ACTA ARITHMETICA
XCV.2 (2000)

On continued fractions and diophantine approximation in power series fields

by

WOLFGANG M. SCHMIDT (Boulder, CO)

1. Continued fractions in fields of series. While some deep work has been done on continued fractions in power series fields, there does not seem to exist a general introduction, or an easily accessible account of Serret's Theorem or Lagrange's Theorem in this case. We therefore will start with the (obvious) definitions, and set some notation. But see also [12]. Whenever possible, we will try to stay close to the approach in Perron's classical treatise [19]. We define rational functions in variables $Z_0, Z_1, \ldots$ by ([1])

$$[Z_0] = Z_0, \quad [Z_0, Z_1] = Z_0 + 1/Z_1, \tag{1.1}$$

$$[Z_0, Z_1, \ldots, Z_n] = [Z_0, \ldots, Z_{n-2}, Z_{n-1} + 1/Z_n] \quad (n \geq 2). \tag{1.2}$$

Then

$$[Z_0, Z_1, \ldots, Z_n] = [Z_0, \ldots, Z_{m-1}, [Z_m, \ldots, Z_n]] \quad (1 \leq m \leq n). \tag{1.3}$$

Setting $P_{-2} = 0$, $Q_{-2} = 1$, $P_{-1} = 1$, $Q_{-1} = 0$,

$$P_n = Z_n P_{n-1} + P_{n-2}, \quad Q_n = Z_n Q_{n-1} + Q_{n-2} \quad (n \geq 0), \tag{1.4}$$

we observe that P_n, Q_n are polynomials in $Z_0, Z_1, \ldots, Z_n$ $(n \geq 0)$, that

$$Q_n P_{n-1} - P_n Q_{n-1} = (-1)^n \quad (n \geq -1), \tag{1.5}$$

and

$$[Z_0, Z_1, \ldots, Z_n] = P_n/Q_n. \tag{1.6}$$

In the function field case we will be considering, the rôles of $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ in the classical theory of continued fractions are played by $k[X], k(X), k((X^{-1}))$

2000 *Mathematics Subject Classification*: 11J61, 11J70.
Supported in part by NSF grant DMS–9710314.

([1]) I had been tempted to introduce a different convention, with the + signs in (1.1), (1.2) replaced by −, and setting $Q_{-2} = -1$. This would make some sense in the function field case (where positivity of partial quotients is not defined) by simplifying a number of formulas (e.g., the right hand side of (1.5) would become 1). I gave up on this idea since it would have made it harder to connect with the existing literature.

respectively, where k is a field. Elements of k will be denoted by a, b, c, elements of $k[X]$ by $A, B, \ldots$, and elements of $k((X^{-1}))$ by $\alpha, \beta, \ldots$ When $\alpha = a_t X^t + a_{t-1} X^{t-1} + \ldots$ with $a_t \neq 0$, we set $|\alpha| = e^t$, and we set $|0| = 0$.

An element of $k(X)$ may be uniquely expressed as

$$[A_0, A_1, \ldots, A_n] = P_n/Q_n, \tag{1.7}$$

where each $A_i \in k[X]$ and where $\deg A_i > 0$ for $i > 0$. Here $P_n = P_n(A_0, \ldots, A_n)$, $Q_n = Q_n(A_0, \ldots, A_n)$. This is in contrast to the classical theory, where rational numbers have two "regular" continued fraction expansions. By (1.5), the polynomials P_n, Q_n in $k[X]$ are relatively prime. Note that the pair P_n, Q_n is determined by (1.7) and being relatively prime only up to a common factor in $k^\times$: aP_n, aQ_n with $a \in k^\times$ have the same properties. When $A_0, A_1, \ldots$ are given with $\deg A_i > 0$ for $i > 0$, then $[A_0, A_1, \ldots, A_n]$ as $n \to \infty$ converges with respect to $|\cdot|$ to an element of $k((X^{-1}))$ which will be denoted by

$$[A_0, A_1, \ldots]. \tag{1.8}$$

Every element $\alpha \in k((X^{-1}))\backslash k(X)$ can be uniquely expressed as such an infinite "regular" continued fraction. Writing $\alpha = [A_0, A_1, \ldots]$, we call P_n/Q_n as given by (1.7) a *convergent*, and we call A_n a *partial quotient*, $\alpha_n = [A_n, A_{n+1}, \ldots]$ a *complete quotient*. We have

$$\alpha = [A_0, \ldots, A_{m-1}, \alpha_m] = \frac{\alpha_m P_{m-1} + P_{m-2}}{\alpha_m Q_{m-1} + Q_{m-2}} \qquad (m \geq 1)$$

by (1.3), (1.4) and letting n go to infinity.

It is easily checked that when B, C are nonzero in $k[X]$, then

$$C[BA_0, CA_1, BA_2, \ldots] = B[CA_0, BA_1, CA_2, \ldots]. \tag{1.9}$$

In particular, when $a \in k^\times$ and $\alpha = [A_0, A_1, A_2, \ldots]$, then

$$a\alpha = [aA_0, a^{-1}A_1, aA_2, \ldots]. \tag{1.10}$$

This and similar relations may be interpreted appropriately for finite continued fractions (1.7) as well as for infinite continued fractions (1.8).

With $P_n = P_n(A_0, \ldots, A_n)$, $Q_n = Q_n(A_0, \ldots, A_n)$, we have

$$|Q_n| = |A_n| \cdot |Q_{n-1}| = |A_n| \cdot |A_{n-1}| \ldots |A_1| \qquad (n \geq 1), \tag{1.11}$$

$$|\alpha - P_{n-1}/Q_{n-1}| = 1/(|Q_{n-1}| \cdot |Q_n|) = 1/(|A_n| \cdot |Q_{n-1}|^2) \quad (n \geq 1). \tag{1.12}$$

The following version of Legendre's Theorem holds: *If* $|\alpha - P/Q| < 1/|Q|^2$, *then* P/Q *is a convergent to* α. For if $|Q_n| \leq |Q| < |Q_{n+1}|$, we have $|\alpha - P/Q| < 1/(|Q| \cdot |Q_n|)$, $|\alpha - P_n/Q_n| = 1/(|Q_n| \cdot |Q_{n+1}|) < 1/(|Q| \cdot |Q_n|)$, so that $|P/Q - P_n/Q_n| < 1/(|QQ_n|)$, whence $P/Q = P_n/Q_n$.

2. A version of Serret's Theorem. We will write $(k^\times)^2$ for the subgroup of squares in $k^\times$; its cosets are $a(k^\times)^2$ with $a \in k^\times$.

LEMMA 1. *Suppose*

$$\alpha = (A\beta + B)/(C\beta + D)$$

where $|D| < |C|$ *and* $AD - BC = a \in k^\times$. *Suppose* α, β *are not in* $k(X)$, *and* $|\beta| > 1$. *Let* P_n/Q_n *and* α_n $(n \geq 0)$ *be the convergents and complete quotients of* α. *Then for some* n,

$$A/C = P_n/Q_n, \qquad B/D = P_{n-1}/Q_{n-1},$$

and $\beta = b\alpha_{n+1}$ *with* $b \in (-1)^{n+1}a(k^\times)^2$.

Proof. Write A/C as a finite continued fraction: $A/C = [A_0, \ldots, A_n] = P_n^*/Q_n^*$ (where the star indicates it is the (last) convergent of A/C). Since A, C are coprime,

$$A = c^{-1}P_n^*, \qquad C = c^{-1}Q_n^*$$

with $c \in k^\times$. Thus

$$P_n^*D - Q_n^*B = c(AD - BC) = ac = ac(-1)^n(Q_n^*P_{n-1}^* - P_n^*Q_{n-1}^*)$$

by (1.5). Hence

$$P_n^*(D + (-1)^n acQ_{n-1}^*) = Q_n^*(B + (-1)^n acP_{n-1}^*).$$

Since P_n^*, Q_n^* are coprime, $Q_n^* \mid (D + (-1)^n acQ_{n-1}^*)$. But $|D| < |C| = |Q_n^*|$, $|Q_{n-1}^*| < |Q_n^*|$, so that $D + (-1)^n acQ_{n-1}^* = 0$. We obtain

$$D = (-1)^{n+1}acQ_{n-1}^*, \qquad B = (-1)^{n+1}acP_{n-1}^*. \tag{2.1}$$

Thus

$$\alpha = \frac{c^{-1}P_n^*\beta + (-1)^{n+1}acP_{n-1}^*}{c^{-1}Q_n^*\beta + (-1)^{n+1}acQ_{n-1}^*} = \frac{((-1)^{n+1}c^{-2}a^{-1}\beta)P_n^* + P_{n-1}^*}{((-1)^{n+1}c^{-2}a^{-1}\beta)Q_n^* + Q_{n-1}^*}.$$

Therefore $\alpha = [A_0, \ldots, A_n, (-1)^{n+1}a^{-1}c^{-2}\beta]$. Since $|\beta| > 1$, we see that $A_0, \ldots, A_n$ are the first partial quotients of α, and $(-1)^{n+1}a^{-1}c^{-2}\beta = \alpha_{n+1}$. Hence the convergent P_n/Q_n of α is $[A_0, \ldots, A_n] = P_n^*/Q_n^*$, so that in fact $P_n/Q_n = A/C$, and similarly $P_{n-1}/Q_{n-1} = B/D$ by (2.1). Finally $\beta = b\alpha_{n+1}$ with $b = (-1)^{n+1}ac^2 \in (-1)^{n+1}a(k^\times)^2$. ∎

Write $\alpha \approx \beta$ if there is a relation

$$\beta = (R\alpha + S)/(T\alpha + U) \tag{2.2}$$

with R, S, T, U in $k[X]$ having $RU - ST \in k^\times$. This is an equivalence relation. Note that $\alpha \approx \alpha_n$ $(n = 1, 2, \ldots)$ by (1.5). When $\beta \in k(X)$ and $\alpha \approx \beta$, then $\alpha \in k(X)$. On the other hand, any two elements of $k(X)$ are equivalent under $\approx$: Since $[A_0, \ldots, A_n] = P_n/Q_n = (A_nP_{n-1} + P_{n-2})/(A_nQ_{n-1} + Q_{n-2}) \approx A_n$ by (1.5), it suffices to check that any polynomial A is equivalent to 1. But $A = (1 \cdot 1 + (A - 1))/(0 \cdot 1 + 1)$, hence indeed $A \approx 1$.

THEOREM 1. *Suppose α, β are not in $k(X)$. Then $\alpha \approx \beta$ iff for some m, n and some $a \in k^\times$ we have*

$$\beta_m = a\alpha_n,$$

so that the expansions are as follows:

$$\alpha = [A_0, \ldots, A_{n-1}, C_1, C_2, \ldots], \tag{2.3}$$

$$\beta = [B_0, \ldots, B_{m-1}, aC_1, a^{-1}C_2, \ldots]. \tag{2.4}$$

This is our version of a theorem of Serret [22]. The theorem is also proved in [12], Section IV.3.

Proof. Assuming (2.3) and (2.4), we have $\alpha_n = [C_1, C_2, \ldots]$, $\beta_m = [aC_1, a^{-1}C_2, \ldots]$. Then by (1.10), $\beta_m = a\alpha_n = (a\alpha_n + 0)/(0\alpha_n + 1) \approx \alpha_n$. Since $\alpha \approx \alpha_n$, $\beta \approx \beta_m$, we have indeed $\alpha \approx \beta$.

