Université de Montréal

The distribution of k-tuples of reduced residues

par

Farzad Aryan

Département de mathématiques et de statistique

Faculté des arts et des sciences

Mémoire présenté à la Faculté des études supérieures

en vue de l'obtention du grade de

Maître ès sciences (M.Sc.)
en Mathématique

August

Université de Montréal

Faculté des études supérieures

Ce mémoire intitulé

The distribution of k-tuples of reduced residues

présenté par

Farzad Aryan

a été évalué par un jury composé des personnes suivantes :

Andrew Granville

(président-rapporteur)

Andrew Granville

(directeur de recherche)

Mémoire accepté le:

CONTENTS

ACKNOWLEDGMENTS

I am grateful to my advisor Andrew Granville for his support and helpful comments. Additionally, I would like to thank Dimitris Koukoulopoulos and Vorrapan Chandee for their careful analysis of my thesis and their useful suggestions. Special appreciation for the Université de Montréal's staff, my colleagues and friends, Mohammad Bardestani, Francois Charette, Dimitri Dias, Daniel Fiorilli, Kevin Henriot, and Marzieh Mehdizadeh, for their support. Finally, I would like to dedicate my thesis in loving memory of my father whom I lost last spring.

ABSTRACT

En 1940, Paul Erdős énonça une conjecture sur la distribution des classes inversibles modulo un entier. La présente thèse étudie la distribution des k-uplets de classes inversibles et propose une preuve de la conjecture d'Erdős étendue au cas des k-uplets.

INTRODUCTION

In 1936 Cramer [**1**], assuming the Riemann hypothesis (RH), showed that

$$\sum_{p_n<x}(p_{n+1}-p_n)^2 \ll x(\log x)^{3+\epsilon} \tag{0.1}$$

from which he deduced $p_{n+1}-p_n = O(\sqrt{p_n}\log p_n)$. Based on his probabilistic model for the primes he also conjectured that

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

Taking into account various sieve estimates in Cramer's probabilistic model, Granville [**2**] in 1995 conjectured that

$$\limsup_{n\to\infty}\frac{p_{n+1}-p_n}{(\log p_n)^2}\geq 2e^{-\gamma},$$

which is bigger than 1. Note that γ is the Euler constant. Proving (0.1) unconditionally seems quite deep, which led P. Erdős to make an analogous conjecture:

Conjecture (Erdős [**3**])**.** *Let q be a natural number, and let $P = \phi(q)/q$ be the probability that a randomly chosen integer is relatively prime to q. Let*

$$1 = a_1 < a_2 < \cdots$$

be the integers co-prime to q in increasing order, and let

$$V_\lambda(q)=\sum_{i=1}^{\phi(q)}(a_{i+1}-a_i)^\lambda.$$

then

$$V_2(q) \ll \phi(q)P^{-2} = qP^{-1}.$$

More generally

$$V_\lambda(q) \ll qP^{1-\lambda}.$$

For a heuristic of Erdős conjecture note that if a_i are uniformly distributed $\Big(a_i - a_{i-1} = P^{-1}\Big)$, then

$$\sum_{i=1}^{\phi(q)} (a_{i+1} - a_i)^\lambda = \phi(q)P^{-\lambda} = qP^{1-\lambda}.$$

For $\lambda < 2$ this was derived by Hooley [**4**]. Hausman and Shapiro [**5**] gave weaker upper bounds for V_2. Finally Montgomery and Vaughan [**6**] in 1986 proved the conjecture for all λ (and another easier proof appeared in the paper of Montgomery and Soundararajan [**8**]).

Investigating the distribution of prime numbers and the objects that behave like them is always an interesting subject for analytic number theorists. Also studying the behavior of subsets of prime numbers like primes in an arithmetic progression or s-tuples of primes is of huge interest. In this thesis we investigate the distribution of of s-tuples of reduced residues which in some sense are similar to s-tuples of primes and we prove the analogy of Erdős's conjecture for s-tuple reduced residues.

Let $\mathcal{D} = \{h_1, h_2, \cdots, h_s\}$ and $\nu_p(\mathcal{D})$ be the number of distinct elements in $\mathcal{D}$ mod p. $\mathcal{D}$ is called *admissible* if $\nu_p(\mathcal{D}) < p$ for all primes p. We call $a + h_1, \ldots, a + h_s$ an *s-tuple of reduced residues* if they are each coprime with q.

Theorem 0.1. *Let q be a square-free number and $\mathcal{D} = \{h_1, h_2, \cdots, h_s\}$ be a fixed admissible set of integers. Let $a_1 < a_2 < \cdots$ be those integers for which*

$a_i + h_1, \ldots, a_i + h_s$ is an s-tuple of reduced residues. Then

$$V_\lambda^{\mathcal{D}}(q) := \sum_{i=1}^{\phi_{\mathcal{D}}(q)} (a_{i+1} - a_i)^\lambda \ll \phi_{\mathcal{D}}(q) P^{-s\lambda}$$

where $\phi_{\mathcal{D}}(q) := \prod_{p|q}(p - \nu_p(\mathcal{D}))$, and the implied constant depends on $\mathcal{D}$ and λ.

The theorem follows immediately for q non-square-free as well, by considering the result for $Q = \prod_{p|q} p$. Motivated by Theorem 0.1 the analogy of this result for primes is

Conjecture. *Let $p_1, \cdots$ be the set of primes for which $p_i + h_j$ are prime for all $h_j \in \mathcal{D}$. We have*

$$\sum_{p_n < x} (p_{n+1} - p_n)^\lambda \ll_{\mathcal{D}} x(\log x)^{s(\lambda-1)+\epsilon}$$

CHAPTER 1

AN EXPONENTIAL SUM ESTIMATE

In this chapter we prove a preliminary estimate about the distribution of s-tuples of reduced residues, using exponential sums. The estimate we derive here is valid for every choice of q, but this estimate is not the best we will give. We will prove a better estimate, using this exponential sum estimate, in chapter 3.

Lemma 1.1. *Define $k_q(m)$ as follows:*

$$k_q(m) = \begin{cases} 1 & \textit{if } \gcd(m,q) = 1, \\ 0 & \textit{otherwise.} \end{cases}$$

Then we have

$$k_q(m) = P \sum_{r|q} \Bigg(\sum_{\substack{0 \le a < r \\ (a,r)=1}} e\bigg(m \frac{a}{r}\bigg) \Bigg) \frac{\mu(r)}{\phi(r)}$$

Proof. We have

$$k_q(m) = \sum_{s|(m,q)} \mu(s) = \sum_{s|q} \frac{\mu(s)}{s} \sum_{0 \le b < s} e\bigg(m \frac{b}{s}\bigg),$$

therefore

$$k_q(m) = \sum_{r|q} \Bigg(\sum_{\substack{0 < a \le r \\ (a,r)=1}} e\bigg(m \frac{a}{r}\bigg) \Bigg) \Bigg(\sum_{\substack{s \\ r|s|q}} \frac{\mu(s)}{s} \Bigg).$$

Since

$$\sum_{\substack{s \\ r|s|q}} \frac{\mu(s)}{s} = P \frac{\mu(r)}{\phi(r)},$$

we can deduce

$$k_q(m) = P\sum_{r|q}\Bigg(\sum_{\substack{0<a\leq r\\(a,r)=1}} e\Big(m\frac{a}{r}\Big)\Bigg)\frac{\mu(r)}{\phi(r)}.$$

This completes the proof of the Lemma. □

Remark 1.1. Important to note that $\nu_p(\mathcal{D}) \leq s$ with equality if $p > h_s - h_1$. Also if $\mathcal{D}$ is admissible, then

$$\frac{1}{p} \leq 1 - \frac{\nu_p(\mathcal{D})}{p} \leq 1 - \frac{1}{p}$$

and we have that

$$\prod_{p\leq h_s-h_1}\frac{1}{p} \leq \prod_{p\leq h_s-h_1}\Big(1-\frac{\nu_p(\mathcal{D})}{p}\Big) \leq \prod_{p\leq h_s-h_1}\Big(1-\frac{1}{p}\Big).$$

