ROCKY MOUNTAIN
JOURNAL OF MATHEMATICS
Volume 6, Number 1, Winter 1976

ON THE APPROXIMATION OF INVARIANT MEASURES FOR CONTINUED FRACTIONS*

M. S. WATERMAN

ABSTRACT. Kuzmin's theorem gives a sequence of functions which converge to the density of the invariant measure of n-dimensional continued fractions. The convergence is uniform and geometric. This paper gives bounds on the rate of convergence for two natural approximations to the sequence of functions given by Kuzmin's theorem.

1. **Introduction.** The metric theory of continued fractions, for $n > 1$, does not at this time include the form of the absolutely continuous invariant measure for the associated shift transformation. Numerical as well as theoretical results have been obtained for this measure.

In [1] some ergodic computations were performed in order to approximate the invariant measure for Jacobi's algorithm (the 2-dimensional continued fraction). Although an approximation was obtained, certain measure theoretic difficulties made an estimation of error in the approximation impossible. Thus it is of interest to find a technique where an estimate of the error is possible.

After Schweiger proved the existence of the invariant measure, Kuzmin's theorem was proved for $n > 1$ in [6]. Kuzmin's theorem, generalized to n-dimensions, gives a sequence of approximates to the density of the invariant measure which converges uniformly and geometrically. Many metric results, such as a geometric rate of mixing, follow from this theorem.

The purpose of this paper is to give bounds on the rate of convergence of some natural approximations to the sequence of functions in Kuzmin's theorem. The evaluation of these functions involves summing over a countable set. To illustrate our techniques in a simpler setting and to provide some bounds for Gauss' measure, we deal with the one-dimensional case first. It should be pointed out that each of our approximation theorems depends on the measure of sets whose continued fraction expansions have bounded partial quotients.

Received by the editors on March 26, 1974.

*Work supported in part by National Science Foundation grants GP-28313 and GP-28313 #1. The work was partly performed under the auspices of the U.S. Atomic Energy Commission while the author was a faculty participant of the Associated Western Universities at Los Alamos Scientific Laboratory.

2. **Gauss' Measure.** In this section we consider one-dimensional continued fractions. For $x \in (0, 1)$ and $[\ ]$ denoting the greatest integer function, define

$$T(x) = \frac{1}{x} - \left[\frac{1}{x}\right],$$

$$a_1(x) = \left[\frac{1}{x}\right], \tag{1}$$

$$a_k(x) = a_1(T^{k-1}(x)), \qquad k \geqq 2.$$

The integers $a_1(x), a_2(x), \cdots$ are the partial quotients in the continued fraction expansion of x. We define the ν^{th} order cylinder generated by x to be $B_\nu(x) = \{y : a_i(y) = a_i(x), i = 1, 2, \cdots, \nu\}$. We will use $[a_1, a_2, \cdots]$ to denote the finite or infinite continued fraction $1/a_1 + 1/a_2 + \cdots$. If $T^n(x) = 0$, then the continued fraction expansion of x is finite. Note that T maps $(0, 1)$ into $(0, 1)$. The mapping T does not preserve Lebesgue measure λ, but does preserve Gauss' measure μ. This measure μ is defined by

$$\mu(A) = \frac{1}{\log 2} \int_A \frac{1}{1 + x} dx. \tag{2}$$

Gauss found that $\lambda\{\alpha : T^\nu(\alpha) < x\}$ has the limiting value, as $\nu \to \infty$, $\log_2(1 + x)$. It is easy to see that $\mu(0, x) = \log_2(1 + x)$. This result can be obtained from the individual ergodic theorem. Gauss posed the problem of estimating the difference between the approximate and limiting values. Kuzmin in 1928 solved this problem by a consideration of the sequence of functions defined in the next theorem from [3].