Conversely, suppose $\alpha \approx \beta$, and write $\alpha = [A_0, A_1, \ldots]$, so that

$$\alpha = [A_0, \ldots, A_{n-1}, \alpha_n] = (P_{n-1}\alpha_n + P_{n-2})/(Q_{n-1}\alpha_n + Q_{n-2}) \quad (n \geq 0).$$

Then when (2.2) holds,

$$\beta = \frac{R(P_{n-1}\alpha_n + P_{n-2}) + S(Q_{n-1}\alpha_n + Q_{n-2})}{T(P_{n-1}\alpha_n + P_{n-2}) + U(Q_{n-1}\alpha_n + Q_{n-2})} = \frac{A\alpha_n + B}{C\alpha_n + D} \tag{2.5}$$

with

$$\begin{aligned} A &= RP_{n-1} + SQ_{n-1}, \quad & B &= RP_{n-2} + SQ_{n-2}, \\ C &= TP_{n-1} + UQ_{n-1}, \quad & D &= TP_{n-1} + UQ_{n-2}. \end{aligned} \tag{2.6}$$

Now $|\alpha - P_{n-1}/Q_{n-1}| < 1/|Q_{n-1}|^2$ by (1.12), so that $P_{n-1} = \alpha Q_{n-1} + \delta$ with $|\delta| < 1/|Q_{n-1}|$, and

$$C = (T\alpha + U)Q_{n-1} + \delta T.$$

Here $|T\alpha + U| \neq 0$ and $|T|$ are fixed, so that $|C| = |T\alpha + U| \cdot |Q_{n-1}|$ when n is large. Similarly $|D| = |T\alpha + U| \cdot |Q_{n-2}|$, therefore $|D| < |C|$. By the preceding lemma, with $\beta, \alpha_n, m-1$ playing the respective rôles of α, β, n, we have $\alpha_n = b\beta_m$ for some n, with $b \in k^\times$. Hence $\beta_m = a\alpha_n$ with $a \in k^\times$. ∎

An expansion of the type

$$\begin{aligned} [A_0, \ldots, A_{n-1}, B_1, \ldots, B_{2t}, aB_1, a^{-1}B_2, \ldots, a^{-1}B_{2t}, \\ a^2B_1, \ldots, a^{-2}B_{2t}, a^3B_1, \ldots] \end{aligned} \tag{2.7}$$

with $a \in k^\times$ will be called *pseudoperiodic*. When a is a root of 1, such an expansion is in fact periodic.

LEMMA 2. *Suppose $\alpha \notin k(X)$. Then α has a pseudoperiodic continued fraction expansion if and only if it satisfies a relation*

$$\alpha = (R\alpha + S)/(T\alpha + U) \tag{2.8}$$

where

$$\begin{pmatrix} R & S \\ T & U \end{pmatrix} \tag{2.9}$$

has determinant in $k^\times$ *and is not a multiple of the identity matrix.*

Proof. Suppose α has the expansion (2.7). Then $\alpha_{n+2t} = a\alpha_n$. Recall that $\alpha = (P_{l-1}\alpha_l + P_{l-2})/(Q_{l-1}\alpha_l + Q_{l-2})$ $(l \geq 0)$, and set

$$\mathcal{M}_l = \begin{pmatrix} P_{l-1} & P_{l-2} \\ Q_{l-1} & Q_{l-2} \end{pmatrix}.$$

Then (2.8) holds with

$$\begin{pmatrix} R & S \\ T & U \end{pmatrix} = \mathcal{M}_{n+2t} \begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix} \mathcal{M}_n^{-1}.$$

The determinant here is $a \in k^\times$. If the matrix were a multiple of the identity matrix, it would be a constant multiple, say b times the identity matrix, and

$$\mathcal{M}_{n+2t} \begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix} = \mathcal{M}_n \begin{pmatrix} b & 0 \\ 0 & b \end{pmatrix}.$$

Since the entry aQ_{n+2t-1} of the matrix on the left has larger degree than the corresponding entry bQ_{n-1} on the right, this is impossible.

Now suppose we have (2.8). Suppose we follow the proof of the second part of Theorem 1, with $\beta = \alpha$. We end up with $\alpha_n = b\alpha_m$ for some m. We just need to show that $m \neq n$. But the argument depended on Lemma 1, which (with m playing the rôle of $n+1$) also gives $A/C = P_{m-1}/Q_{m-1}$, $B/D = P_{m-2}/Q_{m-2}$. If this were true with $m = n$, then $A = uP_{n-1}$, $C = uQ_{n-1}$, $B = vP_{n-2}$, $D = vQ_{n-2}$ with $u, v \in k^\times$, and substitution into (2.5) with $\beta = \alpha$ gives

$$\alpha = (uP_{n-1}\alpha_n + vP_{n-2})/(uQ_{n-1}\alpha_n + vQ_{n-2}).$$

Since also $\alpha = (P_{n-1}\alpha_n + P_{n-2})/(Q_{n-1}\alpha_n + Q_{n-2})$, we may infer that $v = u$, so that

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = u \begin{pmatrix} P_{n-1} & P_{n-2} \\ Q_{n-1} & Q_{n-2} \end{pmatrix} = u\mathcal{M}_n.$$

But by (2.6),

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} R & S \\ T & U \end{pmatrix} \mathcal{M}_n.$$

Since (2.9) is not a multiple of the identity matrix, this is impossible. ∎

3. On Lagrange's Theorem. When α has a pseudoperiodic expansion, there is by Lemma 2 a relation (2.8), so that

$$T\alpha^2 + (U - R)\alpha - S = 0. \tag{3.1}$$

Since (2.9) is not a multiple of the identity matrix, not all coefficients T, $U-R, -S$ are zero. Since $\alpha \notin k(X)$ we see that α is quadratic over $k(X)$.

Conversely, suppose α is quadratic over $k(X)$. When k is finite, an obvious adoption of the standard proof of Lagrange's Theorem shows that α has a periodic continued fraction. Thus Lagrange's Theorem holds when k is finite. Hence it is true when $\operatorname{char} k = p$ and k is algebraic over the prime field $\mathbb{F}_p$.

When $\operatorname{char} k = 0$, or $\operatorname{char} k = p$ and k is transcendental over $\mathbb{F}_p$, there exist elements $a \in k^\times$ which are not roots of 1, and hence there exist pseudoperiodic fractions which are not periodic. Therefore Lagrange's Theorem is not true in the form that every α quadratic over $k(X)$ has a periodic continued fraction expansion. But one may ask whether a quadratic α always has a pseudoperiodic expansion. We will show that in general this is not the case.

THEOREM 2. *Suppose* $\operatorname{char} k \neq 2$. *Suppose α is quadratic over $k(X)$ and satisfies*

$$A\alpha^2 + B\alpha + C = 0 \tag{3.2}$$

where A, B, C are relatively prime. The discriminant of this equation is $D = B^2 - 4AC$. Then α has a pseudoperiodic continued fraction if and only if the relation

$$Y^2 - DZ^2 \in k^\times \tag{3.3}$$

has a nontrivial solution, i.e., a solution Y, Z in $k[X]$ with $Z \neq 0$.

We may regard (3.3) as a "Pell relation". When k is closed under taking square roots, the nontrivial solubility of (3.3) is equivalent to the nontrivial solubility of $Y^2 - DZ^2 = 1$. In the next section we will show a theorem already known to Abel [1] (see also Schinzel [20], [21]) that when (3.3) has a nontrivial solution, then $\sqrt{D}$ has in fact a periodic (not just pseudoperiodic) expansion.

Proof. In view of Lemma 2, we have to show that α satisfies a relation (2.8) with matrix (2.9) as specified in Lemma 2, precisely if (3.3) has a nontrivial solution.

In (3.1), the triple $T, U-R, -S$ must be proportional to A, B, C in (3.2), say

$$T = ZA, \quad U - R = ZB, \quad -S = ZC \tag{3.4}$$

with $Z \in k[X]$, $Z \neq 0$. In view of $RU - ST \in k^\times$ we obtain

$$R(R + ZB) + ACZ^2 = a \tag{3.5}$$

with $a \in k^\times$. Considered as a quadratic equation in R, the discriminant must be a square in $k[X]$, i.e.,

$$Z^2B^2 - 4ACZ^2 + 4a = DZ^2 + 4a \tag{3.6}$$

must be a square, say Y^2, with $Y \in k[X]$. Thus

$$Y^2 - DZ^2 = 4a \in k^\times. \tag{3.7}$$

Conversely, (3.3) gives (3.7) with $a \in k^\times$, so that (3.6) is a square, and the equation (3.5) in R has the solution $\frac{1}{2}(-ZB+(DZ^2+4a)^{1/2}) = \frac{1}{2}(-ZB+Y)$. Defining T, U, S by (3.4) we obtain (3.1). Moreover, $RU - ST = a$, and $T = ZA \neq 0$, so that the matrix (2.9) is not a multiple of the identity matrix. Finally (3.1) yields (2.8). ∎

In our context, when α with (3.2) is in $k((X^{-1}))$, there is a square root of $D = D(X)$ in $k((X^{-1}))$, and this is true precisely when D is of even degree and its leading coefficient is a square in $k^\times$. Clearly this is also necessary for the solubility of (3.3). Further facts about the solubility of the Pell relation (3.3) will be given in Section 5.

4. Pseudoperiodic elements. An element $\alpha \in k((X^{-1}))$ with a pseudoperiodic continued fraction (2.7) will itself be called *pseudoperiodic*, briefly *pp.*, and we will use the notation

$$\alpha = [A_0, \ldots, A_{n-1}, \overline{B_1, \ldots, B_{2t}}^a]. \tag{4.1}$$

In particular, such α is quadratic over $k(X)$. We will call

$$\overline{B_1, \ldots, B_{2t}}^a \tag{4.2}$$

a *pseudoperiod* of α, and $A_0, \ldots, A_{n-1}$ a *preperiod*. Further $n(\alpha)$ will denote the smallest n such that α can be written with a preperiod of n terms. Finally, α will be called *purely pseudoperiodic*, briefly *ppp.*, if $n(\alpha) = 0$.

When α has pseudoperiod (4.2), each element $[\overline{a^\nu B_1, a^{-\nu} B_2, \ldots, a^{-\nu} B_{2t}}^a]$ with $\nu \in \mathbb{Z} \geq 0$ is some complete denominator α_l of α. Therefore when $\beta \approx \alpha$, there is by Theorem 1 some ν, and some $b \in k^\times$, such that $[\overline{ba^\nu B_1, b^{-1}a^{-\nu} B_2, \ldots, b^{-1}a^{-\nu} B_{2t}}^a]$ is a complete denominator of β. Writing $c = ba^\nu$, we see that β may be written as

$$\beta = [C_0, \ldots, C_{m-1}, \overline{cB_1, c^{-1}B_2, \ldots, c^{-1}B_{2t}}^a]$$

with certain $C_0, \ldots, C_{m-1}$. Hence also β is pp.

We will now suppose that α is given by (4.1). We know it to be quadratic, and furthermore it is separable over $k(X)$. (For when $\operatorname{char} k = p > 0$ and $\beta_1, \ldots, \beta_n \in k((X^{-1}))$ are linearly independent over $k(X)$, then so are $\beta_1^p, \ldots, \beta_n^p$.) It has a conjugate $\alpha' \in k((X^{-1}))$ with $\alpha' \neq \alpha$.

LEMMA 3. *Suppose* $|\alpha| > 1$. *Then*

(i) α *is ppp. precisely if* $|\alpha'| < 1$.
(ii) $n(\alpha) = 1$ *precisely if* $|\alpha'| \geq 1$, $|\alpha - \alpha'| > 1$.

