Metric discrepancy theory, functions of bounded variation and GCD sums

Christoph Aistleitner[*], Philipp A. Mayer[†], Volker Ziegler[‡]

Abstract

Let $f(x)$ be a 1-periodic function of bounded variation having mean zero, and let $(n_k)_{k\geq1}$ be an increasing sequence of positive integers. Then a result of Baker implies the upper bound $\left|\sum_{k=1}^N f(n_kx)\right| = \mathcal{O}\left(\sqrt{N}(\log N)^{3/2+\varepsilon}\right)$ for almost all $x \in (0,1)$ in the sense of the Lebesgue measure. We show that the asymptotic order of $\left|\sum_{k=1}^N f(n_kx)\right|$ is closely connected with certain number-theoretic properties of the sequence $(n_k)_{k\geq1}$, namely a certain function involving the greatest common divisor function. More exactly, we give an upper bound for the asymptotic order of $\left|\sum_{k=1}^N f(n_kx)\right|$ in terms of the function

$$h_N(n_1,\dots,n_N) = \sum_{1\leq k_1,k_2\leq N} \frac{\gcd(n_{k_1},n_{k_2})}{\max(n_{k_1},n_{k_2})}.$$

1 Introduction

Given a sequence of real numbers $(x_k)_{k\geq1}$, the value

$$D_N = D_N(x_1,\dots,x_N) = \sup_{0\leq a<b<1} \left| \frac{\sum_{k=1}^N \mathbb{1}_{[a,b)}(\langle x_k\rangle)}{N} - (b-a) \right|,$$

is called the *discrepancy* of the first N elements of $(x_k)_{k\geq1}$. Here $\mathbb{1}_{[a,b)}$ is the indicator function of the interval $[a,b)$ and $\langle x\rangle$ denotes the fractional part of x. A sequence $(x_k)_{k\geq1}$ is called *uniformly distributed modulo 1* (u.d. mod 1) if and only if

$$D_N \to 0 \quad \text{as} \quad N \to \infty.$$

[*]Graz University of Technology, Institute of Mathematics A, Steyrergasse 30, 8010 Graz, Austria. e-mail: aistleitner@finanz.math.tugraz.at. Research supported by the Austrian Research Foundation (FWF), Project S9603-N23. This paper was written during a stay at the Alfréd Rényi Institute of Mathematics of the Hungarian Academy of Sciences, which was made possible by an MOEL scholarship of the *Österreichische Forschungsgemeinschaft* (ÖFG).

[†]Graz University of Technology, Institute of Mathematics A, Steyrergasse 30, 8010 Graz, Austria. e-mail: mayer@finanz.math.tugraz.at.

[‡]Graz University of Technology, Institute of Mathematics A, Steyrergasse 30, 8010 Graz, Austria. e-mail: ziegler@finanz.math.tugraz.at. Research supported by the Austrian Research Foundation (FWF), Project S9603-N23.

Mathematics Subject Classification: Primary 26A45,11K38, 60F15, 10A05

Keywords: functions of bounded variation, discrepancy, gcd sums

By a classical theorem of Weyl [18] (see also Kuipers and Niederreiter [14, p. 32]) for any sequence $(n_k)_{k\geq 1}$ of distinct positive integers

$$D_N(n_k x) \to 0 \qquad \text{a.e.},$$

which means that for almost all $x \in (0,1)$ (with respect to the Lebesgue measure) the sequence $(n_k x)_{k\geq 1}$ is u.d. mod 1 (for abbreviation we write $D_N(n_k x)$ instead of $D_N(n_1 x, \dots, n_N x)$). The exact asymptotic order of $D_N(n_k x)$ is only known in very few special cases: in the case $n_k = k$, where Kesten [12] proved

$$\frac{N D_N(x)}{\log N \log\log N} \to \frac{2}{\pi^2} \qquad \text{in measure},$$

and in the case of *lacunary* $(n_k)_{k\geq 1}$, i.e. $(n_k)_{k\geq 1}$ satisfying the *Hadamard gap condition*

$$\frac{n_{k+1}}{n_k} > q > 1, \quad k \geq 1,$$

where Philipp [17] proved the law of the iterated logarithm

$$\frac{1}{4} \leq \frac{N D_N(n_k x)}{\sqrt{2N \log\log N}} \leq C_q \quad \text{a.e.}$$

(for precise results see Aistleitner [1] and Fukuyama [9]).

