ON A PROBLEM OF W. J. LEVEQUE CONCERNING METRIC DIOPHANTINE APPROXIMATION

MICHAEL FUCHS

Abstract. We consider the diophantine approximation problem

$$\left|x - \frac{p}{q}\right| \le \frac{f(\log q)}{q^2}$$

where f is a fixed function satisfying suitable assumptions. Suppose that x is randomly chosen in the unit interval. In a series of papers that appeared in earlier issues of this journal, LeVeque raised the question whether or not the central limit theorem holds for the solution set of the above inequality (compare also with some work of Erdős). Here, we are going to extend and solve LeVeque's problem.

1. Introduction and Result

Suppose f is a positive real-valued function defined on the non-negative real numbers satisfying the following conditions

$$f \downarrow 0, \quad \sum_{k=1}^{\infty} f(k) = \infty, \tag{1}$$

$$\sum_{k=1}^{n} f(k)k^{-\delta} \ll \left(\sum_{k=1}^{n} f(k)\right)^{1/2}, \tag{2}$$

$$\sum_{k=1}^{n} f(k)^2 \ll \left(\sum_{k=1}^{n} f(k)\right)^{1/2}, \tag{3}$$

where $0 < \delta < 1/2$.

We are concerned with the diophantine approximation problem

$$\left|x - \frac{p}{q}\right| \le \frac{f(\log q)}{q^2}. \tag{4}$$

According to a result of Szüsz [10] (which extends a famous result due to Khintchine [4]), it is known that inequality (4) has infinitely many integer solutions $\langle p, q\rangle$ subjected to the conditions

$$q > 0, \quad q \equiv s \bmod r, \quad s, r \in \mathbb{N}$$

for almost all $x \in [0, 1]$ (in the sense of Lebesgue measure which we are going to denote by λ).

Date: October 23, 2002.

2000 *Mathematics Subject Classification.* 11J83, 60F05.

Key words and phrases. Continued fractions, metric diophantine approximation, dependent random variables, central limit theorem.

This work was supported by the Austrian Science Foundation FWF, grant S8302-MAT.

We are interested in the statistical behavior of the following sequence of random variables

$$X_n(x) := \#\{\langle p,q\rangle | 1 \leq q \leq n, q \equiv s \bmod r, p/q \text{ is a solution of (4)}\}$$

that encodes the number of solutions of (4) for a fixed $x \in [0,1]$.

In [5], LeVeque claimed that he had proved a central limit theorem for the above sequence of random variables (actually, LeVeque considered a more restrictive class of functions f and defined the random variables without the restriction that denominators have to be in an arithmetic progression). This result turned out to be wrong. In fact, LeVeque had proved a central limit theorem under the additional restriction of p and q being coprime (as it was observed by Erdős). Therefore, LeVeque wrote a second paper (see [6]) where he once more tried to obtain the result he originally had in mind. Although he could prove the strong law of large numbers even under much weaker assumptions on f, he failed in proving the desired central limit theorem and so, he had to leave this problem open.

LeVeque's central limit theorem [5] with the additional restriction of p and q being coprime was generalized a few years later by Philipp (see [8]) to more or less to the setting introduced above. Furthermore in the same paper, Philipp stated a theorem that apparently solved LeVeque's problem and outlined a proof. However, there is an uncorrectable mistake in this sketch (to be more precise Lemma 3.3.1 on page 62 in [8] for the approximating sequence of random variables $\tilde{\eta}_\nu$ cannot hold because they have infinite variance) and so, LeVeque's problem remained to be unsolved.

The main result of this paper is the next theorem which extends (in the flavor of Philipp) and finally gives an answer to LeVeque's question.

Theorem 1. *Set*

$$F(n) := \sum_{k=1}^{n} \frac{f(\log k)}{k}.$$

Then, we have

$$\lim_{n\longrightarrow\infty} \lambda\left[X_n \leq \sigma_1 F(n) + \omega\left(\sigma_2 F(n)\log F(n)\right)^{1/2}\right] = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\omega} e^{-u^2/2}du,$$

where

$$\sigma_1 = \frac{2}{r}, \quad \sigma_2 = \frac{12(s,r)\varphi(r)}{\pi^2 r C(s,r)},$$

and

$$C(s,r) = r^2 \prod_{p|r}\left(1-\frac{1}{p}\right)\prod_{p|\frac{r}{(s,r)}}\left(1+\frac{1}{p}\right).$$

Remark 1. Note that the result is different from the claimed result in [8]. Especially, the variance is of order $F(n)\log F(n)$ whereas in [8] the variance was claimed to be of order $F(n)$.

Remark 2. Furthermore note, that the above conditions on f are not exactly the same than the conditions possed by Philipp [8]; especially (3) is added. This is done in order to shorten the proof on the one hand and to make the arguments more lucid on the other hand. However, by combining Philipp's method with the method introduced in this paper, it is straightforward to avoid this additional assumption.

By using the new condition (3), the proofs in [8] can be considerably simplified as well. Thereby, the main simplification emerges from the possibility of direct application of limit theorems for mixing sequences of random variables. This especially makes the elaborate blocking arguments avoidable.

We conclude the introduction by giving a short plan of the paper: in the next section, we prepare the proof of the main result by approximating the sequence of random variables $(X_n)_{n\geq 1}$ several times. In Section 3, we prove the corresponding central limit theorem for the approximating sequence of random variables and finally, Section 4 is used to finish the proof of our main result by showing that the approximation is good enough to entail asymptotic normality of $(X_n)_{n\geq 1}$ from the one of the approximating sequence.

