^{f_2}(\varepsilon_1,\varepsilon_2) = e_{f_2}(d) K_{d,f_2}(a_2) \Omega_{d,f_1}(\epsilon).$$

Proof. Note that

$$e_f(d) = e_{f_1}(d) e_{f_2}(d)$$

and

$$K_{d,f}(a) = K_{d,f_1}(a_1) K_{d,f_2}(a_2).$$

It follows that

$$\begin{aligned}
\Omega_{a_2}^{f_2}(\varepsilon_1,\varepsilon_2) &= \sum_{a_1 \in \Phi_\epsilon(f_1,d)} e_f(d) K_{d,f}(a) \\
&= e_{f_2}(d) K_{d,f_2}(a_2) e_{f_1}(d) \sum_{a_1 \in \Phi_\epsilon(f_1,d)} K_{d,f_1}(a_1) \\
&= e_{f_2}(d) K_{d,f_2}(a_2) \Omega_{d,f_1}(\epsilon),
\end{aligned}$$

as desired. $\qquad\square$

The following lemma provides a simpler description of $\Omega_{d,f}(\varepsilon_1, \varepsilon_2)$ that is useful for numerical computations:

Lemma 4.3. *Fix $d \in 2\mathbb{N}$ and $f \in \mathbb{N}$ to be odd and square-free. Then*

$$\Omega_{d,f}(\varepsilon_1, \varepsilon_2) = \frac{\mathbf{e}_f(d)}{4} \left(c(f) + \varepsilon_1 c_1(f) + \varepsilon_2 c_2(f) + \varepsilon_1 \varepsilon_2 c_{12}(f) \right),$$

where c, c_1, c_2, c_{12} are multiplicative functions defined on primes $q \in \mathbb{N}$ by

$$c(q) := \begin{cases} q - 1 - \frac{2}{q} + \frac{1}{q^2} & \text{if } q \mid d \\ q - 2 - \frac{1}{q} & \text{if } q \mid d \pm 1 \\ q - 2 - \frac{2}{q} & \text{if } q \nmid d(d \pm 1), \end{cases}$$

$$c_1(q) := \begin{cases} -\frac{2}{q} + \frac{1}{q^2} & \text{if } q \mid d \\ -1 - \frac{1}{q}\left(\frac{1-d}{q}\right) & \text{if } q \mid d+1 \\ -\left(\frac{-d}{q}\right) - \frac{1}{q}\left(1 + \left(\frac{1-d}{q}\right)\right) & \text{if } q \nmid d(d+1), \end{cases}$$

$$c_2(q) := \begin{cases} -\frac{2}{q} + \frac{1}{q^2} & \text{if } q \mid d \\ -1 - \frac{1}{q}\left(\frac{1+d}{q}\right) & \text{if } q \mid d-1 \\ -\left(\frac{d}{q}\right) - \frac{1}{q}\left(1 + \left(\frac{1+d}{q}\right)\right) & \text{if } q \nmid d(d-1), \end{cases}$$

$$c_{12}(q) := \begin{cases} q - 1 - \frac{2}{q} + \frac{1}{q^2} & \text{if } q \mid d \\ -1 - \frac{1}{q}\left(\left(\frac{1+d}{q}\right) + \left(\frac{1-d}{q}\right)\right) & \text{if } q \nmid d. \end{cases}$$

Proof. Note that

$$|\Phi_{\varepsilon_1,\varepsilon_2}(f,d)| = \frac{1}{4} \sum_{\substack{1 \le \alpha \le f \\ (\alpha,f)=1 \\ (\alpha+d,f)=1}} \left(1 + \varepsilon_1\left(\frac{\alpha}{f}\right)\right)\left(1 + \varepsilon_2\left(\frac{\alpha+d}{f}\right)\right),$$

and therefore that

$$\Omega_{d,f}(\varepsilon_1,\varepsilon_2) = \mathbf{e}_f(d) \sum_{a \in \Phi_{\varepsilon_1,\varepsilon_2}(f,d)} \prod_{p|(a-1,f)} \left(1 - \frac{1}{p}\right) \prod_{p|(a+d-1,f)} \left(1 - \frac{1}{p}\right)$$

