ON THE CENTRAL LIMIT THEOREM FOR RANDOM VARIABLES RELATED TO THE CONTINUED FRACTION EXPANSION

BY

C. FAIVRE (MARSEILLE)

1. Introduction. The continued fraction expansion of an irrational number \( x \in [0,1] \) will be denoted by \( x = [0; a_1(x),\ldots,a_n(x),\ldots] \) and \( p_n(x)/q_n(x) = [0; a_1(x),\ldots,a_n(x)] \) (or \( p_n/q_n \) if there is no confusion) will be as usual the \( n \)th convergent. The continued fraction expansion is related to the transformation \( T : [0,1] \to [0,1] \) defined by \( T(0) = 0 \) and \( T(x) = 1/x - \lfloor 1/x \rfloor \) for \( x \in (0,1] \). It is well known that \( ([0,1],T,\nu) \) is an ergodic system [2], where \( \nu \) is the Gauss measure on \([0,1]\) defined by the invariant density

\[
h(x) = \frac{1}{\log 2} \frac{1}{1 + x},
\]

with respect to the Lebesgue measure. Hence for sequences of random variables \( X_1, X_2, \ldots \) with \( X_n(x) = f(T^{n-1}(x)) \) (for an integrable \( f \)) the ergodic theorem can be used to show that for almost all \( x \in [0,1] \), as \( n \to \infty \),

\[
\frac{1}{n} (X_1(x) + \ldots + X_n(x)) \to \int_0^1 f \, d\nu.
\]

For example, the particular choices \( f(x) = \log a_1(x) \) and \( f(x) = 1_{\{p\}}(a_1(x)) \) (where \( p \geq 1 \) is an integer and \( 1_{\{p\}} \) denotes the indicator function of \( \{p\} \)) yield in a simple way the celebrated formulas of Khinchin and Lévy

\[
\lim_{n \to \infty} \sqrt[n]{a_1(x)\ldots a_n(x)} = \prod_{k=1}^{\infty} \left( \frac{(k+1)^2}{k(k+2)} \right)^{\log k / \log 2},
\]

\[
\lim_{n \to \infty} \frac{1}{n} \sharp \{1 \leq i \leq n : a_i(x) = p\} = \frac{1}{\log 2} \log \frac{(p+1)^2}{p(p+2)},
\]

which hold for almost all \( x \). Unfortunately, many interesting sequences \( X_1, X_2, \ldots \) of random variables related to the continued fraction expansion cannot always be expressed as \( f(T^{n-1}x) \), for some function \( f \).

1991 Mathematics Subject Classification: Primary 11K50.
For example, the quantity $X_n = \theta_n$, $n \geq 1$, defined by

$$\left| x - \frac{p_n}{q_n} \right| = \frac{\theta_n}{q_n^2},$$

has this property. The reason is that $\theta_n(x)$ involves the whole continued fraction expansion of $x$, i.e. $\theta_n(x)$ depends on the whole sequence $a_1(x), a_2(x), \ldots$ and not only on $a_n(x), a_{n+1}(x), \ldots$ as would be the case if $\theta_n(x) = f(T^{n-1}x)$. However, the $\theta_n$ can be expressed by means of the $n$th iterate of $W$, the natural extension of $T$, which is the map defined by

$$W : [0, 1]' \times [0, 1] \to [0, 1]' \times [0, 1],$$

$$W(x, y) := \left( T^n x, \frac{1}{a_1(x) + y} \right),$$

where $[0, 1]'$ denotes the set of irrational numbers in $[0, 1]$. To see this, notice that

$$x = [0; a_1, \ldots, a_n + T^n x]$$

yields that

$$(2) \quad x = \frac{p_{n-1}T^n x + p_n}{q_{n-1}T^n x + q_n}.$$ 

Now from (1), (2) and the well known relation $p_{n-1}q_n - p_nq_{n-1} = (-1)^n$ it follows that

$$\theta_n = \frac{T^n x}{T^n x q_{n-1}/q_n + 1}.$$ 

From

$$W^n(x, y) = (T^n x, [0; a_n(x), \ldots, a_1(x) + y])$$

and $q_{n-1}/q_n = [0; a_n, \ldots, a_1]$ we find

$$W^n(x, 0) = (T^n x, q_{n-1}/q_n).$$

Therefore $\theta_n(x) = f(W^n(x, 0))$ with $f(x, y) = x/(xy + 1)$.

