



Contents lists available at ScienceDirect

Journal of Number Theory

www.elsevier.com/locate/jnt



General Section

Modular hyperbolas and Beatty sequences



Marc Technau

*Institute of Analysis and Number Theory, Graz University of Technology,
Kopernikusgasse 24, 8010 Graz, Austria*

ARTICLE INFO

Article history:

Received 1 August 2018

Received in revised form 8 July 2019

Accepted 29 July 2019

Available online 27 August 2019

Communicated by L. Smajlovic

MSC:

primary 11B83

secondary 11L05, 11L07, 11D79

Keywords:

Beatty sequence

Modular hyperbola

Kloosterman sum

ABSTRACT

Bounds for $\max\{m, \tilde{m}\}$ subject to $m, \tilde{m} \in \mathbb{Z} \cap [1, p)$, p prime, z indivisible by p , $m\tilde{m} \equiv z \pmod{p}$ and m belonging to some fixed Beatty sequence $\{\lfloor n\alpha + \beta \rfloor : n \in \mathbb{N}\}$ are obtained, assuming certain conditions on α . The proof uses a method due to Banks and Shparlinski. As an intermediate step, bounds for the discrete periodic autocorrelation of the finite sequence $0, e_p(y\bar{1}), e_p(y\bar{2}), \dots, e_p(y(p-1))$ on average are obtained, where $e_p(t) = \exp(2\pi it/p)$ and $m\bar{m} \equiv 1 \pmod{p}$. The latter is accomplished by adapting a method due to Kloosterman.

© 2019 Elsevier Inc. All rights reserved.

1. Introduction

Consider the *modular hyperbola*

$$\mathcal{H}_{z \bmod p} = \{(m, \tilde{m}) \in \mathbb{Z}^2 \cap [1, p)^2 : m\tilde{m} \equiv z \pmod{p}\} \quad (\text{with } p \nmid z), \quad (1.1)$$

where the letter p denotes a prime here and throughout. The distribution of points on these modular hyperbolas has attracted wide interest and the interested reader is

E-mail address: mtechnau@math.tugraz.at.

referred to [28] for a survey of questions related to this topic and various applications. Our particular starting point is the following intriguing property of such hyperbolas:

Theorem 1.1. *For any prime p and z coprime to p there is always a point $(m, \tilde{m}) \in \mathcal{H}_{z \bmod p}$ with $\max\{m, \tilde{m}\} \leq 2(\log p)p^{3/4}$.*

Proof. See, e.g., [17, p. 382].¹ $\square$

Loosely speaking, the above theorem states that, when sampling enough points

$$(1, ?), (2, ?'), (3, ?''), \dots \in \mathcal{H}_{z \bmod p},$$

at least one amongst the second coordinates will not be too large.

Here we shall investigate analogous distribution phenomena when one of the coordinates is required to belong to some fixed Beatty set. The *Beatty set* $\mathcal{B}(\alpha, \beta)$ associated with real *slope* $\alpha \geq 1$ and *shift* $\beta \geq 0$ is given by

$$\mathcal{B}(\alpha, \beta) = \{\lfloor n\alpha + \beta \rfloor : n \in \mathbb{N}\},$$

where $\lfloor x \rfloor$ denotes the largest integer less than or equal to x . When thinking of $\mathcal{B}(\alpha, \beta)$ as ordered according to size of its elements, we call it a *Beatty sequence*. Such sequences appeared first in the astronomical studies of Johann III Bernoulli [12] as a means to control the accumulative error when calculating successive multiples of an approximation to some given number. Later, Beatty sequences were studied by Elwin Bruno Christoffel [14, 15] with respect to their arithmetical significance with the goal of easing the discomfort present during that time when working with irrational numbers. The name “Beatty sequence” is in the honour of Samuel Beatty who popularised these sequences by posing a problem for solution in *The American Mathematical Monthly* [8, 9]; the theorem to be proved there appears to be due to John William Strutt (3rd Baron Rayleigh) [29] though. A reader interested in more recent work on Beatty sequences is invited to start his or her journey by tracing the references in [2], for instance.

Returning to modular hyperbolas, the question we seek to answer may be enunciated as follows:

Question A. If the first coordinate of $(m, \tilde{m})$ in Theorem 1.1 is additionally required to belong to a fixed Beatty set with irrational slope, can one still prove a result like Theorem 1.1, i.e., is it impossible for a Beatty set to contain only those elements $m < p$ with ‘large’ corresponding $\tilde{m}$ from (1.1)?

¹ The proof itself gives much more. Indeed, instead of restricting $(m, \tilde{m})$ to the square box $(0, 2(\log p)p^{3/4}]^2$, any rectangular box $\mathcal{R} \subseteq (0, p]^2$ with sufficiently large area can be seen to have a non-empty intersection with $\mathcal{H}_{z \bmod p}$.

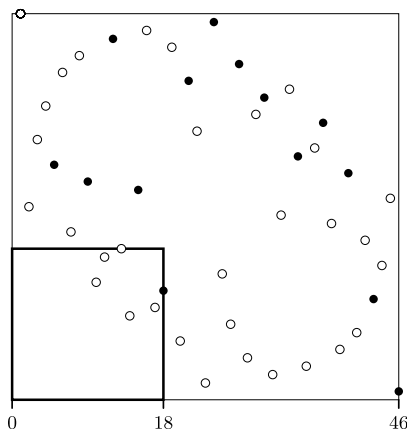


Fig. 1. An illustration of the points $(m, \tilde{m}) \in \mathcal{H}_{-1 \bmod 47}$. The points $(m, \tilde{m})$ with m belonging to the Beatty set $\mathcal{B}(\pi, e)$ are filled.

More specifically, for fixed irrational $\alpha > 1$ and non-negative β , we shall be interested in solving

$$m\tilde{m} \equiv z \pmod{p} \quad \text{with} \quad m \in \mathcal{B}(\alpha, \beta) \quad \text{and} \quad 1 \leq m, \tilde{m} < p \quad (1.2)$$

whilst keeping $\max\{m, \tilde{m}\}$ as small as possible. Thus, we shall want to bound

$$F(z \bmod p) = \min\{\max\{m, \tilde{m}\} : (m, \tilde{m}) \text{ such that (1.2) holds}\}. \quad (1.3)$$

As an illustration, consider Fig. 1: therein, bounding $F(-1 \bmod 47)$ is equivalent to asking how large one must take the side length of a square with lower left corner positioned at $(0, 0)$, in order to be guaranteed to find a black point inside. The smallest such square is sketched thick in the figure.

