Journal of Integer Sequences, Vol. 14 (2011),
Article 11.8.6

On Error Sum Functions Formed by Convergents of Real Numbers

Carsten Elsner and Martin Stein
Fachhochschule für die Wirtschaft Hannover
Freundallee 15
30173 Hannover
Germany
Carsten.Elsner@fhdw.de

Abstract

Let p_m/q_m denote the m-th convergent ($m \geq 0$) from the continued fraction expansion of some real number α. We continue our work on error sum functions defined by $\mathcal{E}(\alpha) := \sum_{m\geq 0} |q_m\alpha - p_m|$ and $\mathcal{E}^*(\alpha) := \sum_{m\geq 0}(q_m\alpha - p_m)$ by proving a new density result for the values of $\mathcal{E}$ and $\mathcal{E}^*$. Moreover, we study the function $\mathcal{E}$ with respect to continuity and compute the integral $\int_0^1 \mathcal{E}(\alpha)\,d\alpha$. We also consider generalized error sum functions for the approximation with algebraic numbers of bounded degrees in the sense of Mahler.

1 Introduction and statement of the main results

Recently the first author [2] introduced two error sums: Let $\alpha = [a_0; a_1, a_2, \ldots]$ be the continued fraction expansion of a real number α, which may be finite in the case of a rational number α. Let

$$\frac{p_m}{q_m} = [a_0; a_1, \ldots, a_m] \qquad (m \geq 0)$$

denote the convergents of α. The error sum functions $\mathcal{E}(\alpha)$ and $\mathcal{E}^*(\alpha)$ are defined by

$$\begin{aligned}\mathcal{E}(\alpha) &= \sum_{m\geq 0} |\alpha q_m - p_m| = \sum_{m\geq 0} (-1)^m(\alpha q_m - p_m)\,,\\ \mathcal{E}^*(\alpha) &= \sum_{m\geq 0}(\alpha q_m - p_m)\,.\end{aligned}$$

Both functions do not depend on the integer part a_0 of α. So we may restrict their domains on the interval $[0,1)$.

The first author [2] proved that

$$0 \le \mathcal{E}(\alpha) \le \rho = \frac{1+\sqrt{5}}{2} \qquad \text{and} \qquad 0 \le \mathcal{E}^*(\alpha) \le 1 \qquad (\alpha \in \mathbb{R}).$$

The series $\sum_{m\ge0} |q_m\alpha - p_m| \in [0,\rho]$ measures the approximation properties of α on average. The smaller this series is, the better rational approximations α has. Nevertheless, α can be a Liouville number and $\sum_{m\ge0} |q_m\alpha - p_m|$ takes a value close to ρ. So, it may be interesting to question on the average value of $\mathcal{E}$ and $\mathcal{E}^*$, respectively. We compute the average value of $\mathcal{E}$, see Theorem 5. The error sum functions $\mathcal{E}$ and $\mathcal{E}^*$ have various interesting properties. In [2], applications are discussed for certain transcendental numbers and for quadratic irrational numbers. For instance, we have

$$\begin{aligned}
\mathcal{E}(\exp(1)) &= \sum_{m\ge0} |q_m e - p_m| = 2e\int_0^1 \exp(-t^2)\,dt - e = 1.3418751\ldots, \\
\mathcal{E}^*(\exp(1)) &= \sum_{m\ge0} (q_m e - p_m) = 2\int_0^1 \exp(t^2)\,dt - 2e + 3 = 0.4887398\ldots, \\
\mathcal{E}(\sqrt{7}) &= \sum_{m\ge0} |q_m\sqrt{7} - p_m| = \frac{7+5\sqrt{7}}{14} = 1.444911182\ldots, \\
\mathcal{E}^*(\sqrt{7}) &= \sum_{m\ge0} (q_m\sqrt{7} - p_m) = \frac{21-5\sqrt{7}}{14} = 0.555088817\ldots.
\end{aligned}$$

It is clear that for any rational number α the series for $\mathcal{E}(\alpha)$ and $\mathcal{E}^*(\alpha)$ become finite sums and therefore belong to $\mathbb{Q}$. In the case of quadratic irrational numbers α we have $\mathcal{E}(\alpha) \in \mathbb{Q}(\alpha)$ and $\mathcal{E}^*(\alpha) \in \mathbb{Q}(\alpha)$ ([2, Theorem 3]). But for quadratic irrationals $\mathcal{E}(\alpha) \in \mathbb{Q}(\alpha) \setminus \mathbb{Q}$ does not hold in general. For example, $\mathcal{E}((3-\sqrt{5})/2) = 1$ (see [3, Lemma 8]). On the other hand $\mathcal{E}(\alpha) \in \mathbb{Q}(\alpha)$ is not true for all real numbers α. For $\alpha = e = \exp(1)$ we have $\mathcal{E}(e) \notin \mathbb{Q}(e)$, since e and $\int_0^1 \exp(-t^2)\,dt$ are algebraically independent over $\mathbb{Q}$. This follows from a remark on page 193 in [8]. Similarly, one can show that $\mathcal{E}^*(e) \notin \mathbb{Q}(e)$.

