Higher-Dimensional Kloosterman Sums and the Greatest Prime Factor of Integers of the Form $a_1 a_2 \cdots a_{k+1} + 1$

by

Shengli Wu

A thesis
presented to the University of Waterloo
in fulfilment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Pure Mathematics

Waterloo, Ontario, Canada, 2007

I hereby declare that I am the sole author of this thesis.

I authorize the University of Waterloo to lend this thesis to other institutions or individuals for the purpose of scholarly research.

Abstract

Let N be a positive integer and let A and B be subsets of $\{1, 2, ..., N\}$. What can be deduced about the arithmetical character of the integers of the form $ab + 1$ with a in A and b in B from just information about the cardinalities of A and B? In 2000, Sárközy and Stewart conjectured,

For each positive number ϵ there are positive real numbers $N_0(\epsilon)$ and $C(\epsilon)$ such that if N exceeds $N_0(\epsilon)$ and $|A| > \epsilon N$ and $|B| > \epsilon N$. where ϵ is a real number with $0 < \epsilon < 1$ then there are a in A and b in B with

$$P(ab + 1) > C(\epsilon)N^2.$$

In this thesis we shall discuss progress towards this conjecture. We have improved or extended the work of Sárközy and Stewart in this context in some special cases. For example we proved

Let N be a positive integer and put $S = \{1, 2, \cdots, N\}$. Let α be a real number with $\alpha > 10$. When N is large enough, then there are integers $a, b \in S$, such that

$$P(ab + 1) > \frac{1}{2}\frac{N^2}{(\log N)^\alpha}.$$

We also considered the following problem,

Let N and k be positive integers and let ϵ be a real number with $0 < \epsilon < 1$. When N is large enough, for each prime p with $p < N^{2\frac{k+1}{k+2}-\epsilon}$ there exist integers a_i, $i = 1, 2, \cdots, k + 1$ with $1 \leq a_i \leq N$ such that

$$p \mid a_1 a_2 \cdots a_{k+1} + 1.$$

Acknowledgments

I would like to thank, first and foremost, my supervisor, Cameron Stewart, for his guidance and support through my doctoral studies. Especially during the final stage of my thesis. He spent hours and hours to discuss the problems involved in my thesis with me. At several occasions when I can not move on, he pointed me in the right direction and greatly helped me in getting the results in my thesis. He is an English teacher with responsibility to correct my spelling, grammar, amongst other things. On the other hand, I was impressed by his attitude toward teaching. I have taken five courses taught by him. I also want to thank the reading committee for their work and the time they spent on my thesis. I thank them for providing me with helpful suggestions to improve the thesis.

I would like to thank the staff of the Pure Mathematics Department for their help. I would like to thank Pure Mathematics Department. It is really a great experience to study here. I would like to thank all the friends, for their help and advice. Finally I would like to thank my family for continued support.

Contents

Notation

Let f be any real or complex-valued function, and let g be a positive function. We write

$$f = O(g)$$

if there exists a constant $C > 0$ such that

$$|f(x)| \leq Cg(x)$$

for all sufficiently large x in the domain of f.

For any real number α, we write $e(\alpha) = e^{2\pi i\alpha}$. Furthermore, we write $\|\alpha\|$ to denote the distance of α to the nearest integer, and $[\alpha]$ for the greatest integer less or equal to α.

We write (a, b) to denote the greatest common divisor of a and b.

Let S be a subset of $\{1, 2, \cdots, N\}$. Let $|S|$ denote the cardinality of S. If there exist a real number ϵ with $0 < \epsilon < 1$ such that

$$|S| \geq \epsilon N$$

Then we say S is a dense subset of $\{1, 2, \cdots, N\}$.

Further notations will be introduced as needed.

Chapter 1

Introduction

Let N be a positive integer and let A and B be subsets of $\{1, 2, ..., N\}$. A basic question of combinatorial number theory is the following. What can be deduced about the arithmetical character of the integers of the form $a+b$ with a in A and b in B from information about the cardinalities of A and B? There is an extensive literature addressing this problem. For example, in 1986, by means of the Hardy-Littlewood method, A. Sárközy and C. Stewart [24] proved the following:

For any set X let $|X|$ denote the cardinality of X and for any integer n, larger than one, let $P(n)$ denote the greatest prime factor of n. Let ϵ be a positive real number and suppose that

$$|A| > \epsilon N \text{ and } |B| > \epsilon N. \tag{1.1}$$

Then there is a positive number $C(\epsilon)$, which is effectively computable in terms of ϵ, such that if (1.1) holds then there exist integers a in A and b in B with

$$P(a + b) > C(\epsilon)N. \tag{1.2}$$

This result is best possible up to a determination of $C(\epsilon)$ since $a + b$ is at most $2N$. In 1992 Ruzsa [20] gave a different proof of (1.2).

One might like to study the multiplicative analogues of sum set results. One way of doing this, proposed by Sárközy [26], is to replace the sums $a + b$ by the numbers $ab + 1$(see also [13, 27]). However, it should be noted that the first result on the arithmetic properties of numbers $ab+1$ is due, probably , to Vinogradov(see Chapter V of [31]). Let p be a prime number and n be an integer coprime with

p. Let $(\frac{n}{p})$ denote the Legendre symbol of n over p. Vinogradov established the estimate

$$\left|\sum_{a\in A}\sum_{b\in B}\left(\frac{ab+k}{p}\right)\right| \leq (2|A||B|p)^{\frac{1}{2}}. \tag{1.3}$$

This result can be considered as the multiplicative analogue of a result of Friedlander and Iwaniec [11] on sums of the form $\sum_{a\in A}\sum_{b\in B}\chi(a+b)$ where χ is a non-principal character modulo a prime p.

Put

$$Z = min(|A|, |B|). \tag{1.4}$$

In 2000, Sárközy and Stewart [25] proved the following,

Theorem A*(A. Sárközy, C. Stewart) For each positive real number ϵ there are numbers $N_1(\epsilon)$ and $C(\epsilon)$ which are effectively computable in terms of ϵ such that if N exceeds $N_1(\epsilon)$ and*

$$Z > C(\epsilon)\frac{N}{\log N}, \tag{1.5}$$

then there are a in A and b in B such that

$$P(ab+1) > (1-\epsilon)Z\log N. \tag{1.6}$$

In 2001, C. Stewart [30] sharpened the lower bound. By making use of estimates for Kloosterman sums and Selberg's upper bound sieve, he proved the following theorem,

Theorem B*(C. Stewart) There are effectively computable positive numbers c_1, c_2 and c_3 such that if N exceeds c_1 and*

$$Z > c_2\frac{N}{\sqrt{(\log N)/\log\log N}}, \tag{1.7}$$

then there are a in A and b in B such that

$$P(ab+1) > N^{1+c_3(Z/N)^2}. \tag{1.8}$$

This is the best result so far on this problem.

We want to know what the best possible lower bound is for $P(ab+1)$. In their joint work in 2000, Sárközy and Stewart [25] made the following conjecture,

Conjecture 1. *(A. Sárközy, C. Stewart) For each positive number ϵ there are positive real numbers $N_0(\epsilon)$ and $C(\epsilon)$ such that if N exceeds $N_0(\epsilon)$ and (1.1) holds, then there are a in A and b in B with*

$$P(ab+1) > C(\epsilon)N^2.$$

Conjecture 2. *(A. Sárközy, C. Stewart) For each positive real number ϵ and each integer k with $k > 2$, there are positive real numbers $N_0(\epsilon, k)$ and $c(\epsilon, k)$ such that if N exceeds $N_0(\epsilon, k)$ and (1.1) holds, then there are a in A and b in B and a prime p with*

$$p^k | ab+1 \text{ and } p^k > c(\epsilon, k)N^2.$$

We will give an explicit c_3 for Theorem B of Stewart. We shall prove that one may take $c_3 = 0.055$.

Theorem 1. *Let N be a positive integer, let A and B be two subsets of $\{1, 2, \cdots, N\}$ and put $Z = min(|A|, |B|)$. There are effectively computable positive numbers c_1 and c_2 such that if N exceeds c_1 and (1.7) holds, then there are a in A and b in B such that*

$$P(ab+1) > N^{1+0.055(Z/N)^2}. \tag{1.9}$$

We are unable to improve Stewart's result about the lower bound of $P(ab+1)$ for dense subsets A and B of the set $\{1, 2, \cdots, N\}$. However we are able to give a much better lower bound in some special cases. Instead of looking at the dense subsets A and B, we consider a and b both from $\{1, 2, \cdots, N\}$. In this case we may consider the primes in the arithmetic progression

$$mq + 1$$

with m varying from 1 up to N.

Define

$$\psi(x; q, a) = \sum_{\substack{n \le x \\ n \equiv a (mod\ q)}} \Lambda(n),$$

where $\Lambda(n)$ is the von Mangoldt Function. I.e.

$$\Lambda(n) = \begin{cases} \log n & \text{if n is a power of a prime number,} \\ 0 & \text{otherwise.} \end{cases}$$

We quote one version of the prime number theorem for arithmetic progressions in the following.

Theorem C *(The prime number theorem for arithmetic progressions) Let x be a positive real number, q and a be positive integers and suppose*

$$q \leq (\log x)^{1-\delta}, \tag{1.10}$$

where δ is a fixed real number such that $0 < \delta < 1$. Then

$$\psi(x; q, a) = \frac{x}{\phi(q)} + O\left(xe^{-c(\log x)^{\frac{1}{2}}}\right), \tag{1.11}$$

where c is an absolute constant.

This is a weak result but it is effective, in the sense that, if δ is given a numerical value, both c and the constant implied by the symbol O can be given numerical values. Applying this result, we deduced the following,

Let N be a positive integer and let ϵ be a real number such that $0 < \epsilon < 1$. There exists a positive integer N_1 such that when $N > N_1$, there exists a prime number p and integers a and b with $1 \leq a, b \leq N$ such that $p = ab + 1$ and

$$p > \frac{1}{2} N(\log N)^{1-\epsilon}. \tag{1.12}$$

To have a better estimate for the size of p in (1.12), we need better estimates for the number of primes in an arithmetic progression. Let χ be a Dirichlet character (mod q). The Dirichlet L-function is defined by

$$L(\chi, s) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s} \tag{1.13}$$

for every complex number s with real part bigger than 1. By analytic continuation, this function can be extended to a meromorphic function defined on the whole complex plane. Here we quote the following result which was probably formulated for the first time by Piltz in 1884:

The Generalized Riemann Hypothesis (GRH) asserts that for every Dirichlet character χ and every complex number s with $L(\chi, s) = 0$: if the real part of s is between 0 and 1, then it is actually 1/2.

The case $\chi(n) = 1$ for all n gives the ordinary Riemann Hypothesis.

Theorem D *If GRH is true, then for $q \leq x$,*

$$\pi(x; q, a) = \frac{Li(x)}{\phi(q)} + O\left(x^{\frac{1}{2}} \log x\right) \tag{1.14}$$

where

$$Li(x) = \int_2^x \frac{1}{\log u} du,$$

and

$$\pi(x; q, a) = |\ \{\ p \mid p \leq x,\ p\ \equiv a\ (mod\ q)\ \}\ |.$$

By Theorem D, that is, with the assumption of GRH, we can prove the following result.

Theorem 2. *Let N be a positive integer, let α be a real number with $\alpha > 4$, and let $q = [\frac{N}{(\log N)^\alpha}]$. Put $x_1 = Nq + 1$ and $x_2 = \frac{N}{2}q + 1$. If GRH is true, we have*

$$\begin{aligned} & \pi(x_1; q, 1) - \pi(x_2; q, 1) \\ \geq\ & \frac{1}{4}\frac{N}{\log N} + O\left(\frac{N \log\log N}{(\log N)^2} + \frac{N}{(\log N)^{\frac{\alpha}{2}-1}}\right). \end{aligned}$$

We have the following corollary,

Corollary 1. *Let N be an integer and put $S = \{1, 2, \cdots, N\}$. Let α be a real number with $\alpha > 4$. Assume GRH. When N is large enough there are integers $a, b \in S$ such that*

$$P(ab + 1) > \frac{1}{2}\frac{N^2}{(\log N)^\alpha}.$$

Note upon choosing suitable α, the factor $\frac{1}{2}$ may be removed.

Furthermore, we may appeal to the Bombieri-Vinogradov theorem on the average of primes in arithmetic progressions. We have unconditionally proved a slightly weaker result than Theorem 2. It is proved without GRH. We put

$$E(x; q, a) = \psi(x; q, a) - \frac{x}{\phi(q)}$$

for $(a,q)=1$, we let

$$E(x;q)=\max_{\substack{a\\(a,\ q)=1}}|E(x;q,a)|,$$

and

$$E^*(x,q)=\max_{y\le x}E(y,q).$$

Theorem E *(Bombieri-Vinogradov) Let $\alpha>0$ be fixed. Then*

$$\sum_{q\le Q}E^*(x,q)\ll x^{\frac{1}{2}}Q(\log x)^5$$

provided that $x^{\frac{1}{2}}(\log x)^{-\alpha}\le Q\le x^{\frac{1}{2}}$.

Proof. See Davenport's Multiplicative Number Theory [10]. p161.

Theorem 3. *Let α and β be two positive real numbers with $\alpha<\beta$. Put*

$$x_1=\frac{N^2}{(\log N)^\beta}+1$$

and

$$x_2=\frac{N^2}{2(\log N)^\beta}+1.$$

Then we have

$$\begin{aligned}&\sum_{\frac{N}{(\log N)^\beta}<q\le\frac{N}{(\log N)^\alpha}}(\psi(x_1;q,1)-\psi(x_2;q,1))\\&\ge\frac{\beta-\alpha}{2}\frac{N^2}{\log^\beta N}(\log\log N)+O\left(\frac{N^2}{\log^{\frac{\beta}{2}+\alpha-5}}\right),\end{aligned}\tag{1.15}$$

If $\frac{\beta}{2}+5<\alpha$, then the main term in (1.15) dominates. That is, there is a prime number p with $p>x_2$ in an arithmetic progression $mq+1$ with m between 1 and N and q in the range

$$\frac{N}{(\log N)^\beta}<q\le\frac{N}{(\log N)^\alpha}.$$

If N is large enough, then since $p > x_2$ we have

$$p > \frac{1}{2}\frac{N^2}{\log^\beta N}.$$

Note $\frac{\beta}{2} + 5 < \alpha < \beta$ so $\beta > 10$ so this is a weaker result than Corollary 1 which we proved under the assumption of GRH.

Corollary 2. *Let N be an integer and put $S = \{1, 2, \cdots, N\}$. Let β be a real number with $\beta > 10$. When N is large enough there are integers $a, b \in S$ such that*

$$P(ab+1) > \frac{1}{2}\frac{N^2}{(\log N)^\beta}.$$

Equivalently if β is a real number larger than 10 then for N sufficiently large

$$P\left(\prod_1^N \prod_1^N (ab+1)\right) > \frac{1}{2}\frac{N^2}{(\log N)^\beta}.$$

Note again the factor $\frac{1}{2}$ may be removed.