2. Preparatory results

In order to fix notation let $x = [0, a_1, \ldots]$ be the continued fraction expansion of $x \in [0,1]$ and denote by

$$\frac{p_k}{q_k} = [0, a_1, \ldots, a_k]$$

the k-th convergent. Furthermore put

$$\varphi_k = [a_{k+1}, a_{k+2}, a_{k+3}, \ldots] + [0, a_k, a_{k-1}, \ldots, a_1],$$

and

$$\xi_k = [0, a_{k+1}, a_{k+2}, \ldots].$$

We consider the following sequence of random variables

$$Y_k(x) := \#\{1 \leq c \leq a_{k+1} | cq_k \equiv s \bmod r, c^2 \leq \varphi_k f(\log cq_k)\}$$

which, by [10], can be used to approximate the sequence of random variables introduced in the first section

$$\sum_{q_{k+1} \leq n} Y_k(x) \leq X_n(x) \leq \sum_{q_k \leq n} Y_k(x). \tag{5}$$

As it was pointed out in [10], it is not possible to prove some mixing condition for the sequence $(Y_k)_{k\geq 0}$ and therefore, we have to approximate once more. In order to do this, we need the following result of the metric theory of continued fraction expansion due to Gordin and Reznik (see [2])

Lemma 1. *For almost all $x \in [0,1]$, we have*

$$\limsup_{k \longrightarrow \infty} \frac{|\log q_k - k \log \gamma|}{\sqrt{2\sigma^2 k \log \log k}} = 1,$$

where $\sigma > 0$ *and* $\gamma = \exp\big(\pi^2/(12\log 2)\big)$ *is the Khintchine-Levy constant.*

By this Lemma, we have for each $\epsilon > 0$ that there exists κ large enough such that

$$k\log\gamma - \kappa k^{1-\delta} \le \log q_k \le k\log\gamma + \kappa k^{1-\delta}, \quad k \ge 1 \tag{6}$$

for a subset F of $[0,1]$ with $\lambda(F) \ge 1-\epsilon$. Using this, we get

$$\begin{aligned} f((k+1)\log\gamma + \kappa(k+1)^{1-\delta}) &\le f(\log q_{k+1}) \le f(\log cq_k) \\ &\le f(\log q_k) \le f(k\log\gamma - \kappa k^{1-\delta}) \end{aligned}$$

for $x \in F$ and $1 \le c \le a_{k+1}$. Next, we set

$$f_1(k) := f((k+1)\log\gamma + \kappa(k+1)^{1-\delta}), \quad f_2(k) := f((k\log\gamma - \kappa k^{1-\delta}),$$

and define the following random variables

$$Z_k^{(i)}(x) := \#\{1 \le c | cq_k \equiv s \bmod r, c^2 \le (a_{k+1} + 2\delta_{2,i})f_i(k)\} \quad i = 1,2,$$

where $\delta_{i,j}$ is the Kronecker function. It is easy to see that we have

$$Z_k^{(1)}(x) \le Y_k(x) \le Z_k^{(2)}(x) \tag{7}$$

for all $x \in F$.

Our main goal is to obtain the central limit theorem for the sequences $(Z_k^{(i)})_{k\ge 0}$. By a theorem of Szüsz (see [11]), it is easy to see that these sequences are ψ-mixing with an exponential mixing rate, i.e. there exists a constant $q \in (0,1)$ such that

$$|\lambda(AB) - \lambda(A)\lambda(B)| \ll q^n\lambda(A)\lambda(B)$$

for all $A \in \mathcal{F}_0^k, B \in \mathcal{F}_{k+n}^\infty, k \ge 0, n \ge 1$ (here, $\mathcal{F}_a^b$ denotes the σ-algebra generated by $(Z_k^{(i)})_{a\le k\le b}$). However, central limit theorems for ψ-mixing sequences of random variables cannot be applied directly because the variance of $(Z_k^{(i)})_{k\ge 0}$ is infinite (this was overlooked in the sketch of [8]; compare with the introduction). Therefore, we use truncation in order to approximate the sequences $(Z_k^{(i)})_{k\ge 0}$ by a double sequence of random variables.

We put

$$F_i(n) := \sum_{k=1}^{n} f_i(k), \quad i = 1,2,$$

and define the following double sequence of random variables

$$Z_{k,n}^{(i)}(x) := \#\{1 \le c \le \phi_n | cq_k \equiv s \bmod r, c^2 \le (a_{k+1} + 2\delta_{2,i})f_i(k)\}, \quad i = 1,2,$$

where $\phi_n = [(F_i(n))^{1/2}(\log F_i(n))^{1/2-\rho}]$.

Next, we recall a theorem due to Szüsz (see [11]) that will enable us to compute moments of the above random variables.

Lemma 2. *For $t \in [0,1]$ and $a, b \in \mathbb{N}_0$ define*

$$m_k(a,b,t) := \lambda\{x \in [0,1] | q_{k-1} \equiv a \ mod\ r, q_k \equiv b \ mod\ r, \xi_k \leq t\}.$$

Then, we have

$$m_k(a,b,t) = \begin{cases} \frac{1}{C(r)} \frac{\log(1+t)}{\log 2}(1 + O(q^k)) & (a,b,r) = 1 \\ 0 & (a,b,r) \neq 1, \end{cases}$$

where $C(r) = r^2 \prod_{p|r} \left(1 - \frac{1}{p^2}\right)$, $q < 1$ is a constant, and the constant implied in the error term only depends on r.

Furthermore, we need the following identity observed by Philipp (see [8]).

Lemma 3. *Set*

$$K(d) := \frac{r\varphi((d,r))}{C(r)(d,r)},$$

where $C(r)$ is defined as above. Then, we have

$$\sum_{0 \leq d < r} \sum_{\substack{1 \leq c \\ cd \equiv s\ (r)}} K(d) \frac{1}{c^2} = \frac{\pi^2}{6r}.$$

Finally, we need two technical lemmas. The first one can easily be proved by using ideas of [6] (compare with Lemma 4 there).

Lemma 4. *We have, as $n \longrightarrow \infty$,*

$$\sum_{k=1}^{n} f_i(k)^2 \ll F_i(n)^{1/2}, \quad i = 1, 2.$$

Lemma 5. *We have*

$$\sum_{0 \leq d < r, (d,r)|s} (d,r)\varphi((d,r)) = (s,r)\varphi(r) \prod_{p|r, p\nmid \frac{r}{(s,r)}} \left(1 + \frac{1}{p}\right). \tag{8}$$

Proof. We start by observing that

$$\sum_{0 \leq d < r, (d,r)|s} (d,r)\varphi((d,r)) = \sum_{k|(s,r)} k\varphi(k) \sum_{0 \leq d < r, (d,r) = k} 1 = \sum_{k|(s,r)} k\varphi(k)\varphi\left(\frac{r}{k}\right).$$

Furthermore, we put

$$r = \prod_{i=1}^{n} p_i^{e_i}, \quad (s,r) = \prod_{i=1}^{n} p_i^{f_i}$$

for the prime number decomposition of r resp. (s,r). We prove the claimed result by induction on n.