The authors [3] studied the value distribution of the error sum functions in more detail. They constructed two algorithms which prove that the set of values of $\mathcal{E}$ is dense in the interval $I_{\mathcal{E}} = [0,\rho]$, and that the set of values of $\mathcal{E}^*$ is dense in the interval $I_{\mathcal{E}^*} = [0,1]$ (see [3, Theorems 1, 2]). But, given any uniformly modulo one distributed sequence $(\alpha_\nu)_{\nu\ge1}$ of real numbers, the sequences $(\mathcal{E}(\alpha_\nu))_{\nu\ge1}$ and $(\mathcal{E}^*(\alpha_\nu))_{\nu\ge1}$ are not uniformly distributed in $I_{\mathcal{E}}$ and $I_{\mathcal{E}^*}$, respectively (see [3, Theorems 3, 4]). In this paper we show that any dense subset of $(0,1)$ is mapped by $\mathcal{E}(\alpha)$ and $\mathcal{E}^*(\alpha)$ into a set which is dense in $I_{\mathcal{E}}$ and $I_{\mathcal{E}^*}$, respectively. Then, we continue to study the analytic properties of the error sum functions. The function $\mathcal{E}^*$ has already been investigated by Ridley and Petruska [7]. Among other things they showed that $\mathcal{E}^*(\alpha)$ is continuous at every irrational point α, and discontinuous when α is rational. Moreover, they computed the integral $\int_0^1 \mathcal{E}^*(\alpha)\,d\alpha$ by applying the functional equation

$$\mathcal{E}^*(\alpha) + \mathcal{E}^*(1-\alpha) = \max\{\alpha, 1-\alpha\} \quad \text{except at} \quad \alpha = 0 \quad \text{and} \quad \alpha = \frac{1}{2}.$$

Inspired by the work of Ridley and Petruska, we prove similar results for the error sum function $\mathcal{E}$. We compute the integral $\int_0^1 \mathcal{E}(\alpha)\, d\alpha$ by using a multiple sum, which expresses the integral in terms of denominators of convergents. Unfortunately, the functional equation

$$\mathcal{E}(\alpha) - \mathcal{E}(1-\alpha) = \begin{cases} \alpha - 1, & \text{if} \quad 0 < \alpha < 1/2; \\ \alpha, & \text{if} \quad 1/2 < \alpha < 1; \end{cases}$$

cannot be used to evaluate the integral $\int_0^1 \mathcal{E}(\alpha)\, d\alpha$.

The main results of this paper are given by the following theorems.

Theorem 1. *Let $(\alpha_n)_{n\geq 1}$ be a sequence of real numbers forming a dense set $\{\alpha_n : n \in \mathbb{N}\}$ in $(0,1)$. Then the set $\{\mathcal{E}(\alpha_n) : n \in \mathbb{N}\}$ is dense in $(0,\rho)$, and the set $\{\mathcal{E}^*(\alpha_n) : n \in \mathbb{N}\}$ is dense in $(0,1)$.*

Theorem 2. *The function $\mathcal{E}(\alpha)$ is discontinuous at every rational point α, and it is continuous at every irrational point α.*

Example 3. Let n, k be integers with $n, k \geq 3$. For $x = 1/n$ we have

$$\begin{aligned} \mathcal{E}\Big(\frac{1}{n} + \frac{1}{n^k}\Big) &= \frac{2}{n} + \frac{3}{n^k} \quad \rightarrow \quad \frac{2}{n} \quad (k \to \infty)\,, \\ \mathcal{E}\Big(\frac{1}{n} - \frac{1}{n^k}\Big) &= \frac{1}{n} - \frac{1}{n^k} + \frac{2}{n^{k-1}} \quad \rightarrow \quad \frac{1}{n} \quad (k \to \infty)\,, \\ \mathcal{E}^*\Big(\frac{1}{n} + \frac{1}{n^k}\Big) &= \frac{2}{n^{k-1}} - \frac{1}{n^k} \quad \rightarrow \quad 0 \quad (k \to \infty)\,, \\ \mathcal{E}^*\Big(\frac{1}{n} - \frac{1}{n^k}\Big) &= \frac{1}{n} - \frac{3}{n^k} \quad \rightarrow \quad \frac{1}{n} \quad (k \to \infty)\,. \end{aligned}$$

These expressions are obtained by using the identities

$$\begin{aligned} \frac{1}{n} + \frac{1}{n^k} &= [\,0; n-1, 1, n^{k-2}-1, n\,]\,, \\ \frac{1}{n} - \frac{1}{n^k} &= [\,0; n, n^{k-2}-1, 1, n-1\,]\,. \end{aligned}$$

□

Let $m \geq 1$, and let $a_1, \ldots, a_m$ be positive integers. Set

$$\frac{p_m}{q_m} = [\,0; a_1, \ldots, a_m\,]\,,$$

where p_m and q_m with $q_m > 0$ are coprime integers.

Theorem 4. *We have*

$$\int_0^1 \mathcal{E}(\alpha)\, d\alpha = \frac{1}{2} + \frac{1}{2} \sum_{m=1}^{\infty} \sum_{a_1=1}^{\infty} \cdots \sum_{a_m=1}^{\infty} \frac{1}{q_m (q_m + q_{m-1})^2}\,,$$

and

$$\int_0^1 \mathcal{E}^*(\alpha)\, d\alpha = \frac{1}{2} + \frac{1}{2} \sum_{m=1}^{\infty} \sum_{a_1=1}^{\infty} \cdots \sum_{a_m=1}^{\infty} \frac{(-1)^m}{q_m (q_m + q_{m-1})^2}\,.$$

With the first identity from the preceding theorem, we compute the mean value of the function $\mathcal{E}$.

Theorem 5. *We have*

$$\int_0^1 \mathcal{E}(\alpha)\, d\alpha = -\frac{5}{8} + \frac{3\zeta(2)\log 2}{2\zeta(3)} = 0.79778798\ldots\,,$$

where $\zeta(s)$ denotes the Riemann zeta function.

Remark 6. Ridley and Petruska [7] proved that

$$\int_0^1 \mathcal{E}^*(\alpha)\, d\alpha = \frac{3}{8}\,.$$

We point out that by Theorem 5 and Remark 6 the mean values of $\mathcal{E}$ and $\mathcal{E}^*$ are less than half of the maximum value of $\mathcal{E}$ and $\mathcal{E}^*$, respectively.

