ODD MOMENTS IN THE DISTRIBUTION OF PRIMES

VIVIAN KUPERBERG

Abstract. Montgomery and Soundararajan showed that the distribution of \( \psi(x + H) - \psi(x) \), for \( 0 \leq x \leq N \), is approximately normal with mean \( \sim H \) and variance \( \sim H \log(N/H) \), when \( N^\delta \leq H \leq N^{1-\delta} \). Their work depends on showing that sums \( R_k(h) \) of \( k \)-term singular series are \( \mu_k(-h \log h + Ah)^{k/2} + O_k(h^{k/2-1/(7k)+\epsilon}) \), where \( A \) is a constant and \( \mu_k \) are the Gaussian moment constants. We study lower-order terms in the size of these moments. We conjecture that when \( k \) is odd, \( R_k(h) \sim h^{(k-1)/2(\log h)(k+1)/2} \). We prove an upper bound with the correct power of \( h \) when \( k = 3 \), and prove analogous upper bounds in the function field setting when \( k = 3 \) and \( k = 5 \). We provide further evidence for this conjecture in the form of numerical computations.

1. Introduction

What is the distribution of primes in short intervals? Cramér [2] modeled the indicator function of the sequence of primes by independent random variables \( X_n \), for \( n \geq 3 \), which are 1 ("\( n \) is prime") with probability \( \frac{1}{\log n} \), and 0 ("\( n \) is composite") with probability \( 1 - \frac{1}{\log n} \). Cramér’s model predicts that the distribution of \( \psi(n + h) - \psi(n) \), a weighted count of the number of primes in an interval of size \( h \) starting at \( n \), follows a Poisson distribution when \( n \) varies in \([1, N]\) and when \( h \sim \log N \). Gallagher [5] proved that this follows from a quantitative version of the Hardy-Littlewood prime \( k \)-tuple conjecture: namely, that if \( \mathcal{D} = \{d_1, d_2, \ldots, d_k\} \) is a set of \( k \) distinct integers, then

\[
\sum_{n \leq N} \prod_{i=1}^{k} \Lambda(n + d_i) = (\mathcal{G}(\mathcal{D}) + o(1))N,
\]

where \( \mathcal{G}(\mathcal{D}) \) is the singular series, a constant dependent on \( \mathcal{D} \) given by

\[
\mathcal{G}(\mathcal{D}) = \prod_{p} \left( 1 - \frac{1}{p} \right)^{-k} \left( 1 - \frac{\nu_p(\mathcal{D})}{p} \right),
\]

where \( \nu_p(\mathcal{D}) \) denotes the number of distinct residue classes modulo \( p \) among the elements of \( \mathcal{D} \). The singular series is also given by the formula

\[
(1) \quad \mathcal{G}(\mathcal{D}) = \sum_{q_1, \ldots, q_k} \left( \prod_{i=1}^{k} \frac{\mu(q_i)}{\phi(q_i)} \right) \sum_{a_1, \ldots, a_k \in \mathbb{Z}} e \left( \sum_{i=1}^{k} \frac{a_i d_i}{q_i} \right).
\]

The author is supported by NSF GRFP grant DGE-1656518, and would like to thank Kannan Soundararajan for many helpful comments and discussions, as well as Régis de la Bretèche, Alexandra Florea, Zeev Rudnick, and Yuval Wigderson for helpful feedback.
The Hardy-Littlewood prime $k$-tuple conjectures give us a better lens through which to understand the distribution of primes: by understanding sums of singular series. For example, Gallagher used the estimate that

$$\sum_{D \in [1, H]} \mathcal{S}(D) \sim \sum_{D \in [1, H]} 1$$

to prove that the Hardy-Littlewood conjectures imply Poisson behavior in intervals of logarithmic length. Our concern is the distribution of primes in somewhat longer intervals; namely, those of size $H$ where $H = o(N)$ and $H/\log N \to \infty$ as $N \to \infty$. In this setting, the Cramér model would predict that the distribution of $\psi(n + H) - \psi(n)$ for $n \leq N$ is approximately normal, with mean $\sim H$ and variance $\sim H \log N$. However, the Hardy-Littlewood prime $k$-tuple conjecture gives a different answer in this case. In [12], Montgomery and Soundararajan provide evidence based on the Hardy-Littlewood prime $k$-tuple conjectures that the distribution ought to be approximately normal with variance $\sim H \log N$. They consider the $K$th moment $M_K(N; H)$ of the distribution of primes in an interval of size $H$,

$$M_K(N; H) = \sum_{n=1}^{N} (\psi(n + H) - \psi(n) - H)^K.$$

They conjecture that these moments should be given by the Gaussian moments

$$M_K(N; H) = (\mu_K + o(1))N \left( H \log \frac{N}{H} \right)^{K/2},$$

where $\mu_K = 1 \cdot 3 \cdots (K - 1)$ if $K$ is even and 0 if $K$ is odd, uniformly for $(\log N)^{1+\delta} \leq H \leq N^{1-\delta}$. Their technique relies on more refined estimates of sums of the singular series constants $\mathcal{S}(D)$. Instead of the von Mangoldt function $\Lambda(n)$, they consider sums of $\Lambda_0(n) = \Lambda(n) - 1$, where the main term has been subtracted from the beginning. The corresponding form of the Hardy-Littlewood conjecture states that

$$\sum_{n \leq N} \prod_{i=1}^{k} \Lambda_0(n + d_i) = (\mathcal{S}_0(D) + o(1))N$$

as $N \to \infty$, where $\mathcal{S}_0(D)$ is given by

$$\mathcal{S}_0(D) = \sum_{\mathcal{J} \subseteq D} (-1)^{|D \setminus \mathcal{J}|} \mathcal{S}(\mathcal{J}),$$

and in turn

$$\mathcal{S}(D) = \sum_{\mathcal{J} \subseteq D} \mathcal{S}_0(\mathcal{J}).$$

We can combine this with Equation 1 to see that

$$\mathcal{S}_0(D) = \sum_{q_1, \ldots, q_k \leq \sqrt{q}} \left( \prod_{i=1}^{k} \frac{\mu(q_i)}{\phi(q_i)} \right) \sum_{\substack{a_1, \ldots, a_k \leq q_i \leq q \atop (a_i, q_i) = 1}} e \left( \sum_{i=1}^{k} \frac{a_i d_i}{q_i} \right).$$
Montgomery and Soundararajan considered the sum
\[ R_k(h) := \sum_{d_1, \ldots, d_k \leq h \atop d_i \text{ distinct}} \mathcal{S}_0(D), \]
showing that for any nonnegative integer \( k \), for any \( h > 1 \), and for any \( \varepsilon > 0 \),
\[ R_k(h) = \mu_k(-h \log h + A h)^{k/2} + O_k(h^{k/2 - 1/(7k) + \varepsilon}), \]
where \( A = 2 - \gamma - \log 2\pi \). Their estimate on \( R_k(h) \) implies their bound on the moments. For more on the distribution of primes in short intervals, see for example [1] and [6], as well as [12].

For all \( k \), the optimal error term in (4) is expected to be smaller. In the case of the variance, this was studied in [11]. In this paper, we restrict our attention to the cases when \( k \) is odd. We conjecture the following, which was mentioned by Lemke Oliver and Soundararajan in [10].

**Conjecture 1.1.** Let \( k \geq 3 \) be an odd integer, and let \( h > 1 \). With \( R_k(h) \) defined as above,
\[ R_k(h) \approx h^{(k-1)/2}(\log h)^{(k+1)/2}. \]

The conjectured power of \( \log h \) here comes from numerical evidence, which we present in Section 5. For \( k \) odd, we do not know, even heuristically, which terms contribute to the main term in \( R_k(h) \); for this reason, we do not know what the constant should be in front of the asymptotic in Conjecture 1.1. Nevertheless, our goal in this paper is to provide evidence for Conjecture 1.1. When \( k = 3 \), we can show an upper bound with the correct power of \( h \).

**Theorem 1.2.** For \( h \geq 2 \) and \( R_3 \) defined in (3),
\[ R_3(h) \ll h(\log h)^{4}(\log \log h)^{2}. \]

Another source of evidence for Conjecture 1.1 is the analog of this problem in the function field setting, which is also studied in [9]. As we discuss in Section 3, we can consider analogous questions over \( \mathbb{F}[T] \) where \( \mathbb{F} \) is a finite field, instead of over \( \mathbb{Z} \). To state the analog, we first revisit the techniques of Montgomery and Soundararajan in the integer case. Upon expanding Equation 3 using Equation 2, we get
\[ R_k(h) = \sum_{d_1, \ldots, d_k \leq h \atop d_i \text{ distinct}} \sum_{q_1, \ldots, q_k \leq h \atop 1 \leq q_i < \infty} \left( \prod_{i=1}^{k} \frac{\mu(q_i)}{\phi(q_i)} \right) \sum_{a_1, \ldots, a_k \leq q_i \atop (a_i, q_i) = 1 \atop \sum a_i/q_i \in \mathbb{Z}} e \left( \sum_{i=1}^{k} \frac{a_i d_i}{q_i} \right) \]
\[ = \sum_{q_1, \ldots, q_k \leq h \atop 1 \leq q_i < \infty} \left( \prod_{i=1}^{k} \frac{\mu(q_i)}{\phi(q_i)} \right) \sum_{a_1, \ldots, a_k \leq q_i \atop (a_i, q_i) = 1 \atop \sum a_i/q_i \in \mathbb{Z}} E \left( \frac{a_i}{q_i} \right), \]
where \( E(\alpha) = \sum_{m=1}^{h} e(m \alpha) \). The sums \( E(\alpha) \) approximately detect when \( \|\alpha\| \leq \frac{1}{h} \).

This expression for \( R_k(h) \) is closely related to a quantity studied by Montgomery and Vaughan in [13]. They considered the related problem of the \( k \)th moment of reduced residues.
modulo a fixed $q$, given by

$$m_k(q; h) = \sum_{n=1}^q \left( \sum_{1 \leq m \leq h \atop (m+n,q)=1} 1 - h \frac{\phi(q)}{q} \right)^k.$$ 

The moment $m_k$ satisfies $m_k(q; h) = q \left( \frac{\phi(q)}{q} \right)^k V_k(q; h)$, where $V_k(q; h)$ is the “singular series sum,”

$$V_k(q; h) = \sum_{d_1, \ldots, d_k} \sum_{1 \leq d_i \leq h \atop 1 \leq q_i | q} \left( \prod_{i=1}^k \mu(q_i) \right) \sum_{a_1, \ldots, a_k \atop 1 \leq a_i \leq q_i \atop (a_i,q_i) = 1} e \left( \sum_{i=1}^k a_i d_i q^{-i} \right),$$

which differs from $R_k(h)$ only in that the $q_i$ are now constrained to divide a fixed $q$. In this paper as well as in the work of Montgomery and Soundararajan, estimating $V_k(q; h)$ is a key step towards estimating $R_k(h)$. Similarly, understanding $m_k(q; h)$ is closely related to understanding $R_k(h)$. In the function field setting, we study the moments $m_k(q; h)$.

Let $\mathbb{F}_q$ be a finite field with $q$ elements, and let $Q$ be a fixed monic polynomial in $\mathbb{F}_q[T]$. Note that $Q$ in the function field case serves the same role as $q$ in the integer case, since $q$ now represents the size of the field. For an appropriate function field analog of the $k$th moment of reduced residues in short intervals $m_k(Q; h)$, we can adapt the methods of Montgomery-Vaughan to prove a bound on $m_k(Q; h)$ that has the same shape as the bounds of Montgomery-Vaughan and Montgomery-Soundararajan.

**Theorem 1.3.** For any fixed $k \geq 3$ and for $Q \in \mathbb{F}_q[T]$ squarefree,

$$m_k(Q; h) \ll \begin{cases} |Q|(q^h)^{k/2} \left( \frac{\phi(Q)}{|Q|} \right)^{1/2} \left( 1 + (q^h)^{-1/(k-2)} \frac{\phi(Q)}{|Q|} \right)^{-2k+k/2} & \text{if } k \text{ is even} \\ |Q|((q^h)^{k/2-1/2} + (q^h)^{k/2-1/(k-2)}) \left( \frac{\phi(Q)}{|Q|} \right)^{-2k+k/2} & \text{if } k \text{ is odd} \end{cases}$$

The function field exponential sums are cleaner than their integer analogs, making this proof more streamlined than the proof of Montgomery-Vaughan. As a result, the bound is tighter; in fact, for $k = 3$, Theorem 1.3 already yields a bound where the exponent of $q^h$ is 1. This is of the same shape as Theorem 1.2, where the exponent of $h$ is 1.

Using a more involved argument we can achieve a bound on the fifth moment of reduced residues in short intervals.

**Theorem 1.4.** Let $Q = \prod_{|P| \leq q^{2h}} P$. For all $\varepsilon > 0$,

$$m_5(Q; h) \ll |Q| q^{2h+\varepsilon}.$$ 

The analog of Conjecture 1.1 would predict that $m_5(Q; h) \approx |Q| q^{2h} (\log q^h)^3$. Our techniques do not quite succeed in proving such a bound for any higher odd moments, as we note in Section 4. However, we do get as a corollary the following bound on sums of singular series in function fields.
Corollary 1.5. Let $Q = \prod_{P \text{ irreducible}} P$. Then
\[ R_3(q^h) \ll V_3(Q; h) + q^h \left( \frac{|Q|}{\phi(Q)} \right)^2 \ll q^h \left( \frac{|Q|}{\phi(Q)} \right)^{19/2}, \]
and for all $\varepsilon > 0$,
\[ R_5(q^h) \ll V_5(Q; h) + \left( \frac{|Q|}{\phi(Q)} \right)^{21/2} q^{2h} \ll q^{2h+\varepsilon}. \]

The rest of this paper is organized as follows. In Section 2 we prove Theorem 1.2. In Section 3, we discuss the analogous problem in $\mathbb{F}_q[T]$, and adapt the framework of Montgomery and Vaughan to the function field setting to prove Theorem 1.3. In Section 4 we prove Theorem 1.4. Finally, in Section 5 we provide numerical evidence for Conjecture 1.1, and in Section 6 we discuss toy problems, further directions of inquiry, and possible applications of these questions.

2. THREE-TERM INTEGER SUMS: PROOF OF THEOREM 1.2

Our goal is to bound
\[ R_3(h) = \sum_{d_1, d_2, d_3 \leq h, \, d_i \text{ distinct}} \mathcal{S}_0(D). \]
Expanding $\mathcal{S}_0(D)$ as an exponential sum yields
\[ R_3(h) = \sum_{d_1, d_2, d_3 \leq h, \, d_i \text{ distinct}} \sum_{q_1, q_2, q_3 \leq q, \, q_i < \infty} \left( \prod_{i=1}^3 \frac{\mu(q_i)}{\phi(q_i)} \right) \sum_{a_1, a_2, a_3 \leq q, \, (a_i, q_i) = 1} e \left( \sum_{i=1}^3 a_i d_i \right). \]

Our argument will follow the same thread as that of Montgomery and Soundararajan [12], which in turn relies on the analysis of Montgomery and Vaughan [13] of the distribution of reduced residues. To that end, we consider $V_3(q; h)$, which is approximately the third centered moment of the number of reduced residues mod $q$ in an interval of length $h$. Precisely, $V_3(q; h)$ is given by
\[ V_3(q; h) = \sum_{d_1, \ldots, d_3 \leq h, \, 1 < q_i \leq q} \left( \prod_{i=1}^3 \frac{\mu(q_i)}{\phi(q_i)} \right) \sum_{a_1, a_2, a_3 \leq q, \, (a_i, q_i) = 1} e \left( \sum_{i=1}^3 a_i d_i \right) \]

This is very similar to the above expression for $R_3(h)$; the two differences are that the outer sum in $R_3(h)$ is taken over distinct $d_i$’s, whereas the outer sum for $V_3(q; h)$ is not, and that the summands $q_i$ range over all integers for $R_3(h)$, but are restricted to factors of $q$ for $V_3(q; h)$.

Theorem 2.1. Let $q$ be the product of primes up to $h^4$. Then
\[ V_3(q; h) \ll h (\log h)^4 (\log \log h)^2. \]
We then use Theorem 2.1 to establish Theorem 1.2. In order to derive Theorem 1.2, it suffices to show that terms arising from transforming $V_3(q; h)$ into $R_3(h)$ do not contribute more than $O(h(\log h)^4(\log \log h)^2)$; in fact they contribute on the order of $h(\log h)^2$, which is the conjectured asymptotic size of $R_3(h)$. We begin with this derivation of Theorem 1.2 from Theorem 2.1

In order to account for terms where $d_1, d_2, d_3$ are not necessarily distinct, we make the following definition.

