A GEOMETRICAL APPROACH TO APPROXIMATIONS BY CONTINUED FRACTIONS

H. G. KOPETZKY and F. J. SCHNITZER

Abstract

By simple geometrical considerations new proofs for some classical results are given and also new theorems about approximation by continued fractions are derived. This geometrical approach presents an instructive visualisation of the nature of approximation theorems.

1980 *Mathematics subject classification (Amer. Math. Soc.)*: 10 F 20.

In our note we present a geometrical method which will lead to very simple proofs of well-known results about approximations by continued fractions. Furthermore we shall be able to obtain a number of new theorems almost effortlessly. Our geometrical ideas are quite different from those of older authors such as Klein, Humbert or Züllig (see Koksma's book [8] where additional references can be found). Contrary to those methods we are not interested in geometrical represen- tations of a certain continued fraction but in the study of regions which correspond to numbers with certain approximation properties given by their continued fraction expansion. At the same time this geometrical approach pre- sents an instructive visualisation of the nature of these approximation theorems.

Let us begin with some notation. For a real number ξ let $\xi = [a_0, a_1, \ldots]$ denote the regular continued fraction expansion of ξ and let p_n/q_n stand for the nth convergent to ξ. We shall be mainly concerned with the study of numbers λ_n which describe the approximation behaviour of continued fractions and which are defined by the expression

$$\left| \xi - \frac{p_n}{q_n} \right| = \frac{1}{\lambda_n q_n^2}$$

where p_n, q_n, λ_n obviously depend on ξ. We shall frequently use the elementary but fundamental relation

$$\lambda_n = a_{n+1} + [0, a_n, a_{n-1}, \ldots, a_1] + [0, a_{n+2}, a_{n+3}, \ldots].$$

In the sequel the real number ξ will be assumed to be irrational. This simplifies our formulations; if rational ξ's are not excluded one has to assume the existence of sufficiently many convergents to ξ.

The well known theorem of Hurwitz [7] which states that the inequality $|\xi - p/q| < 1/\sqrt{5}\, q^2$ has infinitely many integer solutions p, q for any real irrational number ξ can be proved by Borel's theorem [2]: At least one of three consecutive convergents to ξ satisfies the inequality above. In our notation Borel's result can be formulated in this way:

THEOREM 1. *For any $n \geq 1$ one has*
$$\max(\lambda_{n-1}, \lambda_n, \lambda_{n+1}) > \sqrt{5}\, .$$

There exist several generalizations of this theorem two of which are due to Humbert [6] and Fujiwara [3]. A very general form of it is due to Bagemihl and McLaughlin [1]:

THEOREM 2. *If $a_{n+1} \geq k$, for some $n \geq 1$, then*
$$\max(\lambda_{n-1}, \lambda_n, \lambda_{n+1}) > \sqrt{k^2 + 4}\, .$$

There exist also results involving only two convergents to ξ, of which the oldest one (due to Vahlen [13]) says:

THEOREM 3. *For any $n \geq 1$ one has*
$$\max(\lambda_{n-1}, \lambda_n) > 2.$$

This theorem was extended by Fujiwara [4] and brought into this final form (again in [1]) by Bagemihl and McLaughlin:

THEOREM 4. *If $a_n \geq k$ or $a_{n+1} \geq k$ for some $n \geq 1$, then*
$$\max(\lambda_{n-1}, \lambda_n) > k + \frac{1}{k}\, .$$

This is often formulated in the following equivalent form

THEOREM 4'. *If $a_{n+1} \geq k$ for some $n \geq 1$, then*
$$\max(\lambda_n, \min(\lambda_{n-1}, \lambda_{n+1})) > k + \frac{1}{k}\, .$$

The important question whether the constants involved are best possible can be decided in all these theorems. In Theorems 1 to 3 no improvements are possible as the irrational $\xi = [k, k, \ldots]$ shows, a fact first noticed by Hurwitz [7] and the authors mentioned above. That the constant in Theorem 4 is best possible, is proved in [9].

There exist several proofs for these results by many different authors which underlines the interest this topic has found.

Recently J. Tong [10] searched for estimations of the λ_n from above. He found results corresponding to Theorems 1 and 2 and called what he obtained the "conjugate property" of these theorems.