Since h_s and h_1 are fixed integers, we therefore have

$$\prod_{\substack{p\leq h_s-h_1\\p|q}}\Big(1-\frac{\nu_p(\mathcal{D})}{p}\Big) \asymp_{\mathcal{D}} \prod_{\substack{p\leq h_s-h_1\\p|q}}\Big(1-\frac{1}{p}\Big)^s.$$

Moreover, if $p > h_s - h_1$ then $1 - \frac{\nu_p(\mathcal{D})}{p} = 1 - \frac{s}{p}$, so that

$$\prod_{\substack{p>h_s-h_1\\p|q}}\Big(1-\frac{\nu_p(\mathcal{D})}{p}\Big) = \prod_{\substack{p>h_s-h_1\\p|q}}\Big(1-\frac{s}{p}\Big) \asymp_{\mathcal{D}} \prod_{\substack{p>h_s-h_1\\p|q}}\Big(1-\frac{1}{p}\Big)^s.$$

Putting these together we deduce that

$$\frac{\phi_{\mathcal{D}}(q)}{q} \asymp_{\mathcal{D}} \Big(\frac{\phi(q)}{q}\Big)^s = P^s.$$

Now we state the theorem which we will prove at the end of this chapter:

Theorem 1.1. *Let*

$$M_k^{\mathcal{D}}(q,h) = \sum_{n=0}^{q-1}\Bigg(\sum_{m=1}^{h} k_q(n+m+h_1)\cdots k_q(n+m+h_s) - h\prod_{p|q}\Big(1-\frac{\nu_p(D)}{p}\Big)\Bigg)^k.$$

Then we have that

$$M_k^{\mathcal{D}}(q,h) \ll qh^{k/2}P^{-2^{ks}+ks},$$

where the implicit constant depends on k and s.

In order to go toward the proof we use exponential sums to better understand the admissible set $\mathcal{D} = \{h_1, h_2, \cdots, h_s\}$. Also we need to prove some lemmas. We have that

$$k_q(m) = P\sum_{r|q}\Bigg(\sum_{\substack{0<a\leq r\\(a,r)=1}} e\Big(m\frac{a}{r}\Big)\Bigg)\frac{\mu(r)}{\phi(r)}.$$

by lemma 1.1. Thus,

$$k_q(m+h_1) = P\sum_{r|q}\Bigg(\sum_{\substack{0<a\leq r\\(a,r)=1}} e\Big(m\frac{a}{r}+h_1\frac{a}{r}\Big)\Bigg)\frac{\mu(r)}{\phi(r)},$$

.

.

.

$$k_q(m+h_s) = P\sum_{r|q}\Bigg(\sum_{\substack{0<a\leq r\\(a,r)=1}} e\Big(m\frac{a}{r}+h_s\frac{a}{r}\Big)\Bigg)\frac{\mu(r)}{\phi(r)}.$$

So we deduce that

$$\begin{aligned} &k_q(m+h_1)\cdots k_q(m+h_s) \qquad (1.1)\\ &= P^s\sum_{r_1,r_2,\cdots,r_s|q}\frac{\mu(r_1)\cdots\mu(r_s)}{\phi(r_1)\cdots\phi(r_s)}\sum_{\substack{0<a_i\leq r_i\\(a_i,r_i)=1\\1\leq i\leq s}} e\Bigg(m\sum_{i=1}^{s}\frac{a_i}{r_i}\Bigg)e\Bigg(\sum_{i=1}^{s}h_i\frac{a_i}{r_i}\Bigg).\end{aligned}$$

By summing the left-hand side of (1.1) we have

$$\begin{aligned}&\sum_{m=1}^{h} k_q(n+m+h_1)\cdots k_q(n+m+h_s)\\ &= P^s\sum_{r_1,r_2,\cdots,r_s|q}\frac{\mu(r_1)\cdots\mu(r_s)}{\phi(r_1)\cdots\phi(r_s)}\sum_{\substack{0<a_i\leq r_i\\(a_i,r_i)=1\\1\leq i\leq s}}\Bigg(\sum_{m=1}^{h} e\Bigg(m\sum_{i=1}^{s}\frac{a_i}{r_i}\Bigg)e\Bigg(\sum_{i=1}^{s}h_i\frac{a_i}{r_i}\Bigg)\Bigg)\,e\Bigg(n\Bigg(\sum_{i=1}^{s}\frac{a_i}{r_i}\Bigg)\Bigg)\\ &= P^s\sum_{r_1,r_2,\cdots,r_s|q}\frac{\mu(r_1)\cdots\mu(r_s)}{\phi(r_1)\cdots\phi(r_s)}\sum_{\substack{0<a_i\leq r_i\\(a_i,r_i)=1\\1\leq i\leq s}}\Bigg(E_h\Bigg(\sum_{i=1}^{s}\frac{a_i}{r_i}\Bigg)e\Bigg(\sum_{i=1}^{s}h_i\frac{a_i}{r_i}\Bigg)\Bigg)\,e\Bigg(n\Bigg(\sum_{i=1}^{s}\frac{a_i}{r_i}\Bigg)\Bigg),\end{aligned}$$

where

$$E_h(x) = \sum_{m=1}^{h} e(mx).$$

To proceed with the argument we have to consider the case $\sum_{i=1}^{s} \frac{a_i}{r_i} \in \mathbb{Z}$ to extract the main term from the sum. We have that

$$P^s \sum_{r_1, r_2, \cdots, r_s | q} \frac{\mu(r_1) \cdots \mu(r_s)}{\phi(r_1) \cdots \phi(r_s)} \sum_{\substack{0 < a_i \leq r_i \\ (a_i, r_i) = 1 \\ 1 \leq i \leq s \\ \sum_{i=1}^{s} \frac{a_i}{r_i} \in \mathbb{Z}}} \left(E_h \left(\sum_{i=1}^{s} \frac{a_i}{r_i} \right) e \left(\sum_{i=1}^{s} h_i \frac{a_i}{r_i} \right) \right) e \left(n \left(\sum_{i=1}^{s} \frac{a_i}{r_i} \right) \right)$$

$$= hP^s \sum_{r_1, r_2, \cdots, r_s | q} \frac{\mu(r_1) \cdots \mu(r_s)}{\phi(r_1) \cdots \phi(r_s)} \sum_{\substack{0 < a_i \leq r_i \\ (a_i, r_i) = 1 \\ 1 \leq i \leq s \\ \sum_{i=1}^{s} \frac{a_i}{r_i} \in \mathbb{Z}}} e \left(\sum_{i=1}^{s} h_i \frac{a_i}{r_i} \right),$$

since $E_h(r) = h$ for all integers r. Now, we need to use Lemma 3 of [**8**] (due to Hardy and Littlewood). Hardy and Littlewood proved that

$$\mathfrak{S}(D) = \sum_{r_1, r_2, \cdots, r_s < \infty} \frac{\mu(r_1) \cdots \mu(r_s)}{\phi(r_1) \cdots \phi(r_s)} \sum_{\substack{0 < a_i \leq r_i \\ (a_i, r_i) = 1 \\ 1 \leq i \leq s \\ \sum_{i=1}^{s} \frac{a_i}{r_i} \in \mathbb{Z}}} e \left(\sum_{i=1}^{s} h_i \frac{a_i}{r_i} \right)$$

where $\mathfrak{S}$ is the singular series

$$\mathfrak{S}(D) = \prod_{p} \left(1 - \frac{1}{p} \right)^{-s} \left(1 - \frac{\nu_p(\mathcal{D})}{p} \right).$$

Lemma 1.2. (Hardy and Littlewood) *Let $r_1, r_2 \ldots, r_s$ be square-free integers. Then*

$$A(r_1, \cdots, r_s) = \sum_{\substack{0 < a_i \leq r_i \\ (a_i, r_i) = 1 \\ 1 \leq i \leq s \\ \sum_{i=1}^{s} \frac{a_i}{r_i} \in \mathbb{Z}}} e \left(\sum_{i=1}^{s} h_i \frac{a_i}{r_i} \right)$$