**Definition 2.2.** Let $k \geq 2$, and let $D = \{d_1, \ldots, d_k\}$ be a $k$-tuple of not necessarily distinct integers, and fix $q$ a squarefree integer. Then the singular series at $D$ with respect to $q$ is given by

$$
\mathcal{S}(D; q) := \sum_{q_1, \ldots, q_k|q} \left( \prod_{i=1}^{k} \frac{\mu(q_i)}{\phi(q_i)} \right) \sum_{a_1, \ldots, a_k \leq q_i, (a_i, q_i) = 1} e \left( \sum_{i=1}^{k} \frac{a_i d_i}{q_i} \right).
$$

Just as for $\mathcal{S}(D)$, one can subtract off the main term of $\mathcal{S}(D; q)$ to define

$$
\mathcal{S}_0(D; q) := \sum_{J \subset D} (-1)^{|D \setminus J|} \mathcal{S}(J; q).
$$

Combining this with the definition for $\mathcal{S}(D; q)$ yields the formula

$$
(5) \quad \mathcal{S}_0(D; q) = \sum_{1 < q_1, \ldots, q_k|q} \left( \prod_{i=1}^{k} \frac{\mu(q_i)}{\phi(q_i)} \right) \sum_{a_1, \ldots, a_k \leq q_i, (a_i, q_i) = 1} e \left( \sum_{i=1}^{k} \frac{a_i d_i}{q_i} \right).
$$

If the $d_i$ are not all distinct, this expression converges for any fixed $q$ but not in the $q \to \infty$ limit. The singular series at $D$ with respect to $q$ is equal to a finite Euler product.

**Lemma 2.3.** Let $k \geq 2$, and let $D = \{d_1, \ldots, d_k\}$ be a $k$-tuple of not necessarily distinct integers, and fix $q$ a squarefree integer. Then

$$
\mathcal{S}(D; q) = \prod_{p|q} \left( 1 - \frac{1}{p} \right)^{-k} \left( 1 - \frac{\nu_p(D)}{p} \right),
$$

where $\nu_p(D)$ is the number of distinct residue classes mod $p$ occupied by elements of $D$.

This lemma is proven in Lemma 3 of [12]; it is stated there for sets with distinct elements, but their proof holds in this setting as well. They note first that $\mathcal{S}(D; q)$ is multiplicative in $q$, so that it suffices to check the lemma for primes $p$. For a given prime $p$, they express the condition that $\sum_{i=1}^{k} \frac{a_i}{q_i} \in \mathbb{Z}$ in terms of additive characters mod $p$, and then rearrange to get the result.

Consider the following expression for $\mathcal{S}_0$, which is equation (45) in [12]. For all $y \geq h$,

$$
\mathcal{S}_0(D) = \sum_{q_1, q_2, q_3, y \geq 1} \prod_{i=1}^{3} \frac{\mu(q_i)}{\phi(q_i)} A(q_1, q_2, q_3; D) + O \left( \frac{(\log y)^2}{y} \right),
$$
where

\[ A(q_1, q_2, q_3; \mathcal{D}) = \sum_{a_1, a_2, a_3, 1 \leq a_i \leq q_i} e \left( \sum_{i=1}^{3} \frac{d_i a_i}{q_i} \right). \]

Apply this to \( R_3(h) \) with \( y = h^2 \) and \( q = \prod_{p \leq y} p \) to get

\[ R_3(h) = \sum_{q_1, q_2, q_3} \prod_{q_i > 1}^{3} \mu(q_i) \phi(q_i) S(q_1, q_2, q_3; h) + O(1), \]

where

\[ S(q_1, q_2, q_3; h) := \sum_{d_1, d_2, d_3, 1 \leq d_i \leq h} A(q_1, q_2, q_3; \{d_1, d_2, d_3\}) = \sum_{d_1, d_2, d_3, 1 \leq d_i \leq h} e \left( \sum_{i=1}^{3} \frac{d_i a_i}{q_i} \right). \]

If the condition that the \( d_i \) should be distinct were omitted, then the main term in \( R_3(h) \) would be exactly \( V_3(h) \). So, it suffices to remove this condition.

Put \( \delta_{i,j} = 1 \) if \( d_i = d_j \) and \( 0 \) otherwise, so that

\[ \prod_{1 \leq i < j \leq 3} (1 - \delta_{i,j}) = \begin{cases} 1 & \text{if the } d_i \text{ are distinct} \\ 0 & \text{otherwise}, \end{cases} \]

and

\[ S(q_1, q_2, q_3; h) = \sum_{d_1, d_2, d_3} \left( \prod_{1 \leq i < j \leq 3} (1 - \delta_{i,j}) \right) \sum_{d_1, d_2, d_3, 1 \leq d_i \leq h} e \left( \sum_{i=1}^{3} \frac{d_i a_i}{q_i} \right). \]

Expanding the product yields

\[ 1 - \delta_{1,2} - \delta_{1,3} - \delta_{2,3} + \delta_{1,2}\delta_{2,3} + \delta_{1,3}\delta_{1,2} + \delta_{2,3}\delta_{1,3} - \delta_{1,2}\delta_{2,3}\delta_{1,3}. \]

Note that the last four terms each require precisely that \( d_1 = d_2 = d_3 \) in order to be nonzero; each of these can be written as \( \delta_{1,2,3} \), so that their sum is \( 2\delta_{1,2,3} \). The following lemma addresses the contribution of these last four terms.

**Lemma 2.4.** We have

\[ 2 \sum_{d \leq h} \sum_{q_1, q_2, q_3 > 1} \prod_{q_i > 1}^{3} \frac{\mu(q_i)}{\phi(q_i)} \sum_{a_1, a_2, a_3, 1 \leq a_i \leq q_i} e \left( \sum_{i=1}^{3} \frac{d a_i}{q_i} \right) = 2h \left( \frac{q}{\phi(q)} \right)^2 - 6h \frac{q}{\phi(q)} + 4h. \]
Proof. Note that the left-hand expression is precisely \( 2 \sum_{d \leq h} \mathcal{S}_0(d, d; d; q) \). Expanding \( \mathcal{S}_0 \) and applying Lemma 2.3 yields
\[
2 \sum_{d \leq h} \mathcal{S}_0(d, d, d; q) = 2 \sum_{d \leq h} \left[ \mathcal{S}(d, d, d; q) - 3 \mathcal{S}(d, d; q) + 3 \mathcal{S}(d; q) - 1 \right]
= 2 \sum_{d \leq h} \left( \prod_{p | q} \left( 1 - \frac{1}{p} \right)^{-2} - 3 \prod_{p | q} \left( 1 - \frac{1}{p} \right)^{-1} + 2 \right)
= 2h \frac{q^2}{\phi(q)^2} - 6h \frac{q}{\phi(q)} + 4h,
\]
as desired. \( \square \)

Now consider the contribution to \( R_3(h) \) from the terms \(-\delta_{1,2}, -\delta_{1,3}, \) and \(-\delta_{2,3} \). Via relabeling, it suffices to only consider the term with \(-\delta_{1,2} \), which is nonzero when \( d_1 = d_2 \) and otherwise 0.

**Lemma 2.5.** We have
\[
\sum_{d, d_3 \leq h} \sum_{q_1, q_2, q_3} \prod_{i=1}^{3} \frac{\mu(q_i)}{\phi(q_i)} \sum_{\substack{a_1, a_2, a_3 \in \mathbb{Z} \setminus \{0\} \atop (a_i, q_i) = 1 \atop i = 1}} e \left( d \left( \frac{a_1}{q_1} + \frac{a_2}{q_2} \right) \right) e \left( \frac{d a_3}{q_3} \right)
= \left( \frac{q}{\phi(q)} - 2 \right) \left( h \frac{q}{\phi(q)} - h \log h + Bh + O(h^{1/2 + \varepsilon}) \right)
\]

Proof. As in the previous lemma, we note that the left-hand side is \( \sum_{d, d_3 \leq h} \mathcal{S}_0(d, d, d_3; q) \). We again expand and apply Lemma 2.3, to get
\[
\sum_{d, d_3 \leq h} \mathcal{S}_0(d, d, d_3; q) = \sum_{d, d_3 \leq h} \left[ \mathcal{S}(d, d, d_3; q) - 2 \mathcal{S}(d_3; q) - \mathcal{S}(d; q) + 2 \right]
= \left( \frac{q}{\phi(q)} - 2 \right) \left( \sum_{d, d_3 \leq h} \mathcal{S}(d, d_3; q) - h^2 \right).
\]

By Lemma 4 of [12],
\[
\sum_{d, d_3 \leq h} \mathcal{S}(d, d_3; q) = \sum_{q_1 | q} \mu(q_1)^2 \sum_{\substack{1 \leq a \leq q_1 \atop (a, q_1) = 1}} \left| E \left( \frac{a}{q_1} \right) \right|^2 = h \frac{q}{\phi(q)} + h^2 - h \log h + Bh + O(h^{1/2 + \varepsilon}),
\]
with \( B = 1 - \gamma - \log 2\pi \). Thus our expression becomes
\[
= \left( \frac{q}{\phi(q)} - 2 \right) \left( h \frac{q}{\phi(q)} - h \log h + Bh + O(h^{1/2 + \varepsilon}) \right),
\]
as desired. \( \square \)
Combining these computations yields

\[ R_3(h) = V_3(q; h) + 2h \left( \frac{q}{\phi(q)} \right)^2 - 6h \frac{q}{\phi(q)} + 4h \]

\[ - 3 \left( \frac{q}{\phi(q)} - 2 \right) \left( h \frac{q}{\phi(q)} - h \log h + Bh + O(h^{1/2+\varepsilon}) \right) \]

\[ = V_3(q; h) - h \left( \frac{q}{\phi(q)} \right)^2 + 3h \log h \frac{q}{\phi(q)} - 3Bh \frac{q}{\phi(q)} \]

\[ - 6h \log h + 6Bh + 4h + O \left( h^{1/2+\varepsilon} \frac{q}{\phi(q)} \right) \]

By Theorem 2.1, \( V_3(q; h) \ll h(\log h)^4(\log \log h)^2 \), so \( R_3(h) \ll h(\log h)^4(\log \log h)^2 \), which completes the proof of Theorem 1.2.

2.1. **Preparing for the proof of Theorem 2.1.** The rest of this section will be devoted to the proof of Theorem 2.1; here we begin by fixing some notation and proving several preparatory lemmas.

Throughout, we will parametrize by variables \( g, x, y, z \), where \( g = \gcd(q_1, q_2, q_3) \), \( x = \gcd(q_2/g, q_3/g) \), \( y = \gcd(q_1/g, q_3/g) \), and \( z = \gcd(q_1/g, q_2/g) \); note that \( q \) is squarefree by assumption, so each \( q_i \) is as well, and thus also \( g, x, y \), and \( z \). Thus \( q_1 = gyz, q_2 = gxz, \) and \( q_3 = gxy \), with \( g, x, y, z \) pairwise coprime.

Then

\[ V_3(q; h) = \sum_{g,x,y,z \mid q, \text{gcd}(g,xyz) = 1} \frac{\mu(g)\mu(gxyz)^2}{\phi(g)\phi(gxyz)^2} \sum_{a_1,a_2,a_3} E \left( \frac{a_1}{gyz} \right) E \left( \frac{a_2}{gxz} \right) E \left( \frac{a_3}{gxy} \right). \]

We discard the alternation from \( \mu(g) \) and use the bound that for all \( 0 \leq \alpha < 1, |E(\alpha)| \leq F(\alpha) \), where

\[ F(\alpha) := \min \{ h, \| \alpha \|^{-1} \}, \]

so that

\[ V_3(q; h) \ll \sum_{g,x,y,z \mid q, \text{gcd}(g,xyz) = 1} \frac{\mu(gxyz)^2}{\phi(g)\phi(gxyz)^2} \sum_{a_1,a_2,a_3} F \left( \frac{a_1}{gyz} \right) F \left( \frac{a_2}{gxz} \right) F \left( \frac{a_3}{gxy} \right). \]

We now split the sum \( V_3(q; h) \) into three different sums, addressed separately. Let \( T_1 \) consist of all terms \( g, x, y, z \) in (6) with \( gx \geq h \). Let \( T_2 \) consist of all terms \( g, x, y, z \) in (6) with \( gx < h, gy < h, \) and \( gz < h, \) and \( \left\| \frac{a_2}{q_2} \right\|, \left\| \frac{a_1}{q_1} \right\| > \frac{1}{h} \). In particular, for terms in \( T_2 \), at most one of the three fractions \( a_i/q_i \) is close to \( 0 \). Finally, let \( T_3 \) consist of all terms \( g, x, y, z \) in (6) with \( gx < h, gy < h, \) and \( gz < h \) as well as the constraints that \( \left\| \frac{a_1}{gyz} \right\| \leq \frac{2}{h}, \left\| \frac{a_2}{gxz} \right\| \leq \frac{2}{h}, \) and \( \left\| \frac{a_3}{gxy} \right\| \leq \frac{2}{h}. \)
Note that if two of the three fractions $\frac{a_i}{q_i}$ satisfy $\|\frac{a_i}{q_i}\| \leq \frac{1}{h}$ (say $i = 1, 2$), then the third one must satisfy $\|\frac{a_i}{q_i}\| \leq \frac{2}{h}$. Thus after permuting the names of variables as necessary, each term $g, x, y, z, a_1, a_2, a_3$ is contained in sums for $T_1, T_2, T_3$. In particular,

$$V_3(q; h) \ll T_1 + T_2 + T_3.$$ 

We will show in Lemmas 2.9, 2.10, and 2.11 respectively that $T_1 \ll h(\log h)^4$, that $T_2 \ll h(\log h)^2(\log \log h)^2$, and that $T_3 \ll h(\log h)^4(\log \log h)^2$, which completes the proof of Theorem 2.1.

We will make use of the following lemmas concerning sums of $F(\alpha)$. We write $\tilde{q} := \min\{q, h\}$, so that $F(a/q) \leq 2\|n(a, q)/\tilde{q}\|$.

**Lemma 2.6.** Fix an integer $\nu \in \mathbb{N}$ and let $\alpha_1 \geq \alpha_2 \geq 1$ be two positive real numbers, and let $h \in \mathbb{N}$. As $h \to \infty$, 

$$\sum_{1 \leq n_1 \leq h/(2\alpha_1)} \sum_{1 \leq n_2 \leq h/(2\alpha_2)} \sum_{1 \leq \nu + a_1 n_1 - a_2 n_2 \leq h/2} \frac{1}{\alpha_1 n_1 \alpha_2 n_2 \nu + \alpha_1 n_1 - \alpha_2 n_2} = O\left(\frac{1}{\alpha_1^2 \alpha_2}\right)$$

*Proof.** First rewrite the inner fraction via 

$$\frac{1}{\alpha_1 n_1 \alpha_2 n_2 \nu + \alpha_1 n_1 - \alpha_2 n_2} = \frac{1}{\alpha_1 n_1 (\nu + \alpha_1 n_1) \alpha_2 n_2} + \frac{1}{\alpha_1 n_1 (\nu + \alpha_1 n_1) (\nu + \alpha_1 n_1 - \alpha_2 n_2)}.$$ 

and then move the sum over $n_2$ inside to get

$$\sum_{1 \leq n_1 \leq h/(2\alpha_1)} \frac{1}{\alpha_1 n_1 (\nu + \alpha_1 n_1)} + \sum_{\nu + a_1 n_1 - \nu/2 \leq a_2 n_2 \leq \nu + a_1 n_1 - 1} \frac{1}{\alpha_2 n_2} + \frac{1}{\alpha_1 n_1 (\nu + \alpha_1 n_1 - \alpha_2 n_2)}.$$ 

Note that if $\nu + a_1 n_1 - 1 < \alpha_2$ or $\nu + a_1 n_1 - h/2 > h/2$, the inner sum is empty and thus the sum is 0. If not, then $2 \leq 1 + \alpha_2 \leq \nu + \alpha_1 n_1 \leq h$ and the inner sum is bounded by

$$\frac{2}{\alpha_2} \log(\frac{1}{\alpha_2} \min\{h/2, \nu + a_1 n_1 - 1\}) - \frac{2}{\alpha_2} \log(\frac{1}{\alpha_2} \max\{1, \nu + a_1 n_1 - h/2\}) + O(1/h).$$

Thus the expression is bounded by

$$\ll \frac{1}{\alpha_2 \alpha_1} \sum_{1 \leq n_1 \leq h/(2\alpha_1)} \log(a_1 n_1 + \nu - 1) + O(1/h) + \frac{1}{\alpha_2 \alpha_1} \sum_{1 \leq n_1 \leq h/(2\alpha_1)} \log(\frac{h/2}{a_1 n_1 + \nu - h/2}) + O(1/h).$$

For fixed $\nu$ and as $h \to \infty$, each of these sums is $O(1/\alpha_1)$, which gives the result. 