$$= \frac{\mathbf{e}_f(d)}{4} \sum_{\substack{1 \le a \le f \\ (a,f)=1 \\ (a+d,f)=1}} \left(1 + \varepsilon_1\left(\frac{a}{f}\right)\right)\left(1 + \varepsilon_2\left(\frac{a+d}{f}\right)\right) \sum_{k|(a-1,f)} \sum_{l|(a+d-1,f)} \frac{\mu(k)}{k}\frac{\mu(l)}{l}$$

$$= \frac{\mathbf{e}_f(d)}{4} \sum_{k,l|f} \frac{\mu(k)}{k}\frac{\mu(l)}{l} \sum_{\substack{1 \le a \le f \\ (a,f)=1 \\ (a+d,f)=1 \\ a \equiv 1 \pmod{k} \\ a \equiv 1-d \pmod{l}}} \left(1 + \varepsilon_1\left(\frac{a}{f}\right) + \varepsilon_2\left(\frac{a+d}{f}\right) + \varepsilon_1\varepsilon_2\left(\frac{a}{f}\right)\left(\frac{a+d}{f}\right)\right).$$

Split the inner sum into four parts, and then consider the case when $f = q$ is prime. Employing that fact that

$$(4.18) \qquad \sum_{\substack{1 \leq \alpha \leq q \\ (\alpha+d,q)=1}} \left(\frac{\alpha}{q}\right) = \sum_{1 \leq \alpha \leq q} \left(\frac{\alpha}{q}\right) - \left(\frac{-d}{q}\right) = \begin{cases} 0 & \text{if } q|d \\ -\left(\frac{-d}{q}\right) & \text{if } q \nmid d, \end{cases}$$

as well as that

$$(4.19) \qquad \begin{aligned} \sum_{1 \leq \alpha \leq q} \left(\frac{\alpha}{q}\right)\left(\frac{\alpha+d}{q}\right) &= \sum_{1 \leq \alpha < q} \left(\frac{\alpha^2}{q}\right)\left(\frac{\alpha^{-1}(\alpha+d)}{q}\right) \\ &= \sum_{1 \leq \alpha < q} \left(\frac{1+\alpha^{-1}d}{q}\right) = \begin{cases} q-1 & \text{if } q|d, \\ -1 & \text{if } q \nmid d, \end{cases} \end{aligned}$$

we find that the sums are equal to $c(q)$, $c_1(q)$, $c_2(q)$ and $c_{12}(q)$, respectively. By CRT, we moreover find that for $f = f_1 f_2$, where $(f_1, f_2) = 1$,

$$\mathbf{e}_f(d) \sum_{(a,f)=1} K_{d,f}(a) = \mathbf{e}_{f_1}(d)\mathbf{e}_{f_2}(d) \sum_{\substack{(a_1,f_1)=1 \\ (a_1+d,f_1)=1}} K_{d,f_1}(a_1) \sum_{\substack{(a_2,f_2)=1 \\ (a_2+d,f_2)=1}} K_{d,f_2}(a_2).$$

In other words, upon extending $c(q)$ (and similarly c_1, c_2 and c_{12}) to a multiplicative function on $\mathbb{N}$, we obtain the desired count. $\qquad\square$

4.3. Computing the Correction Factor. For $g \in \mathcal{H}_1$, define

$$(4.20)$$
$$\mathbf{c}_g(d) := 2 \times \begin{cases} 2 \cdot \Omega_{d,\tilde{g}}(-1,-1) & \text{if } g_0 \equiv 1(4) \\ \Omega_{d,\tilde{g}}(-1,-1) + \Omega_{d,\tilde{g}}(1,1) & \text{if } g_0 \equiv 3(4) \text{ and } d \equiv 0(4) \\ \Omega_{d,\tilde{g}}(-1,1) + \Omega_{d,\tilde{g}}(1,-1) & \text{if } g_0 \equiv 3(4) \text{ and } d \equiv 2(4) \end{cases}$$

when g_0 is odd, and
$$(4.21)$$
$$\mathbf{c}_g(d) := \begin{cases} 2\left(\Omega_{d,\tilde{g}}(-1,-1) + \Omega_{d,\tilde{g}}(1,1)\right) & \text{if } d \equiv 0(8) \\ 2\left(\Omega_{d,\tilde{g}}(-1,1) + \Omega_{d,\tilde{g}}(1,-1)\right) & \text{if } d \equiv 4(8) \\ \Omega_{d,\tilde{g}}(-1,1) + \Omega_{d,\tilde{g}}(1,1) + \Omega_{d,\tilde{g}}(1,-1) + \Omega_{d,\tilde{g}}(-1,-1) & \text{if } d \equiv 2,6(8) \end{cases}$$

when g_0 is even.