If $r_i = r_i' r_i''$ with $(\prod r_i', \prod r_i'') = 1$, then

$$A(r_1, \cdots, r_s) = A(r_1', \cdots, r_k') A(r_1'', \cdots, r_k'').$$

Also, we have that

$$\sum_{r_1,r_2,\cdots,r_s|q}\frac{\mu(r_1)\cdots\mu(r_s)}{\phi(r_1)\cdots\phi(r_s)}\sum_{\substack{0<a_i\le r_i\\(a_i,r_i)=1\\1\le i\le s\\\sum_{i=1}^s\frac{a_i}{r_i}\in\mathbb{Z}}}e\Bigg(\sum_{i=1}^s h_i\frac{a_i}{r_i}\Bigg)=\mathfrak{S}_q(D),$$

where we define the singular series

$$\mathfrak{S}_q(D):=\prod_{p|q}\left(1-\frac{1}{p}\right)^{-s}\left(1-\frac{\nu_p(\mathcal{D})}{p}\right)$$

This lemma allows us to partition the sum over $A(r_1,\cdots,r_s)$ into sums $A\Big(p_1^{e_1},\cdots,p_s^{e_s}\Big)$ corresponding to the primes $p_i|r_i$.

Proof. For proving that

$$A(r_1,\cdots,r_s)=A(r_1',\cdots,r_k')A(r_1'',\cdots,r_k''),$$

we write

$$\frac{a_i}{r_i}\equiv\frac{a_i'}{r_i'}+\frac{a_i''}{r_i''}\ (\mathrm{mod}\ 1).$$

By the Chinese Remainder Theorem, each reduced residue a_i modulo r_i corresponds to a pair a_i',a_i'' of reduced residues modulo r_i',r_i''. Hence $A(.,\dots,.)$ is a multiplicative function in all of its variables and therefore, to prove

$$\sum_{r_1,r_2,\cdots,r_s|q}\frac{\mu(r_1)\cdots\mu(r_s)}{\phi(r_1)\cdots\phi(r_s)}A(r_1,\cdots,r_s)=\mathfrak{S}_q(D),$$

it suffices to only prove the case $q=p$ for p prime (since the r_i's are squarefree). In this case

$$\sum_{r_1,r_2,\cdots,r_s|q}\frac{\mu(r_1)\cdots\mu(r_s)}{\phi(r_1)\cdots\phi(r_s)}A(r_1,\cdots,r_s)=\sum_{\substack{I\subseteq\{1,\cdots,s\}\\|I|\ge1}}\frac{(-1)^{|I|}}{(p-1)^{|I|}}A_I(p),$$

where

$$A_I(p)=\sum_{\substack{i\in I\\1\le a_i\le p-1\\\sum_{i\in I}a_i\equiv0\ (\mathrm{mod}\ p)}}e\Bigg(\sum_{i\in I}h_i\frac{a_i}{p}\Bigg).$$

Let

$$B_I(p) = \sum_{\substack{i\in I \\ 0\le a_i\le p-1 \\ \sum_{i\in I} a_i \equiv 0 \pmod p}} e\Bigg(\sum_{i\in I} h_i \frac{a_i}{p}\Bigg).$$

We have that:

1) $B_I(p) = \sum_{J\subseteq I} A_J(p)$, because we can write

$$B_I(p) = \sum_{J\subseteq I} \sum_{\substack{i\in J \\ 1\le a_i\le p-1 \\ \sum_{i\in J} a_i \equiv 0 \pmod p \\ a_j=0 \\ j\in I\setminus j}} e\Bigg(\sum_{i\in I} h_i \frac{a_i}{p}\Bigg)$$

2) $B_I(p) = p^{|I|-1}$ if $p|h_i - h_j$ for all $i, j \in I$, otherwise $B_I(p) = 0$. To see this note that if $h_i \equiv h_j \pmod p$ for each i, j, then we have

$$B_I(p) = \sum_{\substack{i\in I \\ 0\le a_i\le p-1 \\ \sum_{i\in I} a_i \equiv 0 \pmod p}} e\Bigg(h_{i_1} \sum_{i\in I} \frac{a_i}{p}\Bigg) = \sum_{\substack{i\in I \\ 0\le a_i\le p-1 \\ \sum_{i\in I} a_i \equiv 0 \pmod p}} 1.$$

This equals $p^{|I|-1}$ since the condition $\sum_{i\in I} a_i \equiv 0 \pmod p$ implies that $\sum_{i\in I\setminus\{i_1\}} a_i \equiv -a_{i_1} \pmod p$, which means that we can choose $0 \le a_i \le p-1$ for $i \in I \setminus \{i_1\}$, as we wish and then a_{i_1} is fixed. Now if p does not divide $h_i - h_j$ for some $i, j \in I$, we eliminate the condition

$$\sum_{i\in I} a_i \equiv 0 \pmod p$$

with replacing a_j by $-\sum_{m\in I\setminus\{j\}} a_m$, which proves $B_I(p) = 0$.

Also let $B_\varnothing(p) = 1$. Therefore if we write $I = I_1 \cup \cdots \cup I_p$, which is a partition of values of h_i mod p into congruence classes ($I_j = \{i : h_i \equiv j \pmod p\}$), then by inclusion-exclusion we find that

$$A_I(p) = \sum_{J\subseteq I} (-1)^{|I|-|J|} B_J(p) = (-1)^{|I|}\Bigg(1 + \sum_{k=1}^{p} \sum_{\substack{J\subseteq I_k \\ |J|\ge 1}} (-1)^{|J|} p^{|J|-1}\Bigg)$$

$$= (-1)^{|I|}\Bigg(1 + \frac{1}{p}\sum_{k=1}^{p}\Big((1-p)^{|I_k|} - 1\Big)\Bigg) = \frac{(-1)^{|I|}}{p}\sum_{k=1}^{p}(1-p)^{|I_k|}.$$

But then

$$\sum_{p^{e_1},\cdots,p^{e_s}|p} \frac{\mu(p^{e_1})\cdots\mu(p^{e_s})}{\phi(p^{e_1})\cdots\phi(p^{e_s})} A_I(p)$$

for $I = \{i : r_i = p\}$ is equal to

$$\sum_{I\subseteq\{1,\cdots,s\}} \left(\frac{-1}{p-1}\right)^{|I|} \frac{(-1)^{|I|}}{p} \sum_{k=1}^{p} (1-p)^{|I_k|}. \tag{1.2}$$

Now let $H_1 \cup \cdots \cup H_p$ be partition of $\{1, \cdots, s\}$ given by $H_j := \{i : h_i \equiv j \pmod{p}\}$. Hence $I_k = I \cap H_k$ and we have that the sum in (1.2) equals

$$\frac{1}{p}\sum_{\substack{I_r\subseteq H_r\\ 1\le r\le p}} \frac{1}{(p-1)^{|I_1|+\cdots+|I_p|}} \sum_{k=1}^{p} (1-p)^{|I_k|} = \frac{1}{p} \sum_{k=1}^{p} \sum_{I_k\subseteq H_k} (-1)^{|I_k|} \prod_{\substack{j=1\\ j\neq k}}^{p} \sum_{I_j\subseteq H_j} \frac{1}{(p-1)^{|I_j|}}$$

$$= \frac{1}{p} \sum_{\substack{k=1\\ H_k=\varnothing}}^{p} \left(1+\frac{1}{p-1}\right)^{s} = \left(1-\frac{\nu_p(\mathcal{D})}{p}\right)\left(\frac{p}{p-1}\right)^{s}.$$

