COMPLEX NUMBERS WITH BOUNDED PARTIAL QUOTIENTS

WIEB BOSMA AND DAVID GRUENEWALD

ABSTRACT. Conjecturally, the only real algebraic numbers with bounded partial quotients in their regular continued fraction expansion are rationals and quadratic irrationals. We show that the corresponding statement is not true for complex algebraic numbers in a very strong sense, by constructing for every even degree d algebraic numbers of degree d that have bounded complex partial quotients in their Hurwitz continued fraction expansion. The Hurwitz expansion is the generalization of the nearest integer continued fraction for complex numbers. In the case of real numbers the boundedness of regular and nearest integer partial quotients is equivalent.

1. INTRODUCTION

Real numbers admit regular continued fraction expansions that are unique except for an ambiguity in the ultimate partial fraction of rational numbers. The same is true for nearest integer continued fraction expansions. The nearest integer expansion is easily obtained from the regular expansion, by applying a certain modification for partial quotients that equal 1, as a result of which some partial quotients are incremented by 1, and some minus signs are introduced. As a consequence, for questions concerning the boundedness of partial quotients of real numbers, the behaviour of regular and nearest integer continued fractions is alike.

In both cases, finite expansions occur precisely for the rational numbers, and ultimately periodic expansions occur precisely for quadratic irrational numbers. Not much is known about the partial quotients for other algebraic irrationalities. There exist transcendental numbers with bounded partial quotients, and also transcendental numbers with unbounded partial quotients.

Date: September 12, 2011.

Key words and phrases. continued fraction, bounded partial quotients.

The frequency of partial quotients for almost all real numbers is well-understood, and the usual behaviour is that arbitrarily large partial quotients do occur very occasionally. More precisely, for almost all real numbers (in the Lebesgue sense), partial quotient k appears with frequency $2 \log \left(1 + \frac{1}{k(k+2)}\right)$, according to the Theorem of Gauss, Kuzmin and Levy. A theorem by Borel and Bernstein states that real numbers with bounded partial quotient have measure 0, and implies that $a_n > n \log n$ infinitely often for almost all x.

On the other hand, it is easy to construct real numbers with bounded partial quotients (and there are uncountably many), but although we do not know very much about their partial quotients, it seems impossible to construct *algebraic* numbers this way avoiding finite expansions (rational numbers) or ultimately periodic expansions (quadratic irrationals). For all this, and much more, see [8].

Conjecture 1.1. The only real algebraic numbers for which the partial quotients in their regular or nearest integer continued fraction expansion are bounded, are rational numbers and quadratic irrational numbers.

If true, this means that non-periodic expansions using bounded partial quotients only occur for transcendental numbers.

In this paper we consider the corresponding question for complex continued fractions. The reason we insisted on mentioning the nearest integer expansion for the real case is that it admits an immediate generalization to the complex case, as first studied by Hurwitz. It is much harder to generalize the regular continued fraction to the complex case; see also the next section.

Surprisingly, Doug Hensley [2] found examples of complex numbers that are algebraic of degree 4 over $\mathbb{Q}(i)$ and have bounded complex partial quotients (in the Hurwitz expansion). This paper attempts to collect and tidy up the examples and proofs of Hensley, and to generalize them to obtain the following theorem.

Theorem 1.2. For every even integer d there exist algebraic elements $\alpha \in \mathbb{C} \setminus \mathbb{R}$ of degree d over \mathbb{Q} for which the Hurwitz continued fraction expansion has bounded partial quotients.

The numbers we construct all lie on certain circles in the complex plane; the only real numbers on these circles have degree 2 over \mathbb{Q} and although they too have bounded partial quotients, they are of no help in refuting the above Conjecture.

On the other hand, it will also be easy to construct *transcendental* numbers on the same circles.

 $\mathbf{2}$

2. HURWITZ CONTINUED FRACTIONS

For a real number x, the nearest integer continued fraction expansion

$$x = a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{\ddots}}},$$

can be found by applying the operator

$$x_{k+1} = \mathcal{N}x_k = \frac{1}{x_k} - a_k = \frac{1}{x_k} - \lfloor \frac{1}{x_k} \rceil,$$

for $k \ge 0$, where $a_0 = \lfloor x \rceil$ and $x_0 = x - a_0$. Here $a_k = \lfloor \frac{1}{x_k} \rceil$ is an integer, with $|a_k| \ge 2$ for $k \ge 1$, obtained by rounding to the nearest integer, and $-\frac{1}{2} \le x_k \le \frac{1}{2}$ for $k \ge 0$. The continued fraction stops, and becomes finite, if and only if $x_k = 0$ for some $k \ge 0$, which is the case if and only if x is rational. Note that we allow negative partial quotients a_k here, but the 'numerators' are all 1; alternatively, one often chooses a_k positive but allows numerators ± 1 . Also note that this continued fraction operator only differs from the regular one in the way of rounding: one obtains the regular operator \mathcal{T} by always rounding down, $a_k = \lfloor \frac{1}{x_k} \rfloor$.