The main result of this paper is that, for any fixed α satisfying certain Diophantine conditions and arbitrary (not necessarily fixed) non-negative β , the quantity $F(z \bmod p)$ is ‘not too large’ in the sense that

$$F(z \bmod p) \leq \alpha p^{684/727} \log p + \beta \quad (1.4)$$

for all sufficiently large p (in terms of α) and arbitrary z indivisible by p . (Here the main point is that the exponent of p in (1.4) is strictly smaller than 1.) We refer to Theorem 2.1 for the precise formulation.

2. Main results

2.1. A partial answer to Question A

In order to describe the numbers α for which we obtain results, recall that the *type* τ of an irrational number α is defined by

$$\tau = \sup \left\{ \eta \in \mathbb{R} : \liminf_{q \rightarrow \infty} q^\eta \|\alpha q\| = 0 \right\},$$

where $\|\varrho\| = \min_{n \in \mathbb{Z}} |\varrho - n|$ denotes the distance to a nearest integer. (Note that Dirichlet's approximation theorem ensures that $\tau \geq 1$.) We say that α is of *finite type* if $\tau < \infty$.

Our main result may now be enunciated as follows:

Theorem 2.1. *Let $\epsilon > 0$ and β be non-negative. Moreover, assume that $\alpha > 1$ is irrational and of finite type $\leq \frac{512}{43} - \epsilon$. Then there is a number $p_0(\alpha, \epsilon)$ such that, for all primes $p \geq p_0(\alpha, \epsilon)$ and z coprime to p , there is a point $(m, \tilde{m}) \in \mathcal{H}_{z \bmod p}$ with $m \in \mathcal{B}(\alpha, \beta)$ and*

$$\max\{m, \tilde{m}\} \leq \alpha p^{684/727} \log p + \beta. \quad (2.1)$$

Remark 2.2. After posting an initial draft of this article on ‘the arXiv’, I. E. Shparlinski kindly pointed out to the author that a generalised form of Weil’s bound also applies directly to the sum (2.4). This leads to a chain of improvements in Proposition 4.2, Theorem 4.1, and Theorem 2.1. We sketch the relevant changes in Section 6 and close with a discussion of two directions in which our results may potentially be generalised.

Remark 2.3. If $(m, \tilde{m})$ is a solution of (1.2), then $m \geq \min \mathcal{B}(\alpha, \beta) = \lfloor \alpha + \beta \rfloor > \beta$. Hence, the dependence of the right hand side of (2.1) on β is not a deficiency in the proof, but instead occurs naturally.

2.2. Intermediate results and discussion of the methods

The proof of Theorem 2.1 is carried out in Section 5 via a discrete Fourier transform method. It is an adaptation of the argument which is used in [17] to prove Theorem 1.1. Here one naturally encounters the discrete Fourier transform of the characteristic function of points whose (integer) coordinates are inverse to each other modulo p , namely *Kloosterman sums*:

$$S(x, y; p) = \sum_{1 \leq m < p} e_p(xm + y\overline{m}). \quad (2.2)$$

(Here x and y are integers, $e_p(r) = \exp(2\pi i r/p)$, and $\overline{m}$ is the unique positive integer below p which is inverse to m modulo p .) These sums were first introduced by Kloosterman [21] in his seminal refinement of the Hardy–Littlewood circle method to handle diagonal quadratic forms in four variables. He proved the bound

$$\sum_{1 \leq m < p} e_p(xm + y\overline{m}) \ll p^{3/4} \gcd(x, y, p)^{1/4},$$

which was later improved by Weil [31] to

$$\left| \sum_{1 \leq m < p} e_p(xm + y\overline{m}) \right| \leq 2p^{1/2} \gcd(x, y, p)^{1/2},$$

the latter bound being asymptotically optimal (see [18, Section 11.7]).

Upon using such a Fourier transform approach to bound (1.3) without imposing the restriction $m \in \mathcal{B}(\alpha, \beta)$ in (1.2), one naturally has to deal with *incomplete* Kloosterman sums:

$$\sum_{1 \leq m < M} e_p(xm + y\overline{m}) \quad (\text{with } 1 \leq M \leq p).$$

Using a completing technique, bounds for such sums can be derived from bounds for the complete sum (2.2). In our setting—reinstating the restriction $m \in \mathcal{B}(\alpha, \beta)$ in (1.2)—one has to deal with incomplete Kloosterman sums along the Beatty sequence in question:

$$K_{\alpha, \beta}(x, y; p, N) = \sum_{\substack{1 \leq n \leq N \\ p \nmid [n\alpha + \beta]}} e_p(x[n\alpha + \beta] + y\overline{[n\alpha + \beta]}). \quad (2.3)$$

We are able to bound such sums (see Theorem 4.1 below) by means of a method due to Banks and Shparlinski [3,4,5]. In order to succeed in our particular case, the method requires bounds for sums of the shape

$$S(x, y, w; p) = \sum_{\substack{1 \leq m < p \\ p \nmid (m+w)}} e_p(x\overline{m} + y\overline{m+w}) \quad (2.4)$$

and this input is obtained by adapting Kloosterman's original method [21] for bounding his sums (see Proposition 4.2 below, and Remark 5.4 for more context on the similarity to Kloosterman's argument). The quality of the bounds we obtain depends on the Diophantine properties of the slope α of the Beatty sequence $\mathcal{B}(\alpha, \beta)$.

2.3. Plan of the paper

The rest of the paper is structured as follows: first, we fix some notation and recall some basic facts related to Diophantine properties of irrational numbers α . Then, in Section 4, we state our results concerning Kloosterman sums along Beatty sequences (2.3). In Section 5, we provide the proofs for our results. Finally, in Section 6, we sketch how the improvements hinted at in Remark 2.2 can be obtained by adjusting the relevant parts of our arguments. The last part of Section 6 contains a discussion of potential generalisations of our work.

3. Preliminaries

3.1. Notation

Before proceeding, we fix some notation used throughout the paper. We use the Vinogradov symbol $f(x) \ll g(x)$ and the Big Oh notation $f(x) = O(g(x))$ to mean $|f(x)| \leq Cg(x)$ for some absolute constant $C > 1$. In cases where the constant depends on a parameter (usually α and sometimes $\epsilon > 0$), we indicate this with subscripts.