**Lemma 2.7.** Let $d_1$ and $d_2$ be positive integers with $d_1 < d_2 < h$. Then

$$\sum_{1 \leq n_1 < d_1} \sum_{1 \leq n_2 < d_2} F\left(\frac{n_1}{d_1}\right) F\left(\frac{n_2}{d_2}\right) F\left(\frac{n_1 - n_2}{d_1} \frac{d_2}{d_1} \right) \ll \begin{cases} \frac{h d_1^2}{d_2} & \text{if } \min(n_1, n_2) \frac{n_1}{d_1} - \frac{n_2}{d_2} < \frac{1}{h} \\ \frac{d_2^2}{d_1^2} & \text{otherwise.} \end{cases}$$

10
Proof. Since \( d_1 < h \), \( F(n_1/d_1) = \|n_1/d_1\|^{-1} \), and similarly \( F(n_2/d_2) = \|n_2/d_2\|^{-1} \).

As for \( F\left( \frac{n_1}{d_1} - \frac{n_2}{d_2} \right) \),
\[
\left\| \frac{n_1}{d_1} - \frac{n_2}{d_2} \right\| < \frac{1}{h} \Leftrightarrow \left\| \frac{n_1}{d_1} - \frac{n_2}{d_2} \right\| < \frac{1}{h} \\
\Leftrightarrow \left\| \frac{d_2}{d_1} n_1 - n_2 \right\| < \frac{d_2}{h}.
\]

Since \( \frac{d_2}{h} < 1 \), for each value of \( n_1 \), there are at most 2 values of \( n_2 \) satisfying this constraint: namely, \( n_2 \in \left\{ \left\lfloor \frac{d_2}{d_1} n_1 \right\rfloor, \left\lceil \frac{d_2}{d_1} n_1 \right\rceil \right\} \). We first tackle these terms, using the trivial bound \( F(x) \leq h \), and getting
\[
\sum_{1 \leq n_1 < d_1} h \left\| \frac{n_1}{d_1} \right\|^{-1} \left( \left\| \frac{d_2}{d_1} \frac{n_1}{d_1} \right\|^{-1} + \left\| \frac{d_2}{d_1} \frac{n_1}{d_1} \right\|^{-1} \right) \ll h \sum_{1 \leq n_1 < d_1} \left\| \frac{n_1}{d_1} \right\|^{-2} \\
\ll h \sum_{1 \leq n_1 < d_1/2} \frac{d_1^2}{n_1^2} \ll h d_1^2.
\]

The remaining terms are given by the sum
\[
\sum_{1 \leq n_1 < d_1} \left\| \frac{n_1}{d_1} \right\|^{-1} \left\| \frac{n_2}{d_2} \right\|^{-1} \left\| \frac{d_2 n_1 - d_1 n_2}{d_1 d_2} \right\|^{-1} \ll \sum_{1 \leq n_1 < d_1/2} \frac{d_1 d_2}{n_1 n_2} \left\| \frac{d_2}{d_1} \frac{n_1}{d_1} - \frac{d_1}{d_2} \frac{n_2}{d_2} \right\|^{-1},
\]
which by applying Lemma 2.6 with \( \alpha_2 = d_1 \) and \( \alpha_1 = d_2 \) is \( \ll d_1^2 d_2 \). Since \( d_2 n_1 - d_1 n_2 \) is an integer, if \( |d_2 n_1 - d_1 n_2| > d_1 d_2/h \) it must in fact be at least 1, so the lemma applies. This completes the proof. \( \square \)

In what follows, it will be helpful for us to approximate fractions \( \frac{a}{q} \) by a nearest multiple of \( \frac{1}{h} \); to do so, we make the following definition.

**Definition 2.8.** Let \( q > 1 \) and let \( 1 \leq a < q \) with \( (a, q) = 1 \). If \( q \geq h \), the \( h \)-approximate numerator \( n(a, q) \) is defined to as
\[
(7) \quad \frac{a}{q} = \left\lfloor \frac{ha}{q} \right\rfloor \quad \text{if } \frac{a}{q} \leq \frac{1}{2}
\]
\[
(7) \quad \frac{a}{q} = \left\lceil \frac{ha}{q} \right\rceil \quad \text{if } \frac{a}{q} > \frac{1}{2}.
\]

Meanwhile, if \( q < h \), the \( h \)-approximate numerator \( n(a, q) \) is defined to be \( a \) itself.

The key consequence of this definition is that for \( F(\alpha) = \min\{h, \|\alpha\|^{-1}\} \),
\[
(8) \quad F\left( \frac{a}{q} \right) \leq 2 \left\| \frac{n(a, q)}{\min\{q, h\}} \right\|^{-1}.
\]
2.2. Bounding $T_1$: terms with $gx \geq h$. Define

$$T_1 = \sum_{g,x,y,z \mid q, gx \geq h} \frac{\mu(gxyz)^2}{\phi(g)^3 \phi(xyz)^2} \sum_{1 \leq n_1, n_2, n_3 \leq \tilde{q}_i - 1} \sum_{1 \leq n_1, n_2, n_3 \leq \tilde{q}_i - 1} F \left( \frac{a_1}{qyz} \right) F \left( \frac{a_2}{gxz} \right) F \left( \frac{a_3}{gxy} \right).$$

We want to show the following lemma.

**Lemma 2.9.** Let $q$ be the product of primes up to $h^4$, and define $T_1$ by (9). Then

$$T_1 \ll h(\log h)^4.$$  

**Proof.** Recall that $q_1 = gyz$, $q_2 = gxz$, and $q_3 = gxy$. Since $gx \geq h$, $gxy$ and $gxyz$ (i.e., $q_2$ and $q_3$) must also both be $\geq h$. Recall the notation that $\tilde{q}_i = \min\{q_i, h\}$, so that $\tilde{q}_2 = \tilde{q}_3 = h$.

Since $\sum_i \frac{a_i}{q_i} \in \mathbb{Z}$, the sum $\sum_i \frac{n(a_i, q_i)}{q_i}$ of $h$-approximate numerators must be small, and in particular $\left| \sum_i \frac{n(a_i, q_i)}{q_i} \right| = O(1/h)$. We can then rewrite the sum as

$$T_1 = \sum_{g,x,y,z \mid q, gx \geq h} \frac{\mu(gxyz)^2}{\phi(g)^3 \phi(xyz)^2} \sum_{1 \leq n_1, n_2, n_3 \leq \tilde{q}_i - 1} \sum_{1 \leq n_1, n_2, n_3 \leq \tilde{q}_i - 1} F \left( \frac{a_1}{q_1} \right) F \left( \frac{a_2}{q_2} \right) F \left( \frac{a_3}{q_3} \right).$$

By (8), the inside summand can be bounded solely in terms of the $n_i$, with no dependence on the $a_i$'s themselves. The inside sum can then be evaluated by counting triplets $a_1, a_2, a_3$ with $n(a_i, q_i) = n_i$ for all $i$, $(a_i, q_i) = 1$, and $\sum_i a_i/q_i \in \mathbb{Z}$. There are $\ll \frac{q_i}{h} + 1$ choices of $a_i$ with $n(a_i) = n_i$.

The constraint that $\sum_i \frac{n_i}{q_i} \in \mathbb{Z}$, after multiplying out denominators, is equivalent to the constraint that

$$a_1 x + a_2 y + a_3 z \equiv 0 \mod gxyz.$$  

Once the $q_i$'s are fixed, a choice of $a_1$ determines $a_2 \mod z$. Thus fixing $a_2$ is equivalent to choosing a congruence class mod $gx$ for $a_2$; there are $\ll \frac{q_i}{h} + 1 \ll \frac{q_i}{h}$ choices of this congruence class with $n(a_2, q_2) = n_2$, since $gx \geq h$. Once $a_1$ and $a_2$ have been fixed, $a_3$ is entirely determined. Thus the total number of triplets $a_1, a_2, a_3$ satisfying all constraints is $\ll \left( \frac{q_i}{h} + 1 \right) \frac{q_x}{h}$.

Thus $T_1$ is bounded by

$$T_1 \ll \sum_{g,x,y,z \mid q, gx \geq h} \frac{\mu(gxyz)^2}{\phi(g)^3 \phi(xyz)^2} \sum_{1 \leq n_1, n_2, n_3 \leq \tilde{q}_i - 1} \sum_{1 \leq n_1, n_2, n_3 \leq \tilde{q}_i - 1} \left\| \frac{n_1}{q_1} \right\|^{-1} \left\| \frac{n_2}{h} \right\|^{-1} \left\| \frac{n_3}{h} \right\|^{-1} \left( \frac{q_i}{h} + 1 \right) \frac{gx}{h}.$$  

By Lemma 2.6,

$$\sum_{1 \leq n_1, n_2, n_3 \leq \tilde{q}_i - 1} \left\| \frac{n_1}{q_1} \right\|^{-1} \left\| \frac{n_2}{h} \right\|^{-1} \left\| \frac{n_3}{h} \right\|^{-1} \ll O \left( \tilde{q}_ih^2 \right).$$

The application of Lemma 2.6 here is slightly subtle. In the application of the Lemma, each $n_i$ is replaced with $n_i - \tilde{q}_i$ as necessary so that $n_i \in [-\tilde{q}_i/2, \tilde{q}_i/2]$ and $n_1 + n_2 + n_3 = O(1)$. Unless all three of $n_1, n_2$, and $n_3$ are very small, in which case there are $O(1)$ options for
\(n_1, n_2, n_3\) and (11) holds regardless, it must be the case that not all of \(n_1, n_2,\) and \(n_3\) have the same sign; by permuting the variable names one can then apply Lemma 2.6.

Thus

\[
T_1 \ll \sum_{g,x,y,z|q \atop gx \geq h} \frac{\mu(gxyz)^2}{\phi(g)^3 \phi(xyz)^2}k^3 \min \left\{ \frac{q_1}{h}, 1 \right\} \left( \frac{q_1}{h} + 1 \right) \frac{gx}{h}
\]

\[
\ll \sum_{g,x,y,z|q \atop gx \geq h} \frac{\mu(gxyz)^2}{\phi(g)^3 \phi(xyz)^2}hg_1gx
\]

\[
\ll h \sum_{g,x,y,z|q \atop gx \geq h} \frac{\mu(gxyz)^2}{\phi(g)\phi(gxyz)^2}g^2xyz, \text{ since } q_1 = gyz
\]

\[
\ll h(\log h)^4,
\]

as desired. \(\Box\)

2.3. Bounding \(T_2\): terms with \(gx, gy, gz\) small and \(a_2, a_3\) large. We now consider \(T_2\), which is the sum of terms in (6) where \(gx, gy,\) and \(gz\) are all \(< h\) and \(\left\| \frac{a_2}{gxz} \right\| \geq \frac{1}{h}\), and \(\left\| \frac{a_3}{gxy} \right\| \geq \frac{1}{h}\). That is, define

\[
(12) \quad T_2 := \sum_{g,x,y,z|q \atop x,y,z < h/g} \frac{\mu(gxyz)^2}{\phi(g)^3 \phi(xyz)^2} \sum_{a_1,a_2,a_3 \atop (a_1,gyz) = \ldots = 1, \atop a_1/gyz \ldots eZ, \atop |a_2/gxz| > 1/h, \atop |a_3/gxy| > 1/h} F \left( \frac{a_1}{gyz} \right) F \left( \frac{a_2}{gxz} \right) F \left( \frac{a_3}{gxy} \right).
\]

Lemma 2.10. Let \(q\) be the product of primes \(p \leq h^4\), and let \(T_2\) be defined as in (12). Then

\[
T_2 \ll h(\log h)^2(\log \log h)^2.
\]

Proof. Write

\[
a_1 \equiv b \equiv c \equiv a_2 \equiv c \equiv a_3 \equiv a - b \equiv gyz \equiv gy \equiv gz \equiv gxz \equiv x \equiv gxy \equiv x \equiv gy,
\]

where the equalities hold modulo 1. We pick \(1 \leq b \leq gy, 1 \leq c \leq gz\) relatively prime to \(gy, gz\) respectively so that the last two equalities hold, and then the first equality is given automatically.

Upon moving the sums over \(y\) and \(z\) in (12) inside, we get

\[
T_2 = \sum_{g,x|q \atop x < h/g} \frac{\mu(gx)^2}{\phi(g)^3 \phi(x)^2} \sum_{a \atop (a,x) = 1} S_2(g,x,a),
\]

where \(S_2(g,x,a)\) denotes the sum

\[
(13) \quad S_2(g,x,a) = \sum_{y,z|q \atop y,z < h/g} \frac{\mu(gxyz)^2}{\phi(yz)^2} \sum_{b,c \atop (b,gyz) = (c,gz) = 1} \left( \frac{a}{x} - \frac{b}{gy} \right) F \left( \frac{b}{gy} - \frac{c}{gz} \right) F \left( \frac{c}{gz} - \frac{a}{x} \right).
\]
Since $gy < h$ and $gz < h$, the product $yz$ is less than $h^2$, so that
\[
\frac{yz}{\phi(yz)} \ll \log \log h^2 \ll \log \log h.
\]
Thus we can replace the expression $\phi(yz)^2$ in the denominator in $S_2(g, x, a)$ with $y^z z^2$ while losing only a factor of $(\log \log h)^2$.

Let $\ell$ and $m$ be such that $2^\ell < \frac{a}{x} \leq 2^{\ell + 1}$ and $2^m < \frac{b}{y} \leq 2^{m + 1}$, and further define $n_\ell$ and $n_m$ to be variables ranging from 1 to $g2^\ell$ and 1 to $g2^m$ respectively.

Crucially, if \( \frac{n_\ell}{g2^\ell} \leq \frac{a}{x} - \frac{b}{gy} \leq \frac{n_\ell + 1}{g2^{\ell+1}} \), then \( F\left(\frac{a}{x} - \frac{b}{gy}\right) \ll F\left(\frac{n_\ell}{g2^\ell}\right) \), a bound which depends only on $\ell$ and $n_\ell$, and does not depend on $b$ or $y$. The same situation holds with $q$ and $z$. The range $1 \leq n_\ell \leq g2^\ell$ covers all possibilities for $\frac{a}{x} - \frac{b}{gy}$ because of our assumption that $\frac{a}{x} - \frac{b}{gy} \geq \frac{1}{h}$, which in particular rules out the case that $n_\ell = 0$. Similarly, the case that $n_m = 0$ is ruled out by our assumptions on $\frac{x}{g} - \frac{a}{x}$.

Thus
\[
S_2(g, x, a) \ll (\log \log h)^2 \sum_{\ell,m=1}^{\log_2 h/g} \sum_{n_\ell=1}^{2^\ell+1} \sum_{n_m=1}^{2^m} \frac{1}{2^{2\ell+2m}} \times F\left(\frac{n_\ell}{g2^\ell}\right) F\left(\frac{n_m}{g2^m}\right) F\left(\frac{n_\ell 2^m - n_m 2^\ell}{g2^{2\ell+m+1}}\right) \sum_{2^\ell < \frac{b}{y} \leq 2^{\ell+1}, \frac{n_\ell}{g2^\ell} < \frac{a}{x} - \frac{b}{gy} < \frac{n_\ell + 1}{g2^{\ell+1}}, \frac{n_m}{g2^m} < \frac{b}{y} < \frac{n_m + 1}{g2^{m+1}}}
\]

Define
\[
C_{\ell,n_\ell} = \# \left\{ b, y : \frac{b}{y} \in \left(\frac{\frac{a}{x} - \frac{n_\ell + 1}{y}, \frac{\frac{a}{x} - \frac{n_\ell}{y}}{y} \right) \cup \left(\frac{\frac{a}{x} + \frac{n_\ell}{y}, \frac{\frac{a}{x} + \frac{n_\ell + 1}{y}}{y}\right), \frac{1}{2^\ell + 1}, \frac{2^{\ell + 1}}{2^\ell}\right\},
\]
and define $C_{m,n_m}$ in the same way, so that the inside sum of $S_2(g, x, a)$ is given by $C_{\ell,n_\ell} C_{m,n_m}$. The minimum spacing of two distinct points $\frac{b_1}{y_1}$ and $\frac{b_2}{y_2}$ with denominators $y_i \leq 2^{\ell + 1}$ is $O(2^{-2\ell})$, so that
\[
C_{\ell,n_\ell} \ll 2^{2\ell + 1},
\]
and similarly $C_{m,n_m} \ll 1$. This implies that
\[
S_2(g, x, a) \ll (\log \log h)^2 \sum_{\ell,m=1}^{\log_2 h/g} \sum_{n_\ell=1}^{2^\ell+1} \sum_{n_m=1}^{2^m} \frac{1}{2^{2\ell+2m}} F\left(\frac{n_\ell}{g2^\ell}\right) F\left(\frac{n_m}{g2^m}\right) F\left(\frac{n_\ell 2^m - n_m 2^\ell}{g2^{2\ell+m+1}}\right).
\]

Applying Lemma 2.7 to the inner sum gives
\[
S_2(g, x, a) \ll (\log \log h)^2 \sum_{\ell,m=1}^{\log_2 h/g} \frac{1}{2^{2\ell+2m}} h g^2 2^\ell \ll h (\log \log h)^2 2^\ell \sum_{\ell,m=1}^{\log_2 h/g} \frac{1}{2^{2m}} = h (\log \log h)^2 g^2,
\]
and thus
\[ T_2 \ll h(\log \log h)^2 \sum_{g,x \in [q \mod h]} \frac{\mu(g)^2}{\phi(g)^3 \phi(x)^2} \sum_{a \in [x]} g^2 \]
\[ \ll h(\log \log h)^2 \sum_{g,x \in [q \mod h]} \frac{\mu(g)^2 g^2}{\phi(g)^3 \phi(x)} < h(\log h)^2 (\log \log h)^2. \]

2.4. Bounding $T_3$: terms with $gx, gy, gz$ small and each $a_i$ small. All that remains is to analyze the sum $T_3$, which consists of the terms in (6) where $gx, gy$, and $gz < h$, and for each $i$, $\left\| \frac{a_i}{q} \right\| < \frac{2}{h}$. Precisely, we define

(14) \[ T_3 := \sum_{g,x,y,z \in [q \mod h]} \frac{\mu(gxyz)^2}{\phi(g)^3 \phi(xyz)^2} \sum_{a_1,a_2,a_3 \in \mathbb{Z}} \phi(a_1 gyz) \phi(a_2 gxz) \phi(a_3 gxy). \]

Lemma 2.11. Let $q$ be the product of all primes $p \leq h^4$ and define $T_3$ by (14). Then

$T_3 \ll h(\log h)^4 (\log \log h)^2$.

