Continued Fractions and Linear Recurrences

By W. H. Mills

Abstract. Let \( t_0, t_1, t_2, \cdots \) be a sequence of elements of a field \( F \). We give a continued fraction algorithm for \( t_0 + t_1 x + t_2 x^2 + t_3 x^3 + \cdots \). If our sequence satisfies a linear recurrence, then the continued fraction algorithm is finite and produces this recurrence.

More generally, the algorithm produces a nontrivial solution of the system

\[
\sum_{j=0}^{s} t_{i+j} y_j \quad 0 \leq i \leq s - 1,
\]

for every positive integer \( s \).

1. Let \( t_0, t_1, t_2, \cdots \) be a sequence of elements of a field \( F \). Set

\[
T = \sum_{j=0}^{\infty} t_j x^j.
\]

Let \( d \) be a nonnegative integer. We say that \( T^* \) is an approximation of \( T \) of degree \( d \) if there exist polynomials \( V \) and \( W \) such that \( T^* = V/W \), \( \deg V < d \), \( \deg W < d \), \( x \nmid W \), and \( x^d | WT - V \).

We shall give a continued fraction expansion for \( xT \). This yields polynomials \( V_i, W_i \), and integers \( d_i \), \( 0 = d_1 < d_2 < d_3 < \cdots \), such that \( (V_i, W_i) = 1 \) and \( V_i/W_i \) is an approximation of \( T \) of degree \( d_i \). Suppose \( T^* \) is any approximation of \( T \) of some degree \( d \). Then \( T^* = V_i/W_i \) for that value of \( i \) such that \( d_i \leq d < d_{i+1} \).

If the sequence of the \( t_i \) satisfies a linear recurrence of degree \( d \), but not one of smaller degree, then there is an \( m \) such that \( d_m = d \) and the linear recurrence is given by the polynomial \( W_m \). In this case, \( W_m T = V_m \), the continued fraction expansion terminates at \( i = m \), and we can determine \( W_m \) from the first \( 2d \) of the \( t_i \), i.e., from those \( t_j \) such that \( 0 \leq j < 2d \).

Our algorithm is closely related to Zierler’s version of Berlekamp’s algorithm [1].

2. We consider continued fraction expansions of the form

\[
\alpha = N_1 + \cfrac{1}{N_2 + \cfrac{1}{N_3 + \cdots}},
\]

where \( N_1, N_2, N_3, \cdots \) are elements from some field \( E \). We can write

Received January 28, 1974.

\[ \alpha = N_1 + R_1, \quad 1/R_1 = N_2 + R_2, \quad 1/R_2 = N_3 + R_3, \ldots. \]

If \( R_m = 0 \) for some \( m \), then the continued fraction terminates with \( N_m \). Otherwise it is an infinite continued fraction.

In the classical case, \( \alpha \) is a real number, the \( N_i \) are integers, and \( 0 \leq R_i < 1 \) for all \( i \). We are interested in a different case.

We set

\[ P_0 = 1, \quad Q_0 = 0; \quad P_1 = N_1, \quad Q_1 = 1, \]

(1)

\[ P_i = N_i P_{i-1} + P_{i-2}, \quad i \geq 2, \]

and

\[ Q_i = N_i Q_{i-1} + Q_{i-2}, \quad i \geq 2. \]

(2)

It is well known, and easy to show, that

\[ P_1/Q_1 = N_1, \quad P_2/Q_2 = N_1 + 1/N_2, \]

\[ P_3/Q_3 = N_1 + 1/(N_2 + 1/N_3), \ldots. \]

The sequence \( P_1/Q_1, P_2/Q_2, P_3/Q_3, \ldots \) converges to \( \alpha \) in many cases, including the classical case.

We put

\[ A_i = a Q_i - P_i, \quad i \geq 0. \]

Then we have

\[ \Delta_0 = -1, \quad \Delta_1 = \alpha - N_1 \]

and

\[ \Delta_i = N_i \Delta_{i-1} + \Delta_{i-2}, \quad i \geq 2. \]

(3)

Clearly \( R_1 = \alpha - N_1 = -\Delta_1/\Delta_0 \). Since \( R_{i+1} = -N_{i+1} + 1/R_i \) it follows from (3), by induction on \( i \), that

\[ R_i = -\Delta_i/\Delta_{i-1}, \quad i \geq 1. \]

(4)