In the general case of a sequence $(n_k)_{k\geq 1}$ of distinct positive integers Cassels [5] and Erdős and Koksma [8] proved independently that for any $\varepsilon > 0$

$$N D_N(n_k x) = \mathcal{O}\left(\sqrt{N} (\log N)^{5/2+\varepsilon}\right) \quad \text{a.e.}$$

(see also Drmota and Tichy [6, p. 154-159]). For strictly increasing $(n_k)_{k\geq 1}$ this was improved by Baker [3] to

$$N D_N(n_k x) = \mathcal{O}\left(\sqrt{N} (\log N)^{3/2+\varepsilon}\right) \quad \text{a.e.}, \tag{1}$$

which is the best upper bound known to date.

Now, there is a close connection between discrepancy discussed above and sums of functions of bounded variation, which is revealed by Koksma's inequality: For a function f, satisfying $f(x+1) = f(x)$ (we say f is *1-periodic*) and having total variation $\mathrm{Var}_{[0,1]} f$ in the unit interval, and a sequence $(x_k)_{k\geq 1}$ of reals from the unit interval, Koksma's inequality states that

$$\left| \int_0^1 f(x)\, dx - \frac{1}{N} \sum_{k=1}^N f(x_k) \right| \leq D_N(x_k) \,\mathrm{Var}_{[0,1]} f, \qquad N \geq 1.$$

Thus Baker's result implies that for any 1-periodic function f of bounded variation, satisfying

$$\int_0^1 f(x)\, dx = 0,$$

and for any increasing sequence $(n_k)_{k\geq 1}$ of positive integers we have (for any $\varepsilon > 0$)

$$\left| \sum_{k=1}^N f(n_k x) \right| = \mathcal{O}\left(\sqrt{N} (\log N)^{3/2+\varepsilon}\right) \quad \text{a.e.}$$

On the other hand, Berkes and Philipp [4] constructed an increasing sequence $(\bar{n}_k)_{k\geq 1}$ of integers, for which

$$\limsup_{N\to\infty} \frac{\left|\sum_{k=1}^N \cos 2\pi \bar{n}_k x\right|}{\sqrt{N \log N}} = \infty \quad \text{a.e.},$$

which implies (again by means of Koksma's inequality)

$$\limsup_{N\to\infty} \frac{|ND_N(\bar{n}_k x)|}{\sqrt{N \log N}} = \infty \quad \text{a.e.}$$

We can summarize these results as follows:

- For any increasing sequence $(n_k)_{k\geq 1}$ of positive integers,

  $$ND_N(n_k x) = \mathcal{O}\left(\sqrt{N} (\log N)^{3/2+\varepsilon}\right) \quad \text{a.e.}$$

  and the exponent of the logarithmic term can in general not be reduced below 1/2.

- For any increasing sequence $(n_k)_{k\geq 1}$ of positive integers, and any 1-periodic function f of bounded variation, satisfying $\int_0^1 f(x)\ dx = 0$,

  $$\left|\sum_{k=1}^N f(n_k x)\right| = \mathcal{O}\left(\sqrt{N} (\log N)^{3/2+\varepsilon}\right) \quad \text{a.e.},$$

  and the exponent of the logarithmic term can in general not be reduced below 1/2.

The purpose of this paper is to show that there exists a strong connection between the asymptotic order of sums of the form

$$\sum_{k=1}^N f(n_k x) \tag{2}$$

and number-theoretic properties of the sequence $(n_k)_{k\geq 1}$. More exactly, we will show that the asymptotic order of (2) can be estimated in terms of the value of $h_N = h_N(n_1, \ldots, n_N)$, which is defined by

$$h_N(n_1, \ldots, n_N) = \sum_{1\leq k_1, k_2 \leq N} \frac{\gcd(n_{k_1}, n_{k_2})}{\max(n_{k_1}, n_{k_2})}, \qquad N \geq 1.$$

Observe that trivially always $h_N \geq N$, since

$$h_N(n_1, \ldots, n_N) \geq \sum_{1\leq k\leq N} \frac{\gcd(n_k, n_k)}{\max(n_k, n_k)} = \sum_{1\leq k\leq N} 1 = N, \qquad N \geq 1.$$

Our main result is the following:

Theorem 1 *Let $f(x)$ be a function satisfying*

$$f(x+1) = f(x), \qquad \int_0^1 f(x)\ dx = 0, \qquad \mathrm{Var}_{[0,1]} f < \infty, \tag{3}$$

and let $(n_k)_{k\geq 1}$ be a strictly increasing sequence of positive integers. Let $g(N)$ be a nondecreasing positive function satisfying $\sup_{N\geq 1} g(2N)/g(N) < \infty$. Then for almost every $x \in (0,1)$ we have

$$\left| \sum_{k=1}^{N} f(n_k x) \right| = \mathcal{O}\left(\sqrt{N} g(N)\right), \tag{4}$$

provided

$$\sum_{N\geq 1} \frac{\left(\log\left(\frac{(\log N) h_N(n_1,\dots,n_N)}{N}\right)\right)^2}{N g(N)^2} < \infty.$$

The result summed up in Theorem 1 obviously improves the better the asymptotic for h_N can be calculated. A related problem, namely to find the best possible bound for

$$\sum_{1\leq k_1,k_2\leq N} \frac{(\gcd(n_{k_1}, n_{k_2}))^2}{n_{k_1} n_{k_2}}. \tag{5}$$

was solved by Gál [10]. He showed that for any sequence of distinct positive integers $(n_1,\dots,n_N)$ the value of (5) is bounded by

$$cN(\log\log N)^2$$

for some absolute constant c, and that this upper bound is best possible. By the way, Koksma [13] was the first who observed that this implies

$$\int_0^1 \left(\sum_{k=1}^{N} f(n_k x)\right)^2 dx \leq cN(\log\log N)^2 \tag{6}$$

for some constant c for any function f satisfying (3).

Finding a precise general upper bound for h_N, i.e. a function $H(N)$ such that

$$h_N(n_1,\dots,n_N) \leq H(N)$$

for any sequence $(n_1,\dots,n_N)$ of distinct integers, appears to be much more difficult than finding an upper bound for (5). Defining a function $\tilde{h}_N$ by

$$\tilde{h}_N(n_1,\dots,n_N) = \sum_{1\leq k_1,k_2\leq N} \frac{\gcd(n_{k_1}, n_{k_2})}{\sqrt{n_{k_1} n_{k_2}}},$$

it is easy to see that always $h_N \leq \tilde{h}_N$. The sum in $\tilde{h}_N$ seems to be more tractable than the one in h_N, but it is still very difficult to handle it. The first upper bound for the function

$\tilde{h}_N$ (which is still the best known upper bound to date) was given by Dyer and Harman [7] in 1986, who proved, for non-negative sequences $(u_k)_{k\geq 1}$ and $(v_k)_{k\geq 1}$ of reals,

$$\begin{aligned} \sum_{1\leq k_1,k_2\leq N} \frac{\sqrt{u_{k_1}v_{k_2}}\gcd(n_{k_1},n_{k_2})}{\sqrt{n_{k_1}n_{k_2}}} &\ll \left(\sum_{1\leq k_1,k_2\leq N} u_{k_1}v_{k_2} e^{\frac{c\log(k_1k_2)}{\log\log(k_1k_2)}} \right)^{1/2} \\ &\ll \left(\sum_{k=1}^N u_k e^{\frac{c\log k}{\log\log k}} \right)^{1/2} \left(\sum_{k=1}^N v_k e^{\frac{c\log k}{\log\log k}} \right)^{1/2}. \end{aligned} \tag{7}$$

Applied to our situation, where $u_k = 1, v_k = 1, k \geq 1$, this gives

$$\sum_{1\leq k_1,k_2\leq N} \frac{\gcd(n_{k_1},n_{k_2})}{\sqrt{n_{k_1}n_{k_2}}} \ll N e^{\frac{c\log N}{\log\log N}}$$

Dyer and Harman also showed that the factor

$$e^{\frac{c\log N}{\log\log N}}$$

in (7) can not be replaced by a factor smaller than

$$e^{\frac{c\sqrt{\log N}}{\log\log N}} \tag{8}$$

and in his book [11, p. 62] Harman writes that *it is tempting to conjecture that the factor in (8) gives the "correct" maximal order of magnitude.*

Using the result of Dyer and Harman we see that the factor

$$\left(\log\left(\frac{(\log N)h_N(n_1,\dots,n_N)}{N}\right)\right)^2 \tag{9}$$

in our theorem is at most $\left(\frac{c\log N}{\log\log N}\right)^2$. Thus we get the following corollary to Theorem 1 :