Proof. Since $\left\| \frac{a_i}{q} \right\| < \frac{2}{h}$, this implies that

\[ \frac{1}{gy} < \frac{2}{h}, \]

so if $y < \sqrt{\frac{h^2}{2g}}$, then $x > \sqrt{\frac{h^2}{2g}}$. Thus at most one of $x, y, z$ can be $< \sqrt{\frac{h^2}{2g}}$. By relabeling if necessary, we get that

$T_3 \ll \sum_{g,x,y,z \in [q \mod h]} \frac{\mu(gxyz)^2}{\phi(g)^3 \phi(xyz)^2} \sum_{a_1,a_2,a_3 \in \mathbb{Z}} \phi(a_1 gyz) \phi(a_2 gxz) \phi(a_3 gxy).$

As in the proof of Lemma 2.10,

\[ \frac{a_1}{gy} = \frac{b}{gy} - \frac{c}{gy}, \quad \frac{a_2}{gx} = \frac{c}{gx} - \frac{a}{x}, \quad \frac{a_3}{gx} = \frac{a}{x} - \frac{b}{gy}, \]

and we can reparametrize $T_3$ in terms of sums over $a, b, c$ instead of $a_1, a_2, a_3$. Doing so, and moving the sums over $b, y, c$, and $z$ inside, we get that

$T_3 \ll h^3 \sum_{g,x \in [q \mod h]} \frac{\mu(g)^2}{\phi(g)^3 \phi(x)} \sum_{a \in [x \mod h]} S_3(g,x,a).$
where
\[ S_3(g, x, a) = \sum_{\substack{h/y \leq y \leq h/g \leq h/\ell, \ h/g \leq z \leq h/g \leq h/\ell}} \frac{\mu(yz)^2}{\phi(yz)^2} \# \left\{ b, c : \frac{b}{gy}, \frac{c}{gz} \in \left( \frac{a}{x} - \frac{1}{h}, \frac{a}{x} + \frac{1}{h} \right) \right\}. \]

Since \( y, z \leq h \), the product \( yz \leq h^2 \), and thus \( \frac{1}{\phi(yz)^2} \ll \frac{(\log \log h)^2}{y^2z^2} \), so we can replace \( \phi(yz)^2 \) in the denominator of \( S_3(g, x, a) \) with \( y^2z^2 \) and in doing so lose a factor of \( (\log \log h)^2 \). In order to estimate \( S_3(g, x, a) \), we split the sums over \( y \) and \( z \) dyadically, defining \( \ell \) such that \( 2^\ell < y \leq 2^{\ell+1} \) and \( 2^m < z \leq 2^{m+1} \).

Then
\[ S_3(g, x, a) \ll (\log \log h)^2 \sum_{\ell, m = \frac{1}{2}(\log_2 (h/g))}^{\log_2 (h/g)} \frac{C_{\ell}C_m}{2^{2\ell}2^{2m}}, \]

where
\[ C_\ell = \# \left\{ b, y : \frac{b}{y} \in \left( \frac{ga}{x} - \frac{g}{h}, \frac{ga}{x} + \frac{g}{h} \right), 1 \leq b < y, y \leq 2^{\ell+1} \right\}, \]

and similarly for \( C_m \). The minimum spacing of two distinct points \( \frac{b_1}{y_1} \) and \( \frac{b_2}{y_2} \) with denominators at most \( 2^{\ell+1} \) is \( O \left( \frac{1}{2^\ell} \right) \), so then \( C_\ell \ll 2^{2\ell}g^2 \), \( 2^\ell g \geq 1 \), so in particular \( C_\ell \ll 2^{2\ell}g^2 \), and similarly for \( C_m \).

Plugging this in gives
\[ S_3(g, x, a) \ll (\log \log h)^2 \sum_{\ell, m = \frac{1}{2}(\log_2 (h/g))}^{\log_2 (h/g)} \frac{2^{2\ell}2^{2m}g^2}{2^{2\ell}2^{2m}h^2} \ll \frac{g^2}{h^2} (\log (h/g))^2 (\log \log h)^2, \]

so that
\[ T_2 \ll h(\log \log h)^2 \sum_{\substack{g, x \mid q \quad g \leq h}} \frac{\mu(gx)^2g^2}{\phi(g)^3\phi(x)^2} \sum_{\substack{a \leq x \quad (a, x) = 1}} (\log (h/g))^2 \]
\[ \ll h(\log \log h)^2 \sum_{\substack{g, x \mid q \quad g \leq h}} \frac{\mu(gx)^2g^2}{\phi(g)^3\phi(x)} (\log (h/g))^2 \]
\[ \ll h(\log \log h)^3 \sum_{g \leq h} \frac{1}{g} (\log (h/g))^3 \]
\[ \ll h(\log h)^4 (\log \log h)^3. \]

□

Putting Lemmas 2.9, 2.10, and 2.11 together yields Theorem 2.1.

3. Function Field Analogues: Proof of Theorem 1.3

We consider analogous questions when working in \( \mathbb{F}_q[T] \) rather than in \( \mathbb{Z} \). To begin with, let’s set up the situation in the function field case. Fix a finite field \( \mathbb{F}_q \). Rather than primes in \( \mathbb{N} \), consider monic irreducible polynomials in \( \mathbb{F}_q[T] \).
The norm of a polynomial \( F \in \mathbb{F}_q[T] \) is defined as as \( |F| = q^\deg F \). We consider intervals in norm, where the interval \( I(F, h) \) of degree \( h \) is defined as

\[
I(F, h) := \{ G \in \mathbb{F}_q[T] : |F - G| < q^h \}
\]

For a fixed monic polynomial \( Q \), we denote

\[
\mathcal{C}(Q) := \left\{ \frac{A}{Q} \in \mathbb{F}_q[T] : |A| < |Q| \right\}
\]

\[
\mathcal{R}(Q) := \left\{ \frac{A}{Q} \in \mathbb{F}_q[T] : |A| < |Q|, (A, Q) = 1 \right\}
\]

\[
\mathcal{N}(Q) := \left\{ \frac{A}{Q} \in \mathbb{F}_q[T] : |A| < |Q|, Q \nmid A \right\}
\]

For \( Q = 1 \), we instead for convenience define \( \mathcal{C}(Q) = \{1\} = \mathcal{R}(Q) \). If \( \deg Q > 0 \), the set of polynomials \( F \) with \( \deg F < \deg Q \) is a canonical set of representatives of \( \mathbb{F}_q[T]/(Q) \); in what follows, we will identify \( \{ F \in \mathbb{F}_q[T] : \deg F < \deg Q \} \) with \( \mathbb{F}_q[T]/(Q) \). If \( Q = 1 \), we will take 1 to represent the unique equivalence class of \( \mathbb{F}_q[T]/(Q) \).

Our problem is then to consider the \( k \)th moment of the distribution of irreducible polynomials in intervals \( I(F, h) \). As in the integer case, we begin by considering the related quantity of the distribution of reduced residues modulo a squarefree monic polynomial \( Q \).

That is, for \( Q \) a fixed squarefree monic polynomial, we consider

\[
m_k(Q; h) = \sum_{F \mod Q} \left( \sum_{G \in I(F, h); (G, Q) = 1} 1 \right) - \frac{q^h \phi(Q)}{|Q|} \right)^k.
\]

Here we are taking the centered moment \( m_k(Q; h) \) by subtracting \( \frac{q^h \phi(Q)}{|Q|} \).

As in the integer case, we can express the moment \( m_k(Q; h) \) in terms of exponential sums. For \( \alpha \in \mathbb{F}_q(T) \) a rational function, let \( \text{res}(\alpha) \) be the coefficient of \( \frac{1}{T} \) when \( \alpha \) is written as a Laurent series with finitely many positive terms. Then we define

\[
e(\alpha) := e_q(\text{res}(\alpha)) = \exp(2\pi i \cdot \text{tr}(\text{res}(\alpha))/p),
\]

where \( q \) is a power of the prime \( p \) and \( \text{tr} : \mathbb{F}_q \to \mathbb{F}_p \) is the trace function. This exponential function, like its integer analog, satisfies the crucial property that for a monic polynomial \( F \in \mathbb{F}_q[T] \),

\[
\sum_{\alpha \in \mathcal{C}(F)} e(\alpha) = \begin{cases} 1 & \text{if } F = 1 \\ 0 & \text{otherwise}. \end{cases}
\]

We then have the following lemma, analogous to Lemma 2 in [13].

**Lemma 3.1.**

\[
m_k(Q; h) = |Q| \left( \frac{\phi(Q)}{|Q|} \right)^k V_k(Q; h),
\]
where
\[ V_k(Q; h) = \sum_{R_i, \ldots, R_k | Q, i=1}^{k} \mu(R_i) \prod_{|R_i| > 1} E \left( \frac{\rho_1}{R_i} \right) \cdots E \left( \frac{\rho_k}{R_k} \right), \]
and where, for \( \alpha \in \mathbb{F}_q(t) \) a rational function,
\[ E(\alpha) = \sum_{M \in I(0, h)} e(M\alpha). \]

The proof follows that of Lemma 2 in [13] very closely.

**Proof.** Let \( \kappa(R) = 1 \) when \( (R, Q) = 1, \kappa(R) = 0 \) otherwise. Then
\[ \kappa(R) = \sum_{S \mid (R, Q)} \mu(S) = \sum_{S \mid Q} \frac{\mu(S)}{|S|} \sum_{\sigma \in C(S)} e(R\sigma) \]
\[ = \sum_{T \mid Q} \left( \sum_{A \in C(T) \mid (A, T) = 1} e(RA) \right) \left( \sum_{T \mid S} \frac{\mu(S)}{|S|} \right). \]

Here the second factor is \( \frac{\phi(Q)}{|Q|} \frac{\mu(T)}{|T|} \). The function \( \kappa(R) \) has mean value \( \frac{\phi(Q)}{|Q|} \), so we subtract \( \frac{\phi(Q)}{|Q|} \) from both sides, which deletes the term when \( T = 1 \). We then replace \( R \) by \( M + N \), and sum over \( M \) to see that
\[ \sum_{|M| < q^h} 1 - h \frac{\phi(Q)}{|Q|} = \sum_{|R| > 1} \frac{\mu(R)}{|R|} \sum_{A \in C(R) \mid (A, R) = 1} E \left( \frac{A}{R} \right) e(NA/R). \]

The argument is completed upon raising both sides to the \( k \)th power, summing over \( N \), multiplying out the right hand side, and appealing to the fact that
\[ \sum_{|N| < q^d} e(N(\alpha_1 + \cdots + \alpha_k)) = \begin{cases} q^d & \text{if } \sum \alpha_i \in \mathbb{Z} \\ 0 & \text{else.} \end{cases} \]

One important difference between the integer setting and the function field setting is the behavior of the sums \( E(\alpha) \), which are particularly well-behaved in \( \mathbb{F}_q[T] \). These sums have also been studied by Hayes in [8, Theorem 3.5].

**Lemma 3.2.** Let \( \alpha \in \mathbb{F}_q(t) \) be a rational function with \( \deg \alpha \leq -1 \). Then
\[ E(\alpha) = \begin{cases} q^h & \text{if } \deg \alpha < -h \\ 0 & \text{if } \deg \alpha \geq h. \end{cases} \]

**Proof.** Let \( \mathcal{P}_h \subseteq \mathbb{F}_q[t] \) be the set of polynomials of degree less than \( h \). Assume first that \( \deg \alpha < -h \). Then for all \( M \in \mathcal{P}_h \), \( \deg M\alpha = \deg M + \deg \alpha \leq h - 1 - h - 1 = -2 \), so the Laurent series for \( M\alpha \) has no \( \frac{1}{i} \) term, and thus \( \text{res}(M\alpha) = 0 \). But then
\[ E(\alpha) = \sum_{M \in \mathcal{P}_h} e(M\alpha) = \sum_{M \in \mathcal{P}_h} e_q(\text{res}(M\alpha)) = \sum_{M \in \mathcal{P}_h} 1 = q^h. \]
Now assume that \( \deg \alpha \geq -h \). Consider the map \( \text{res}_{\alpha} : \mathcal{P}_h \to \mathbb{F}_q \) which at a polynomial \( M \) returns the residue of \( M \alpha \). This map is linear over \( \mathbb{F}_q \), so its image is either 0 or all of \( \mathbb{F}_q \). Let \( M = t^{-\deg \alpha - 1} \). Since \( -h \leq \deg \alpha \leq -1 \), we have \( 0 \leq -\deg \alpha - 1 \leq h - 1 \), so \( M \) indeed is a polynomial in \( \mathcal{P}_h \). On the other hand, \( \text{res}(M\alpha) \) is precisely the leading coefficient of \( \alpha \), which must be nonzero. Thus the image of \( \text{res}_{\alpha} \) is nonzero, so it is all of \( \mathbb{F}_q \). In particular, \( \text{res}_{\alpha}(M) \) takes each value in \( \mathbb{F}_q \) equally often. Thus

\[
E(\alpha) = \sum_{M \in \mathcal{P}_h} e_q(\text{res}(M\alpha))
\]

is a balanced exponential sum, which has sum 0.

This fact and other properties of the sums \( E(\alpha) \) mean that the analysis of Montgomery and Vaughan in [13] in the function field setting is more streamlined. In fact, their work automatically gives the analog of our desired bound for the third moment in the function field case.

3.1. The analog of [13] in the function field setting. We begin with the following fundamental lemma, with an identical proof to the integer case.

**Lemma 3.3** (Fundamental Lemma). Let \( R_1, \ldots, R_k \) be squarefree monic polynomials with \( R = [R_1, \ldots, R_k] \). Suppose for all irreducible \( P \mid R \), \( P \) divides at least two \( R_i \)'s. Let \( G_i \) be positive real-valued function defined on \( \mathcal{C}(R_i) \). Then

\[
\left| \sum_{\substack{A_i \mod R_i \\ \sum_i A_i \equiv 0 \mod R}} G_1 \left( \frac{A_1}{R_i} \right) \cdots G_k \left( \frac{A_k}{R_k} \right) \right| \leq \frac{1}{|R|} \prod_{i=1}^k \left| R_i \right| \sum_i \left| G_i \left( \frac{A_i}{R_i} \right) \right|^2 \right)^{1/2}.
\]

The proof follows Montgomery-Vaughan very closely.

**Proof.** We proceed by induction on \( k \).

Assume first that \( k = 2 \). Then we must have \( R_1 = R_2 = R \). By Cauchy-Schwarz,

\[
\left| \sum_{|A| < |R|} G_1 \left( \frac{A}{R} \right) G_2 \left( \frac{A}{R} \right) \right| \leq \left( \sum_{|A| < |R|} \left| G_1 \left( \frac{A}{R} \right) \right|^2 \right)^{1/2} \left( \sum_{|A| < |R|} \left| G_2 \left( \frac{A}{R} \right) \right|^2 \right)^{1/2},
\]

which after a bit of rearranging gives the desired result.

For arbitrary \( k \), set \( D = (R_1, R_2) \), and write \( D = ST \) with \( S|R_3 \cdots R_k \) and \( (T, R_3 \cdots R_k) = 1 \). Furthermore, write \( R_1 = DR_1' \) and \( R_2 = DR_2' \). Consider any term in the sum. Since \( \sum_i \frac{A_i}{R_i} = 0 \), we have \( T \left| \left( \frac{A_1}{R_1} + \frac{A_2}{R_2} \right) \right. \). Thus \( \frac{A_1}{STR_1'} + \frac{A_2}{STR_2'} \) can be expressed as a fraction \( \frac{A}{RSTR_3 \cdots R_k} \).

By the Chinese Remainder theorem, \( \frac{A_1}{STR_1'} = \frac{a_1}{R_1} + \frac{\beta_1}{ST} \) and \( \frac{A_2}{STR_2'} = \frac{a_2}{R_2} + \frac{\beta_2}{ST} \), where \( \beta_1 = -\frac{\beta_1}{ST} + \frac{\gamma}{S} \) because \( T \left| \left( \frac{A_1}{R_1} + \frac{A_2}{R_2} \right) \right. \). Thus \( \frac{A_1}{R_1} \) and \( \frac{A_2}{R_2} \) can be written as \( \frac{A_1}{R_1} = \frac{A'_1}{R_1} + \frac{\delta}{D} \) and \( \frac{A_2}{R_2} = \frac{A'_2}{R_2} + \frac{\gamma}{S} - \frac{\delta}{D} \), with each rational function of degree less than 0.