An easy calculation gives the case $n = 1$. Therefore, suppose that (8) is proved for all integers with at most $n-1$ prime numbers in the prime number decomposition. Denote by

$$\bar{r} = \prod_{i=1}^{n-1} p_i^{e_i} = \frac{r}{p_n^{f_n}}$$

and notice that

$$\sum_{k|(s,r)} k\varphi(k)\varphi\left(\frac{r}{k}\right) = \sum_{l=0}^{f_n} \sum_{k|(s,\bar{r})} kp_n^l \varphi(kp_n^l)\varphi\left(\frac{\bar{r}}{k}p_n^{e_n-l}\right)$$
$$= \left(\sum_{k|(s,\bar{r})} k\varphi(k)\varphi\left(\frac{\bar{r}}{k}\right)\right)\left(\sum_{k|p_n^{f_n}} k\varphi(k)\varphi\left(\frac{p_n^{e_n}}{k}\right)\right).$$

Applying twice the induction hypothesis yields

$$\sum_{0\le d<r,(d,r)|s} (d,r)\varphi((d,r)) = (s,\bar{r})\varphi(\bar{r}) \prod_{p|\bar{r},p\nmid\frac{\bar{r}}{(s,\bar{r})}} \left(1+\frac{1}{p}\right)\cdot$$
$$\cdot\, p_n^{f_n}\varphi(p_n^{e_n}) \prod_{p_n\nmid p_n^{e_n-f_n}} \left(1+\frac{1}{p}\right)$$
$$= (s,r)\varphi(r) \prod_{p|r,p\nmid\frac{r}{(s,r)}} \left(1+\frac{1}{p}\right)$$

which is the desired result. □

By these lemmas, we obtain the mean value and the variance of the double sequence introduced above.

Lemma 6. *For the random variables $Z_{k,n}^{(i)}$ introduced above, we have*

$$\mu_{i,n} := \mathbf{E}\sum_{k\le n} Z_{k,n}^{(i)} = \frac{\pi^2}{6r\log 2}F_i(n) + O\left(F_i(n)^{1/2}\right), \tag{9}$$

and

$$\tau_{i,n}^2 := \mathbf{V}\sum_{k\le n} Z_{k,n}^{(i)} \sim \sigma F_i(n)\log F_i(n) \tag{10}$$

with

$$\sigma = \frac{(s,r)\varphi(r)}{rC(s,r)\log 2}$$

where $C(s,r)$ is as in the introduction.

Proof. First, we observe

$$Z_{k,n}^{(i)}(x) = \sum_{0\le d<r} \sum_{\substack{1\le c\le \phi_n \\ cd\equiv s\ (r)}} \xi_k^{(d,c,i)}(x), \tag{11}$$

where

$$\xi_k^{(d,c,i)}(x) = \begin{cases} 1 & \text{if } q_k \equiv d\ (r) \text{ and } c^2 \le (a_{k+1}+2\delta_{i,2})f_i(k) \\ 0 & \text{otherwise} \end{cases}.$$

Therefore

$$\mathbf{E}Z_{k,n}^{(i)} = \sum_{0\le d<r} \sum_{\substack{1\le c\le \phi_n \\ cd\equiv s\ (r)}} \mathbf{E}\xi_k^{(d,c,i)}. \tag{12}$$

Next notice that

$$c^2 \le (a_{k+1}+2\delta_{i,2})f_i(k) \Longleftrightarrow \xi_k \le 1/\left(\left\lceil \frac{c^2-2\delta_{i,2}f_i(k)}{f_i(k)} \right\rceil\right)$$

and together with Lemma 2

$$\mathbf{E}\xi_k^{(d,c,i)} = \frac{K(d)}{\log 2}\left(\frac{f_i(k)}{c^2} + \mathrm{O}\left(\frac{f_i(k)^2}{c^2}\right)\right)(1+\mathrm{O}(q^k)). \tag{13}$$

Hence, by combining (12) and (13)

$$\mathbf{E}Z_{k,n}^{(i)} = \frac{1}{\log 2}(1+\mathrm{O}(q^k))\Bigg(f_i(k)\sum_{0\le d<r} \sum_{\substack{1\le c \\ cd\equiv s\ (r)}} \frac{K(d)}{c^2} -$$

$$- f_i(k)\sum_{0\le d<r} \sum_{\substack{\phi_n < c \\ cd\equiv s\ (r)}} \frac{K(d)}{c^2} + f_i(k)^2 \sum_{0\le d<r} \sum_{\substack{1\le c\le \phi_n \\ cd\equiv s\ (r)}} \mathrm{O}\left(\frac{K(d)}{c^2}\right)\Bigg).$$

Using Lemma 3 for the first double sum and simple estimates for the second and third implies

$$\mathbf{E}Z_{k,n}^{(i)} = \frac{1}{\log 2}\left(\frac{\pi^2}{6r}f_i(k) + \mathrm{O}\left(\frac{f_i(k)}{\phi_n}\right) + \mathrm{O}(f_i(k)^2)\right)(1+\mathrm{O}(q^k)). \tag{14}$$

Finally, summing up and applying Lemma 4 gives (9).

In order to compute the variance, we need to consider the second moment of $Z_{k,n}^{(i)}$. By (11), it is not hard to see that

$$\mathbf{E}(Z_{k,n}^{(i)})^2 = \sum_{0\le d<r} \sum_{\substack{1\le c\le \phi_n \\ cd\equiv s\ (r)}} k_{c,d}\mathbf{E}\xi_k^{(d,l,i)}, \tag{15}$$

where $k_{c,d}$ is the number of pairs $\langle k_1, k_2\rangle$ of solutions of $xd \equiv s\ (r)$ with $k_1 \le c, k_2 \le c$ and either $k_1 = c$ or $k_2 = c$. Furthermore

$$k_{c,d} = \begin{cases} 2\frac{(d,r)}{r}c + c_d & \text{if } (d,r)|s \\ 0 & \text{otherwise} \end{cases}, \tag{16}$$

with a suitable constant c_d. Combining (13),(15), and (16) implies

$$\mathbf{E}(Z_{k,n}^{(i)})^2 = \frac{1}{\log 2}(1+\mathrm{O}(q^k)) \sum_{\substack{0\le d<r\\ (d,r)|s}} \sum_{\substack{1\le c\le \phi_n\\ cd\equiv s\ (r)}} K(d)\cdot$$
$$\cdot\left(2\frac{(d,r)}{r}c + c_d\right)\left(\frac{f_i(k)}{c^2} + \mathrm{O}\left(\frac{f_i(k)^2}{l^2}\right)\right).$$