On the other hand, we also discuss an interesting implication from Stewart's proof [30]. To prove Lemma 2 in his paper, Stewart introduced the following set:

$$U_t(N) = \{\ (m,n) \in \mathbb{Z} \times \mathbb{Z} \mid 1 \le m \le N,\ 1 \le n \le N,\ t \mid mn+1\ \} \tag{1.16}$$

where $\mathbb{Z}$ denotes the set of integers.

If t divides N then we may decompose $\{(m,n)|\ 1 \le m,\ n \le N\}$ into $(N/t)^2$ blocks consisting of the Cartesian product of two complete sets of residues modulo t. Thus if t divides N, $|U_t(N)| = \phi(t)(N/t)^2$. In general we deduce that

$$|U_t(N)| = \phi(t)\left(\frac{N}{t} + O(1)\right)^2 = \frac{\phi(t)}{t^2}N^2 + O(N). \tag{1.17}$$

By a simple calculation we find that the above estimate (1.17) is dominated by the main term for t up to N, which is trivial. To get a sharper estimate Stewart appealed to Weil's estimates for Kloosterman sums [33]. For any positive integer n let $d(n)$ denote the number of divisors of n. He proved,

Theorem F *(C. Stewart)*

$$|U_t(N)| = \frac{\phi(t)}{t^2}N^2 + O\left(t^{\frac{1}{2}}d(t)^{3/2}(\log t)^2 + \frac{Nd(t)\log t}{t}\right). \tag{1.18}$$

Proof. See [30]. □

This estimate is dominated by the main term for t up to $N^{(\frac{4}{3}-\epsilon)}$. In other words, for all the integers t up to $N^{(\frac{4}{3}-\epsilon)}$, we can find integers m and n with $1 \leq m, n \leq N$ such that $t \mid mn+1$. This assures us that for all the prime numbers p up to $N^{(\frac{4}{3}-\epsilon)}$, we can find integers m and n with $1 \leq m, n \leq N$ such that $p|mn+1$. In particular,

$$P(mn+1) \geq N^{4/3-\epsilon}.$$

The proof of Theorem F involves the estimate of a special sum called a Kloosterman sum.

For each integer a coprime with t let $\bar{a}$ denote the integer from $\{1, \cdots, t\}$ for which $a\bar{a} \equiv 1(mod\ t)$. For integers g, h and t, the Kloosterman sum $S(g,h;t)$ is defined by

$$S(g,h;t) = \sum_{\substack{a=1 \\ a\bar{a}\equiv 1(mod\ t)}}^{t} e\left(\frac{ga+h\bar{a}}{t}\right). \tag{1.19}$$

They are named for the Dutch mathematician Hendrik Kloosterman, who introduced them in 1926 [18] when he adapted the Hardy-Littlewood circle method to solve a problem of Ramanujan of representing sufficiently large numbers in the form $ax^2 + by^2 + cz^2 + dt^2$. In his paper he proved that

$$S(g,h;t) = O(t^{\frac{3}{4}+\epsilon}(g,t)^{\frac{1}{4}}). \tag{1.20}$$

These sums turn out to have close connections with modular forms, and various analytic number theory techniques are used to provide estimates for the coefficients of modular forms starting with estimates for Kloosterman sums.

In 1931, H. Salié [23] proved that if $m \geq 2$ then

$$|S(g,h;p^m)| \leq Cp^{\frac{1}{2}m}, \tag{1.21}$$

where C is an absolute constant. But his argument gave no information for $m = 1$.

In 1933, Davenport [7] proved that for any prime p with $(gh,p) = 1$,

$$S(g,h;p) = O(p^{\frac{2}{3}}). \tag{1.22}$$

In 1934, Hasse [15] showed that if the Riemann Hypothesis for algebraic function fields in one variable over finite fields is true, then

$$|S(g,h;p)| \leq 2\sqrt{p}. \tag{1.23}$$

In 1941,Weil [34] established the truth of this hypothesis and thereby established (1.23) and he published the proof [33] in 1948.

Using the Hasse-Weil result in (1.23) and Salié's result in (1.21), in 1961, T. Estermann [9] (see also Hooley [16]) proved that

$$|S(g,h;t)| \leq d(t)\ t^{\frac{1}{2}}\ (g,t)^{\frac{1}{2}} \tag{1.24}$$

for all g, h and t.

In a paper in 1958, Mordell [19] introduced a generalization of the Kloosterman sum and conjectured that it satisfies an upper bound analogous to (1.23). In 1977, Deligne [4] established such a bound for the Mordell generalized Kloosterman sum. Indeed, Deligne's proof appeals to his 1974 work [5] in which he established the last portion of the Weil Conjectures [32], i.e., the Riemann Hypothesis for algebraic varieties over finite fields.

Let $l_i,\ i = 1, \cdots, k+1$ be integers such that $(l_i, p) = 1$. The k-dimensional Kloosterman sum is defined as follows:

$$K(l_1, \cdots, l_{k+1}; p) = \sum_{\substack{a_1,\cdots,a_{k+1}=1 \\ a_1 \cdots a_{k+1} \equiv 1 (mod\ p)}}^{p} e\left(\frac{l_1 a_1 + \cdots + l_{k+1} a_{k+1}}{p}\right),$$

where $a_i,\ i = 1, \cdots, k+1$ are integers.

When $k = 1$, we have

$$S(l_1, l_2; p) = \sum_{\substack{a_1,a_2=1 \\ a_1 a_2 \equiv 1 (mod\ p)}}^{p} e\left(\frac{l_1 a_1 + l_2 a_2}{p}\right) = \sum_{\substack{a=1 \\ (a,p)=1}}^{p} e\left(\frac{l_1 a + l_2 \bar{a}}{p}\right),$$

where $\bar{a}$ is such that $1 \leq \bar{a} < p$ and $a\bar{a} \equiv 1 (mod\ p)$. This is the standard Kloosterman sum. It is 1-dimensional.

For the 2-dimensional Kloosterman sum, Mordell [19] conjectured that for all primes p,

$$|K(l_1, l_2, l_3; p)| \leq cp \qquad if \quad p \nmid l_1 l_2 l_3,$$

where c is an absolute constant.

This may be considered the analogue of the 1-dimensional upper bound in (1.23). Using (1.23), Carlitz [3] proved that there is an absolute constant c such that

$$K(l_1, l_2, l_3; p) \leq cp^{\frac{5}{4}} \qquad if \quad p \nmid l_1 l_2 l_3.$$

And finally, Deligne proved [4](see also Bombieri [2] and Serre [28]) that if $p \nmid l_1 \cdots l_{k+1}$ then

$$|K(l_1, \cdots, l_{k+1}; p)| \leq (k+1)p^{\frac{k}{2}} \tag{1.25}$$

holds for all k, p.

A result analogous to Estermann's result in (1.24) was given by Smith [29] in 1979 and we shall state it below. Smith deduced his result from Deligne's Theorem.

For each prime p, and any $(a_1, a_2, \cdots, a_{n+1}) \in Z^{n+1}$, there exists a unique integer $\gamma \geq 0$ such that

$$(a_1, \cdots, a_{n+1}) = p^{\gamma}(b_1, \cdots, b_{n+1})$$

for some $(b_1, \cdots, b_{n+1}) \in Z^{n+1} - pZ^{n+1}$. Let s denote the number of b_i, $i = 1, \cdots, n+1$ which are divisible by p so that $0 \leq s \leq n$. For each $\alpha \geq 0$, define

$$(a_1, \cdots, a_{n+1}; p^{\alpha+1})_n = p^{\sigma_n(a_1, \cdots, a_{n+1}; p^{\alpha}),}$$

where

$$\sigma_n(a_1, \cdots, a_{n+1}; p^{\alpha}) = \begin{cases} \alpha & \text{if} \gamma \geq \alpha \\ \gamma & \text{if} \gamma < \alpha - 1 \\ \gamma & \text{if} \gamma = \alpha - 1 \text{ and } s = 0 \\ \gamma - 1 + \frac{2(s-1)}{n} & \text{if } \gamma = \alpha - 1 \text{ and } 1 \leq s \leq n. \end{cases}$$

(Note that $\sigma_n(a_1, \cdots, a_{n+1}; p^{\alpha}) \leq \alpha$ for all $(a_1, \cdots, a_{n+1})$).

For any $q \geq 1$ we now define

$$(a_1, \cdots, a_{n+1};\ q)_n = \prod_{p^{\alpha} \| q} (a_1, \cdots, a_{n+1};\ p^{\alpha})_n.$$

We have

Lemma 1. $(a_1, \cdots, a_{n+1}; p^\alpha)_n^{\frac{n}{2}} \leq (a_1, p^\alpha)^{\frac{1}{2}} \cdots (a_{n+1},\ p^\alpha)^{\frac{1}{2}}$.

Proof. Let

$$(a_1, \cdots, a_{n+1}) = p^\gamma (b_1, \cdots, b_{n+1})$$

where

$$(b_1, \cdots, b_{n+1}) \in Z^{n+1} - pZ^{n+1}.$$

If $\gamma \geq \alpha$, then according to the definition of $\sigma_n(a_1, \cdots, a_{n+1};\ p^\alpha)_n$,

$$(a_1, \cdots, a_{n+1}; p^\alpha)_n^{\frac{n}{2}} = (p^\alpha)^{\frac{n}{2}} \leq (a_1, p^\alpha)^{\frac{1}{2}} \cdots (a_{n+1}, p^\alpha)^{\frac{1}{2}} = (p^\alpha)^{\frac{n+1}{2}}.$$

If $\gamma < \alpha - 1$ then

$$(a_1, \cdots, a_{n+1}; p^\alpha)_n^{\frac{n}{2}} = (p^\gamma)^{\frac{n}{2}},$$

but $(a_i, p^\alpha)^{\frac{1}{2}} \geq (p^\gamma)^{\frac{1}{2}}$ for $i = 1, \cdots, n+1$. Hence

$$(a_1, \cdots, a_{n+1}; p^\alpha)_n^{\frac{n}{2}} \leq (a_1, p^\alpha)^{\frac{1}{2}} \cdots (a_{n+1}, p^\alpha)^{\frac{1}{2}}.$$

Similarly we have if $\gamma = \alpha - 1$ and $s = 0$ (s is the number of b_i, $i = 1, \cdots, n+1$, such that $p|b_i$)

$$(a_1, \cdots, a_{n+1}; p^\alpha)_n^{\frac{n}{2}} < (a_1, p^\alpha)^{\frac{1}{2}} \cdots (a_{n+1}, p^\alpha)^{\frac{1}{2}}.$$

If $\gamma = \alpha - 1$ and $1 \leq s \leq n$

$$(a_1, \cdots, a_{n+1}; p^\alpha)_n^{\frac{n}{2}} = (p^{\gamma - 1 + \frac{2(s-1)}{n}})^{\frac{n}{2}} = p^{\frac{n\gamma - n + 2s - 2}{2}}.$$

But since

$$(a_1, p^\alpha)^{\frac{1}{2}} \cdots (a_{n+1}, p^\alpha)^{\frac{1}{2}} \geq (p^\gamma)^{\frac{n+1}{2}} (p^s)^{\frac{1}{2}} = p^{\frac{n\gamma+\gamma}{2} + \frac{s}{2}},$$

We once again have

$$(a_1, \cdots, a_{n+1}; p^\alpha)_n^{\frac{n}{2}} < (a_1, p^\alpha)^{\frac{1}{2}} \cdots (a_{n+1}, p^\alpha)^{\frac{1}{2}}$$

This finishes the proof of Lemma 1. □

Let n and q be positive integers and define

$$d_{n+1}(q) = |\{(d_1, d_2, \cdots, d_{n+1}) : \ 1 \leq d_i \leq q, \ d_i | q, \ 1 \leq i \leq n+1\}|.$$

That is, $d_{n+1}(q)$ denotes the number of representations of q as a product of $n+1$ ordered positive factors.

Theorem G(R. Smith) *For all $n, q \geq 1$, and all $(a_1, \cdots, a_{n+1}) \in Z^{n+1}$, we have*

$$|K(a_1, \cdots, a_{n+1}; q)| \leq \ q^{\frac{n}{2}} (a_1, \cdots, a_{n+1}; q)_n^{\frac{n}{2}} \ d_{n+1}(q).$$

Proof. See [29]. □

Although the result of Deligne (1.25) is deep, we have not been able to find many applications of it. In the following, we looked at a "higher dimensional" case of the set $U_t(N)$ in (1.16). By making use of Deligne's and Smith's results, we are able to give multi-dimensional analogues of Stewart's result (1.18).

Let N and k be two integers and let t be a positive integer. Define

$$U_t(N, k) = \{(a_1, \cdots, a_{k+1}) | 1 \leq a_1, \cdots, a_{k+1} \leq N, t | a_1 \cdots a_{k+1} + 1\},$$

where a_i, $i = 1, \cdots, k+1$ are integers.

Let $\binom{n}{l}$ be the binomial coefficient. By making use of Deligne's result (1.25) we have proved the following:

Theorem 4. *Let N and k be positive integers and let p be a prime. We have*

$$|U_p(N, k)| = \frac{1}{p-1} \left(N - \left[\frac{N}{p} \right] \right)^{k+1} + O(kp^{\frac{k}{2}} \log^{k+1} p).$$

Further, if $p \mid N$ then

$$|U_p(N, k)| = \left(\frac{N}{p} \phi(p) \right)^k \frac{N}{p} = \frac{N^{k+1}}{p^{k+1}} \phi^k(p) = \frac{N^{k+1}}{p-1} (1 - \frac{1}{p})^{k+1}.$$

It follows from a paper of Fouvry and Katz [10] that if $p \geq N$ one has

$$|U_p(N,k)| = \frac{N^{k+1}}{p^{k+1}}(p-1)^k + O\left(kp^{\frac{k}{2}}\log^{k+1} p + \frac{N^k}{\sqrt{p}}(\log p)^{-k}\right).$$

So our result is an improvement of theirs. We have a more explicit main term and as a consequence we have a more precise result.

We shall get asymptotic estimates for $|U_t(N,k)|$ by using Smith's result on higher-dimensional Kloosterman sums. However, we shall first look at the case when t is a prime power.

Theorem 5. *Let N, k and m be positive integers and let p be a prime. Then we have*

$$|U_{p^m}(N,k)| = \frac{1}{\phi(p^m)}\left(N - \left[\frac{N}{p}\right]\right)^{k+1} + O\left(d_{k+1}(p^m)(p^m)^{\frac{k}{2}}\log^{k+1} p^m\right).$$

The main term dominates for $p^m < N^{\frac{2(k+1)}{k+2}-\epsilon}$.

Secondly we have the general result. It is weaker than Theorem 4 and Theorem 5. We shall discuss this later in the proof.