The Hurwitz continued fraction operator \mathcal{H} is obtained by a straightforward generalization to complex arguments. Let z be a complex number, and define

$$z_{k+1} = \mathcal{H}z_k = \frac{1}{z_k} - \alpha_k = \frac{1}{z_k} - \lfloor \frac{1}{z_k} \rceil,$$

for $k \ge 0$, with $\alpha_0 = \lfloor z \rceil \in \mathbb{Z}[i]$ and $z_0 = z - \alpha_0$. Now $\lfloor z \rceil$ denotes rounding to the nearest element of the ring of Gaussian integers, $\mathbb{Z}[i]$, with respect to the ordinary 'Euclidean' distance in the complex plane. Then obviously $|\alpha_k| \ge 2$ for $k \ge 1$, as it is easy to see that z_k lies in the symmetric 'unit box'

$$\mathcal{B} = \{ z \in \mathbb{C} \mid -\frac{1}{2} \le \Im w, \Re w \le \frac{1}{2} \}.$$

Again, one takes $\mathcal{H}0 = 0$, and the expansion becomes finite for elements of $\mathbb{Q}(i)$, but infinite otherwise:

$$z = \alpha_0 + \frac{1}{\alpha_1 + \frac{1}{\alpha_2 + \frac{1}{\ddots}}},$$

which is an expansion for z in $\mathbb{Q}(i)$ in the sense that always

$$\alpha_0 + \frac{1}{\alpha_1 + \frac{1}{\alpha_2 + \frac{1}{\ddots \frac{1}{\alpha_n}}}} \longrightarrow z, \text{ for } n \to \infty$$

and the finite continued fraction on the left is an element $\frac{r_n}{s_n} \in \mathbb{Q}(i)$.

Also, the behaviour on quadratic irrationalities is analogous: the nearest integer continued fraction (the Hurwitz continued fraction) of an element $x \in \mathbb{R} \setminus \mathbb{Q}$ (an element $z \in \mathbb{C} \setminus \mathbb{Q}[i]$) is ultimately periodic if and only if x is a quadratic irrationality over \mathbb{Q} (z is quadratic irrational over $\mathbb{Q}[i]$).

So, the Hurwitz operator very nicely generalizes the nearest integer case to the complex case. For the regular case there is no such straightforward generalization, as the square of complex numbers with both real and imaginary parts between 0 and 1 does not lie completely within the unit circle. The best attempt, in some sense, that we know of, is the, rather cumbersome, construction by A. Schmidt [7].

3. Generalized circles

In the proof of the Main Theorem, certain circles in the complex plane, and their images under the Hurwitz continued fraction operator, will play an important role. We fix the notation and list the relevant properties here.

Definition 3.1. A generalized circle, or g-circle for short, is the set of complex solutions to an equation of the form

$$Aw\bar{w} + Bw + B\bar{w} + D = 0$$

in the complex variable w (where $\bar{}$ denotes complex conjugation), for real coefficients A, D and a complex coefficient B satisfying $B\bar{B}-AD \geq$ 0. We will denote a g-circle by the matrix

$$\begin{pmatrix} A & \bar{B} \\ B & D \end{pmatrix},$$

as

$$Aw\bar{w} + Bw + \bar{B}\bar{w} + D = \begin{pmatrix} \bar{w} & 1 \end{pmatrix} \begin{pmatrix} A & \bar{B} \\ B & D \end{pmatrix} \begin{pmatrix} w \\ 1 \end{pmatrix}.$$

The motivation for this definition is that the set of solutions in the complex w = x + yi-plane form an ordinary circle with centre $-\bar{B}/A$ and radius $\sqrt{|B|^2 - AD}/|A|$ when $A \neq 0$, whereas for A = 0 they form

a line ax - by = -D/2, with $a = \Re(B)$ and $b = \Im(B)$; in any case it passes through the origin precisely when D = 0.