We extend the last product and break the double sum into four parts. For one part, we have

$$\sum_{\substack{0\le d<r\\ (d,r)|s}} \sum_{\substack{1\le c\le \phi_n\\ cd\equiv s\ (r)}} 2cK(d)\frac{(d,r)}{r}\frac{f_i(k)}{c^2}$$
$$= \frac{2f_i(k)}{C(r)} \sum_{\substack{0\le d<r\\ (d,r)|s}} \varphi(d,r) \sum_{\substack{1\le c\le \phi_n\\ cd\equiv s\ (r)}} \frac{1}{c}$$
$$= f_i(k)\sigma\log 2\log F_i(n) + \mathrm{O}(f_i(k)\log\log F_i(n))$$

where Lemma 5 was used. It is easy to see that the other parts are bounded by either $f_i(k)$, $f_i(k)^2$, or $f_i(k)^2\log F_i(n)$. Hence

$$\sum_{k\le n}\mathbf{E}(Z_{k,n}^{(i)})^2 = \sigma F_i(n)\log F_i(n) + \mathrm{O}(F_i(n)\log\log F_i(n))$$
$$+\mathrm{O}(F_i(n)) + \mathrm{O}\left(\sum_{k=1}^{n} f_i(k)^2\right) + \mathrm{O}\left(\log F_i(n)\sum_{k=1}^{n} f_i(k)^2\right),$$

and taking Lemma 4 into account gives

$$\sum_{k\le n}\mathbf{E}(Z_{k,n}^{(i)})^2 \sim \sigma F_i(n)\log F_i(n). \tag{17}$$

Next observe, by (14) and Lemma 4

$$\sum_{k\le n}(\mathbf{E}Z_{k,n}^{(i)})^2 \ll (F_i(n))^{1/2}, \tag{18}$$

and we finish the proof by considering the covariance of $Z_{k_1,n}^{(i)}$ and $Z_{k_2,n}^{(i)}$ (with $k_1 \le k_2$).

By [10], it is plain that $(Z_{k,n}^{(i)})_{k\le n}$ is ψ-mixing with an exponential mixing rate and that q of Lemma 2 is the basis of the mixing rate. Hence, by Lemma

1.2.1 in [8]

$$|\mathbf{E}Z_{k_1,n}^{(i)}Z_{k_2,n}^{(i)}-\mathbf{E}Z_{k_1}^{(i)}\mathbf{E}Z_{k_2}^{(i)}|\ll q^{k_2-k_1}\mathbf{E}Z_{k_1,n}^{(i)}\mathbf{E}Z_{k_2,n}^{(i)} \ll q^{k_2-k_1}f_i(k_1)f_i(k_2), \tag{19}$$

where the last estimation follows from (14). Combining (17), (18), and (19) immediately gives (10). □

The next step is normalization. Therefore, we define

$$\eta_{k,n}^{(i)} := (Z_{k,n}^{(i)}-\mathbf{E}Z_{k,n}^{(i)})/\tau_{i,n}.$$

In order to prepare the proof of the central limit theorem of these double sequences of random variables, we gather some useful properties.

Lemma 7. *The double sequence $\eta_{k,n}^{(i)}$ satisfies the following properties*

(i) *$\eta_{k,n}^{(i)}$ is uniformly strong mixing (see [7] for a definition),*

(ii)
$$\mathbf{V}\sum_{k\le n}\eta_{k,n}^{(i)}=1,$$

(iii)
$$|\eta_{k,n}^{(i)}|\le\epsilon_n \quad \text{and} \quad \epsilon_n\longrightarrow 0,\ \text{as } n\longrightarrow\infty,$$

(iv)
$$\mathbf{V}\sum_{k\in I}\eta_{k,n}^{(i)}=\sum_{k\in I}\mathbf{V}\eta_{k,n}^{(i)}+O\left(\psi_1(n)\right),$$

where $I\subseteq\{0,1,\ldots,n\}$, the implied constant doesn't depend on I, and $\psi_1(n)\longrightarrow 0$, as $n\longrightarrow\infty$,

(v)
$$\mathbf{E}\left(\sum_{k\in I}\eta_{k,n}^{(i)}\right)^4\ll\sum_{k\in I}\mathbf{E}(Z_{k,n}^{(i)})^4+\left(\sum_{k\in I}\mathbf{E}(Z_{k,n}^{(i)})^2\right)^2+O(\psi_2(n)),$$

where $I\subseteq\{0,1,\ldots,n\}$, the implied constant doesn't depend on I, and $\psi_2(n)\longrightarrow 0$, as $n\longrightarrow\infty$.

Remark 3. Inequality (v) is - except of the error term - an analogue of the famous Rosenthal inequality.

Proof. (i) is an easy consequence of a theorem of Szüsz (see [11]) and (ii) follows immediately from the definition of $\eta_{k,n}^{(i)}$. Furthermore, by (10), (14), and the definition of $Z_{k,n}^{(i)}$ property (iii) is obvious.

In order to prove (iv), observe

$$\mathbf{V}\sum_{k\in I}Z_{k,n}^{(i)}=\sum_{k\in I}\mathbf{V}Z_{k,n}^{(i)}+2\sum_{\substack{0\le k_1<k_2\le n\\ k_1,k_2\in I}}(\mathbf{E}Z_{k_1,n}^{(i)}Z_{k_2,n}^{(i)}-\mathbf{E}Z_{k_1,n}^{(i)}\mathbf{E}Z_{k_2,n}^{(i)}).$$

Applying Lemma 4 and using (19) yields

$$2\sum_{\substack{0\le k_1<k_2\le n\\ k_1,k_2\in I}}(\mathbf{E}Z_{k_1,n}^{(i)}Z_{k_2,n}^{(i)}-\mathbf{E}Z_{k_1,n}^{(i)}\mathbf{E}Z_{k_2,n}^{(i)})\ll\sum_{k=1}^{n}f_i(k)^2\ll F_i(n)^{1/2}$$

and hence (iv) follows.