Theorem 6. *Let N, k and t be positive integers, then*

$$|U_t(N,k)| = \frac{N^{k+1}}{t^{k+1}}\phi^k(t) + O\left(\frac{N^k}{t}d(t)\log t + t^{\frac{k}{2}}2^{k+1}d^{\frac{k+1}{2}}(t)d_{k+1}(t)\log^{k+1} t\right).$$

Here once again, the main term dominates for $t < N^{2\frac{k+1}{k+2}-\epsilon}$. We have the following corollary,

Corollary 3. *Let N, k be integers. Let ϵ be a real number with $0 < \epsilon < 1$. For each positive integer t with $t < N^{2\frac{k+1}{k+2}-\epsilon}$, we can find integers $a_1, \cdots, a_{k+1}$ with $1 \leq a_i \leq N$, $i = 1, 2, \cdots, k+1$ such that*

$$t \mid a_1 \cdots a_{k+1} + 1.$$

Chapter 2

An Explicit Constant

As mentioned above, our first goal is to find an explicit c_3 in Stewart's theorem. For comparison, we repeat Stewart's theorem here again.

Theorem B *(C. Stewart) There are effectively computable positive numbers c_1, c_2 and c_3 such that if N exceeds c_1 and*

$$Z > c_2 \frac{N}{\sqrt{(\log N)/\log\log N}},$$

then there are a in A and b in B such that

$$P(ab+1) > N^{1+c_3(Z/N)^2}.$$

Stewart's proof employed a strategy first introduced by C. Hooley [17] for his proof that $P(n^2+1)$ exceeds $n^{\frac{11}{10}}$ for infinitely many integers n. More exactly, he used estimates of Kloosterman sums and Selberg's upper bound sieve. The application of the sieve method is similar to that of Greaves [12] in that they sieve a subset of $\mathbb{Z}\times\mathbb{Z}$ not a set of integers.

I shall repeat most of his proof. However we need to expand the proof to get an explicit result. We have

Theorem 1 *Let N be a positive integer, let A and B be two subsets of $\{1, 2, \cdots, N\}$ and put $Z = min(|A|, |B|)$. There are effectively computable positive numbers c_1 and c_2 such that if N exceeds c_1 and* (1.7) *holds, then there are a in A and b in B such that*

$$P(ab+1) > N^{1+0.055(Z/N)^2}. \tag{2.1}$$

Define E by

$$E = \prod_{a \in A, b \in B} (ab + 1) \tag{2.2}$$

and put

$$E_1 = \prod_{p \leq N} p^{ord_p E}, \tag{2.3}$$

where the product is taken over all primes p up to N and $ord_p E$ denotes the highest power of p that divides E.

Lemma 2. *Let $\epsilon > 0$, there exists a positive number $N_0(\epsilon)$, which is effectively computable in terms of ϵ, such that for $N > N_0$,*

$$\log E_1 < (1 + \epsilon) Z^2 \log N. \tag{2.4}$$

Proof. See lemma 1 in [25]. □

Let t and N be positive integers and let z be an integer with $1 \leq z \leq N$. We put $U_t(N, z)$ to be the following set

$$\left\{ (m, n) | 1 \leq m, n \leq N, t | mn + 1, \ all \ prime \ factors \ of \ \frac{mn+1}{t} \ exceed \ z \right\}$$

In [30], Stewart gave an estimate of $|U_t(N, z)|$ by means of Selberg's upper bound sieve. Before we state his result, we need the following:

Let $t > 0$ be an integer. We define $f_t(d)$ for each integer d by

$$f_t(d) = \frac{d}{\prod_{\substack{p|d \\ p \nmid t}} (1 - \frac{1}{p}).} \tag{2.5}$$

Observe that $f_t(d)$ is multiplicative and, for each positive integer n, put

$$g_t(n) = f_t(n) \prod_{p|n} (1 - \frac{1}{f_t(p)}). \tag{2.6}$$

Since

$$f_t(p) = \begin{cases} p & \text{if } p|t \\ \frac{p^2}{p-1} & \text{if } p \nmid t, \end{cases}$$

we see that

$$g_t(n) = n\prod_{\substack{p|n\\ p|t}}(1-\frac{1}{p})\prod_{\substack{p|n\\ p\nmid t}}(1+\frac{1}{p(p-1)}). \tag{2.7}$$

So

$$\begin{aligned}
g_t(n) &\le n\prod_{\substack{p|n\\ p\nmid t}}(1+\frac{1}{p(p-1)})\\
&\le n\prod_{p|n}(1+\frac{1}{p(p-1)})\\
&\le n\prod_{\substack{p|n\\ p\le 41}}(1+\frac{1}{p(p-1)})\prod_{\substack{p|n\\ p>41}}\left(1+\frac{1+\frac{1}{p-1}}{p^2}\right)\\
&\le n\prod_{\substack{p|n\\ p\le 41}}(1+\frac{1}{p(p-1)})\prod_{\substack{p|n\\ p>41}}(1+\frac{1}{p^2})^2\\
&< n\prod_{\substack{p|n\\ p\le 41}}(1+\frac{1}{p(p-1)})\prod_{p>41}(1+\frac{1}{p^2})^2\\
&< n\prod_{\substack{p|n\\ p\le 41}}(1+\frac{1}{p(p-1)})\prod_{p}(1+\frac{1}{p^2})^2\prod_{p\le 41}(1+\frac{1}{p^2})^{-2}\\
&= n\prod_{\substack{p|n\\ p\le 41}}(1+\frac{1}{p(p-1)})(\frac{\pi^2}{6})^2\prod_{p\le 41}(1+\frac{1}{p^2})^{-2}\\
&\le 2.287743n.
\end{aligned} \tag{2.8}$$

Lemma 3. *For each integer $z \ge 2$ and each integer $t > 0$, we put*

$$V_t(z) = \sum_{n\le z}\frac{\mu^2(n)}{g_t(n)}. \tag{2.9}$$

We have

$$V_t(z) > \frac{2.6226}{\pi^2}\log z \tag{2.10}$$

provided that z is large enough.

Proof. From (2.8) and (2.9) we get

$$V_t(z) = \sum_{n\le z} \frac{\mu^2(n)}{g_t(n)} \ge \sum_{n\le z} \frac{\mu^2(n)}{2.287743n}, \tag{2.11}$$

By Abel's summation formula, we further have

$$\begin{aligned} V_t(z) &\ge \frac{1}{2.287743} \sum_{n\le z} \frac{\mu^2(n)}{n} \\ &= \frac{1}{2.287743} \left(\frac{1}{z} \sum_{n\le z} \mu^2(n) + \int_1^z \frac{\sum_{n\le u} \mu^2(n)}{u^2} \right) \\ &= \frac{1}{2.287743} \left(\frac{1}{z} \left(\frac{6}{\pi^2} z + O(\sqrt{z}) \right) \right) + \frac{1}{2.287743} \int_1^z \frac{\frac{6}{\pi^2} u + O(\sqrt{u})}{u^2} \\ &= \frac{1}{2.287743} \frac{6}{\pi^2} \log z + \frac{1}{2.287743} \frac{6}{\pi^2} + O(\frac{1}{\sqrt{z}}) \\ &> \frac{2.6226}{\pi^2} \log z, \end{aligned} \tag{2.12}$$

provided that z is large enough. Here we have used the known result that

$$\sum_{n\le x} \mu^2(n) = \frac{6}{\pi^2} x + O(\sqrt{x}).$$

(See theorem 334 in [14]) □

We can now state Stewart's upper bound for $U_t(N, z)$.

Lemma 4. *Let $\epsilon > 0$ and let N, t and z be positive integers with $t > N^{\frac{2}{3}}$ and $z \ge 2$. Then*

$$|U_t(N, z)| \le \frac{N^2 \phi(t)}{V_t(z) t^2} + O_\epsilon \left((t^{\frac{1}{2}} z^3)^{1+\epsilon} \right). \tag{2.13}$$

Proof. See lemma 2 in [30]. □

Now we are ready to prove Theorem 1.

Let $\epsilon > 0$ and let $N_0, N_1, N_2, \cdots$, denote positive numbers which are effectively computable in terms of ϵ. Define E by (2.2) and E_1 by (2.3) and put $E_2 = E/E_1$. The proof proceeds by a comparison of estimates for E.

Proof. Note that

$$\begin{aligned} E &\geq \prod_{\substack{a\in A\\ a\geq \frac{\epsilon Z}{10}}} \prod_{\substack{b\in B\\ b\geq \frac{\epsilon Z}{10}}} \left((\frac{\epsilon Z}{10})^2 + 1 \right) \\ &\geq (\frac{\epsilon Z}{10})^{2(|A|-\frac{\epsilon Z}{10})(|B|-\frac{\epsilon Z}{10})} \\ &\geq (\frac{\epsilon Z}{10})^{2(1-\frac{\epsilon}{10})^2 Z^2}, \end{aligned}$$

Thus by (1.7),

$$\log E > (2-\epsilon) Z^2 \log N,$$

for $N > N_0$. By Lemma 2,

$$\log E_1 < (1+\epsilon) Z^2 \log N$$

for $N > N_1$. Hence, for $N > N_2$,

$$\log E_2 > (1-2\epsilon) Z^2 \log N. \tag{2.14}$$

Let P denote the greatest prime factor of E, then

$$E_2 \leq \prod_{N\leq p\leq P} p^{ord_p G} \tag{2.15}$$

where

$$G = \prod_{1\leq m,n\leq N} (mn+1). \tag{2.16}$$

Put $P = NY$ and note that

$$\sum_{N<p\leq NY} ord_p G \ \log p = \sum_{N<p\leq NY} \sum_{\substack{1\leq m,n\leq N\\ p|mn+1}} \log p. \tag{2.17}$$

Observe that

$$\begin{aligned} &\sum_{N<p\leq NY} \sum_{\substack{1\leq m,n\leq N\\ p|mn+1\\ mn+1\leq \frac{N^2}{(\log N)^2}}} \log p \\ &\leq 2 \sum_{\substack{1\leq m,n\leq N\\ mn+1\leq \frac{N^2}{(\log N)^2}}} \log N \\ &= O(N^2) = o(Z^2 \log N) \end{aligned} \tag{2.18}$$

and so, by (2.14), (2.15), (2.16) ,(2.17) and (2.18) , for $N > N_2$,

$$\sum_{N<p\le NY} \sum_{\substack{1\le m,n\le N \\ p|mn+1 \\ mn+1>\frac{N^2}{(\log N)^2}}} \log p > (1-3\epsilon)Z^2 \log N. \tag{2.19}$$

Put

$$S_t(N) = \sum_{\substack{1\le m,n\le N \\ mn+1=tp \\ N<p\le NY \\ mn+1>\frac{N^2}{(\log N)^2}}} 1.$$

If $Y > N^{\frac{1}{20}}$ then $P = NY = N^{21/20}$ then the result holds and so we may suppose that $Y \le N^{\frac{1}{20}}$. Then by (2.19)

$$\sum_{\frac{N}{Y(\log N)^2}<t\le N} \log(\frac{N^2+1}{t}) S_t(N) > (1-3\epsilon)Z^2 \log N, \tag{2.20}$$

and so, for $N > N_3$

$$\sum_{\frac{N}{Y(\log N)^2}<t\le N} S_t(N) > \frac{19.5}{21} Z^2. \tag{2.21}$$

For each real number z with $2 \le z \le N$

$$S_t(N) \le U_t(N,N) \le U_t(N,z). \tag{2.22}$$

Let $c_4, c_5, \cdots$ denote effectively computable positive numbers. By Lemma 3, we have

$$V_t(z) > \frac{2.6226}{\pi^2} \log z. \tag{2.23}$$

We now apply Lemma 4 with $z = N^{\frac{1}{7}}$ and $\epsilon = \frac{1}{20}$ to conclude from (2.22) and (2.23), that for $N > N_3$ and $\frac{N}{Y(\log N)^2} < t \le N$,

$$\begin{aligned} S_t(N) &\leq \frac{N^2\phi(t)}{V_t(z)t^2} + O_\epsilon\left((t^{\frac{1}{2}}z^3)^{1+\epsilon}\right) \\ &< \frac{N^2}{\frac{2.6226}{\pi^2}\log z}\cdot\frac{\phi(t)}{t^2} + O_\epsilon\left((t^{\frac{1}{2}}z^3)^{1+\epsilon}\right) \\ &< \frac{N^2}{\frac{1}{7}\frac{2.6225}{\pi^2}\log N}\cdot\frac{\phi(t)}{t^2} \\ &= \frac{7\pi^2N^2}{2.6225\log N}\cdot\frac{\phi(t)}{t^2} \end{aligned}$$

and so

$$\sum_{\frac{N}{Y(\log N)^2}<t\leq N} S_t(N) < \frac{7\pi^2N^2}{2.6225\log N}\sum_{\frac{N}{Y(\log N)^2}<t\leq N}\frac{\phi(t)}{t^2} \tag{2.24}$$

Note that (See [1])

$$\sum_{t\leq x}\frac{\phi(t)}{t^2} = \frac{6}{\pi^2}\log x + A + O\left(\frac{\log x}{x}\right)$$

where A is a constant. Therefore we have

$$\begin{aligned} \sum_{\frac{N}{Y(\log N)^2}<t\leq N} S_t(N) &< \frac{7\pi^2N^2}{2.6225\log N}\left(\frac{6}{\pi^2}(\log Y + 2\log\log N) + O(\frac{Y(\log N)^3}{N})\right) \\ &< \frac{42N^2}{2.6225\log N}(\log Y + 2.1\log\log N) \end{aligned} \tag{2.25}$$

provided that N is large enough.

We now can choose suitable c_2 in (1.7) such that

$$\frac{2.1\cdot 42N^2}{2.6225\log N}\log\log N < \frac{0.5}{21}Z^2. \tag{2.26}$$

Hence the first term in (2.24), from (2.21), (2.24) and (2.25), satisfies

$$\frac{42N^2}{2.6225\log N}\log Y > \frac{19}{21}Z^2$$

and this gives us

$$Y > N^{\frac{19 \cdot 2.6225}{21 \cdot 42}(Z/N)^2},$$

where $\frac{19}{21} \cdot \frac{2.6225}{42} > 0.055$

This finishes the proof of Theorem 1. □

Chapter 3

Primes in Arithmetical Progressions

3.1 Our plan

As we mentioned before, we are unable to improve the lower bound for the greatest prime factor of $ab+1$ with a and b from two dense subsets of $\{1, 2, \cdots, N\}$ apart from the constant in the exponent. It may be worthwhile to mention here that C. Hooley [17] proved that $P(n^2+1)$ exceeds $n^{\frac{11}{10}}$ infinitely many times. Later, it was refined to $n^{\frac{12}{10}}$ by J. Deshouillers and H. Iwaniec [6]. However, even if the conjecture that there are infinitely many primes of the form n^2+1 is true, it does not imply the conjecture of Sárközy and Stewart. The conjecture will be difficult to prove.

This leads us to look at the case when $A = B = \{1, 2, \cdots, N\}$. In other words, we want to know what kind of lower bound we could have for $P(ab+1)$ when a and b are both from the set $\{1, 2, \cdots, N\}$. We then can immediately look at the primes in the arithmetical progression

$$mq + 1$$

where q is fixed such that $1 \leq q \leq N$ and $m = 1, 2, \cdots, N$.