The map $w \mapsto \frac{1}{w}$ maps g-circles to g-circles. Indeed, the image of $\mathcal{C} = \begin{pmatrix} A & \bar{B} \\ B & D \end{pmatrix}$ under this involution is $\mathcal{C} = \begin{pmatrix} D & B \\ \bar{B} & A \end{pmatrix}$. Of course any translation of the complex plane also maps g-circles to g-circles; as a consequence, the composed map $\mathcal{H}w = \frac{1}{w} - \alpha$ maps g-circle $\mathcal{C} = \begin{pmatrix} A & \bar{B} \\ B & D \end{pmatrix}$ to another g-circle $\mathcal{H}\mathcal{C}$ given by $\begin{pmatrix} \bar{\alpha} & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} A & \bar{B} \\ B & D \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & \alpha \end{pmatrix} = \begin{pmatrix} D & B + \alpha D \\ \bar{B} + \bar{\alpha}D & A + \alpha \bar{B} + \bar{\alpha}B + \alpha \bar{\alpha}D \end{pmatrix}.$

Note that \mathcal{H} leaves the determinant $AD - B\overline{B}$ of the matrix corresponding to \mathcal{C} invariant.

Let z be a complex number, which we will assume to be irrational to avoid notational complications arising from terminating continued fractions, and let $\alpha_0 = \lfloor z \rfloor$ and $z_0 = z - \alpha_0$. Also, let the circle C have centre $-\alpha_0$ and radius |z|. This is given by

$$||w + \alpha_0|| = (w + \alpha_0)(\overline{w + \alpha_0}) = |z|^2,$$

so as a g-circle this is

$$\mathcal{C}_0 = \begin{pmatrix} 1 & \alpha_0 \\ \bar{\alpha}_0 & |\alpha_0|^2 - |z|^2 \end{pmatrix}.$$

For $n \geq 1$ define $\alpha_n = \lfloor \frac{1}{z_{n-1}} \rfloor$ and $z_n = \frac{1}{z_{n-1}} - \alpha_n$. Then $[\alpha_0, \alpha_1, \ldots]$ is the Hurwitz continued fraction expansion of z. By definition, z_0 lies on g-circle \mathcal{C}_0 , and also lies in the unit box \mathcal{B} .

If we apply $\mathcal{H}_1: w \mapsto \frac{1}{w} - \alpha_1$ to z_0 we obtain $z_1 \in \mathcal{B}$, while applying \mathcal{H}_1 to \mathcal{C}_0 we obtain a g-circle \mathcal{C}_1 as above, with z_1 lying on it. Repeating this, we find g-circles $\mathcal{C}_0, \mathcal{C}_1, \mathcal{C}_2, \ldots$, with corresponding matrices $\begin{pmatrix} A_j & \bar{B}_j \\ B_j & D_j \end{pmatrix}$ for $j \geq 0$, and complex numbers $z_j \in \mathcal{C}_j \cap \mathcal{B}$. Moreover, $A_j D_j - B_j \bar{B}_j = A_0 D_0 - B_0 \bar{B}_0 = -|z|^2$ for $j \geq 1$. We call the \mathcal{C}_j the sequence of g-circles corresponding to the Hurwitz exapansion $z = [\alpha_0, \alpha_1, \ldots]$.

Lemma 3.2. If $|z|^2 = n \in \mathbb{Z}$ then for the g-circles $C_j = \begin{pmatrix} A_j & B_j \\ B_j & D_j \end{pmatrix}$ corresponding to the Hurwitz continued fraction expansion of z it holds that $A_j, D_j \in \mathbb{Z}$ and $B_j \in \mathbb{Z}[i]$, and $B_j \overline{B}_j - A_j D_j = n$. *Proof.* The statement is true for j = 0, as $C_0 = \begin{pmatrix} 1 & \alpha_0 \\ \overline{\alpha}_0 & |\alpha_0|^2 - |z|^2 \end{pmatrix}$ with $\alpha_0 \in \mathbb{Z}[i]$ the nearest Gaussian integer to z. For j > 0 it then follows inductively from the action of \mathcal{H} .

4. Main Theorem

Theorem 4.1. Let z be a complex number. If $n = |z|^2 \in \mathbb{Z}_{>0}$ then the sequence C_0, C_1, C_2, \ldots of g-circles corresponding to the Hurwitz expansion of z consists of finitely many different q-circles.