In Section 5 we generalize the error sum function $\mathcal{E}$ to the approximation with algebraic numbers of bounded degree. Here, the Mahler function $w_n(H,\alpha)$ will be involved.

2 Proof of Theorem 1

We will only prove the statement concerning the values of the function $\mathcal{E}$, since there are no additional arguments for the function $\mathcal{E}^*$.

It is shown in the proof of Theorem 1 in [3] that the set $\{\mathcal{E}(\alpha) : \alpha \in \mathbb{Q} \cap (0,1)\}$ is dense in $(0,\rho)$. Hence, for any real number $\eta \in (0,\rho)$ and for any $\delta > 0$ there is a rational number $r \in (0,1)$ satisfying

$$\big|\eta - \mathcal{E}(r)\big| < \frac{\delta}{3}\,. \tag{1}$$

By

$$r = [0; a_1, a_2, \ldots, a_t] = \frac{p_t}{q_t}$$

we denote the continued fraction expansion of r. Without loss of generality we may assume that t satisfies

$$\frac{1+\sqrt{2}}{\big(\sqrt{2}\big)^{t-1}} < \frac{\delta}{3}\,. \tag{2}$$

This can be seen by the following argument: For any number $r' = [0; a_1, \ldots, a_{t'}]$ satisfying $|\eta - \mathcal{E}(r')| < \delta/3$ and $t' < t$ we construct a number $r = [0; a_1, a_2, \ldots, a_t]$ with $a_{t'+1} = \cdots = a_t = b$, such that t satisfies (2) and b is sufficiently large (see [3, Lemma 1]). Namely, for r_k

defined by $r_k := [0; a_1, \dots, a_{t'}, \underbrace{b, \dots, b}_{k}]$ we have

$$\begin{aligned}
\big|\mathcal{E}(r) - \mathcal{E}(r')\big| &= \big|\mathcal{E}(r_{t-t'}) - \mathcal{E}(r_0)\big| = \Big| \sum_{k=0}^{t-t'-1} \mathcal{E}(r_{k+1}) - \mathcal{E}(r_k) \Big| \\
&\le \sum_{k=0}^{t-t'-1} \big| \mathcal{E}(r_{k+1}) - \mathcal{E}(r_k) \big| < \sum_{k=0}^{t-t'-1} \frac{1}{b} = \frac{t-t'}{b} \\
&< \frac{t}{b} \to 0 \qquad (b \to \infty)\,.
\end{aligned}$$

Since the set $\{\alpha_n : n \in \mathbb{N}\}$ is dense in $(0,1)$ by the assumption in the theorem, there is a positive integer m satisfying

$$\alpha_m = [0; a_1, a_2, \dots, a_t, a_{t+1}, \dots]$$

and

$$\big|r - \alpha_m\big| < \frac{\delta}{3(t+1)q_t}\,. \tag{3}$$

Let p_ν/q_ν be the convergents of α_m. Then, by applying the inequalities (1), (3) and (2) we have

$$\begin{aligned}
\big|\eta - \mathcal{E}(\alpha_m)\big| &= \big|\eta - \mathcal{E}(r) + \mathcal{E}(r) - \mathcal{E}(\alpha_m)\big| \le \big|\eta - \mathcal{E}(r)\big| + \big|\mathcal{E}(r) - \mathcal{E}(\alpha_m)\big| \\
&< \frac{\delta}{3} + \Big| \sum_{\nu=0}^{t} |q_\nu r - p_\nu| - \sum_{\nu\ge 0} |q_\nu \alpha_m - p_\nu| \Big| \\
&\le \frac{\delta}{3} + \sum_{\nu=0}^{t} |r - \alpha_m| q_t + \sum_{\nu \ge t+1} |q_\nu \alpha_m - p_\nu| \\
&\le \frac{\delta}{3} + \sum_{\nu=0}^{t} \frac{\delta}{3(t+1)} + \sum_{\nu \ge t+1} \frac{1}{q_\nu} \\
&\le \frac{2\delta}{3} + \sum_{\nu \ge t} \frac{1}{(\sqrt{2})^\nu} \\
&= \frac{2\delta}{3} + \frac{1+\sqrt{2}}{(\sqrt{2})^{t-1}} \\
&< \delta\,,
\end{aligned}$$

which completes the proof of Theorem 1. □

3 Proof of Theorem 2

Since the function $\mathcal{E}$ is periodic of period one, it suffices to prove Theorem 2 for $\alpha \in [0,1)$. We will prove the statement on continuity first. Let $\eta \in [0,1)$ be a real irrational number, say

$$\eta = [0; a_1, a_2, \dots]\,,$$

and let $(\xi_n)_{n\geq 1}$ be a sequence of real numbers converging to η. By $I_m = I_m(a_1, \dots, a_m)$ we denote the interval defined uniquely by

$$[0; b_1, b_2, \dots] \in I_m \qquad \Longleftrightarrow \qquad (b_1 = a_1 \wedge \dots \wedge b_m = a_m)\,. \tag{4}$$

The boundary points of I_m are rational numbers, and therefore the irrational number η lies in the interior of I_m for any $m \geq 1$. With $\lim_{n\to\infty} \xi_n = \eta$ we conclude on

$$\xi_n \in I_m \quad (n \geq n_0)$$

for some positive integer $n_0 = n_0(m)$. Hence, by (4), we have

$$\xi_n = [0; a_1, \dots, a_m, \dots]. \tag{5}$$

Let p_ν/q_ν for $\nu \geq 0$ be the convergents of η and let $p_\nu^{(n)}/q_\nu^{(n)}$ be the convergents of ξ_n. Then, from (5), it follows that

$$\frac{p_\nu}{q_\nu} = \frac{p_\nu^{(n)}}{q_\nu^{(n)}} \quad (0 \leq \nu \leq m)\,.$$