3. We now take \( E \) to be the field of all series of the form \( \sum_{j=k}^{\infty} a_j x^j \), where the \( a_j \) are elements of the field \( F \) and \( k \) is a rational integer which may be negative. For convenience let \( y = 1/x \). We set \( \alpha = xT \) and \( N_1 = 0 \). Then \( R_1 = \alpha = xT \). We now define the \( N_i \) and \( R_i \) inductively using

\[ 1/R_{i-1} = N_i + R_i, \quad i \geq 2, \]

where we take \( N_i \) to be a polynomial in \( y \) and \( x|R_i \). Thus if

\[ 1/R_{i-1} = \sum_{j=k}^{\infty} a_j x^j, \quad a_k \neq 0, \]
it turns out that $k < 0$ and we have
\[ N_i = \sum_{j=k}^{0} a_j x^j = \sum_{u=0}^{-k} a_u x^u \quad \text{and} \quad R_i = \sum_{j=1}^{\infty} a_j x^j. \]

This determines the $N_i$ and $R_i$ uniquely. If $R_m = 0$ for some $m$, then the process terminates at this point. The $P_i$, $Q_i$, and $A_i$ are now determined by (1), (2), (3), (4), and (5).

We shall write $x^r \| A$ if $x^r$ divides $A$, but $x^{r+1}$ does not divide $A$. This means that $A$ is of the form $A = \sum_{j=r}^{\infty} a_j x^j$ with $a_r \neq 0$. Let $x^r \| R_i$, $i \geq 1$. If $R_m = 0$, we set $r_m = \infty$. Then $r_i \geq 1$ for $i \geq 1$. For $i \geq 2$, $N_i$ is a polynomial in $y$ of degree $r_{i-1}$. Set

\[ (8) \quad d_i = \sum_{j=1}^{i-1} r_j. \]

Then we have $0 = d_1 < d_2 < d_3 < \cdots$. It follows from (1) and (3), by induction on $i$, that $Q_i$ is a polynomial in $y$ of degree $d_i$. Similarly, for $i \geq 2$, $P_i$ is a polynomial in $y$ of degree $d_i - r_1$. Set

\[ V_i = x^{d_i-1} P_i, \quad W_i = x^{d_i} Q_i. \]

Then $V_i$ and $W_i$ are polynomials in $x$, deg $V_i < d_i$, and deg $W_i \leq d_i$. Moreover, $W_i$ has a nonzero constant term so that $x^r \| W_i$. Now

\[ TW_i - V_i = x^{d_i-1} (\alpha Q_i - P_i) = x^{d_i-1} \Delta_i. \]

Since $\Delta_0 = -1$, (6) gives us

\[ \Delta_i = (-1)^{i+1} \prod_{j=1}^{i} R_j. \]

Since $x^r \| R_j$, we have

\[ (9) \quad x^{d_i+1} \| \Delta_i \]

by (8). Hence

\[ (10) \quad x^{d_i+d_i+1-1} \| TW_i - V_i. \]

Therefore, $x^{2d_i+1} | TW_i - V_i$ so that $V_i/W_i$ is an approximation of $T$ of degree $d_i$.

**Lemma 1.** Let $T^*$ be an approximation of $T$ of degree $d$. Let $i$ be the integer such that $d_i \leq d < d_{i+1}$. Then $T^* = V_i/W_i$.

**Proof.** We have $T^* = V_i/W_i$ where deg $W \leq d$, deg $V < d$, and $x^{2d} | WT - V$. Now $d + d_i \leq 2d$ so that $x^{d+d_i} | WT - V$. Moreover, $d + d_i \leq d_i + d_i+1 - 1$ so that $x^{d+d_i} | W_iT - V_i$ by (10). Since

\[ V_i W - V W_i = W_i (WT - V) - W_i(W_iT - V_i), \]
we have
\[ x^{d+d_{i}} | V_{i}W - VW_{i}. \]
Now the degree of \( V_{i}W - VW_{i} \) is less than \( d + d_{i} \). Therefore \( V_{i}W - VW_{i} = 0 \), so that
\[ T^* = V/W = V_{i}/W_{i}. \]

**Lemma 2.** If \( V_{i}/W_{i} = V_{j}/W_{j} \), then \( i = j \).

**Proof.** Suppose \( V_{i}/W_{i} = V_{j}/W_{j} \). Then we have \( V_{i} = VD, W_{i} = WD, V_{j} = VE, W_{j} = WE \) for suitable polynomials \( V, W, D, E \) with \( (V, W) = 1 \). Since \( x^{d_{i}} | W_{i} \), we have \( x^{d_{i}} | D \) so that (10) yields
\[ x^{d_{i} + d_{i+1} - 1} | TW - V. \]
Similarly
\[ x^{d_{j} + d_{j+1} - 1} | TW - V. \]
Hence
\[ d_{i} + d_{i+1} - 1 = d_{j} + d_{j+1} - 1. \]

Therefore, \( i = j \).