Another example is given by

$$r_n(x) = \left| x - \frac{p_n}{q_n} \right| \left| x - \frac{p_{n-1}}{q_{n-1}} \right|,$$

which measures the approximation of $x$ by its $n$th convergent $p_n/q_n$ compared with the approximation by the $(n-1)$th. In this case one can show that

$$r_n(x) = \frac{q_{n-1}}{q_n} T^n x$$

(see [1]). Since $T^n x = [0; a_{n+1}, a_{n+2}, \ldots]$ and $q_{n-1}/q_n = [0; a_n, a_{n-1}, \ldots, a_1]$ we show as for $\theta_n$ that the quantity $r_n$ involves also the whole continued fraction expansion of $x$ and we have $r_n(x) = f(W^n(x, 0))$ with $f(x, y) = xy$ this time. Other examples can also be given which show that many quantities may be expressed as functions of $T^n x$ and $q_{n-1}/q_n$, i.e. as $f(W^n(x, 0))$ for some $f$. 
It is known that \( W \) preserves the probability measure on \([0,1]' \times [0,1]\) defined by
\[
d\mu(x,y) := \frac{1}{\log 2} \frac{dxdy}{(1+xy)^2},
\]
and that \((W,\mu)\) is an ergodic system [5] (and even a \(K\)-system). From the ergodicity of \(W\), Bosma, Jager and Wiedijk [1] have shown that the \(\theta_n\) and the \(r_n\) satisfy a strong law of large numbers. Their proof can easily be adapted to show that for a large class of functions \(f\) a strong law of large numbers holds for \(X_n(x) = f(W^n(x,0))\). Evidently random variables of the form \(f(T^{n-1}x)\) are special cases of those of the form \(h(W^{n-1}(x,0))\).

The aim of this note is to derive a central limit theorem for the random variables \(X_n(x) = f(W^n(x,t))\), where \(t\) is a fixed number in the interval \([0,1]\). This generalizes the case \(X_n(x) = f(T^{n-1}x)\). Classically the central limit theorem for the \(f \circ T^{n-1}\) is investigated using general results about the central limit theorem for dependent variables (see [6] and [3]), since the sequence \(a_1,a_2,\ldots\) of partial quotients is known to be \(\psi\)-mixing ([2], p. 50). For another approach based on the spectral properties of the Perron–Frobenius operator associated with \(T\), see [4].

2. The results. From the definition of \(T\) it follows immediately that
\[
T[0; \alpha_1, \alpha_2, \ldots] = [0; \alpha_2, \alpha_3, \ldots],
\]
that is, \(T\) corresponds to the one-sided shift. Now if we denote by \(\ldots, \alpha_{-1}, \alpha_0; \alpha_1, \ldots\) (where the \(\alpha_i\) are integers \(\geq 1\)) the pair \((x,y)\) with
\[
x = [0; \alpha_1, \alpha_2, \ldots] \quad \text{and} \quad y = [0; \alpha_0, \alpha_{-1}, \ldots],
\]
then
\[
W([\ldots, \alpha_{-1}, \alpha_0; \alpha_1, \ldots]) = [\ldots, \alpha_0, \alpha_1; \alpha_2, \ldots],
\]
in other words, \(W\) is the bilateral shift. Obviously \(W\) is a bijection on \(R = [0,1]' \times [0,1]'\). For \(n \in \mathbb{Z}\) we define random variables \(A_n(z)\) on \(R\) by
\[
A_n(z) = \begin{cases} 
  a_n(x) & \text{if } n \geq 1, \\
  a_{-n+1}(y) & \text{if } n \leq 0,
\end{cases}
\]
for \(z = (x,y)\). Thus
\[
z = [\ldots, A_{-1}(z), A_0(z); A_1(z), \ldots],
\]
and \(A_n = A_0 \circ W^n\) for all \(n \in \mathbb{Z}\). Therefore the process \(\ldots, A_{-1}, A_0, A_1, \ldots\) is stationary (of course we put on \(R\) the probability measure \(\mu\)). The \(A_i\) can be seen as the partial quotients of the “two-sided continued fraction expansion of \(z\)". In the following we will denote by \(C_1(\alpha_1, \ldots, \alpha_p)\) (where \(q \geq 1\)) the set of irrational numbers \(x \in [0,1]\) such that \(a_q(x) = \alpha_1, \ldots, a_{q+p-1}(x) = \alpha_p\), and similarly \(C_2(\alpha_1, \ldots, \alpha_p)\) (with \(q \in \mathbb{Z}\) this time) will denote the set of \(z \in R\) such that \(A_q(z) = \alpha_1, \ldots, A_{q+p-1}(z) = \alpha_p\). Lastly, for all \(k \in \mathbb{Z}\)
we set $\mathcal{F}_-^k = \sigma(\ldots, A_k)$ (i.e. the sigma-field generated by the random variables $\ldots, A_{k-1}, A_k$ and $\mathcal{F}_-^\infty = \sigma(A_k, \ldots)$.