THEOREM 5. *If* $a_{n+1} \leqslant k$ *for some* $n \geqslant 1$, *then*

$$\min(\lambda_{n-1}, \lambda_n, \lambda_{n+1}) < \sqrt{k^2 + 4}.$$

Further results which can also be treated by our method have been derived by J. Tong in [11].

After these preliminaries we first explain our geometrical approach in case of Theorems 1 and 2. For $\xi = [a_0, a_1, \ldots]$ irrational we use the abbreviations

$$x_n = [0, a_n, a_{n-1}, \ldots, a_1],$$
$$y_n = [0, a_{n+2}, a_{n+3}, \ldots].$$

Thus x_n is rational while y_n is always irrational by our assumption on ξ and hence $0 < y_n < 1$. For convenience we assume $0 < x_n < 1$. The case $x_n = 1$ can occur only for $n = 1$ and $a_1 = 1$ and can be handled easily; thus $x_n < 1$ is no serious restriction. Conversely, for any pair of reals (α, β) with $0 < \alpha, \beta < 1$ and for arbitrary small $\varepsilon > 0$ there exist a real number ξ and a nonnegative integer $n = n(\varepsilon)$ such that $y_n = \beta$ and $|x_n - \alpha| < \varepsilon$. We shall use this correspondence repeatedly. Obviously we have the relation

$$\lambda_n = a_{n+1} + x_n + y_n.$$

Using the definition of continued fractions one obtains the relations

$$\lambda_{n-1} = \frac{1}{x_n} + \frac{1}{a_{n+1} + y_n} \quad \text{and} \quad \lambda_{n+1} = \frac{1}{y_n} + \frac{1}{a_{n+1} + x_n}.$$

Next, we define the following regions in the plane

$$B_0 = B_0(\lambda, k) = \{(x, y) \mid k + x + y \geqslant \lambda\},$$

$$B_{-1} = B_{-1}(\lambda, k) = \left\{(x, y) \mid \frac{1}{x} + \frac{1}{k + y} \geqslant \lambda\right\},$$

$$B_1 = B_1(\lambda, k) = \left\{(x, y) \mid \frac{1}{y} + \frac{1}{k + x} \geqslant \lambda\right\}.$$

Denote the boundary curves of these regions B_i by $C_i = C_i(\lambda, k)$, $i = -1, 0, 1$.

It is clear that with $k = a_{n+1}$ one has $\lambda_n \geqslant \lambda$ if and only if the corresponding point (x_n, y_n) is in $B_0(\lambda, k)$. The same holds for λ_{n-1} together with B_{-1} and for λ_{n+1} together with B_1.

Therefore, since $0 < x, y < 1$, Theorems 1 and 2 (where equality still is possible; this possibility will be excluded by the argument below) will be proved if one shows that the open unit square $S = \{(x, y) \mid 0 < x, y < 1\}$ can be covered by the union of the regions $B_i(\sqrt{k^2 + 4}, k)$, $k = 1, 2, \ldots$, $i = -1, 0, 1$. (Actually the whole plane will be covered).

The following figures reveal the simple geometrical nature of the necessary considerations whilst at the same time they present an instructive visualisation of the content of the theorems.
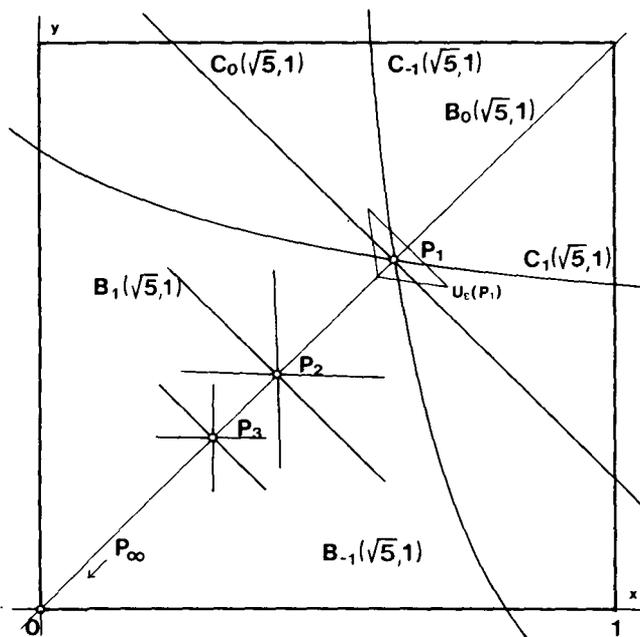