This completes the proof of Lemma. □

Using Lemma 1.2 we have

$$\sum_{m=1}^{h} k_q(n+m+h_1)\cdots k_q(n+m+h_s) - h\prod_{p|q}\left(1-\frac{\nu_p(\mathcal{D})}{p}\right)$$

$$= P^s \sum_{r_1,r_2,\cdots,r_s|q} \frac{\mu(r_1)\cdots\mu(r_s)}{\phi(r_1)\cdots\phi(r_s)} \sum_{\substack{0<a_i\le r_i\\ (a_i,r_i)=1\\ 1\le i\le s\\ \sum_{i=1}^{s}\frac{a_i}{r_i}\notin\mathbb{Z}}} \left(E_h\left(\sum_{i=1}^{s}\frac{a_i}{r_i}\right) e\left(\sum_{i=1}^{s} h_i\frac{a_i}{r_i}\right)\right) e\left(n\left(\sum_{i=1}^{s}\frac{a_i}{r_i}\right)\right)$$

and, consequently,

$$\left(\sum_{m=1}^{h} k_q(n+m+h_1)\cdots k_q(n+m+h_s) - h\prod_{p|q}\left(1-\frac{\nu_p(\mathcal{D})}{p}\right)\right)^{k} \tag{1.3}$$

$$= P^{ks} \sum_{\substack{r_{i,j}|q\\ 1\le i\le k\\ 1\le j\le s}} \left(\prod_{i,j}\frac{\mu(r_{i,j})}{\phi(r_{i,j})}\right) \sum_{\substack{0<a_{i,j}\le r_{i,j}\\ (a_{i,j},r_{i,j})=1\\ 1\le j\le s\\ \sum_{j=1}^{s}\frac{a_{i,j}}{r_{i,j}}\notin\mathbb{Z}\\ 1\le i\le k}} \left(E_h\left(\sum_{i=1}^{s}\frac{a_{1,j}}{r_{1,j}}\right)\cdots E_h\left(\sum_{j=1}^{s}\frac{a_{k,j}}{r_{k,j}}\right) e\left(\sum_{i,j} h_j\frac{a_{i,j}}{r_{i,j}}\right)\right)$$

$$\times\, e\left(n\left(\sum_{i,j}\frac{a_{i,j}}{r_{i,j}}\right)\right).$$

Summing (1.3) over n mod q and using the fact that when $q\sum_i \rho_i \in \mathbb{Z}$

$$\sum_{n=0}^{q-1} e\Big(n\Big(\sum_i \rho_i\Big)\Big) = 0$$

unless $\sum_i \rho_i \in \mathbb{Z}$, we have that

$$\sum_{n=0}^{q-1}\left(\sum_{m=1}^{h} k_q(n+m+h_1)\cdots k_q(n+m+h_s) - h\prod_{p|q}\left(1-\frac{\nu_p(\mathcal{D})}{p}\right)\right)^k$$
$$= qP^{ks}\sum_{\substack{r_{i,j}|q\\ 1\le i\le k\\ 1\le j\le s}}\left(\prod\frac{\mu(r_{i,j})}{\phi(r_{i,j})}\right)\sum_{\substack{1\le i\le k\\ 0<a_{i,j}\le r_{i,j}\\ (a_{i,j},r_{i,j})=1\\ \sum_{j=1}^{s}\frac{a_{i,j}}{r_{i,j}}\notin\mathbb{Z}\\ \sum_{i,j}\frac{a_{i,j}}{r_{i,j}}\in\mathbb{Z}}}\left(E_h\Big(\sum_{j=1}^{s}\frac{a_{1,j}}{r_{1,j}}\Big)\cdots E_h\Big(\sum_{j=1}^{s}\frac{a_{k,j}}{r_{k,j}}\Big)e\Big(\sum_{i,j} h_j\frac{a_{i,j}}{r_{i,j}}\Big)\right).$$

Let $F(x) = \min(h, \frac{1}{\|x\|})$ where $\|x\|$ is the distance between x and the closest integer to x. We will show that $|E_h(x)| \le F(x)$. In order to do this note that

$$\left|\sum_{m=1}^{h} e(mx)\right| = \left|\frac{e^{2\pi i h x}-1}{e^{2\pi i x}-1}\right| = \left|\frac{e^{2\pi i\frac{h}{2}x}-e^{-2\pi i\frac{h}{2}x}}{e^{\pi i x}-e^{-\pi i x}}\right| = \left|\frac{\sin(\pi h x)}{\sin(\pi x)}\right|$$

for $-\frac{1}{2} < x < \frac{1}{2}$ and $x \neq 0$, we have $|\sin(x)| > \frac{2}{\pi}x$, thus for $-\frac{1}{2} \le x \le \frac{1}{2}$ we have

$$\left|\frac{\sin(\pi h x)}{\sin(\pi x)}\right| \le \left|\frac{1}{\sin(\pi x)}\right| < \left|\frac{1}{\frac{2}{\pi}\pi x}\right| = \frac{1}{2x}.$$

We deduce that for arbitrary x we have $E_h(x) \le \frac{1}{\|x\|}$, and obviously $E_h(x) \le h$. Consequently we have

$$\sum_{n=0}^{q-1}\left(\sum_{m=1}^{h} k_q(n+m+h_1)\cdots k_q(n+m+h_s) - h\prod_{p|q}\left(1-\frac{\nu_p(D)}{p}\right)\right)^k$$
$$\ll qP^{ks}\sum_{\mathbf{r}|q}\ \sum_{[r_{1,1},r_{1,2},\cdots r_{k,s}]=\mathbf{r}}\frac{S(\{r_{i,j}\}_{i,j})}{(\prod\phi(r_{i,j}))} \tag{1.4}$$

where

$$S(\{r_{i,j}\}_{i,j}) = \sum_{\substack{0<a_{i,j}\le r_{i,j}\\ (a_{i,j},r_{i,j})=1\\ \sum_{j=1}^{s}\frac{a_{i,j}}{r_{i,j}}\notin\mathbb{Z}\\ \sum_{i,j}\frac{a_{i,j}}{r_{i,j}}\in\mathbb{Z}}} F\Big(\sum_{j=1}^{s}\frac{a_{1,j}}{r_{1,j}}\Big)\cdots F\Big(\sum_{j=1}^{s}\frac{a_{k,j}}{r_{k,j}}\Big)$$

Lemma 1.3. *Every element of the form*

$$\sum_{j=1}^{s} \frac{a_{i,j}}{r_{i,j}} \quad \textit{where} \quad 0 < a_{i,j} \leq r_{i,j}$$

can be written as

$$\frac{a}{[r_{i,1}, r_{i,2}, \cdots r_{i,s}]}(\text{ mod } 1), \quad \textit{where} \quad 1 \leq a \leq [r_{i,1}, r_{i,2}, \cdots r_{i,s}],$$

and each fraction that has such a representation has exactly $\frac{r_{i,1}r_{i,2}\cdots r_{i,s}}{[r_{i,1},r_{i,2},\cdots r_{i,s}]}$ *representations.*

By $\frac{r_{i,1}r_{i,2}\cdots r_{i,s}}{[r_{i,1},r_{i,2},\cdots r_{i,s}]}$ representations we mean that the equation

$$\sum_{j=1}^{s} \frac{a_{i,j}}{r_{i,j}} = \tau \text{ (mod 1)}$$

has exactly $\frac{r_{i,1}r_{i,2}\cdots r_{i,s}}{[r_{i,1},r_{i,2},\cdots r_{i,s}]}$ different answers, if it has any.

