

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

MICHAEL FUCHS

ABSTRACT. We consider the diophantine approximation problem

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

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 of 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 nonnegative real numbers satisfying the following conditions:

- (1) $f \downarrow 0, \quad \sum_{k=1}^{\infty} f(k) = \infty,$
- (2) $\sum_{k=1}^n f(k)k^{-\delta} \ll (\sum_{k=1}^n f(k))^{1/2},$
- (3) $\sum_{k=1}^n f(k)^2 \ll (\sum_{k=1}^n f(k))^{1/2},$

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

We are concerned with the diophantine approximation problem

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

According to a result of Szűs [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

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

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

We are interested in the statistical behaviour of the following sequence of random variables:

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

Received by the editors February 7, 2002 and, in revised form, September 18, 2002.

2000 *Mathematics Subject Classification.* Primary 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.

which 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 Erdős observed). 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 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 flavour 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

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

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]. In particular, 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 as the conditions posed by Philipp [8]; in particular, (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. 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 this 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 that of the approximating sequence.

2. PREPARATORY RESULTS

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

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

the k -th convergent. Furthermore, put

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

and

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

We consider the following sequence of random variables:

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

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

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

As was pointed out in [10], it does not seem to be 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*

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

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

By this lemma, for each $\epsilon > 0$ there exists κ large enough so that

$$(6) \quad k \log \gamma - \kappa k^{1-\delta} \leq \log q_k \leq k \log \gamma + \kappa k^{1-\delta}, \quad k \geq 1,$$

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

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

for $x \in F$ and $1 \leq c \leq a_{k+1}$. 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 random variables

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

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

$$(7) \quad Z_k^{(1)}(x) \leq Y_k(x) \leq Z_k^{(2)}(x)$$

for all $x \in F$.

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

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

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

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

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

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

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 $n \rightarrow \infty$,*

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

Lemma 5. *We have*

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

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

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

and notice that

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

Applying the induction hypothesis twice yields

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

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*

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

and

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

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 that

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

where

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

Therefore

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

Next notice that

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

and, together with Lemma 2,

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

Hence, by combining (12) and (13) we get

$$\begin{aligned} \mathbf{E}Z_{k,n}^{(i)} &= \frac{1}{\log 2} (1 + O(q^k)) \left(f_i(k) \sum_{0 \leq d < r} \sum_{\substack{1 \leq c \\ cd \equiv s \pmod{r}}} \frac{K(d)}{c^2} \right. \\ &\quad \left. - f_i(k) \sum_{0 \leq d < r} \sum_{\substack{\phi_n < c \\ cd \equiv s \pmod{r}}} \frac{K(d)}{c^2} + f_i(k)^2 \sum_{0 \leq d < r} \sum_{\substack{1 \leq c \leq \phi_n \\ cd \equiv s \pmod{r}}} O\left(\frac{K(d)}{c^2}\right) \right). \end{aligned}$$

Using Lemma 3 for the first double sum and simple estimates for the second and third, we obtain

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

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

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

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

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

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

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

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

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

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

$$\begin{aligned} \sum_{k \leq n} \mathbf{E}(Z_{k,n}^{(i)})^2 &= \sigma F_i(n) \log F_i(n) + O(F_i(n) \log \log F_i(n)) \\ &\quad + O(F_i(n)) + O\left(\sum_{k=1}^n f_i(k)^2\right) + O\left(\log F_i(n) \sum_{k=1}^n f_i(k)^2\right), \end{aligned}$$

and taking Lemma 4 into account gives

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

Next observe that, by (14) and Lemma 4,

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

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

By [10], it is plain that $(Z_{k,n}^{(i)})_{k \leq n}$ 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],

$$(19) \quad \begin{aligned} |\mathbf{E}Z_{k_1,n}^{(i)}Z_{k_2,n}^{(i)} - \mathbf{E}Z_{k_1}^{(i)}\mathbf{E}Z_{k_2}^{(i)}| &\ll q^{k_2-k_1} \mathbf{E}Z_{k_1,n}^{(i)} \mathbf{E}Z_{k_2,n}^{(i)} \\ &\ll q^{k_2-k_1} f_i(k_1)f_i(k_2), \end{aligned}$$

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

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 $\eta_{k,n}^{(i)}$ has the following properties:*

- (i) $\eta_{k,n}^{(i)}$ is uniformly strong mixing (see [7] for a definition).
- (ii) $\mathbf{V} \sum_{k \leq n} \eta_{k,n}^{(i)} = 1$.
- (iii) $|\eta_{k,n}^{(i)}| \leq \epsilon_n$ and $\epsilon_n \rightarrow 0$, as $n \rightarrow \infty$.
- (iv) $\mathbf{V} \sum_{k \in I} \eta_{k,n}^{(i)} = \sum_{k \in I} \mathbf{V} \eta_{k,n}^{(i)} + O(\psi_1(n))$, where $I \subseteq \{0, 1, \dots, n\}$, the implied constant does not depend on I , and $\psi_1(n) \rightarrow 0$ as $n \rightarrow \infty$.
- (v)

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

where $I \subseteq \{0, 1, \dots, n\}$, the implied constant does not depend on I , and $\psi_2(n) \rightarrow 0$ as $n \rightarrow \infty$.