Let ϵ be a fixed real number such that $0 < \epsilon < 1$ and $q = [(\log N)^{1-\epsilon}]$. Let $x = Nq + 1$, then

$$q < (\log x)^{1-\epsilon},$$

By (1.11), the prime number theorem for arithmetic progressions, we have

$$\begin{aligned} & \psi(x;q,1)-\psi(\frac{x}{2};q,1) \\ = & \frac{x}{\phi(q)}-\frac{\frac{x}{2}}{\phi(q)}+O\left(xe^{-c(\log x)^{\frac{1}{2}}}\right) \\ = & \frac{\frac{x}{2}}{\phi(q)}+O\left(xe^{-c(\log x)^{\frac{1}{2}}}\right) \end{aligned}$$

where c is an absolute constant. Note that the main term dominates the error term, so we deduce that there exists at least one prime of the form $aq+1$ between $\frac{x}{2}$ and x.

To improve the lower bound for the size of the prime of the form $ab+1$, we need a better estimate for $\psi(x;q,a)$. One way of doing this is with the assumption of the GRH. The other substitute is by using the Bombieri-Vinogradov theorem.

3.2 Proof of Theorem 2 and its corollary

Theorem 2 *Let N and q be positive integers and let α be a real number with $\alpha > 4$ and $q = [\frac{N}{(\log N)^{\alpha}}]$. Put $x_1 = Nq+1$ and $x_2 = \frac{N}{2}q+1$. If GRH is true, we have*

$$\begin{aligned} & \pi(x_1;q,1)-\pi(x_2;q,1) \\ \geq & \frac{1}{4}\frac{N}{\log N}+O\left(\frac{N\log\log N}{(\log N)^2}+\frac{N}{(\log N)^{\frac{\alpha}{2}-1}}\right) \end{aligned}$$

Proof. Under GRH, if $q \leq x$, by theorem D, we have,

$$\pi(x;q,a)=\frac{Lix}{\phi(q)}+O(x^{\frac{1}{2}}\log x) \tag{3.1}$$

We shall now estimate the number of primes in the arithmetic progression $mq+1$ where $\frac{N}{2} < m \leq N$. By (3.1) we have

$$\pi(Nq+1;q,1)=\frac{Li(Nq+1)}{\phi(q)}+O((Nq)^{\frac{1}{2}}\log(Nq)) \tag{3.2}$$

and

$$\pi(\frac{N}{2}q+1;q,1) \;=\; \frac{Li(\frac{N}{2}q+1)}{\phi(q)}+O((Nq)^{\frac{1}{2}}\log(Nq)). \tag{3.3}$$

Hence from (3.2) and (3.3),

$$\begin{aligned}
&\pi(Nq+1;q,1)-\pi(\frac{N}{2}q+1;q,1)\\
&=\frac{Li(Nq+1)-Li(\frac{N}{2}q+1)}{\phi(q)}+O((Nq)^{\frac{1}{2}}\log(Nq))\\
&=\frac{\int_{\frac{N}{2}q+1}^{Nq+1}\frac{1}{\log u}du}{\phi(q)}+O((Nq)^{\frac{1}{2}}\log(Nq))\\
&=\frac{\frac{u}{\log u}|_{\frac{N}{2}q+1}^{Nq+1}+\int_{\frac{N}{2}q+1}^{Nq+1}\frac{1}{(\log u)^2}}{\phi(q)}+O((Nq)^{\frac{1}{2}}\log(Nq))\\
&=\frac{\frac{Nq+1}{\log(Nq+1)}-\frac{\frac{N}{2}q+1}{\log(\frac{N}{2}q+1)}+O(\frac{Nq}{(\log Nq)^2})}{\phi(q)}+O((Nq)^{\frac{1}{2}}\log(Nq))\\
&=\frac{\frac{Nq}{\log Nq}-\frac{\frac{N}{2}q}{\log Nq}+O(\frac{Nq}{(\log Nq)^2})}{\phi(q)}+O((Nq)^{\frac{1}{2}}\log(Nq))\\
&=\frac{\frac{1}{2}\frac{Nq}{\log Nq}+O(\frac{Nq}{(\log Nq)^2})}{\phi(q)}+O((Nq)^{\frac{1}{2}}\log(Nq))\\
&=\frac{1}{2}\frac{Nq}{\phi(q)\log Nq}+O\left(\frac{Nq}{\phi(q)(\log Nq)^2}\right)+O((Nq)^{\frac{1}{2}}\log(Nq))\\
&\geq\frac{1}{2}\frac{N}{\log Nq}+O\left(\frac{Nq}{\phi(q)(\log Nq)^2}\right)+O((Nq)^{\frac{1}{2}}\log(Nq)).
\end{aligned}\tag{3.4}$$

For any positive integer q, there exists an absolute constant C such that $\phi(q)\geq C\frac{q}{\log\log q}$. Note that since $q=[\frac{N}{(\log N)^{\alpha}}]$, we may deduce that

$$\begin{aligned}
&\pi(x_1;q,1)-\pi(x_2;q,1)\\
\geq\;&\frac{1}{4}\frac{N}{\log N}+O\left(\frac{N\log\log N}{(\log N)^2}+\frac{N}{(\log N)^{\frac{\alpha}{2}-1}}\right).
\end{aligned}$$

This completes the proof. □

This implies that there is a prime lying between $[\frac{N}{2}]q+1$ and $Nq+1$. i.e. when N is large, there is a prime number p such that

$$p \geq \frac{1}{2}Nq + 1.$$

We have the following corollary:

Corollary 1 *Let N be an integer and put $S = \{1, 2, \cdots, N\}$. Let α be a real number with $\alpha > 4$. Assume GRH. When N is large enough, there are integers $a, b \in S$ such that*

$$P(ab+1) > \frac{1}{2}\frac{N^2}{(\log N)^\alpha}.$$

By choosing a suitable α, we may remove the factor $\frac{1}{2}$.

3.3 Proof of Theorem 3 and its corollary

Theorem 3 *Let α and β be two positive numbers with $\alpha < \beta$. Put*

$$x_1 = \frac{N}{(\log N)^\beta}N + 1,$$

and

$$x_2 = \frac{N}{(\log N)^\beta}\frac{N}{2} + 1.$$

Then we have

$$\begin{aligned} &\sum_{\frac{N}{(\log N)^\beta} < q \leq \frac{N}{(\log N)^\alpha}} (\psi(x_1; q, 1) - \psi(x_2; q, 1)) \\ &\geq \frac{\beta - \alpha}{2}\frac{N^2}{\log^\beta N}(\log\log N) + O\left(\frac{N^2}{(\log N)^{\frac{\beta}{2}+\alpha-5}}\right). \end{aligned} \tag{3.5}$$

Proof. Let x_1 and x_2 be as in the statement of the theorem. We have

$$\begin{aligned} & \psi(x_1; q, 1) \quad -\psi(x_2; q, 1) \\ = \ & \psi(x_1; q, 1) \quad -\frac{x_1}{\phi(q)} + \frac{x_1}{\phi(q)} - \psi(x_2; q, 1) \\ = \ & \tfrac{x_1 - x_2}{\phi(q)} + \quad \psi(x_1; q, 1) - \frac{x_1}{\phi(q)} + \frac{x_2}{\phi(q)} - \psi(x_2; q, 1) \end{aligned}$$

Therefore for $\frac{N}{(\log N)^\beta} < q \leq \frac{N}{(\log N)^\alpha}$,

$$\sum_{\frac{N}{(\log N)^\beta}<q\leq\frac{N}{(\log N)^\alpha}} (\psi(x_1;q,1)-\psi(x_2;q,1))$$

$$\begin{aligned} &= \sum_{\frac{N}{(\log N)^\beta}<q\leq\frac{N}{(\log N)^\alpha}} \frac{x_1-x_2}{\phi(q)} \\ &+ \sum_{\frac{N}{(\log N)^\beta}<q\leq\frac{N}{(\log N)^\alpha}} (\psi(x_1;q,1)-\frac{x_1}{\phi(q)}) \\ &+ \sum_{\frac{N}{(\log N)^\beta}<q\leq\frac{N}{(\log N)^\alpha}} (\frac{x_2}{\phi(q)}-\psi(x_2;q,1)), \end{aligned} \tag{3.6}$$

where

$$\begin{aligned} \sum_{\frac{N}{(\log N)^\beta}<q\leq\frac{N}{(\log N)^\alpha}} \frac{x_1-x_2}{\phi(q)} &> (x_1-x_2) \sum_{\frac{N}{(\log N)^\beta}<q\leq\frac{N}{(\log N)^\alpha}} \frac{1}{q} \\ &= (\beta-\alpha)(x_1-x_2)(\log\log N)+O(\frac{(\log N)^\beta}{N}). \end{aligned} \tag{3.7}$$

By the Bombieri-Vinogradov theorem, Theorem E, we have

$$\begin{aligned} &| \sum_{\frac{N}{(\log N)^\beta}<q\leq\frac{N}{(\log N)^\alpha}} (\psi(x_1;q,1)-\frac{x_1}{\phi(q)})| \\ \leq\ & \sum_{\frac{N}{(\log N)^\beta}<q\leq\frac{N}{(\log N)^\alpha}} |\psi(x_1;q,1)-\frac{x_1}{\phi(q)}| \\ =\ & O\left(x_1^{\frac{1}{2}}\frac{N}{(\log N)^\alpha}(\log x_1)^5\right) \end{aligned} \tag{3.8}$$

and

$$\begin{aligned} & \Big| \sum_{\frac{N}{(\log N)^\beta} < q \le \frac{N}{(\log N)^\alpha}} (\psi(x_2; q, 1) - \frac{x_2}{\phi(q)}) \Big| \\ \le & \sum_{\frac{N}{(\log N)^\beta} < q \le \frac{N}{(\log N)^\alpha}} |\psi(x_2; q, 1) - \frac{x_2}{\phi(q)}| \\ = & O\left(x_1^{\frac{1}{2}} \frac{N}{(\log N)^\alpha} (\log x_1)^5 \right). \end{aligned} \tag{3.9}$$

Thus by (3.6), (3.7) (3.8)and (3.9),

$$\begin{aligned} & \sum_{\frac{N}{(\log N)^\beta} < q \le \frac{N}{(\log N)^\alpha}} (\psi(x_1; q, 1) - \psi(x_2; q, 1)) \\ > & (\beta - \alpha)(x_1 - x_2)(\log \log N) + O\left(x_1^{\frac{1}{2}} \frac{N}{(\log N)^\alpha} (\log x_1)^5 \right) \\ = & \frac{\beta - \alpha}{2} \frac{N^2}{\log^\beta N} (\log \log N) + O\left(\frac{N^2}{(\log N)^{\frac{\beta}{2} + \alpha - 5}} \right). \end{aligned} \tag{3.10}$$

This finishes the proof. □

Note that by (3.10) the main term is of the size

$$\frac{\beta - \alpha}{2} \frac{N^2}{(\log N)^\beta} (\log \log N)$$

and the error term is of the size

$$\frac{N^2}{(\log N)^{\frac{\beta}{2} + \alpha - 5}}.$$

So long as we choose α, β with $\alpha < \beta$ and

$$\alpha + \frac{\beta}{2} - 5 > \beta,$$

i.e.

$$\alpha > \frac{\beta}{2} + 5,$$

we have that (3.10) is dominated by the main term. If we then also choose β so that $\frac{\beta}{2}+5<\beta$,i.e., $\beta>10$, We obtain the following corollary.

Corollary 2 *Let N be a positive integer and put $S=\{1,2,\cdots,N\}$. Let β be a real number with $\beta>10$, when N is large enough, then there are integers $a,b\in S$, such that*

$$P(ab+1)>\frac{1}{2}\frac{N^2}{(\log N)^\beta}.$$

Equivalently if β is a real number larger than 10 then for N sufficiently large

$$P\left(\prod_1^N\prod_1^N(ab+1)\right)>\frac{1}{2}\frac{N^2}{(\log N)^\beta}.$$

Chapter 4

All the Primes and Higher Dimensional Kloosterman Sums

4.1 An Observation

To prove Lemma 3, Stewart introduced the following set:

$$U_t(N) = \{\ (m,n) \in Z \times Z \mid 1 \le m \le N,\ 1 \le n \le N, t \mid mn+1\ \}, \tag{4.1}$$

If t divides N then we may decompose $\{(m,n)|\ 1 \le m,\ n \le N\}$ into $(N/t)^2$ blocks consisting of the Cartesian product of two complete sets of residues modulo t. Thus if t divides N, $|U_t(N)| = \phi(t)(N/t)^2$. In general we deduce that

$$|U_t(N)| = \phi(t)\left(\frac{N}{t} + O(1)\right)^2 = \frac{\phi(t)}{t^2}N^2 + O(N).$$

By a simple calculation we find that the above estimate is dominated by the main term for t up to N, which is trivial. To get sharper estimates Stewart appealed to Weil's estimates for Kloosterman sums [33]. For any positive integer n let $d(n)$ denote the number of divisors of n. He proved,

Theorem F *(C. Stewart)*

$$|U_t(N)| = \frac{\phi(t)}{t^2}N^2 + O\left(t^{\frac{1}{2}}d(t)^{3/2}(\log t)^2 + \frac{Nd(t)\log t}{t}\right).$$

Proof. See [30]. □

This estimate is dominated by the main term for t up to $N^{(\frac{4}{3}-\epsilon)}$. In other words, for all the integers t up to $N^{(\frac{4}{3}-\epsilon)}$, we can find integers m and n with $1 \leq m, n \leq N$ such that $t \mid mn+1$. In particular, if t is a prime number p, we immediately get the following from Theorem F.

$$|U_p(N)| = \frac{p-1}{p^2}N^2 + O\left(p^{\frac{1}{2}}(\log p)^2 + \frac{N\log p}{p}\right).$$

This assures us that for all the prime numbers p up to $N^{(\frac{4}{3}-\epsilon)}$, we can find integers m and n with $1 \leq m, n \leq N$ such that $p \mid mn+1$. In other words, we have

$$P(mn+1) \geq N^{4/3-\epsilon}.$$

for some integers m and n with $1 \leq m, n \leq N$.

We shall prove the "higher-dimensional" analogue of this result in the following sections.

4.2 Proof of Theorem 4

Let N and k be two integers and let t be a positive integer. Recall that

$$U_t(N,k) = \{(a_1, \cdots, a_{k+1}) | 1 \leq a_1, \cdots, a_{k+1} \leq N, t \mid a_1 \cdots a_{k+1} + 1\},$$

where a_i, $i = 1, \cdots, k+1$ are integers.