3.2. Discrepancy

For a real number ϱ we denote its *fractional part* by $\{\varrho\} = \varrho - \lfloor \varrho \rfloor \in [0, 1)$. We denote the *discrepancy* of the finite sequence $(\{n\alpha + \beta\})_{n \leq N}$ by

$$D_{\alpha, \beta}(N) = \sup_{0 \leq x < y < 1} \left| \frac{\#\{n \leq N : \{n\alpha + \beta\} \in [x, y)\}}{N} - (y - x) \right|, \quad (3.1)$$

and we put $D_{\alpha}(N) = D_{\alpha, 0}(N)$. The discrepancy (3.1) is a measure for how well distributed the finite sequence $(\{n\alpha + \beta\})_{n \leq N}$ is. A drawback of the above definition is that it is needlessly dependent on the shift β (consider, for instance, $(\alpha, N) = (\frac{15}{8}, 2)$ and $\beta \in \{0, \frac{1}{8}\}$). This dependency could be removed by defining the discrepancy as, e.g., Montgomery [24] does, but we refrain from doing so for the purpose of being able to use results from [22] below. Nonetheless, the dependence of (3.1) on β can be easily sidestepped by means of the following well-known fact:

Lemma 3.1. *For any $\alpha, \beta \in \mathbb{R}$ and $N \in \mathbb{N}$, we have $D_{\alpha, \beta}(N) \leq 8D_{\alpha}(N)$.*

Proof. This follows easily from [22, Chapter 2, Theorem 1.3] and [24, Chapter 1, Eq. (8)]. $\square$

In view of the above, it suffices to consider $D_{\alpha}(N)$ in the sequel. It is a classic result due to Bohl, Sierpiński and Weyl that the discrepancy of $(\{n\alpha\})_{n \leq N}$ tends to zero (see [32] and the references therein). Namely, we have:

Lemma 3.2 ([32, Satz 2]). *Let α be irrational. Then*

$$D_{\alpha}(N) = o_{\alpha}(1), \quad \text{as } N \rightarrow \infty,$$

where the implied constant only depends on α .

For α of finite type as defined in Section 2, Lemma 3.2 can be sharpened considerably:

Lemma 3.3 ([22, Chapter 2, Theorem 3.2]). *Let α be of finite type τ . Then, for every $\epsilon > 0$,*

$$D_\alpha(N) \ll_{\alpha, \epsilon} N^{-1/\tau + \epsilon},$$

where the implied constant only depends on α and ϵ .

4. Bounds for Kloosterman sums along Beatty sequences

In this section we state our bounds for (2.3) (see Theorem 4.1 below). However, before doing so, we give a short survey of previous work on sums along Beatty sequences. In fact, as a general heuristic, for any fixed Beatty sequence $\mathcal{B}(\alpha, \beta)$ and ‘reasonable’ arithmetical function f , one should expect that

$$\sum_{\substack{1 \leq m \leq x \\ m \in \mathcal{B}(\alpha, \beta)}} f(m) = \frac{1}{\alpha} \sum_{1 \leq m \leq x} f(m) + \{\text{error term}\} \quad \text{as } x \rightarrow \infty, \quad (4.1)$$

with some error term which one expects to be able to bound non-trivially. Indeed, this heuristic principle is substantiated by a sizable list of particular results:

- d , the *divisor function* (i.e., the function giving the number of positive divisors of its argument), starting with Abercrombie [1], improved by Begunts [10], and work on generalised divisor functions by Zhai [33] and Lü and Zhai [23].
- Certain multiplicative functions like $n \mapsto n^{-1}\varphi(n)$, where φ is *Euler’s totient*, due to Begunts [11], and a certain class of multiplicative f in the work of Güloğlu and Nevans [16].
- Dirichlet-characters and special functions related to the orbits of elements $g \bmod m$ along Beatty sequences, treated by Banks and Shparlinski [3] (see also [4]), with improvements when f is the Legendre-Symbol due to Banks, Garaev, Heath-Brown, and Shparlinski [7].
- ω , the function counting the number of distinct prime divisors of its argument (without multiplicity), and $n \mapsto (-1)^{\Omega(n)}$, where $\Omega(n)$ counts the number of distinct prime divisors of n with multiplicity, due to Banks and Shparlinski [5].
- Λ , the *von Mangoldt lambda function*, studied by Banks and Shparlinski [6] who attribute earlier results to Ribenboim [26, Chapter 4.V], although it seems that such observations were already apparent to Heilbronn in 1954, as evidenced in [30, Notes to Chapter XI].

Moreover, by a result due to Abercrombie, Banks, and Shparlinski [2], the heuristic principle is seen to hold in surprising generality for a large class of functions f in a metric sense.

However, the problem of bounding (2.3) is a little more subtle, as the function

$$f(m) = \begin{cases} e_p(xm + y\overline{m}) & p \nmid m, \\ 0 & p \mid m, \end{cases}$$

for which we might want to invoke some result of the type given in (4.1), also depends on x , y and p , and we lack the necessary uniformity in those parameters. Nonetheless, we obtain the following:

Theorem 4.1. *Suppose that p is a prime, $\beta \geq 0$ and $x, y \in \mathbb{Z}$ such that $p \nmid y$. Then, for any irrational $\alpha > 1$ and $N \leq p$, the sum $K_{\alpha,\beta}(x, y; p, N)$ given by (2.3) satisfies the bound*

$$|K_{\alpha,\beta}(x, y; p, N)| \ll_{\alpha} N^{297/512} p^{43/128} + ND_{\alpha}(N), \quad (4.2)$$

where the implied constant only depends on α .