FIGURE 1. Theorems 1, 2

The curves $C_i(\lambda, k)$, $i = -1, 0, 1$, intersect in exactly one point $P = P_k$, $k = 1, 2, \ldots$, in S if and only if $\lambda = \sqrt{k^2 + 4}$ since, by symmetry, we may assume $x = y$ and then obtain

$$\lambda = k + 2x, \qquad \lambda = \frac{1}{x} + \frac{1}{k + x}$$

which gives immediately $\lambda^2 = k^2 + 4$. Substituting the coordinates of suitable points shows that the regions B_j lie on appropriate sides of the curves C_i, $i = -1, 0, 1$.

The point P_k has the coordinates

$$x = y = \tfrac{1}{2}\left(-k + \sqrt{k^2 + 4}\right) = [0, \overline{k}]$$

(the bar indicates periodicity of the continued fraction expansion). From this expression we learn three facts:

(i) The point P_k cannot correspond to an admissible number because x is not rational as it would have to be. Thus equality can be excluded in the formulation of the theorem.

(ii) If $\lambda = \sqrt{k^2 + 4}$ is replaced by $\lambda = \sqrt{k^2 + 4} + \varepsilon$, $\varepsilon > 0$, a certain neighborhood $U_\varepsilon(P_k)$ of P_k will not be covered by the union of the three regions which shows that the constants in Theorems 1 and 2 are best possible.

(iii) Numbers which correspond to points in $U_\varepsilon(P_k)$ are exactly those for which the constant in the theorem cannot be larger then $\sqrt{k^2 + 4} + \varepsilon$. Since $\varepsilon > 0$ is arbitrary we see for example that the constant $\sqrt{5}$ in the theorem of Hurwitz cannot be improved for exactly those reals which are equivalent to $\tfrac{1}{2}(\sqrt{5} - 1) = [0, \overline{1}]$.

At the same time Theorem 5 (the "conjugate property" of Theorem 2) is obtained by the observation that the square S is also covered by the complements of the B_i together with the C_i for $i = -1, 0, 1$.

By the same method we can obtain an "asymmetric" version of Theorems 1 and 2.

THEOREM 6. *For $a_{n+1} \geqslant k$ for some $n \geqslant 1$ and real numbers λ', λ'' which satisfy*

$$k < \lambda' \leqslant k + 2, \qquad \lambda'' = \frac{2}{\lambda' - k} + \frac{2}{\lambda' + k}$$

one has

$$\lambda_n > \lambda' \quad or \quad \max(\lambda_{n-1}, \lambda_{n+1}) > \lambda''.$$

PROOF. If we choose an admissible λ' instead of λ in the definition of B_0 and the corresponding λ'' instead of λ in the definitions of B_{-1} and B_1 then the claims follow since S is again covered by the union of the three regions which can be seen by taking $x = y$ as above (a picture would look quite similar to Figure 1 with the points P_k shifted on the line $x = y$ depending on the chosen value of λ').

The "conjugate property" is again obvious. Numbers for which no improvement of the constant is possible (in the sense that $\lambda' \geqslant \lambda_0'$, $\lambda'' \geqslant \lambda_0''$ with given λ_0' and λ_0'' and related as in the theorem) depend on λ_0' and can be determined from the point of intersection as before. Clearly one can obtain a similar theorem in which three different constants are involved.

The numbers $\sqrt{5}$ (in Theorems 1 and 2) and $\sqrt{8}$ (in Theorem 2) are the two largest values of the socalled Markoff spectrum which is the set of numbers (see [8])

$$\mu(\xi) = \limsup_{n \to \infty} \lambda_n.$$

To prove this and to obtain the third value of the Markoff spectrum, namely $\sqrt{221}/5$, additional considerations are necessary. To apply our method we need appropriate expressions for λ_{n-2} and λ_{n+2} where additional partial quotients will occur. We get from the fundamental relation for λ_n

$$\lambda_{n-2} = \cfrac{1}{-a_n + \cfrac{1}{x_n}} + \cfrac{1}{a_n + \cfrac{1}{a_{n+1} + y_n}},$$

$$\lambda_{n+2} = \cfrac{1}{-a_{n+2} + \cfrac{1}{y_n}} + \cfrac{1}{a_{n+2} + \cfrac{1}{a_{n+1} + x_n}}.$$