Proof. With α pp., each complete denominator α_m is pp. We will write α'_m for the conjugate of α_m (which need not be the mth denominator of α'). In view of $\alpha = (P_{m-1}\alpha_m + P_{m-2})/(Q_{m-1}\alpha_m + Q_{m-2})$ we have $\alpha' = (P_{m-1}\alpha'_m + P_{m-2})/(Q_{m-1}\alpha'_m + Q_{m-2})$, so that

$$\alpha'_m = \left(-\frac{Q_{m-2}}{Q_{m-1}}\right)\left(\frac{\alpha' - P_{m-2}/Q_{m-2}}{\alpha' - P_{m-1}/Q_{m-1}}\right).$$

The second factor approaches $(\alpha'-\alpha)/(\alpha'-\alpha) = 1$ as $m \to \infty$, and therefore $|\alpha'_m| < 1$ when m is large.

(i) When α is ppp., then $\alpha_{2tl} = a^l\alpha$ $(l = 1, 2, \ldots)$, hence $\alpha'_{2tl} = a^l\alpha'$, therefore $|\alpha'_{2tl}| = |\alpha'|$. Since $|\alpha'_{2tl}| < 1$ for large l, we obtain $|\alpha'| < 1$.

Conversely, when $|\alpha'| < 1$, we claim that each $|\alpha'_m| < 1$. For with the changed notation $\alpha = [A_0, A_1, \ldots]$, we have $\alpha_m = A_m + 1/\alpha_{m+1}$, hence $\alpha'_m = A_m + 1/\alpha'_{m+1}$. Here each $|A_m| > 1$ (since $|\alpha| > 1$). So if $|\alpha'_m| < 1$, then $|1/\alpha'_{m+1}| = |A_m| > 1$, and $|\alpha'_{m+1}| < 1$, which proves our claim by induction. Write again α as (4.1) with $n = n(\alpha)$; then $\alpha_{n+2t} = a\alpha_n$. If we had $n(\alpha) > 0$, then

$$\alpha_{n-1} - a\alpha_{n+2t-1} = A_{n-1} + 1/\alpha_n - a(B_{2t} + 1/\alpha_{n+2t}) = A_{n-1} - aB_{2t},$$

therefore

$$\alpha'_{n-1} - a\alpha'_{n+2t-1} = A_{n-1} - aB_{2t}.$$

The left hand side has absolute value < 1, the right hand side is a polynomial, so that both sides are 0. Therefore $A_{n-1} = aB_{2t}$, so that

$$\alpha = [A_0, \ldots, A_{n-2}, \overline{aB_{2t}, B_1, \ldots, B_{2t-1}}^{a^{-1}}],$$

which gives $n(\alpha) < n = n(\alpha)$, a contradiction.

(ii) When $n(\alpha) > 0$, then $|\alpha'| \geq 1$ by (i). When $n(\alpha) = 1$, then, again by (i), $|\alpha_1| > 1$, $|\alpha'_1| < 1$. But $\alpha = A_0 + 1/\alpha_1$, $\alpha' = A_0 + 1/\alpha'_1$, so that $|\alpha - \alpha'| = |1/\alpha_1 - 1/\alpha'_1| = |1/\alpha'_1| > 1$.

Conversely, if $|\alpha'| \geq 1$, $|\alpha - \alpha'| > 1$, then $n(\alpha) > 0$ by (i). Further $|1/\alpha_1 - 1/\alpha'_1| = |\alpha - \alpha'| > 1$, and since $|1/\alpha_1| < 1$, we have $|1/\alpha'_1| > 1$, $|\alpha'_1| < 1$. Thus $n(\alpha_1) = 0$ by (i), hence $n(\alpha) = 1$. ∎

LEMMA 4. *Suppose* α *is pp. with pseudoperiod* (4.2). *Then* α' *has a pseudoperiod*

$$\overline{cB_{2t}, c^{-1}B_{2t-1}, \ldots, c^{-1}B_1}^{a^{-1}} \tag{4.3}$$

with $c \in k^\times$.

Proof. Suppose initially that α is ppp., so that

$$\alpha = [\overline{B_1, \ldots, B_{2t}}^{a}].$$

We have $\alpha_i = B_{i+1} + 1/\alpha_{i+1}$ $(i = 0, \ldots, 2t-1)$, where we set $\alpha_0 = \alpha$. Hence $\alpha'_i = B_{i+1} + 1/\alpha'_{i+1}$ $(i = 0, \ldots, 2t-1)$, so that $\beta_j := -1/\alpha'_{2t-j}$ $(j = 0, \ldots, 2t)$ has

$$\beta_j = B_{2t-j} + 1/\beta_{j+1} \qquad (j = 0, \ldots, 2t-1).$$

But $\alpha_{2t} = a\alpha$, so that $\alpha'_{2t} = a\alpha'$ and $\beta_{2t} = a^{-1}\beta_0$. We obtain

$$-1/(a\alpha') = -1/\alpha'_{2t} = \beta_0 = [\overline{B_{2t}, B_{2t-1}, \ldots, B_1}^{a^{-1}}]. \tag{4.4}$$

Now let α be a general pp. element given by (4.1). By what we have just shown, $-1/(a\alpha'_n)$ equals the right hand side of (4.4). But $\alpha \approx \alpha_n$, hence $\alpha' \approx \alpha'_n \approx -1/(a\alpha'_n)$. Therefore by what we said in the second paragraph of this section, α' has a pseudoperiod (4.3). ∎

A pseudoperiod (4.2) with $a = 1$ will be called a *period*, and, as is customary, will be denoted by $\overline{B_1, \ldots, B_{2t}}$. But of course the length h of a period $\overline{B_1, \ldots, B_h}$ need not be even.

THEOREM 3. *Suppose α is pp. The following two conditions are equivalent.*

(a) $|\alpha| > 1$, $|\alpha'| \geq 1$, $|\alpha - \alpha'| > 1$, *and* $\alpha + \alpha' \in k[X]$.

(b) α *has a periodic expansion* $[B_0, \overline{B_1, \ldots, B_h}]$ *with* $B_h \neq B_0$, $\deg B_0 > 0$, *and*

$$B_i = B_{h-i} \qquad (i = 1, \ldots, h-1). \tag{4.5}$$

When these conditions hold, $\alpha + \alpha' = 2B_0 - B_h$.

Proof. When (a) holds, we have $n(\alpha) = 1$ by Lemma 3, so that

$$\alpha = [B_0, \overline{B_1, \ldots, B_{2t}}^{a}] \tag{4.6}$$

with $a \in k^\times$ and $B_{2t} \neq aB_0$. Then $1/(\alpha - B_0) = [\overline{B_1, \ldots, B_{2t}}^{a}]$. Applying (4.4) to $1/(\alpha - B_0)$ in place of α, we obtain

$$(-1/a)(\alpha' - B_0) = [\overline{B_{2t}, \ldots, B_1}^{a^{-1}}],$$

$$-\alpha' + B_0 = a[\overline{B_{2t}, \ldots, B_1}^{a^{-1}}] = [\overline{aB_{2t}, a^{-1}B_{2t-1}, \ldots, a^{-1}B_1}^{a^{-1}}].$$

Setting $R = \alpha + \alpha'$, we observe that

$$\begin{aligned} \alpha &= (R - B_0) - \alpha' + B_0 \\ &= [aB_{2t} + R - B_0, \overline{a^{-1}B_{2t-1}, aB_{2t-2}, \ldots, a^{-1}B_1, B_{2t}}^{a}]. \end{aligned} \tag{4.7}$$

Comparison with (4.6) yields

$$B_1 = a^{-1}B_{2t-1}, \qquad B_2 = aB_{2t-2}, \qquad \ldots, \qquad B_{2t-1} = a^{-1}B_1.$$

In particular, $B_t = a^\varepsilon B_t$, where $\varepsilon = 1$ if t is even, $\varepsilon = -1$ if t is odd. Therefore $a = 1$, and $\alpha = [B_0, \overline{B_1, \ldots, B_h}]$ with $h = 2t$, and (4.5) holds. Comparison of the first terms of (4.6), (4.7) yields $B_0 = B_h + R - B_0$, therefore $\alpha + \alpha' = R = 2B_0 - B_h$.

When (b) holds, so that $n(\alpha) = 1$ by $B_0 \neq B_h$, then $|\alpha| > 1$, $|\alpha'| \geq 1$, $|\alpha - \alpha'| > 1$ by Lemma 3. Further $1/(\alpha - B_0) = [\overline{B_1, \ldots, B_h}]$ is purely periodic, so that $-(\alpha' - B_0) = [\overline{B_h, \ldots, B_1}]$ by applying (4.4) to $1/(\alpha - B_0)$. Then

$$-\alpha' = [B_h - B_0, \overline{B_{h-1}, \ldots, B_1, B_h}] = [B_h - B_0, \overline{B_1, \ldots, B_h}] = \alpha + B_h - 2B_0$$

by (4.5). Hence $\alpha + \alpha' = 2B_0 - B_h \in k[X]$. ∎

Suppose $\operatorname{char} k \neq 2$. We have $\alpha + \alpha' = 0$ precisely when $B_h = 2B_0$. But $\alpha + \alpha' = 0$ means that $\alpha = \sqrt{D}$ where $D \in k(X)$ (not necessarily in $k[X]$), and where D is not a square in $k(X)$. We therefore obtain the following corollary, already known to Abel [1] when $k = \mathbb{C}$.

COROLLARY. *Suppose* $\operatorname{char} k \neq 2$. *The following two conditions on a pp.* $\alpha \in k((X^{-1}))$ *are equivalent.*

(i) $\alpha = \sqrt{D}$ *where* $D \in k(X)$ *but* D *is not a square in* $k(X)$, *and* $|D| > 1$.
(ii) α *has an expansion* $\left[\frac{1}{2}B_h, \overline{B_1, \ldots, B_h}\right]$ *with* (4.5).

REMARK. Suppose $\alpha = [B_0, \overline{B_1, \ldots, B_h}]$ as in Theorem 3. Let $l = l(\alpha)$ be the least length of a period of α. Then $l \mid h$ and

$$\alpha = [B_0, \overline{B_1, \ldots, B_l}].$$

Since now $B_i = B_j$ when $i, j > 0$ and $i \equiv j \pmod{l}$, we have $B_l = B_h \neq B_0$, and by (4.5)

$$B_i = B_{l-i} \quad (i = 1, \ldots, l-1), \tag{4.8}$$

since $B_{l-i} = B_{h-i}$.

For $\alpha = \sqrt{D}$ as in the Corollary, the continued fraction is periodic, and the situation is almost exactly as in the classical situation. In particular, the complete quotients α_n may be written as

$$\alpha_n = (R_n + \sqrt{D})/S_n$$

with $R_n, S_n \in k[X]$. (See Perron [19], p. 67, where $R_n, S_n \in \mathbb{Z}$. Our notation $\alpha_n, P_n, Q_n, R_n, S_n$ corresponds to Perron's $\xi_n, A_n, B_n, P_n, Q_n$, respectively.) By loc.cit., p. 92, formula (2),

$$P_{n-1}^2 - DQ_{n-1}^2 = (-1)^n S_n \quad (n = 1, 2, \ldots). \tag{4.9}$$

Since $|\alpha Q_{n-1} - P_{n-1}| < 1$, we have

$$|\alpha Q_{n-1} + P_{n-1}| = |\alpha Q_{n-1}| = |D|^{1/2}|Q_{n-1}|,$$

therefore

$$\begin{aligned}|S_n| &= |P_{n-1}^2 - DQ_{n-1}^2| = |\alpha Q_{n-1} - P_{n-1}| \cdot |\alpha Q_{n-1} + P_{n-1}| \\ &= (1/|Q_{n-1}| \cdot |A_n|)|D|^{1/2}|Q_{n-1}|,\end{aligned}$$

so that

$$|S_n| = |D|^{1/2}/|A_n|. \tag{4.10}$$

In our situation where $\sqrt{D} = \frac{1}{2}B_h + 1/\alpha_1$, it is easily seen that $\alpha_1, -1/\alpha_1'$ have reversed periods, so that Satz 3.6, hence Satz 3.7 of [19] applies to $\xi_0 = \alpha_1$, and then in turn by Satz 3.10, the sequence $S_0, \ldots, S_l$ is symmetric, and so is the sequence $R_1, \ldots, R_l$, that is,

$$S_{l-i} = S_i \ (i = 0, \ldots, l), \qquad R_{l-i} = R_{i+1} \ (i = 0, \ldots, l-1). \tag{4.11}$$

Hence when l is odd, $S_{(l-1)/2} = S_{(l+1)/2}$, and when l is even, $R_{l/2} = R_{l/2+1}$. On the other hand, by loc.cit., Satz 3.11, when $S_n = S_{n+1}$ with $0 \le n < l$, then l is odd and $n = (l-1)/2$, and when $R_n = R_{n+1}$ with $1 \le n < l$, then l is even and $n = l/2$.