The proof of this result is given in Section 5 and makes use of an estimate for the periodic autocorrelation of the finite sequence

$$0, e_p(y\overline{1}), e_p(y\overline{2}), \dots, e_p(y\overline{(p-1)})$$

on average. Such a bound is furnished by the next result applied with $x = -y$:

Proposition 4.2. *Suppose that p is a prime and $x, y \in \mathbb{Z}$ such that $p \nmid y$. If $S(x, y, w; p)$ is given by (2.4), then, for any set $\mathcal{W}$ of pairwise incongruent numbers modulo p , we have*

$$\sum_{w \in \mathcal{W}} |S(x, y, w; p)| \ll (\#\mathcal{W})^{3/4} p.$$

Turning back to Theorem 4.1 and recalling its assumption that $N \leq p$, we find that the first term on the right hand side of (4.2) is dominant, provided that the type of α is strictly smaller than $\frac{512}{43}$. As an immediate consequence we may state:

Corollary 4.3. *Let $\epsilon > 0$. Under the hypotheses of Theorem 4.1, and restricting to only those irrational $\alpha > 1$ of finite type $< \frac{512}{43}$, we have the bound*

$$|K_{\alpha,\beta}(x, y; p, N)| \ll_{\alpha} N^{297/512} p^{43/128}. \quad (4.3)$$

In order to put this into context, note that the famous result of Khintchine [20] asserts that almost all real numbers are of type 1 and the celebrated theorem of Roth [27] establishes that all real algebraic irrational numbers are of type 1; for numbers of finite type larger than 1, the Lebesgue measure fails to provide useful information about

their abundance. However, by the Jarník–Besicovitch theorem [19,13], we know that the Hausdorff dimension of all real numbers of type $> \tau$ is

$$\dim_{\mathbb{H}}\{x \in \mathbb{R} \text{ of type } > \tau\} = \frac{2}{\tau + 1}.$$

In particular, bounds for the Hausdorff dimension of the set of those α , for which we obtain no results, are immediate.

Furthermore, observe that Theorem 4.1 is only non-trivial for p and N in some range of the type

$$N \leq p \ll_{\epsilon} N^{5/4-\epsilon}.$$

However, in view of Lemma 3.2, for fixed irrational α , Theorem 4.1 always beats the trivial bound $|K_{\alpha,\beta}(x, y; p, N)| \leq N$ in the above range provided that N is large enough.

5. Proofs

5.1. Proof of Theorem 2.1

We start by fixing some irrational $\alpha > 1$ and non-negative β . To detect pairs $(m, \tilde{m})$ satisfying (1.2), we adapt the arguments given in [17, p. 382]. Taking

$$f_M(\tilde{m}) = \begin{cases} 1 & \text{if there exists } 1 \leq m \leq M \text{ such that (1.2) holds,} \\ 0 & \text{otherwise,} \end{cases} \quad (5.1)$$

one immediately observes that $F(z \bmod p)$ from (1.3) may be written as

$$F(z \bmod p) = \min \left\{ 1 \leq M < p : \sum_{1 \leq \tilde{m} \leq M} f_M(\tilde{m}) > 0 \right\}.$$

Therefore, in order to bound $F(z \bmod p)$, it suffices to establish the positivity of

$$\sum_{1 \leq \tilde{m} \leq M} f_M(\tilde{m}) \quad (5.2)$$

for some $M \in [\alpha + \beta, p)$, which one would like to be as small as possible. To achieve this, one notes that (5.2) should be well approximated by the average

$$\frac{M}{p} \sum_{1 \leq \tilde{m} < p} f_M(\tilde{m}) = \frac{M}{p} \#\{m \leq M : m \in \mathcal{B}(\alpha, \beta)\} > \frac{M(M - \alpha - \beta)}{p\alpha}. \quad (5.3)$$

This can be made precise by a standard completing technique. We employ this in the following form given in [17]:

Lemma 5.1. *Let $(A_m)_m$ be a sequence of complex numbers with period q , and let $\hat{A}_k$ be the discrete Fourier transform $\sum_{m=1}^q A_m e_q(mk)$. Then, for any integers $a < b$, we have*

$$\left| \sum_{a < m \leq b} A_m - \frac{b-a}{q} \hat{A}_0 \right| \leq (\log q) \max_{1 \leq k < q} |\hat{A}_k|.$$

Proof. A slightly more general statement along with its proof can be found in [18, Lemma 12.1]. $\square$

Returning to our analysis of (5.2) and applying Lemma 5.1, we infer that

$$\left| \sum_{1 \leq \tilde{m} \leq M} f_M(\tilde{m}) - \frac{M}{p} \sum_{1 \leq \tilde{m} < p} f_M(\tilde{m}) \right| \leq (\log p) \max_{1 \leq k < p} \left| \sum_{1 \leq \tilde{m} < p} f_M(\tilde{m}) e_p(k\tilde{m}) \right|. \quad (5.4)$$

Clearly, if the right hand side of (5.4) does not exceed the right hand side of (5.3), then (5.2) is guaranteed to be positive. For $1 \leq k < p$, we have

$$\sum_{1 \leq \tilde{m} < p} f_M(\tilde{m}) e_p(k\tilde{m}) = \sum_{1 \leq \tilde{m} < p} \sum_{\substack{1 \leq m \leq M \\ m \in \mathcal{B}(\alpha, \beta) \\ m\tilde{m} \equiv z \pmod{p}}} e_p(k\tilde{m}) = \sum_{\substack{1 \leq m \leq M \\ m \in \mathcal{B}(\alpha, \beta)}} e_p(kz\overline{m}).$$

Thus, in conclusion, if $M < p$ is a positive integer such that

$$\max_{1 \leq y < p} \left| \sum_{\substack{1 \leq m \leq M \\ m \in \mathcal{B}(\alpha, \beta)}} e_p(y\overline{m}) \right| \leq \frac{M(M - \beta - \alpha)}{p(\log p)\alpha}, \quad (5.5)$$

then $F(z \bmod p) \leq M$. (It is perhaps worthwhile to observe at this point that the bound on $F(z \bmod p)$ produced by this method is wholly independent of $z \not\equiv 0 \pmod{p}$. Indeed, this is a recurring feature in many problems which are similar to Question A, as is perhaps already foreshadowed by the footnote on page 149. We refer to [28] for more information.)

We now put $M = \lfloor N\alpha + \beta \rfloor$ (with $N \geq 3$ as to have $M \geq \alpha + \beta$, as required earlier) in (5.5) and employ Corollary 4.3, tacitly assuming its hypotheses in the process. Then, after a short calculation, we find that (5.5) is satisfied provided that

$$(\log p) N^{-2+297/512} p^{1+43/128} \ll_{\alpha, \epsilon} 1. \quad (5.6)$$

Letting

$$N = \lfloor p^{684/727} \log p \rfloor,$$

the condition (5.6) will be satisfied for p sufficiently large in terms of α and ϵ . This completes the proof of Theorem 2.1.