Theorem 4 *Let N and k be positive integers and let p be a prime. We have*

$$|U_p(N,k)| = \frac{1}{p-1}\left(N - \left[\frac{N}{p}\right]\right)^{k+1} + O(kp^{\frac{k}{2}}\log^{k+1} p).$$

Further, if $p \mid N$ then

$$|U_p(N,k)| = \left(\frac{N}{p}\phi(p)\right)^k \frac{N}{p} = \frac{N^{k+1}}{p^{k+1}}\phi^k(p) = \frac{N^{k+1}}{p-1}(1-\frac{1}{p})^{k+1}.$$

Actually Theorem 4 can be deduced from our next result Theorem 5. But Theorem 4 can be proved by using Deligne's result directly whereas for the proof of Theorem 5 Smith's result is needed. Also, the result is much simpler in the case $p > N$ and the argument is easier. Notice in particular that if $p > N$ then

$$|U_p(N,k)| = \frac{N^{k+1}}{p-1} + O(kp^{\frac{k}{2}} \log^{k+1} p).$$

Proof. The simplest case is when $p \mid N$, for each k-tuple $(a_1, a_2, \cdots, a_k)$ with $1 \leq a_1, \cdots, a_{k+1} \leq N$ and $p \nmid a_1 a_2 \cdots a_k$. Then there exists exactly one choice of a_{k+1} modulo p. Hence there exist $\frac{N}{p}$ choices of a_{k+1} such that

$$a_1 a_2 \cdots a_{k+1} \equiv -1 \ (mod\ p).$$

Thus for $p \mid N$,

$$|U_p(N,k)| = \left(\frac{N}{p}\phi(p)\right)^k \frac{N}{p} = \frac{N^{k+1}}{p^{k+1}}\phi^k(p).$$

In the following, we shall consider the remaining cases, $p > N$ and $p < N$ but $p \nmid N$.
Recall that

$$\sum_{g=1}^{m} e(\frac{gn}{m}) = \begin{cases} m & m|n \\ 0 & m \nmid n. \end{cases} \tag{4.2}$$

We have

$$|U_p(N,k)| = \frac{1}{p^{k+1}} \sum_{l_1,\cdots,l_{k+1}=1}^{p} \sum_{\substack{a_1,\cdots,a_{k+1}=1 \\ a_1\cdots a_{k+1} \equiv -1 (mod\ p)}}^{p}$$
$$\sum_{b_1,\cdots,b_{k+1}=1}^{N} e\left(\frac{l_1(a_1-b_1)+\cdots+l_{k+1}(a_{k+1}-b_{k+1})}{p}\right).$$

This is because for each (k+1)-tuple $(a_1, a_2, \ldots, a_{k+1})$ such that

$$a_1 a_2 \cdots a_{k+1} \equiv -1 \ \ (mod\ p),$$

if there is a (k+1)-tuple $(b_1, b_2, \cdots, b_{k+1})$ such that

$$a_i \equiv b_i, \quad i = 1, 2, \cdots, k+1, \tag{4.3}$$

then the sums involving this (k+1)-tuple $(b_1, b_2, \cdots, b_{k+1})$ contribute p^{k+1} to $|U_p(N,k)|$ and the (k+1)-tuples not satisfying (4.3) contribute 0 to $|U_p(N,k)|$.

The terms with $l_1 = l_2 = \cdots = l_{k+1} = p$ contribute

$$\frac{N^{k+1}}{p^{k+1}}(p-1)^k$$

to $|U_p(N,k)|$. Thus we have

$$|U_p(N,k)| = \frac{N^{k+1}}{p^{k+1}}(p-1)^k + R, \tag{4.4}$$

where R denotes the sum over those l_i's such that not all l_i's are equal to p.

Let R_i be the subsums of R such that i of the l_j's equal p. Then we have

$$R = R_0 + \cdots + R_k$$

We now will give an estimate for each R_i, $i = 0, 1, \cdots, k$. If no l_i equals p, we have

$$\begin{aligned}
R_0 &= \frac{1}{p^{k+1}} \sum_{l_1,\cdots,l_{k+1}=1}^{p-1} \sum_{\substack{a_1,\cdots,a_{k+1}=1 \\ a_1\cdots a_{k+1} \equiv -1 \pmod p}}^{p} \\
&\quad \sum_{b_1,\ \cdots,\ b_{k+1}=1}^{N} e\left(\frac{l_1(a_1-b_1)+\cdots+l_{k+1}(a_{k+1}-b_{k+1})}{p}\right) \\
&= \frac{1}{p^{k+1}} \sum_{l_1,\cdots,l_{k+1}=1}^{p-1} \sum_{\substack{a_1,\cdots,a_{k+1}=1 \\ a_1\cdots a_{k+1} \equiv -1 \pmod p}}^{p} e\left(\frac{l_1a_1+\cdots+l_{k+1}a_{k+1}}{p}\right) \\
&\quad \sum_{b_1,\cdots,b_{k+1}=1}^{N} e\left(\frac{-l_1b_1-\cdots-l_{k+1}b_{k+1}}{p}\right) \\
&= \frac{1}{p^{k+1}} \sum_{l_1,\cdots,l_{k+1}=1}^{p-1} \sum_{\substack{a_1,\cdots,a_{k+1}=1 \\ a_1\cdots a_{k+1} \equiv 1 \pmod p}}^{p} e\left(\frac{l_1a_1+\cdots-l_{k+1}a_{k+1}}{p}\right) \\
&\quad \sum_{b_1,\cdots,b_{k+1}=1}^{N} e\left(\frac{-l_1b_1-\cdots-l_{k+1}b_{k+1}}{p}\right)
\end{aligned}$$

$$= \frac{1}{p^{k+1}} \sum_{l_1,\cdots,l_{k+1}=1}^{p-1} K(l_1,\cdots,-l_{k+1};p) \sum_{b_1,\cdots,b_{k+1}=1}^{N} e\left(-\frac{l_1b_1+\cdots+l_{k+1}b_{k+1}}{p}\right).$$

By Deligne's theorem (1.25), we have

$$|R_0| = O\left(\frac{1}{p^{k+1}} kp^{\frac{k}{2}} \sum_{l_1,\cdots,l_{k+1}=1}^{p-1} \left| \sum_{b_1,\cdots,b_{k+1}=1}^{N} e\left(-\frac{l_1b_1+\cdots+l_{k+1}b_{k+1}}{p}\right)\right|\right)$$

$$= O\left(\frac{1}{p^{k+1}} kp^{\frac{k}{2}} \sum_{\substack{l_1,\cdots,l_{k+1}=-\frac{p-1}{2} \\ l_1\cdots l_{k+1}\neq 0}}^{\frac{p-1}{2}} \left| \sum_{b_1,\cdots,b_{k+1}=1}^{N} e\left(-\frac{l_1b_1+\cdots+l_{k+1}b_{k+1}}{p}\right)\right|\right)$$

$$= O\left(\frac{1}{p^{k+1}} kp^{\frac{k}{2}} \sum_{\substack{l_1,\cdots,l_{k+1}=-\frac{p-1}{2} \\ l_1\cdots l_{k+1}\neq 0}}^{\frac{p-1}{2}} \frac{p}{|l_1|}\cdots\frac{p}{|l_{k+1}|}\right)$$

$$= O\left(kp^{\frac{k}{2}}\log^{k+1} p\right), \tag{4.5}$$

where we have used the well known inequality

$$\sum_{1\le s\le M} e(\frac{gs}{t}) = O(\frac{t}{|g|}),\ for\ 1\le |g|\le \frac{t}{2}.$$

Let $R_{1,i}$ denote the subsum in R_1 such that only $l_i = p$, then

$$R_1 = R_{1,1} + R_{1,2} + \cdot + R_{1,k+1}. \tag{4.6}$$

Now assume that only $l_1 = p$. We have

$$\begin{aligned} R_{1,1} &= \frac{1}{p^{k+1}} \sum_{\substack{l_2,\cdots,l_{k+1}=1 \\ l_1=p}}^{p-1} \sum_{\substack{a_1,\cdots,a_{k+1}=1 \\ a_1\cdots a_{k+1}\equiv -1(mod\ p)}}^{p} \\ &\quad \sum_{b_1,\cdots,b_{k+1}=1}^{N} e\left(\frac{l_2(a_2-b_2)+\cdots+l_{k+1}(a_{k+1}-b_{k+1})}{p}\right) \\ &= \frac{1}{p^{k+1}} \sum_{l_2,\cdots,l_{k+1}=1}^{p-1} \sum_{\substack{a_1,\cdots,a_{k+1}=1 \\ a_1\cdots a_{k+1}\equiv -1(mod\ p)}}^{p} e\left(\frac{l_2a_2+\cdots+l_{k+1}a_{k+1}}{p}\right) \\ &\quad \sum_{b_1,\cdots,b_{k+1}=1}^{N} e\left(\frac{-l_2b_2-\cdots-l_{k+1}b_{k+1}}{p}\right). \end{aligned}$$

Note that for fixed $a_2, \cdots, a_{k+1}$, the solution for

$$xa_2\cdots a_{k+1} \equiv -1(mod\ p),$$

with $1 \le x < p$, is unique if the a_i, $i = 2, \cdots, k+1$ are coprime to p. Hence assuming that $p > N$, we have

$$\begin{aligned} R_{1,1} &= \frac{1}{p^{k+1}} \sum_{l_2,\cdots,l_{k+1}=1}^{p-1} \sum_{\substack{a_1,\cdots,a_{k+1}=1 \\ a_1\cdots a_{k+1}\equiv -1(mod\ p)}}^{p} e\left(\frac{l_2a_2+\cdots+l_{k+1}a_{k+1}}{p}\right) \\ &\quad \sum_{b_1,\cdots,b_{k+1}=1}^{N} e\left(\frac{-l_2b_2-\cdots-l_{k+1}b_{k+1}}{p}\right) \\ &= \frac{1}{p^{k+1}} \sum_{l_2,\cdots,l_{k+1}=1}^{p-1} \sum_{a_2=1}^{p-1} e\left(\frac{l_2a_2}{p}\right) \sum_{a_3=1}^{p-1} e\left(\frac{l_3a_3}{p}\right)\cdots \\ &\quad \sum_{a_{k+1}=1}^{p-1} e\left(-\frac{l_{k+1}a_{k+1}}{p}\right) \sum_{b_1,\cdots,b_{k+1}=1}^{N} e\left(\frac{-l_2b_2-\cdots-l_{k+1}b_{k+1}}{p}\right) \end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{p^{k+1}} \sum_{l_2,\cdots,l_{k+1}=1}^{p-1} (-1)^k \sum_{b_1,\cdots,b_{k+1}=1}^{N} e\left(\frac{-l_2b_2-\cdots-l_{k+1}b_{k+1}}{p}\right)\\
&= \frac{1}{p^{k+1}}(-1)^k \sum_{b_1,\cdots,b_{k+1}=1}^{N} \sum_{l_2,\cdots,l_{k+1}=1}^{p-1} e\left(\frac{-l_2b_2-\cdots-l_{k+1}b_{k+1}}{p}\right)\\
&= \frac{1}{p^{k+1}}(-1)^k \sum_{b_1,\cdots,b_{k+1}=1}^{N} \sum_{l_2=1}^{p-1} e\left(\frac{-l_2b_2}{p}\right)\cdots\sum_{l_{k+1}=1}^{p-1} e\left(\frac{-l_{k+1}b_{k+1}}{p}\right)\\
&= \frac{1}{p^{k+1}}(-1)^k \sum_{b_1,\cdots,b_{k+1}=1}^{N} (-1)^k\\
&= \frac{N^{k+1}}{p^{k+1}}.
\end{aligned}
$$

Similarly we have

$$R_{1,2} = R_{1,3} = \cdots = R_{1,k+1} = \frac{N^{k+1}}{p^{k+1}}.$$

Thus by (4.6) we get

$$R_1 = \binom{k+1}{1}\frac{N^{k+1}}{p^{k+1}}. \tag{4.7}$$

If two of the l_i's are equal to p then, since the sum is symmetric, we may assume that $l_1 = l_2 = p$. Thus,

$$
\begin{aligned}
R_2 &= \frac{\binom{k+1}{2}}{p^{k+1}} \sum_{l_3,\cdots,l_{k+1}=1}^{p-1} \sum_{\substack{a_1,\cdots,a_{k+1}=1\\ a_1\cdots a_{k+1}\equiv -1 (mod\ p)}}^{p}\\
&\qquad \sum_{b_1,\cdots,b_{k+1}=1}^{N} e\left(\frac{l_3(a_3-b_3)+\cdots+l_{k+1}(a_{k+1}-b_{k+1})}{p}\right)\\
&= \frac{\binom{k+1}{2}}{p^{k+1}} \sum_{l_3,\cdots,l_{k+1}=1}^{p-1} \sum_{\substack{a_1,\cdots,a_{k+1}=1\\ a_1\cdots a_{k+1}\equiv -1 (mod\ p)}}^{p} e\left(\frac{l_3a_3+\cdots+l_{k+1}a_{k+1}}{p}\right)\\
&\qquad \sum_{b_1,\cdots,b_{k+1}=1}^{N} e\left(-\frac{l_3b_3+\cdots+l_{k+1}b_{k+1}}{p}\right)
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{\binom{k+1}{2}}{p^{k+1}}(p-1)\sum_{a_3=1}^{p-1} e\left(\frac{l_3 a_3}{p}\right)\cdots\sum_{a_{k+1}=1}^{p-1} e\left(-\frac{l_{k+1}a_{k+1}}{p}\right)\\
&\quad \sum_{b_1,\cdots,b_{k+1}=1}^{N}\ \sum_{l_3,\cdots,l_{k+1}=1}^{p-1} e\left(-\frac{l_3 b_3+\cdots+l_{k+1}b_{k+1}}{p}\right)\\
&= \frac{N^{k+1}}{p^{k+1}}\binom{k+1}{2}(p-1). \qquad (4.8)
\end{aligned}
$$

In general, if j of these l_i's are equal to p, we have

$$
R_j = \frac{N^{k+1}}{p^{k+1}}\binom{k+1}{j}(p-1)^{j-1}. \qquad (4.9)
$$

Combining all the results for the R_i, by (4.5), (4.7), (4.8) and (4.9) we have

$$
R = \frac{N^{k+1}}{p^{k+1}}\sum_{j=1}^{k}\binom{k+1}{j}(p-1)^{j-1} + O(kp^{\frac{k-1}{2}}\log^k p). \qquad (4.10)
$$

Thus, by (4.4) and (4.10), if $p > N$,

$$
\begin{aligned}
|U_p(N,k)| &= \frac{N^{k+1}}{p^{k+1}}(p-1)^k + \sum_{j=1}^{k}\frac{N^{k+1}}{p^{k+1}}\binom{k+1}{j}(p-1)^{j-1} + O(kp^{\frac{k}{2}}\log^{k+1} p)\\
&= \sum_{j=1}^{k+1}\frac{N^{k+1}}{p^{k+1}}\binom{k+1}{j}(p-1)^{j-1} + O(kp^{\frac{k}{2}}\log^{k+1} p)\\
&= \frac{N^{k+1}}{p^{k+1}}\sum_{j=1}^{k+1}\binom{k+1}{j}(p-1)^{j-1} + O(kp^{\frac{k}{2}}\log^{k+1} p)\\
&= \frac{N^{k+1}}{p^{k+1}}\left(\sum_{j=0}^{k+1}\binom{k+1}{j}(p-1)^{j-1} - \frac{1}{p-1}\right) + O(kp^{\frac{k}{2}}\log^{k+1} p)\\
&= \frac{N^{k+1}}{p^{k+1}}\left(\frac{(1+p-1)^{k+1}}{p-1} - \frac{1}{p-1}\right) + O(kp^{\frac{k}{2}}\log^{k+1} p)\\
&= \frac{N^{k+1}}{p^{k+1}}\left(\frac{p^{k+1}-1}{p-1}\right) + O(kp^{\frac{k}{2}}\log^{k+1} p).
\end{aligned}
$$