Let $\alpha = \sqrt{D}$ be as in the Corollary, and let $l' = l'(\alpha)$ be the least $l' > 0$ such that $S_{l'} \in k^\times$. Since $S_0 = S_l$, we have $l' \le l$.

The following result of Lozach is in the unpublished manuscript [7].

LEMMA 5. *$l' = l$, except that possibly $l' = l/2$ when $l \equiv 2 \pmod 4$.*

Proof. We have $\alpha_n - B_n = (R_n + \sqrt{D})/S_n - B_n$, and for $n \le l$,

$$\alpha_{l-n} - B_{l-n} = (R_{l-n} + \sqrt{D})/S_{l-n} - B_{l-n} = (R_{n+1} + \sqrt{D})/S_n - B_n$$

by virtue of (4.8), (4.11). We obtain

$$(\alpha_n - B_n) - (\alpha_{l-n} - B_{l-n}) = (R_n - R_{n+1})/S_n.$$

The left hand side is of modulus < 1, therefore $|R_n - R_{n+1}| < |S_n|$. Thus when $S_n \in k^\times$, so that $|S_n| = 1$, we have $|R_n - R_{n+1}| < 1$, hence $R_n = R_{n+1}$. By what we said above, when $n < l$ this implies that l is even and $n = l/2$. Thus $l' < l$ yields $l' = l/2$ with l even.

It remains for us to show that $l' = l/2$ is odd. Set $R_{l'} = R$, $S_{l'} = c \in k^\times$, so that $\alpha_{l'} = (R + \sqrt{D})/c = R/c + \alpha/c$. This yields by (1.10)

$$\alpha_{l'+i} = c^{(-1)^{i-1}}\alpha_i \qquad (i = 1, 2, \ldots),$$

hence $B_{l'+i} = c^{\pm 1}B_i$ $(i = 1, 2, \ldots)$. If l' were even, we could set $i = l'/2$ to obtain $B_{3l'/2} = c^{\pm 1}B_{l'/2} = c^{\pm 1}B_{3l'/2}$, by (4.8) and since $2l' = l$. Then $c = 1$, hence $\alpha_{l'+i} = \alpha_i$, therefore $B_{l'+i} = B_i$ $(i = 1, 2, \ldots)$, contradicting the fact that l was the least period of α. ∎

5. The Pell relation. For advice on this section I am indebted to A. Schinzel and U. Zannier (*). I will report on results about the solubility of (3.3) when $D \in k[X]$, but is not a square in $k[X]$. I will suppose that $\operatorname{char} k \neq 2$. For effective algorithms, see [5]. Since there is no nontrivial solution otherwise, we will suppose throughout that D is of even degree $2d$, with leading coefficient which is a square in $k^\times$.

LEMMA 6. *Let K be an extension field of k, and $Y', Z' \in K[X]$ a nontrivial solution of*

$$Y'^2 - DZ'^2 \in K^\times. \tag{5.1}$$

Then Y'/Z' is a convergent of $\sqrt{D}$ or $-\sqrt{D}$. In particular, $Y' = cY$, $Z' = cZ$ where $Y, Z \in k[X]$ and $c \in K^\times$.

Proof. The absolute value $|\cdot|$ can be extended in the obvious way to an absolute value of $K((X^{-1}))$, and then $|Y' + \sqrt{D}Z'| \cdot |Y' - \sqrt{D}Z'| = 1$. Not both factors here can be 1, for this would yield $|\sqrt{D}Z'| \le 1$, contradicting $\deg D > 0$. Suppose that

$$|Y' - \sqrt{D}Z'| < 1 < |Y' + \sqrt{D}Z'|. \tag{5.2}$$

Since $Y' + \sqrt{D}Z' = (Y - \sqrt{D}Z') + 2\sqrt{D}Z'$, we have $|Y' + \sqrt{D}Z'| = |\sqrt{D}Z'| > |Z'|$, therefore

$$|\sqrt{D} - Y'/Z'| < 1/|Z'|^2.$$

By Legendre's Theorem, Y'/Z' is a convergent of $\sqrt{D}$. Since $\sqrt{D} \in k((X^{-1}))$, such a convergent P_n/Q_n has $P_n, Q_n \in k[X]$. Thus $Y' = cY$, $Z' = cZ$ with $Y, Z \in k[X]$ and $c \in K^\times$, where Y, Z satisfy (3.3). An analogous conclusion can be reached when the inequalities in (5.2) are reversed. ∎

The solutions of (3.3) form a group: when solutions (Y, Z) are identified with $Y + \sqrt{D}Z \in k(X, \sqrt{D})$, composition is by multiplication. Assuming there is a nontrivial solution, i.e., a solution with $Z \neq 0$, let $Y_1 + Z_1\sqrt{D}$ be such a solution with minimal $|Z_1|$. Then the general solution is $c(Y_1 + Z_1\sqrt{D})^m$ with $c \in k^\times$, $m \in \mathbb{Z}$. By (4.9), $Y_1/Z_1 = P_{l'-1}/Q_{l'-1}$, where $l' = l'(\sqrt{D})$ is the least $l' > 0$ with $S_{l'} \in k^\times$.

When $\deg D = 2$, (3.3) has a solution with $Z = 1$ and Y linear. We will then suppose that $\deg D \ge 4$. We have seen above that (3.3) has a nontrivial solution precisely if $\alpha = \sqrt{D}$ has a periodic continued fraction expansion. Let $\mathcal{H}$ be the hyperelliptic curve

$$W^2 = D(X)$$

in the (X, W)-plane, and ∞_1, ∞_2 its two points at infinity.

(*) *Added in proof.* I now also became aware of the paper by R. Paysant-Le Roux in Arch. Math. (Basel) 61 (1993), 46–58.

LEMMA 7. *Suppose $D \in k[X]$ of degree $2d \geq 4$ is square free.*

(a) $\infty_1 - \infty_2$ *is of finite order on the Jacobian precisely if α has a periodic continued fraction.*

(b) *When this is the case, then*

$$l' + d - 1 \leq m \leq 1 + l'(d-1) \tag{5.3}$$

where $l' = l'(\alpha)$ and m is the order of $\infty_1 - \infty_2$.

Part (a) is implicit in Abel [1]. For the case of D quartic see also [6], p. 592. Part (b) is due to Y. Hellegouarch and M. Lozach in the unpublished manuscript [7].

Proof. (a) Suppose Y, Z is a nontrivial solution of (3.3). Then the functions $F = Y + ZW$, $\widehat{F} = Y - ZW$ on $\mathcal{H}$ have no poles in the finite (X, W)-plane, hence can have poles only at ∞_1, ∞_2. In view of $F\widehat{F} = Y^2 - DZ^2 \in k^\times$, also their zeros can only be at ∞_1, ∞_2. The divisor of F is therefore of the form $m\infty_1 + n\infty_2$, and since it is a principal divisor, it is $m(\infty_1 - \infty_2)$ with some $m \neq 0$. Conversely, suppose $\infty_1 - \infty_2$ is torsion, and $m(\infty_1 - \infty_2)$ the divisor of some function F on $\mathcal{H}$, say $F = Y' + Z'W$ with $Y', Z' \in K(X)$ where K is an overfield of k. Then $\widehat{F} = Y' - Z'W$ has the divisor $m(\infty_2 - \infty_1)$, so that $F\widehat{F} \in K^\times$, i.e., $Y'^2 - Z'^2 D \in K^\times$. Neither F nor $\widehat{F}$ has a finite pole, hence neither do Y' and $Z'W$. Therefore $Y' \in K[X]$ and $Z'^2 W^2 = Z'^2 D \in K[X]$, so that $Z' \in K[X]$ since D is square-free. By Lemma 6, there is then some $Y = c^{-1}Y'$, $Z = c^{-1}Z'$ in $k[X]$ with (3.3).

(b) When again $F\widehat{F} \in k^\times$, the order of $F + \widehat{F} = 2Y$ at ∞_1 is $\deg Y$, so that F or $\widehat{F}$ has order $\deg Y$ at ∞_1. Thus when $\infty_1 - \infty_2$ has order $m \in \mathbb{N}$, we can pick $F = Y + ZW$ with divisor $m(\infty_1 - \infty_2)$ and $\deg Y = m$. Here Y, Z will be a solution of (3.3) with $|Y|$ of minimal degree, so that $Y = cP_{l'-1}$ with $c \in k^\times$. Therefore

$$m = \deg P_{l'-1}. \tag{5.4}$$

Here $\deg P_{l'-1} = \deg A_0 + \deg A_1 + \ldots + \deg A_{l'-1}$. Clearly $\deg A_0 = d$, and $1 \leq \deg A_i < d$ for $i = 1, \ldots, l'-1$, by (4.10) and since $|S_i| > 0$ for i in this range. Therefore

$$d + l' - 1 \leq \deg P_{l'-1} \leq d + (l'-1)(d-1) = 1 + l'(d-1),$$

and (5.3) follows. ∎

In the case when $d = 2$, $\mathcal{H}$ is an elliptic curve. When $k = \mathbb{Q}$, Mazur [14] proved that the order of $\infty_1 - \infty_2$, when finite, is among 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12. By (5.3), the only possible values for $l' = l'(\alpha)$ are 1, 2, 3, 4, 5, 6, 7, 8, 9, 11. The length $l = l(\alpha)$ of the shortest period is l', or possibly $2l'$ when l' is odd, so that the only possibilities for l are 1, 2, 3, 4, 5, 6, 7, 8, 9,

10, 11, 14, 18, 22. These values were already mentioned by Schinzel [21] as a consequence of a conjecture of Nagell [17].

When $d = 2$ and k is an algebraic number field of degree δ, the order of $\infty_1 - \infty_2$, when finite, is $\le c(\delta)$ according to Merel [15]. Therefore $l'(\alpha)$, $l(\alpha)$ are below some constant $\widetilde{c}(\delta)$.

Schinzel [21] proved the following

LEMMA 8. *Suppose $D(X) \in k[X]$ is irreducible over any quadratic extension of k. Then* (3.3) *has no nontrivial solution.*

We will reproduce Schinzel's argument, a special case of which was essentially given by Tchebycheff [23].