5.2. Proof of Proposition 4.2

We start by deducing Proposition 4.2 from the following result:

Proposition 5.2. *Under the assumptions of Proposition 4.2, we have*

$$\sum_{w \in \mathcal{W}} |S(x, y, w; p)|^4 \ll p^4. \quad (5.7)$$

Proof of Proposition 4.2. Note that, by Hölder's inequality, we have

$$\sum_{w \in \mathcal{W}} |S(x, y, w; p)| \leq (\#\mathcal{W})^{3/4} \cdot \left(\sum_{w \in \mathcal{W}} |S(x, y, w; p)|^4 \right)^{1/4}.$$

The result now follows after applying Proposition 5.2. $\square$

To prove Proposition 5.2, we adapt Kloosterman's original argument for bounding his sums (see [21, Section 2.43] and the comments made in Remark 5.4 below). The argument is based on using a transformation property of the sums $S(x, y, w; p)$ (see (5.13) below) in order to find a large contribution of $|S(x, y, w; p)|^4$ in an average over the first three parameters (see (5.14)). This average can be seen to count the number of solutions to certain congruence equations (see (5.15)), and this number can be bounded non-trivially.

We start with an analysis of the congruence equations in question. To this end, write $\mathbf{m}$ for an integer vector (m_1, m_2, m_3, m_4) . For any integer u , we let

$$A(\mathbf{m}, u) = \sum_{1 \leq k \leq 4} (-1)^k \overline{m_k + u} \quad (5.8)$$

if $p \nmid (m_1 + u)(m_2 + u)(m_3 + u)(m_4 + u)$ and, for notational simplicity, we put $A(\mathbf{m}, u) = 1$ otherwise. The key result is the following:

Lemma 5.3. *Let p be a prime and $A(\mathbf{m}, u)$ be given by (5.8) and suppose that*

$$\mathcal{X}_u = \{\mathbf{m} \in \{1, \dots, p\}^4 : A(\mathbf{m}, 0) \equiv A(\mathbf{m}, u) \equiv 0 \pmod{p}\}. \quad (5.9)$$

Then

$$\sum_{1 \leq u \leq p} \#\mathcal{X}_u \ll p^3.$$

Proof. Consider $\mathbf{m} = (m_1, m_2, m_3, m_4) \in \mathcal{X}_u$. Certainly we have

$$p \nmid \prod_{1 \leq k \leq 4} m_k(m_k + u). \quad (5.10)$$

Additionally, from $p \mid A(\mathbf{m}, 0)$ and $p \mid A(\mathbf{m}, u)$, we deduce that

$$\begin{cases} \overline{m_1} + \overline{m_3} \equiv \overline{m_2} + \overline{m_4} & \text{mod } p, \\ \overline{m_1 + u} + \overline{m_3 + u} \equiv \overline{m_2 + u} + \overline{m_4 + u} & \text{mod } p. \end{cases} \quad (5.11)$$

This is satisfied trivially if (m_1, m_3) is a permutation of (m_2, m_4) . There are at most $2(p-1)^2$ such trivial solutions. Assume next that $\mathbf{m}$ is a non-trivial solution to the system (5.11) and, additionally, that neither expression in (5.11) is $\equiv 0$; this additional assumption excludes at most $2(p-1)^2$ values of $\mathbf{m}$. From (5.11) it follows that

$$\begin{aligned} & (\overline{m_1} + \overline{m_3})m_1m_3(m_2 + m_4) \\ & \equiv (m_2 + m_4)(m_1 + m_3) \\ & \equiv (m_1 + m_3)m_2m_4(\overline{m_2} + \overline{m_4}) \pmod{p}. \end{aligned}$$

Upon cancelling $\overline{m_1} + \overline{m_3} \equiv \overline{m_2} + \overline{m_4} \not\equiv 0 \pmod{p}$, we find that

$$m_1m_3(m_2 + m_4) \equiv (m_1 + m_3)m_2m_4 \pmod{p},$$

which in turn may be rearranged to

$$m_1(m_2m_3 + m_3m_4 - m_2m_4) \equiv m_2m_3m_4 \pmod{p}. \quad (5.12)$$

Note that the term $m_2m_3 + m_3m_4 - m_2m_4 \pmod{p}$ cannot vanish, for otherwise (5.12) implies $p \mid m_2m_3m_4$, in contradiction to (5.10). Hence, for any $\mathbf{m}$ satisfying the above assumptions, we may compute $m_1 \pmod{p}$ from (m_2, m_3, m_4) . Indeed, by (5.12),

$$m_1 \equiv m_2m_3m_4 \overline{(m_2m_3 + m_3m_4 - m_2m_4)} \pmod{p}.$$

Next, we claim that $\mathbf{m}$ belongs to at most two of the sets $\mathcal{X}_u$ ($u = 1, \dots, p$). To see this, first observe that, along similar lines as the deduction of (5.12) from (5.11), we have

$$\begin{aligned} & (m_1 + u)(m_3 + u)(m_2 + m_4 + 2u) \\ & \equiv (m_1 + m_3 + 2u)(m_2 + u)(m_4 + u) \pmod{p}. \end{aligned}$$

Recalling that $\mathbf{m}$ was assumed to be non-trivial, the difference of the left and right hand side of the above congruence is seen to be a *non-zero* polynomial of degree at most two in u . Since $\mathbb{Z}/p\mathbb{Z}$ is an integral domain, the claim follows. From this, and taking the trivial solutions into account, we conclude that

$$\sum_{u < p} \# \mathcal{X}_u \leq p(2p^2 + 4(p-1)^2) \ll p^3. \quad \square$$

Proof of Proposition 5.2. First recall (2.4) and observe that

$$S(x, y, w; p) = S(ax, ay, aw; p) \quad (5.13)$$

for any a indivisible by p . Thus, the average

$$\Sigma = \sum_{1 \leq r, t \leq p} \sum_{1 \leq u < p} |S(r, t, u; p)|^4$$

contains $p - 1$ copies of $|S(x, y, w; p)|^4$. Hence,

$$(p - 1) \sum_{w \in \mathcal{W}} |S(x, y, w; p)|^4 \leq \Sigma. \quad (5.14)$$

To estimate Σ , write $\mathbf{m} = (m_1, m_2, m_3, m_4)$ and start by squaring out each term $|S(r, t, u; p)|^4$ twice to obtain

$$\Sigma = \sum_{1 \leq r, t \leq p} \sum_{1 \leq u, m_1, m_2, m_3, m_4 < p} \sum_{p \nmid (m_1+u)} \sum_{p \nmid (m_2+u)} \sum_{p \nmid (m_3+u)} \sum_{p \nmid (m_4+u)} e_p(rA(\mathbf{m}, 0) + tA(\mathbf{m}, u)),$$

where $A(\mathbf{m}, u)$ is given by (5.8).