According to these expressions we define the regions B_{-2}, B_2 (and as above the corresponding boundary curves C_{-2}, C_2):

$$B_{-2} = B_{-2}(\lambda, j, k) = \left\{ (x, y) \mid \cfrac{1}{-j + \cfrac{1}{x}} + \cfrac{1}{j + \cfrac{1}{k + y}} \geqslant \lambda \right\}$$

$$B_2 = B_2(\lambda, k, l) = \left\{ (x, y) \mid \cfrac{1}{-l + \cfrac{1}{y}} + \cfrac{1}{l + \cfrac{1}{k + x}} \geqslant \lambda \right\}.$$

If $a_i \geqslant 3$ then $\lambda_{i+1} \geqslant 3 > \sqrt{221}/5$; hence we may assume $a_i < 3$ for all i involved. The special cases when all partial quotients are either 1 or 2 have already been treated. Therefore we assume that both 1 and 2 occur. We take

$a_{n+1} = k = 1$ (otherwise rename the subscripts) and obtain the three possibilities: $(j, l) = (1, 2), (2, 1), (2, 2)$.

In the third case one has $\lambda_{n+1} > 2 + [0, 1, 2] + [0, 1, 1] > 3$ or $\lambda_{n+1} > 2 + [0, 1, 2] + [0, 2, 1] = 3$ according to $a_{n+3} = 1$ or $= 2$.

The second case is the first in reverse order and can thus be treated in the same way (since B_2 can be obtained from B_{-2} by changing both the variables x, y and the parameters j, l). In the first case ($j = 1, l = 2$) we take the curves $C_i(\sqrt{211}/5)$, $i = -2, 1, 2$ and go through similar considerations as before.

But we have to observe that assumptions on $a_n (= j)$ and $a_{n+2} (= l)$ possibly reduce the range of x and y. In our case we obtain from $l = 2$ the inequality $y_n < 1/2$ and have therefore to consider only the rectangle $S^* = \{(x, y) | 0 < x < 1, 0 < y < 1/2\}$ (from $j = 1$ we get $1/2 < x$; but this is of no importance). With slightly more calculation we find that these curves intersect in exactly one point in S^* whose coordinates are $((\sqrt{211} - 5)/14, (\sqrt{211} - 9)/14)$. S^* is in fact covered by the three regions; this can be seen again by substituting three appropriate points into the defining inequalities. The situation is shown in Figure 2.
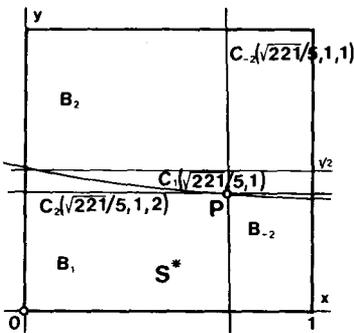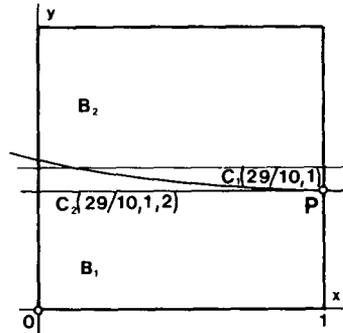
FIGURE 2. Theorem 7   FIGURE 3. Theorem 9

Thus we have proved

THEOREM 7. *If $a_{n+1} = 1$ and $a_{n+2} \geq 2$ for some $n \geq 2$, then*
$$\max(\lambda_{n-2}, \lambda_{n+1}, \lambda_{n+2}) > \sqrt{221}/5,$$
and if $a_{n+1} = 1$ and $a_n \geq 2$ for some $n \geq 2$, then
$$\max(\lambda_{n-2}, \lambda_{n-1}, \lambda_{n+2}) > \sqrt{221}/5.$$

The "conjugate property" is obvious again. (To write down that property one has to exchange the functions max and min and to alter the directions of the

inequalities.) By using the coordinates of the point of intersection we get the numbers for which the constant λ is best possible and obtain

$$y = \left(\sqrt{221} - 9/14\right) = \left[0, \overline{2,2,1,1}\right].$$

The content of Theorem 7 can then be used for the straightforward derivation of the following theorem (which has clearly a "conjugate property" again).