Proof. Let $d = (r_1, r_2)$ and we call $r'_i = \frac{r_i}{d}$ for $i = 1, 2$. For fixed a, b we are interested in the number of solutions for the equation

$$\frac{a}{r_1} + \frac{b}{r_2} = \frac{x}{r_1} + \frac{y}{r_2} \text{ (mod 1)}$$

where $1 \leq x \leq r_1$ and $1 \leq y \leq r_2$, which leads us to the number of solutions of

$$ar'_2 + br'_1 \equiv xr'_2 + yr'_1 \text{ (mod } r'_1 r'_2 d). \tag{1.5}$$

We have $a \equiv x \pmod{r'_1}$ and $b \equiv y \pmod{r'_2}$. Let $x = a + ir'_1$ and $y = b + jr'_2$. Then by using (1.5) we have

$$(a - x)r'_2 \equiv (y - b)r'_1 \text{ (mod } r'_1 r'_2 d).$$

Therefore we have $i + j \equiv 0 \pmod{d}$, which has exactly d solutions. So we conclude that, given a and b, there are exactly d solutions (x, y) with $1 \leq x \leq r_1$ and $1 \leq y \leq r_2$ to the equation

$$\frac{a}{r_1} + \frac{b}{r_2} = \frac{x}{r_1} + \frac{y}{r_2} \text{ (mod } \mathbb{Z}).$$

Obviously $\frac{a}{r_1}+\frac{b}{r_2}$ (mod 1) $\in \left\{\frac{t}{[r_1,r_2]} : 0 \le t \le \ [r_1,r_2]\right\}$ and as we showed above, each element is repeated exactly $d = \frac{r_1r_2}{[r_1,r_2]}$ times. This proves the lemma for $s=2$. Using induction, we have that

$$\frac{a_{i,1}}{r_{i,1}}+\cdots+\frac{a_{i,k-1}}{r_{i,k-1}} = \frac{a}{[r_{i,1},r_{i,2},\cdots r_{i,k-1}]},$$

with exactly $\frac{r_{i,1}r_{i,2}\cdots r_{i,k-1}}{[r_{i,1},r_{i,2},\cdots r_{i,k-1}]}$ repetitions each. And, by the first part of the proof there are exactly

$$\frac{[r_{i,1},r_{i,2},\cdots r_{i,k-1}]r_{i,k}}{[r_{i,1},r_{i,2},\cdots r_{i,k}]}$$

ways to write $\frac{a_{i,1}}{r_{i,1}}+\cdots+\frac{a_{i,k}}{r_{i,k}}$ as $\frac{a}{[r_{i,1},r_{i,2},\cdots r_{i,k-1}]}+\frac{a_{i,k}}{r_{i,k}}$ (mod1). Now the total number of repetitions is

$$\frac{[r_{i,1},r_{i,2},\cdots r_{i,k-1}]r_{i,k}}{[r_{i,1},r_{i,2},\cdots r_{i,k}]}\cdot\frac{r_{i,1}r_{i,2}\cdots r_{i,k-1}}{[r_{i,1},r_{i,2},\cdots r_{i,k-1}]} = \frac{r_{i,1}r_{i,2}\cdots r_{i,k}}{[r_{i,1},r_{i,2},\cdots r_{i,k}]}$$

□

Now our task is to bound (1.4), for which we need to use the idea of Montgomery and Vaughan's Fundamental Lemma [**6**], slightly modified. In order to do that we use Lemma 2 from [**8**, Page 596].

Lemma 1.4. *Let $q_1,\cdots,q_k$ be square-free integers, each one strictly greater than 1, and put $d=[q_1,...,q_k]$. Let G be a complex-valued function defined on $(0,1)$, and suppose that G_0 is a nondecreasing function on the positive integers such that*

$$\sum_{a=1}^{q-1}|G(a/q)|^2 \le qG_0(q),$$

for all square-free integers $q>1$. Then

$$\left|\sum_{\substack{a_1,\cdots,a_k\\ 0<a_i<q_i\\ \sum\frac{a_i}{q_i}\in\mathbb{Z}}}\prod_{i=1}^{k}G(a_i/q_i)\right| \le \frac{1}{d}\prod_{i=1}^{k}q_iG_0(q_i)^{1/2}.$$

We now need to verify that F satisfies the requirements for G in the Lemma 1.4. Lemma 4 of [**6**] asserts that

$$\sum_{0<a<q}F\left(\frac{a}{q}\right)^2 \ll q\min(q,h).$$

Since $\min(q,h)$ is obviously a non-decreasing function of q, we can use Lemma 1.4 with F and $\min(q,h)$ in place of G and G_0 respectively. About the condition $q_i > 1$, note that, since we apply Lemma 1.4 for $q_i = [r_{i,1},\cdots,r_{i,s}]$ and we have $\sum_{j=1}^{s}\frac{a_{i,j}}{r_{i,j}} \notin \mathbb{Z}$, then $q_i = [r_{i,1},\cdots,r_{i,s}] \neq 1$. From Lemma 1.3 we have

$$
\begin{aligned}
S(\{r_{i,j}\}_{i,j}) &= \sum_{\substack{0<a_{i,j}\le r_{i,j}\\ (a_{i,j},r_{i,j})=1\\ \sum_{j=1}^{s}\frac{a_{i,j}}{r_{i,j}}\notin\mathbb{Z}\\ \sum_{i,j}\frac{a_{i,j}}{r_{i,j}}\in\mathbb{Z}}} F\left(\sum_{j=1}^{s}\frac{a_{1,j}}{r_{1,j}}\right)\cdots F\left(\sum_{j=1}^{s}\frac{a_{k,j}}{r_{k,j}}\right)\\
&\le T \sum_{\substack{0<a_i<[r_{i,1},\cdots,r_{i,s}]\\ \sum_{i=1}^{s}\frac{a_i}{[r_{i,1},\cdots,r_{i,s}]}\in\mathbb{Z}}} F\left(\frac{a_1}{[r_{1,1},\cdots,r_{1,s}]}\right)\cdots F\left(\frac{a_k}{[r_{k,1},\cdots,r_{k,s}]}\right), \qquad (1.6)
\end{aligned}
$$

where

$$
T = \frac{r_{1,1}\cdots r_{1,s}}{[r_{1,1},\cdots,r_{1,s}]}\cdots\frac{r_{k,1}\cdots r_{k,s}}{[r_{k,1},\cdots,r_{k,s}]}.
$$

Now using Lemma 1.4 with $G = F$ and $q_i = [r_{i,1},\cdots,r_{i,s}]$, we have that

$$
S(\{r_{i,j}\}_{i,j}) \ll \frac{r_{1,1}\cdots r_{k,s}}{\mathbf{r}}h^{k/2} \qquad (1.7)
$$

Now we are ready to prove our Theorem:

Proof of Theorem 1.1. We prove the result with q square-free. Then the theorem follows for q non-square-free immediately by considering the result for $Q = \prod_{p|q} p$. Now using relations (1.4) and (1.7), we have that

$$
\begin{aligned}
&\sum_{n=0}^{q-1}\left(\sum_{m=1}^{h} k_q(n+m+h_1)\cdots k_q(n+m+h_s) - h\prod_{p|q}\left(1-\frac{\nu_p(D)}{p}\right)\right)^k\\
&\ll qP^{ks}\sum_{r|q}\frac{1}{r}\sum_{[r_{1,1},r_{1,2},\cdots,\,r_{k,s}]=r}\frac{r_{1,1}\cdots r_{k,s}}{\phi(r_{1,1})\cdots\phi(r_{k,s})}h^{k/2}\\
&\le qP^{ks}\sum_{r|q}\frac{1}{r}\left(\sum_{r'|r}\frac{r'}{\phi(r')}\right)^{ks}h^{k/2} = qh^{k/2}P^{ks}\prod_{p|q}\left(1+\frac{1}{p}\left(2+\frac{1}{p-1}\right)^{ks}\right)\\
&\ll qh^{k/2}P^{-2^{ks}+ks}.
\end{aligned}
$$

The last inequality is valid since we have that

$$\prod_{p|q}\left(1+\frac{1}{p}\Big(2+\frac{1}{p-1}\Big)^{ks}\right) = P^{-2^{ks}}\prod_{p|q}\left(1-\frac{1}{p}\right)^{2^{ks}}\left(1+\frac{1}{p}\Big(2+\frac{1}{p-1}\Big)^{ks}\right)$$

and for $p > (ks)^2$ we have

$$\Big(2+\frac{1}{p-1}\Big)^{ks} = 2^{ks}e^{O(\frac{ks}{p})}.$$

Thus, from $e^x > 1+x$, we find that

$$\left(1+\frac{1}{p}\Big(2+\frac{1}{p-1}\Big)^{ks}\right) = 1+\frac{2^{ks}}{p}\left(1+O\Big(\frac{ks}{p}\Big)\right) < e^{\frac{2^{ks}}{p}},$$

and, consequently

$$\left(1-\frac{1}{p}\right)^{2^{ks}}\left(1+\frac{1}{p}\Big(2+\frac{1}{p-1}\Big)^{ks}\right) < 1.$$

If $p \leq (ks)^2$, then

$$\left(1-\frac{1}{p}\right)^{2^{ks}}\left(1+\frac{1}{p}\Big(2+\frac{1}{p-1}\Big)^{ks}\right) < e^{-\frac{2^{ks}}{p}}3^{ks} < 1.$$

□

CHAPTER 2

A PROBABILISTIC ESTIMATE

In this chapter we prove an estimate about the distribution of s-tuples of reduced residues using a probabilistic method. The estimate derived here is valid only when q is not divisible by any small prime, and in this case it is the best possible we can have. In particular, it's much better than our earlier exponential sum estimate in this range.