THEOREM 1. *Define* Ψ_ν *recursively by*

$$\Psi_{\nu+1}(x) = \sum_{k=0}^{\infty} \Psi_\nu\left(\frac{1}{k + x}\right)\frac{1}{(k + x)^2}, \quad \nu \geqq 1, \tag{3}$$

where Ψ_0 *satisfies* $0 < m \leqq \Psi_0(x) \leqq M$ *and* $|\Psi_0(x) - \Psi_0(y)| \leqq N|x - y|$. *Then we have*

$$\left|\Psi_\nu(x) - \frac{a}{(1 + x)\log 2}\right| < b\sigma(\nu), \tag{4}$$

where $a = \int_0^1 \Psi_0(t)\, dt$ *and* b *are fixed constants and* $\sigma(\nu) = \operatorname{ess\,sup}_{0<t<1} \lambda(B_\nu(t))$.

The motivation for this theorem is that $h(x)$ is the density of the invariant measure for T if and only if $h(x)$ satisfies Kuzmin's equation:

$$h(x) = \sum_{k=0}^{\infty} h\left(\frac{1}{k+x}\right)\frac{1}{(k+x)^2}.$$

It can be shown ([4]) that $\sigma(\nu) < 3\,(2/(3+\sqrt{5}))^\nu$.

Let $Q = \{1, 2, \cdots\}$. A useful form of (3) is contained in the following lemma from [3].

LEMMA 1. *If $\Psi_\nu(x)$ is defined as in Theorem 1, then*

$$(5)\qquad \Psi_\nu(x) = \sum_{Q^\nu} \Psi_0(f_\nu(x))|f_\nu'(x)|,$$

where $f_\nu(x) = [a_1, a_2, \cdots, a_\nu + x]$ for each $\mathbf{a} = (a_1, a_2, \cdots, a_\nu) \in Q^\nu$.

For the remainder of this section we take $\Psi_0(x) = 1$, so that (5) becomes

$$\Psi_\nu(x) = \sum_{Q^\nu} |f_\nu'(x)|.$$

Applying the chain rule for derivatives to $f_\nu'(x)$, we obtain

$$(6)\qquad \Psi_\nu(x) = \sum_{Q_\nu} \prod_{i=1}^{\nu} \left(\frac{1}{a_i + [a_{i+1}, \cdots, a_\nu + x]}\right)^2.$$

Since Q is infinite, it is of some interest to approximate $\Psi_\nu(x)$ by summing (6) over a selected finite subset of Q_ν. We first consider

$$(7)\qquad A_\nu^N(x) = \sum_{\{1,2,\cdots,N\}^\nu} |f_\nu'(x)|.$$

The next theorem gives bounds on the rate of convergence of Ψ_ν.

THEOREM 2. *If $\Psi_\nu(x)$ and $A_\nu^N(x)$ are defined as in* (6) *and* (7),

$$(8)\qquad \begin{aligned} \frac{1}{2}\left(1 - \left(1 - \frac{1}{3(N+1)}\right)^{\nu-1}\right) &< \Psi_\nu(x) - A_\nu^N(x) \\ &< 2\left(1 - \left(1 - \frac{1}{N+1}\right)\left(1 - \frac{1}{N}\right)^{\nu-1}\right). \end{aligned}$$

PROOF. Let $I_N^\nu = \{1, 2, \cdots, N\}^\nu$ and $\tilde{I}_N^\nu = Q^\nu \sim I_N^\nu$. Then

$$\Psi_\nu(x) - A_\nu^N(x) = \sum_{\tilde{I}_N^\nu} |f_\nu'(x)|.$$

According to Renyi [4], we have

$$\frac{\sup_x|f_\nu'(x)|}{\inf_x|f_\nu'(x)|} \leqq 2.$$

Thus

$$|f_\nu'(x)| \leqq \sup|f_\nu'(x)| \leqq 2\inf|f_\nu'(x)| \leqq 2\int_0^1 |f_\nu'(x)|\,dx = 2\lambda(B_\nu),$$

and

$$|f_\nu'(x)| \geqq \inf|f_\nu'(x)| \geqq \tfrac{1}{2}\sup|f_\nu'(x)| \geqq \tfrac{1}{2}\int_0^1 |f_\nu'(x)|\,dx = \tfrac{1}{2}\lambda(B_\nu).$$