Proof. Let $Y, Z \in k[X]$ be a nontrivial solution of (3.3), with Y of minimum degree. After multiplication by a nonzero constant we may suppose that Y has leading coefficient 1. Now (3.3) yields

$$DZ^2 = Y^2 - a = (Y - b)(Y + b)$$

with $b^2 = a$. Since by hypothesis D is irreducible over $K = k(b)$, we have $D \mid (Y - b)$ or $D \mid (Y + b)$; say $D \mid (Y - b)$, i.e., $Y - b = DV$ with $V \in K[X]$. Then $Z^2 = V(Y + b)$, and since $Y - b$, $Y + b$ are coprime, with leading coefficient 1, we have $V = Z'^2$, $Y + b = Y'^2$ with $Y', Z' \in K[X]$. Here $Y'^2 - DZ'^2 = 2b$, so that Y', Z' is a nontrivial solution of (5.1). By Lemma 6, Y', Z' is proportional to a nontrivial solution $\widetilde{Y}, \widetilde{Z}$ of (3.3) with coefficients in k. The relation $\deg \widetilde{Y} = \deg Y' = \frac{1}{2} \deg Y$ now contradicts the minimality of $\deg Y$. ∎

When $k = \mathbb{Q}$, many polynomials D of degree > 2 have the property stipulated in Lemma 8. When $D \in \mathbb{Q}[X]$ is this type and $k \supset \mathbb{Q}$, a solution Y', Z' with components in $k[X]$ of (5.1) by Lemma 6 would give rise to a solution $\widetilde{Y}, \widetilde{Z} \in \mathbb{Q}[X]$ of (3.3), which is impossible. We may conclude that given a field k of characteristic 0, there are quadratic elements in $k((X^{-1}))$ whose continued fraction is not pseudoperiodic. The same is true when $\operatorname{char} k = p$ and k is transcendental over $\mathbb{F}_p$.

6. The approximation spectrum. Let $\alpha \notin k(X)$ be given. When $|\alpha - P/Q| < |Q|^{-2}$ with coprime P, Q having $|Q| > 1$, define $c = c(P/Q)$ by $|\alpha - P/Q| = |Q|^{-1-c}$, i.e., by

$$|\alpha Q - P| = |Q|^{-c}. \tag{6.1}$$

Endow $[1, \infty]$ with the usual topology, where in particular the sets $(r, \infty]$ are a neighborhood basis of ∞. We define the *approximation spectrum* $S(\alpha)$ to consist of the elements $u \in [1, \infty]$ such that every neighborhood of u contains $c(P/Q)$ for infinitely many fractions P/Q. In view of Legendre's

Theorem, $S(\alpha)$ can be defined in terms of the convergents of α: it is the set of limit points of the sequence

$$c_m := c(P_m/Q_m) \quad (m = 0, 1, 2, \ldots), \tag{6.2}$$

where P_m/Q_m are the convergents. It is a closed subset of $[1, \infty]$. Borrowing from classical language, we might say that α is a "Liouville element" if $\infty \in S(\alpha)$. It is well known that Liouville's Theorem holds in our context: when α is of degree d over $k(X)$, then $S(\alpha) \subset [1, d-1]$. "Roth's Theorem" holds for algebraic α precisely when $S(\alpha) = \{1\}$. We define the *approximation exponent* $r(\alpha)$ as the maximum element of $S(\alpha)$ (possibly ∞). Then when $r(\alpha)$ is finite, given $\varepsilon > 0$ there are infinitely many approximations P/Q with $|\alpha - P/Q| < |Q|^{-r(\alpha)-1+\varepsilon}$, but only finitely many with $|\alpha - P/Q| < |Q|^{-r(\alpha)-1-\varepsilon}$.

We introduce an equivalence relation coarser than $\approx$: We set $\alpha \sim \beta$ if there is a relation (2.2) with $R, S, T, U \in k[X]$ and $RU - ST \neq 0$.

LEMMA 9. *Suppose $\alpha \notin k(X)$ and $\alpha \sim \beta$. Then $S(\alpha) = S(\beta)$.*

Proof. Suppose (2.2) holds. Let P, Q be coprime with $|Q| > 1$ and $|\alpha - P/Q| < |Q|^{-2}$. We have

$$\begin{aligned} &|(R\alpha + S)(TP + UQ) - (T\alpha + U)(RP + SQ)| \\ &\qquad = |RU - ST| \cdot |\alpha Q - P| = |RU - ST| \cdot |Q|^{-c(P/Q)}. \end{aligned} \tag{6.3}$$

If $|Q|$ is large, the right hand side will be small, and $|R\alpha + S| \cdot |TP + UQ| = |T\alpha + U| \cdot |RP + SQ|$. Hence $P' = RP + SQ$, $Q' = TP + UQ$ will have $|P'| = |\beta| \cdot |Q'|$. From this and the first equation of

$$RQ' - TP' = Q(RU - ST), \qquad SQ' - UP' = -P(RU - ST), \tag{6.4}$$

we may infer that

$$\max(|R| \cdot |Q'|, |T| \cdot |\beta| \cdot |Q'|) \geq |Q| \cdot |RU - ST|,$$

hence

$$|Q'| \geq s_1 |Q|$$

with a constant $s_1 > 0$ independent of Q. On the other hand, $|P| = |Q| \cdot |\alpha|$, so that $|Q'| \leq s_2 |Q|$. By (6.4), $\gcd(P', Q')$ divides $RU - ST$, so that if $P'/Q' = P''/Q''$ with coprime P'', Q'', then

$$s_3 |Q| \leq |Q''| \leq s_2 |Q|. \tag{6.5}$$

Now (6.3) yields

$$|Q'\beta - P'| = |RU - ST| \cdot |T\alpha + U|^{-1} |Q|^{-c(P/Q)}. \tag{6.6}$$

Hence as long as $c(P/Q)$ remains in a fixed (bounded) interval,

$$|Q''\beta - P''| = t|Q''|^{-c(P/Q)}$$

with t bounded above and below, depending on the interval. Therefore if $c_\beta(P''/Q'')$ is the exponent with respect to β, then

$$c_\beta(P''/Q'') = c(P/Q) + v/\log|Q''|$$

where v is bounded. Thus if we have a sequence of fractions P/Q whose $c(P/Q)$ tends to some finite $u \in S(\alpha)$, then the corresponding sequence $c_\beta(P''/Q'')$ will also tend to u. Therefore when finite $u \in S(\alpha)$, then $u \in S(\beta)$. In a similar way, (6.5), (6.6) imply that $\infty \in S(\alpha)$ implies $\infty \in S(\beta)$. ∎

7. Algebraic elements in the characteristic p case. From now on we will suppose that $\operatorname{char} k = p > 0$. Further q will denote a positive power of p.

Suppose $\alpha \in k((X^{-1}))\backslash k(X)$ is algebraic over $k(X)$. Following Lasjaunias [8], we will say that

$$\alpha \text{ is of Class I} \quad \text{if} \quad \alpha \sim \alpha^q \text{ for some } q.$$

Otherwise we will say that α is of Class II. We introduce a subclass of Class I as follows:

$$\alpha \text{ is of Class IA} \quad \text{if} \quad \alpha \approx \alpha^q \text{ for some } q.$$

THEOREM 4. *$\alpha \approx \alpha^q$ precisely if the continued fraction of α is of the form*

$$\alpha = [A_0, \ldots, A_{n-1}, C_1, C_2, \ldots] \tag{7.1}$$

where for some $t \in \mathbb{N}$ and some $a \in k^\times$ we have

$$C_{j+t} = \begin{cases} aC_j^q & \textit{when } j \textit{ is odd}, \\ a^{-1}C_j^q & \textit{when } j \textit{ is even}. \end{cases} \tag{7.2}$$

Hence when t is even, the continued fraction is

$$[A_0, \ldots, A_{n-1}, \underleftrightarrow{C_1, \ldots, C_t}, \underleftrightarrow{aC_1^q, a^{-1}C_2^q, \ldots, a^{-1}C_t^q}, \underleftrightarrow{a^{q+1}C_1^{q^2}, \ldots, a^{-q-1}C_t^{q^2}}, \underleftrightarrow{a^{q^2+q+1}C_1^{q^3}, \ldots}, \ldots], \tag{7.3}$$

and when t is odd, it is

$$[A_0, \ldots, A_{n-1}, \underleftrightarrow{C_1, \ldots, C_t}, \underleftrightarrow{aC_1^q, a^{-1}C_2^q, \ldots, aC_t^q}, \underleftrightarrow{a^{q-1}C_1^{q^2}, \ldots, a^{q-1}C_t^{q^2}}, \underleftrightarrow{a^{q^2-q+1}C_1^{q^3}, \ldots}, \ldots]. \tag{7.4}$$

Possibly there are no initial terms $A_0, \ldots, A_{n-1}$.

Proof. The map $\alpha \mapsto \alpha^q$ is an isomorphism of $k((X^{-1}))$ into itself. Therefore

$$[B_0, B_1, \ldots, B_m]^q = [B_0^q, B_1^q, \ldots, B_m^q],$$

and a corresponding relation holds for infinite continued fractions.

With α given by (7.1), (7.2), we have $\alpha_n = [C_1, C_2, \ldots]$, $\alpha_{n+t} = [C_{t+1}, C_{t+2}, \ldots]$. Then $\alpha_n^q = [C_1^q, C_2^q, \ldots]$, and (7.2) yields $\alpha_{n+t} = [aC_1^q, a^{-1}C_2^q, \ldots] = a\alpha_n^q$, so that $\alpha_{n+t} \approx \alpha_n^q$. Since $\alpha \approx \alpha_n$, hence $\alpha^q \approx \alpha_n^q$, and since $\alpha \approx \alpha_{n+t}$, we obtain $\alpha \approx \alpha^q$.

Conversely, when $\alpha \approx \alpha^q$, then by Theorem 1 there are n, m, and $b \in k^\times$ such that

$$\alpha = [A_0, \ldots, A_{n-1}, C_1, C_2, \ldots], \qquad \alpha^q = [B_0, \ldots, B_{m-1}, bC_1, b^{-1}C_2, \ldots].$$

Then $\alpha^q = [A_0^q, \ldots, A_{n-1}^q, C_1^q, C_2^q, \ldots]$. If $m \geq n$, say $m = n+t$, we have $C_1^q = B_n, \ldots, C_t^q = B_{m-1}$, $C_{t+1}^q = bC_1, C_{t+2}^q = b^{-1}C_2, \ldots$ Therefore $\deg C_1 = q \deg C_{t+1} = q^2 \deg C_{2t+1} = \ldots$, which is impossible. We may conclude that $n > m$, say $n = m + t$. We obtain $bC_1 = A_m^q, \ldots, aC_t = A_{n-1}^q, a^{-1}C_{t+1} = C_1^q, aC_{t+2} = C_2^q, \ldots$, where $a = b$ if t is odd, $a = b^{-1}$ if t is even. Thus indeed (7.2) holds. ∎

Since not only $\alpha \approx \alpha_n$, but in fact α, α_n are connected by a fractional linear transformation of determinant $(-1)^n$, and similar for $\alpha^q \approx \alpha_n^q$, and in $\alpha \approx \alpha_{n+t}$, the implied transformation is of determinant $(-1)^{n+t}$, and the first half of the proof of Theorem 4 shows that (6.1), (6.2) imply

$$\alpha = (R\alpha^q + S)/(T\alpha^q + U) \tag{7.5}$$

with $RU - ST = (-1)^t a$. This can be generalized as follows.

LEMMA 10. *Suppose $t \in \mathbb{N}$, and $C_1, \ldots, C_t, B, D$ in $k[X]$ are given, with $BC_1, C_2, \ldots, C_t$ of positive degree. Set $\varepsilon = 1$ when t is even, $\varepsilon = 0$ when t is odd. Let B, D have*

$$D \mid C_i^{q-\varepsilon} \text{ when } i \text{ is odd}, \qquad 1 \leq i \leq t,$$
$$B \mid C_i^{q-\varepsilon} \text{ when } i \text{ is even}, \qquad 1 \leq i \leq t.$$

Set

$$C_{j+t} = \begin{cases} (B/D)C_j^q & \text{when } j \in \mathbb{N} \text{ is odd}, \\ (D/B)C_j^q & \text{when } j \in \mathbb{N} \text{ is even}. \end{cases}$$

Then each C_j $(j = 1, 2, \ldots)$ is a polynomial of positive degree, and α as given by (7.1) satisfies a relation (7.5) with $RU - ST = (-1)^t BD$.