Let X_i, for $1 \leq i \leq h$, be independent identically distributed random variables such that

$$\mathrm{Prob(X_i = 1) = 1 - Prob(X_i = 0) = P.}$$

Then

$$X = X_1 + \cdots + X_h$$

is called a *binomial random variable.* Given such a random variable X, we denote with $\mu_k(h, P)$ its k-th moment about its mean, that is to say,

$$\mu_k(h, P) := \mathbb{E}\Big((X - hP)^k\Big).$$

We will use these random variables in the proof of the following Theorem:

Theorem 2.1. *Let A be a set of h integers and $h_1 < \cdots < h_s$. Suppose that for each prime divisor p of q we have $p > \max A - \min A + h_s - h_1$. Suppose also* *that $p > y$ for all $p|q$. Then for $y > h^k$ and for each fixed even $k > 1$*

$$M_k^{\mathcal{D}}(q,h) = \sum_{n=0}^{q-1}\Bigg(\sum_{\substack{m\in A\\(n+m+h_i,q)=1\\1\leq i\leq s}} 1 - h\left(\frac{\phi(q)}{q}\right)^s\Bigg)^k \ll q\left(h\left(\frac{\phi(q)}{q}\right)^s\right)^{[k/2]} + qh\left(\frac{\phi(q)}{q}\right)^s,$$

which the implicit constant depends on k and $|h_1 - h_s|$.

Remark 2.1. Under the conditions of Theorem 2.1, it provides a better estimate than Theorem 1.1. Indeed Theorem 1.1 yields the estimate

$$M_k^{\mathcal{D}}(q,h) \ll qh^{k/2}P^{-2^{ks}+ks}$$

whereas by Theorem 2.1 we have that

$$M_k^{\mathcal{D}}(q,h) \ll q\left(h\left(\frac{\phi(q)}{q}\right)^s\right)^{[k/2]} + qh\left(\frac{\phi(q)}{q}\right)^s.$$

Comparing two bounds and using the fact that $q\left(h\left(\frac{\phi(q)}{q}\right)^s\right)^{[k/2]} \leq qh^{k/2}P^{-2^{ks}+ks}$ proves the point.

Proof. The proof is similar to the proof of Lemma 9 in [**6**], with a small variation which we explain. We have that

$$\sum_{\substack{m\in A\\(m+h_i,q)=1\\1\leq i\leq s}} 1 - h\left(\frac{\phi(q)}{q}\right)^s = \sum_{1\leq j\leq H}\Bigg(\sum_{\substack{m\in A_j\\(m+h_i,q)=1\\1\leq i\leq s}} 1 - |A_j|\left(\frac{\phi(q)}{q}\right)^s\Bigg)$$

where

$$A_j = \{m \in A : m \equiv j \pmod H\},$$

where $H = |h_s - h_1| + 1$. From Hölder's inequality with, $\frac{1}{k} + \frac{1}{\frac{k}{k-1}} = 1$, we have that

$$\left|\sum_{i=1}^{H} a_i\right| \leq H^{\frac{k-1}{k}}\left(\sum_{i=1}^{H}|a_i|^k\right)^{\frac{1}{k}}$$

and, consequently,

$$\Bigg(\sum_{\substack{m\in A\\(m+h_i,q)=1\\1\le i\le s}} 1-h\left(\frac{\phi(q)}{q}\right)^s\Bigg)^k \le H^{k-1}\sum_{1\le j\le H}\Bigg(\sum_{\substack{m\in A_j\\(m+h_i,q)=1\\1\le i\le s}} 1-|A_j|\left(\frac{\phi(q)}{q}\right)^s\Bigg)^k.$$

Now we focus on

$$S_j=\sum_{n=0}^{q-1}\Bigg(\sum_{\substack{m\in A_j\\(n+m+h_i,q)=1\\1\le i\le s}} 1-|A_j|\left(\frac{\phi(q)}{q}\right)^s\Bigg)^k. \tag{2.1}$$

We note that

$$S_j=\sum_n\sum_r\binom{k}{r}\Bigg(\sum_{\substack{m\in A_j\\(n+m+h_i,q)=1\\1\le i\le s}} 1\Bigg)^r\Bigg(-|A_j|\left(\frac{\phi(q)}{q}\right)^s\Bigg)^{k-r}.$$

Moreover, we have that

$$\Bigg(\sum_{\substack{m\in A_j\\(n+m+h_i,q)=1\\1\le i\le s}} 1\Bigg)^r=\sum_{\substack{m_1,\cdots,m_r\in A_j\\(n+m_l+h_i,q)=1\\1\le i\le s\\1\le l\le r}} 1$$

We will show that $m_l+h_i\neq m_{l'}+h_{i'}$ for $m_l\neq m_{l'}$. Without loss of generality, we assume that $m_l<m_{l'}$ and therefore $m_l+h_i<m_{l'}+h_{i'}$. This is true since $m_l-m_{l'}\equiv 0 \pmod H$ and thus $|m_l-m_{l'}|\ge H>|h_i-h_{i'}|$. Now we claim that $m_l+h_i\not\equiv m_{l'}+h_{i'}$ (mod p) for all $p|q$. Assume, on the contrary, that

$$m_l+h_i\equiv m_{l'}+h_{i'} \pmod p$$

for some $p|q$. Then we have that $p|m_l+h_i-\Big(m_{l'}+h_{i'}\Big)$. We already have shown $m_l+h_i-\Big(m_{l'}+h_{i'}\Big)\neq 0$, therefore

$$p\le|m_l-m_{l'}|+|h_i-h_{i'}|,$$

which contradicts our assumption that $p > \max A - \min A + h_s - h_1$.

Applying these facts and changing the order of summation in S_j, we have that

$$\sum_{\substack{n=0 \\ (n+m_j+h_i,q)=1 \\ 1\leq i\leq s \\ 1\leq j\leq r}}^{q-1} 1 = \prod_{p|q}(p - st), \tag{2.2}$$

where $t = \#\left\{m_1, \cdots, m_r\right\}$. Let $S(r,t)$ denote the Stirling number of the second kind, i.e. the number of ways of partitioning a set of cardinality r into exactly t non-empty subsets. Following the proof of Lemma 9 in [**6**], $S(r,t)t!$ is the number of surjective maps from a set of cardinality r to a set of cardinality t. We set $S(r,0) = 0$ so that we have

$$\sum_{n=0}^{q-1} \sum_{\substack{m_1,\cdots,m_r\in A_j \\ (n+m_k+h_i,q)=1 \\ 1\leq i\leq s \\ 1\leq k\leq r}} 1 = \sum_{t=0}^{r} \sum_{\substack{\mathcal{B}\subseteq A_j \\ \mathrm{card}(\mathcal{B})=t}} S(r,t)t! \prod_{p|q}(p - st)$$