Conjecture 1.2 now follows from Conjecture 2.1, together with the following Corollary:

Corollary 4.4. *Fix $d \in 2\mathbb{N}$ and $g \in \mathcal{H}_1$. If Conjecture 3.1 holds, then*

$$W_{d,g}(x) = \mathbf{A}(d)\mathbf{c}_g(d)\mathfrak{S}(d)\mathrm{Li}_2(x) + O\left(\frac{x}{(\log x)^C}\right),$$

Proof. By Lemma 4.1 and Corollary 3.4, we find that

$$W_{d,g}(x) = 4c_{d,g} \cdot \mathbf{A}(d)\mathfrak{S}(d)\mathrm{Li}_2(x) + O\left(\frac{x}{(\log x)^C}\right),$$

where

$$c_{d,g} := \begin{cases} \Omega_1^2(-1,-1) & \text{if } g_0 \equiv 1(4) \\ \sum_{t=1,3} \Omega_t^4(T_1(t), T_1(t+d)) & \text{if } g_0 \equiv 3(4) \\ \sum_{t=1,3,5,7} \Omega_t^8(T_2(t), T_2(t+d)) & \text{if } g_0 \equiv 2(8) \\ \sum_{t=1,3,5,7} \Omega_t^8(T_3(t), T_3(t+d)) & \text{if } g_0 \equiv 6(8). \end{cases}$$

The desired result now follows from Lemma 4.2. $\qquad\square$

5. Determining Cases for Which $\mathbf{c}_g(d) = 0$

5.1. Computing $|\Phi_{\varepsilon_1,\varepsilon_2}(\tilde{g}, d)|$.

Lemma 5.1. *Fix $\tilde{g} \in \mathbb{N}_{>1}$ to be odd and square-free, and write $\tilde{g} = q_1 q_2 \ldots q_r$, where the q_i's are pairwise distinct prime numbers. Then*

$$|\Phi_{\varepsilon_1,\varepsilon_2}(\tilde{g}, d)| = \frac{1}{4} \times \begin{cases} |\varphi_d(\tilde{g})| + (-1)^r \left(\left(\frac{-d}{\tilde{g}}\right)\varepsilon_1 + \left(\frac{d}{\tilde{g}}\right)\varepsilon_2 + \varepsilon_1\varepsilon_2 \right) & \text{if } (\tilde{g}, d) = 1 \\ \\ |\varphi_d(\tilde{g})| + (-1)^{r-s}\varepsilon_1\varepsilon_2 \prod_{i=1}^s (q_i - 1) & \text{if } (\tilde{g}, d) \neq 1, \end{cases}$$

where $(\tilde{g}, d) = q_1 \ldots q_s$.

Proof. Note that

$$4 \cdot |\Phi_{\varepsilon_1,\varepsilon_2}(\tilde{g}, d)| = \sum_{\substack{1 \leq \alpha \leq \tilde{g} \\ (\alpha,\tilde{g})=1 \\ (\alpha+d,\tilde{g})=1}} \left(1 + \varepsilon_1\left(\frac{\alpha}{\tilde{g}}\right) \right) \left(1 + \varepsilon_2\left(\frac{\alpha+d}{\tilde{g}}\right) \right)$$

(5.1)

$$= |\varphi_d(\tilde{g})| + \varepsilon_1 \sum_{\substack{1 \leq \alpha \leq \tilde{g} \\ (\alpha+d,\tilde{g})=1}} \left(\frac{\alpha}{\tilde{g}}\right) + \varepsilon_2 \sum_{\substack{1 \leq \alpha \leq \tilde{g} \\ (\alpha,\tilde{g})=1}} \left(\frac{\alpha+d}{\tilde{g}}\right) + \varepsilon_1\varepsilon_2 \sum_{1 \leq \alpha \leq \tilde{g}} \left(\frac{\alpha}{\tilde{g}}\right)\left(\frac{\alpha+d}{\tilde{g}}\right).$$