Proof. The expansion of α is like (7.3) or (7.4), but with B/D in place of a. Setting $\delta = 1$ when t is even, $\delta = -1$ when t is odd, we have, for $l = 1, 2, \ldots$,

$$C_{j+lt} = \begin{cases} (B/D)^{q^{l-1}+\delta q^{l-2}+\ldots+\delta^{l-1}} C_j^{q^l} & \text{when } j \text{ is odd}, \\ (D/B)^{q^{l-1}+\delta q^{l-2}+\ldots+\delta^{l-1}} C_j^{q^l} & \text{when } j \text{ is even}. \end{cases}$$

Since $(q^{l-1} + \delta q^{l-2} + \ldots + \delta^{l-1})(q - \varepsilon) < q^l$, we see that C_{j+lt} is a polynomial divisible by C_j $(1 \leq j \leq t,\ l \in \mathbb{N})$, and in fact C_{1+lt} is divisible by BC_1. The

complete denominators α_n, α_{n+t} have

$$\alpha_{n+t} = [(B/D)C_1^q, (D/B)C_2^q, \ldots] = (B/D)[C_1^q, C_2^q, \ldots] = (B/D)\alpha_n^q.$$

Thus α_{n+t}, α_n^q are connected by a fractional linear transformation of determinant BD. The desired conclusion follows. ∎

EXAMPLE. $n = 0$, $t = 2$, $C_1 = 1$, $C_2 = X$, $B = X^{q-1}$, $D = 1$. Here

$$\alpha = [1, X, X^{q-1}, X, X^{q^2-1}, X, X^{q^3-1}, \ldots], \tag{7.6}$$

and α has

$$\alpha = ((X^q + X^{q-1})\alpha^q + 1)/(X^q\alpha^q + 1).$$

8. The approximation spectrum of elements of Class IA

THEOREM 5. *Suppose α is as in Theorem 4. Set $s_j = \deg C_j$ $(j = 1, \ldots, t)$, $s_{t+1} = qs_1$, and $q_j = s_{j+1}/s_j$ $(j = 1, \ldots, t)$. Extend $\{q_j\}_{1 \le j \le t}$ periodically by $q_{j+t} = q_j$ $(j = 1, 2, \ldots)$, and put*

$$u_j = 1 + (q-1)/(1 + q_j + q_jq_{j+1} + \ldots + q_jq_{j+1}\ldots q_{j+t-2}) \quad (j = 1, \ldots, t). \tag{8.1}$$

Then the sequence $\{c_m\}$ introduced in (6.2) *has*

$$\lim_{l\to\infty} c_{n-2+j+lt} = u_j \quad (j = 1, \ldots, t), \tag{8.2}$$

so that $S(\alpha)$ consists of $u_1, \ldots, u_t$ (which are not necessarily distinct). Finally,

$$u_1 \ldots u_t = q. \tag{8.3}$$

Proof. Write α also as $[R_0, R_1, \ldots]$ and set $r_m = \deg R_m$, so that $\deg Q_m = r_1 + \ldots + r_m$, and by (1.12),

$$c_m = \deg Q_{m+1}/\deg Q_m = 1 + r_{m+1}/(r_1 + \ldots + r_m). \tag{8.4}$$

When

$$m = n - 1 + j + lt \quad (j = 1, \ldots, t), \tag{8.5}$$

we have

$$r_m = s_jq^l, \quad r_{m+1} = s_{j+1}q^l. \tag{8.6}$$

Therefore, setting $u = r_1 + \ldots + r_{n-1}$ when $n > 1$, $u = 0$ when $n = 1$, and $u = -r_0$ when there are no terms $A_0, \ldots, A_{n-1}$, we have

$$\begin{aligned} r_1 + \ldots + r_m &= u + (s_1 + \ldots + s_t)(1 + q + \ldots + q^{l-1}) + (s_1 + \ldots + s_j)q^l \\ &= q^l((s_1 + \ldots + s_t)/(q-1) + s_1 + \ldots + s_j) + O(1) \\ &= (q^l/(q-1))(q(s_1 + \ldots + s_j) + s_{j+1} + \ldots + s_t) + O(1), \end{aligned}$$

with $s_{j+1}+\ldots+s_t$ interpreted as 0 when $j=t$. This equals

$$\begin{aligned}(s_{j+1}q^l/(q-1))&(1+q_{j+1}+q_{j+1}q_{j+2}+\ldots+q_{j+1}q_{j+2}\ldots q_{j+t-1})+O(1)\\ &=s_{j+1}q^l/(u_{j+1}-1)+O(1)=r_{m+1}/(u_{j+1}-1)+O(1),\end{aligned}$$

where we have set $u_{t+1}=u_1$. Therefore as m runs through the numbers (8.5) with $l=1,2,\ldots$,

$$r_{m+1}/(r_1+\ldots+r_m)\to u_{j+1}-1,$$

and (8.4) gives

$$c_{n-1+j+lt}\to u_{j+1}.$$

This proves (8.2).

By (8.6),

$$\begin{aligned}\deg Q_{n-1+lt}&=u+(s_1+\ldots+s_t)(1+q+\ldots+q^{l-1})\\ &=q^l(s_1+\ldots+s_t)/(q-1)+O(1).\end{aligned}\tag{8.7}$$

Therefore

$$\begin{aligned}\prod_{j=1}^{t}c_{n-2+j+lt}&=\prod_{j=1}^{t}(\deg Q_{n-1+j+lt}/\deg Q_{n-2+j+lt})\\ &=\deg Q_{n-1+(l+1)t}/\deg Q_{n-1+lt}.\end{aligned}$$

By (8.7), this tends to q as $l\to\infty$. On the other hand, by (8.2), the limit is $u_1\ldots u_t$. Now (8.3) follows. ∎

By the theorem, the approximation exponent $r(\alpha)=\max(u_1,\ldots,u_t)$. Hence when $t=1$ we have $r(\alpha)=q$, whereas for $t>1$ we have

$$q^{1/t}\le r(\alpha)<q$$

by (8.3), and since each $u_i>1$.

LEMMA 11. *Suppose $t>1$ and r is a rational number in $q^{1/t}\le r<q$. Then polynomials $C_1,\ldots,C_t$ can be chosen such that the α of Theorem 5 has $r(\alpha)=r$.*

Proof. Initially suppose that $r=q^{1/t}$, so that $r\in\mathbb{N}$, $r>1$. Pick $C_1,\ldots,C_t$ with $\deg C_j=q^{j/t}=r^j$. In the notation of Theorem 5, $s_j=r^j$, $q_j=r$ $(1\le j\le t)$. Therefore $u_1=\ldots=u_t$, whence by (8.3), $u_j=q^{1/t}=r$ $(1\le j\le t)$, and $r(\alpha)=r$.

Now suppose that $r>q^{1/t}$. Given positive reals x,y set

$$q_1=y,\quad q_2=\ldots=q_{t-1}=x,\quad q_t=q/(q_1\ldots q_{t-1}).\tag{8.8}$$

(When $t=2$ we have $q_1=y$, $q_2=q/q_1=q/y$, and no dependency on x.) Then each u_j as given by (8.1) becomes a continuous function $u_j=u_j(x,y)$ of x,y. Note that

$$u_1(x,y)=1+(q-1)/(1+y(1+x+\ldots+x^{t-2})),$$

and pick $y = y(r, x)$ with $u_1(x, y) = r$. When $x = q^{1/t}$ we have $u_1(x, q^{1/t}) = u_1(q^{1/t}, q^{1/t}) = q^{1/t}$. Therefore, since $u_1(x, y)$ is a decreasing function in y, and since $r > q^{1/t}$, we have $y(r, q^{1/t}) < q^{1/t}$. Thus with $x = q^{1/t}$, $y = y(r, q^{1/t})$, we have $q_1 < q_2 = \ldots = q_{t-1} < q_t$ by (8.8), hence

$$u_1 > \max(u_2, \ldots, u_t) \tag{8.9}$$

by (8.1). When x is close to $q^{1/t}$ and $y = y(r, x)$, we still have (8.9). Choose x rational with this property; then $q_1, \ldots, q_t$ will be rational. Set $s_1 = m$, $s_2 = mq_1, \ldots, s_t = mq_1 \ldots q_{t-1}$, with $m \in \mathbb{N}$ chosen so that $s_1, \ldots, s_t$ lie in $\mathbb{N}$. Now if $C_1, \ldots, C_t$ are polynomials of respective degrees $s_1, \ldots, s_t$, the quantities $u_1, \ldots, u_t$ of Theorem 5 have (8.9), and $r(\alpha) = u_1 = u_1(x, y) = r$. ∎

For elements of Class IA see also Voloch [24], where approximations to, e.g., α with $\alpha = \alpha^q + B(X)$, $B(X) \in k[X]$, are being considered.

9. Examples of elements of Class I. Elements of Class IA are now well understood, but general elements of Class I show a much more complicated behavior. Voloch in important work [24], [25] showed in particular that for such elements α,

$$|\alpha Q - P| \cdot |Q|^{-r(\alpha)} \tag{9.1}$$

is bounded from below by a positive constant. According to de Mathan [13], there is a constant $c(\alpha)$ such that there are P, Q with arbitrarily large $|Q|$ having $|\alpha Q - P| \cdot |Q|^{-r(\alpha)} < c(\alpha)$. Further the members of $S(\alpha)$ are rational, and the only possible accumulation point of $S(\alpha)$ is 1. Further results were recently obtained by Lasjaunias [9].

When $d = 3$, then $1, \alpha, \alpha^q, \alpha^{q+1}$ are linearly dependent over $k(X)$, so that α is necessarily of Class I. Baum and Sweet [2] gave an example of a cubic over $\mathbb{F}_2(X)$ with partial denominators of degree ≤ 2. They characterized [3] elements of $\mathbb{F}_2((X^{-1}))$ having $A_0 = 0$ and $\deg A_n = 1$ for $n \geq 1$. Mills and Robbins [16] gave a complete description of the continued fraction of the Baum–Sweet cubic, which shows in particular that it is of Class I but not IA. They explicitly constructed continued fractions of algebraic elements over $\mathbb{F}_p(X)$ for an arbitrary prime p whose partial denominators $A_1, A_2, \ldots$ have degree 1. Perhaps more importantly, they presented an efficient algorithm to compute the continued fraction of elements of Class I, which in some cases allows one to find an explicit description of its partial denominators.

Here I present a particularly simple example obtained with their algorithm. We consider $\alpha = k((X^{-1}))$ having

$$\alpha = (X/a)\alpha^q/(b\alpha^q + a) \tag{9.2}$$

with $a, b \in k^\times$, so that

$$b\alpha^q - (X/a)\alpha^{q-1} + a = 0.$$

It is easily seen that this equation has a unique solution $\alpha = (X/ab) + \ldots$ in $k((X^{-1}))$, and since the equation is irreducible over $k(X)$, $\deg \alpha = q$.