As there are $\binom{|A_j|}{t}$ possible choices for $\mathcal{B}$, the above is

$$q\sum_{t=1}^{r}\binom{|Aj|}{t}S(r,t)t!\left(\frac{\phi(q)}{q}\right)^{st}\prod_{p|q}\left(1-\frac{st}{p}\right)\left(1-\frac{1}{p}\right)^{-st}$$

and, since $p > y > h_s - h_1$ we have that

$$\prod_{p|q}\left(1-\frac{st}{p}\right)\left(1-\frac{1}{p}\right)^{-st} = 1 + O_{st}\left(\frac{1}{y}\right)$$

From Lemma 9 in [**6**, page.326] we have that

$$S_j = q\sum_{r=0}^{k}\binom{k}{r}(-|A_j|P^s)^{k-r}\sum_{t=0}^{r}\binom{|A_j|}{t}S(r,t)t!(P)^{st}\Big(1+O_{st}(\frac{1}{y})\Big)$$

and

$$q\sum_{r=0}^{k}\binom{k}{r}(-|A_j|P^s)^{k-r}\sum_{t=0}^{r}\binom{|A_j|}{t}S(r,t)t!(P)^{st} = \mu_k(|A_j|, P^s)$$

using [**6**, page.327]. Thus

$$\begin{aligned} S_j =& q\sum_{r=0}^{k}\binom{k}{r}\left(-|A_j|\left(\frac{\phi(q)}{q}\right)^s\right)^{k-r}\sum_{t=0}^{r}\binom{|A_j|}{t}S(r,t))t!\left(\frac{\phi(q)}{q}\right)^{st}\left(1+O_{st}\left(\frac{1}{y}\right)\right) \\ &= q\mu_k\left(|A_j|, \left(\frac{\phi(q)}{q}\right)^s\right) + O\left\{\frac{q}{y}\left(h\left(\frac{\phi(q)}{q}\right)^s\right)^k + h\left(\frac{\phi(q)}{q}\right)^s\right\}, \end{aligned}$$

using the fact that $|A_j| \leq h$. For the error term the dependence of the implicit constant on t can be considered to be a dependence on k, since $t < s < k$ we also use

$$\frac{q}{y}\sum_{t=0}^{r}\binom{|A_j|}{t}\left(\frac{\phi(q)}{q}\right)^{st} \ll \frac{q}{y}\left((h\left(\frac{\phi(q)}{q}\right)^{s})^{r} + h\left(\frac{\phi(q)}{q}\right)^{s}\right)$$

Next note that Lemma 11 of [**6**] states that, for any fixed integer $k > 0$, $\mu_k(h, P) \ll (hP)^{[k/2]} + hP$, uniformly for $0 < P < 1$, $h = 1, 2, 3, ...$ So

$$\mu_k(|A_j|, P^s) \ll (|A_j|P^s)^{[k/2]} + |A_j|P^s \leq (hP^s)^{[k/2]} + hP^s.$$

Using this and and our assumption that $y > h^k$, we find that

$$\sum_{n=0}^{q-1}\left(\sum_{\substack{m\in A\\ (n+m+h_i,q)=1\\ 1\leq i\leq s}} 1 - h\left(\frac{\phi(q)}{q}\right)^{s}\right)^{k} \ll q\left(h\left(\frac{\phi(q)}{q}\right)^{s}\right)^{[k/2]} + qh\left(\frac{\phi(q)}{q}\right)^{s}, \tag{2.3}$$

which concludes the proof of the Theorem. □

CHAPTER 3

THE PRINCIPAL ESTIMATE

In this chapter we will prove our principal estimate about the distribution of s-tuples of reduced residues, by combining both our probabilistic and exponential sum estimates. The new estimate that we derive here is valid for every q and it is better than our exponential sum estimate. Using the principal estimate, we will be able to prove Theorem 0.1.

Theorem 3.1. *Let k be a given even number, and fix constant $A > k$. Let $q_1 = \prod_{\substack{p|q \\ p \leq y}} p$ and $q_2 = \prod_{\substack{p|q \\ p > y}} p$, where $h^A > y > h^k$. Correspondingly we set $P_i = \frac{\phi(q_i)}{q_i}$ for $i = 1, 2$. For $h > P^{-1}$ we have*

$$M_k^{\mathcal{D}}(q,h) \ll q(hP^s)^{[k/2]} + qh(P)^s + qh^{k/2}P_1^{-2^{ks}+ks}P_2^{sk}.$$

And the implicit constant depends on k and s.

Remark 3.1. We call the result above principal estimate. It is better than Theorem 1.1 $\left(M_k^{\mathcal{D}}(q,h) \ll qh^{k/2}P^{-2^{ks}+ks}\right)$ because of the factor P_2^{sk} in

$$qh^{k/2}P_1^{-2^{ks}+ks}P_2^{sk}.$$

Proof. Since q is square-free we have $q = q_1q_2$ and $(q_1, q_2) = 1$. By the Chinese Remainder Theorem we have that

$$M_k^{\mathcal{D}}(q,h) = \sum_{n_1=0}^{q_1-1}\sum_{n_2=0}^{q_2-1} D(n_1,n_2)^k,$$

where

$$D(n_1,n_2) = \sum_{\substack{m=1\\(n_i+m+h_j,q_i)=1\\1\le j\le s\\i=1,2}}^{h} 1 - h\prod_{p|q}\left(1-\frac{\nu_p(\mathcal{D})}{p}\right).$$

Following [**6**], we may write $D = D_1 + D_2$ where

$$D_1 = \prod_{p|q_2}\left(1-\frac{\nu_p(\mathcal{D})}{p}\right)\sum_{\substack{m=1\\(n_1+m+h_j,q_1)=1\\1\le j\le s}}^{h} 1 - h\prod_{p|q}\left(1-\frac{\nu_p(\mathcal{D})}{p}\right)$$

$$D_2 = \sum_{\substack{m=1\\(n_i+m+h_j,q_i)=1\\1\le j\le s\\i=1,2}}^{h} 1 - \prod_{p|q_2}\left(1-\frac{\nu_p(\mathcal{D})}{p}\right)\sum_{\substack{m=1\\(n_1+m+h_j,q_1)=1\\1\le j\le s}}^{h} 1$$

From Holder's inequality we have $D^k \le 2^k\Big(D_1^k + D_2^k\Big)$, and consequently

$$M_k^{\mathcal{D}}(q,h) \ll \sum_{n_1}\sum_{n_2} D_1^k + \sum_{n_1}\sum_{n_2} D_2^k.$$

Since D_1 is independent of n_2, we have that

$$\sum_{n_1}\sum_{n_2} D_1^k \ll q_2 P_2^{sk} M_k^{\mathcal{D}}(q_1,h),$$

which by Theorem 1.1 leads to

$$\sum_{n_1}\sum_{n_2} D_1^k \ll_k q_2 P_2^{sk} q_1 h^{k/2} P_1^{-2^{ks}+ks} = q h^{k/2} P_1^{-2^{ks}+ks} P_2^{sk}.$$

To estimate $\sum_{n_1}\sum_{n_2} D_2^k$ let

$$A_{n_1} = \left\{1 \le m \le h : (n_1+m+h_j,q_1) = 1, 1 \le j \le s\right\}.$$

Note that the size of A_{n_1} is

$$\sum_{\substack{m=1\\(n_1+m+h_j,q_1)=1\\1\le j\le s}}^{h} 1,$$

which, by a simple sieve argument, is $\ll hP_1^s$. Therefore

$$\sum_{n_2} D_2^k = \sum_{n_2}\Bigg(\sum_{\substack{m\in A_{n_1}\\(n_2+m+h_j,q_2)=1\\1\le j\le s}} 1 - \prod_{p|q_2}\left(1-\frac{\nu_p(\mathcal{D})}{p}\right)|A_{n_1}|\Bigg)^k.$$