We are left with the proof of property (v). Therefore, we expand the left hand side of (v) by the multinomial theorem

$$\mathbf{E}\left(\sum_{k\in I}\eta_{k,n}^{(i)}\right)^4=\sum_{e_1+\dots e_t=4}\binom{4}{e_1,\dots,e_t}\mathbf{E}(\eta_{k_1,n}^{(i)})^{e_1}\cdots(\eta_{k_t,n}^{(i)})^{e_t}, \tag{20}$$

where $I=\{k_1,\dots,k_t\}$. We split the sum into several parts according to the powers on the right hand side.

The part with the 4th-powers immediately gives the first sum on the right hand side of (v).

Next, we consider the part where two 2nd-powers occur and take the mixing property of $\eta_{k,n}^{(i)}$ into account

$$\sum_{\substack{0\le l_1<l_2\le n\\ l_1,l_2\in I}}\binom{4}{2,2}\mathbf{E}(\eta_{l_1,n}^{(i)})^2(\eta_{l_2,n}^{(i)})^2=$$

$$\sum_{\substack{0\le l_1<l_2\le n\\ l_1,l_2\in I}}\binom{4}{2,2}\mathbf{E}(\eta_{l_1,n}^{(i)})^2\mathbf{E}(\eta_{l_2,n}^{(i)})^2(1+\mathrm{O}(q^{l_2-l_1}))\ll\left(\sum_{k\in I}\mathbf{E}(\eta_{k,n}^{(i)})^2\right)^2.$$

Hence, we have the second sum on the right hand side of (v).

In order to finish the proof, we have to show that all other parts are bounded by a function which tends to 0 as n tends to infinity. We use ideas of [9]. First consider

$$\sum_{\substack{0\le l_1<l_2<l_3<l_4\le n\\ l_1,l_2,l_3,l_4\in I}}\mathbf{E}\eta_{l_1,n}^{(i)}\eta_{l_2,n}^{(i)}\eta_{l_3,n}^{(i)}\eta_{l_4,n}^{(i)}.$$

We break the sum into two parts $\sum=\sum^*+\sum^{**}$ according to whether $l_2-l_1\le[F_i(n)^{1/2}]$ and $l_3-l_2\le[F_i(n)^{1/2}]$ or not. The first part can be estimated as follows

$$\begin{aligned}\sum\nolimits^*\mathbf{E}\eta_{l_1,n}^{(i)}\eta_{l_2,n}^{(i)}\eta_{l_3,n}^{(i)}\eta_{l_4,n}^{(i)}&\ll\sum\nolimits^*\mathbf{E}|\eta_{l_1,n}^{(i)}\eta_{l_2,n}^{(i)}\eta_{l_3,n}^{(i)}|\mathbf{E}|\eta_{l_4,n}^{(i)}|q^{l_4-l_3}\\&\ll\sum\nolimits^*\mathbf{E}|\eta_{l_1,n}^{(i)}|\mathbf{E}|\eta_{l_2,n}^{(i)}|\mathbf{E}|\eta_{l_3,n}^{(i)}|\mathbf{E}|\eta_{l_4,n}^{(i)}|q^{l_4-l_3}\\&\ll\frac{1}{\tau_{i,n}^4}\sum\nolimits^*\mathbf{E}Z_{l_1,n}^{(i)}\mathbf{E}Z_{l_2,n}^{(i)}\mathbf{E}Z_{l_3,n}^{(i)}\mathbf{E}Z_{l_4,n}^{(i)}q^{l_4-l_3}.\end{aligned}$$

Using (10) and (14) yields

$$\sum\nolimits^{*} \mathbf{E}\eta^{(i)}_{l_1,n}\eta^{(i)}_{l_2,n}\eta^{(i)}_{l_3,n}\eta^{(i)}_{l_4,n} \ll \frac{F_i(n)}{(F_i(n)\log F_i(n))^2}\sum_{k=1}^{n} f_i(k) = \frac{1}{(\log F_i(n))^2}.$$

In order to estimate the second part, we break the sum again into two parts $\sum^{**} = \sum' + \sum''$ according to which of the two condition is violated. We consider the first part

$$\begin{aligned}\sum\nolimits' \mathbf{E}\eta^{(i)}_{l_1,n}\eta^{(i)}_{l_2,n}\eta^{(i)}_{l_3,n}\eta^{(i)}_{l_4,n} &\ll \sum\nolimits' \mathbf{E}|\eta^{(i)}_{l_1,n}|\mathbf{E}|\eta^{(i)}_{l_2,n}\eta^{(i)}_{l_3,n}\eta^{(i)}_{l_4,n}|q^{l_2-l_1}\\ &\ll \frac{q^{(F_i(n))^{1/2}}}{\tau^4_{i,n}}\sum\nolimits' \mathbf{E}Z^{(i)}_{l_1,n}\mathbf{E}Z^{(i)}_{l_2,n}\mathbf{E}Z^{(i)}_{l_3,n}\mathbf{E}Z^{(i)}_{l_4,n}\\ &\ll \frac{q^{(F_i(n))^{1/2}}}{\tau^4_{i,n}}\left(\sum_{k\le n}\mathbf{E}Z^{(i)}_{k,n}\right)^4.\end{aligned}$$