**Lemma 3.** \( (V_{i}, W_{i}) = 1 \).

**Proof.** Suppose \( (V_{i}, W_{i}) = D \) where \( \deg D > 0 \). Then \( V_{i} = VD, W_{i} = WD \) for suitable polynomials \( V, W \) such that \( x^{d_{i}} | W, \deg W < d_{i} \), and \( \deg V < d_{i} - 1 \). Moreover \( x^{d_{i}} | D \) so that \( x^{2d_{i}} | TW - V. \) Hence \( V/W \) is an approximation of \( T \) of degree less than \( d_{i} \). By Lemma 1 we have \( V/W = V_{j}/W_{j} \) for some \( j < i \). This contradicts Lemma 2.

**Lemma 4.** For any particular value of \( i \) we have either \( \deg V_{i} = d_{i} - 1 \) or \( \deg W_{i} = d_{i} \).

**Proof.** Since \( \deg W_{1} = 0 = d_{1} \), we may suppose \( i > 1 \). If the result is false, then \( V_{i}/W_{i} \) is an approximation of \( T \) of degree less than \( d_{i} \). By Lemma 1 this implies that \( V_{i}/W_{i} = V_{j}/W_{j} \) for some \( j < i \), which contradicts Lemma 2.

4. Let \( \{t_{j}\} = \{t_{0}, t_{1}, \cdots, t_{n-1}\} \) be a finite sequence of elements of \( F \), and set
\[ T = \sum_{j=0}^{n-1} t_{j} x^{j}. \]
Let \( W \) be a polynomial of degree \( s \) with a nonzero constant term. Thus \( W = \sum_{j=0}^{s} w_{j} x^{j} \), where the \( w_{j} \) are elements of \( F \), \( w_{0} \neq 0 \), \( w_{s} \neq 0 \). The linear recurrence given by \( W \) is
\[ \sum_{i=0}^{s} w_{i} t_{k-i} = 0. \]
If (11) holds for a particular value \( k_{0} \) of \( k \), we say that the linear recurrence \( W \) holds
for \( k_0 \). If (11) holds for all values of \( k \) for which the left side is defined, i.e., for \( s \leq k \leq n - 1 \), then we say that the sequence \( \{t_j\} \) satisfies the linear recurrence \( W \).

Whenever we speak of a linear recurrence \( W \) we shall mean a polynomial \( W \) with a nonzero constant term. The degree of the linear recurrence is defined to be the degree of this polynomial.

In order to determine \( W \), up to a multiplicative constant, we must have (11) satisfied by at least \( s \) values of \( k \). Hence we must have \( 2s \leq n \). Our problem is to determine whether or not the sequence \( \{t_j\} \) satisfies a linear recurrence of degree \( s \leq n/2 \), and if so to determine the linear recurrence of lowest degree that \( \{t_j\} \) satisfies.

Let \( h = [n/2] \). Thus \( h \) is an integer and either \( n = 2h \) or \( n = 2h + 1 \). Let \( xT \) be expanded in a continued fraction as indicated in Section 2 and Section 3. This gives us polynomials \( V_j \) and \( W_j \) and integers \( d_j \). Let \( m \) be the integer such that \( dm < n < dm + x \). This is equivalent to

\[
2d_m \leq n < 2d_{m+1}.
\]

Now suppose that the sequence \( \{t_j\} \) satisfies a linear recurrence \( W \) of degree \( s \), where \( s \leq n/2 \). Thus \( s \leq h \). We suppose \( W \) chosen so that \( s \) is minimal. Set

\[
V = \sum_{j=0}^{s-1} v_j x^j,
\]

where

\[
v_j = \sum_{i=0}^{j} w_{j-i}.
\]

Then \( x^n |TW - V \) by (11) so that \( V/W \) is an approximation of \( T \) of degree \( h \). More precisely it is an approximation of \( T \) of degree \( d \) for any \( d \) such that \( s \leq d \leq h \). By Lemma 1 and the choice (12) of \( m \) we have \( V/W = V_m/W_m \). Since \( W \) is of minimal degree, we have \( (V, W) = 1 \). Moreover \( (V_m, W_m) = 1 \) by Lemma 3, so that \( W = \lambda W_m \) for some nonzero element \( \lambda \) of \( F \).