Let \( R^* = R'_1R'_2S \). For each \( A^* \) with \( |A^*| < |R^*| \), \( \frac{A^*}{R^*} \) is uniquely of the form \( \frac{A'_1}{R'_1} + \frac{A'_2}{R'_2} + \frac{\gamma}{S} \). Define
Proof. If \( \deg E < \) which completes the first portion.

Fix a rational function \( \alpha \). For all \( \frac{S}{R} \), \( E \left( \frac{S}{R} + \alpha \right) \) is unchanged by replacing \( \alpha \) with its fractional part; i.e, subtracting off the polynomial portion of \( \alpha \) so that \( |\alpha| < 1 \), including the possibility that \( \alpha = 0 \).

Lemma 3.4. For any polynomial \( R \),

\[
\sum_{S \mod R} E \left( \frac{S}{R} \right)^2 = \max\{q^{2h}, |R|q^h\}
\]

Moreover, for any polynomial \( R \) and any rational function \( \alpha \),

\[
\sum_{S \mod R} E \left( \frac{S}{R} + \alpha \right)^2 = \begin{cases} 
\max\{q^{2h}, |R|q^h\} & \text{if } |\alpha| < q^{-h} \\
\leq |R|q^{h-1} & \text{if } |\alpha| \geq q^{-h}
\end{cases}
\]

Proof. If \( \deg R \ll h \), then for all \( S \) with \( 0 \neq |S| < |R|, h \geq \deg R - \deg S \), and thus \( E \left( \frac{S}{R} \right)^2 = 0 \). Meanwhile, \( E(0)^2 = q^{2h} \), so in this case the sum is \( q^{2h} \).

Now suppose \( \deg R > h \). Then \( E \left( \frac{S}{R} \right) \) is nonzero if and only if deg \( S < \deg R - h \). Thus

\[
\sum_{S \mod R} E \left( \frac{S}{R} \right)^2 = \sum_{S \mod R} E \left( \frac{S}{R} \right)^2 = \sum_{S \mod R} q^{2h} = |R|q^h,
\]

which completes the first portion.
If a term \(E \left( \frac{S}{R} + \alpha \right)\) is nonzero, then \(\left| \frac{S}{R} + \alpha \right| < q^{-h}\). We’ll split into two cases, when \(|\alpha| < q^{-h}\) and when \(|\alpha| \geq q^{h}\). First, if \(|\alpha| < q^{-h}\), then \(\left| \frac{S}{R} + \alpha \right| < q^{-h}\) if and only if \(\left| \frac{S}{R} \right| < q^{-h}\).

If \(|R| \geq q^{h}\), there are \(|R|/q^{h}\) values of \(S\) satisfying this; if not, there is 1 value. Thus if \(|\alpha| < q^{-h}\), we have \(\sum_{S \mod R} E \left( \frac{S}{R} + \alpha \right)^2 E \left( \frac{S}{R} + \alpha \right) = \max(q^{2h}, |R|q^{h})\).

Meanwhile if \(|\alpha| \geq q^{-h}\), then for \(\left| \frac{S}{R} + \alpha \right| < q^{-h}\) we must have \(\left| \frac{S}{R} \right| = |\alpha| \geq q^{-h}\). Thus the first deg \(\alpha + h + 1\) terms of \(\frac{S}{R}\) are fixed, by the fact that they must cancel with the corresponding terms of \(\alpha\) to yield a rational function of small enough degree. Correspondingly, the first deg \(\alpha + h + 1\) terms of \(S\) are determined. We have \(|S| = |Ra|\), so there are at most \(|R|a\cdot \frac{1}{|\alpha|q^{h-1}} = |R|q^{h-1}\) nonzero terms here. Thus in this case we have \(\sum_{S \mod R} E \left( \frac{S}{R} + \alpha \right)^2 \ll |R|q^{h-1}\).

The following lemma corresponds to Lemma 6 of Montgomery-Vaughan.

**Lemma 3.5.** Let \(R\) be a polynomial, and let \(\alpha, \beta\) be rational functions. Then

\[
\sum_{S \mod R} E \left( \frac{S}{R} + \alpha \right) E \left( \frac{S}{R} + \beta \right) \ll E(\alpha - \beta)q^{-h} \sum_{S \mod R} E \left( \frac{S}{R} + \alpha \right)^2
\]

*Proof.* Again, we split into two cases. Assume first that \(|\alpha - \beta| \geq q^{-h}\). Then either \(\left| \frac{S}{R} + \beta \right| \geq q^{-h}\), or \(\left| \frac{S}{R} + \alpha \right| \geq q^{-h}\). Thus for each \(\frac{S}{R}\), either \(E \left( \frac{S}{R} + \alpha \right) = 0\) or \(E \left( \frac{S}{R} + \beta \right) = 0\), so the product must be 0, and thus the sum is 0.

Now assume that \(|\alpha - \beta| < q^{-h}\), so \(E(\alpha - \beta) = q^{h}\). Then \(E \left( \frac{S}{R} + \alpha \right) \neq 0\) if and only if \(E \left( \frac{S}{R} + \beta \right) \neq 0\), so

\[
\sum_{S \mod R} E \left( \frac{S}{R} + \alpha \right) E \left( \frac{S}{R} + \beta \right) = \sum_{S \mod R} E \left( \frac{S}{R} + \alpha \right)^2,
\]

which gives the result. \(\square\)

With these preliminaries in hand, we prove the following lemma, which corresponds to Lemma 7 of [13].

**Lemma 3.6.** Let \(k \geq 3\), and let \(R_1, \ldots, R_k\) be squarefree polynomials with \(|R_i| > 1\) for all \(i\). Let \(R = [R_1, \ldots, R_k]\). Let \(D = (R_1, R_2)\) and \(D = ST\) with \(S|R_3 \cdots R_k\) and \((T, R_3 \cdots R_k) = 1\). Write \(R_1 = DR_1, R_2 = DR_2,\) and \(R^* = R_1R_2S\). Define

\[
S(R_1, \ldots, R_k) := \sum_{(A_i \mod R_i) \neq 1 \cup \cup_{i=1}^k E \left( \frac{A_i}{R_i} \right)}
\]

If there exists an \(i\) such that \(|R_i| \leq q^h\), then \(S(R_1, \ldots, R_k) = 0\). Otherwise,

\[
S(R_1, \ldots, R_k) \ll |R_1 \cdots R_k| \cdot |R|^{-1}(q^h)^{k/2}(X_1 + X_2 + X_3),
\]
where

\[ X_1 = q^{-h/2} \]
\[ X_2 = \begin{cases} |D|^{-1} & \text{if } |R'_1| > q^h \\ 0 & \text{otherwise} \end{cases} \]
\[ X_3 = \begin{cases} |S|^{-1/2} & \text{if } R_1 = R_2 \\ 0 & \text{otherwise} \end{cases} \]

**Proof.** Assume first that for some \( i \), \(|R_i| \leq q^h\). Then \( E(A_i/R_i) = 0\) whenever \( A_i \neq 0\), so in particular for all \( A_i \) with \( (A_i, R_i) = 1 \), the sum is 0. Assume from now on that \(|R_i| > q^h\) for all \( i \).

We now return to the proof of the Fundamental Lemma. For \( \frac{A^*}{R^*} = \frac{A'_1}{R'_1} + \frac{A'_2}{R'_2} + \frac{\sigma}{S} \), we define

\[
G^* \left( \frac{A^*}{R^*} \right) = \sum_{\substack{\left(D A'_1 + \delta R'_1, R_1 \right) = 1 \\
\left(D A'_2 R'_2, R_1 \right) = 1}} E \left( \frac{A'_1}{R'_1} + \frac{\delta}{D} \right) E \left( \frac{A'_2}{R'_2} + \frac{\sigma - \delta}{S} \right).
\]

For this sum to be nonempty, \((A'_1, R'_1) = (A'_2, R'_2) = 1\). Then

\[
S(R_1, \ldots, R_k) \leq \left| \frac{T}{R} \right| \left( |R^*| \sum_{A^* \mod R^*} G^* \left( \frac{A^*}{R^*} \right) \right)^{1/2} \prod_{i=3}^{k} |R_i| \sum_{\substack{A_i \mod R_i \\
(A_i, R_i) = 1}} 1 \right)^{1/2}
\]

By Lemma 3.4, the product is \( \ll |R_3 \cdots R_k|q^{hk/2-h} \). Thus it suffices to show that

\[
\sum_{A^* \mod R^*} G^* \left( \frac{A^*}{R^*} \right)^2 \ll |R_1| \cdot |R_2| \cdot |S|q^{2h}(X_1^2 + X_2^2 + X_3^2).
\]

By Lemma 3.5,

\[
G^* \left( \frac{A^*}{R^*} \right) \ll E \left( \frac{A^*}{R^*} \right) q^{-h} \sum_{\delta \mod D} E \left( \frac{\delta}{D} + \frac{A'_1}{R'_1} \right).
\]

Lemma 3.4 then gives us the following:

\[
G^* \left( \frac{A^*}{R^*} \right) \ll \begin{cases} E \left( \frac{A^*}{R^*} \right) \max\{q^h, |D|\} & \text{if } |A'_1/R'_1| < q^{-h} \\ E \left( \frac{A^*}{R^*} \right) |D|q^{-1} & \text{if } |A'_1/R'_1| \geq q^{-h} \end{cases}
\]

Summing over \( A^* \) then gives

\[
\sum_{A^* \mod R^*} G^* \left( \frac{A^*}{R^*} \right)^2 \ll \sum_{A^* \mod R^*} E \left( \frac{A^*}{R^*} \right)^2 \max\{q^{2h}, |D|^2\} + \sum_{A^* \mod R^*} E \left( \frac{A^*}{R^*} \right)^2 |D|^2
\]

In the equation above,
Here as in the definition of $G^*$, for any nonzero term we must have $(A_1', R_1') = (A_2', R_2') = 1$. In particular, $A_1'$ can only be congruent to 0 if $R_1' = 1$. We now split into cases based on whether or not $|R^*| > q^h$ and whether or not $|R_1'| > q^h$. First assume that $|R^*| > q^h$ and $|R_1'| > q^h$. Then we have

$$\sum_{A^* \text{ mod } R^*} G^* \left( \frac{A^*}{R^*} \right)^2 \ll \max\{q^{2h}, |D|^2\} |A^*|^{2h} \frac{|R_1'|}{q^h} + |D|^2 \sum_{A^* \text{ mod } R^* \atop |A^*|/R^* < q^{-h} \atop |A_1'/R_1'| \geq q^h} E \left( \frac{A^*}{R^*} \right)^2$$

$$\ll \max\{q^{2h}, |D|^2\} |R^*| + |D|^2 |R^*| |q^h$$

Now assume that $|R^*| > q^h$ but $|R_1'| \leq q^h$. Then the first sum is empty unless $R_1' = 1$ and $A_1' = 0$. If $R_1' = 1$, then $R_1 = D$, so $|D| > q^h$. The above bound then yields

$$\sum_{A^* \text{ mod } R^*} G^* \left( \frac{A^*}{R^*} \right)^2 \ll q^{2h} |D|^2 + \left| \frac{|R^*|}{q^h} q^{2h} |D|^2 \right| |R_1 R_2 S| q^{2h} \left( \frac{1}{|R^*|} + q^{-h} \right) \ll |R_1 R_2 S| q^{2h} (X_1^2)$$

If $R_1' \neq 1$, then the first sum is empty, so we get

$$\sum_{A^* \text{ mod } R^*} G^* \left( \frac{A^*}{R^*} \right)^2 \ll \frac{|R^*|}{q^h} q^{2h} |D|^2 = |R_1 R_2 S| q^{2h} (X_1^2)$$

Finally, assume that $|R^*| \leq q^h$ and thus $|R_1'| \leq q^h$. In this case the only nonzero term in either sum is when $A^* = 0$, which forces $A_1' = A_2' = \sigma = 0$. But then since $(A_1', R_1') = (A_2', R_2') = 1$, we must have $R_1' = R_2' = 1$, so that $R_1 = R_2 = D$, which has magnitude $> q^h$. In this case

$$\sum_{A^* \text{ mod } R^*} G^* \left( \frac{A^*}{R^*} \right)^2 \ll q^{2h} |D|^2 = |R_1 R_2 S| q^{2h} \cdot |S|^{-1} = |R_1 R_2 S| q^{2h} X_3^2.$$

We now turn to the proof of Theorem 1.3, which corresponds to Lemma 8 of [13]. The main strategy here is a careful application of Lemma 3.6, keeping in mind that we can choose which variables play the roles of $R_1$ and $R_2$.

**Lemma 3.7.** For any fixed $k \geq 3$, for $Q \in \mathbb{F}_q[T]$ squarefree,

$$m_k(Q; h) \ll |Q| (q^h)^{k/2} \left( \frac{\phi(Q)}{|Q|} \right)^{k/2} \left( 1 + ((q^h)^{-1/2} + (q^h)^{-1/(k-2)}) \left( \frac{\phi(Q)}{|Q|} \right)^{-2k+k/2} \right)$$

**Proof.** We begin with the bound that

$$m_k(Q; h) \ll |Q| \left( \frac{\phi(Q)}{|Q|} \right)^{k} \sum_{R \mid Q} \sum_{R \mid Q} \frac{S(R_1, \ldots, R_k)}{\phi(R_1) \cdots \phi(R_k)}$$

23
where $S(R_1, \ldots, R_k) = \sum_{A_i \mod R_i} \prod_{i=1}^{k} E(A_i/R_i)$. We will then apply Lemma 3.6, but we’ll use the fact that we have flexibility in how we label $R_1, \ldots, R_k$ in our application of Lemma 3.6. We’ll choose $R_1$ and $R_2$ as follows.

If for any $i$ we have $|R_i| < q^h$, then $S(R_1, \ldots, R_k)$ must be 0, so assume that $|R_i| \geq q^h$ for all $i$. Let $R_{ij} = (R_i, R_j)$. Since $R_i | \prod_{j \neq i} R_j$, $R_i | \prod_{i \neq j} R_{ij}$ as well, for all $i$. Thus for all $i$, there exists some $j \neq i$ such that $|R_{ij}| \geq |R_i|^{1/(k-1)}$. If there exists some $i, j$ so that $|R_{ij}| \geq |R_i|^{1/(k-1)}$ but $R_i \neq R_j$, then pick $R_1$ and $R_2$ to be $R_i$ and $R_j$, respectively.

Meanwhile, if no such $i$ exists, then for each $i$, there is some $j \neq i$ with $R_i = R_j$. If there exists any triple $R_i = R_j = R_k$, then pick $R_1$ and $R_2$ to be $R_i$. If not, then the $R_i$’s must be equal in pairs and otherwise disjoint, and $k$ must be even. Without loss of generality, say that $R_1 = R_2, R_3 = R_4, \ldots, R_{k-1} = R_k$, before the relabeling. Write $R = UV$, where $V$ is the product of all primes dividing at least two $R_{2i}$’s, and $U$ is the product of all primes dividing exactly one $R_{2i}$. Then

$$V^{2k/2} \prod_{i=1}^{k/2} \left( R_{2i}, \prod_{j \neq i} R_{2j} \right),$$

so there exists some $i$ with $\left( R_{2i}, \prod_{j \neq i} R_{2j} \right) \geq |V|^{4/k}$. Take $R_1$ and $R_2$ to be $R_{2i}$ and $R_{2i+1}$.

Now we return to our bound on $m_k(Q; h)$. We have

$$m_k(Q; h) \ll |Q| \left( \frac{\phi(Q)}{|Q|} \right)^k \left( q^h \right)^{k/2} \sum_{\substack{R | Q \text{ monic} \ R \neq Q}} \frac{1}{|R|} \sum_{\substack{R_i | Q \text{ monic} \ |R_i| > 1 \ |R_1, \ldots, R_k| = R}} \frac{|R_1 \cdots R_k|}{\phi(R_1) \cdots \phi(R_k)} (X_1 + X_2 + X_3),$$

where the $X_i$ arise by use of Lemma 3.6 as described above.