Using (9) and (10) implies

$$\sum\nolimits' \mathbf{E}\eta^{(i)}_{l_1,n}\eta^{(i)}_{l_2,n}\eta^{(i)}_{l_3,n}\eta^{(i)}_{l_4,n} \ll \frac{F_i(n)^4 q^{(F_i(n))^{1/2}}}{(F_i(n)\log F_i(n))^2} = \frac{F_i(n)^2 q^{(F_i(n))^{1/2}}}{(\log F_i(n))^2},$$

which tends to 0 as n tends to ∞. The second part is treated similarly

$$\begin{aligned}\sum\nolimits'' \mathbf{E}\eta^{(i)}_{l_1,n}\eta^{(i)}_{l_2,n}\eta^{(i)}_{l_3,n}\eta^{(i)}_{l_4,n} &\ll \sum\nolimits'' \mathbf{E}\eta^{(i)}_{l_1,n}\eta^{(i)}_{l_2,n}\mathbf{E}\eta^{(i)}_{l_3,n}\eta^{(i)}_{l_4,n}\\ &\quad + \sum\nolimits'' \mathbf{E}|\eta^{(i)}_{l_1,n}\eta^{(i)}_{l_2,n}|\mathbf{E}|\eta^{(i)}_{l_3,n}\eta^{(i)}_{l_4,n}|q^{l_3-l_2}\\ &\ll \sum\nolimits'' \mathbf{E}|\eta^{(i)}_{l_1,n}|\mathbf{E}|\eta^{(i)}_{l_1,n}|q^{l_2-l_1}\mathbf{E}|\eta^{(i)}_{l_3,n}|\mathbf{E}|\eta^{(i)}_{l_4,n}|q^{l_4-l_3}\\ &\quad + q^{(F_i(n))^{1/2}}\sum\nolimits'' \mathbf{E}|\eta^{(i)}_{l_1,n}|\mathbf{E}|\eta^{(i)}_{l_2,n}|\mathbf{E}|\eta^{(i)}_{l_3,n}|\mathbf{E}|\eta^{(i)}_{l_4,n}|\\ &\ll \frac{1}{\tau^4_{i,n}}\left(\sum_{k\le n}\mathbf{E}Z^{(i)}_{k,n}\right)^2 + \frac{q^{(F_i(n))^{1/2}}}{\tau^4_{i,n}}\left(\sum_{k\le n}\mathbf{E}Z^{(i)}_{k,n}\right)^4.\end{aligned}$$

Using once more (9) and (10) yields

$$\sum\nolimits'' \mathbf{E}\eta^{(i)}_{l_1,n}\eta^{(i)}_{l_2,n}\eta^{(i)}_{l_3,n}\eta^{(i)}_{l_4,n} \ll \frac{1}{(\log F_i(n))^2} + \frac{F_i(n)^2 q^{(F_i(n))^{1/2}}}{(\log F_i(n))^2},$$

which again tends to 0 as n tends to ∞.

All remaining parts of the right hand side of (20), namely $\sum_{l_1<l_2}\mathbf{E}(\eta^{(i)}_{l_1,n})^3\eta^{(i)}_{l_2,n}$, $\sum_{l_1<l_2}\mathbf{E}\eta^{(i)}_{l_1,n}(\eta^{(i)}_{l_2,n})^3$, $\sum_{l_1<l_2<l_3}\mathbf{E}(\eta^{(i)}_{l_1,n})^2\eta^{(i)}_{l_2,n}\eta^{(i)}_{l_3,n}$,

$\sum_{l_1<l_2<l_3} \mathbf{E}\eta^{(i)}_{l_1,n}(\eta^{(i)}_{l_2,n})^2\eta^{(i)}_{l_3,n}$, and $\sum_{l_1<l_2<l_3} \mathbf{E}\eta^{(i)}_{l_1,n}\eta^{(i)}_{l_2,n}(\eta^{(i)}_{l_3,n})^2$ can be treated in the same manner and therefore, we are done. □

In the next section, we shall use the properties of the last lemma together with standard techniques in proving central limit theorems for weakly dependent random variables to prove asymptotic normality of $\eta^{(i)}_{k,n}$.

3. The Central Limit Theorem for the double sequence $\eta^{(i)}_{k,n}$

Throughout this section, we suppress the dependence on i. In order to prove the central limit theorem for the sequence $\eta_{k,n}$, we closely follow the proof of the main result in [7] and as in this paper, we proceed in several steps.

3.1. *Step 1: Blocking.*

Because of $\epsilon_n \longrightarrow 0$, $q^n \longrightarrow 0$, and $\psi_i(n) \longrightarrow 0, i = 1,2$ as $n \longrightarrow \infty$ there exists a sequence c_n with the following properties, as $n \to \infty$,

$$c_n \longrightarrow \infty, \tag{21}$$
$$c_n\epsilon_n \longrightarrow 0, \tag{22}$$
$$c_n q^{[\epsilon_n^{-1}]} \longrightarrow 0, \tag{23}$$
$$c_n\psi_1(n) \longrightarrow 0, \tag{24}$$
$$c_n\psi_2(n) \longrightarrow 0. \tag{25}$$

We fix n and define a sequence of integers by

$$m_{0,n} := 0,$$

and for $l = 0, 1, 2, \dots$ by

$$m_{2l+1,n} := \min\{m > m_{2l} | \sum_{k=m_{2l}+1}^{m} \mathbf{V}\eta_{k,n} \geq c_n^{-1}\}, \tag{26}$$
$$m_{2l+2,n} := m_{2l+1} + [\epsilon_n^{-1}]. \tag{27}$$

There are two possibilities how this inductive procedure can stop: on the one hand if we have constructed (27) and there are no random variables for (26) left or on the other hand if the sum of the variances of the remaining random variables is too small to be at least c_n^{-1}. In the first case, we put $m_{2l+2} := n$ and in the second case, we increase m_{2l} by the number of remaining random variables.

Next define

$$I_{l,n} = \{k | m_{2l} < k \leq m_{2l+1}\},$$
$$J_{l,n} = \{k | m_{2l+1} < k \leq m_{2l+2}\},$$

and finally

$$\begin{aligned}\xi_{l,n} &= \sum_{k\in I_{l,n}} \eta_{k,n},\\ \zeta_{l,n} &= \sum_{k\in J_{l,n}} \eta_{k,n},\end{aligned}$$

where $0 \le l < w_n$ and w_n is the number of $I_{l,n}$ (resp. $J_{l,n}$) obtained from the construction above.

We start with an easy observation. By Lemma 7 (iv) and the definition of $I_{l,n}$, we have

$$c_n^{-1} w_n \le \sum_{l=0}^{w_n-1} \sum_{k\in I_{l,n}} \mathbf{V}\eta_{k,n} \le \sum_{k\le n} \mathbf{V}\eta_{k,n} \le K,$$

for a suitable constant K. Hence

$$w_n \le K c_n. \tag{28}$$

The next step is to show that it is sufficient to prove the central limit theorem for the double sequence $\xi_{l,n}$.