If $p < N$ and $p \nmid N$ then

$$\begin{aligned}
R_1 &= \binom{k+1}{1}\frac{1}{p^{k+1}} \sum_{\substack{l_2,\cdots,l_{k+1}=1 \\ l_1=p}}^{p-1} \sum_{\substack{a_1,\cdots,a_{k+1}=1 \\ a_1\cdots a_{k+1}\equiv -1 (mod\ p)}}^{p} \\
&\quad \sum_{b_1,\cdots,b_{k+1}=1}^{N} e\left(\frac{l_2(a_2-b_2)+\cdots+l_{k+1}(a_{k+1}-b_{k+1})}{p}\right) \\
&= \binom{k+1}{1}\frac{1}{p^{k+1}} \sum_{l_2,\cdots,l_{k+1}=1}^{p-1} \sum_{\substack{a_1,\cdots,a_{k+1}=1 \\ a_1\cdots a_{k+1}\equiv -1 (mod\ p)}}^{p} e\left(\frac{l_2a_2+\cdots+l_{k+1}a_{k+1}}{p}\right) \\
&\quad \sum_{b_1,\cdots,b_{k+1}=1}^{N} e\left(\frac{-l_2b_2-\cdots-l_{k+1}b_{k+1}}{p}\right) \\
&= \binom{k+1}{1}\frac{1}{p^{k+1}} \sum_{l_2,\cdots,l_{k+1=1}}^{p-1} \sum_{a_2=1}^{p-1} e\left(\frac{l_2a_2}{p}\right) \sum_{a_3=1}^{p-1} e\left(\frac{l_3a_3}{p}\right)\cdots \\
&\quad \sum_{a_{k+1}=1}^{p-1} e\left(-\frac{l_{k+1}a_{k+1}}{p}\right) \sum_{b_1,\cdots,b_{k+1}=1}^{N} e\left(\frac{-l_2b_2-\cdots-l_{k+1}b_{k+1}}{p}\right) \\
&= \binom{k+1}{1}\frac{1}{p^{k+1}} \sum_{l_2,\cdots,l_{k+1}=1}^{p-1} (-1)^k \sum_{b_1,\cdots,b_{k+1}=1}^{N} e\left(\frac{-l_2b_2-\cdots-l_{k+1}b_{k+1}}{p}\right)
\end{aligned}$$

$$\begin{aligned}
&= \binom{k+1}{1}\frac{(-1)^k}{p^{k+1}} \sum_{b_1,\cdots,b_{k+1}=1}^{N} \sum_{l_2,\cdots,l_{k+1}=1}^{p-1} e\left(\frac{-l_2b_2-\cdots-l_{k+1}b_{k+1}}{p}\right) \\
&= \binom{k+1}{1}\frac{N(-1)^k}{p^{k+1}} \left(\sum_{b=1}^{N}\sum_{l=1}^{p-1} e\left(-\frac{lb}{p}\right)\right)^k \\
&= \binom{k+1}{1}\frac{N(-1)^k}{p^{k+1}} \left(\sum_{\substack{b=1 \\ p|b}}^{N}(p-1) + \sum_{\substack{b=1 \\ p\nmid b}}^{N}(-1)\right)^k \\
&= \binom{k+1}{1}\frac{N(-1)^k}{p^{k+1}} \left((p-1)\left[\frac{N}{p}\right] - (N-\left[\frac{N}{p}\right])\right)^k \\
&= \frac{N}{p^{k+1}}\binom{k+1}{1}\left(N-p\left[\frac{N}{p}\right]\right)^k. \qquad (4.11)
\end{aligned}$$

Similarly, we have

$$R_2 = \frac{N^2}{p^{k+1}}\phi(p)\binom{k+1}{2}\left(N - p\left[\frac{N}{p}\right]\right)^{k-1} \tag{4.12}$$

$$\vdots$$

$$R_k = \frac{N^k}{p^{k+1}}\phi^{k-1}(p)\binom{k+1}{k}\left(N - p\left[\frac{N}{p}\right]\right). \tag{4.13}$$

Thus, by(4,4), (4.5), (4.11),(4.12) and (4.13) we have that if $p < N$ and $p \nmid N$, then

$$\begin{aligned}
& |U_p(N,k)| \\
= & \frac{\phi^{-1}(p)}{p^{k+1}}\sum_{j=1}^{k+1}\binom{k+1}{j}N^j\phi^j(p)\left(N - p\left[\frac{N}{p}\right]\right)^{k+1-j} + O(kp^{\frac{k}{2}}\log^{k+1}p) \\
= & \frac{\phi^{-1}(p)}{p^{k+1}}\left(\left(N\phi(p) + N - p\left[\frac{N}{p}\right]\right)^{k+1} - \left(N - p\left[\frac{N}{p}\right]\right)^{k+1}\right) \\
& + O(kp^{\frac{k}{2}}\log^{k+1}p) \\
= & \frac{1}{p-1}\left(N - \left[\frac{N}{p}\right]\right)^{k+1} + O(kp^{\frac{k}{2}}\log^{k+1}p).
\end{aligned}$$

This finishes the proof of theorem 4. □

4.3 Proof of Theorem 5

Let N and k be two integers and let t be a positive integer. Again recall

$$U_t(N,k) = \{(a_1,\cdots,a_{k+1}) | 1 \le a_1,\cdots,a_{k+1} \le N, t \mid a_1\cdots a_{k+1}+1\},$$

where a_i, $i = 1,\cdots,k+1$ are integers. We have

Theorem 5 *Let N, k and m be positive integers and let p be a prime, then we have*

$$|U_{p^m}(N,k)| = \frac{1}{\phi(p^m)}\left(N - \left[\frac{N}{p}\right]\right)^{k+1} + O\left(d_{k+1}(p^m)(p^m)^{\frac{k}{2}}\log^{k+1}p^m\right).$$

We remark that when $n = 1$ we recover the first part of Theorem 4 since $d_{k+1}(p) = k+1$.

Proof. By (4.2) we have

$$\begin{aligned} & |U_{p^m}(N,k)| \\ = & \frac{1}{(p^m)^{k+1}} \sum_{l_1,\cdots,l_{k+1}=1}^{p^m} \sum_{\substack{a_1,\cdots,a_{k+1}=1 \\ a_1\cdots a_{k+1}\equiv -1 \ (mod\ p^m)}}^{p^m} \\ & \sum_{b_1,\cdots,b_{k+1}=1}^{N} e\left(\frac{l_1(a_1-b_1)+\cdots+l_{k+1}(a_{k+1}-b_{k+1})}{p^m}\right). \end{aligned} \tag{4.14}$$

This is because for each (k+1)-tuple $(a_1, a_2, \ldots, a_{k+1})$ such that

$$a_1 a_2 \cdots a_{k+1} \equiv -1 \ (mod\ p^m),$$

if there is a (k+1)-tuple $(b_1, b_2, \cdots, b_{k+1})$ such that

$$a_i \equiv b_i \ (mod\ p^m), \ i = 1, 2, \cdots, k+1, \tag{4.15}$$

then the sums involving this (k+1)-tuple $(b_1, b_2, \cdots, b_{k+1})$ contribute $(p^m)^{k+1}$ to $|U_{p^m}(N,k)|$, and the (k+1)-tuples that do not satisfy (4.15) contribute 0.

The terms with $l_1 = l_2 = \cdots = l_{k+1} = p^m$ contribute

$$\frac{1}{(p^m)^{k+1}} \phi^k(p^m) N^{k+1} \tag{4.16}$$

to the right hand side of the (4.14). Thus we have

$$|U_{p^m}(N,k)| = \frac{1}{(p^m)^{k+1}} \phi^k(p^m) N^{k+1} + R, \tag{4.17}$$

where

$$R = R_0 + R_1 + \cdots + R_k \tag{4.18}$$

and R_i denotes the sum over those terms in $|U_{p^m}(N,k)|$ such that i of l_j's are equal to p^m, the others sum over 1 to $p^m - 1$.

It is easier to compute $R_1, R_2, \cdots$, and R_k than to compute R_0 since in the former cases we do not have to deal with the conditions

$$a_1 \cdots a_{k+1} \equiv -1 \ (mod\ p^m).$$

We first compute R_1. Note that it is symmetric over l_i, $i = 1, \cdots, k+1$. We get

$$\begin{aligned} R_1 &= \frac{\binom{k+1}{1}}{(p^m)^{k+1}} \sum_{l_2,\cdots,l_{k+1}=1}^{p^m-1} \sum_{\substack{a_1,\cdots,a_{k+1}=1 \\ a_1\cdots a_{k+1}\equiv -1 \ (mod\ p^m)}}^{p^m} \\ &\quad \sum_{b_1,\cdots,b_{k+1}=1}^{N} e\left(\frac{l_2(a_2-b_2)+\cdots+l_{k+1}(a_{k+1}-b_{k+1})}{p^m}\right) \\ &= \frac{\binom{k+1}{1}}{(p^m)^{k+1}} \sum_{l_2,\cdots,l_{k+1}=1}^{p^m-1} \sum_{\substack{a_1,\cdots,a_{k+1}=1 \\ a_1\cdots a_{k+1}\equiv -1 \ (mod\ p^m)}}^{p^m} e\left(\frac{l_2a_2+\cdots+l_{k+1}a_{k+1}}{p^m}\right) \\ &\quad \sum_{b_1,\cdots,b_{k+1}=1}^{N} e\left(-\frac{l_2b_2+\cdots+l_{k+1}b_{k+1}}{p^m}\right) \end{aligned}$$

Note that $(a_1\cdots a_{k+1}, p^m) = 1$, once $a_2, \cdots, a_{k+1}$ are given, we then have a unique choice for $1 \le a_1 \le p^m$ and since R_1 is symmetric over l_i and also symmetric over b_j we may compute R_1 in the following way. We have

$$\begin{aligned} R_1 &= \frac{\binom{k+1}{1}}{(p^m)^{k+1}} \sum_{l_2,\cdots,l_{k+1}=1}^{p^m-1} \sum_{\substack{a_2=1 \\ (a_2,p^m)=1}}^{p^m} e(\frac{l_2a_2}{p^m}) \cdots \sum_{\substack{a_{k+1}=1 \\ (a_{k+1},p^m)=1}}^{p^m} e(\frac{l_{k+1}a_{k+1}}{p^m}) \\ &\quad \sum_{b_1,\cdots,b_{k+1}=1}^{N} e(-\frac{l_2b_2+\cdots+l_{k+1}b_{k+1}}{p}) \\ &= \frac{\binom{k+1}{1}}{(p^m)^{k+1}} \sum_{l_2=1}^{p^m-1} (l_2,p^m)\mu(\frac{p^m}{(l_2,p^m)}) \cdots \sum_{l_{k+1}=1}^{p^m-1} (l_{k+1},p^m)\mu(\frac{p^m}{(l_{k+1},p^m)}) \\ &\quad \sum_{b_1,\cdots,b_{k+1}=1}^{N} e(-\frac{l_2b_2+\cdots+l_{k+1}b_{k+1}}{p^m}), \end{aligned}$$

where

$$\begin{aligned}\sum_{\substack{a=1\\(a,p^m)=1}}^{p^m} e(\frac{la}{p^m}) &= \sum_{\substack{a=1\\(a,p^m)=1}}^{p^m} e\left(\frac{\frac{l}{(l,p^m)}a}{\frac{p^m}{(l,p^m)}}\right)\\ &= (l,p^m)\sum_{\substack{a=1\\(a,p^m)=1}}^{\frac{p^m}{(l,p^m)}} e\left(\frac{\frac{l}{(l,p^m)}a}{\frac{p^m}{(l,p^m)}}\right)\\ &= (l,p^m)\mu(\frac{p^m}{(l,p^m)}).\end{aligned}$$

Here we have used the well known fact that the sum of the n-th primitive roots of unity equals $\mu(n)$. (See [14], Theorem 272.)

If $1 \le l \le p^m - 1$, then

$$\mu\left(\frac{p^m}{(l,p^m)}\right) = \begin{cases} -1, & p^{m-1}|l \\ 0, & p^{m-1} \nmid l. \end{cases}$$

Thus,

$$\begin{aligned}R_1 &= \frac{\binom{k+1}{1}}{(p^m)^{k+1}} \sum_{\substack{l_2=1\\p^{m-1}|l_2}}^{p^m-1} p^{m-1}(-1)\cdots \sum_{\substack{l_{k+1}=1\\p^{m-1}|l_{k+1}}}^{p^m-1} p^{m-1}(-1)\\ &\quad \sum_{b_1,\cdots,b_{k+1}=1}^{N} e(-\frac{l_2b_2+\cdots+l_{k+1}b_{k+1}}{p})\\ &= \frac{\binom{k+1}{1}}{(p^m)^{k+1}}(-1)^k(p^{m-1})^k \sum_{l_2=1}^{p-1}\cdots\sum_{l_{k+1}=1}^{p-1}\\ &\quad \sum_{b_1,\cdots,b_{k+1}=1}^{N} e(-\frac{l_2b_2+\cdots+l_{k+1}b_{k+1}}{p})\\ &= \frac{N\binom{k+1}{1}}{(p^m)^{k+1}}(-1)^k(p^{m-1})^k\left(\sum_{l_2=1}^{p-1}\sum_{b_2=1}^{N} e\left(-\frac{l_2b_2}{p}\right)\right)^k.\end{aligned}$$

Note that

$$\begin{aligned}\sum_{l_2=1}^{p-1}\sum_{b=1}^{N} e\left(-\frac{l_2 b}{p}\right) &= \sum_{\substack{b=1\\ p|b}}^{N}(p-1)+\sum_{\substack{b=1\\ p\nmid b}}^{N}(-1)\\ &= (p-1)\left[\frac{N}{p}\right]-(N-\left[\frac{N}{p}\right])\\ &= p[\frac{N}{p}]-N. \end{aligned} \tag{4.19}$$

By (4.19) and the above result for R_1 we further have,

$$\begin{aligned} R_1 &= \frac{N\binom{k+1}{1}}{(p^m)^{k+1}}(-1)^k(p^{m-1})^k\left(p\left[\frac{N}{p}\right]-N\right)^k\\ &= \frac{N\binom{k+1}{1}}{(p^m)^{k+1}}(p^{m-1})^k\left(N-p\left[\frac{N}{p}\right]\right)^k. \end{aligned} \tag{4.20}$$