We now conclude

(9) $$\tfrac{1}{2}\sum_{\bar{I}_N^\nu} \lambda(B_\nu) \leqq \Psi_\nu(x) - A_\nu^N(x) \leqq 2\sum_{\bar{I}_N^\nu} \lambda(B_\nu).$$

To bound $\sum_{\bar{I}_N^\nu}\lambda(B_\nu)$ we follow Khintchine [3, p. 70], who derives a lower bound. Let $B_\nu^{(k)}$ be a cylinder of order ν such that $a_\nu = k$ and $B_\nu^{(k)} \subset B_{\nu-1}$. Then [3, p. 69]

$$\frac{1}{3k^2}\lambda(B_{\nu-1}) < \lambda(B_\nu^k) < \frac{1}{k^2}\lambda(B_{\nu-1}),$$

$$\frac{\lambda(B_{\nu-1})}{3}\sum_{k\geqq N+1}\frac{1}{k^2} < \sum_{k\geqq N+1}\lambda(B_\nu^k) < \lambda(B_{\nu-1})\sum_{k\geqq N+1}\frac{1}{k^2},$$

and, noting

$$\frac{1}{N+1} = \int_{N+1}^\infty \frac{1}{x^2}\,dx < \sum_{k\geqq N+1}\frac{1}{k^2} < \int_N^\infty \frac{1}{x^2}dx = \frac{1}{N},$$

$$\lambda(B_{\nu-1})\left(1 - \frac{1}{N}\right) < \sum_{k\leqq N}\lambda(B_\nu^k) < \lambda(B_{\nu-1})\left(1 - \left(\frac{1}{3(N+1)}\right)\right).$$

This argument can be iterated, as in Khintchine, to obtain

$$\left(1 - \frac{1}{N+1}\right)\left(1 - \frac{1}{N}\right)^{\nu-1} = \left(1 - \frac{1}{N}\right)^{\nu-1}\sum_{I_N^1}\lambda(B_1)$$

$$< \sum_{I_N 1}\lambda(B_\nu) < \left(1 - \frac{1}{3(N+1)}\right)^{\nu-1}\sum_{\bar{I}_N^1}\lambda(B_1)$$

$$\leqq \left(1 - \frac{1}{3(N+1)}\right)^{\nu-1}.$$

Using $\sum_{\bar{I}_N^\nu}\lambda(B_\nu) = 1 - \sum_{I_N^\nu}\lambda(B_\nu)$ and (9), we obtain (8).

Clearly the upper bound of (8) is a decreasing function of N. If $\nu - 1 < N$, the binomial expansion gives

$$\Psi_\nu(x) - A_\nu^N(x) < 2\,\nu/N.$$

To make this approximation error equal $b\sigma(\nu)$ of Theorem 1, we need

$$3b \exp\left[-\log\left(\frac{3+\sqrt{5}}{2}\right)\nu\right] = \frac{2\nu}{N},$$

or

$$N = \frac{2}{3b}\nu \exp\left[\log\left(\frac{3+\sqrt{5}}{2}\right)\nu\right].$$

Then the number of points summed over in $A_\nu^N(x)$ is

$$\left(\frac{2}{3b}\nu \exp\left[\log\left(\frac{3+\sqrt{5}}{2}\right)\nu\right]\right)^\nu.$$

Corollary 1. *If N is fixed,* $\lim_{\nu\to\infty} A_\nu^N(x) = 0.$

Proof. By the proof of Theorem 1,

$$A_\nu^N(x) < 2\left(1 - \frac{1}{3(N+1)}\right)^{\nu-1}.$$

In order to find another approximation, we note that

$$\sum_{Q^\nu} \prod_{i=1}^{\nu} \frac{1}{(a_i+1)^2} \leqq \Psi_\nu(x) \leqq \sum_{Q^\nu} \prod_{i=1}^{\nu} \frac{1}{(a_i)^2}.$$

Therefore the "best" $\mathbf{a} = (a_1, \cdots, a_\nu) \in Q^\nu$ to sum over are those $\mathbf{a}$ which have the smallest product. Accordingly we let