Consider the contribution from each $X_i$. Since $X_1 = q^{-h/2}$, the $X_1$ terms contribute

$$\ll |Q| \left( \frac{\phi(Q)}{|Q|} \right)^k \left( q^h \right)^{k/2 - 1/2} \sum_{\substack{R | Q \text{ monic} \ R \neq Q}} \frac{1}{|R|} \sum_{\substack{R_i | Q \text{ monic} \ |R_i| > 1 \ |R_1, \ldots, R_k| = R}} \frac{|R_1 \cdots R_k|}{\phi(R_1) \cdots \phi(R_k)}$$

$$\ll |Q| \left( \frac{\phi(Q)}{|Q|} \right)^k \left( q^h \right)^{k/2 - 1/2} \prod_{P | Q} \left( 1 + \frac{1}{|P|} \left( 2 + \frac{1}{|P| - 1} \right)^k \right)$$

$$\ll |Q| \left( q^h \right)^{k/2 - 1/2} \left( \frac{\phi(Q)}{|Q|} \right)^{2k + k}.$$

Now consider the $X_2$ contribution. If $X_2 \neq 0$, then $|R_i| > q^h$. By our choice of $R_1, R_2$, this implies that $|D| \geq |R_i|^{1/(k-1)}$. But then $|D| \geq |R_1' \cdot D|^{1/(k-1)}$, so $|D|^{-1} \leq q^{-h/(k-2)}$. By the same logic as for the $X_1$ terms, the $X_2$ terms contribute

$$\ll |Q| \left( q^h \right)^{k/2 - 1/(k-2)} \left( \frac{\phi(Q)}{|Q|} \right)^{2k + k}.$$

Finally we consider $X_3$. If $X_3 \neq 0$, then $R_1 = R_2$. By our choice of $R_1$ and $R_2$ for the application of Lemma 3.6, in this case each $R_i$ is equal to some $R_j$. If there exists some
$R_i = R_1 = R_2$, with $i \geq 3$, then $S = R_1 = R_2$, so $|S| > q^h$, and thus for these terms we get a saving of $q^{-h/2}$ and the bound for $X_1$ applies. If not, then $k$ is even and the $R_i$'s must be equal in pairs and otherwise disjoint. Let $R = UV$ as above, where $U$ is the product of irreducibles $P$ dividing exactly one pair of $R_i$'s, and $V$ is the product of all other irreducibles $P$ dividing $R$. Write $R_i = U_iV_i$, where $U_i = (R_i, U)$ and $V_i = (R_i, V)$. For fixed $U, V$, let $C(U, V)$ be the set of $R_1, \ldots, R_k$ yielding that $U$ and $V$. Then there are at most $d_{k/2}(U)$ choices for $U_2, U_4, \ldots, U_k$, and since $V_i|V$, there are at most $d(V)^{k/2}$ choices for $V_2, V_4, \ldots, V_k$. Thus $\#|C(U, V)| \leq d_{k/2}(U)d(V)^{k/2}$. Since in our application of Lemma 3.6 we have $|S| \geq |V|^{4/k}$, we have

$$
\sum_{V|Q \text{monic}} \frac{1}{|UV|} \sum_{(R_1, \ldots, R_k) \in C(U, V)} \left( \prod_{i=1}^{k} \frac{|R_i|}{\phi(R_i)} \right) X_3 \ll \sum_{V|Q \text{monic}} d_{k/2}(U)(|U|/\phi(U))^2d(V)^{k/2}(|V|/\phi(V))^k
$$

$$
= \prod_{P|Q} \left( 1 + \frac{k|P|}{2(|P| - 1)^2} + \frac{2^{k/2}(|P|/(|P| - 1))^k}{|P|^{1+2/k}} \right) \right)^{-k/2} \left( \frac{\phi(Q)}{|Q|} \right),
$$

so the $X_3$ terms contribute $\ll |Q|(q^h)^{k/2} \left( \frac{\phi(Q)}{|Q|} \right)^{k/2}$, as desired. \hfill \Box

The final contribution of $X_3$ only arises when $k$ is even, so when $k$ is odd we have the estimate

$$
m_k(Q; h) \ll |Q|((q^h)^{k/2-1/2} + (q^h)^{k/2-1/(k-2)}) \left( \frac{\phi(Q)}{|Q|} \right)^{k/2-2^k}.
$$

For $k = 3$ this implies that

$$
m_3(Q; h) \ll |Q|q^h \left( \frac{\phi(Q)}{|Q|} \right)^{-5}.
$$

In the case when $k = 5$, we can bound $m_5(Q; h)$ via a more involved argument.

4. The Fifth Moment of Reduced Residues in the Function Field Setting:

Proof of Theorem 1.4

In this section we derive a stronger bound for $m_5(Q; h)$ than what is given in Lemma 3.7, which implies that

$$
m_5(Q; h) \ll |Q|(q^h)^{13/6} \left( \frac{\phi(Q)}{|Q|} \right)^{-29.5}
$$

We will prove that for all $\varepsilon > 0$,

$$
m_5(Q; h) \ll |Q|(q^h)^{2+\varepsilon},
$$

where Conjecture 1.1 would predict a bound where the power of $q^h$ is 2.

We would like to apply Lemma 3.6 to bound the size of $S(R_1, \ldots, R_5)$. But when applying this lemma, we can freely choose which of the $R_i$'s plays the role of $R_1$ and $R_2$. If any $R_i$ satisfies $|R_i| < q^h$, the choice is immaterial, so assume that $|R_i| \geq q^h$ for all $i$. If there is any triple $R_i, R_j, R_k$ with $R_i = R_j = R_k$, pick $R_1 = R_i$ and $R_2 = R_j$. In this case $X_2$ will
have no contribution, and $X_3$ and $X_1$ will each be $\widetilde{O}(q^{-h/2})$. If there is no such triple, but
there exists $R_i \neq R_j$ with either $\left| \frac{R_i}{(R_i, R_j)} \right| < q^h$, or $\left| \frac{R_i}{(R_i, R_j)} \right| \geq q^h$ and $|(R_i, R_j)| \geq q^{h/2}$, then we choose $R_1 = R_i$ and $R_2 = R_j$. In this case we have $X_3 = 0$ and $X_1, X_2$ each contributing
$\widetilde{O}(q^{-h/2})$. So, it remains to bound what happens in the remaining cases. We first show that
in those cases, up to some reordering, certain factors of $R_2$ and $R_3$ are bounded below.

**Lemma 4.1.** Let $(R_1, \ldots, R_5)$ be a tuple of divisors of $Q$ such that
- $|R_i| \geq q^h$ for all $i$,
- no three $R_i$’s are equal,
- for any $R_i, R_j$, either $R_i = R_j$, or $\left| \frac{R_i}{(R_i, R_j)} \right| \geq q^h$ and $|(R_i, R_j)| < q^{h/2}$, and
- $R_1, R_2, R_3$ are all distinct.

Then
- $\left| \frac{R_2}{(R_1, R_2)} \right| \geq q^h$, and
- $\left| \frac{R_3}{(R_1, R_2)} \right| \geq q^{h/2}$.

**Remark.** The bound on $\left| \frac{R_3}{(R_1, R_2)} \right|$ above is worse than the bound on $\left| \frac{R_2}{(R_1, R_2)} \right|$. In order to
apply Lemma 4.3 below, we will need both of them to be at least of size $q^{h/2}$, so the bound
on $\left| \frac{R_2}{(R_1, R_2)} \right|$ is better than necessary.

However, the fact that these bounds get worse is precisely what prevents us from applying
our technique to bound higher moments. If instead we applied the same argument to a 7-tuple $(R_1, \ldots, R_7)$ of divisors of $Q$, we would not be able to guarantee that $\left| \frac{R_3}{(R_4, R_5, R_6, R_7)} \right| \geq q^{h/2}$,
even if we weaken the conditions to allow reordering.

**Proof.** The first conclusion, that $\left| \frac{R_2}{(R_1, R_2)} \right| \geq q^h$, follows directly from the third assumption,
since we have $R_1 \neq R_2$.

For the second conclusion, let $R_{123} = (R_1, R_2, R_3)$ and let $R_{13} = \frac{(R_1 R_3)}{(R_1, R_2 R_3)}$ and $R_{23} = \frac{(R_2 R_3)}{(R_1 R_2 R_3)}$ so that $R_{13}$ is the product of all primes dividing $R_1$ and $R_3$ but not $R_2$, and $R_{23}$ is
analogous. Then $(R_3, R_1 R_2) = R_{13} R_{23} R_{123}$. By assumption, $|(R_2, R_3)| < q^{h/2}$, so $|R_{23} R_{123}| < q^{h/2}$, and in particular $|R_{23}| < q^{h/2}$. Now assume by contradiction that $\left| \frac{R_3}{(R_3, R_1, R_2)} \right| < q^{h/2}$. Then
$$\left| \frac{R_3}{(R_1, R_3)} \right| = \left| \frac{R_3}{R_{13} R_{23} R_{123}} \right| = \left| \frac{R_3}{R_{13} R_{23} R_{123}} \right| \cdot |R_{23}| < \frac{q^{h/2}}{q^{h/2}} = q^h,$$
which contradicts the assumption because $R_1 \neq R_3$. \qed

The following auxiliary lemma provides a standard bound on $\tau_k$ in the function field
setting. We will also use that $\phi(F) \gg \frac{|F|}{\log \log |F|}$ for all $F \in \mathbb{F}_q[T]$.

**Lemma 4.2.** Let $M = \max_{b \geq 1} (\tau_k(T^b))^{1/b}$. For any fixed $k$,
$$\limsup_{\deg F \to \infty} \frac{\log \tau_k(F) \log \log |F|}{\log |F|} = \log M,$$
and thus for all $\varepsilon > 0$, $\tau_k(F) = O_\varepsilon(|F|^\varepsilon)$. 26
Proof. The proof of the above lemma follows closely along the lines of Shiu [14]. We will show one direction of the statement, adapted to our setting; the other direction also follows very closely, so we omit it. Note first that
\[ 1 \leq (\tau_k(T^b))^{1/b} = \left( \frac{b + k - 1}{b} \right)^{1/b} < \left( \frac{(b + k - 1)e}{k - 1} \right)^{(k-1)/b} \rightarrow 1 \]
as \( b \to \infty \), so \( M \) exists.

We now show that \( \lim \sup \frac{\log \tau_k(F) \log \log |F|}{\log |F|} \geq \log M \). Fix \( b \) such that \( \tau_k(T^b) = M^b \).

Let
\[ F = \prod_{\deg P = d \atop P \text{ irredu}} P^b, \]
so that \( \tau_k(F) = \prod_{\deg P = d} \tau_k(P^b) = (\tau_k(T^b))^\pi(d;\mathbb{F}_q) = M^{b\pi(d;\mathbb{F}_q)} \). We have that \( \pi(d;\mathbb{F}_q) \sim \frac{q^d}{d} \) as \( d \to \infty \), so that
\[ \log |F| = bd \log q \pi(d;\mathbb{F}_q) \sim bq^d \log q, \]
and
\[ \log \log |F| = d \log q + O(1). \]

Thus as \( d \to \infty \),
\[ \log \tau_k(F) = b \pi(d;\mathbb{F}_q) \log M \sim b \log M \cdot \frac{q^d}{d} \sim \frac{\log M \log |F|}{\log \log |F|}, \]
so \( \lim \sup \frac{\log \tau_k(F) \log \log |F|}{\log |F|} \geq \log M \).

As mentioned above, \( \lim \sup \frac{\log \tau_k(F) \log \log |F|}{\log |F|} \leq \log M \) also follows Shiu’s proof in [14] closely.

The above bound implies that for all \( \varepsilon > 0 \), \( \tau_k(F) = |F|^{O(1/\log \log |F|)} = O_\varepsilon(|F|^\varepsilon) \).

Here we have a final preparatory lemma before the main proposition leading to the bound on \( m_\zeta(Q;h) \). In what follows, our main strategy will be carefully isolating factors of the \( R_i \)'s in order to bound the number of terms in our sum. In doing so, we will make use of the following bound.

Lemma 4.3. Let \( Q \) be a squarefree polynomial, and let \( n \in \mathbb{N}_{\geq 2} \). Let \( \mathcal{I} \subseteq \mathbb{F}_q(t) \) be an interval of size \( q^{-h} \). That is to say, for some rational function \( \alpha \in F_q(T) \), let \( \mathcal{I} := \{ \beta \in \mathbb{F}_q(T) : |\alpha - \beta| < q^{-h} \} \). Assume in the following that \( X_i,Y_i \in \mathbb{F}_q[T] \) for all \( i \). Then
\[ \sum_{\substack{Y_1,...,Y_n|Q \\mod Y_i \atop (X_i,Y_i)=1, \sum_{X_i/Y_i \in \mathcal{I}}} \frac{\mu((\prod_i Y_i)^2}{\prod_i \phi(Y_i)^2} \ll_h q^{-h} \sum_{m=h/2} \frac{2h}{m} q^{m/\log m} (\log m)^2. \]

Proof. For given \( X_1,...,X_n \) and \( Y_1,...,Y_n \), let \( X \) and \( Y \) be defined so that \( Y = \prod_i Y_i \) and \( X = \sum_i \frac{X_i}{Y_i} \). Then for all tuples considered in the sum, \( \frac{X}{Y} \in \mathcal{I} \) and \( q^{h/2} \leq |Y| \leq q^{2h} \). We’ll proceed by counting the number of possibilities for \( \frac{X}{Y} \) satisfying this constraint, which is bounded above by the number of points in \( \mathcal{I} \) with denominator smaller than \( q^{2h} \), and
finally count the number of ways of splitting $Y$ up into $Y_1, \ldots, Y_n$. However, we want to also consider the weighting in the sum of $\frac{1}{\phi(Y)^2}$, so we start by splitting the sum up into different sizes of $Y$, and then applying bounds on $\phi(Y)$.

To begin with, we rewrite the sum in terms of $X$ and $Y$. Note that all $Y_i$ in our sum are relatively prime, because of the Möbius factor. Thus $Y$ is squarefree and $\phi(Y) = \prod_i \phi(Y_i)$. Moreover, a choice of $X, Y$, and a decomposition $Y = Y_1 \cdots Y_n$ determines $X_i$ for each $i$ by the Chinese Remainder Theorem. Our sum is thus equal to

$$\sum_{Y \mid Q} \sum_{X \in \mathcal{R}(Y)} \frac{\mu(Y)^2}{\phi(Y)^2} \# \{Y_1, \ldots, Y_n : Y_1 \cdots Y_n = Y \}.$$ 

We now split the sum up according to $|Y|$, defining $m := \deg Y$. The sum is then equal to

$$\sum_{m=h/2}^{2h} \sum_{Y \mid q^m} \sum_{X \in \mathcal{R}(Y)} \frac{\mu(Y)^2}{\phi(Y)^2} \# \{Y_1, \ldots, Y_n : Y_1 \cdots Y_n = Y \}$$

$$\ll_n \sum_{m=h/2}^{2h} q \frac{m}{\log m} \left( \frac{\log m}{q^m} \right)^{-2} \# \{X, Y : X / Y \in \mathcal{I}, |Y| = q^m \},$$

by Lemma 4.2 and the fact that $\phi(Y)^{-2} \ll \left( \frac{|Y|}{\log \log |Y|} \right)^{-2}$. Note that we can further relax the condition that $|Y| = q^m$ to the condition that $|Y| \leq q^m$. The number of $X/Y$ in this interval is $q^{2m-h} + O(1)$; since $m \geq h/2$, $1$ is $O(q^{2m-h})$, so the number of points in the interval is $O(q^{2m-h})$. Thus the sum is

$$\ll \sum_{m=h/2}^{2h} q \frac{m}{\log m} \left( \frac{\log m}{q^m} \right)^{-2} q^{2m-h}$$

$$\ll q^{-h} \sum_{m=h/2}^{2h} q \frac{m}{\log m} (\log m)^2,$$

as desired. 

Note that the bound in the lemma is $O(\varepsilon q^{-h+\varepsilon h})$ for all $\varepsilon > 0$.

We now turn to bounding the contribution to the fifth moment $m_5(Q; h)$ coming from tuples $(R_1, \ldots, R_5)$ satisfying the conclusions of Lemma 4.1.

**Proposition 4.4.** Let $S$ be the set of tuples $(R_1, \ldots, R_5)$ such that

- $R_i \mid Q$ for all $i$,
- $q^h \leq |R_i| \leq q^{2h}$ for all $i$,
- $\left| \frac{R_3}{(R_1, R_2)} \right| \geq q^{h/2}$, and
- $\left| \frac{R_1}{(R_3, R_1 R_2)} \right| \geq q^{h/2}$.


Then for all \( \varepsilon > 0 \),
\[
\sum_{(R_1, \ldots, R_5) \in S} \prod_{i=1}^{5} \frac{1}{\phi(R_i)} \sum_{A_i \mod \mathbb{R}_i} q^{5h} \ll q^{2h+\varepsilon} \frac{|Q|}{\phi(Q)}
\]

Proof. We’ll begin by sketching an overview of the strategy. For each subset \( I \subseteq [5] \), let \( R_I = \prod_{P|R_i \forall i \in I} P \) be the product of the irreducible factors dividing \( R_i \) if and only if \( i \in I \). Note that these \( R_I \)'s must be pairwise relatively prime.

We start by using the constraint that \( \left| \frac{A_i}{R_i} \right| < q^{-h} \). We will count the total number of rational functions in this interval with denominator of degree at most \( 2h \). For each option of \( \frac{A_i}{R_i} \), we can decompose \( R_1 = \prod_{I \neq 1} R_I \), so the number of ways to decompose \( R_1 \) into these \( R_i \)'s factors is \( \tau_{2k-1-1}(R_1) \), which we can bound based on the degree of \( R_1 \). We then also get \( \frac{A_1}{R_1} = \sum_{I \neq 1} \frac{A_i}{R_i} \), where the \( A_i \)'s are determined by the Chinese Remainder Theorem.