Proof. According to Lemma 3.2 the matrix entries for the g-circles C_j corresponding to the Hurwitz expansion of z satisfy: $A_j, D_j \in \mathbb{Z}$ and $B_j \in \mathbb{Z}[i]$, while also $A_j D_j - B_j \overline{B}_j = -|z|^2 = -n$. The finiteness of the number of different g-circles among the C_j will follow from the observation that there are only finitely many solutions to this equation with the additional property that the g-circle intersects the unit box \mathcal{B} , a condition imposed by the fact that the remainder $z_j \in C_j \cap \mathcal{B}$.

For the case $A_j = 0$ this is clear: the g-circle is then a line $r_j x - i_j y = -D_j/2$, where $r_j = \Re B_j$ and $i_j = \Im B_j$ are rational integers satisfying $r_j^2 + i_j^2 = n$; this admits only finitely many solutions for B_j , and for the line to intersect the unit box one needs $D_j \leq |r_j| + |i_j|$.

For the case $A_j \neq 0$ we proceed by induction on j, to show that the radius R_j satisfies $R_j^2 > 1/8$ for all j. For j = 0 this holds, as the g-circle C_0 has radius $R_0 = \sqrt{n}$. The induction hypothesis (which will only be used in the final subcase below) is that if C_{j-1} is a proper circle, then its radius satisfies $R_{j-1}^2 > 1/8$.

Suppose that g-circle C_j happens to pass through the origin, for some $j \geq 1$; that means that $D_j = 0$. This implies that g-circle C_{j-1} is a line not passing through the origin; but it has to intersect the unit box, so the point P on it that is closest to the origin is at distance less than $1/\sqrt{2}$ from the origin. But under \mathcal{H} the point P of C_{j-1} gets mapped to the point diametrically opposed from the origin on C_j and will be at distance at least $\sqrt{2}$. Hence the square R_j^2 of the radius of C_j will be at least 1/2.

In the remaining cases, both A_j and D_j are non-zero integers, and so are A_{j-1} and D_{j-1} .

First suppose that A_{j-1} and D_{j-1} have opposite signs. This means that the origin is in interior point of the g-circle C_{j-1} . Also, $z_{j-1} \in C_{j-1} \cap \mathcal{B}$ is at distance at most $1/\sqrt{2}$ from the origin. The image C of C_{j-1} under \mathcal{H}_0 is a g-circle that also has the origin as an interior point, that has the same radius as C_j , and that contains $1/z_{j-1}$, which is at distance at least $\sqrt{2}$ from the origin. This implies that the radius of C_j is at least $\sqrt{2}/2$, so $R_j^2 \ge 1/2$.

Finally, suppose that A_{j-1} and D_{j-1} have the same sign. In this case the origin is an exterior point of both \mathcal{C}_{j-1} and of \mathcal{C}_j . However, the point P on \mathcal{C}_{j-1} nearest to the origin is at distance $c < 1/\sqrt{2}$ from the origin (as there is at least one point in $\mathcal{C}_{j-1} \cap \mathcal{B}$. The diametrically opposed point Q on \mathcal{C}_{j-1} is at distance c+d from the origin, with d the diameter of \mathcal{C}_{j-1} . Now using the induction hypothesis that $d > 1/\sqrt{2}$, we infer that the diameter of the image of \mathcal{C}_{j-1} under \mathcal{H}_0 , and hence \mathcal{C}_j , has diameter

$$\frac{1}{c} - \frac{1}{c+d} = \frac{d}{c(c+d)} > \frac{\frac{1}{\sqrt{2}}}{c(c+\frac{1}{\sqrt{2}})} > \frac{c}{c(c+\frac{1}{\sqrt{2}})} > \frac{1}{\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}} = \frac{1}{\sqrt{2}},$$

and therefore $R_i^2 > 1/8$.

We conclude that in any case $R_i^2 > 1/8$.

As

$$R_j^2 = (B_j \bar{B}_j - A_j D_j) / A_j^2 = n / A_j^2,$$

this leaves only finitely many possibilities for the integer A_j . For C_j to intersect the unit box, its center can not be too far from the origin:

$$\left|\frac{-\bar{B}_j}{A_j}\right| \le \frac{1}{2\sqrt{2}} + \frac{\sqrt{n}}{|A_j|},$$

and this leaves only finitely many possibilities for B_j , for each A_j . Since D_j is completely determined by A_j and B_j , the proof is complete. \Box

Corollary 4.2. Let $z \in \mathbb{C}$ be such that its norm $n = |z|^2 \in \mathbb{Z}_{>0}$ is not the sum of two squares of integers. Then the partial quotients in the Hurwitz continued fraction of z are bounded.