For a fixed positive integer m and any $n \geq n_0$ we estimate

$$\begin{aligned}
|\mathcal{E}(\eta) - \mathcal{E}(\xi_n)| &= \Big| \sum_{\nu\geq 0} |q_\nu \eta - p_\nu| - \sum_{\nu \geq 0} |q_\nu^{(n)} \xi_n - p_\nu^{(n)}| \Big| \\
&\leq \Big| \sum_{\nu=0}^{m} (-1)^\nu q_\nu (\eta - \xi_n) \Big| + \sum_{\nu \geq m+1} |q_\nu \eta - p_\nu| + \sum_{\nu\geq m+1} |q_\nu^{(n)} \xi_n - p_\nu^{(n)}| \\
&\leq \Big| \sum_{\nu=0}^{m} (-1)^\nu q_\nu (\eta - \xi_n) \Big| + \sum_{\nu \geq m+1} \frac{1}{q_\nu} + \sum_{\nu\geq m+1} \frac{1}{q_\nu^{(n)}} \\
&\leq \Big| \sum_{\nu=0}^{m} (-1)^\nu q_\nu (\eta - \xi_n) \Big| + \sum_{\nu \geq m+1} \frac{1}{2^{(\nu-1)/2}} + \sum_{\nu\geq m+1} \frac{1}{2^{(\nu-1)/2}} \\
&= \Big| \sum_{\nu=0}^{m} (-1)^\nu q_\nu (\eta - \xi_n) \Big| + \frac{2\sqrt{2}}{\sqrt{2}-1} \cdot \frac{1}{(\sqrt{2})^m}\,.
\end{aligned}$$

Since m can be chosen arbitrary large and ξ_n tends to η for increasing n, we conclude on

$$\lim_{n\to\infty} \mathcal{E}(\xi_n) = \mathcal{E}(\eta)\,.$$

This proves that the function $\mathcal{E}(\alpha)$ is continuous at every irrational point α.

To prove the statement on discontinuity we shall at first discuss the case when η is a rational number in $(0,1)$. Let

$$\eta = [0; a_1, a_2, \dots, a_m]$$

for some integers $m \geq 1$ and $a_m > 1$. Moreover, let $(\xi_n^{(1)})_{n\geq 2}$ and $(\xi_n^{(2)})_{n\geq 2}$ be two sequences of rationals defined by

$$\xi_n^{(1)} = [0; a_1, \ldots, a_m, n] \quad \text{and} \quad \xi_n^{(2)} = [0; a_1, \ldots, a_m - 1, 1, n] \quad (n \geq 2)\,.$$

Obviously we have

$$\lim_{n\to\infty} \xi_n^{(1)} = \eta = \lim_{n\to\infty} \xi_n^{(2)}\,. \tag{6}$$

Let $p_\nu^{(1)}/q_\nu^{(1)}$ for $\nu = 0, \ldots, m+1$ be the convergents of $\xi_n^{(1)}$. By $p_\nu^{(2)}/q_\nu^{(2)}$ for $\nu = 0, \ldots, m+2$ we denote the convergents of $\xi_n^{(2)}$. Then we have

$$\frac{p_\nu^{(1)}}{q_\nu^{(1)}} = \frac{p_\nu^{(2)}}{q_\nu^{(2)}} \quad (0 \leq \nu \leq m-1)\,.$$

Therefore we may set $p_\nu := p_\nu^{(1)} = p_\nu^{(2)}$ and $q_\nu := q_\nu^{(1)} = q_\nu^{(2)}$ for $\nu = 0, \ldots, m-1$. We compute

$$\begin{aligned}
\mathcal{E}(\xi_n^{(2)}) - \mathcal{E}(\xi_n^{(1)}) - &\sum_{\nu=0}^{m-1} (-1)^\nu (\xi_n^{(2)} - \xi_n^{(1)}) q_\nu \\
&= (-1)^m((a_m - 1)q_{m-1} + q_{m-2})\xi_n^{(2)} - (-1)^m((a_m - 1)p_{m-1} + p_{m-2}) \\
&\quad + (-1)^{m+1}(a_m q_{m-1} + q_{m-2})\xi_n^{(2)} - (-1)^{m+1}(a_m p_{m-1} + p_{m-2}) \\
&\quad - (-1)^m(a_m q_{m-1} + q_{m-2})\xi_n^{(1)} + (-1)^m(a_m p_{m-1} + p_{m-2}) \\
&= (-1)^m(\xi_n^{(2)} - \xi_n^{(1)})(a_m q_{m-1} + q_{m-2}) + (-1)^m(p_{m-1} - q_{m-1}\xi_n^{(2)}) \\
&\quad + (-1)^{m+1}(a_m q_{m-1} + q_{m-2})\xi_n^{(2)} - (-1)^{m+1}(a_m p_{m-1} + p_{m-2})\,.
\end{aligned}$$

For $n \to \infty$, by (6) and with $\eta = p_m^{(1)}/q_m^{(1)}$ we obtain the limit

$$\begin{aligned}
\lim_{n\to\infty} \left(\mathcal{E}(\xi_n^{(2)}) - \mathcal{E}(\xi_n^{(1)})\right) &= (-1)^m \left[(p_{m-1} - q_{m-1}\eta) + (p_m^{(1)} - q_m^{(1)}\eta)\right] \\
&= (-1)^m \frac{p_{m-1}^{(1)} q_m^{(1)} - p_m^{(1)} q_{m-1}^{(1)}}{q_m^{(1)}} = \frac{1}{q_m^{(1)}}\,.
\end{aligned}$$

In particular, by $1/q_m^{(1)} \neq 0$, this proves that the function $\mathcal{E}$ is discontinuous at η.