THEOREM 8. *If* $(a_n, a_{n+1}, a_{n+2}) \neq (k, k, k)$ *for some* $n \geqslant 2$ *with* $k = 1, 2$, *then*
$$\max(\lambda_i) > \sqrt{221}/5, \qquad i = n - 2, \ldots, n + 2.$$

From this and earlier remarks we infer that $\sqrt{221}/5$ is the third value of the Markoff spectrum.

Theorems 7 and 8 were first proved by Fukasawa [5].

The region B_{-2} is needed to cover only a very small portion of S^*: If we decrease λ somewhat we are able to cover our square S by the regions B_1 and B_2 alone (see Figure 2 again). The curves C_1 and C_2 intersect then on the line $x = 1$ with the common value $\lambda = 29/10$ for $k = 1$ and $l = 2$. If we observe that $a_{n+2} \geqslant 3$ implies $\lambda_{n+1} > 3$ we have

THEOREM 9. *If* $a_{n+1} = 1$ *and* $a_{n+2} \geqslant 2$ *for some* $n \geqslant 0$, *then*
$$\max(\lambda_{n+1}, \lambda_{n+2}) > 29/10.$$

This result can also be obtained by the use of λ_n and λ_{n+1} observing that $1/2 \leqslant x < 1$ holds since in that case one has to assume $a_n = 1$.

By interchanging x and y and renumbering the subscripts we get

THEOREM 9'. *If* $a_{n+1} \geqslant 2$ *and* $a_n = a_{n+2} = 1$ *for some* $n \geqslant 1$, *then*
$$\max(\lambda_n, \min(\lambda_{n-1}, \lambda_{n+1})) > 29/10.$$

These theorems are closely related to the theorems of Vahlen and Fujiwara and generalizations of these. We shall now discuss them from our geometrical point of view.

We have to cover our square S by the intersection of B_{-1} and B_1 and by B_0 for an appropriate value of λ. If $a_{n+1} = k$ we find a common point of $C_{-1}(\lambda)$ and $C_0(\lambda)$ on the line $y = 0$ for $\lambda = k + 1/k$ with $x = 1/k$. The second point of intersection for this λ is $x = k$ which lies outside of S. By symmetry we have an analogous situation for C_1 and C_0. Therefore Theorem 4' (and 4) is proved if one observes that $(0, 0) \in B_{-1} \cap B_1$, $(1, 1) \in B_0$ and $(1, 0) \in B_1$.

To obtain the "conjugate property" of Theorem 4' a slight modification is necessary. Simple considerations show that one needs the intersection of C_{-1} and

C_0 on the line $y = 1$ and the intersection of C_1 and C_0 on the line $x = 1$ which gives $\lambda = k + 1 + 1/(k + 1)$. As above the second point of intersection lies outside of S; the corresponding regions are on the same sides of the curves just as before. Hence we have

THEOREM 10. *If* $a_{n+1} \geqslant k$ *for some* $n \geqslant 1$, *then*

$$\min(\lambda_n, \max(\lambda_{n-1}, \lambda_{n+1})) < k + 1 + \frac{1}{k + 1}.$$
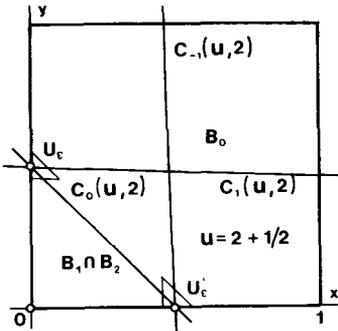
FIGURE 4. Theorems 4′, 4″   FIGURE 5. Theorem 10
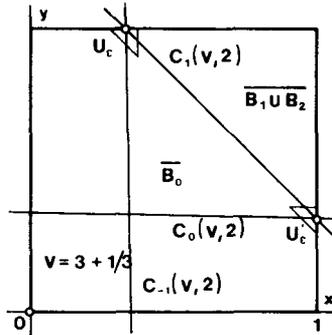
It should now be clear how one can obtain an "asymmetric" version of Theorem 4′ (and of Theorem 10, too). The corresponding result in the first case is

THEOREM 11. *If* $a_{n+1} \geqslant k$ *for some* $n \geqslant 1$, *then for any pair of real numbers* λ' *and* λ'' *with*

$$k < \lambda' < k + 2, \quad \lambda'' = \frac{1}{k} + \frac{1}{\lambda' - k}$$

one has

$$\lambda_n > \lambda' \quad or \quad \max(\lambda_{n-1}, \lambda_{n+1}) > \lambda''.$$

It is not difficult to determine the numbers for which (in case of Theorems 4, 4′, 9, 9′ and 10) the constants cannot be improved. These numbers are again given by the points of intersection of the curves C_i. But now one has to observe that these points have rational coordinates. Since we are dealing only with irrational numbers we have to choose points with at least one irrational coordinate in a sufficiently small neighbourhood. From this remark we obtain the assertions but formulate only the case of Theorem 4′ (first proved in [9] by a different method; but see also the paper [12] by J. Tong which contains a minor correction to [9]).