Corollary 1 *Let $f(x)$ be a function satisfying (3) and let $(n_k)_{k\geq 1}$ be a strictly increasing sequence of positive integers. Let $g(N)$ be a nondecreasing positive function satisfying* $\sup_{N\geq 1} g(2N)/g(N) < \infty$. *Then for almost every $x \in (0,1)$*

$$\left| \sum_{k=1}^N f(n_k x) \right| = \mathcal{O}\left(\sqrt{N}g(N)\right),$$

provided

$$\sum_{N\geq 1} \frac{(\log N)^2}{Ng(N)^2(\log\log N)^2} < \infty.$$

Baker's result (1) is based on the estimate

$$\left(\int_0^1 \left(\max_{1\leq M\leq N} MD_M(n_k x) \right)^2 dx \right)^{1/2} \ll \sqrt{N}(\log N),$$

or, formulated for $f(n_k x)$, on the estimate

$$\left(\int_0^1 \left(\max_{1\le M\le N}\left|\sum_{k=1}^{M} f(n_k x)\right|\right)^2 dx\right)^{1/2} \ll \sqrt{N}(\log N),$$

which yields

$$\left|\sum_{k=1}^{N} f(n_k x)\right| = \mathcal{O}\left(\sqrt{N} g(N)\right) \quad \text{a.e.}$$

provided

$$\sum_{N\ge 1} \frac{(\log N)^2}{N g(N)^2} < \infty.$$

In particular this implies

$$\left|\sum_{k=1}^{N} f(n_k x)\right| = \mathcal{O}\left(\sqrt{N}(\log N)^{3/2}(\log\log N)^{1/2+\varepsilon}\right) \quad \text{a.e.}$$

for $\varepsilon > 0$. The proof of our theorem contains the estimate

$$\left(\int_0^1 \left(\max_{1\le M\le N}\left|\sum_{k=1}^{M} f(n_k x)\right|\right)^2 dx\right)^{1/2} \ll \sqrt{N}(\log N)(\log\log N)^{-1}$$

which implies the following (use $g(N) = ((\log N)^{3/2}(\log\log N)^{-1/2+\varepsilon}$ in Corollary 1)

Corollary 2 *Let $f(x)$ be a function satisfying (3) and let $(n_k)_{k\ge 1}$ be a strictly increasing sequence of positive integers. Then for almost every $x \in (0,1)$*

$$\left|\sum_{k=1}^{N} f(n_k x)\right| = \mathcal{O}\left(\sqrt{N}(\log N)^{3/2}(\log\log N)^{-1/2+\varepsilon}\right) \tag{10}$$

This means that compared to Baker's result our theorem gives a minimal gain of a factor $(\log\log N)$ in the general case. On the other hand, if the conjecture of Dyer and Harman is true and in fact

$$h_N(n_1,\dots,n_N) \ll N e^{\frac{c\sqrt{\log N}}{\log\log N}} \tag{11}$$

for any sequence $(n_1,\dots,n_N)$ and any $N \ge 1$, then our method would give

$$\left(\int_0^1 \left(\max_{1\le M\le N}\left|\sum_{k=1}^{M} f(n_k x)\right|\right)^2 dx\right)^{1/2} \ll \sqrt{N}\sqrt{\log N}(\log\log N)^{-1}$$

and

$$\left|\sum_{k=1}^{N} f(n_k x)\right| = \mathcal{O}\left(\sqrt{N}(\log N)(\log\log N)^{-1/2+\varepsilon}\right) \quad \text{a.e.} \tag{12}$$

This is obviously better than (10) (the difference is a factor $\sqrt{\log N}$), but still does not answer the question of the best possible exponent for the logarithmic term in such estimates. It is possible that a direct estimate for h_N (instead of an estimate for $\tilde{h}_N$) could yield an even stronger upper bound than (11), and may result in an further improvement of (12).

We strongly believe that (11) is true, but despite considerable efforts we have not been able to prove it.

The condition $\sup_{N \geq 1} g(2N)/g(N) < \infty$ in the statement of the theorem and the corollary is a technical condition which is used in the proof. The considerations above show that typical examples for "interesting" values for $g(N)$ are in the region of $g(N) = \log N$ or $g(N) = (\log N)^{3/2}$. Thus it is no real restriction to require $\sup_{N \geq 1} g(2N)/g(N) < \infty$, which is satisfied for slowly growing functions $g(N)$. The factor $(\log N)$ in (9) is only relevant for "small" values of h_N, i.e. if h_N is close to N. If $h_N > N \log N$ the additional factor $\log N$ in (9) is negligible.