To compute R_i, $i=2,\cdots,k+1$. We first choose $l_1=\cdots=l_i=p^m$. Denote the sum by R_{i1}. We have

$$\begin{aligned} R_{i1} &= \frac{N^i\phi^{i-1}(p^m)}{(p^m)^{k+1}}\sum_{l_{i+1},\cdots,l_{k+1}=1}^{p^m-1}\ \sum_{\substack{a_{i+1},\cdots,a_{k+1}=1\\ (a_{i+1}\cdots a_{k+1},\ p)=1}}^{p^m} e\left(\frac{l_{i+1}a_{i+1}+\cdots+l_{k+1}a_{k+1}}{p^m}\right)\cdots\\ &\quad\cdots\sum_{b_{i+1},\cdots,b_{k+1}=1}^{N} e\left(-\frac{l_{i+1}b_{i+1}+\cdots+l_{k+1}b_{k+1}}{p^m}\right)\\ &= \frac{\phi^{i-1}(p^m)N^i}{(p^m)^{k+1}}\sum_{b_{i+1},\cdots,b_{k+1}=1}^{N}\ \sum_{l_{i+1}=1}^{p^m-1}(l_{i+1},p^m)\mu\left(\frac{p^m}{(l_{i+1},p^m)}\right)e\left(-\frac{l_{i+1}b_{i+1}}{p^m}\right)\cdots\\ &\quad\cdots\sum_{l_{k+1}=1}^{p^m-1}(l_{k+1},p^m)\mu\left(\frac{p^m}{(l_{k+1},p^m)}\right)e\left(-\frac{l_{k+1}b_{k+1}}{p^m}\right)\\ &= \frac{N^i}{(p^m)^{k+1}}\phi^{i-1}(p^m)\left(\sum_{l=1}^{p-1}(-p^{m-1})\sum_{b=1}^{N}e(-\frac{lb}{p})\right)^{k+1-i}. \end{aligned}$$

By (4.19), we have

$$R_{i1} = \frac{N^i}{(p^m)^{k+1}} \phi^{i-1}(p^m)(p^{m-1})^{k+1-i} \left(N - p \left[\frac{N}{p} \right] \right)^{k+1-i}. \tag{4.21}$$

Combining all the cases of different choices of i of the l_j's that are equal to p^m, by (4.21) we have

$$R_i = \binom{k+1}{i} \frac{N^i}{(p^m)^{k+1}} \phi^{i-1}(p^m)(p^{m-1})^{k+1-i} \left(N - p \left[\frac{N}{p} \right] \right)^{k+1-i}. \tag{4.22}$$

It remains to compute R_0, where no l_i equals p^m. We shall apply Smith's theorem on estimates of the higher-dimensional Kloosterman sums and Lemma 1. We have

$$\begin{aligned} R_0 &= \frac{1}{(p^m)^{k+1}} \sum_{l_1,\cdots,l_{k+1}=1}^{p^m-1} \sum_{\substack{a_1,\cdots,a_{k+1}=1 \\ a_1\cdots a_{k+1} \equiv -1 \ (mod\ p^m)}}^{p^m} e\left(\frac{l_1a_1 + \cdots + l_{k+1}a_{k+1}}{p^m} \right) \\ &\quad \sum_{b_1,\cdots,b_{k+1}=1}^{N} e\left(-\frac{l_1a_1 + \cdots + l_{k+1}b_{k+1}}{p^m} \right) \\ &= \frac{1}{(p^m)^{k+1}} \sum_{l_1,\cdots,l_{k+1}=1}^{p^m-1} K(l_1, l_2, \cdots, -l_{k+1}; p^m) \\ &\quad \sum_{b_1,\cdots,b_{k+1}=1}^{p^m} e\left(-\frac{l_1b_1 + \cdots + l_{k+1}b_{k+1}}{p^m} \right) \end{aligned}$$

By Theorem G and Lemma 1 we get

$$\begin{aligned} & |K(l_1, \cdots, l_k, -l_{k+1}; p^m)| \\ \leq\ & (p^m)^{\frac{k}{2}} (l_1, \cdots, l_k, -l_{k+1}; p^m)_k^{\frac{k}{2}}\ d_{k+1}(p^m) \\ \leq\ & (p^m)^{\frac{k}{2}} (l_1, p^m)^{\frac{1}{2}} \cdots (l_{k+1}, p^m)^{\frac{1}{2}}\, d_{k+1}(p^m). \end{aligned} \tag{4.23}$$

Thus, by R_0 from the above and (4.23), we have

$$\begin{aligned}
|R_0| &\leq \frac{(p^m)^{\frac{k}{2}}}{(p^m)^{k+1}} \sum_{l_1,\cdots,l_{k+1}=1}^{p^m-1} (l_1,\ \cdots, l_k,\ -l_{k+1};\ p^m)_k^{\frac{k}{2}}\, d_{k+1}(p^m) \\
&\qquad \left| \sum_{b_1,\cdots,b_{k+1}=1}^{N} e\left(-\frac{l_1b_1+\cdots+l_{k+1}b_{k+1}}{p^m}\right)\right| \\
&\leq d_{k+1}(p^m)\ \frac{(p^m)^{\frac{k}{2}}}{(p^m)^{k+1}} \sum_{l_1=1}^{p^m-1} (l_1,\ p^m)^{\frac{1}{2}} \left|\sum_{b_1=1}^{N} e\left(-\frac{l_1b_1}{p^m}\right)\right| \cdots \\
&\qquad \cdots \sum_{l_{k+1}=1}^{p^m-1} (l_{k+1},\ p^m)^{\frac{1}{2}} \left|\sum_{b_{k+1}=1}^{N} e\left(-\frac{l_{k+1}b_{k+1}}{p^m}\right)\right| \\
&= d_{k+1}(p^m)\ \frac{(p^m)^{\frac{k}{2}}}{(p^m)^{k+1}} \left(\sum_{l=1}^{p^m-1} (l,\ p^m)^{\frac{1}{2}} \left|\sum_{b=1}^{N} e\left(-\frac{lb}{p^m}\right)\right|\right)^{k+1} \qquad (4.24)
\end{aligned}$$

Note that if α does not equal an integer, then

$$\left|\sum_{b=1}^{N} e(\alpha b)\right| \leq \frac{1}{\|\alpha\|}.$$

Thus,

$$\begin{aligned}
&\sum_{l=1}^{p^m-1} (l,p^m)^{\frac{1}{2}} \left|\sum_{b=1}^{N} e\left(-\frac{lb}{p^m}\right)\right| \\
\leq\ &\sum_{l=1}^{p^m-1} (l,p^m)^{\frac{1}{2}} \frac{1}{\left\|-\frac{l}{p^m}\right\|}
\end{aligned}$$

$$
\begin{aligned}
&\leq 2\sum_{l=1}^{\frac{p^m-1}{2}} \frac{p^m}{l}(l,p^m)^{\frac{1}{2}} \\
&\leq 2p^m \sum_{l=1}^{\frac{p^m-1}{2}} \frac{(l,p^m)^{\frac{1}{2}}}{l} \\
&\leq 2p^m \sum_{d|p^m} \sum_{\substack{l=1 \\ (l,p^m)=d}}^{\frac{p^m-1}{2}} \frac{d^{\frac{1}{2}}}{l} \\
&\leq 2p^m \sum_{d|p^m} \sum_{\substack{l=1 \\ (l,\frac{p^m}{d})=1}}^{[\frac{p^m-1}{2d}]} \frac{1}{ld^{\frac{1}{2}}} \\
&= O\left(p^m \log p^m \sum_{d|p^m} \frac{1}{d^{\frac{1}{2}}}\right) \\
&= O\left(p^m \log p^m \sum_{n=1}^{d(p^m)} \frac{1}{n^{\frac{1}{2}}}\right) \\
&= O(p^m \log p^m).
\end{aligned}
$$

Thus, by (4.24) we get

$$
R_0 = O\left(d_{k+1}(p^m)(p^m)^{\frac{k}{2}} \log^{k+1} p^m\right). \tag{4.25}
$$

Combining all the estimates for R_0, R_1, $\cdots$, and R_k, by (4.17), (4.18),(4.20), (4.22) and (4.25), we have

$$
\begin{aligned}
& |U_{p^m}(N,k)| \\
= \quad & \frac{\phi^k(p^m)}{(p^m)^{k+1}}N^{k+1} + \frac{1}{(p^m)^{k+1}}\sum_{i=1}^{k}\binom{k+1}{i}N^i(p^{m-1})^{k+1-i} \\
& \phi^{i-1}(p^m)\left(N - p\left[\frac{N}{p}\right]\right)^{k+1-i} + O\left(d_{k+1}(p^m)(p^m)^{\frac{k}{2}}\log^{k+1}p^m\right) \\
= \quad & \frac{1}{(p^m)^{k+1}}\sum_{i=0}^{k+1}\binom{k+1}{i}N^i(p^{m-1})^{k+1-i}\phi^{i-1}(p^m) \\
& (N - p\left[\frac{N}{p}\right])^{k+1-i} + O\left(d_{k+1}(p^m)(p^m)^{\frac{k}{2}}\log^{k+1}p^m\right) \\
= \quad & \frac{1}{(p^m)^{k+1}}\sum_{i=0}^{k+1}\binom{k+1}{i}N^i(p^{m-1})^{k+1-i}\phi^{i-1}(p^m)(N - p\left[\frac{N}{p}\right])^{k+1-i} \\
& -\frac{1}{\phi(p^m)}\left(\frac{N}{p} - \left[\frac{N}{p}\right]\right)^{k+1} + O\left(d_{k+1}(p^m)(p^m)^{\frac{k}{2}}\log^{k+1}p^m\right) \\
= \quad & \frac{1}{\phi(p^m)}\left(\left(N - \left[\frac{N}{p}\right]\right)^{k+1} - \left(\frac{N}{p} - \left[\frac{N}{p}\right]\right)^{k+1}\right) \\
& +O\left(d_{k+1}(p^m)(p^m)^{\frac{k}{2}}\log^{k+1}p^m\right).
\end{aligned}
$$

This finishes the proof of theorem 5. □

4.4 Proof of Theorem 6 and its corollary

Theorem 6 *Let N, k and t be positive integers, then we have*

$$
|U_t(N,k)| = \frac{N^{k+1}}{t^{k+1}}\phi^k(t) + O\left(\frac{N^k}{t}d(t)\log t + t^{\frac{k}{2}}2^{k+1}d^{\frac{k+1}{2}}(t)d_{k+1}(t)\log^{k+1}t\right). \tag{4.26}
$$

This estimate of $|U_t(N,k)|$ is dominated by the main term for k comparatively much smaller than N and for t of the size less than $N^{\frac{2(k+1)}{k+2}-\epsilon}$. But in comparing to Theorem 4 and Theorem 5, the main term in Theorem 6 is less explicit. This

is due to the difficulties involved in estimating the sums in $|U_t(N,k)|$ when t has more than one prime factor. More importantly, the first term in the error term of (4.26), which very likely can be calculated explicitly, would dominate the second error term of (4.26) for small t. So we indeed have a weaker result here for the general case.

Proof. Recall that

$$\sum_{g=1}^{m} e(\frac{gn}{m}) = \begin{cases} m & m|n \\ 0 & m \nmid n. \end{cases}$$

We have

$$\begin{aligned} & |U_t(N,k)| \\ = \quad & \frac{1}{t^{k+1}} \sum_{l_1,\cdots,l_{k+1}=1}^{t} \sum_{\substack{a_1,\cdots,a_{k+1}=1 \\ a_1\cdots a_{k+1}\equiv -1(mod\ t)}}^{t} \\ & \sum_{b_1,\cdots,b_{k+1}=1}^{N} e\left(\frac{l_1(a_1-b_1)+\cdots+l_{k+1}(a_{k+1}-b_{k+1})}{t}\right) \end{aligned} \tag{4.27}$$

This is again because for each (k+1)-tuple $(a_1, a_2, \ldots, a_{k+1})$ such that

$$a_1 a_2 \cdots a_{k+1} \equiv -1(mod\ t),$$

if there is a (k+1)-tuple $(b_1, b_2, \cdots, b_{k+1})$ such that

$$a_i \equiv b_i, \quad i = 1, 2, \cdots, k+1, \tag{4.28}$$

then the sums involving this (k+1)-tuple $(b_1, b_2, \cdots, b_{k+1})$ contribute t^k to $|U_t(N,k)|$, and the k-tuples not satisfying (4.26) contribute 0.

Taking $l_1 = \cdots = l_{k+1} = t$, we get the main term for $|U_t(N,k)|$, which is

$$\frac{N^{k+1}}{t^{k+1}}\phi^k(t).$$

Thus we have by (4.25),

$$|U_t(N,k)| = \frac{N^{k+1}}{t^{k+1}}\phi^k(t) + R, \tag{4.29}$$

where R denotes sum of those terms such that not all of the l_i's are equal to t.

Letting R_i be the sum of those terms in R such that i of the l_j's are equal to t, then we have

$$R = R_0 + R_1 + \cdots + R_k. \tag{4.30}$$

We will compute R_1 first, where there is only one l_i equal to t. For the case that $l_i = t$ we denote it by R_{1i}. Hence

$$R_1 = R_{11} + \cdots + R_{1(k+1)} \tag{4.31}$$

and

$$\begin{aligned} R_{11} &= \frac{1}{t^{k+1}} \sum_{l_2,\cdots,l_{k+1}=1}^{t-1} \sum_{\substack{a_1,\cdots,a_{k+1}=1 \\ a_1\cdots a_{k+1}\equiv -1 (mod\ t)}}^{t} \sum_{b_1,\cdots,b_{k+1}=1}^{N} \\ &\qquad e\left(\frac{l_2(a_2-b_2)+\cdots+l_{k+1}(a_{k+1}-b_{k+1})}{t}\right) \\ &= \frac{N}{t^{k+1}} \sum_{l_2,\cdots,l_{k+1}=1}^{t-1} \sum_{\substack{a_2,\cdots,a_{k+1}=1 \\ (a_2\cdots a_{k+1},\ t)=1}}^{t} e\left(\frac{l_2a_2+\cdots+l_{k+1}a_{k+1}}{t}\right) \\ &\qquad \sum_{b_2,\cdots,b_{k+1}=1}^{N} e\left(-\frac{l_2b_2+\cdots+l_{k+1}b_{k+1}}{t}\right) \\ &= \frac{N}{t^{k+1}} \sum_{l_2,\cdots,l_{k+1}=1}^{t-1} \sum_{\substack{a_2=1 \\ (a_2,t)=1}}^{t} e\left(\frac{l_2a_2}{t}\right)\cdots \sum_{\substack{a_{k+1}=1 \\ (a_{k+1},t)=1}}^{t} e\left(\frac{l_{k+1}a_{k+1}}{t}\right) \\ &\qquad \sum_{b_2,\cdots,b_{k+1}=1}^{N} e\left(-\frac{l_2b_2+\cdots+l_{k+1}b_{k+1}}{t}\right). \end{aligned}$$