More generally, suppose only that the linear recurrence \( W \) holds for those \( k \) such that \( h \leq k \leq n - 1 \), that \( \deg W \leq h \), and that \( W \) is a linear recurrence of minimal degree with these properties. As above there is a polynomial \( V \) such that \( V/W \) is an approximation of \( T \) of degree \( h \), \( (V, W) = 1 \), and \( W = \lambda W_m \) for some nonzero \( \lambda \) in \( F \).

It is easy to see that there need not be such a linear recurrence. For example, we may take \( \{t_j\} = \{0, 0, \cdots, 0, 1 \} \). However, we have shown that if there is one, then it must be \( W_m \), up to a multiplicative constant.

Now

\[
x^{dm+dm+1-1} ||TW_m - V_m
\]

by (10). Hence if \( n \geq dm + dm + 1 \), then \( \{t_j\} \) does not satisfy the linear recurrence \( W_m \), in fact \( W_m \) fails to hold for \( dm + dm + 1 - 1 \). Thus we have the following result:

**Theorem 1.** If \( dm + dm + 1 \leq n < 2dm + 1 \), then the sequence \( \{t_j\} \) does
not satisfy any linear recurrence of degree $\leq n/2$. In fact, there is no linear recurrence of degree $\leq n/2$ that holds for all $k$ such that $h \leq k \leq n - 1$.

Now suppose that $n < dm + dm + x$. Then the linear recurrence $W_m$ holds for all $k$ in the range $d_m \leq k \leq n - 1$. We have $\deg W_m \leq d_m$. If $\deg W_m = d_m$, then $\{t_j\}$ satisfies the linear recurrence $W_m$. However, if $\deg W_m < d_m$, then $\deg V_m = d_m - 1$ by Lemma 4, and, therefore, the linear recurrence $W_m$ fails to hold at $d_m - 1$. Thus we have the following result:

**Theorem 2.** Suppose $2d_m \leq n < dm + dm + x$. If $\deg W_m = d_m$, then $W_m$ is a linear recurrence of minimal degree satisfied by $\{t_j\}$. If $\deg W_m < d_m$, then there is no linear recurrence of degree $\leq n/2$ which is satisfied by $\{t_j\}$. However, $W_m$ is a linear recurrence of minimal degree that holds for all $k$ such that $h \leq k \leq n - 1$. It holds for all $k$ in the range $d_m \leq k \leq n - 1$, and fails to hold for $d_m - 1$.

5. In this section, we shall describe an efficient method of computing the polynomial $W_m$. As before, let $\{t_j\} = \{t_0, t_1, \cdots, t_{n-1}\}$ be the finite sequence we are interested in. We start with $N_0 = 0$, $A_0 = -1$, and $A_1 = (T, t_1, 1) = t_0 x^{t_1 + 1}$.

For $i \geq 2$, (6) and (7) give us

$$N_i + R_i = xT - N_1 = \sum_{j=0}^{n-1} t_j x^{j+1}.$$  

where $xR_i$ and $N_i$ is a polynomial in $y$, $y = 1/x$. Thus $N_i$ can be obtained from $\Delta_{i-2}$ and $\Delta_{i-1}$ by an ordinary division process. Then $\Delta_i$ is given by (5):

$$\Delta_i = N_i \Delta_{i-1} + \Delta_{i-2}.$$  

In this way, the $N_i$ and the $\Delta_i$ can be successively obtained. We must continue this out to $i = m$ where $2d_m \leq n < 2d_m + 1$. Since $x^{d_i} || \Delta_{i-1}$ by (9), we know at once when we have reached $i = m$. If $d_m + dm + x < n$, then there is no solution. If $d_m + dm + x > n$, then we calculate $Q_m$ from the $N_i$ and the relations $Q_0 = 0, Q_1 = 1, Q_i = N_i Q_{i-1} + Q_{i-2}$.

If $Q_m$ has a nonzero constant term, then $\deg W_m = d_m$ and $W_m = x^{d_m} Q_m$ is the required linear recurrence. If $Q_m$ has no constant term, then $\deg W_m < d_m$ and $\{t_j\}$ does not satisfy a linear recurrence of degree $\leq n/2$. However, in this case, $W_m = x^{d_m} Q_m$ is a linear recurrence that holds for all $k$ such that $d_m \leq k \leq n - 1$.

We note that $x^{d_i} || \Delta_{i-1}$, $x^{d_{i-1}} || \Delta_{i-2}$, and $d_i = r_{i-1} + d_{i-1}$. Hence in performing the division $\Delta_{i-2}/\Delta_{i-1}$ we need only use the first $r_{i-1} + 1$ terms of $\Delta_{i-2}$ and the same number of terms of $\Delta_{i-1}$. This is sufficient to determine $N_i$ completely.