It remains to prove that $\mathcal{E}$ is discontinuous at $\eta = 0$. Let $\xi_n^{(1)} := [0; n]$ and $\xi_n^{(2)} := [-1; 1, n]$. Then both sequences $(\xi_n^{(1)})_{n\geq 1}$ and $(\xi_n^{(2)})_{n\geq 1}$ tend to 0 for increasing n, but

$$\mathcal{E}(\xi_n^{(1)}) = \frac{1}{n} \to 0 \quad (n \to \infty),$$

wheras $\mathcal{E}(\xi_n^{(2)}) = 1$ holds for every positive integer n. Hence, Theorem 2 is proven. □

4 Proofs of Theorem 4 and Theorem 5

Proof of Theorem 4. Let m and $a_1, \ldots, a_m$ be positive integers. Set

$$\xi_1 = [0; a_1, \ldots, a_{m-1}, a_m]\,, \qquad \xi_2 = [0; a_1, \ldots, a_{m-1}, a_m + 1]\,.$$

Then we have $\xi_1 < \xi_2$ for even m and $\xi_2 < \xi_1$ otherwise. We define $I_m := (\xi_1, \xi_2)$ for even m and $I_m := (\xi_2, \xi_1)$ for odd m, which depend on $a_1, \dots, a_m$. The intervals I_m are disjoint for different m-tuples $(a_1, \dots, a_m)$. For any fixed m the union of all closed intervals $\overline{I}_m$ gives the interval $[0, 1]$. With this decomposition of $[0, 1]$ we obtain

$$\begin{aligned}
\int_0^1 \mathcal{E}(\alpha)\, d\alpha &= \int_0^1 \sum_{m=0}^{\infty} (-1)^m (q_m\alpha - p_m)\, d\alpha \\
&= \sum_{m=0}^{\infty} (-1)^m \int_0^1 (q_m\alpha - p_m)\, d\alpha \\
&= \frac{1}{2} + \sum_{m=1}^{\infty} (-1)^m \sum_{a_1=1}^{\infty} \cdots \sum_{a_m=1}^{\infty} \int_{I_m} (q_m\alpha - p_m)\, d\alpha \\
&= \frac{1}{2} + \sum_{m=1}^{\infty} \sum_{a_1=1}^{\infty} \cdots \sum_{a_m=1}^{\infty} \int_{\xi_1}^{\xi_2} (q_m\alpha - p_m)\, d\alpha \qquad (7)
\end{aligned}$$

and

$$\begin{aligned}
\int_0^1 \mathcal{E}^*(\alpha)\, d\alpha &= \int_0^1 \sum_{m=0}^{\infty} (q_m\alpha - p_m)\, d\alpha \\
&= \frac{1}{2} + \sum_{m=1}^{\infty} \sum_{a_1=1}^{\infty} \cdots \sum_{a_m=1}^{\infty} \int_{I_m} (q_m\alpha - p_m)\, d\alpha \\
&= \frac{1}{2} + \sum_{m=1}^{\infty} (-1)^m \sum_{a_1=1}^{\infty} \cdots \sum_{a_m=1}^{\infty} \int_{\xi_1}^{\xi_2} (q_m\alpha - p_m)\, d\alpha \qquad (8)
\end{aligned}$$

Every point $\alpha \in I_m$ satisfies $\alpha = [0; a_1, \dots, a_{m-1}, a_m, \dots]$, hence the convergents p_ν/q_ν for $\nu \le m$ depend on I_m, but not on $\alpha \in I_m$. Therefore, we derive

$$\int_{\xi_1}^{\xi_2} (q_m\alpha - p_m)\, d\alpha = (\xi_2 - \xi_1)\, \frac{(\xi_2 + \xi_1)q_m - 2p_m}{2}\,.$$

Using

$$\xi_1 = \frac{p_m}{q_m} \quad \text{and} \quad \xi_2 = \frac{(a_m + 1)p_{m-1} + p_{m-2}}{(a_m + 1)q_{m-1} + q_{m-2}}$$

we compute the expressions

$$\xi_2 - \xi_1 = \frac{(-1)^m}{(q_m + q_{m-1})q_m}$$

and

$$\xi_2 + \xi_1 = \frac{p_{m-1}q_m + q_{m-1}p_m + 2p_mq_m}{(q_m + q_{m-1})q_m}\,,$$

which give

$$\int_{\xi_1}^{\xi_2} (q_m\alpha - p_m)\, d\alpha = \frac{1}{2q_m(q_m + q_{m-1})^2}\,.$$

Substituting this integral into (7) and (8), we finally get the formulas stated in the theorem. □

Proof of Theorem 5. First we show that

$$\frac{1}{2}+\frac{1}{2}\sum_{m=1}^{\infty}\sum_{a_1=1}^{\infty}\cdots\sum_{a_m=1}^{\infty}\frac{1}{q_m\left(q_m+q_{m-1}\right)^2}=-\frac{3}{8}+\sum_{a=1}^{\infty}\frac{1}{a}\sum_{\substack{b=0\\ \gcd(a,b)=1}}^{a-1}\frac{1}{(a+b)^2}\,. \tag{9}$$