By (4.18) and multiplicativity, it follows that

(5.2)
$$\sum_{\substack{1 \leq \alpha \leq \tilde{g} \\ (\alpha+d,\tilde{g})=1}} \left(\frac{\alpha}{\tilde{g}}\right) = \begin{cases} 0 & \text{if } (\tilde{g}, d) \neq 1, \\ (-1)^r\left(\frac{-d}{\tilde{g}}\right) & \text{if } (\tilde{g}, d) = 1, \end{cases}$$

and similarly that

(5.3)
$$\sum_{\substack{1 \leq \alpha \leq \tilde{g} \\ (\alpha,\tilde{g})=1}} \left(\frac{\alpha+d}{\tilde{g}}\right) = \begin{cases} 0 & \text{if } (\tilde{g}, d) \neq 1, \\ (-1)^r\left(\frac{d}{\tilde{g}}\right) & \text{if } (\tilde{g}, d) = 1. \end{cases}$$

By (4.19) and multiplicativity, we moreover obtain

(5.4)
$$\sum_{1 \leq \alpha \leq \tilde{g}} \left(\frac{\alpha}{\tilde{g}}\right)\left(\frac{\alpha+d}{\tilde{g}}\right) = (-1)^{r-s} \prod_{i=1}^s (q_i - 1).$$

Lemma 5.1 now follows upon inserting (5.2), (5.3), and (5.4), into (5.1). $\quad\square$

We are particularly interested in determining the cases for which $|\Phi_{\varepsilon_1,\varepsilon_2}(\tilde{g}, d)| = 0$.

Lemma 5.2. *Fix $d \in 2\mathbb{N}$ and $\tilde{g} \in \mathbb{N}_{>1}$ to be odd and square-free. Then*
- (1) $|\Phi_{-1,-1}(\tilde{g}, d)| = 0$ *if and only if either*
  - $\tilde{g} = 5$ *and* $d \equiv 2, 3 \pmod 5$, *or*
  - $\tilde{g} \nmid d$ *and* $\tilde{g} | 3d$
- (2) $|\Phi_{1,1}(\tilde{g}, d)| = 0$ *if and only if*
  - $\tilde{g} = 5$ *and* $d \equiv 1, 4 \pmod 5$, *or*
  - $\tilde{g} \nmid d$ *and* $\tilde{g} | 3d$,
- (3) $|\Phi_{-1,1}(\tilde{g}, d)| = 0$ *if and only if* $\tilde{g} | d$,
- (4) $|\Phi_{1,-1}(\tilde{g}, d)| = 0$ *if and only if* $\tilde{g} | d$.

Proof. As above, write $\tilde{g} = q_1 \ldots q_r$, where q_i are pairwise distinct prime numbers, and let $(\tilde{g}, d) = q_1 \ldots q_s$.

1) If $(\tilde{g}, d) = 1$, then by Lemma 5.1,

$$|\Phi_{-1,-1}(\tilde{g}, d)| = 0 \Leftrightarrow |\varphi_d(\tilde{g})| + (-1)^r \left(1 - \left(\frac{-d}{\tilde{g}}\right) - \left(\frac{d}{\tilde{g}}\right)\right) = 0,$$

which occurs if and only if, either: *i*): $\tilde{g} = 5$ and $d \equiv 2, 3 \pmod 5$, or *ii*): $\tilde{g} = 3$ and $d \equiv 1, 2 \pmod 3$. If $(\tilde{g}, d) > 1$, then

$$|\Phi_{-1,-1}(\tilde{g}, d)| = 0 \Leftrightarrow |\varphi_d(\tilde{g})| + (-1)^{r-s} \prod_{i=1}^{s}(q_i - 1) = 0,$$

which occurs if only if $r = s + 1$ and $q_r = 3$. Since $\tilde{g}$ is square-free, this is equivalent to $\tilde{g} | 3d \wedge \tilde{g} \nmid d$.