(10) $$B_\nu^N(x) = \sum_{\substack{\mathbf{a}\in Q^\nu \\ \prod_{i=1}^{\nu} a_i \leqq N-1}} \prod_{i=1}^{\nu} |f_i'(x)|.$$

Theorem 3. *If $\Psi_\nu(x)$ and $B_\nu^N(x)$ are defined as above, then*

(11) $$\Psi_\nu(x) - B_\nu^N(x) < \frac{2^\nu}{N} \sum_{i=0}^{\nu-1} \frac{(\log N)^i}{i!}.$$

Proof.

$$\Psi_\nu(x) - B_\nu^N(x) = \sum_{\substack{\mathbf{a}\in Q^\nu \\ \Pi a_i \geqq N}} \prod_{i=1}^{\nu} \left(\frac{1}{a_i + [a_{i+1}, \cdots, a_\nu + x]}\right)^2$$

$$\leqq \sum_{\substack{\mathbf{a}\in Q^\nu \\ \Pi a_i \geqq N}} \prod_{i=1}^{\nu} \frac{1}{a_i^2}.$$

The required bound for the last quantity is found in Khintchine [3, p. 75].

The next corollary makes further use of Khintchine's results.

COROLLARY 2. *Let* $N = e^{A\nu}$ *where* $A > 1$ *satisfies* $A - \log A - \log 2 - 1 > 0$. *Then*

$$\Psi_\nu(x) - B_\nu^N(x) < (\sqrt{2}\pi)^{-1}\sqrt{\nu}\, e^{-\nu(A-\log A-\log 2-1)}. \tag{12}$$

PROOF. See Khintchine [3, p. 77].

3. **The Jacobi-Perron Algorithm.** Next we consider the n-dimensional continued fraction (the Jacobi-Perron algorithm) where $n \geqq 2$. Most of the work of this section is analogous with that of section 2, although somewhat harder to accomplish. Identical symbols will be used in these sections and similar calculations and reasonings will be utilized. We now define the Jacobi-Perron algorithm. Let $x \in (0,1)^n$ and $[\ ]$ denote the greatest integer function. Define

$$\begin{aligned} T(x) &= \left(\frac{x_2}{x_1} - \left[\frac{x_2}{x_1}\right], \frac{x_3}{x_1} - \left[\frac{x_3}{x_1}\right], \cdots, \frac{1}{x_1} - \left[\frac{1}{x_1}\right]\right), \\ a^{(1)}(x) &= \left(\left[\frac{x_2}{x_1}\right], \left[\frac{x_3}{x_1}\right], \cdots, \left[\frac{1}{x_1}\right]\right), \\ a^{(\nu)}(x) &= a^{(1)}(T^{\nu-1}(x)),\ \nu \geqq 2. \end{aligned} \tag{13}$$

The expansion of a.a. $x \in (0,1)^n$ is accomplished by

$$x = \lim_{\nu\to\infty} F(a^{(1)}(x) + F(a^{(2)}(x) + \cdots + F(a^{(\nu)}(x)) \cdots)),$$

where $F(x) = (1/x_n, x_1/x_n, \cdots, x_{n-1}/x_n)$.

The invariant measure for T, which is absolutely continuous with respect to Lebesgue measure λ, is known to exist but has not been found yet. Therefore a Kuzmin theorem giving a uniform rate of approximation to $\rho(x)$, the density of this measure, is of some importance. This theorem is stated next [6]. The symbol J_g denotes the Jacobian of the function g.

THEOREM 4. *The set* $(0,1)^n$ *is partitioned in* $n!$ *simplices* A_i *defined by the intersection of the sets* $\{y \in (0,1)^n : y_j < y_{j+h}\}$. *Let* $\mathcal{E}_{i,\nu} = \{k = (k^{(1)}, \cdots, k^{(\nu)}) : T^\nu B(k) \supset A_i\}$. *Define* Ψ_ν *recursively by*

$$\Psi_\nu(x) = \sum_{k\in\mathcal{E}_{i,1}} \Psi_\nu(f_k(x))\,|J_{f_k}(x)|,\quad x \in A_i, \tag{14}$$

where $f_k(x) = F(x+k)$ *and* $0 < m \leqq \Psi_0(x) \leqq M$, $|\Psi_0(x) - \Psi_0(y)| \leqq N\|x-y\|$. *Then*

$$|\Psi_\nu(x) - a\rho(x)| < b\sigma(\nu), \tag{15}$$

where ρ *is the density of the invariant measure for* T, $a = \int \Psi_0 \, d\lambda$ *and* b *are constants, and* $\sigma(\nu) = \text{ess sup}_t \text{diam } B_\nu(t)$.