An interesting problem that we did not touch so far is to which extent results like (1) or (4) remain valid if the condition that $(n_k)_{k \geq 1}$ is increasing is dropped (we shall assume that the n_k's are distinct instead), or how large the best possible exponent for the logarithmic term is in this case. It is somehow natural to expect that the best possible upper bound will be the same for increasing $(n_k)_{k \geq 1}$ as in the general case of not necessarily increasing $(n_k)_{k \geq 1}$. However, for the proof of Baker's results as well as for the proof of our Theorem 1 the Carleson-Hunt inequality is a crucial tool and in this last inequality the condition of increasing n_k's is essential. As far as we know, in the case of not necessarily increasing $(n_k)_{k \geq 1}$ the best upper bound for $D_N(n_k x)$ known to date is

$$N D_N(n_k x) = \mathcal{O}\left(\sqrt{N} (\log N)^{5/2+\varepsilon}\right) \quad \text{a.e.}$$

(cf. [5],[8]), and for $f(n_k x)$ the estimate

$$\left| \sum_{k=1}^{N} f(n_k x) \right| = \mathcal{O}\left(\sqrt{N} (\log N)^{3/2+\varepsilon}\right) \quad \text{a.e.}$$

can be deduced from (6), Lemma 1 below and the Borel-Cantelli lemma.

Some remarks on the notation: Throughout this paper c will stand for appropriate positive numbers, not always the same, which may not depend on N or M (but may depend on f, g and $(n_k)_{k \geq 1}$), $\log x$ will mean the maximum of $\log x$ and 1, and $\mathbb{P}$ denotes the Lebesgue measure on $(0,1)$.

2 Preparations

We will use the following inequality of Móricz, Serfling and Stout [15, Corollary 3.1]:

Lemma 1 *Let $X_1, \ldots, X_N$ be random variables, and let $s(k_1, k_2)$ be a superadditive function, i.e. a function satisfying*

- $s(M_1, M_2) \geq 0, \quad 1 \leq M_1 \leq M_2 \leq N$
- $s(M_1, M_2) \leq s(M_1, M_2 + 1), \quad 1 \leq M_1 \leq M_2 < N$

- $s(M_1, M_2) + s(M_2+1, M_3) \leq s(M_1, M_3), \quad 1 \leq M_1 \leq M_2 < M_3 \leq N.$

Suppose that

$$\mathbb{E}\left(\sum_{k=M_1}^{M_2} X_k\right)^2 \leq s(M_1, M_2), \qquad 1 \leq M_1 \leq M_2 \leq N.$$

Then

$$\mathbb{E}\left(\max_{1 \leq M \leq N}\left(\sum_{k=1}^{M} X_k\right)^2\right) \leq s(1, N)(1 + \log_2 N)^2.$$

The following two lemmas will be used in the proof of Theorem 1.

Lemma 2 *Let $f(x)$ be a function satisfying (3), and let $(n_1, \dots, n_N)$ be a sequence of distinct positive integers. Let $J \geq 1$ and write $r(x)$ for the J-th remainder term of the Fourier series of f, i.e.*

$$f(x) = \sum_{j=1}^{\infty} a_j \cos 2\pi j x + b_j \sin 2\pi j x, \qquad r(x) = \sum_{j=J}^{\infty} a_j \cos 2\pi j x + b_j \sin 2\pi j x.$$

Then

$$\int_0^1 \left(\sum_{k=1}^{N} r(n_k x)\right)^2 dx \leq \frac{h_N(n_1, \dots, n_N)}{J}.$$

Lemma 3 *Let $f(x)$ be a function satisfying (3), and let $(n_1, \dots, n_N)$ be an increasing sequence of distinct positive integers. Then*

$$\int_0^1 \max_{1 \leq M \leq N}\left(\sum_{k=1}^{M} f(n_k x)\right)^2 dx \leq c\left(\log\left(\frac{(\log N) h_N(n_1, \dots, n_N)}{N}\right)\right)^2 N.$$