3.2. *Step 2: $\zeta_{l,n}$ is negligible.* First observe, by Lemma 7 (iii),(iv), and the definition of $J_{w_n-1,n}$

$$\begin{aligned}\mathbf{V}\zeta_{w(n)-1,n} = \mathbf{V}\sum_{k\in J_{w_n-1,n}} \eta_{k,n} &= \sum_{k\in J_{w_n-1,n}} \mathbf{V}\eta_{k,n} + \mathrm{O}(\psi_1(n))\\ &\le \epsilon_n^{-1} \max_{k\le n} \mathbf{E}\eta_{k,n}^2 + c_n^{-1} + \mathrm{O}(\psi_1(n))\\ &= \epsilon_n + c_n^{-1} + \mathrm{O}(\psi_1(n)),\end{aligned} \tag{29}$$

which tends to 0 as n tends to ∞. Taking once more Lemma 7 (iii),(iv), and the definition of $J_{l,n}$ together with (28) into account yields

$$\begin{aligned}\mathbf{V}\sum_{l<w(n)-1} \zeta_{l,n} &= \sum_{l<w(n)-1} \sum_{k\in J_{l,n}} \mathbf{V}\eta_{l,n} + \mathrm{O}(\psi_1(n))\\ &\le w(n)\epsilon_n^{-1} \max_{k\le n} \mathbf{E}\eta_{k,n}^2 + \mathrm{O}(\psi_1(n))\\ &\ll c_n\epsilon_n + \mathrm{O}(\psi_1(n)),\end{aligned} \tag{30}$$

which - because of (22) - also tends to 0 as n tends to ∞. Next, observe

$$\mathbf{V}\sum_{l<w_n} \zeta_{l,n} = \mathbf{V}\zeta_{w(n)-1,n} + \mathbf{V}\sum_{l<w(n)-1} \zeta_{l,n} + \mathrm{O}(\psi_1(n)),$$

which together with (29) and (30) implies, as $n \to \infty$,

$$\mathbf{V}\sum_{l<w_n} \zeta_{l,n} \to 0. \tag{31}$$

Therefore, as $n \to \infty$,

$$\lambda \left[\sum_{l<w_n} \zeta_{l,n} \geq \epsilon \right] \leq \left(\mathbf{V} \sum_{l<w_n} \zeta_{l,n} \right) / \epsilon^2 \to 0,$$

and the second assertion is proved.

It follows that asymptotic normality of the double sequence $\xi_{l,n}$ is sufficient to imply the central limit theorem for $\eta_{l,n}$. In the next step we use standard tools in order to approximate by independent random variables.

3.3. *Step 3: Approximation by Independent Random Variables.* First by Lemma 7 (iv) and (31), it is plain that

$$\mathbf{V} \sum_{l<w_n} \xi_{l,n} \longrightarrow 1, \quad \text{as } n \longrightarrow \infty. \tag{32}$$

Next consider

$$\sum_{l<w_n} \mathbf{V}\xi_{l,n} = \mathbf{V} \sum_{l<w_n} \xi_{l,n} + \mathrm{O}(c_n \psi_1(n)), \tag{33}$$

where (28) and again Lemma 7 (iv) was used. Hence, by combining (32) and (33) and taking (24) into account, we get

$$a_n := \sum_{l<w_n} \mathbf{V}\xi_{l,n} \longrightarrow 1, \quad \text{as } n \longrightarrow \infty.$$

By a standard argument (see for instance [3]) and using (22) an approximation of the double sequence $\xi_{l,n}$ by an independent double sequence $\bar{\xi}_{l,n}$ with the same distribution can be obtained. Therefore, we are left with proving asymptotic normality of $\bar{\xi}_{l,n}$.

3.4. *Step 4: Asymptotic Normality of* $\bar{\xi}_{l,n}$. First, we normalize

$$\tilde{\xi}_{l,n} := \bar{\xi}_{l,n}/a_n.$$

Notice that it is sufficient to verify Lyapunov's condition for $\tilde{\xi}_{l,n}$, that is

$$\sum_{l<w_n} \mathbf{E}\tilde{\xi}_{l,n}^{2+\tau} \longrightarrow 0, \quad \text{as } n \longrightarrow \infty,$$

where $\tau > 0$ is a real constant. Therefore, observe

$$\sum_{k \in I_l} \mathbf{V}\eta_{k,n} \leq c_n^{-1} + \max_{k \leq n} \mathbf{E}\eta_{k,n}^2 \leq c_n^{-1} + \epsilon_n^2, \tag{34}$$

by the definition of I_l and Lemma 7 (iii). Next, we have

$$\mathbf{E}\eta_{k,n}^4 \leq \epsilon_n^2 \mathbf{E}\eta_{k,n}^2, \tag{35}$$

where again Lemma 7 (iii) was used. Hence Lemma 7 (v), (34), and (35) implies

$$\begin{aligned}\mathbf{E}\xi_{l,n}^4 &= \mathbf{E}\left(\sum_{k\in I_l}\eta_{k,n}\right)^4 \\ &\ll \sum_{k\in I_l}\mathbf{E}\eta_{k,n}^4 + \left(\sum_{k\in I_l}\mathbf{E}\eta_{k,n}^2\right)^2 + \mathrm{O}(\psi_2(n)) \\ &\ll (\epsilon_n + c_n^{-1})^2 + \mathrm{O}(\psi_2(n)).\end{aligned}$$

Furthermore notice

$$\sum_{l<w_n}\mathbf{E}\tilde{\xi}_{l,n}^4 = \frac{1}{a_n^4}\sum_{l<w_n}\mathbf{E}\xi_{l,n}^4 \ll w_n(\epsilon_n + c_n^{-1})^2 + \mathrm{O}(w_n\psi_2(n)),$$

and together with (24),(25), and (28), we get

$$\sum_{l<w_n}\mathbf{E}\tilde{\xi}_{l,n}^4 \longrightarrow 0, \quad \text{as } n \longrightarrow \infty,$$

which proves the asymptotic normality of $\tilde{\xi}_{l,n}$

$$\sum_{l<w_n}\tilde{\xi}_{l,n} \rightharpoonup \mathcal{N}(0,1), \quad \text{as } n \to \infty.$$

Since $a_n \longrightarrow 1,$ as $n \longrightarrow \infty$, and because of step 2 and step 3, we obtain

$$\sum_{k\le n}\eta_{k,n} \rightharpoonup \mathcal{N}(0,1), \quad \text{as } n \to \infty,$$

so that the proof of the central limit theorem for the double sequence $\eta_{k,n}$ is finished.