For the denominators of two subsequent convergents of the continued fraction expansion of $\alpha=[0;a_1,\ldots,a_m,\ldots]$ it is well-known that $\gcd(q_m,q_{m-1})=1$. For fixed $q_m=a$ we count the solutions of $q_{m-1}=b$ with $\gcd(a,b)=1$ and $0\le b\le a-1$ in the multiple sum on the left-hand side of (9). It is necessary to distinguish the cases $m\ge 2$ and $m=1$.
<u>Case 1:</u> $m\ge 2$. First let $a_1=1$. Then,

$$\frac{q_{m-1}}{q_m}=[0;a_m,\ldots,a_2,1]=[0;a_m,\ldots,a_2+1]\,.$$

For $a_1\ge 2$ we have

$$\frac{q_{m-1}}{q_m}=[0;a_m,\ldots,a_2,a_1]=[0;a_m,\ldots,a_2,a_1-1,1]\,.$$

<u>Case 2:</u> $m=1$. For $a_1=1$ we have a unique representation of the fraction

$$\frac{q_{m-1}}{q_m}=\frac{q_0}{q_1}=\frac{1}{a_1}=\frac{1}{1}=[0;1]\,,$$

since the integer part $a_0=0$ must not be changed. For $a_1\ge 2$ there are again two representations:

$$\frac{q_{m-1}}{q_m}=\frac{q_0}{q_1}=\frac{1}{a_1}=[0;a_1]=[0;a_1-1,1]\,.$$

Therefore it is clear that for fixed $q_m=a$ every b with $\gcd(a,b)=1$ and $0\le b\le a-1$ occurs exactly two times in the multiple sum on the left-hand side of (9), except for $m=1$ and $a_1=1$. For this exceptional case we separate the term

$$\frac{1}{2q_1(q_1+q_0)^2}=\frac{1}{8}$$

from the multiple sum. Then we obtain

$$\begin{aligned}
&\frac{1}{2}+\frac{1}{2}\sum_{m=1}^{\infty}\sum_{a_1=1}^{\infty}\cdots\sum_{a_m=1}^{\infty}\frac{1}{q_m(q_m+q_{m-1})^2}\\
&\qquad=\frac{1}{2}+\frac{1}{2}\sum_{m=2}^{\infty}\sum_{a_1=1}^{\infty}\cdots\sum_{a_m=1}^{\infty}\frac{1}{q_m(q_m+q_{m-1})^2}+\frac{1}{2}\sum_{a_1=2}^{\infty}\frac{1}{q_1(q_1+1)^2}+\frac{1}{8}\\
&\qquad=\frac{1}{2}+\sum_{a=1}^{\infty}\frac{1}{a}\sum_{\substack{b=1\\ \gcd(a,b)=1}}^{a-1}\frac{1}{(a+b)^2}+\frac{1}{8}\\
&\qquad=\frac{1}{2}+\sum_{a=1}^{\infty}\frac{1}{a}\sum_{\substack{b=0\\ \gcd(a,b)=1}}^{a-1}\frac{1}{(a+b)^2}-1+\frac{1}{8}\\
&\qquad=-\frac{3}{8}+\sum_{a=1}^{\infty}\frac{1}{a}\sum_{\substack{b=0\\ \gcd(a,b)=1}}^{a-1}\frac{1}{(a+b)^2},
\end{aligned}$$

which proves the identity in (9).

Next we treat the double sum on the right-hand side of (9). Let μ denote the Möbius function. Then we derive

$$\begin{aligned}
\sum_{a=1}^{\infty}\frac{1}{a}\sum_{\substack{b=0\\ \gcd(a,b)=1}}^{a-1}\frac{1}{(a+b)^2}&=\sum_{a=1}^{\infty}\frac{1}{a}\sum_{b=0}^{a-1}\sum_{\substack{d>0\\ d\mid\gcd(a,b)}}\frac{\mu(d)}{(a+b)^2}=\sum_{a=1}^{\infty}\sum_{b=0}^{a-1}\sum_{\substack{d>0\\ d\mid a\,\wedge\, d\mid b}}\frac{\mu(d)}{a(a+b)^2}\\
&=\sum_{d=1}^{\infty}\sum_{\substack{a=1\\ d\mid a}}^{\infty}\sum_{\substack{b=0\\ d\mid b}}^{a-1}\frac{\mu(d)}{a(a+b)^2}=\sum_{d=1}^{\infty}\sum_{n=1}^{\infty}\sum_{m=0}^{n-1/d}\frac{\mu(d)}{nd(nd+md)^2}
\end{aligned}$$

$$\begin{aligned}
&=\sum_{d=1}^{\infty}\frac{\mu(d)}{d^3}\sum_{n=1}^{\infty}\sum_{m=0}^{n-1}\frac{1}{n(n+m)^2}=\frac{1}{\zeta(3)}\sum_{a=1}^{\infty}\sum_{b=0}^{a-1}\frac{1}{a(a+b)^2}\\
&=\frac{1}{\zeta(3)}\sum_{a=1}^{\infty}\frac{1}{a}\sum_{c=a}^{2a-1}\frac{1}{c^2}=\frac{1}{\zeta(3)}\sum_{c=1}^{\infty}\frac{1}{c^2}\sum_{a=\lfloor c/2\rfloor+1}^{c}\frac{1}{a}\\
&=\frac{1}{\zeta(3)}\sum_{c=1}^{\infty}\frac{1}{c^2}\sum_{a=1}^{c}\frac{(-1)^{a+1}}{a}\\
&=\frac{1}{\zeta(3)}\sum_{c=1}^{\infty}\frac{1}{c^2}\sum_{a=1}^{c-1}\frac{(-1)^{a+1}}{a}+\frac{1}{\zeta(3)}\sum_{c=1}^{\infty}\frac{(-1)^{c+1}}{c^3}\\
&=-\frac{\zeta(2,-1)}{\zeta(3)}+\frac{3}{4},
\end{aligned}$$

where

$$\zeta(2,-1)=\sum_{c=1}^{\infty}\frac{1}{c^2}\sum_{a=1}^{c-1}\frac{(-1)^a}{a}=\sum_{c>a>0}\frac{(-1)^a}{ac^2}$$

is a special case of the multivariate zeta function (see [1, Section 2.6]), satisfying

$$\zeta(2,-1)=\zeta(3)-\frac{3}{2}\zeta(2)\log 2\,.$$

Collecting together we obtain from (9) that

$$\int_0^1 \mathcal{E}(\alpha)\,d\alpha=-\frac{3}{8}-1+\frac{3}{4}+\frac{3}{2}\frac{\zeta(2)\log 2}{\zeta(3)}=-\frac{5}{8}+\frac{3\zeta(2)\log 2}{2\zeta(3)}\,,$$

which completes the proof of the theorem.