LEMMA 12. $\alpha = [A_0, A_1, \ldots]$ *where*

$$A_{2l} = (-1)^l X/ab, \tag{9.3}$$

$$A_{2l+1} = \begin{cases} -b^2 A_l^q/X & \textit{when } l \textit{ is even}, \\ a^2 A_l^q/X & \textit{when } l \textit{ is odd}. \end{cases} \tag{9.4}$$

Proof. When $\alpha = (R\beta+S)/(T\beta+U)$, we will use the notation $\alpha = \mathcal{M}\beta$ with

$$\mathcal{M} = \begin{pmatrix} R & S \\ T & U \end{pmatrix}.$$

We will write $\alpha_0 = \alpha$, and we will show that

$$\alpha_{2l} = \mathcal{M}_l \alpha_l^q, \tag{9.5}$$

$$\alpha_{2l+1} = \mathcal{N}_l \alpha_l^q, \tag{9.6}$$

$$\alpha_{2l+1} = \mathcal{L}_l \alpha_{l+1}^q, \tag{9.7}$$

where for l even

$$\mathcal{M}_l = \begin{pmatrix} X/a & 0 \\ b & a \end{pmatrix}, \quad \mathcal{N}_l = \begin{pmatrix} b & a \\ 0 & -X/b \end{pmatrix}, \quad \mathcal{L}_l = \begin{pmatrix} bA_l^q + a & b \\ -X/b & 0 \end{pmatrix},$$

and for l odd,

$$\mathcal{M}_l = \begin{pmatrix} -X/b & 0 \\ a & b \end{pmatrix}, \quad \mathcal{N}_l = \begin{pmatrix} a & b \\ 0 & X/a \end{pmatrix}, \quad \mathcal{L}_l = \begin{pmatrix} aA_l^q + b & a \\ X/a & 0 \end{pmatrix}.$$

Note that by (9.2), we have (9.5) for $l = 0$. Suppose that either $l = 0$, or $l > 0$ and (9.5) is true for l, and $A_0, \ldots, A_{2l-1}$ are determined by the recursions (9.3), (9.4). Then a Mills–Robbins step of Type ([2]) 1 leads to (9.6) for l, as well as to (9.3) for l, determining A_{2l}.

When (9.6) is true for l, a Mills–Robbins step of Type 2 leads to (9.7).

When (9.7) is true for l and $A_0, \ldots, A_{2l}$ are determined by (9.3), (9.4), a Mills–Robbins step of Type 1 leads to (9.5) for $l+1$ in place of l, as well as to (9.4) for l, determining A_{2l+1}. ∎

We wish to determine the approximation spectrum of the α of Lemma 12. But first we will deal with the much easier element α of the example at the end of Section 7. Writing $\alpha = [R_0, R_1, \ldots]$ with $\deg R_m = r_m$, we observe (8.4). By (7.6), $r_{2n-1} = 1$, $r_{2n} = q^n - 1$. Hence $c_{2n} = 1 + 1/(r_1 + \ldots + r_{2n})$ tends to 1. On the other hand, $r_1 + \ldots + r_{2n-1} = 1 + q + \ldots + q^{n-1} =$

([2]) Steps of Type 1, 2 have nothing to do with Classes I, II. Note that a Mills–Robbins step of Type 1 requires $T \neq 0$, whereas a Type 2 step can be performed when $T = 0$.

$(q^n-1)/(q-1)$, so that $c_{2n-1}=1+(q^n-1)/((q^n-1)/(q-1))=q$. We may conclude that $S(\alpha)=\{1,q\}$, $r(\alpha)=q$. If α were as in Theorem 5 with $t=1$, we would have $S(\alpha)=\{q\}$, and if it were with $t>1$, we would have $r(\alpha)<q$. Therefore α is of Class I, but not IA.

LEMMA 13. *The approximation spectrum of α given by* (9.2) *consists of* $1, q-1$, *and the numbers*

$$1+(q-2)q^{-v_s}/((q-1)(1+q^{-v_1}+\ldots+q^{-v_{s-1}})+q^{-v_s}) \tag{9.8}$$

where $s>0$, $0<v_1<\ldots<v_s$.

In particular, the spectrum is infinite, so that α is of Class I, but not IA.

Proof. Write $r_m=\deg A_m$, so that

$$r_m=\begin{cases}1 & \text{when } m \text{ is even,}\\ qr_{(m-1)/2}-1 & \text{when } m \text{ is odd.}\end{cases} \tag{9.9}$$

We claim that for odd m,

$$r_m=q^l-q^{l-1}-\ldots-q-1 \quad \text{when } 2^l \parallel (m+1). \tag{9.10}$$

When $l=1$, then $(m-1)/2$ is even, so that by (9.9), indeed $r_m=q-1$. When (9.10) is true for $l-1$, and $2^l\parallel(m+1)$, then $2^{l-1}\parallel((m-1)/2+1)$, so that $r_{(m-1)/2}=q^{l-1}-q^{l-2}-\ldots-1$, and $r_m=q(q^{l-1}-q^{l-2}-\ldots-1)-1=q^l-q^{l-1}-\ldots-1$ by (9.9).

Now suppose

$$n+1=2^{u_0}+2^{u_1}+\ldots+2^{u_s} \tag{9.11}$$

with $u_0>u_1>\ldots>u_s>0$. Then $2^{u_s}\parallel(n+1)$, so that

$$r_n=q^{u_s}-q^{u_s-1}-\ldots-q-1=((q-2)q^{u_s}+1)/(q-1). \tag{9.12}$$

On the other hand,

$$r_0+\ldots+r_{n-1}=s_0+s_1(q-1)+s_2(q^2-q-1)+\ldots+s_{u_0}(q^{u_0}-q^{u_0-1}-\ldots-1)$$

where s_j is the number of m, $0\le m<n$, with $2^j\parallel(m+1)$, i.e., it is the number of m, $1\le m\le n$, with $2^j\parallel m$. Our expression equals

$$\begin{aligned}q^{u_0}s_{u_0}+q^{u_0-1}(s_{u_0-1}-s_{u_0})+q^{u_0-2}(s_{u_0-2}-s_{u_0-1}-s_{u_0})+\ldots\\ +q(s_1-s_2-\ldots-s_{u_0})+(s_0-s_1-\ldots-s_{u_0}).\end{aligned}$$

But $s_j=t_j-t_{j+1}$, where t_j is the number of m, $1\le m\le n$, with $2^j\mid m$, and therefore

$$\begin{aligned}s_j-s_{j+1}-\ldots-s_{u_0}&=t_j-t_{j+1}-(t_{j+1}-t_{j+2})-\ldots-(t_{u_0}-t_{u_0+1})\\&=t_j-2t_{j+1},\end{aligned}$$

since $t_{u_0+1}=0$. In view of (9.11), $t_j=[n/2^j]$ is given by

$$t_j = \begin{cases} 2^{u_0-j} + \ldots + 2^{u_i-j} & \text{when } u_{i+1} < j \le u_i \ (i = 0, 1, \ldots, s-1), \\ 2^{u_0-j} + \ldots + 2^{u_s-j} - 1 & \text{when } 0 \le j \le u_s. \end{cases}$$

We may infer that

$$t_j - 2t_{j+1} = \begin{cases} 1 & \text{if } j = u_i \ (i = 0, \ldots, s-1), \text{ or if } j < u_s, \\ 0 & \text{otherwise.} \end{cases}$$

Therefore $r_1 + \ldots + r_{n-1} = r_0 + \ldots + r_{n-1} - 1$ equals

$$q^{u_0-1} + q^{u_0-2} + \ldots + 1 - 1 = (q^{u_0} - q)/(q-1) \tag{9.13}$$

when $s = 0$, and

$$\begin{aligned} q^{u_0} + q^{u_1} + \ldots + q^{u_{s-1}} + (q^{u_s-1} + \ldots + q^2 + q) \\ = q^{u_0} + q^{u_1} + \ldots + q^{u_{s-1}} + (q^{u_s} - q)/(q-1) \end{aligned} \tag{9.14}$$

when $s > 0$.

We now apply (8.4) with $m = n - 1$. When $s = 0$, so that $n = 2^{u_0} - 1$, then (9.12), (9.13) yield

$$c_{n-1} = 1 + ((q-2)q^{u_0} + 1)/(q^{u_0} - q).$$

Therefore as n runs through the numbers $q^{u_0} - 1$ $(u_0 = 1, 2, \ldots)$, then c_{n-1} tends to $1 + q - 2 = q - 1$. When $s > 0$, we have

$$c_{n-1} = 1 + ((q-2)q^{u_s} + 1)/((q-1)/(q^{u_0} + \ldots + q^{u_{s-1}} + (q^{u_s} - q)/(q-1))).$$

Set $u_i = u_0 - v_i$ $(i = 1, \ldots, s)$ with $0 < v_1 < \ldots < v_s$. Then (9.11) becomes

$$n = 2^{u_0}(1 + 2^{-v_1} + \ldots + 2^{-v_s}) - 1, \tag{9.15}$$

and

$$c_{n-1} = 1 + ((q-2)q^{-v_s} + q^{-u_0})/((q-1)(1 + q^{-v_1} + \ldots + q^{-v_{s-1}}) + q^{-v_s} - q^{1-u_0}).$$

When $v_1, \ldots, v_s$ are given and n runs through the numbers (9.15) with $u_0 = v_s + 1, v_s + 2, \ldots$, then c_{n-1} tends to (9.8). When v_s is under a fixed bound, so that also s is bounded, there are finitely many choices for s and $v_1, \ldots, v_s$, and the limit points are among the numbers (9.8). On the other hand, as $v_s \to \infty$ (where also s may vary), then also $u_0 > v_s$ tends to ∞, and c_{n-1} approaches 1. ∎

Say $\alpha = [A_0, A_1, \ldots]$ is an *l-fraction* if the degrees $r_m = \deg A_m$ have

$$r_m = \begin{cases} 1 & \text{when } m \not\equiv 1 \pmod{l}, \\ qr_{(m-1)/l} - 1 & \text{when } m \equiv 1 \pmod{l}. \end{cases}$$

By (9.9), our α with (9.2) is a 2-fraction. I conjecture that when the field k is sufficiently large, then for every $l > 1$ there are l-fractions α with $\alpha \sim \alpha^q$.

10. The Mills–Robbins continued fraction. Let $\Omega = \emptyset$, $\Omega_1 = X$, and for $n \geq 2$ let Ω_n be the finite sequence of polynomials

$$\Omega_n = \Omega_{n-1}, -X, \Omega_{n-2}^{(3)}, -X, \Omega_{n-1},$$

where commas signify juxtaposition of sequences, and $\Omega_k^{(3)}$ is obtained by cubing every element of Ω_k. Since Ω_{n-1} appears as the initial segment of Ω_n, we may define Ω_∞ as the infinite sequence having each Ω_n as an initial segment. Set

$$\alpha = [0, \Omega_\infty].$$

Mills and Robbins [16] had conjectured and Buck and Robbins [4] proved the following intriguing theorem:

Suppose $k = \mathbb{F}_3$. Then α is the unique root in $k((X^{-1}))$ with

$$\alpha^4 + \alpha^2 - X\alpha + 1 = 0. \tag{10.1}$$

Lasjaunias [8] gave another proof in a wider setting. Here we will rearrange ideas of [4] to give a rather short argument.

Writing

$$\mathcal{N}_n = \begin{pmatrix} P_{n-1} & P_n \\ Q_{n-1} & Q_n \end{pmatrix} \qquad (n = -1, 0, 1, \ldots),$$

we have

$$\mathcal{N}_{-1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \mathcal{N}_n = \mathcal{N}_{n-1} \begin{pmatrix} 0 & 1 \\ 1 & A_n \end{pmatrix} \qquad (n = 0, 1, \ldots)$$

where $P_l = P_l(A_0, \ldots, A_l)$, $Q_l = Q_l(A_0, \ldots, A_l)$ $(l = 0, 1, \ldots)$.