Note that Ramanujan's sum equals(See [14])

$$\sum_{\substack{a=1 \\ (a,t)=1}}^{t} e\left(\frac{la}{t}\right) = \sum_{\substack{a=1 \\ (a,t)=1}}^{t} e\left(\frac{\frac{l}{(l,t)}a}{\frac{t}{(l,t)}}\right) = (l,t) \sum_{\substack{a=1 \\ (a,t)=1}}^{\frac{t}{(l,t)}} e\left(\frac{\frac{l}{(l,t)}a}{\frac{t}{(l,t)}}\right) = (l,t)\mu\left(\frac{t}{(l,t)}\right),$$

and if α does not equal an integer, then

$$\left|\sum_{b=1}^{N} e(\alpha b)\right| \leq \frac{1}{\|\alpha\|},$$

where $\|\alpha\|$ denotes the distance between α and the nearest integer.
We get

$$\begin{aligned}
|R_{11}| &= \frac{N}{t^{k+1}} \left| \sum_{l_2,\cdots,l_{k+1}=1}^{t-1} (l_2,t)\mu\left(\frac{t}{(l_2,t)}\right)(l_3,t)\mu\left(\frac{t}{(l_3,t)}\right)\cdots \right. \\
&\quad \left. (l_{k+1},t)\ \mu\left(\frac{t}{(l_{k+1},t)}\right) \sum_{b_2,\cdots,b_{k+1}=1}^{N} e\left(-\frac{l_2b_2+\cdots+l_{k+1}b_{k+1}}{t}\right)\right| \\
&\leq \frac{N}{t^{k+1}} \sum_{l_2=1}^{t-1}(l_2,t)\left|\sum_{b_2=1}^{N} e\left(-\frac{l_2b_2}{t}\right)\right| \sum_{l_3=1}^{t-1}(l_3,t)\left|\sum_{b_3=1}^{N} e\left(-\frac{l_3b_3}{t}\right)\right|\cdots \\
&\quad \sum_{l_{k+1}=1}^{t-1}(l_{k+1},t)\left|\sum_{b_{k+1}=1}^{N} e\left(-\frac{l_{k+1}b_{k+1}}{t}\right)\right| \\
&\leq \frac{N}{t^{k+1}}\left|\sum_{l_2=1}^{t-1}(l_2,t)\frac{1}{\|-\frac{l_2}{t}\|}\right|^k
\end{aligned}$$

Note that $\|-\frac{l}{t}\| = \|\frac{l}{t}\|$ and if $1 \leq l \leq t$, then $(l,t) = (t-l,t)$. We further have

$$\begin{aligned}
|R_{11}| &\leq \frac{N}{t^{k+1}}t^k\left|\sum_{d|t}\sum_{\substack{l=1\\(l,t)=d}}^{t-1}\frac{(l,t)}{l}\right|^k \\
&= \frac{N}{t}\left|\sum_{d|t}\sum_{\substack{l=1\\(l,\frac{t}{d})=1}}^{\frac{t}{d}-1}\frac{1}{l}\right|^k \\
&= O\left(\frac{N}{t}d^k(t)\log^k t\right).
\end{aligned}$$

Similarly we compute $R_{1,2},\cdots,R_{1,k+1}$. Thus

$$R_1 = O\left(\binom{k+1}{1}\frac{N}{t}(d(t))^k log^k t\right).$$

To compute R_i, $i = 2, \cdots, k$. We first choose $l_1 = \cdots = l_i = t$. Denote the sum by R_{i1}. We have

$$\begin{aligned}
|R_{i1}| &= \frac{\phi^{i-1}(t)N^i}{t^{k+1}} \sum_{l_{i+1},\cdots,l_{k+1}=1}^{t-1} \sum_{\substack{a_{i+1},\cdots,a_{k+1}=1 \\ (l_{i+1\,k+1},t)=1}}^{t} e\left(\frac{l_{i+1}a_{i+1}+\cdots+l_{k+1}a_{k+1}}{t}\right)\cdots \\
&\quad \cdots \sum_{b_{i+1},\cdots,b_{k+1}=1}^{N} e\left(-\frac{l_{i+1}b_{i+1}+\cdots+l_{k+1}b_{k+1}}{t}\right) \\
&\leq \frac{\phi^{i-1}(t)N^i}{t^{k+1}} \sum_{l_{i+1}=1}^{t-1} (l_{i+1},t)\mu\left(\frac{t}{(l_{i+1},t)}\right) e\left(-\frac{l_{i+1}b_{i+1}}{t}\right)\cdots \\
&\quad \cdots \sum_{l_{k+1}=1}^{t-1} (l_{k+1},t)\mu\left(\frac{t}{(l_{k+1},t)}\right) e\left(-\frac{l_{k+1}b_{k+1}}{t}\right) \\
&= O\left(\frac{N^i}{t^{k+1}}\phi^{i-1}(t)t^{k+1-i}(d(t))^{k+1-i}\log^{k+1-i}t\right) \\
&= O\left(\frac{N^i}{t}d(t)^{k+1-i}\log^{k+1-i}t\right).
\end{aligned}$$

Combining all the cases of different choices of i of the l_j's that are equal to t, we have

$$R_i = O\left(\binom{k+1}{i}\frac{N^i}{t}(d(t))^{k+1-i}\log^k t\right).$$

It remains to compute R_0,

Recall that

$$\begin{aligned} R_0 &= \frac{1}{t^{k+1}} \sum_{l_1,\cdots,l_{k+1}=1}^{t-1} \sum_{\substack{a_1,\cdots,a_{k+1}=1 \\ a_1\cdots a_{k+1}\equiv -1 (mod\ t)}}^{t} e\left(\frac{l_1a_1+\cdots+l_{k+1}a_{k+1}}{t}\right) \\ &\quad \sum_{b_1,\cdots,b_{k+1}}^{N} e\left(-\frac{l_1b_1+\cdots+l_{k+1}b_{k+1}}{t}\right) \\ &= \frac{1}{t^{k+1}} \sum_{l_1,\cdots,l_{k+1}=1}^{t-1} K(l_1,\cdots,l_{k+1},-l_{k+1};t) \\ &\quad \sum_{b_1,\cdots,\ b_{k+1}=1}^{N} e\left(-\frac{l_1b_1+\cdots+l_{k+1}b_{k+1}}{t}\right). \end{aligned}$$

Thus by Theorem G, we get

$$|R_0| \le \frac{1}{t^{k+1}} \sum_{l_1,\cdots,l_{k+1}=1}^{t-1} t^{\frac{k}{2}} (l_1,\cdots,l_k,-l_{k+1};t)_k^{\frac{k}{2}}\ d_{k+1}(t)\ \frac{1}{\|\frac{l_1}{t}\|}\cdots\frac{1}{\|\frac{l_k}{t}\|}\frac{1}{\|-\frac{l_{k+1}}{t}\|}.$$

Note that by Lemma 1

$$(l_1,\cdots,l_k,-l_{k+1};t)_k^{\frac{k}{2}} \le (l_1,t)^{\frac{1}{2}}\cdots(l_{k+1},t)^{\frac{1}{2}}$$

We further have

$$\begin{aligned} & |R_0| \\ \le\ & \frac{1}{t^{k+1}} \sum_{l_1,\cdots,l_{k+1}=1}^{t-1} t^{\frac{k}{2}} (l_1,t)^{\frac{1}{2}}\cdots(l_{k+1},t)^{\frac{1}{2}}\ d_{k+1}(t)\frac{1}{\|\frac{l_1}{t}\|}\cdots\frac{1}{\|\frac{l_k}{t}\|}\frac{1}{\|-\frac{l_{k+1}}{t}\|} \\ \le\ & \frac{d_{k+1}(t)}{t^{k+1}}\ 2^{k+1} \sum_{l_1=1}^{\frac{t}{2}} \frac{t}{l_1}(l_1,t)^{\frac{1}{2}}\cdots\sum_{l_{k+1}=1}^{\frac{t}{2}} \frac{t}{l_{k+1}}(l_{k+1},t)^{\frac{1}{2}}. \end{aligned}$$

Note also that

$$\begin{aligned}
\sum_{l=1}^{\frac{t}{2}} \frac{(l,t)^{\frac{1}{2}}}{l} &= \sum_{d|t} \sum_{\substack{l=1 \\ (l,t)=d}}^{\frac{t}{2}} \frac{d^{\frac{1}{2}}}{l} \\
&= \sum_{d|t} \sum_{\substack{l=1 \\ (l,\frac{t}{d})=1}}^{\frac{t}{2d}} \frac{1}{ld^{\frac{1}{2}}} \\
&= O\left(\log t \sum_{d|t} \frac{1}{d^{\frac{1}{2}}}\right) \\
&= O\left(\log t \sum_{n=1}^{d(t)} \frac{1}{n^{\frac{1}{2}}}\right) \\
&= O(d^{\frac{1}{2}}(t) \log t).
\end{aligned}$$

Thus we have

$$|R_0| = O(t^{\frac{k}{2}} 2^{k+1} \log^{k+1} t d^{\frac{k+1}{2}}(t) d_{k+1}(t)).$$

Combining the estimate for $R_0,\ R_1,\ \cdots,\ R_k$, we have

$$\begin{aligned}
|U_t(N,k)| &= \frac{N^{k+1}}{t^{k+1}}\phi^k(t) + O\left(\binom{k+1}{1}\frac{N}{t}(d(t))^k \log^k t\right. \\
&\quad + \cdots + \binom{k+1}{i}\frac{N^i}{t}(d(t))^{k+1-i}\log^{k+1-i} t \\
&\quad \left. + \cdots + \binom{k+1}{k}\frac{N^k}{t}d(t)\log t + t^{\frac{k}{2}}2^{k+1}\log^{k+1}(d(t))^{\frac{k+1}{2}} d_{k+1}(t)\right) \\
&= \frac{N^{k+1}}{t^{k+1}}\phi^k(t) + O\left(\frac{N^k}{t}d(t)\log t + t^{\frac{k}{2}}2^{k+1}d^{\frac{k+1}{2}}(t)d_{k+1}(t)\log^{k+1} t\right).
\end{aligned}$$

This finishes the proof of Theorem 6. □

Corollary 3 *Let N, k be integers. Let ϵ be a real number with $0 < \epsilon < 1$. For each integer t with $t < N^{2\frac{k+1}{k+2}-\epsilon}$, we can find integers $a_1, \cdots, a_{k+1}$ with $1 \le a_i \le N$,*

$i = 1, 2, \cdots, k+1$, *such that*

$$t \mid a_1 \cdots a_{k+1} + 1.$$

Proof. Note that $d(t) = O(t^\epsilon)$, $d_{k+1}(t) = O((d(t))^{k+1}) = O(t^{(k+1)\epsilon})$, $2^k = O(t^\epsilon)$ and $\log^{k+1}(t) = O(t^\epsilon)$. The result then follows by comparing the main term and the error term of Theorem 6. □

Note that t could be either a prime or power of a prime number.

Bibliography

[1] T. Apostol, Introduction to analytic number theory, Springer-Verlag,1976

[2] E. Bombieri, On exponential sums in finite fields,II,Invent. Math. 47(1978),29-39.

[3] L. Carlitz, A note on multiple exponential sums, Pacific J. Math. 15(1965),757-765.

[4] P. Deligne, Applications de la formule des traces aux sommes trigonométriques, in Séminaire de géométrie algébrique du Bois-Marie, S.G.A. $4\frac{1}{2}$, pp.168-232, Lecture notes in mathematics, No.569, Springer-Verlag,Berlin/New York,1977.

[5] P. Deligne, La conjecture de Weil,I, Inst. Hautes Études Sci. Publ. Math. 43(1974). 273-307.

[6] J. Deshouillers and H. Iwaniec, On the greatest prime factor of $n^2 + 1$, Ann. Inst. Fourier, Grenoble 32, 4(1982), 1-11.

[7] H. Davenport, On certain exponential sums, J. Reine Angew. Math. 169(1933),158-176.

[8] H. Davenport, Multiplicative number theory, Springer-Verlag, 1980.

[9] T. Estermann, On Kloosterman's sum, Mathematika 8(1961), 83-86.

[10] E. Fouvry and N. Katz, A general stratification theorem for exponential sums, and applications, J. Reine. Angew. Math 540(2001), 115-166.

[11] J. Friedlander and H. Iwaniec, Estimates for character sums; Proc. Amer. Math. Soc. 119(1993), 365-372.

[12] G. Greaves, Large prime factors of binary forms, J. Number Theory 3(1971), 35-39.

[13] K. Györy, A. Sárközy, C. Stewart, On the number of prime factors of the form $ab + 1$, Acta Arith. 74(1996), 365-385.

[14] G. Hardy and E. Wright, An introduction to the Theory of Numbers,third edition, Oxford Press, Clarendon.

[15] H. Hasse, Theorie der relative-zyklischen algebraischen Functionenkörper,insbefondere bei endlichem Konstantenkörper, J. Reine Angew. Math. 172(1934),37-54.

[16] C. Hooley, An asymptotic formula in the theory of numbers, J. London Math. Soc. 7(1957), 396-413.

[17] C. Hooley, On the greatest prime factor of a quadratic polynomial, Acta Math. 117(1967), 281-299.

[18] H.D. Kloosterman, On the representations of a number in the form $ax^2 + by^2 + cz^2 + dt^2$, Acta Math. 49(1926), 407-464.

[19] L.J.Mordell, On a special polynomial congruence and exponential sums, in"Calcutta Math. Soc. Golden Jubilee Commemoration Volume," Part I, pp. 29-32, 1958/1959.

[20] I. Ruzsa, Large prime factors of sums, Studia Sci. Math. Hungar 27(1992), 463-470

[21] Chengdong Pan and Chengbiao, Analytic number theory, Scientific Press,1991.

[22] Chengdong Pan and Chengbiao, Elementary number theory, Peking University Press,1991.

[23] H. Salié, Über die Kloosterman schen summen S(u,v,q), Math. Z.34(1931),91-109.

[24] A. Sárközy and C. Stewart, On divisors of sums of integers, II. J. Reine. Angew. Math 365(1986), 171-191.

[25] A. Sárközy and C. Stewart, On prime factors of integers of the form $ab + 1$, Publ. Math. Debrecen. 56(2000), 559-573.

[26] A. Sárközy, On sums $a + b$ and numbers of the form $ab + 1$ with many prime factors, Grazer Math. Ber. 318(1992), 141-154.

[27] A. Sárközy, On the average value for the number of divisors of numbers of form $ab + 1$, Acta Math. Hungar. 66(1995), 223-245.

[28] J.P.Serre, Majorations des sommes exponential, Astérisque 41/42(1977, 111-122.

[29] R. Smith, On n-dimensional Kloosterman sums, J. Number Theory, 11(1979), 324-343.

[30] C. Stewart, On the greatest prime factor of integers of the form $ab+1$, Periodica Math. Hung. 43(2001),81-91.

[31] I.M. Vinogradov, An introduction to the theory of numbers, Pergamon press, London and New York, 1955.

[32] A. Weil, Number of solutions of equations in finite fields, Bull. Ameri. Math. Soc. 55(1949),497-508.

[33] A. Weil, On some exponential sums, Proc. Nat. Acad. Sci. USA 34(1948), 204-207.

[34] A. Weil, On the Riemann hypothesis in function fields, Proc. Nat. Acad. Sci. USA 27(1941), 345-347.