The remaining cases can be treated analogously. From the continued fraction expansion of the coordinates $x = 1/k + \varepsilon_1$, $y = \varepsilon_2$ and $x = \varepsilon_2$, $y = 1/k + \varepsilon_1$ with $|\varepsilon_1|$ and $\varepsilon_2 > 0$ sufficiently small we obtain

THEOREM 4″. *Theorem 4′ with $a_{n+1} = k$ becomes false if the constant $k + 1/k$ is replaced by $k + 1/k + \delta$, $\delta > 0$, in exactly one of the following cases where the constants A_i have to be choosen sufficiently large depending on the value of δ:*

1. $a_{n+2} > A_1$,   1.1 $a_n = k - 1$, $a_{n-1} = 1$, $a_{n-2} > A_2$,
                        1.2 $a_n = k$, $a_{n-1} > A_3$.
2. $a_n > A_4$,   2.1 $a_{n+2} = k - 1$, $a_{n+3} = 1$, $a_{n+4} > A_5$,
                        2.2 $a_{n+2} = k$, $a_{n+3} > A_6$.

As a final remark it seems worth mentioning that by the use of this method and by more involved computations additional results about approximations by continued fractions should be derivable. Also new proofs for other results which can be found in the references below should be obtainable.

Acknowledgement

References

[1] F. Bagemihl and J. R. McLaughlin, 'Generalization of some classical theorems concerning triples of consecutive convergents to simple continued fractions,' *J. Reine Angew. Math.* **221** (1966), 146–149.

[2] É. Borel, 'Contribution à l'analyse arithmétique du continu', *J. Math. Pures Appl.* (5) **9** (1903), 329–375.

[3] M. Fujiwara, 'Bemerkung zur Theorie der Approximation der irrationalen Zahlen durch rationale Zahlen', *Tôhoku Math. J.* **11** (1916), 239–242.

[4] M. Fujiwara, 'Bemerkung zur Theorie der Approximation der irrationalen Zahlen durch rationale Zahlen', *Tôhoku Math. J.* **14** (1918), 109–115.

[5] S. Fukasawa, 'Über die Kleinsche geometrische Darstellung des Kettenbruchs', *Japan. J. Math.* **2** (1926), 101–114.

[6] G. Humbert, 'Remarques sur certaines suites d'approximation', *J. Math. Pures Appl.* (7) **2** (1916), 155–167.

[7] A. Hurwitz, 'Über die angenäherte Darstellung der Irrationalzahlen durch rationale Brüche', *Math. Ann.* **39** (1891), 279–284.

[8] J. F. Koksma, *Diophantische Approximationen* (Springer-Verlag, Berlin, Heidelberg, New York, 1974 (Reprint)).

[9] H. G. Kopetzky und F. J. Schnitzer, 'Bemerkungen zu einem Approximationssatz für regelmässige Kettenbrüche', *J. Reine Angew. Math.* **293 / 294** (1977), 437–440.

[10] J. Tong, 'The conjugate property of the Borel theorem on Diophantine approximation', *Math. Z.* **184** (1983), 151–153.

[11] J. Tong, 'A generalization of the Borel theorem in Diophantine approximation', *Riv. Mat. Univ. Parma*, *IV. Ser.* **9** (1983), 121–124.

[12] J. Tong, 'On two theorems of Kopetzky and Schnitzer on the approximation of continued fractions', *J. Reine Angew. Math.* **362** (1985), 1–3.

[13] K. Th. Vahlen, 'Über Näherungswerte und Kettenbrüche', *J. Reine Angew. Math.* **115** (1859), 221–233.

Institut für Mathematik
und Angewandte Geometrie
Montanuniversität Leoben
A-8700 Leoben, Austria