R. Fischer [2] has shown that $\sigma(\nu) \leqq \sqrt{n}(1 - 1/(n+1)^n)^{\nu/n}$. For the motivation of Theorem 4 we consider $n = 2$. Then $\rho(x)$ is the density of the invariant measure if and only if it satisfies

$$\rho(x) = \sum_{m \geqq 1} \sum_{0 \leqq \ell \leqq m} \rho\left(\frac{1}{m+x_2}, \frac{\ell + x_1}{m+x_2}\right) \frac{1}{(m+x_2)^3},$$

$$\text{if } x \in A_1 = \{x : 0 < x_1 < 1, x_1 < x_2 < 1\},$$

and

$$\rho(x) = \sum_{m \geqq 1} \sum_{0 \leqq \ell < m} \rho\left(\frac{1}{m+x_2}, \frac{\ell + x_1}{m+x_2}\right) \frac{1}{(m+x_2)^3},$$

$$\text{if } x \in A_2 = \{x : 0 < x_1 < 1, 0 < x_2 \leqq x_1\}.$$

A useful form of (14) is contained in the next lemma from [**6**].

Lemma 2. *If* $\Psi_\nu(x)$ *is defined as in Theorem* 4, *then*

$$\Psi_{\nu+1}(x) = \sum_{k \in \mathcal{E}_{i,\nu}} \Psi_0(f_\nu(x)) |J_{f_\nu}(x)|, \; x \in A_i, \tag{16}$$

where $f_\nu(x) = F(k^{(1)} + F(k^{(2)} + \cdots + F(k^{(\nu)} + x) \cdots))$.

Again we take $\Psi_0 = 1$ so that (16) becomes

$$\Psi_\nu(x) = \sum_{\mathcal{E}_{i,\nu}} |J_{f_\nu}(x)|, \; x \in A_i.$$

The set $\mathcal{E}_{i,\nu}$ is infinite for each i so that we define

$$A_\nu^N(x) = \sum_{I^\nu_{i,N}} |J_{f_\nu}(x)|, \; x \in A_i,$$

where $I^\nu_{i,N} = \{k \in \epsilon_{i,\nu} : k_n{}^{(j)} \leqq N \text{ for } j = 1, 2, \cdots, \nu\}$.

Theorem 5. *If* $\Psi_\nu(x)$ *and* $A_\nu^N(x)$ *are defined as above, we have*

$$\begin{aligned} &C^{-1}\left(1 - \left(1 - \frac{\gamma_1}{N+1}\right)^{\nu-1}\right) < \Psi_\nu(x) - A_\nu^N(x) \\ &< CL^{-1}\left(1 - \left(1 - \frac{1}{N+1}\right)\left(1 - \frac{\gamma^2}{N}\right)^{\nu-1}\right). \end{aligned} \tag{17}$$

Remark. The values of the constants in (17) are $C = (1 + 2n)^{n+1}$, $L = (n!)^{-1}$, $\gamma_1 = (n!(n-1)!\ (2n+1)^{n+1}(n+1)^{n+1})^{-1}$, and $\gamma_2 = n!(2n+1)^{n+1}\,2^{n-1}$.