Proof of Lemma 2: W.l.o.g. we assume that $f(x)$ is an even function, i.e. the Fourier series of f is of the form

$$f(x) \sim \sum_{j=1}^{\infty} a_j \cos 2\pi j x$$

(the proof in the general case is exactly the same). We can also assume w.l.o.g. that $\mathrm{Var}_{[0,1]} f \leq 2$, which implies (see Zygmund [19, p. 48])

$$|a_j| \leq \frac{1}{j}, \quad j \geq 1.$$

Writing

$$r(x) = \sum_{j=J}^{\infty} a_j \cos 2\pi j x,$$

we have, because of the orthogonality of the trigonometric system,

$$\begin{aligned} \int_0^1 \left(\sum_{k=1}^{N} r(n_k x)\right)^2 dx &= \sum_{1 \leq k_1, k_2 \leq N} \sum_{j_1, j_2 \geq J} \left|\frac{a_{j_1} a_{j_2}}{2}\right| \mathbb{1}(j_1 n_{k_1} = j_2 n_{k_2}) \\ &\leq \sum_{1 \leq k_1, k_2 \leq N} \sum_{j_1, j_2 \geq J} \frac{1}{2 j_1 j_2} \mathbb{1}(j_1 n_{k_1} = j_2 n_{k_2}) \end{aligned}$$

For fixed k_1, k_2 (w.l.o.g. $n_{k_1} \leq n_{k_2}$), using

$$\sum_{j \geq J} \frac{1}{j^2} \leq \frac{2}{J},$$

we have

$$\begin{aligned}
\sum_{j_1, j_2 \geq J} \frac{1}{2 j_1 j_2} \ \mathbb{1}(j_1 n_{k_1} = j_2 n_{k_2}) &= \sum_{j \geq \lceil J \gcd(n_{k_1}, n_{k_2})/n_{k_1} \rceil} \frac{(\gcd(n_{k_1}, n_{k_2}))^2}{2 j^2 n_{k_1} n_{k_2}} \\
&\leq \frac{1}{\lceil J \gcd(n_{k_1}, n_{k_2})/n_{k_1} \rceil} \frac{(\gcd(n_{k_1}, n_{k_2}))^2}{n_{k_1} n_{k_2}} \\
&\leq \frac{n_{k_1}}{J \gcd(n_{k_1}, n_{k_2})} \frac{(\gcd(n_{k_1}, n_{k_2}))^2}{n_{k_1} n_{k_2}} \\
&\leq \frac{1}{J} \frac{\gcd(n_{k_1} n_{k_2})}{\max(n_{k_1}, n_{k_2})}.
\end{aligned}$$

Thus

$$\int_0^1 \left(\sum_{k=1}^N r(n_k x) \right)^2 dx \leq \frac{1}{J} \sum_{1 \leq k_1, k_2 \leq N} \frac{\gcd(n_{k_1} n_{k_2})}{\max(n_{k_1}, n_{k_2})} = \frac{h_N(n_1, \dots, n_N)}{J}. \quad \square$$

Proof of Lemma 3: Again w.l.o.g. we may assume that f is even, and we set

$$p(x) = \sum_{j=1}^J a_j \cos 2\pi j x, \qquad r(x) = \sum_{j=J+1}^\infty a_j \cos 2\pi j x,$$

where J will be specified later. By Minkowski's inequality we have

$$\begin{aligned}
&\left(\int_0^1 \max_{1 \leq M \leq N} \left(\sum_{k=1}^M f(n_k x) \right)^2 dx \right)^{1/2} \\
\leq &\left(\int_0^1 \max_{1 \leq M \leq N} \left(\sum_{k=1}^M p(n_k x) \right)^2 dx \right)^{1/2} + \left(\int_0^1 \max_{1 \leq M \leq N} \left(\sum_{k=1}^M r(n_k x) \right)^2 dx \right)^{1/2}. \quad (13)
\end{aligned}$$

Applying Minkowski's inequality again we deduce

$$
\begin{aligned}
&\left(\int_0^1 \max_{1\le M\le N}\left(\sum_{k=1}^M p(n_k x)\right)^2 dx\right)^{1/2}\\
= & \left(\int_0^1 \max_{1\le M\le N}\left(\sum_{k=1}^M \sum_{j=1}^J a_j \cos 2\pi j n_k x\right)^2 dx\right)^{1/2}\\
\le & \sum_{j=1}^J |a_j| \left(\int_0^1 \max_{1\le M\le N}\left(\sum_{k=1}^M \cos 2\pi j n_k x\right)^2 dx\right)^{1/2}\\
\le & \sum_{j=1}^J \frac{1}{j} \left(\int_0^1 \max_{1\le M\le N}\left(\sum_{k=1}^M \cos 2\pi j n_k x\right)^2 dx\right)^{1/2}\\
\le & (1+\log J) \left(\int_0^1 \max_{1\le M\le N}\left(\sum_{k=1}^M \cos 2\pi n_k x\right)^2 dx\right)^{1/2}. \qquad (14)
\end{aligned}
$$