Proof. According to the Theorem, the remainders z_i of the Hurwitz continued fraction operator all lie on a finite number of different gcircles. If such a g-circle C_j passes through the origin, then the entry D_j of its matrix equals 0, and B_j is a Gaussian integer that satisfies $B_j \bar{B}_j = n$ by Lemma 3.2. This is a contradiction, as n is not the sum of two integer squares. Therefore none of the g-circles passes through the origin. This means that there exists a positive constant C (the shortest distance from any of the g-circles to the origin) such that $|z_j| \ge C$, and then $|\alpha_{j+1}| = \lfloor \frac{1}{z_i} \rceil \le \lceil \frac{1}{C} \rceil$.

As an immediate consequence we have a proof of Theorem (1.2): start with any positive integer $n \equiv 3 \mod 4$, and construct elements of norm n; it is easy to construct algebraic numbers of any even degree 2*m* this way, for example using $\sqrt[m]{2} + i\sqrt{n} - \sqrt[m]{4}$). We will carry this out explicitly for n = 7 in the next section.

5. Examples

Corollary 4.2 allows us to construct examples of various types. All examples in this section use the set of g-circles arising from complex numbers of norm n = 7. We have not attempted to determine all g-circles in this case by a straightforward computation, but it is very likely that the complete set consists of the 72 g-circles of which the arcs intersecting the unit box are shown in the first picture.

More generally, we intend to consider the relation with the reduction theory of complex binary quadratic forms of given determinant [4] at another occasion.

We begin with the type of example that Hensley found [2].

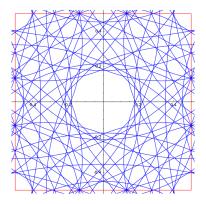
Example 5.1. $z = \sqrt{2} + i\sqrt{5}$.

The Hurwitz continued fraction expansion of $z = \sqrt{2} + i\sqrt{5}$ reads

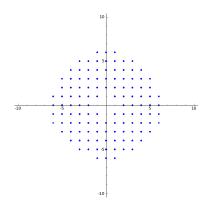
 $z = [2i + 1, -i + 2, i - 5, -i - 2, -4, i - 2, -4, -2, i - 1, -2i, \ldots].$

where the dots do not indicate an obvious continuation.

The first figure below shows the g-circles (or rather: their intersection with the unit box \mathcal{B}) that arise in the first 20000 steps. There are 72 of them; it is very likely that these form the complete set of g-circles, as the same set turns up in the examples below as well, and in each case all 72 circles are hit already after just a couple of 1000s of steps.



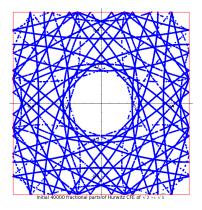
The second picture shows all of the first 20000 partial quotients that occur. There are 118 of them.



When the computations are extended (to 50000 partial quotients) all of the obvious omissions in the picture, like 6+2i, -6+2i) do appear as partial fraction (one around 48000 steps), with the exception of 2-6i.

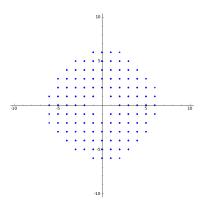
The frequency of the partial quotients in the first 50000 steps varies from around 2100 in 50000 (for the elements of norm 5) to 18 in 50000 (for the elements of norm 27), and 1 in 50000 (for the elements of norm 40). This should be compared (cf. also [1]) with the Gauss-Kuzmin-Levy result in the real regular case.

Also, the frequency with which the various g-circles are visited differs significantly; this is graphically depicted in the third figure, in which the first 40000 remainders are plotted.

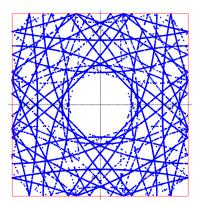


Next consider the family of examples of the form $w_m = \sqrt[m]{2} + i\sqrt{7} - \sqrt[m]{4}$. For odd m, the element w_m is algebraic of degree 2m and norm 7 over \mathbb{Q} . By Corollary 4.2 this gives a proof for Theorem 1.2.

Example 5.2. $\sqrt[3]{2} + i\sqrt{7 - \sqrt[3]{2}}$.