The following proposition shows that the process $(A_n)_{n \in \mathbb{Z}}$ is $\psi$-mixing.

**Proposition 1.** There exist constants $C, q$ with $C > 0$ and $0 < q < 1$ such that for all $k \in \mathbb{Z}$ and $n \geq 1$,

$$|\mu(A \cap B) - \mu(A)\mu(B)| \leq Cq^n\mu(A)\mu(B)$$

for any $A \in \mathcal{F}_-^k$ and $B \in \mathcal{F}_-^{k+n}$.

**Proof.** We shall use in the proof the well known result already stated in the introduction that the process $a_1, a_2, \ldots$ is $\psi$-mixing relative to the Gauss measure $[2]$. More precisely, there exist constants $C, q$ with $C > 0$ and $0 < q < 1$ such that

$$|\nu(C \cap D) - \nu(C)\nu(D)| \leq Cq^n\nu(C)\nu(D)$$

for all $C \in \sigma(a_1, \ldots, a_k)$ and $D \in \sigma(a_{k+n}, \ldots)$. It is enough to prove the proposition when $A$ and $B$ are of the form $A = C_2(\alpha_1, \ldots, \alpha_p)$ with $p + i - 1 = k$ and $B = C_2(\beta_1, \ldots, \beta_j)_{k+n}$. Let $A' = W^{p-1}A = C_2(\alpha_1, \ldots, \alpha_i)_1$ and also $B' = W^{p-1}B = C_2(\beta_1, \ldots, \beta_j)_{i+n}$. Since $W$ is bijective and preserves $\mu$ we have

$$|\mu(A \cap B) - \mu(A)\mu(B)| = |\mu(A' \cap B') - \mu(A')\mu(B')|.$$ 

But

$$A' = C_1(\alpha_1, \ldots, \alpha_i)_1 \times [0, 1]', \quad B' = C_1(\beta_1, \ldots, \beta_j)_{i+n} \times [0, 1]'$$

Thus if $C = C_1(\alpha_1, \ldots, \alpha_i)_1$ and $D = C_1(\beta_1, \ldots, \beta_j)_{i+n}$ we have the equalities

$$\mu(A' \cap B') = \nu(C \cap D), \quad \mu(A') = \nu(C), \quad \mu(B') = \nu(D),$$

since $\int_0^1(1+xy)^{-2}\,dy = 1/(x+1)$. As $|\nu(C \cap D) - \nu(C)\nu(D)| \leq Cq^n\nu(C)\nu(D)$, the result follows. ■

From the previous proposition we are now in a position to apply a general result of Ibragimov on the central limit theorem for processes which are functions of a $\psi$-mixing process ([3], Theorem 2.1). Thus we can now state the following theorem where $\| \cdot \|_2$ denotes the $L^2(\mu)$ norm, $N(0, \sigma^2)$ the normal law with mean 0 and variance $\sigma^2$ (when $\sigma^2 = 0$, $N(0, \sigma^2)$ should be interpreted as the Dirac measure at 0) and finally $\Rightarrow$ will denote the weak convergence of probability measures.

**Theorem 1.** Let $f : [0, 1]' \times [0, 1] \to \mathbb{R}$ in $L^2(\mu)$. If $\sum_{k=0}^\infty v_k < \infty$ where

$$v_k = \| f - E_\mu(f|A_{-k}, \ldots, A_k) \|_2,$$
then the series \( \sigma^2 = \int Y_0^2 \, d\mu + 2 \sum_{k=1}^{\infty} \int Y_0 Y_k \, d\mu \), where \( Y_k = f \circ W^k - \int f \, d\mu \), is absolutely convergent and
\[
\sqrt{n} \left( f + f \circ W + \ldots + f \circ W^{n-1} - n \int f \, d\mu \right) \Rightarrow N(0, \sigma^2),
\]
for all probability measures \( P \) on \([0,1]' \times [0,1]\) absolutely continuous with respect to \( \mu \).

We now state the main theorem. The notations are those of Theorem 1.

**Theorem 2.** Let \( f : [0,1]' \times [0,1] \to \mathbb{R} \) in \( L^2(\mu) \) and let \( t \in [0,1] \).

Assume that:

1. \( \sum_{k=0}^{\infty} v_k < \infty \),
2. \( \sum_{k=0}^{n-1} |f(W^k(x,y)) - f(W^k(x,t))| = o(\sqrt{n}) \) for all \( x \in [0,1]' \) and \( y \in [0,1] \).