4. Proof of Theorem 1

We finish the proof by showing that the central limit theorem for $\eta_{k,n}^{(i)}$ entails the asymptotic normality of X_n (indeed, suitable normalized). In order to do this, we need

Lemma 8. *We have*

$$\mathbf{E}\left|\sum_{k\le n}Z_k^{(i)} - \sum_{k\le n}Z_{k,n}^{(i)}\right| \ll \frac{(F_i(n))^{1/2}}{(\log F_i(n))^{1/2-\rho}}.$$

Proof. First observe

$$\begin{aligned}\mathbf{E}\left|\sum_{k\le n} Z_k^{(i)} - \sum_{k\le n} Z_{k,n}^{(i)}\right| &= \sum_{k\le n}\mathbf{E}(Z_k^{(i)} - Z_{k,n}^{(i)})\\ &\le \sum_{k\le n}\mathbf{E}\#\{\phi_n < c | c^2 \le (a_{k+1}+2\delta_{i,2})f_i(k)\}\\ &= \sum_{k\le n}\sum_{\phi_n<c}\mathbf{E}\xi_k^{(c,i)},\end{aligned} \tag{36}$$

where

$$\xi_k^{(c,i)}(x) := \begin{cases} 1 & \text{if } c^2 \le (a_{k+1}+2\delta_{i,2})f_i(k)\\ 0 & \text{otherwise.}\end{cases}$$

Using Lemma 2 yields

$$\mathbf{E}\xi_k^{(c,i)} \ll \frac{f_i(k)}{c^2},$$

and together with (36)

$$\mathbf{E}\left|\sum_{k\le n} Z_k^{(i)} - \sum_{k\le n} Z_{k,n}^{(i)}\right| \ll F_i(n)\sum_{\phi_n<c}\frac{1}{c^2} \ll \frac{F_i(n)}{F_i(n)^{1/2}(\log F_i(n))^{1/2-\rho}},$$

which proves the assertion. □

The above Lemma together with (10) implies

$$\begin{aligned}\lambda\left[\left(\sum_{k\le n} Z_k^{(i)} - \mathbf{E}Z_{k,n}^{(i)}\right)/\tau_{i,n} - \sum_{k\le n}\eta_{k,n}^{(i)} \ge \epsilon\right]&\\ \le \left(\mathbf{E}\left|\sum_{k\le n} Z_k^{(i)} - \sum_{k\le n} Z_{k,n}^{(i)}\right|\right)/(\epsilon\tau_{i,n})&\\ \ll \frac{1}{(\log F_i(n))^{1-\rho}},&\end{aligned}$$

and hence, as $n \longrightarrow \infty$,

$$\left(\sum_{k\le n} Z_k^{(i)} - \frac{\pi^2}{6r\log 2}F_i(n)\right)/(\sigma F_i(n)\log F_i(n))^{1/2} \rightharpoonup \mathcal{N}(0,1), \tag{37}$$

where (9) and (10) were used.

Finally, we need the following technical Lemma due to Philipp (see [8]).

Lemma 9. *Let g_1 (resp. g_2) be the inverse function of $\gamma^{(k+1)}\exp(\kappa(k+1)^{1-\delta})$ (resp. $\gamma^k exp(-\kappa k^{1-\delta})$). Then, we have*

$$F_i(g_i(n)) = \frac{1}{\log\gamma}F(n) + O(F(n)^{1/2}).$$

By (5),(6),(7), and the definition of g_i, we have

$$\sum_{k\le g_1(n)} Z_k^{(1)}(x) \le \sum_{k\le g_1(n)} Y_k(x) \le \sum_{q_{k+1}\le n} Y_k(x) \le X_n(x),$$

and

$$X_n(x) \le \sum_{q_k\le n} Y_k(x) \le \sum_{k\le g_2(n)} Y_k(x) \le \sum_{k\le g_2(n)} Z_k^{(2)}(x),$$

for $x \in F$.

A standard argument together with (37) and Lemma 9 gives our result.

Acknowledgement. The author is indebted to the anonymous referee for many valuable suggestions that lead to an improvement of the presentation of the paper.

References

[1] P. ERDŐS, Some results on diophantine approximation, *Acta Arith.* **5** (1959) 359-369.
[2] M. I. GORDIN AND M. H. REZNIK, The law of the iterated logarithm for the denominators of continued fractions, *Vestnik Leningrad. Univ.* **25** (1970) 28-33.
[3] I. A. IBRAGIMOV, Some limit theorems for stationary processes, *Teor. Verojatnost. i Primenen.* **7** (1962) 361-392.
[4] A. YA. KHINTCHINE, Zur metrischen Theorie der diophantischen Approximationen, *Math. Z.* **24** (1926) 706-714.
[5] W. J. LEVEQUE, On the frequency of small fractional parts in certain real sequences I, *Trans. Amer. Math. Soc.* **87** (1958) 237-260.
[6] W. J. LEVEQUE, On the frequency of small fractional parts in certain real sequences II, *Trans. Amer. Math. Soc.* **94** (1959) 130-149.
[7] M. PELIGRAD, On the asymptotic normality of sequences of weak dependent random variables, *J. Theoret. Probab.* **9** (1996) 703-715.
[8] W. PHILIPP, Mixing sequences of random variables and probablistic number theory, *Mem. Amer. Math. Soc.* **114** (1971), Providence, Rhode Island.
[9] W. PHILIPP, The central limit problem for mixing sequences of random variables, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **12** (1969) 155-171.
[10] P. SZÜSZ, Über die metrische Theorie der Diophantischen Approximation II, *Acta Arith.* **8** (1963) 225-241.
[11] P. SZÜSZ, Verallgemeinerung und Anwendung eines Kusminschen Satzes, *Acta Arith.* **7** (1962) 149-160.

INSTITUT FÜR GEOMETRIE, TU WIEN, WIEDNER HAUPTSTRASSE 8-10/113, 1040 WIEN, AUSTRIA
Current address: INSTITUTE OF STATISTICAL SCIENCE, ACADEMIA SINICA, TAIPEI, 115, TAIWAN, R.O.C.
E-mail address: **fuchs@stat.sinica.edu.tw**