After moving the summation over r and t to the right and observing that

$$\sum_{1 \leq r \leq p} e_p(rx) = \begin{cases} p & \text{if } p \mid x, \\ 0 & \text{if } p \nmid x, \end{cases}$$

we find that

$$\Sigma = p^2 \sum_{1 \leq u < p} \# \mathcal{X}_u, \quad (5.15)$$

where $\mathcal{X}_u$ is given by (5.9). By Lemma 5.3, $\Sigma \ll p^5$, so that from (5.14) we infer (5.7). $\square$

Remark 5.4. It should be stressed that in the proof of Proposition 5.2 we do not obtain a non-trivial bound for an individual sum $S(x, y, w; p)$; it is only the averaging over $w \in \mathcal{W}$ which makes this result non-trivial. This may be compared with Kloosterman [21], who obtains

$$|S(x, y; p)| \ll p^{3/4} (x, y, p)^{1/4}, \quad (5.16)$$

(recall (2.2)) by exploiting the transformation property

$$S(x, y; p) = S(ax, \bar{a}y; p) \quad (\text{assuming } p \nmid a)$$

in place of (5.13). Notice that only two parameters are transformed, so that when considering the obvious analogue Σ_0 of Σ , one only has to average over two parameters:

$$\Sigma_0 = \sum_{1 \leq r, t \leq p} |S(r, t; p)|^4.$$

Similarly as above one finds that

$$(p-1)|S(x, y; p)|^4 \leq \Sigma_0 = p^2 \#\{\text{solutions } \mathbf{m} \in (\mathbb{Z}/p\mathbb{Z})^4 \text{ to two congruences}\} \quad (5.17)$$

(see [21, pp. 423–236], or [17] for a shorter exposition). Heuristically, one might expect that each congruence condition restricts one of the four variables, so the number of solutions counted on the right hand of (5.17) ought to be $\ll p^2$. Indeed, this can be justified using arguments similar to those used in the proof of Lemma 5.3 (again, see [21, 17]) and from (5.17) one then concludes (5.16). We note that in contrast to the above heuristic, we are unable to show $\#\mathcal{X}_u \ll p^2$ for $\mathcal{X}_u$ given in (5.9), and numerical experiments seem to suggest that this is actually false; Lemma 5.3 only verifies our heuristic on average.

5.3. Proof of Theorem 4.1

We shall need the following result which is an easy consequence of the pigeonhole principle:

Lemma 5.5 ([3, Lemma 3.3]). *Let α be a fixed irrational number. Then, for every positive integer K and real number $\Delta \in (0, 1]$, there exists a real number γ such that*

$$\#\{k \leq K : \{k\alpha + \gamma\} < \Delta\} \geq 0.5 K \Delta.$$

To give a proof of Theorem 4.1, we adapt the reasoning from [3]. Let $K \leq N$ be a positive integer and $\Delta \in (0, 1]$ to be determined later (see (5.22) below). Then, by Lemma 5.5, there is some real number γ such that the set

$$\mathcal{K} = \{k \leq K : \{k\alpha + \gamma\} < \Delta\}$$

has cardinality

$$\#\mathcal{K} \geq 0.5 K \Delta. \quad (5.18)$$

Furthermore, let

$$\mathcal{N} = \{1 \leq n \leq N : \{n\alpha + \beta - \gamma\} < 1 - \Delta\}$$

and $\mathcal{N}^c = \{1, \dots, N\} \setminus \mathcal{N}$. Clearly, recalling (3.1),

$$\#\mathcal{N}^c = N\Delta + O(ND_{\alpha,\beta}(N)). \quad (5.19)$$

Now, writing

$$x_{n,k} = e_p(x[(n+k)\alpha + \beta] + y[\overline{(n+k)\alpha + \beta}])$$

for the moment, for every $k \in \mathcal{K}$, we find that

$$\begin{aligned} K_{\alpha,\beta}(x, y; p, N) &= \sum_{\substack{1 \leq n \leq N \\ p \nmid \lfloor (n+k)\alpha + \beta \rfloor}} x_{n,k} + O(K) \\ &= \sum_{\substack{n \in \mathcal{N} \\ p \nmid \lfloor (n+k)\alpha + \beta \rfloor}} x_{n,k} + O(K + \#\mathcal{N}^c). \end{aligned}$$

Consequently,

$$K_{\alpha,\beta}(x, y; p, N) = \frac{W}{\#\mathcal{K}} + O(K + \#\mathcal{N}^c), \quad (5.20)$$

where

$$W = \sum_{\substack{n \in \mathcal{N} \\ p \nmid \lfloor (n+k)\alpha + \beta \rfloor}} \sum_{k \in \mathcal{K}} e_p(x[(n+k)\alpha + \beta] + y[\overline{(n+k)\alpha + \beta}]).$$

For any $(n, k) \in \mathcal{N} \times \mathcal{K}$, a simple calculation shows that

$$\lfloor (n+k)\alpha + \beta \rfloor = \lfloor n\alpha + \beta - \gamma \rfloor + \lfloor k\alpha + \gamma \rfloor.$$

Next, we apply Cauchy's inequality to W , getting

$$\begin{aligned} |W|^2 &\leq \#\mathcal{N} \cdot \sum_{n \in \mathcal{N}} \left| \sum_{\substack{k \in \mathcal{K} \\ p \nmid \lfloor (n+k)\alpha + \beta \rfloor}} e_p(x[(n+k)\alpha + \beta] + y[\overline{(n+k)\alpha + \beta}]) \right|^2 \\ &\ll_{\alpha} N \cdot \sum_{1 \leq s \leq p} \left| \sum_{\substack{k \in \mathcal{K} \\ p \nmid (s + \lfloor k\alpha + \gamma \rfloor)}} e_p(x(s + \lfloor k\alpha + \gamma \rfloor) + y[\overline{s + \lfloor k\alpha + \gamma \rfloor}]) \right|^2, \end{aligned}$$

where replacing the summation over $n \in \mathcal{N}$ by the summation over $s \leq p$ is allowed, since our assumption $N \leq p$ ensures that

$$\#\{n \in \mathcal{N} : \lfloor n\alpha + \beta - \gamma \rfloor \equiv s \pmod{p}\} < 1 + \alpha \ll_{\alpha} 1.$$