In the case when $A_0, A_1, \ldots$ is the sequence Ω_∞, set $\mathcal{M}_0 = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$, $\mathcal{M}_1 = \left(\begin{smallmatrix} 0 & 1 \\ 1 & X \end{smallmatrix}\right)$,

$$\mathcal{M}_n = \mathcal{M}_{n-1} \begin{pmatrix} 0 & 1 \\ 1 & -X \end{pmatrix} \mathcal{M}_{n-2}^{(3)} \begin{pmatrix} 0 & 1 \\ 1 & -X \end{pmatrix} \mathcal{M}_{n-1} \qquad (n \geq 2), \tag{10.2}$$

where $\mathcal{M}_l^{(3)}$ is obtained from $\mathcal{M}_l$ by cubing each entry. Then

$$\mathcal{M}_n = \mathcal{N}_{r(n)}$$

where $r(n)$ is the number of elements of Ω_n. Since $\mathcal{M}_n$ is symmetric, we may write

$$\mathcal{M}_n = \begin{pmatrix} R_n & S_n \\ S_n & T_n \end{pmatrix}, \tag{10.3}$$

and notice that $R_n/S_n, S_n/T_n$ are consecutive convergents of the continued fraction of α. Therefore

$$\lim_{n\to\infty} R_n/S_n = \lim_{n\to\infty} S_n/T_n = \alpha, \qquad \lim_{n\to\infty} R_n/T_n = \alpha^2.$$

All this is already in [4].

Now set

$$\mathbf{R}_n = \begin{pmatrix} R_n \\ S_n \\ T_n \end{pmatrix}, \quad \mathbf{R}_n^3 = \begin{pmatrix} R_n^3 \\ S_n^3 \\ T_n^3 \end{pmatrix}.$$

Let

$$\mathcal{T} = \begin{pmatrix} -1/X & 1 & 1/X \\ 0 & 0 & 1 \\ 1/X & 1 & X - 1/X \end{pmatrix},$$

and let $Q(\mathbf{R}) = Q(R, S, T)$ be the quadratic form with

$$Q(\mathbf{R}) = (R - T)^2 - XST.$$

LEMMA 14. (i) $Q(\mathbf{R}_n) = 0$ $(n = 0, 1, \ldots)$,
(ii) $\mathbf{R}_{n+1} = \mathcal{T}\mathbf{R}_n^3$ $(n = 0, 1, \ldots)$.

The Buck–Mills–Robbins Theorem follows: we have

$$T_n^{-2} Q(\mathbf{R}_n) = (R_n/T_n)^2 - 2(R_n/T_n) + 1 - X(S_n/T_n) = 0.$$

Taking the limit as $n \to \infty$, we obtain $\alpha^4 - 2\alpha^2 + 1 - X\alpha = 0$, hence (10.1).

Proof (of Lemma 14). Both (i), (ii) are easily checked for $n = 0$. The induction step $n - 1 \Rightarrow n$ is as follows:

$$\begin{aligned} Q(\mathcal{T}\mathbf{R}) &= \left(-\frac{R}{X} + S + \frac{T}{X} - \frac{R}{X} - S - TX + \frac{T}{X}\right)^2 \\ &\quad - XT\left(\frac{R}{X} + S + TX - \frac{T}{X}\right) \\ &= \left(\frac{R}{X} - \frac{T}{X} - TX\right)^2 - RT - STX - T^2X^2 + T^2 \\ &= X^{-2}(R - T)^2 - STX = X^{-2}((R - T)^2 - X^3ST). \end{aligned}$$

From (i), (ii) with $n - 1$ in place of n we get

$$\begin{aligned} Q(\mathbf{R}_n) &= Q(\mathcal{T}\mathbf{R}_{n-1}^3) = X^{-2}((R_{n-1}^3 - T_{n-1}^3) - X^3 S_{n-1}^3 T_{n-1}^3) \\ &= X^{-2} Q(\mathbf{R}_{n-1})^3 = 0, \end{aligned}$$

so that (i) holds for n.

By (10.2) and (10.3) there is a 3-vector of functions $\mathcal{F}(\mathbf{R}, \mathbf{r})$ which is homogeneous of degree 2 and 1 in

$$\mathbf{R} = \begin{pmatrix} R \\ S \\ T \end{pmatrix} \quad \text{and} \quad \mathbf{r} = \begin{pmatrix} r \\ s \\ t \end{pmatrix}$$

respectively, such that

$$\mathbf{R}_n = \mathcal{F}(\mathbf{R}_{n-1}, \mathbf{R}_{n-2}^3) \quad (n = 2, 3, \ldots).$$

Applying this for $n+1$, we may rewrite (ii) as

$$\mathbf{R}_{n+1} = \mathcal{F}(\mathbf{R}_n, \mathbf{R}_{n-1}^3) = \mathcal{T}\mathbf{R}_n^3.$$

In other words, since $\mathbf{R}_{n-1}^3 = \mathcal{T}^{-1}\mathbf{R}_n$ by the case $n-1$ of (ii), we need to check that

$$\mathcal{F}(\mathbf{R}_n, \mathcal{T}^{-1}\mathbf{R}_n) = \mathcal{T}\mathbf{R}_n^3.$$

It suffices to show that

$$\mathcal{F}(\mathbf{R}, \mathcal{T}^{-1}\mathbf{R}) = \mathcal{T}\mathbf{R}^3 \quad \text{whenever } Q(\mathbf{R}) = 0. \tag{10.4}$$

This requires a routine but messy calculation, already implicit at the end of the proof in [4]. From (10.2) we see by matrix multiplication that

$$\mathcal{F}(\mathbf{R}, \mathbf{r}) = \begin{pmatrix} rS^2 + t(R - SX)^2 - sS(R - SX) \\ rST + t(R - SX)(S - TX) + s(S^2 + RT + STX) \\ rT^2 + t(S - TX)^2 - sT(S - TX) \end{pmatrix}.$$

Since

$$\mathcal{T}^{-1}\begin{pmatrix} R \\ S \\ T \end{pmatrix} = \begin{pmatrix} X & 1 + X^2 & -X \\ -1 & X & -1 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} R \\ S \\ T \end{pmatrix} = \begin{pmatrix} RX + S(1 + X^2) - TX \\ -R + SX - T \\ S \end{pmatrix},$$

substitution gives after some calculation that

$$\mathcal{F}(\mathbf{R}, \mathcal{T}^{-1}\mathbf{R}) = \begin{pmatrix} S(STX + S^2 + RT - R^2) \\ ST^2X - RT(R + T) \\ T^3X + S^3 + RST - ST^2 \end{pmatrix}. \tag{10.5}$$

On the other hand,

$$\begin{aligned} \mathcal{T}\mathbf{R}^3 &= \begin{pmatrix} -1/X & 1 & 1/X \\ 0 & 0 & 1 \\ 1/X & 1 & X - 1/X \end{pmatrix}\begin{pmatrix} R^3 \\ S^3 \\ T^3 \end{pmatrix} \\ &= \begin{pmatrix} S^3 - (R - T)^3X^{-1} \\ T^3 \\ R^3X^{-1} + S^3 + T^3(X - X^{-1}) \end{pmatrix}. \end{aligned} \tag{10.6}$$

When $Q(\mathbf{R}) = 0$, the vectors in (10.5), (10.6) are in fact equal. ∎

11. Elements of Class II. For completeness, I just want to recall that Lasjaunias and de Mathan [10] showed that elements α of Class II have $r(\alpha) \le [d/2]$ when $\deg \alpha = d$. In the case when k is finite, they proved [11] the stronger result that $|\alpha Q - P| \ge c|Q|^{-[d/2]}$ with $c = c(\alpha) > 0$; this contains an earlier result of Osgood [18] on elements which satisfy no rational Riccati differential equation.

References

[1] N. H. Abel, *Über die Integration der Differential-Formel $\rho dx/\sqrt{R}$ wenn R und ρ ganze Funktionen sind*, J. Reine Angew. Math. 1 (1826), 185–221.

[2] L. E. Baum and M. M. Sweet, *Continued fractions of algebraic power series in characteristic 2*, Ann. of Math. 103 (1976), 593–610.

[3] —, —, *Badly approximable power series in characteristic 2*, ibid. 105 (1977), 573–580.

[4] M. W. Buck and D. P. Robbins, *The continued fraction expansion of an algebraic power series satisfying a quartic equation*, J. Number Theory 50 (1995), 335–344.

[5] J. H. Davenport, *On the Integration of Algebraic Functions*, Lecture Notes in Comput. Sci. 102, Springer, 1981.

[6] G. H. Halphen, *Traité des fonctions elliptiques et de leurs applications*, Vol. II, Paris, 1891.

[7] Y. Hellegouarch et M. Lozach, *Équation de Pell et points d'ordre fini*, avec "Appendice" de M. Lozach, manuscript, Caen, 1981.

[8] A. Lasjaunias, *Diophantine approximation and continued fraction expansions of algebraic power series in positive characteristic*, J. Number Theory 65 (1997), 206–225.

[9] —, *Continued fractions for algebraic power series over a finite field*, Finite Fields Appl. 5 (1999), 46–56.

[10] A. Lasjaunias and B. de Mathan, *Thue's Theorem in positive characteristic*, J. Reine Angew. Math. 473 (1996), 195–206.

[11] —, —, *Differential equations and diophantine approximation in positive characteristic*, Monatsh. Math. 473 (1996), 195–206.

[12] B. de Mathan, *Approximations diophantiennes dans un corps local*, Bull. Soc. Math. France, Suppl. Mém. 21 (1970).

[13] —, *Approximation exponents of algebraic functions in positive characteristic*, Acta Arith. 60 (1992), 359–370.

[14] B. Mazur, *Rational points on modular curves*, in: Modular Functions of One Variable, V (Bonn, 1976), Lecture Notes in Math. 601, Springer, 1977, 107–148.

[15] L. Merel, *Bornes pour la torsion des courbes elliptiques sur les corps de nombres*, Invent. Math. 124 (1996), 437–449.

[16] W. H. Mills and D. P. Robbins, *Continued fractions for certain algebraic power series*, J. Number Theory 23 (1986), 388–404.

[17] T. Nagell, *Problems in the theory of exceptional points on plane cubics of genus one*, Den 11te Skandinaviske Matematikerkongress, Trondheim, 1949, 71–76.

[18] C. F. Osgood, *Effective bounds on the "diophantine approximation" of algebraic functions over fields of arbitrary characteristic and applications to differential equations*, Indag. Math. 37 (1975), 105–119.

[19] O. Perron, *Die Lehre von den Kettenbrüchen*, Vol. I, 3rd ed., Teubner, Stuttgart, 1954.

[20] A. Schinzel, *On some problems of the arithmetical theory of continued fractions*, Acta Arith. 6 (1961), 393–413.

[21] —, *On some problems of the arithmetical theory of continued fractions II*, ibid. 7 (1962), 287–298. Corrigendum, ibid. 47 (1987), 295.

[22] J. A. Serret, *Handbuch der höheren Algebra*, Vol. I (German translation of French original), 1878.

[23] P. Tchebycheff, *Sur l'intégration des différentielles qui contiennent une racine carrée d'un polynome du troisième ou du quatrième degré*, J. Math. Pures Appl. (2) 9 (1864), 225–246.

[24] J. F. Voloch, *Diophantine approximation in positive characteristic*, Period. Math. Hungar. 19 (1988), 217–225.

[25] —, *Diophantine approximation in characteristic p*, Monatsh. Math. 119 (1995), 321–325.

Department of Mathematics
University of Colorado
Boulder, CO 80309-0395, U.S.A.
E-mail: Schmidt@euclid.colorado.edu

Received on 9.11.1998 (3505)