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

Proof. (i) is an easy consequence of a theorem of Szűsz (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 that

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

Applying Lemma 4 and using (19) yields

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

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:

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

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:

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

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

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

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

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

Using (10) and (14) yields

$$\begin{aligned} \sum^* \mathbf{E} \eta_{l_1, n}^{(i)} \eta_{l_2, n}^{(i)} \eta_{l_3, n}^{(i)} \eta_{l_4, n}^{(i)} &\ll \frac{F_i(n)}{(F_i(n) \log F_i(n))^2} \sum_{k=1}^n f_i(k) \\ &= \frac{1}{(\log F_i(n))^2}. \end{aligned}$$

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

$$\begin{aligned} \sum' \mathbf{E} \eta_{l_1, n}^{(i)} \eta_{l_2, n}^{(i)} \eta_{l_3, n}^{(i)} \eta_{l_4, n}^{(i)} &\ll \sum' \mathbf{E} |\eta_{l_1, n}^{(i)}| \mathbf{E} |\eta_{l_2, n}^{(i)} \eta_{l_3, n}^{(i)} \eta_{l_4, n}^{(i)}| q^{l_2 - l_1} \\ &\ll \frac{q^{(F_i(n))^{1/2}}}{\tau_{i, n}^4} \sum' \mathbf{E} Z_{l_1, n}^{(i)} \mathbf{E} Z_{l_2, n}^{(i)} \mathbf{E} Z_{l_3, n}^{(i)} \mathbf{E} Z_{l_4, n}^{(i)} \\ &\ll \frac{q^{(F_i(n))^{1/2}}}{\tau_{i, n}^4} \left(\sum_{k \leq n} \mathbf{E} Z_{k, n}^{(i)} \right)^4. \end{aligned}$$

Using (9) and (10) implies

$$\begin{aligned} \sum' \mathbf{E} \eta_{l_1, n}^{(i)} \eta_{l_2, n}^{(i)} \eta_{l_3, n}^{(i)} \eta_{l_4, n}^{(i)} &\ll \frac{F_i(n)^4 q^{(F_i(n))^{1/2}}}{(F_i(n) \log F_i(n))^2} \\ &= \frac{F_i(n)^2 q^{(F_i(n))^{1/2}}}{(\log F_i(n))^2}, \end{aligned}$$

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

$$\begin{aligned} \sum'' \mathbf{E}\eta_{l_1,n}^{(i)}\eta_{l_2,n}^{(i)}\eta_{l_3,n}^{(i)}\eta_{l_4,n}^{(i)} &\ll \sum'' \mathbf{E}\eta_{l_1,n}^{(i)}\eta_{l_2,n}^{(i)}\mathbf{E}\eta_{l_3,n}^{(i)}\eta_{l_4,n}^{(i)} \\ &\quad + \sum'' \mathbf{E}|\eta_{l_1,n}^{(i)}\eta_{l_2,n}^{(i)}|\mathbf{E}|\eta_{l_3,n}^{(i)}\eta_{l_4,n}^{(i)}|q^{l_3-l_2} \\ &\ll \sum'' \mathbf{E}|\eta_{l_1,n}^{(i)}|\mathbf{E}|\eta_{l_2,n}^{(i)}|q^{l_2-l_1}\mathbf{E}|\eta_{l_3,n}^{(i)}|\mathbf{E}|\eta_{l_4,n}^{(i)}|q^{l_4-l_3} \\ &\quad + q^{(F_i(n))^{1/2}} \sum'' \mathbf{E}|\eta_{l_1,n}^{(i)}|\mathbf{E}|\eta_{l_2,n}^{(i)}|\mathbf{E}|\eta_{l_3,n}^{(i)}|\mathbf{E}|\eta_{l_4,n}^{(i)}| \\ &\ll \frac{1}{\tau_{i,n}^4} \left(\sum_{k \leq n} \mathbf{E}Z_{k,n}^{(i)} \right)^2 + \frac{q^{(F_i(n))^{1/2}}}{\tau_{i,n}^4} \left(\sum_{k \leq n} \mathbf{E}Z_{k,n}^{(i)} \right)^4. \end{aligned}$$

Using (9) and (10) once more yields

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

which again tends to 0 as n tends to ∞ .