The pictures show the first 20000 partial quotients and the first 40000 g-circles that occur in this case. Of course the example was designed in such a way that the same set of g-circles occurs as in the previous example.



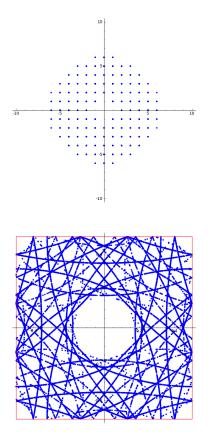
In this case both 6 + 2i and -2 - 6i are still missing after 50000 steps. The frequencey distribution here is similar to the one in the previous case, but there seem to be some differences; for example, the six elements of norm 40 that do occur in the first 50000 steps, do so 6, 4, 3, 2, 2, 1 times, while the seven elements in the previous example each occurred exactly once.

It would be interesting to test the significance of these differences seriously.

Finally, we give a transcendental example on the same set of g-circles.

Example 5.3. $\sqrt{\pi} + \sqrt{7 - \pi}i$.

The plot of the remainders in this case is so similar to the previous cases that we do not reproduce it here.



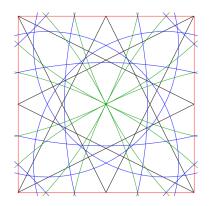
Here only 2 - 6i does not show up among the first 50000 partial quotients.

6. Some Additional Observations

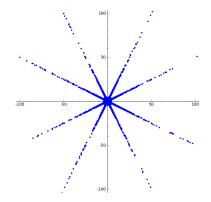
It seems that in the bounded case, the arcs in which the g-circles intersect the unit box always get densely filled. This does not seem to be true in the unbounded case.

Example 6.1. $\sqrt{2} + i\sqrt{3}$

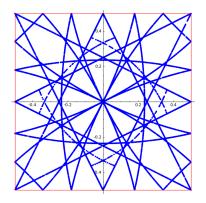
Note that the norm, 5, in this case is the sum of two integral squares. The following pictures nicely illustrate the behaviour in this case:



there are g-circles through the origin, and the partial quotients are the union of a bounded set (coming from the g-circles that are proper circles avoiding the origin) and lattice points near the finite number (4 in this case) of rays corresponding to the g-circles that are lines through the origin.



If, like before, we now plot the first 20000 remainders, we see that certain parts of the g-circle arcs inside the unit box do not get hit.



Also, we conjecture that the following converse of Corollary 4.2 holds.

13

Conjecture 6.2. Let $z \in \mathbb{C}$ be such that its norm $n = |z|^2 \in \mathbb{Z}_{>0}$ is the sum of two squares of integers. Then the partial quotients in the Hurwitz continued fraction of z are unbounded, unless z is in $\mathbb{Q}(i)$ or quadratic over $\mathbb{Q}(i)$.

In this paper we have attempted to answer some questions on complex numbers with bounded partial quotients. Many others may arise; for example: what can be said about numbers defined form more general finite sets of partial quotients than the, very symmetric, finite sets arising from our examples? In general, such finite sets will be a *subset* of a finite set in our construction, and most likely the numbers represented will form a measure zero subset, and questions about algebraicity will be difficult to answer.

References

- [1] Doug Hensley, *The Hurwitz Complex Continued Fraction*, preprint, January 9, 2006.
- [2] Doug Hensley, Continued fractions, Singapore: World Scientific, 2006, Chapter 5: Complex continued fractions.
- [3] A. Hurwitz, Über die Entwicklung complexer Grössen in Kettenbrüche, Acta Math. 11, (1888), 187–200.
- [4] A. Hurwitz, Uber die Anzahl der Classen binärer quadratischen Formen von negativer Determinante, Acta Math. 19, (1895), 351–384.
- [5] J. Hurwitz, Über eine besondere Art der Kettenbruchentwicklung complexer Grössen, Dissertation, Halle, 1895.
- [6] J. Hurwitz, Über die Reduction der binären quadratischen Formen mit complexen Coëfficiënten und Variabeln, Acta Math. 25 (1901), 231–290.
- [7] A. Schmidt, Diophantine approximation of complex numbers, Acta Math. 134 (1975), 1–85.
- [8] J. O. Shallit, Real numbers with bounded partial quotients. A survey., Enseign. Math. 38 (1992), 151–187.

DEPARTMENT OF MATHEMATICS, RADBOUD UNIVERSITEIT, NIJMEGEN NL E-mail address: W.Bosma@math.ru.nl, davidg@maths.usyd.edu.au