We’ll then focus on the constraint that \( \left| \frac{A_i}{R_i} \right| < q^{-h} \). However, \( (R_1, R_2) = \prod_{1 \neq I} R_I \) has already been fixed, so we’ll perform the same analysis on the remaining factors of \( R_2 \). Crucially, \( \frac{R_2}{(R_1, R_2)} \) remains relatively large by assumption, which will ensure that we save enough by doing this. Similarly we’ll then turn to the constraint on \( \frac{A_3}{R_3} \), using our assumption that \( \frac{R_3}{(R_3, R_4, R_5)} \) is large enough.

We begin by rewriting our sum in terms of the \( R_I \). Again, for each subset \( I \subseteq [5] \), and for a fixed \( R_1, \ldots, R_5 \), we will define \( R_I \) to be the product of all primes \( P \) so that \( P \) divides \( R_i \) for each \( i \in I \) and \( P \) does not divide \( R_j \) for all \( j \notin I \). So, for example, \( R_{1,2} \) is the product of all primes dividing \( R_1 \) and \( R_2 \), but \( (R_{1,2}, R_3) = 1 \) for \( j = 3, 4, 5 \). This also implies that \( (R_I, R_J) = 1 \) for each \( I \neq J \subseteq [5] \). In order for the sum over \( A_i \) to be nonempty, \( R_I = 1 \) unless \( |I| \geq 2 \), i.e. each irreducible polynomial dividing an \( R_i \) must divide at least two of them, so from now on we will always assume that \( |I| \geq 2 \). Moreover, each choice of \( A_i \) is equivalent to a choice of \( A_{i,I} \) for all subsets \( I \) containing \( i \), so that \( \frac{A_i}{R_i} = \sum_{I \ni i} \frac{A_i}{R_I} \). Moreover, \( (A_i, R_i) = 1 \) for all \( i \) if and only if \( (A_{i,I}, R_I) = 1 \) for all \( i, I \). The remaining two conditions on each choice of \( A_i \) are that for all \( i \), \( |A_i/R_i| < q^{-h} \) and that \( \sum_{i=1}^{5} A_i/R_i = 0 \). These translate, respectively, to the conditions that

\[
\forall i, \left| \sum_{I \ni i} \frac{A_{i,I}}{R_I} \right| < q^{-h} \text{ and } \forall I, \sum_{i \notin I} A_{i,I} = 0.
\]

Lastly, consider the constraint that \( (R_1, \ldots, R_5) \in S \). This means that each \( R_I \) must divide \( Q \), and for all \( i \), we have \( q^h \leq \left| \prod_{I \ni i} R_I \right| \leq q^{2h} \). In order to address the last two constraints, define \( \ell_I \) to be the minimum element of \( I \). Then \( \frac{R_2}{(R_1, R_2)} = \prod_{I \neq 2} R_I \), and \( \frac{R_3}{(R_3, R_4, R_5)} = \prod_{I \neq 3} R_I \), so that these constraints become

\[
\left| \prod_{I \neq 2} R_I \right| \geq q^{h/2} \text{ and } \left| \prod_{I \neq 3} R_I \right| \geq q^{h/2}.
\]
The sum under consideration is then
\[ \ll q^{5h} \sum_{R_I | Q, I \subseteq [5]} \frac{\mu (\prod_I R_I)^2}{\prod_I \phi(R_I)^{|I|}} \sum_{I, i \in I, A_{i,I} \in \mathcal{R}(R_I)} \frac{1}{\forall i | \sum_{i \leq j} A_{i,I} R_I < q^{-h}} \]

For each \( I \), if \( m_I \) is the largest index contained in \( I \), \( A_{m_I,I} = -\sum_{i < m_I} A_{i,I} \), so if there is a valid choice of \( A_{m_I,I} \), it is completely determined by the other \( A_{i,I} \) values. Thus we can bound our expression above by excluding \( A_{m_I,I} \) whenever \( m_I \) is the largest element of \( I \).

As before, let \( \ell_I \) be the smallest index contained in \( I \). For all \( i \in I \), the number of choices of \( A_{i,I} \) is at most \( \phi(R_I) \), because of the constraint that \( (A_{i,I} R_I) = 1 \). If \( \ell_I < i < m_I \), then we will use this bound to constrain the number of choices of \( A_{i,I} \). We will still need to keep track of our conditions that
\[ \forall i, \left| \sum_{J \ni i} A_{i,J} R_I \right| < q^{-h}, \]
which depend on the choices of \( A_{i,I} \) for \( i > \ell_I \). However, if we fix a choice of some of the terms in the sum above, the remaining terms are still constrained to lie in an interval of size \( q^{-h} \), centered at a possibly-nonzero polynomial. So, for fixed \( i \), and for each choice of \( A_{i,I} \) for all \( I \ni i \) with \( i > \ell_I \), we have that
\[ \left| F + \sum_{j \ni i} A_{i,J} R_I \right| < q^{-h}, \]
where \( F \) depends on our choice of the \( A_{i,I} \). That is to say, the sum \( \sum_{j \ni i} A_{i,J} R_I / R_J \) is always constrained to lie in an interval of size \( q^{-h} \), although the exact choice of interval depends on our choice of \( A_{i,I} \). In what follows, we will only use that each of these sums must lie in an interval of size \( q^{-h} \), without ever using specific information about the interval. Thus, all of the following computations will be equivalent no matter the choice of \( A_{i,I} \), so we can factor out the number of such choices. This yields the following sum.

\[ \ll q^{5h} \sum_{R_I | Q, I \subseteq [5]} \frac{\mu (\prod_I R_I)^2}{\prod_I \phi(R_I)^{|I|}} \prod_I \phi(R_I)^{|I| - 2} \sum_{A_{i,I}, J \ni I, A_{i,I} R_I / R_J < q^{-h}} \frac{1}{\forall i | \sum_{i \leq j} A_{i,I} R_I < q^{-h}} \]

where we are using the interval centered at 0 to denote our constraint that \( \sum_{\ell_J=i} A_{i,J} R_J \) must lie in some interval of size \( q^{-h} \).

Since our only terms \( A_{i,I} \) remaining in our sum are of the form \( A_{\ell_I,I} \), there is only one per subset \( I \), so to simplify our notation we will write \( A_I := A_{\ell_I,I} \) from now on.
Now we will consider subsets $I$ with $\ell_I = 4$. There is only one of these, namely $\{4, 5\}$, so we rewrite the sum as follows:

$$\ll q^{5h} \sum_{R \subseteq [5], I \neq \{4, 5\}} \frac{\mu(\prod_{i} R_i)^2}{\prod_{i} \phi(R_i)^2} \sum_{A \in R(R_i)} \frac{1}{\phi(R(4, 5))^2}$$

In the inside sum, we have dropped the additional constraint that $A_{4, 5}/R_{4, 5}$ must lie in an interval of size $q^{-h}$, since ignoring it only increases the size of the sum. For each $R_{4, 5}$, there are $\phi(R_{4, 5})$ choices of $A_{4, 5}$, so the inner sum becomes

$$\sum_{R_{4, 5}|Q} \frac{1}{\phi(R_{4, 5})} = \frac{|Q|}{\phi(Q)}$$

since $Q$ is squarefree.

Now we will consider subsets $I$ with $\ell_I = 3$, i.e. $\{3, 4\}$, $\{3, 4, 5\}$, and $\{3, 5\}$. We first bookkeep by isolating these terms in the sum, yielding

$$\ll q^{5h} \frac{|Q|}{\phi(Q)} \sum_{R_i|Q} \frac{\mu(\prod_{i} R_i)^2}{\prod_{i} \phi(R_i)^2} \sum_{A_i \in R(R_i)} \frac{1}{\phi(R_i)}$$

We now bound the inner sum using Lemma 4.3. The inner sum comprises three terms $R_i$, so apply the lemma with $n = 3$, to get that the inner sum is $\ll q^{-h} o(q^{eh})$. We repeat the process, now considering subsets $I$ with $\ell_I = 2$. Isolating these terms yields

$$\ll q^{2h} o(q^{eh}) \frac{|Q|}{\phi(Q)} \sum_{R_i|Q} \frac{\mu(\prod_{i} R_i)^2}{\prod_{i} \phi(R_i)^2} \sum_{A_i \in R(R_i)} \frac{1}{\phi(R_i)}$$

Here there are seven $R_i$ terms and seven $A_i$ terms in the inner sum, so, again applying Lemma 4.3, the inner sum is $\ll q^{-h} o(q^{eh})$. Lastly, we address the terms with $\ell_I = 1$:

$$\ll q^{3h} q^{2h} \frac{|Q|}{\phi(Q)} \sum_{R_i|Q} \frac{\mu(\prod_{i} R_i)^2}{\prod_{i} \phi(R_i)^2} \sum_{A_i \in R(R_i)} \frac{1}{\phi(R_i)}$$

We apply Lemma 4.3 one final time, this time with $n = 15$, since there are 15 sets $I \subseteq [5]$ with $|I| \geq 2$ and $\ell_I = 1$. This yields

$$\ll q^{2h} q^{3(h)} \frac{|Q|}{\phi(Q)}$$
as desired.

**Theorem 4.5.** For all $\varepsilon > 0$, for $Q \in \mathbb{F}_q[T]$ squarefree,

$$m_5(Q; h) \ll |Q|q^{2h+\varepsilon} \left( \frac{|Q|}{\phi(Q)} \right)^{27}.$$ 

**Proof.** Using Lemma 3.1, we can express

$$m_5(Q; h) = |Q| \left( \frac{\phi(Q)}{|Q|} \right)^5 V_5(Q; h),$$

where

$$V_5(Q; h) = \sum_{R_1, \ldots, R_5 | Q} \prod_{i=1}^5 \frac{\mu(R_i)}{|R_i|} \sum_{\substack{A_1, \ldots, A_5 \text{ mod } R_i \atop (A_i, R_i) = 1 \atop \sum_i A_i = 0}} E \left( \frac{A_1}{R_1} \right) \cdots E \left( \frac{A_5}{R_5} \right).$$

Now apply Lemma 3.6 to bound the contribution to $V_5(Q; h)$ from many tuples $R_1, \ldots, R_5$. If $|R_i| < q^h$ for any $i$, then these terms contribute 0; assume from now on that $|R_i| \geq q^h$. If for any triple $i, j, k$ we apply Lemma 3.6 with $R_i = R_i$ and $R_2 = R_j$; in this case $X_2 = 0$ and $X_1$ and $X_3$ are $O(q^{-h/2})$, so these terms contribute $O \left( q^h \left( \frac{|Q|}{\phi(Q)} \right)^{32} \right)$. If there exist $R_i \neq R_j$ such that either $\left| \frac{R_i}{(R_i, R_j)} \right| < q^h$ or $|(R_i, R_j)| \geq q^{h/2}$; in this case, $X_3 = 0$, and $X_1$ and $X_2$ are each $O(q^{-h/2})$, so these terms contribute $O \left( q^h \left( \frac{|Q|}{\phi(Q)} \right)^{32} \right)$ as well.

Assume now that $(R_1, R_2, R_3, R_4, R_5)$ does not fall into either of the above cases. Then for all $i$, $|R_i| < q^h$. To see this, assume that $(R_1, R_2, R_3, R_4, R_5)$ has no $i, j, k$ with $R_i = R_j = R_k$, and that for all $R_i \neq R_j$, $\left| \frac{R_i}{(R_i, R_j)} \right| \geq q^h$ and $|(R_i, R_j)| < q^{h/2}$. Assume, relabeling if necessary, that $R_j \geq q^h$. Since $R_1 | \prod_{j \neq 1} (R_1, R_j)$, we must have $|(R_1, R_j)| \geq q^{h/2}$ for some $j \neq 1$. This cannot be true for some $j$ with $R_j \neq R_1$, so we have $R_j = R_1$. At the same time, there can be only one $j \neq 1$ with $R_j = R_1$, so without loss of generality our tuple must be of the form $(R_1, R_1, R_3, R_4, R_5)$. There cannot be an additional equal pair among $R_3, R_4, R_5$; if there is (without loss of generality $R_3 = R_4$), then $R_5 | (R_1, R_3)(R_3, R_5)$, so since $|R_5| \geq q^h$ either $|(R_1, R_3)| \geq q^{h/2}$ or $|(R_3, R_5)| \geq q^{h/2}$, which along with the lack of equal triples yields a contradiction. Now consider $R_3$. Note that $R_3 | (R_1, R_3)(R_4, R_3)(R_5, R_3)$, and that $\left( \frac{R_3}{(R_1, R_3)} \right) | (R_4, R_3)(R_5, R_3)$. But by assumption, $\left| \frac{R_3}{(R_1, R_3)} \right| \geq q^h$ and $|(R_4, R_3)(R_5, R_3)| < (q^{h/2})^2 = q^h$, which yields a contradiction.

So, the only terms remaining are those with $|R_i| < q^h$ for all $i$, no equal triple, and either $\left| \frac{R_i}{(R_i, R_j)} \right| < q^h$ or $|(R_i, R_j)| \geq q^{h/2}$ whenever $R_i \neq R_j$. By Lemma 4.1, $(R_1, \ldots, R_5)$ satisfies the constraints of Proposition 4.4. By Proposition 4.4, these terms contribute $O \left( q^{2h+\varepsilon} \frac{|Q|}{\phi(Q)} \right)$ to $V_5(Q; h)$ for all $\varepsilon > 0$. Thus for all $\varepsilon > 0$,

$$V_5(Q; h) \ll q^{2h+\varepsilon} \frac{|Q|}{\phi(Q)} + q^h \left( \frac{|Q|}{\phi(Q)} \right)^{32},$$

so $m_5(Q; h) \ll |Q|q^{2h+\varepsilon} \left( \frac{|Q|}{\phi(Q)} \right)^{-4} + |Q|q^{2h} \left( \frac{|Q|}{\phi(Q)} \right)^{27}$. □
As in the integer case, we particularly want to consider \( Q \) to be the product of irreducible polynomials \( P \) with \( |P| \leq q^{2h} \). In this case, \( \frac{|Q|}{\phi(Q)} \ll h \), so that we get the following corollary.

**Corollary 4.6.** Fix \( \varepsilon > 0 \) and let \( Q \in \mathbb{F}_q[T] \) be given by \( Q = \prod_{|P| \leq q^{2h}} P \). Then

\[
m_5(Q; h) \ll |Q|^{2h + \varepsilon}.
\]

4.1. **Proof of Corollary 1.5: Bounds on** \( R_k(q^h) \). In this subsection, we discuss the transition from bounds on \( V_k(Q; h) \), from Theorem 1.4 and Lemma 3.7, to bounds on sums of singular series in function fields. Much of this is similar to the integer case discussion in Section 2.