□

Remark 7. Let n be a positive integer. We consider a modified error sum function given by

$$\sum_{m\geq 0}|\alpha q_m-p_m|^n \qquad (0<\alpha<1)\,.$$

By similar methods as used to deduce Theorems 4 and 5 we obtain the following identities:

$$\begin{aligned}\int_0^1\sum_{m\geq 0}|\alpha q_m-p_m|^n\,d\alpha &= \frac{1}{n+1}+\frac{1}{n+1}\sum_{m=1}^{\infty}\sum_{a_1=1}^{\infty}\cdots\sum_{a_m=1}^{\infty}\frac{1}{q_m(q_m+q_{m-1})^{n+1}}\\ &=\frac{1}{n+1}\left(1-\frac{1}{2^{n+1}}-\frac{2\zeta(n+1,-1)}{\zeta(n+2)}\right)\end{aligned}$$

with the multivariate zeta function $\zeta(n+1,-1)$ defined by

$$\zeta(n+1,-1)=\sum_{m_2>m_1>0}\frac{(-1)^{m_1}}{m_1 m_2^{n+1}}\,.$$

This yields an asymptotic expansion, namely

$$\int_0^1\sum_{m\geq 0}|\alpha q_m-p_m|^n\,d\alpha=\frac{1}{n+1}+\mathcal{O}\left(\frac{1}{(n+1)2^n}\right) \qquad (n\to\infty)\,.$$

5 Generalization of the error sum function $\mathcal{E}$

In this section we show that the error sum function $\mathcal{E}$ is the special case of a more general concept involving the theory of approximation with algebraic numbers of bounded degree. We need some notations to recall the definition of the Mahler functions $w_n(H,\alpha)$ and $w_n(\alpha)$. For more details on this function we refer to [5].

For any polynomial $P(x)\in\mathbb{Z}[x]$ we denote by $H(P)$ the *height* of the polynomial P, which is given by the maximum value of the modulus of the coefficients. Let n, H be positive integers and $\alpha \in \mathbb{C}$ with $|\alpha| \le 1/2$ and $\deg \alpha > n$. For α being transcendental we define $\deg \alpha = \infty$. Set

$$w_n(H,\alpha) \quad := \quad \min_{\substack{P \in \mathbb{Z}[x] \setminus \{0\} \\ \deg P \le n \\ H(P) \le H}} \big|P(\alpha)\big| \,,$$

$$w_n(\alpha) \quad := \quad \limsup_{H \to \infty} \frac{-\log w_n(H,\alpha)}{\log H} \,.$$

$w_n(\alpha)$ is the largest positive real number such that for every $\varepsilon > 0$ there are infinitely many polynomials P from $\mathbb{Z}[x]$ of degree at most n satisfying

$$\big|P(\alpha)\big| \,<\, \big(H(P)\big)^{-w_n(\alpha)+\varepsilon} \,.$$

So the function $w_n(H,\alpha)$ is needed to define the important Mahler function $w_n(\alpha)$. From the definition of $w_n(H,\alpha)$ it follows immediately that $w_1(H,\alpha) \ge w_2(H,\alpha) \ge \cdots \ge w_n(H,\alpha)$ holds for all integers $n = 1, 2, \dots$.

Given α and some positive integer n with $\deg \alpha > n$, there is a unique sequence $(H_m)_{m \ge 0}$ of positive integers satisfying the following conditions:

$$\begin{array}{ll} (i) & 1 = H_0 < H_1 < \cdots < H_m < \dots \\ (ii) & w_n(H_0,\alpha) > w_n(H_1,\alpha) > \cdots > w_n(H_m,\alpha) > \dots \\ (iii) & w_n(H_m,\alpha) = w_n(H_{m+1}-1,\alpha) \qquad (m = 0, 1, \dots) \end{array}$$

We define the generalized error sum function

$$\mathcal{E}_n(\alpha) \,:=\, \sum_{m=0}^{\infty} w_n(H_m,\alpha) \,.$$

Note that $\mathcal{E}_n(\alpha) = \mathcal{E}_n(-\alpha)$ holds, since the same is obviously true for the Mahler function: $w_n(H,\alpha) = w_n(H,-\alpha)$. For $n = 1$ and $\alpha \in (-1/2, 1/2) \setminus \mathbb{Q}$ we have $p_0/q_0 \in \{-1/1, 0/1\}$ and $p_1/q_1 = 1/a_1$, where $a_1 = 1$ holds if and only if $-1/2 < \alpha < 0$. This implies that

$$w_1(H_m,\alpha) \;=\; \begin{cases} |q_m\alpha - p_m|, & \text{if } 0 < \alpha < 1/2; \,, \\ |q_{m+1}\alpha - p_{m+1}|, & \text{if } -1/2 < \alpha < 0; \end{cases} \qquad (m = 0, 1, \dots) \,.$$

Therefore,

$$\mathcal{E}_1(\alpha) \;=\; \begin{cases} \mathcal{E}(\alpha), & \text{if } 0 < \alpha < 1/2; \,, \\ \mathcal{E}(\alpha) - \alpha - 1, & \text{if } -1/2 < \alpha < 0; \,, \end{cases}$$

where $\alpha + 1$ equals $q_0\alpha - p_0$ in the second case. Let

$$\mathcal{E}_n \,:=\, \sup\big\{\mathcal{E}_n(\alpha) \,:\, \alpha \in (-1/2, 1/2) \,\wedge\, \deg \alpha > n\big\} \qquad (n = 1, 2, \dots) \,.$$