2) If $(\tilde{g}, d) = 1$, then

$$|\Phi_{1,1}(\tilde{g}, d)| = 0 \Leftrightarrow |\varphi_d(\tilde{g})| + (-1)^r \left(1 + \left(\frac{-d}{\tilde{g}}\right) + \left(\frac{d}{\tilde{g}}\right)\right) = 0,$$

which occurs if and only if either: *i*): $\tilde{g} = 3$ and $3 \nmid d$, or *ii*): $\tilde{g} = 5$ and $d \equiv 1, 4 \pmod 5$. If $(\tilde{g}, d) > 1$, then

$$|\Phi_{1,1}(\tilde{g}, d)| = 0 \Leftrightarrow |\varphi_d(\tilde{g})| + (-1)^{r-s} \prod_{i=1}^{s}(q_i - 1) = 0,$$

which, as before, occurs if and only if $r = s + 1$ and $q_r = 3$.

3) and 4) If $(\tilde{g}, d) = 1$, then

$$|\Phi_{-1,1}(\tilde{g}, d)| = 0 \Leftrightarrow |\varphi_d(\tilde{g})| + (-1)^r \left(1 - \left(\frac{-d}{\tilde{g}}\right) + \left(\frac{d}{\tilde{g}}\right)\right) = 0,$$

which cannot occur. Similarly, we find that $|\Phi_{1,-1}(\tilde{g}, d)| \neq 0$.

If $(\tilde{g}, d) > 1$, then

$$|\Phi_{-1,1}(\tilde{g}, d)| = 0 \Leftrightarrow |\varphi_d(\tilde{g})| - (-1)^{r-s} \prod_{i=1}^{s} (q_i - 1) = 0,$$

which occurs when $r = s$, i.e. when $\tilde{g}|d$. Similarly with $\Phi_{1,-1}(\tilde{g}, d)$.   □

5.2. Proof of Theorem 1.1.

Proof. We wish to identify the instances in which $\mathbf{c}_g(d) = 0$. When $g \equiv 1(4)$,

$$\mathbf{c}_g(d) = \Omega_{d,\tilde{g}}(-1, -1)$$
$$= \mathbf{e}_{\tilde{g}}(d) \sum_{a \in \Phi_{-1,-1}(\tilde{g}, d)} K_{d,\tilde{g}}(a),$$

and therefore since $\mathbf{e}_{\tilde{g}}(d), K_{d,\tilde{g}}(a) > 0$,

$$\mathbf{c}_g(d) = 0 \Leftrightarrow |\Phi_{-1,-1}(\tilde{g}, d)| = 0.$$

Note that by Lemma 4.1, $\Delta_{d,g} : \pi_d \to \{0, 1\}$ is supported on the set

$$\{p \in \pi_d : \Delta_{d,g}(p) \neq 0\} = \{p \in \pi_d : (p, g_s) = (p + d, g_s) = 1 \wedge \delta_{-1,-1}(\tilde{g}, d)(p) = 1\}$$
$$= \bigcup_{a \in \Phi_{-1,-1}(\tilde{g}, d)} \{(p, g_s) = (p + d, g_s) = 1 \wedge p \equiv a \pmod{\tilde{g}}\}.$$

In other words, $\mathbf{c}_g(d) = 0$ implies that $\Delta_{d,g}$ is supported on the empty set. Since $p \in \pi_{d,g} \Rightarrow \Delta_{d,g}(p) = 1$, it further follows that $\pi_{d,g}$ is empty whenever $\mathbf{c}_g(d) = 0$. The other cases follow similarly, and we conclude that $\mathbf{c}_g(d) = 0$ if and only if one of the following cases holds:

- $g_0 \equiv 1 \pmod{4}$ and $|\Phi_{-1,-1}(\tilde{g}, d)| = 0$,
- $g_0 \equiv 3 \pmod{4}$,
  - $d \equiv 0 \pmod{4}$ and $|\Phi_{\pm 1, \pm 1}(\tilde{g}, d)| = 0$,
  - $d \equiv 2 \pmod{4}$ and $|\Phi_{\pm 1, \mp 1}(\tilde{g}, d)| = 0$,
- $g_0 \equiv 2 \pmod{4}$,
  - $d \equiv 0 \pmod{8}$ and $|\Phi_{\pm 1, \pm 1}(\tilde{g}, d)| = 0$,
  - $d \equiv 2, 6 \pmod{8}$ and $|\Phi_{\pm 1, \pm 1}(\tilde{g}, d)| = |\Phi_{\pm 1, \mp 1}(\tilde{g}, d)| = 0$,
  - $d \equiv 4 \pmod{8}$ and $|\Phi_{\pm 1, \mp 1}(\tilde{g}, d)| = 0$.

Theorem 1.1 now follows from Lemma 5.2.   □

Appendix A. Obtaining Numerical Evidence for Conjecture 1.2

The following Mathematica code is used to generate the set of primes $p \in \pi_{d,g}$ up to height 10^9, for the particular case $d = 2$ and $g = 10$. To begin, we generate a list of primes $p \leq 10^9$ ('primes'), and then selecting those for which $p - d$ is also prime ('primePairs'). Next, we select the subset of primes for which g is a primitive root modulo both p and $p - d$ ('primePairsArtin'). The list is subsequently exported into a file.

```
d = 2;
g = 10;
primes = Table[Prime[n], {n, PrimePi[10^9]}];
primePairs = Select[primes, PrimeQ[# - d] &];
primePairsArtin = Select[primePairs,
  MultiplicativeOrder[g, #] == # - 1 &&
    MultiplicativeOrder[g, # - d] == # - d - 1 &];
Export["primePairsArtin", primePairsArtin, "List"]
```

To generate the plot found in Figure 1, we first divide the primes in the above list into bins of size 10^5 ('primePairsBins'), which we then use to count primes up to size $k \cdot 10^5$, for each $k = 1, 2, \ldots$ (('primePairsCounts')). Next, we plot pairs of the form $(k \cdot 10^5, \text{primePairsCounts[[k]]})$, where primePairsCounts[[k]] denotes the number of primes $p \in \pi_{d,g}$ up to height $k \cdot 10^5$. The plot is then graphed alongside $\text{const} \cdot \text{Li}_2(x)$, where $\text{const} = 0.167606 \approx \mathbf{c}_{10}(2)\mathbf{A}(2)\mathfrak{S}(2)$ has been manually inserted.

```
primePairsArtin = Import["primePairsArtin", "List"];
xMax = 10^9;
step = 10^5;
primePairsBins = BinCounts[primePairsArtin, step];
primePairsCounts = Accumulate[primePairsBins];
kMax = Length[primePairsCounts];
primePairsPlot = Table[{k*step, primePairsCounts[[k]]}, {k, kMax}]
plotPrimes = ListPlot[primePairsPlot, PlotStyle -> Blue];
LogInt2[x_] := NIntegrate[1/Log[s]^2, {s, 2, x}] ;
const = 0.167606;
plotEstimate =
  Plot[const*LogInt2[x], {x, 2, xMax},
  PlotStyle -> Red,
  PlotRange -> {{0, xMax}, Automatic}, AxesLabel -> {"x"},
  PlotLegends -> LineLegend[{Blue, Red},
    {"pi_{2, 10}(x)", "C(2, 10)Li_2(x)"}]];
P = Show[plotPrimes, plotEstimate]
```

All the computations were performed on a Windows 10 laptop with an Intel i5-7200U processor and 8 GB RAM.

REFERENCES

[1] S. L. Aletheia-Zomlefer, L. Fukshansky and S. R. Garcia, *The Bateman-Horn conjecture: heuristics, history and applications*, Expos. Math Vol. 38, Issue 4, (Dec. 2020) 430−479.
[2] T. M. Apostol, *Introduction to analytic number theory*, Undergraduate Texts in Mathematics, Springer-Verlag, New York-Heidelberg, 1976.
[3] S. Baier. *On the Bateman–Horn Conjecture*, J. Number Theory **96** (2002), 432–448.
[4] R. Baker, P. Pollack, *Bounded gaps between primes with a given primitive root, II*, Forum Math. **28** no. 4 (2016), 675–687

[5] P. T. Bateman and R. A. Horn, *A heuristic asymptotic formula concerning the distribution of prime numbers*, Math. Comp. **16** (1962), 363–367.