As in the integer case, for \( \mathcal{D} = \{D_1, \ldots, D_k\} \) a set of distinct polynomials in \( \mathbb{F}_q[T] \), we define the singular series

\[
\mathfrak{S}(\mathcal{D}) := \prod_{P \text{ monic, irred.}} \frac{(1 - \nu_P(\mathcal{D})/|P|)}{(1 - 1/|P|)^k} = \sum_{R_1, \ldots, R_k} \left( \prod_{i=1}^k \frac{\mu(R_i)}{\phi(R_i)} \right) \sum_{A_1, \ldots, A_k} e \left( \sum_{i=1}^k \frac{A_i D_i}{R_i} \right),
\]

where \( \nu_P(\mathcal{D}) \) is the number of equivalence classes of \( \mathbb{F}_q[T]/(P) \) occupied by elements of \( \mathcal{D} \). We also define \( \mathfrak{S}_0(\mathcal{D}) \), given by \( \mathfrak{S}_0(\mathcal{D}) := \sum_{J \subseteq \mathcal{D}} (-1)^{|J|} \mathfrak{S}(J) \), and consider

\[
R_k(q^h) := \sum_{D_1, \ldots, D_k \text{ distinct}} \mathfrak{S}_0(\{D_1, \ldots, D_k\}).
\]

Our results on \( m_k(Q; h) \) (and equivalently \( V_k(Q; h) \)) imply bounds on these sums of \( k \)-fold singular series, just as in the integer case in Section 2. We set \( Q \) to be the product of all monic irreducible polynomials of degree at most \( 2h \), so that \( \frac{|Q|}{\phi(Q)} \ll q \). Just as in the integer case, we can truncate the expression for \( \mathfrak{S}_0(\mathcal{D}) \) to only contain terms dividing \( Q \), with an acceptable error term. In particular, we get

\[
R_k(h) = \sum_{D_1, \ldots, D_k \text{ distinct}} \sum_{|D_i| \leq q^h} \prod_{i=1}^k \frac{\mu(R_i)}{\phi(R_i)} \sum_{A_1, \ldots, A_k} e \left( \sum_{i=1}^k \frac{D_i A_i}{R_i} \right) + O(1).
\]

It will again be helpful for us to define the singular series of a \( k \)-tuple \( \mathcal{D} = (D_1, \ldots, D_k) \) relative to the modulus \( Q \). Here the \( k \)-tuple can have repeated elements; since the Euler product is finite, convergence is not a concern. We define

\[
\mathfrak{S}(\mathcal{D}; Q) := \prod_{P \text{ monic}} \frac{(1 - \nu_P(\mathcal{D})/|P|)}{(1 - 1/|P|)^k} = \sum_{R_1, \ldots, R_k \text{ monic}} \left( \prod_{i=1}^k \frac{\mu(R_i)}{\phi(R_i)} \right) \sum_{A_1, \ldots, A_k} e \left( \sum_{i=1}^k \frac{A_i D_i}{R_i} \right).
\]

Note that if \( \mathcal{D} \) has a repeated element, so that \( \mathcal{D} = \{D, D, D_3, \ldots, D_k\} \), then \( \mathfrak{S}(\mathcal{D}; Q) = \frac{|Q|}{\phi(Q)} \mathfrak{S}(\{D, D_3, \ldots, D_k\}; Q) \), so we can remove repeated elements from \( \mathcal{D} \) at the expense of a factor of \( \frac{|Q|}{\phi(Q)} \). We define \( \mathfrak{S}_0(\mathcal{D}; Q) \) to be the alternating sum \( \sum_{J \subseteq \mathcal{D}} (-1)^{|J|} \mathfrak{S}(J; Q) \), so
we have
\[ R_k(q^h) = \sum_{D_1, \ldots, D_k \atop D_i \text{ distinct} \atop |D_i| \leq q^h} \mathcal{S}_0(\{D_1, \ldots, D_k\}; Q) + O(1). \]

This is quite close to the quantity \( V_k(Q; h) \), except with the added constraint that the \( D_i \)'s must be distinct. It suffices to remove this condition. To do so, we put \( \delta_{ij} = 1 \) if \( D_i = D_j \) and 0 otherwise, so that

\[
\sum_{D_1, \ldots, D_k \atop D_i \text{ distinct} \atop |D_i| \leq q^h} \mathcal{S}_0(\{D_1, \ldots, D_k\}; Q) = \sum_{D_1, \ldots, D_k \atop |D_i| \leq q^h} \left( \prod_{1 \leq i < j \leq k} (1 - \delta_{ij}) \right) \mathcal{S}_0(\{D_1, \ldots, D_k\}; Q).
\]

We can expand the product and group terms according to which \( D_i \)'s are required to be equal, noting that, for example, \( \delta_{12}\delta_{23} = \delta_{13}\delta_{23} \). We can also combine terms according to symmetry; the term \( \delta_{12} \) and the term \( \delta_{34} \) will have identical contributions in the final sum.

Let us now proceed with analyzing \( R_5(q^h) \). After some counting, we get that we can expand the product of \( \delta_{ij} \)'s into

\[ 1 - 10\delta_{12} + 20\delta_{12}\delta_{13} + 15\delta_{12}\delta_{34} - 20\delta_{12}\delta_{13}\delta_{45} - 30\delta_{12}\delta_{13}\delta_{14} + 24\delta_{12}\delta_{13}\delta_{14}\delta_{15}. \]

The term 1 gives us precisely \( V_5(Q; h) \), which we have already analyzed. We can then bound each of the remaining six terms by expanding \( \mathcal{S}_0 \) into a sum of \( \mathcal{S} \), removing any repeated terms in the appropriate tuple, and applying Lemma 3.7 to bound \( V_k(Q; h) \) for some \( k < 5 \). For the sake of brevity we omit most of these computations, which are very similar, but we will show that the term corresponding to \( \delta_{12} \) is \( \ll \left( \frac{|Q|}{\phi(Q)} \right)^{21/2} q^{2h} \), since all other terms are smaller.

Assume we have a tuple \( D = \{D, D, D_3, \ldots, D_k\} \), with one repeated term. As mentioned above, \( \mathcal{S}(D; Q) = \frac{|Q|}{\phi(Q)} \mathcal{S}(\{D, D_3, \ldots, D_k\}; Q) \). Expanding \( \mathcal{S}_0 \) and applying this relation shows that

\[
\mathcal{S}_0(D; Q) = \left( \frac{|Q|}{\phi(Q)} - 2 \right) \mathcal{S}_0(\{D, D_3, \ldots, D_k\}; Q) + \left( \frac{|Q|}{\phi(Q)} - 1 \right) \mathcal{S}_0(\{D_3, \ldots, D_k\}; Q),
\]

so in this way we can remove repeated elements from our sum. The term we want to bound is

\[
-10 \sum_{D_1, D_2, D_3, D_4, D_5 \atop |D_i| \leq q^h} \mathcal{S}_0(\{D_1, D_1, D_3, D_4, D_5\}; Q)
\]

\[
= -10 \sum_{D_1, D_2, D_3, D_4, D_5 \atop |D_i| \leq q^h} \left( \frac{|Q|}{\phi(Q)} - 2 \right) \mathcal{S}_0(\{D_1, D_3, D_4, D_5\}; Q) + \left( \frac{|Q|}{\phi(Q)} - 1 \right) \mathcal{S}_0(\{D_3, D_4, D_5\}; Q)
\]

\[
= -10 \left( \frac{|Q|}{\phi(Q)} - 2 \right) V_4(Q; h) + q^h \left( \frac{|Q|}{\phi(Q)} - 1 \right) V_3(Q; h)
\]

\[
\ll \left( \frac{|Q|}{\phi(Q)} \right)^3 q^{2h} + \left( \frac{|Q|}{\phi(Q)} \right)^{21/2} q^{2h},
\]

where in the last step the bounds follow from Lemma 3.7.
All other terms are smaller, so we have shown the following corollary.

**Corollary 4.7.** Let $Q = \prod_{|P| \leq q^h} P$. For all $\varepsilon > 0$,

$$R_5(q^h) \ll V_5(Q; h) + \left( \frac{|Q|}{\phi(Q)} \right)^{21/2} q^{2h} \ll q^{2h+\varepsilon}.$$ 

Performing the same analysis when $k = 3$ yields the bound

**Corollary 4.8.** Let $Q = \prod_{|P| \leq q^h} P$. Then

$$R_3(q^h) \ll V_3(Q; h) + q^h \left( \frac{|Q|}{\phi(Q)} \right)^2 \ll q^h \left( \frac{|Q|}{\phi(Q)} \right)^{19/2}.$$ 

5. **Numerical Evidence for Odd Moments**

Here we present several charts supporting our conjectures on the sizes of the odd moments. To begin with, we have computed $\frac{1}{6} R_3(h) = \sum_{1 \leq d_1 < d_2 < d_3 \leq h} S_0(\{d_1, d_2, d_3\})$. Below, $\frac{1}{6} R_3(h)$ is plotted in black. We expect $R_3(h)$, and thus also $\frac{1}{6} R_3(h)$, to be of the shape $Ah(\log h)^2$, for some constant $A$. We found an experimental best fit value of $A = 0.373727$, and for this $A$ have plotted $Ah(\log h)^2$ alongside $\frac{1}{6} R_3(h)$, as a dashed red line.

![Figure 1](image)

**Figure 1.** $\frac{1}{6} R_3(h)$ for $3 \leq h \leq 20000$

The fit of the theoretical red dashed curve is quite close, but there are lower-order fluctuations; below we plot the difference between $\frac{1}{6} R_3(h)$ and $Ah(\log h)^2$. 


Our analysis above includes relatively little discussion about the moments of the distribution of primes themselves. We have computed several third, fifth, and seventh moments of the distribution of primes. Specifically, we have computed

\[ \tilde{M}_k(N; N^\delta) = \frac{1}{N} \sum_{n=N}^{2N} (\psi(n + N^\delta) - \psi(n) - N^\delta)^k \]

for each of \( \delta = 0.25, 0.5, 0.75 \), and for each of \( k = 3, 5, 7 \). For a fixed \( \delta \) and \( k \), we plot \( \tilde{M}_k(N; N^\delta) \) for values of \( N \) ranging from 1 to \( 10^7 \), and growing exponentially.

Each of the plots below is drawn with both \( x \)- and \( y \)-axes on a logarithmic scale. We expect the \( k \)th moment to be of size approximately \( O(H^{(k-1)/2}(\log \frac{N}{H})^{(k+1)/2}) \), where \( H = N^\delta \), so to give a sense of size, for each plot, \( N^\delta(k-1/2)\log N^{1-\delta}(k+1/2) \) is plotted in dashed red. We have also plotted the reflection of the red dashed curve across the \( x \)-axis, since the odd moments are frequently negative.
6. Toy Models and Open Problems

Throughout, we have studied the sum

$$R_k(h) = \sum_{q_1, \ldots, q_k, 1 < q_i} \left( \prod_{i=1}^{k} \frac{\mu(q_i)}{\phi(q_i)} \right) \sum_{a_1, \ldots, a_k, 1 \leq a_i \leq q_i, \sum_{i, q_i} \frac{a_i}{q_i} \in Z} \prod_{i=1}^{k} E \left( \frac{a_i}{q_i} \right).$$

where $E(\alpha) = \sum_{m=1}^{h} e(m\alpha)$. The sums $E(\alpha)$ approximately detect when $\|\alpha\| \leq \frac{1}{h}$; the analogous sum in the function field case precisely detects when $\alpha$ has small degree. As a result, much of our understanding boils down to answering the following key question.
Question 6.1. Let $\delta > 0$ and let $Q > 1/\delta$. What is

$$\# \left\{ q_1, \ldots, q_k \in [Q,2Q], a_i \mod q_i : \left\| \frac{a_i}{q_i} \right\| \leq \delta, \sum_i a_i \mod q_i \in \mathbb{Z} \right\}?$$

We conjecture that the answer to this question is as follows.

Conjecture 6.2. Let $\delta > 0$ and let $Q > 1/\delta$. Let $S$ be the size of the set in Question 6.1. Then for any $\epsilon > 0$,

$$S \ll \begin{cases} Q^{k+\epsilon} \delta^k/2 & k \text{ even} \\ Q^{k+\epsilon} \delta^{(k+1)/2} & k \text{ odd} \end{cases}.$$ 

As we discussed in the introduction, Montgomery and Vaughan [13] considered the related problem of moments of reduced residues modulo $q$. Their work depends on the following answer to Question 6.1 above.

Theorem 6.3. Let $S$ be the size of the set in Question 6.1. Then

$$S \ll \begin{cases} \delta^{k/2} \sum_{Q \leq r_i \leq 2Q} \frac{r_i^2-r_i^2}{\text{lcm}(r_i)} & k \text{ even} \\ \delta^{k/2-1/2k} \sum_{1 \leq i \leq k \leq 2Q} \frac{r_i - r_i}{\text{lcm}(r_i)} & k \text{ odd} \end{cases}.$$ 

The proof of the above theorem is identical to the proof in [13]. This agrees with Conjecture 6.2 for the case when $k$ is even, but gives a weaker bound when $k$ is odd.

We can also consider generalizations of Question 6.1. For example, instead of specifying that $\left\| \frac{a_i}{q_i} \right\| \leq \delta$, we may ask that it lie in any specified interval.

Question 6.4. Let $Q > 1/\delta$ and let $I_1, \ldots, I_k$ be $k$ intervals in $[0,1]$ with $|I_j| \geq \delta$ for all $j$. What is

$$\# \left\{ q_1, \ldots, q_k \in [Q,2Q], a_i \mod q_i : \left\| \frac{a_i}{q_i} \right\| \in I_i, \sum_i a_i \mod q_i \in \mathbb{Z} \right\}?$$

Answers to these questions would give us more refined understanding of sums of singular series. The conjectures above are related to sums over $\mathcal{G}(\{h_1, \ldots, h_k\})$, where each $h_i$ lies in the same interval $[0,h]$. We can instead ask about sums of singular series restricted to arbitrary intervals, or along arithmetic progressions. We state the following questions using smooth cutoff functions as opposed to intervals.

Question 6.5. Let $\Phi_1, \ldots, \Phi_k$ be smooth functions with compact support on $\mathbb{R}$, and let $H \in \mathbb{R}_{>0}$. What is

$$\sum_{h_1, \ldots, h_k \in \mathbb{Z}} \mathcal{G}_0(\{h_1, \ldots, h_k\}) \Phi_1 \left( \frac{h_1}{H} \right) \cdots \Phi_k \left( \frac{h_k}{H} \right)?$$

Question 6.6. Let $\Phi_1, \ldots, \Phi_k$ be smooth functions with compact support on $\mathbb{R}$, and let $H \in \mathbb{R}_{>0}$. For arithmetic progressions $a_1 \mod q_1, \ldots, a_k \mod q_k$, what is

$$\sum_{h_1, \ldots, h_k \in \mathbb{Z}} \mathcal{G}_0(\{h_1, \ldots, h_k\}) \Phi_1 \left( \frac{h_1}{H} \right) \cdots \Phi_k \left( \frac{h_k}{H} \right)?$$
Question 6.5 addresses the correlations of \( \psi(x + h) - \psi(x) \) and \( \psi(x + h_1 + h) - \psi(x + h_1) \); in other words, the correlations of the number of primes in intervals in different places. Question 6.6 addresses the correlations of the number of primes in distinct arithmetic progressions. For both of these questions, the main term ought to come from diagonal terms where \( h_1 = h_2 \), for example, thus collapsing the weight function, whereas the error term ought to arise from off-diagonal contributions.

In the case when \( k = 2 \), Question 6.6 has been widely studied in the context of prime number races. The “Shanks-Rényi prime number race” is the following problem: let \( \pi(x; q, a) \) denote the number of primes \( p \leq x \) with \( p \equiv a \mod q \). Then for any \( n \)-tuple \((a_1, \ldots, a_n)\) of equivalence classes mod \( q \) that are relatively prime to \( q \), will we have the ordering

\[
\pi(x; q, a_1) > \pi(x; q, a_2) > \cdots > \pi(x; q, a_n)
\]

for infinitely many integers \( x \)? Many aspects of this question have been studied; see for example the expositions of Granville and Martin [7], and Ford and Konyagin [4].

In [3], Ford, Harper, and Lamzouri show that, although any ordering appears infinitely often, for \( n \) large with respect to \( q \), the prime number races among orderings can exhibit large biases. They rely on the fact that counts of primes in distinct progressions have negative correlations, which they arrange to produce a bias. This analysis is also connected to the work of Lemke Oliver and Soundararajan in [10], who use averages of two-term singular series in arithmetic progressions to show bias in the distribution of consecutive primes. It is plausible that a more precise understanding of the questions above would lead to an extension of the work of Lemke Oliver and Soundararajan.

REFERENCES

1. T. H. Chan, A note on primes in short intervals, Int. J. Number Theory 2 (2006), no. 1, 105–110. MR 2217796
2. H. Cramér, On the order of magnitude of the difference between consecutive prime numbers, Acta Arithmetica 2 (1936), no. 1, 23–46.
3. K. Ford, A. J. Harper, and Y. Lamzouri, Extreme biases in prime number races with many contestants, Math. Ann. 374 (2019), no. 1-2, 517–551. MR 3961320
4. K. Ford and S. Konyagin, Chebyshev’s conjecture and the prime number race, IV International Conference “Modern Problems of Number Theory and its Applications”: Current Problems, Part II (Russian) (Tula, 2001), Mosk. Gos. Univ. im. Lomonosova, Mekh.-Mat. Fak., Moscow, 2002, pp. 67–91. MR 1985941
5. P. X. Gallagher, On the distribution of primes in short intervals, Mathematika 23 (1976), no. 1, 4–9. MR 409385
6. A. Granville and A. Lumley, Primes in short intervals: Heuristics and calculations, arXiv:2009.05000, 2020.
7. A. Granville and G. Martin, Prime number races, Amer. Math. Monthly 113 (2006), no. 1, 1–33. MR 2202918
8. D. Hayes, The expression of a polynomial as a sum of three irreducibles, Acta Arithmetica 11 (1966), no. 4, 461–488.
9. J. P. Keating and Z. Rudnick, The variance of the number of prime polynomials in short intervals and in residue classes, Int. Math. Res. Not. IMRN (2014), no. 1, 259–288. MR 3158533
10. R. J. Lemke Oliver and K. Soundararajan, Unexpected biases in the distribution of consecutive primes, Proc. Natl. Acad. Sci. USA 113 (2016), no. 31, E4446–E4454. MR 3624386
11. H. L. Montgomery and K. Soundararajan, Beyond pair correlation, Paul Erdős and his mathematics, I (Budapest, 1999), Bolyai Soc. Math. Stud., vol. 11, János Bolyai Math. Soc., Budapest, 2002, pp. 507–514. MR 1954710
12. ______, Primes in short intervals, Comm. Math. Phys. 252 (2004), no. 1-3, 589–617. MR 2104891
13. H. L. Montgomery and R. C. Vaughan, *On the distribution of reduced residues*, Ann. of Math. (2) **123** (1986), no. 2, 311–333. MR 835765
14. P. Shiu, *The maximum orders of multiplicative functions*, Quart. J. Math. Oxford Ser. (2) **31** (1980), no. 122, 247–252. MR 576341

Department of Mathematics, Stanford University, Stanford, CA, USA  
Email address: viviank@stanford.edu