PROOF. By an analysis similar to that in Theorem 2 (see [7]),

$$C^{-1}\lambda(B_\nu) \leqq |J_\nu(x)| \leqq CL^{-1}\lambda(B_\nu).$$

Let $\Theta(k_n^{(\nu)})$ be the collection of order ν cylinders (with n-th corrdinate of $a^{(\nu)}$ equal to $k^{(\nu)}$) which are contained in a certain cylinder $B_{\nu-1}$. Schweiger [5, p. 78] has shown that

$$\gamma_1/(k_n^{(\nu)})^2 < \lambda\ \theta(k_n^{(\nu)})/\lambda(B_{\nu-1}) < \gamma_2/(k_n^{(\nu)})^2.$$

Summing over $k_n^{(\nu)} \geqq N+1$, we obtain as in Theorem 2,

$$\lambda(B_{\nu-1})\left(1 - \frac{\gamma_2}{N}\right) \leqq \sum_{k_n^{(\nu)} \leqq N} \lambda\Theta(k_n^{(\nu)}) \leqq \lambda(B_{\nu-1})\left(1 - \frac{\gamma_1}{N+1}\right).$$

Noting that $\lambda\{x : a_n^{\{1\}}(x) \leqq N\} = 1 - 1/(N+1)$, we can arrive at (17).

COROLLARY 3. *If N is fixed,* $\lim_{\nu\to\infty} A_\nu^N(x) = 0$.

PROOF. See the proofs of Corollary 1 and Theorem 5.

Our second approximation is

$$B_\nu^N(x) = \sum_{\substack{k\in\mathcal{E}_{i,\nu} \\ \prod_1^\nu k_n^{(i)} \leqq N-1}} |J_{f_\nu}{}^{(x)}|,\ x \in A_i.$$

The results make use of some work of Schweiger.

THEOREM 6. *If $\Psi_\nu(x)$ and $B_\nu^N(x)$ are defined as above, then*

$$\Psi_\nu(x) - B_\nu^N(x) < \frac{(2^{n-1})^\nu}{N} \sum_{i=0}^{\nu-1} \frac{(\log N)^i}{i!}. \tag{18}$$

PROOF. See theorem 3 and [5, p. 84].

COROLLARY 4. *Let $N = e^{A\nu}$ where $A > 1$ satisfies $A - \log A - n \log 2 - 1 > 0$. Then*

$$\Psi_\nu - B_\nu^N(x) < (\sqrt{2\pi})^{-1}\sqrt{\nu}\, e^{-\nu(A - \log A - n\log 2 - 1)}. \tag{19}$$

PROOF. See corollary 2 and [5, p. 85].

4. **Conclusion.** In section 2 and 3 it can be seen that our bounds on the rates of convergence are slightly larger for $n > 1$. In neither case

do we have cause for optimism. Corollaries 2 and 4 give the smallest sets to sum over for good approximation.

If $N = e^{A\nu}$ and $n = 1$, we sum over the set $\{a \in Q^\nu : \prod_{i=1}^{\nu} a_i \leqq N - 1\}$. The number of points in this set is of the same order as $\int_{1 \leqq \prod x_i \leqq N} \prod dx_i$ which is of the same order as $N \log N = A\nu e^{A\nu}$. Thus we have shown that Kuzmin's theorem is not too useful for the numerical approximation of invariant measures.

Bibliography

1. W. A. Beyer and M. S. Waterman, *Ergodic computations with continued fractions and Jacobi's algorithm*. Numer. Math. **19** (1972), 195–205.

2. R. Fischer, *Konvergenzgeschwindigkeit beim Jacobialgorithmus*. (To appear.)

3. A. Va. Khintchine, *Continued Fractions*, P. Noordhoff Ltd. Groningen; The Netherlands, 1963.

4. A. Renyi, *Representations for real numbers and their ergodic properties*. Acta. Math. Acad. Sci. Hungar. **8** (1957), 477–493.

5. F. Schweiger, *Geometrische und elementar metrische Sätze über den Jacobischen Algorithmus*. Österr. Akad. Wiss. Wein. math.-naturw. Kl., S.-Ber. Abt. II, **173** (1964), 59–92.

6. F. Schweiger and M. Waterman, *Some remarks on Kuzmin's Theorem for F-expansions*. J. Number Theory **5** (1973), 123–131.

7. M. S. Waterman, *Some erogodic properties of milti-dimensional F-expansions*. Z. Wahrscheinlichkeitstheorie verw. Geb. **16** (1970), 77–103.

Idaho State University, Pocatello, Idaho 83201