On squaring out the inner sum, we find that

$$|W|^2 \ll_{\alpha} N \cdot \sum_{\substack{1 \leq s \leq p, k, \ell \in \mathcal{K} \\ p \nmid (s + \lfloor k\alpha + \gamma \rfloor) \\ p \nmid (s + \lfloor \ell\alpha + \gamma \rfloor)}} \sum_{k \in \mathcal{K}} \sum_{\ell \in \mathcal{K}} e_p(x(\lfloor k\alpha + \gamma \rfloor + \lfloor \ell\alpha + \gamma \rfloor)) \times \\ \times e_p(y(\overline{s + \lfloor k\alpha + \gamma \rfloor} - \overline{s + \lfloor \ell\alpha + \gamma \rfloor})).$$

This is

$$\ll_{\alpha} N \cdot \sum_{k \in \mathcal{K}} \sum_{\ell \in \mathcal{K}} \left| \sum_{\substack{1 \leq s \leq p \\ p \nmid (s + \lfloor k\alpha + \gamma \rfloor) \\ p \nmid (s + \lfloor \ell\alpha + \gamma \rfloor)}} e_p(y(\overline{s + \lfloor k\alpha + \gamma \rfloor} - \overline{s + \lfloor \ell\alpha + \gamma \rfloor})) \right|.$$

Upon writing

$$\mathcal{W}_k = \{\lfloor \ell\alpha + \gamma \rfloor - \lfloor k\alpha + \gamma \rfloor : \ell \in \mathcal{K}\},$$

we infer

$$|W|^2 \ll_{\alpha} N \sum_{k \in \mathcal{K}} \sum_{w \in \mathcal{W}_k} \left| \sum_{\substack{s \leq p \\ p \nmid (s+w)}} e_p(y(\overline{s} - \overline{s+w})) \right|.$$

Hence, by Proposition 4.2 and using $\#\mathcal{K} = \#\mathcal{W}_k$ ($k \in \mathcal{K}$),

$$|W|^2 \ll_{\alpha} N(\#\mathcal{K})^{7/4} p. \quad (5.21)$$

In view of (5.20), we find that

$$|K_{\alpha, \beta}(x, y; p, N)| \ll_{\alpha} \sqrt{\frac{Np}{(\#\mathcal{K})^{1/4}}} + K + \#\mathcal{N}^c.$$

Consequently, upon gathering (5.18), (5.19) and Lemma 3.1, we obtain the bound

$$|K_{\alpha, \beta}(x, y; p, N)| \ll_{\alpha} (Np)^{1/2} (K\Delta)^{-1/8} + K + N\Delta + ND_{\alpha}(N).$$

We now choose Δ and K . Specifically, we aim at obtaining a bound of the shape $|K_{\alpha, \beta}(x, y; p, N)| = o_{\alpha}(N)$ in a range $N \leq p < N^{\theta}$ with $\theta > 1$ as large as possible. It can be seen that

$$\Delta = N^{-105/128} p^{21/32} \quad \text{and} \quad K = \lceil N\Delta \rceil \quad (5.22)$$

yield the assertion of the theorem provided that $\Delta < 1$, i.e., when

$$(\log p) / \log N < \frac{5}{4}. \quad (5.23)$$

On the other hand, if (5.23) fails to hold, then the theorem asserts nothing more than the trivial bound $|K_{\alpha, \beta}(x, y; p, N)| \ll N$, so the proof is complete. $\square$

6. Improvements and open problems

6.1. Improvements

As noted earlier in Remark 2.2, I. E. Shparlinski pointed out to the author that a generalised form of Weil's bound also applies to the sum $S(x, y, w; p)$ from (2.4): indeed, by applying [25, Theorem 2] to the rational function

$$R(m) = \frac{x}{m} + \frac{y}{m+w} = \frac{(x+y)m + xw}{m(m+w)} \in (\mathbb{Z}/p\mathbb{Z})[m],$$

it is straight-forward to see that

$$S(x, y, w; p) \leq 2\sqrt{p} + 1 \quad (6.1)$$

whenever $p \nmid x+y$ or $p \nmid xw$, for then $R(m)$ is (after cancelling any potential common factors in the above) a non-zero quotient with denominator of degree at most two and numerator of smaller degree. The bound (6.1) should be considered as an improved version of Proposition 4.2.

Consequently, we have a chain of improvements: in the proof of Theorem 4.1, where we have used Proposition 4.2, we instead get

$$|W|^2 \ll_{\alpha} N(\#\mathcal{K})(p + (\#\mathcal{K})\sqrt{p}) \quad (5.21')$$

in place of (5.21). Then, the choices in (5.22) can then be adapted to

$$\Delta = N^{-1/4}p^{1/8} \quad \text{and} \quad K = \lceil N^{1/4}p^{3/8} \rceil, \quad (5.22')$$

provided that we assume that $N \leq p < N^2$. (Recall that the method requires $N \leq p$ and $\Delta < 1$.) This yields

$$|K_{\alpha,\beta}(x, y; p, N)| \ll_{\alpha} N^{3/4}p^{1/8} + ND_{\alpha}(N) \quad (4.2')$$

in place of (4.2).

Similar to our deduction of Corollary 4.3, we may note that Lemma 3.3 then implies the bound

$$|K_{\alpha,\beta}(x, y; p, N)| \ll_{\alpha} N^{3/4}p^{1/8} \quad (4.3')$$

in place of (4.3), provided that α is of finite type < 8 .

Consequently, we have the following improved version of (5.6):

$$(\log p)N^{-2+3/4}p^{1+1/8} \ll_{\alpha,\epsilon} 1. \quad (5.6')$$

Together with the choice $N = \lfloor p^{9/10} \log p \rfloor$ this yields the following improved bound in Theorem 2.1:

$$\max\{m, \tilde{m}\} \leq \alpha p^{9/10} \log p + \beta, \quad (2.1')$$

and the assumption on the type of α can be relaxed to allow for all finite types strictly smaller than 8.

6.2. Scope and open problems

Apart from the obvious advantage of yielding superior results, the above approach via Weil's bound has the disadvantage of failing to generalise to composite moduli in an obvious manner. On the other hand, when studying Question A with p replaced by some composite number, we expect that our approach via Proposition 4.2 should still be capable of yielding non-trivial results. A suitable generalisation of Lemma 5.3 with composite p should be the main ingredient here.