Then it is clear that for $n = 1, 2, \dots$

$$\begin{aligned} \mathcal{E}_n \;&=\; \sup\Big\{\sum_{m=0}^{\infty} w_n(H_m,\alpha) \,:\, \alpha \in (-1/2, 1/2) \,\wedge\, \deg \alpha > n\Big\} \\ &\le\; \sup\Big\{\sum_{m=0}^{\infty} w_1(H_m,\alpha) \,:\, \alpha \in (-1/2, 1/2) \,\wedge\, \deg \alpha > n\Big\} \\ &\le\; \mathcal{E}_1 \,=\, \sup\big\{\mathcal{E}_1(\alpha) \,:\, \alpha \in (-1/2, 1/2) \,\wedge\, \deg \alpha > 1\big\} \\ &\le\; \rho \,. \end{aligned}$$

This bound can be improved by applying two inequalities based on Siegel's Lemma. Let $\alpha \in \mathbb{C}$ with $|\alpha| < 1/2$. For real α and any positive integers n, H we have

$$w_n(H,\alpha) < (n+1)H^{-n}\,. \tag{10}$$

For $\alpha \notin \mathbb{R}$ and any positive integers n, H we have

$$w_n(H,\alpha) < \sqrt{2}(n+1)H^{-(n-1)/2}\,. \tag{11}$$

These inequalities can be found on page 69 in [5], where the constants C_1 and C_2 are given by [5, Hilfssatz 27, Hilfssatz 28]. In what follows we distinguish whether α is real or not.

Case 1: $\alpha \in \mathbb{R}$. By using

$$w_n(1,\alpha) \le \max_{-1/2\le x\le 1/2} |x^n| = \frac{1}{2^n}$$

we obtain with (10) and the Riemann zeta function for $n \ge 2$

$$\begin{aligned}
\mathcal{E}_n(\alpha) &= \sum_{m=0}^{\infty} w_n(H_m,\alpha) = w_n(1,\alpha) + \sum_{m=1}^{\infty} w_n(H_m,\alpha) \\
&\le \frac{1}{2^n} + \sum_{m=1}^{\infty} w_n(m+1,\alpha) \le \frac{1}{2^n} + \sum_{m=1}^{\infty} \frac{n+1}{(m+1)^n} \\
&= \frac{1}{2^n} + (n+1)\big(\zeta(n)-1\big) \quad \to 0 \qquad (\text{for } n\to\infty)\,.
\end{aligned}$$

Case 2: $\alpha \notin \mathbb{R}$. Here we consider the polynomial z^n for $|z| \le 1/2$. Then,

$$w_n(1,\alpha) \le \max_{|z|\le 1/2} |z^n| = \frac{1}{2^n}\,.$$

With (11) we repeat the arguments from Case 1 for $n \ge 4$:

$$\begin{aligned}
\mathcal{E}_n(\alpha) &\le w_n(1,\alpha) + \sum_{m=1}^{\infty} \frac{\sqrt{2}(n+1)}{(m+1)^{(n-1)/2}} \\
&\le \frac{1}{2^n} + \sqrt{2}(n+1)\left(\zeta\left(\frac{n-1}{2}\right) - 1\right) \quad \to 0 \qquad (\text{for } n\to\infty)\,.
\end{aligned}$$

Note that the inequality

$$\frac{1}{2^n} + (n+1)\big(\zeta(n)-1\big) < \rho$$

holds for $n \ge 3$, whereas

$$\frac{1}{2^n} + \sqrt{2}(n+1)\left(\zeta\left(\frac{n-1}{2}\right) - 1\right) < \rho$$

is true for $n \ge 5$.

6 Acknowledgments

The authors are much obliged to Mr. H. A. ShahAli and Professor A. Ustinov. Mr. ShahAli has drawn our attention to the paper [7] of J. N. Ridley and G. Petruska. Professor A. Ustinov, who joined the second author at the Journées Arithmétiques 2011 in Vilnius, has given useful hints to compute the value of the integral in Theorem 5.

References

[1] D. H. Bailey, J. M. Borwein, N. J. Calkin, R. Girgensohn, D. R. Luke, and V. H. Moll, *Experimental Mathematics in Action*, A. K. Peters, 2007.

[2] C. Elsner, Series of error terms for rational approximations of irrational numbers, *J. Integer Sequences* **14** (2011), Article 11.1.4.

[3] C. Elsner and M. Stein, On the value distribution of error sums for approximations with rational numbers, submitted.

[4] A. Khintchine, *Kettenbrüche*, Teubner, 1956.

[5] Th. Schneider, *Einführung in die transzendenten Zahlen*, Springer-Verlag, 1957.

[6] O. Perron, *Die Lehre von den Kettenbrüchen*, Chelsea Publishing Company, 1929.

[7] J. N. Ridley and G. Petruska, The error-sum function of continued fractions, *Indag. Mathem., N.S.*, **11** (2), 2000, 273–282.

[8] A. B. Shidlovskii, *Transcendental Numbers*, de Gruyter, 1989.

2010 *Mathematics Subject Classification*: Primary 11J04; Secondary 11J70, 11B05, 11B39.
Keywords: continued fractions, convergents, approximation of real numbers, error terms, density.

(Concerned with sequences A000045, A007676, A007677, A041008, and A041009.)

Received May 13 2011; revised versions received July 12 2011; August 17 2011. Published in *Journal of Integer Sequences*, September 25 2011.

Return to Journal of Integer Sequences home page.