All remaining parts of the right-hand side of (20), namely

$$\begin{aligned} \sum_{l_1 < l_2} \mathbf{E}(\eta_{l_1,n}^{(i)})^3 \eta_{l_2,n}^{(i)}, & \quad \sum_{l_1 < l_2} \mathbf{E}\eta_{l_1,n}^{(i)}(\eta_{l_2,n}^{(i)})^3, & \quad \sum_{l_1 < l_2 < l_3} \mathbf{E}(\eta_{l_1,n}^{(i)})^2 \eta_{l_2,n}^{(i)} \eta_{l_3,n}^{(i)}, \\ \sum_{l_1 < l_2 < l_3} \mathbf{E}\eta_{l_1,n}^{(i)}(\eta_{l_2,n}^{(i)})^2 \eta_{l_3,n}^{(i)}, & \quad \sum_{l_1 < l_2 < l_3} \mathbf{E}\eta_{l_1,n}^{(i)} \eta_{l_2,n}^{(i)} (\eta_{l_3,n}^{(i)})^2 \end{aligned}$$

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 $\eta_{k,n}^{(i)}$.

3. THE CENTRAL LIMIT THEOREM FOR THE DOUBLE SEQUENCE $\eta_{k,n}^{(i)}$

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 that paper, we proceed in several steps.

3.1. Step 1: Blocking. Because $\epsilon_n \rightarrow 0$, $q^n \rightarrow 0$, and $\psi_i(n) \rightarrow 0, i = 1, 2$, as $n \rightarrow \infty$, there exists a sequence c_n with the following properties, as $n \rightarrow \infty$:

- (21) $c_n \rightarrow \infty,$
- (22) $c_n \epsilon_n \rightarrow 0,$
- (23) $c_n q^{\lceil \epsilon_n^{-1} \rceil} \rightarrow 0,$
- (24) $c_n \psi_1(n) \rightarrow 0,$
- (25) $c_n \psi_2(n) \rightarrow 0.$

We fix n and define a sequence of integers by

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

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

$$(26) \quad m_{2l+1,n} := \min\{m > m_{2l} \mid \sum_{k=m_{2l}+1}^m \mathbf{V}\eta_{k,n} \geq c_n^{-1}\},$$

$$(27) \quad m_{2l+2,n} := m_{2l+1} + \lceil \epsilon_n^{-1} \rceil.$$

There are two possibilities for 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

$$\begin{aligned} I_{l,n} &= \{k \mid m_{2l} < k \leq m_{2l+1}\}, \\ J_{l,n} &= \{k \mid m_{2l+1} < k \leq m_{2l+2}\}, \end{aligned}$$

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 $0 \leq l < w_n$ 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

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

for a suitable constant K . Hence

$$(28) \quad w_n \leq Kc_n.$$

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 that, by Lemma 7 (iii), (iv), and the definition of $J_{w_n-1,n}$,

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

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

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

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

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

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

$$(31) \quad \mathbf{V} \sum_{l < w_n} \zeta_{l,n} \rightarrow 0.$$

Therefore, as $n \rightarrow \infty$,

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

and the second assertion is proved.

It follows that asymptotic normality of the double sequence $\xi_{l,n}$ is sufficient to yield 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

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

Next consider

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

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

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

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,

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

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

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

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

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

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

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

Furthermore, notice that

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

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

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

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

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

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

$$\sum_{k \leq n} \eta_{k,n} \rightarrow \mathcal{N}(0,1), \quad \text{as } n \rightarrow \infty,$$

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*

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

Proof. First observe that

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

where

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

Using Lemma 2 yields

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

and together with (36),

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

which proves the assertion. □

The above lemma together with (10) implies

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

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

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

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

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

and

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

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 led to an improvement of the presentation of the paper.

REFERENCES

1. P. ERDŐS, Some results on diophantine approximation, *Acta Arith.* **5** (1959), 359-369. MR **22**:12091
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), no. 13, 28-33; English transl., *Vestnik Leningrad Univ. Math.* **3** (1976), 207-213. MR **43**:1939
3. I. A. IBRAGIMOV, Some limit theorems for stationary processes, *Teor. Veroyatnost. i Primenen.* **7** (1962) 361-392; English transl. in *Theory Probab. Appl.* **7** (1962), 349-382. MR **26**:5634
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. MR **20**:2314
6. W. J. LEVEQUE, On the frequency of small fractional parts in certain real sequences II, *Trans. Amer. Math. Soc.* **94** (1959) 130-149. MR **22**:12090
7. M. PELIGRAD, On the asymptotic normality of sequences of weak dependent random variables, *J. Theoret. Probab.* **9** (1996) 703-715. MR **97e**:60046
8. W. PHILIPP, Mixing sequences of random variables and probabilistic number theory, *Mem. Amer. Math. Soc.* **114** (1971). MR **55**:10411
9. W. PHILIPP, The central limit problem for mixing sequences of random variables, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **12** (1969) 155-171. MR **39**:7660
10. P. SZÜSZ, Über die metrische Theorie der diophantischen Approximation II, *Acta Arith.* **8** (1963) 225-241. MR **27**:3585
11. P. SZÜSZ, Verallgemeinerung und Anwendung eines Kusminschen Satzes, *Acta Arith.* **7** (1962) 149-160. MR **27**:3584

INSTITUT FÜR GEOMETRIE, TECHNISCHE UNIVERSITÄT 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