Another direction in which our results may be generalised goes as follows: note that, as remarked earlier in Section 4, by Lemma 3.2, Theorem 4.1 is non-trivial for every irrational $\alpha > 1$ provided only that N and p are large and in a suitable range. In spite of this, we obtain a non-trivial answer to Question A (in the form of Theorem 2.1) only for α with some restriction on its type. The reason for this defect can be traced back to the $\log p$ factor appearing in (5.5), which in turn arose from the application of Lemma 5.1. It seems plausible that one should be able to remove this by considering a suitable analogue of Lemma 5.1 where the sharp cut-off at a and b in the sum

$$\sum_{a < m \leq b} A_m$$

is smoothed out. Specifically, working instead with

$$\sum_{1 \leq m \leq q} A_m \chi(m),$$

where $\chi = \chi_0 * \chi_0$ is given as the convolution of the characteristic function χ_0 of an interval with itself, seems promising.

Acknowledgments

Most of this work is part of the author's doctoral dissertation at Würzburg University. The author gratefully acknowledges the encouragement of his advisor, Jörn Steuding. The work on this article was completed while the author was employed at Graz University of Technology. The financial support by both Würzburg University and Graz University of Technology is highly appreciated. The author would like to thank Christoph Aistleitner

for comments on an earlier draft. Furthermore, the valuable criticism of the anonymous referee has lead to various improvements on the exposition and is also greatly appreciated.

References

- [1] A.G. Abercrombie, Beatty sequences and multiplicative number theory, *Acta Arith.* 70 (3) (1995) 195–207.
- [2] A.G. Abercrombie, W.D. Banks, I.E. Shparlinski, Arithmetic functions on Beatty sequences, *Acta Arith.* 136 (1) (2009) 81–89.
- [3] W.D. Banks, I.E. Shparlinski, Non-residues and primitive roots in Beatty sequences, *Bull. Aust. Math. Soc.* 73 (3) (2006) 433–443.
- [4] W.D. Banks, I.E. Shparlinski, Short character sums with Beatty sequences, *Math. Res. Lett.* 13 (4) (2006) 539–547.
- [5] W.D. Banks, I.E. Shparlinski, Prime divisors in Beatty sequences, *J. Number Theory* 123 (2) (2007) 413–425.
- [6] W.D. Banks, I.E. Shparlinski, Prime numbers with Beatty sequences, *Colloq. Math.* 115 (2) (2009) 147–157.
- [7] W.D. Banks, M.Z. Garaev, D.R. Heath-Brown, I.E. Shparlinski, Density of non-residues in Burgess-type intervals and applications, *Bull. Lond. Math. Soc.* 40 (1) (2008) 88–96.
- [8] S. Beatty, Problem 3173, *Amer. Math. Monthly* 33 (1926) 159.
- [9] S. Beatty, A. Ostrowski, J. Hyslop, A.C. Aitken, Solutions to problem 3173, *Amer. Math. Monthly* 34 (1927) 159–160.
- [10] A.V. Begunts, On an analogue of the Dirichlet divisor problem, *Vestnik Moskov. Univ. Ser. I* 2004 (6) (2004) 52–56.
- [11] A.V. Begunts, On the distribution of the values of sums of multiplicative functions on generalized arithmetic progressions, *Chebyshevskii Sb.* 6 (2(14)) (2005) 52–74.
- [12] J. Bernoulli III, Sur une nouvelle espece de calcul, in: *Recueil pour les Astronomes*, vol. I, Berlin, 1772, pp. 255–284.
- [13] A.S. Besicovitch, Sets of fractional dimensions. IV: on rational approximation to real numbers, *J. Lond. Math. Soc.* 9 (1934) 126–131.
- [14] E.B. Christoffel, *Observatio arithmetica*, *Ann. of Math.* (2) 6 (1873) 148–153.
- [15] E.B. Christoffel, *Lehrsätze über arithmetische Eigenschaften der Irrationalzahlen*, *Ann. of Math.* (2) 15 (1887) 253–276.
- [16] A.M. Güloğlu, C.W. Nevans, Sums of multiplicative functions over a Beatty sequence, *Bull. Aust. Math. Soc.* 78 (2) (2008) 327–334.
- [17] D.R. Heath-Brown, Arithmetic applications of Kloosterman sums, *Nieuw Arch. Wiskd.* (5) 1 (4) (2000) 380–384.
- [18] H. Iwaniec, E. Kowalski, *Analytic Number Theory*, American Mathematical Society (AMS), Providence, RI, 2004.
- [19] V. Jarník, Diophantische Approximationen und Hausdorffsches Maß, *Mat. Sb.* 36 (1929) 371–382.
- [20] A. Khintchine, Einige Sätze über Kettenbrüche, mit Anwendungen auf die Theorie der Diophantischen Approximationen, *Math. Ann.* 92 (1924) 115–125.
- [21] H.D. Kloosterman, On the representation of numbers in the form $ax^2 + by^2 + cz^2 + dt^2$, *Acta Math.* 49 (1927) 407–464.
- [22] L. Kuipers, H. Niederreiter, *Uniform Distribution of Sequences*, John Wiley & Sons, New York etc., 1974.
- [23] G. Lü, W. Zhai, The divisor problem for the Beatty sequences, *Acta Math. Sinica (Chin. Ser.)* 47 (6) (2004) 1213–1216.
- [24] H.L. Montgomery, *Ten Lectures on the Interface between Analytic Number Theory and Harmonic Analysis*, American Mathematical Society, Providence, RI, 1994.
- [25] C.J. Moreno, O. Moreno, Exponential sums and Goppa codes. I, *Proc. Amer. Math. Soc.* 111 (2) (1991) 523–531.
- [26] P. Ribenboim, *The New Book of Prime Number Records*, 3rd edition, Springer, New York, 1996.
- [27] K.F. Roth, Rational approximations to algebraic numbers, *Mathematika* 2 (1955) 1–20.
- [28] I.E. Shparlinski, Modular hyperbolas, *Jpn. J. Math.* (3) 7 (2) (2012) 235–294.
- [29] J.W. Strutt, *The Theory of Sound*, second edition, Macmillan, London, 1926.

- [30] I.M. Vinogradov, *The Method of Trigonometrical Sums in the Theory of Numbers*, translated from the Russian, revised and annotated by K.F. Roth and Anne Davenport. Reprint of the 1954 translation, Dover Publications, New York, 2004.
- [31] A. Weil, On some exponential sums, *Proc. Natl. Acad. Sci. USA* 34 (1948) 204–207.
- [32] H. Weyl, Über die Gleichverteilung von Zahlen mod. Eins, *Math. Ann.* 77 (1916) 313–352.
- [33] W. Zhai, A note on a result of Abercrombie, *Kexue Tongbao (Chinese)* 42 (8) (1997) 804–806.