[6] E. Ciolan, *Probability Models for Counting Problems Related to Primitive Roots* [unpublished Masters thesis], University of Bonn, 2017.

[7] C. Cobeli and A. Zaharescu. *On the distribution of primitive roots mod p*, Acta Arith. **83** (1998), 143−153.

[8] H. Davenport, On primitive roots in finite fields, Quart. J. Math. Oxford **8** (1937), 308−312.

[9] C. Frei, P. Koymans and E. Sofos, *Vinogradov's three primes theorem with primes having given primitive roots*. Mathematical Proceedings of the Cambridge Philosophical Society, Vol. 170, Issue 1, Jan. 2021 , 75−110. https://doi.org/10.1017/S0305004119000331

[10] G. H. Hardy and J. E. Littlewood, *Some problems of 'Partitio Numerorum' III: On the expression of a number as a sum of primes*, Acta Math. **44** (1922), 1−70.

[11] D. R. Heath-Brown, *Artin's conjecture for primitive roots*, Quart. J. Math. Oxford **37** (1986), 27−38.

[12] C. Hooley, *Artin's conjecture for primitive roots*, J.Reine Angew. Math. **225** (1967), 209−220.

[13] E. Jensen and M.R. Murty, *Artin's conjecture for polynomials over finite fields*, Number theory, Trends Math., Birkhäuser, Basel (2000), 167–181.

[14] H. W. Lenstra, Jr., *On Artin's conjecture and Euclid's algorithm in global fields*, Invent. Math. **42** (1977), 202−224.

[15] H. W. Lenstra, Jr., P. Stevenhagen and P. Moree, *Character sums for primitive root densities.* Math. Proc. Cambridge Philos. Soc. **157** (2014), 489–511.

[16] P. Moree, *On primes in arithmetic progression having a prescribed primitive root*, J. Number Theory **78** (1999), 85−98.

[17] P. Moree, *Primes in arithmetic progression having a prescribed primitive root. II*, Funct. Approx. Comment. Math. **39** (2008), 133−144.

[18] P. Moree, *Asymptotically exact heuristics for prime divisors of the sequence $\{a^k + b^k\}_{k=1}^{\infty}$*, J. Integer Seq. **9**, no. 2, Article 06.2.8 (2006).

[19] P. Moree, *Asymptotically exact heuristics for (near) primitive roots*, J. Number Theory **83** (2000), 155–181.

[20] P. Moree, *Asymptotically exact heuristics for (near) primitive roots. II*, Japan. J. Math. (N.S.) **29** (2003), 143–157.

[21] P. Pollack, *Bounded gaps between primes with a given primitive root*, Algebra Number Theory **8** (2014), 1769–1786.

[22] F. Pappalardi and I. Shparlinski, *On Artin's conjecture over function fields*, Finite Fields Appl. 1 (1995), 399−404.

[23] Z. Rudnick and A. Zaharescu, *The distribution of spacings between small powers of a primitive root*, Israel J. Math. **120** (2000), 271−287.

[24] P. J. Stephens, *An average result for Artin's conjecture*, Mathematika **16** (1969), 178−188.

[25] P. Stevenhagen *The correction factor in Artin's primitive root conjecture* J. Théor. des Nombres Bordeaux, **15** (2003), 383−391.

[26] M. Zoeteman, *A Siegel-Walfisz theorem for Artin primes and applications* [unpublished bachelor's thesis], University of Leiden, 2017.

Charles University, Faculty of Mathematics and Physics, Department of Algebra, Sokolovská 83, 18600 Praha 8, Czech Republic

Email address: tinkova.magdalena@gmail.com

Technische Universität Dresden, Fakultät Mathematik, Institut für Algebra, Zellescher Weg 12-14, 01062 Dresden, Germany
  Email address: ezrawaxman@gmail.com

Charles University, Faculty of Mathematics and Physics, Department of Mathematical Analysis, Sokolovská 83, 18600 Praha 8, Czech Republic
  Email address: zinmik2@gmail.com