Now by the Carleson-Hunt inequality (see, e.g., the monographs of Arias de Reyna [2] or Mozzochi [16])

$$
\begin{aligned}
\left(\int_0^1 \max_{1\le M\le N}\left(\sum_{k=1}^M \cos 2\pi n_k x\right)^2 dx\right)^{1/2} &\le c \left(\int_0^1 \left(\sum_{k=1}^N \cos 2\pi n_k x\right)^2 dx\right)^{1/2}\\
&\le c\sqrt{N}. \qquad (15)
\end{aligned}
$$

On the other hand for $1 \le M_1 \le M_2 \le N$ Lemma 2 gives

$$
\int_0^1 \left(\sum_{k=M_1}^{M_2} r(n_k x)\right)^2 dx \le \frac{h_{(M_2-M_1+1)}(n_{M_1},\dots,n_{M_2})}{J}.
$$

It is easy to see that $s(M_1,M_2) = h_{(M_2-M_1+1)}(n_{M_1},\dots,n_{M_2})$ is superadditive as a function of M_1, M_2. Thus we can use Lemma 1 and get

$$
\left(\int_0^1 \max_{1\le M\le N}\left(\sum_{k=1}^M r(n_k x)\right)^2 dx\right)^{1/2} \le c(\log N)\left(\frac{h_N(n_1,\dots,n_N)}{J}\right)^{1/2}. \qquad (16)
$$

Together (13), (14), (15) and (16) yield

$$
\left(\int_0^1 \max_{1\le M\le N}\left(\sum_{k=1}^M f(n_k x)\right)^2 dx\right)^{1/2} \le c(\log J)\sqrt{N} + c(\log N)\frac{\sqrt{h_N(n_1,\dots,n_N)}}{\sqrt{J}},
$$

and choosing

$$
J = (\log N)^2 \frac{h_N(n_1,\dots,n_N)}{N}
$$

we finally arrive at

$$\left(\int_0^1 \max_{1\le M\le N}\left(\sum_{k=1}^M f(n_k x)\right)^2 dx\right)^{1/2} \le c\log\left(\frac{(\log N)h_N(n_1,\dots,n_N)}{N}\right)\sqrt{N}. \quad \square$$

3 Proof of Theorem 1

This section is devoted to the proof of Theorem 1. Our method is related to that of Baker [3], but instead of the Erdős-Turán inequality we will use our Lemma 3.

Proof of Theorem 1: Fix some function $g(N)$ satisfying the assumptions in Theorem 1. There exists a constant c_1 such that $c_1 g(N) \ge g\left(2^{M+1}\right)$ for $2^M+1 \le N \le 2^{M+1},\ M \ge 1$. Define

$$A_M := \bigcup_{N=2^M+1}^{2^{M+1}} \left\{x \in (0,1):\ \left|\sum_{k=1}^N f(n_k x)\right| > 2c_1\sqrt{N}g(N)\right\}$$

and

$$B_M := \left\{x \in (0,1):\ \max_{2^M+1\le N\le 2^{M+1}}\left|\sum_{k=1}^N f(n_k x)\right| > \sqrt{2^{M+1}}g\left(2^{M+1}\right)\right\}.$$

Then

$$A_M \subset B_M, \quad M \ge 1.$$

Since by Lemma 3

$$\begin{aligned} &\int_0^1 \max_{2^M+1\le N\le 2^{M+1}}\left(\sum_{k=1}^N f(n_k x)\right)^2 dx \\ \le\ &\int_0^1 \max_{1\le N\le 2^{M+1}}\left(\sum_{k=1}^N f(n_k x)\right)^2 dx \\ \le\ & c\left(\log\left(\frac{M h_{2^{M+1}}(n_1,\dots,n_{2^{M+1}})}{2^{M+1}}\right)\sqrt{2^{M+1}}\right)^2, \end{aligned}$$