Finally we note that it is only necessary to calculate $\Delta_i$ out to the term in $x^{n-d_i}$. This corresponds to the fact that $\Delta = xT$ is known only out to the term in $x^n$. To see this, consider the division of $\Delta_{i-2}$ by $\Delta_{i-1}$. We need $r_{i-1} + 1$ terms of each. More terms of $\Delta_{i-2}$ are assumed known than of $\Delta_{i-1}$. The number of terms of $\Delta_{i-1}$ that we have is $n - d_{i-1} - d_i + 1 = n - 2d_i + r_{i-1} + 1$. Since we
may suppose \( i \leq m \), this is at least \( r_{i-1} + 1 \) terms. Thus \( N_i \) may be computed exactly. Clearly if we know \( \Delta_{i-2} \) out to the term in \( x^{n-d_{i-2}} \) and \( \Delta_{i-1} \) out to the term in \( x^{n-d_{i-1}} \), then once \( N_i \) is known as a polynomial in \( y \) of degree \( r_{i-1} \), we may calculate \( \Delta_i \) out to the term in \( x^{n-d_i} \).

Tables 1 and 2 give examples of the calculation for small \( n \) and \( F = GF(2) \). The unnecessary terms of \( \Delta_i \), i.e., those beyond \( x^{n-d_i} \), are given in parenthesis. In the first example \( n = 12, m = 3, d_3 = 3, d_4 = 7, d_m + d_{m+1} \leq n \), so there is no solution and the \( Q_i \) are not calculated. In the second example, the sequence satisfies the linear recurrence \( x^4 + x + 1 \).

### Table 1

\[
F = GF(2), \ n = 12, \ \{t_i\} = \{100101110111\}
\]

<table>
<thead>
<tr>
<th>(i)</th>
<th>(N_i)</th>
<th>(\Delta_i)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>(x + x^4 + x^6 + x^7 + x^8 + x^{10} + x^{11} + x^{12})</td>
</tr>
<tr>
<td>2</td>
<td>(y)</td>
<td>(x^3 + x^5 + x^6 + x^7 + x^9 + x^{10} + x^{11})</td>
</tr>
<tr>
<td>3</td>
<td>(y^2 + 1)</td>
<td>(x^7 + x^{12})</td>
</tr>
</tbody>
</table>

There is no linear recurrence of degree \( \leq 6 \).

### Table 2

\[
F = GF(2), \ n = 8, \ \{t_i\} = \{11101011\}
\]

<table>
<thead>
<tr>
<th>(i)</th>
<th>(N_i)</th>
<th>(\Delta_i)</th>
<th>(Q_i)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>-1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>(x + x^2 + x^3 + x^5 + x^7 + x^8)</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>(y + 1)</td>
<td>(x^3 + x^4 + x^5 + x^6(x + x^8))</td>
<td>(y + 1)</td>
</tr>
<tr>
<td>3</td>
<td>(y^2)</td>
<td>(x^4 + x^5(x + x^6 + x^7 + x^8))</td>
<td>(y^3 + y^2 + 1)</td>
</tr>
<tr>
<td>4</td>
<td>(y)</td>
<td>((x^7 + x^8))</td>
<td>(y^4 + y^3 + 1)</td>
</tr>
</tbody>
</table>

The linear recurrence is \( x^4(y^3 + y^3 + 1) = x^4 + x + 1 \).

6. We now consider the system

\[
(13) \quad \sum_{j=0}^{s} t_{i+j} \lambda_j, \quad 0 \leq i \leq s - 1,
\]

of \( s \) linear equations in \( s + 1 \) unknowns. This system must have at least one non-trivial solution in \( F \). If we set

\[
\Lambda = \sum_{j=0}^{s} \lambda_j x^{s-j},
\]

then we can write \( \Lambda = x^r W \), where \( W \) is a polynomial with nonzero constant term,
and \( \deg W \leq s - r \). If (13) holds, then there is a polynomial \( V \) such that \( \deg V < s - r \) and \( X^{2s-r}TW - V \). Thus \( V/W \) is an approximation of \( T \) of degree \( s - r \). Hence \( V/W = V_m/W_m \) for some \( m \) with \( d_m \leq s - r \) and \( d_m + d_{m+1} - 1 \geq 2s - r \), so that \( d_m \leq s < d_{m+1} \). Thus we see that our algorithm can be used to solve the system (13) for any positive integer \( s \).

Institute for Defense Analyses
Communications Research Division
Princeton, New Jersey 08540