Now, since $y > h^k$ and $p|q_2$, we have that $p > h^k$ and consequently

$$\prod_{p|q_2}\left(1-\frac{\nu_p(\mathcal{D})}{p}\right)=\prod_{p|q_2}\left(1-\frac{s}{p}\right)=\prod_{p|q_2}\left(1-\frac{1}{p}\right)^s\left(1+O\left(\frac{1}{y}\right)\right).$$

Next we need to use Theorem 2.1 with $A = A_{n_1}$ and $q = q_2$. In order to do this we need to verify that $p > \max A_{n_1} - \min A_{n_1} + h_s - h_1$ for all $p|q_2$. We have that $\max A_{n_1} - \min A_{n_1} \leq h$ and, since for $p|q_2$ we have $p > y > h^k$, it suffices to verify that $h + H < h^k$. This is true because H is fixed, $k \geq 2$ and $h > P^{-1}$. (Note that we may assume $P^{-1} > H$ else, otherwise, P^{-1} is bounded and we can deduce the result desired here from Theorem 1.1.) Using Theorem 2.1, we have that

$$\sum_{n_2} D_2^k \ll q_2\left(|A_{n_1}|\left(\frac{\phi(q_2)}{q_2}\right)^s\right)^{[k/2]}+q_2|A_{n_1}|\left(\frac{\phi(q_2)}{q_2}\right)^s.$$

Since $|A_{n_1}| \ll hP_1^s$, we have that

$$\sum_{n_2} D_2^k \ll q_2\left(h\left(\frac{\phi(q)}{q}\right)^s\right)^{[k/2]}+q_2h\left(\frac{\phi(q)}{q}\right)^s,$$

consequently we have that $\left(\text{with } P=\frac{\phi(q)}{q}\right)$

$$\sum_{n_1}\sum_{n_2} D_2^k \ll q(hP^s)^{[k/2]}+qhP^s.$$

Finally, we arrive at our principal estimate

$$M_k^{\mathcal{D}}(q,h) \ll q(hP^s)^{[k/2]}+qh(P)^s+qh^{k/2}P_1^{-2^{ks}+ks}P_2^{sk}. \tag{3.1}$$

□

CHAPTER 4

PROOF OF THEOREM 0.1

Let $a_1 < a_2 < \cdots$ be the integers, such that $a_i + h_j$ is co-prime to q for each $h_j \in \mathcal{D}$. Let

$$L(x) = \#\left\{i : 1 \leq i \leq \phi_{\mathcal{D}}(q), a_{i+1} - a_i > x\right\}.$$

We have that

$$V_\lambda^{\mathcal{D}}(q) = \lambda \int_0^\infty L(x) x^{\lambda-1} dx.$$

Obviously, $L(x) \leq \prod_{p|q}(p - \nu_p(\mathcal{D})) < CqP^s$ for some constant $C = C(\mathcal{D})$. Therefore for $x < P_{\mathcal{D}}^{-1}$ $\left(\text{with } P_{\mathcal{D}} = \prod_{p|q}\left(1 - \frac{\nu_p(\mathcal{D})}{p}\right)\right)$, since $P_{\mathcal{D}} \asymp P^s$, we have that

$$\lambda \int_0^{P_{\mathcal{D}}^{-1}} L(x) x^{\lambda-1} \ll qP^s \int_0^{P_{\mathcal{D}}^{-1}} x^{\lambda-1} \ll q(P^s)^{1-\lambda}.$$

To bound $L(x)$ for larger x, we note that if $a_{i+1} - a_i > h$, for some integer h. Then

$$\sum_{\substack{m=1 \\ (n+m+h_j,q)=1 \\ 1\leq j \leq s}}^{h} 1 - hP_{\mathcal{D}} = -hP_{\mathcal{D}}$$

for $a_i \leq n < a_{i+1} - h$. Let k be a fixed even integer bigger than 2λ. Then

$$\sum_{\substack{i=1 \\ a_{i+1}-a_i>h}}^{qP_{\mathcal{D}}} (a_{i+1} - a_i - h)(hP_{\mathcal{D}})^k \leq M_k^{\mathcal{D}}(q,h). \tag{4.1}$$

If $h = [\frac{x}{2}]$ and $a_{i+1} - a_i > x$, then $a_{i+1} - a_i - h > h$, so the left-hand side of (4.1) is $\geq h(hP_{\mathcal{D}})^k L(x)$. Combining this with our principal estimate, Theorem (3.1) yields

$$x(xP_{\mathcal{D}})^k L(x) \ll q\Big((xP_{\mathcal{D}})^{k/2} + x^{k/2}P^{ks}P_1^{-2^{ks}}\Big).$$

Now for $x < e^{P^{-\alpha}}$, if $y = x^{2k} + 1$ where, $\alpha = \frac{ks}{2^{ks+1}}$, then we have

$$P_1^{-1} = \Big(\prod_{p<y}(1 - \frac{1}{p})\Big)^{-1} \ll \log y \ll \log x \ll P^{-\alpha}.$$

Therefore

$$P_1^{-2^{ks}} \ll (P^{-\alpha})^{2^{ks}} \ll P^{-\frac{sk}{2}}.$$

So we have

$$L(x) \ll \frac{qP_{\mathcal{D}}}{(xP_{\mathcal{D}})^{\frac{k}{2}+1}}.$$

By integrating both sides we deduce that

$$\int_{P_0^{-1}}^{e^{P^{-\alpha/k}}} L(x)x^{\lambda-1}dx \ll \int_{P_{\mathcal{D}}^{-1}}^{e^{P^{-\alpha/k}}} \frac{qP_{\mathcal{D}}}{(xP_{\mathcal{D}})^{\frac{k}{2}+1}}x^{\lambda-1}dx.$$

Since $\frac{k}{2} + 1 > \lambda$ we have

$$\int_{P_{\mathcal{D}}^{-1}}^{P^{-\alpha/k}} L(x)x^{\lambda-1}dx \ll qP_{\mathcal{D}}^{1-\lambda} \ll q(P^s)^{1-\lambda}.$$

For larger x we use Theorem 1.1, which gives us that

$$M_k^{\mathcal{D}}(q,h) \ll qh^{k/2}P^{-2^{ks}+ks}.$$

Therefore we have

$$L(x) \ll \frac{qP^{-2^{ks}}}{x^{\frac{k}{2}+1}}$$

and

$$\int_{e^{P^{-\alpha/k}}}^{\infty} L(x)x^{\lambda-1}dx \ll qP^{-2^{ks}} \int_{e^{P^{-\alpha/k}}}^{\infty} \frac{x^{\lambda-1}}{x^{\frac{k}{2}+1}}dx$$

So taking $k = 2\lfloor\lambda\rfloor + 1$ implies that

$$\int_{e^{P^{-\alpha/k}}}^{\infty} L(x)x^{\lambda-1}dx \ll q\frac{P^{-2^{ks}}}{e^{P^{-\alpha/k}}} \ll q(P^s)^{1-\lambda},$$

for P^{-1} large enough, which finishes the proof.

BIBLIOGRAPHY

[1] H. Cramer, On the order of magnitude of the difference between consecutive prime numbers. Acta Arithmetica.(1936): 23-46.

[2] A. Granville, Harold Cramer and the distribution of prime numbers. Scandanavian Actuarial. J.1995, no. 1, pages 12- 28.

[3] P. Erdős, The difference of consecutive primes. Duke Math. J. 6, (1940). 438–441

[4] C.Hooley, On the difference of consecutive numbers prime to n. Acta Arith. 8 1962/1963 343–347

[5] M. Hausman and H. Shapiro, On the mean square distribution of primitive roots of unity. Comm. Pure Appl. Math. 26 (1973), 539–547.

[6] H. Montgomery and R. Vaughan, On the distribution of reduced residues. Ann. of Math. (2) 123 (1986), no. 2, 311–333.

[7] H. Halberstam and H. E. Richert, *Sieve Methods*, London Mathematical Society monographs, No 4(London, New York: Academic Press, 1974).

[8] Hugh L. Montgomery and K. Soundararajan, Primes in short intervals, Comm. Math. Phys. 252 (2004), no. 1-3, 589-617.