Chebyshev's inequality gives

$$\mathbb{P}(A_M) \le \mathbb{P}(B_M) \le c\frac{\left(\log\left(\frac{M h_{2^{M+1}}(n_1,\dots,n_{2^{M+1}})}{2^{M+1}}\right)\right)^2}{g(2^{M+1})^2}.$$

Thus we have that

$$\sum_{M=1}^\infty \mathbb{P}(A_M) \ \le\ c\sum_{M=1}^\infty \frac{\left(\log\left(\frac{M h_{2^{M+1}}(n_1,\dots,n_{2^{M+1}})}{2^{M+1}}\right)\right)^2}{g(2^{M+1})^2} < \infty$$

where the last inequality follows from the Cauchy condensation test and the assumption

$$\sum_{N=1}^\infty \frac{\left(\log\left(\frac{(\log N)h_N(n_1,\dots,n_N)}{N}\right)\right)^2}{N g(N)^2} < \infty.$$

Hence the Borel-Cantelli lemma guarantees that all $x \in (0,1)$ are contained in only finitely many A_M's, except for a set of measure zero. Therefore for all $x \in (0,1)$, except a set of measure zero, there exists a constant $c(x)$ such that

$$\left| \sum_{k=1}^{N} f(n_k x) \right| \leq c(x) \sqrt{N} g(N) \quad \text{for} \quad N \geq 1,$$

which proves Theorem 1.

References

[1] C. Aistleitner. On the law of the iterated logarithm for the discrepancy of lacunary sequences. *Trans. Amer. Math. Soc.*, to appear.

[2] J. Arias de Reyna. *Pointwise convergence of Fourier series.* Springer-Verlag, Berlin, 2002.

[3] R.C. Baker. Metric number theory and the large sieve. *J. London Math. Soc. (2)* **24**:34-40, 1981.

[4] I. Berkes and W. Philipp. The size of trigonometric and Walsh series and uniform distribution mod1. *J. London Math. Soc. (2)* **50**:454-464, 1994.

[5] J.W.S. Cassels. Some metrical theorems of Diophantine approximation. III. *Proc. Cambridge Philos. Soc.* **46**:219-225, 1950.

[6] M. Drmota and R. Tichy. *Sequences, discrepancies and applications.* Lecture Notes in Mathematics, 1651. Springer-Verlag, Berlin, 1997.

[7] T. Dyer and G. Harman. Sums involving common divisors. *J. London Math. Soc. (2)* **34**:1-11, 1986.

[8] P. Erdős and J.F. Koksma. On the uniform distribution modulo 1 of sequences $(f(n,\theta))$. *Nederl. Akad. Wetensch., Proc.* **52**:851-854, 1949.

[9] K. Fukuyama. The law of the iterated logarithm for discrepancies of $\{\theta^n x\}$. *Acta Math. Hung.* **118**:155-170, 2008.

[10] I.S. Gál. A theorem concerning Diophantine approximations. *Nieuw. Arch. Wiskunde (2)* **23**:13-38, 1949.

[11] G. Harman. *Metric number theory.* London Mathematical Society Monographs. New Series, 18. The Clarendon Press, Oxford University Press, New York, 1998.

[12] H. Kesten. The discrepancy of random sequences $\{kx\}$. *Acta Arithm.* **10**:183-213, 1964.

[13] J.F. Koksma. On a certain integral in the theory of uniform distribution. *Nederl. Akad. Wetensch., Proc. Ser. A.* **54** = *Indagationes Math.* **13**:285-287, 1951.

[14] L. Kuipers and H. Niederreiter. *Uniform distribution of sequences.* Pure and Applied Mathematics. Wiley-Interscience [John Wiley & Sons], New York-London-Sydney, 1974.

[15] F.A. Móricz, R.J. Serfling and W.F. Stout. Moment and probability bounds with quasi-superadditive structure for the maximum partial sum. *Annals of Probab.* **10 (4)**:1032-1040, 1982.

[16] C.J. Mozzochi. *On the pointwise convergence of Fourier series.* Springer-Verlag, Berlin-New York 1971.

[17] W. Philipp. Limit theorems for lacunary series and uniform distribution mod 1. *Acta Arith.* **26**:241-251, 1975.

[18] H. Weyl. Über die Gleichverteilung von Zahlen mod. Eins. *Math. Ann.* **77**:313-352, 1916.

[19] A. Zygmund, *Trigonometric Series. Vol. I, II. Reprint of the 1979 edition.* Cambridge Mathematical Library. Cambridge University Press, Cambridge, 1988.