Then for any probability measure \( P \) on \([0,1]\), absolutely continuous with respect to the Lebesgue measure \( m \), the sequence of random variables \( X_n(x) = f \circ W^n(x,t) \) defined on \([0,1]' \) satisfies
\[
\sqrt{n} \left( X_1 + \ldots + X_n - na \right) \Rightarrow N(0, \sigma^2),
\]
where \( a = \int f \, d\mu \) and \( \sigma^2 = \int Y_0^2 \, d\mu + 2 \sum_{k=1}^{\infty} \int Y_0 Y_k \, d\mu \). The series is absolutely convergent.

**Proof.** By Theorem 1, the central limit theorem holds for the random variables \( f \circ W^n \) which are defined on the unit square \([0,1]' \times [0,1]\) relative to the probability measure \( \nu_2 = P \otimes m \). Without loss of generality we may suppose that \( \int f \, d\mu = 0 \). Let
\[
Z_n(x,y) = \frac{f(x,y) + \ldots + f \circ W^{n-1}(x,y)}{\sqrt{n}} - \frac{X_1(x) + \ldots + X_n(x)}{\sqrt{n}}.
\]
Then
\[
\int e^{it(f+\ldots+f \circ W^{n-1})/\sqrt{n}} \, d\nu_2 = \int e^{it(X_1+\ldots+X_n)/\sqrt{n}} F_n(x) \, dP(x),
\]
where \( F_n(x) \) is given by
\[
F_n(x) = \int e^{itZ_n(x,y)} \, dm(y).
\]
By (ii), \( Z_n(x,y) \to 0 \) as \( n \to \infty \); then by the dominated convergence theorem we have \( F_n(x) \to 1 \) and it follows that
\[
\int e^{it(X_1+\ldots+X_n)/\sqrt{n}} \, dP(x) \to \frac{1}{\sqrt{2\pi}} e^{-\sigma^2 t^2/2}
\]
as \( n \to \infty \), which proves the result. \( \blacksquare \)
We now show that for suitable functions $f$ conditions (i) and (ii) of the above theorem are satisfied. We denote by $d$ the Euclidean distance on $\mathbb{R}^2$.

**Theorem 3.** If $f$ is Hölder on $R$, i.e. if there exist constants $K, \theta > 0$ such that for all $z, z'$ in $R$, $|f(z) - f(z')| \leq Kd(z, z')^\theta$, then the conclusion of Theorem 2 holds.

**Proof.** Let $z = (x, y) \in R$. For all $k \geq 0$, $z$ belongs to a unique cylinder $C = C_2(\alpha_{-k}, \ldots, \alpha_0, \ldots, \alpha_k)$. From the formula (where $c_1, \ldots, c_n$ are integers $\geq 1$ and $x \in [0, 1]$)

$$[0; c_1, \ldots, c_n + x] = \frac{xp_{n-1}(c_1, \ldots, c_n-1) + p_n(c_1, \ldots, c_n)}{xp_{n-1}(c_1, \ldots, c_n-1) + q_n(c_1, \ldots, c_n)},$$

and from $p_{n-1}q_n - p_nq_{n-1} = (-1)^n$, we deduce the following inequality for all $y \in [0, 1]$:

$$| [0; c_1, \ldots, c_n + x] - [0; c_1, \ldots, c_n + y] | \leq \frac{|y - x|}{q_n^2} \leq \frac{1}{2^{n-1}}.$$

Thus $|x - x'| \leq 2^{-(k-1)}$ and $|y - y'| \leq 2^{-k}$ if $z' = (x', y')$ is another element of $C$. Hence $d(z, z') \leq \sqrt{5}/2^k$. But on the cylinder $C$, $E(f|A_{-k}, \ldots, A_k)$ is constant and equal to $\mu(C)^{-1} \int_C f \, d\mu$. Therefore

$$|f(z) - E(f|A_{-k}, \ldots, A_k)(z)| \leq K\left(\frac{\sqrt{5}}{2^k}\right)^\theta.$$

Thus $v_k \leq K(\sqrt{5}/2^k)^\theta$ and $\sum_{k=0}^{\infty} v_k < \infty$. For (ii) we have, using again (3) and the formula for the iterates of $W$,

$$\frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} |f(W^k(x, y)) - f(W^k(x, t))| \leq \frac{K}{\sqrt{n}} \left[ 1 + \sum_{k=1}^{n-1} \left(\frac{1}{2^{k-1}}\right)^\theta \right],$$

which shows that (ii) is also satisfied.

**Remark.** From Theorem 3 we deduce for example that the conclusion of Theorem 2 holds for the $\theta_n$ and the $r_n$.

**REFERENCES**


Centre de Mathématiques et Informatique
de l'Université de Provence
39, rue Joliot Curie
13453 Marseille Cedex 13, France
E-mail: faivre@gyptis.univ-mrs.fr

Received 18 July 1994;
revised 20 January 1996