THEOREM OF SKOLEM-MAHLER-LECH

CAMERON FRANC

1. LINEAR RECURRENCE SEQUENCES

A linear recurrence sequence (LRS for short) is a sequence of numbers \((a_n)_{n=0}^{\infty}\) satisfying

\[ a_n = c_1a_{n-1} + c_2a_{n-2} + \cdots + c_da_{n-d} \quad \text{for all } n \geq d, \]

where \(c_1\) through \(c_d\) are fixed numbers. Such a sequence is uniquely determined by the values \(a_0\) through \(a_{d-1}\), as well as the constants \(c_1\) through \(c_d\). We will mainly be interested in the case where all the \(a_i\)s and \(c_i\)s are integers. Such a sequence is called an integer LRS.

Example 1. The Fibonacci sequence is an example of an integer LRS.

Example 2. The sequence \(a_n\) defined by \(a_n = n\) is an integer LRS. It satisfies the recurrence relation \(a_n = 2a_{n-1} - a_{n-2}\).

Polynomials transform sequences into other sequences in the following way: let \(f(x) = \sum_{i=0}^{d} \alpha_i x^i\) be a polynomial, and let \(a_n\) be a sequence. We define a new sequence \(f \ast a\) whose \(n\)th term is defined as follows:

\[ (f \ast a)_n := \sum_{i=0}^{d} \alpha_i a_{n+i}. \]

Questions.

(1) What do constant polynomials do to sequences?
(2) What does \(f(x) = x\) do to a sequence?
(3) Let \(f(x) = x^2 - x - 1\), and let \(F\) denote the Fibonacci sequence. What sequence is \(f \ast F\)?

If \(a\) is a sequence then we define a set of polynomials \(I(a)\) as follows:

\[ I(a) := \{ f(x) \in \mathbb{C}[x] \mid f \ast a = (0, 0, 0, \ldots) \}. \]

This set satisfies the following properties:

(a) \(I(a)\) contains the zero polynomial.
(b) If \(f, g \in I(a)\) then also \(f + g \in I(a)\).
(c) If \(f \in I(a)\) and \(g \in \mathbb{C}[x]\) then \(fg \in I(a)\).

A subset of \(\mathbb{C}[x]\) satisfying these properties is called an ideal. One can use the division algorithm for polynomials to show that every ideal containing a nonzero element in fact contains a unique monic polynomial of smallest positive degree. If \(a\) is a sequence such that \(I(a)\) contains a nonzero polynomial, then we call the unique monic polynomial of smallest positive degree the minimal polynomial of \(a\).

Questions.

(1) Prove that a sequence \(a\) is an LRS if and only if \(I(a) \neq \{0\}\).
(2) Give an example of a sequence \(a\) such that \(I(a) = \{0\}\).
If \( a \) is a sequence satisfying the LRS \( a_n = c_1a_{n-1} + c_2a_{n-2} + \cdots + c_d a_{n-d} \) then we call 
\[ f(x) = x^d - c_1x^{d-1} - c_2x^{d-2} - \cdots - c_d x_{n-d} \]
The characteristic polynomial of the recurrence relation. Note that 
\( f(x) \in I(a) \) and the minimal polynomial of \( I(a) \) divides \( f(x) \). The following result can be proved using the theory of the Jordan canonical form from linear algebra.

**Theorem 3.** Let \( f(x) = x^d - c_1x^{d-1} - c_2x^{d-2} - \cdots - c_d \) denote a monic polynomial with \( c_0 \neq 0 \), and suppose that \( f(x) \) factors over \( C \) as 
\[ f(x) = \prod_{i=1}^k (x - r_i)^{m_i} \]
where the \( r_i \) are distinct complex numbers and the \( m_i \) are positive integers satisfying \( \sum_{i=1}^k m_i = d \). Then a sequence \( a \) satisfies a linear recurrence with characteristic polynomial \( f(x) \) if and only if there exist polynomials \( g_1(n), \ldots, g_k(n) \) with \( \deg g_i \leq m_i - 1 \) for all \( i \), such that 
\[ a_n = g_1(n) r_1^n + \cdots + g_k(n) r_k^n \]
for all \( n \).

**Remark 4.** If \( f(x) \) has no repeated roots in Theorem 3 then \( k = d = \deg(f) \) and a sequence satisfies a linear recurrence with characteristic polynomial \( f(x) \) if and only if there exist constants \( g_1, \ldots, g_d \) such that \( a_n = \sum_{i=1}^d g_i r_i^n \) for all \( n \).

2. **The Skolem-Mahler-Lech Theorem**

If \( a \) is a sequence, then we call \( Z(a) := \{ n \in \mathbb{N} | a_n = 0 \} \) the zero set of \( a \).

**Theorem 5** (**Skolem (1934)**). Let \( a \) be an integer LRS. Then \( Z(a) \) is the union of a finite (possibly empty) set and a finite (possibly empty) list of arithmetic progressions.

**Remark 6.** This result was extended by Mahler (1935) to number fields and by Lech (1952) to arbitrary fields of characteristic zero. Other authors have generalized this in a variety of ways. For example, Harm Derksen has considered the case of positive characteristic fields, and Bell-Lagarias have proved an analogue for iterated automorphisms of \( K \)-algebras, where \( K \) is a field of characteristic zero.

**Question.**

Give examples of LRSs whose zero sets are:
(a) the even natural numbers;
(b) the natural numbers congruent to 2 mod 3;
(c) the finite set \( \{1, 3, 5, 7\} \).

One can find a nice discussion of Theorem 5 on Terry Tao’s blog at the link [https://terrytao.wordpress.com/tag/skolem-mahler-lech-theorem/](https://terrytao.wordpress.com/tag/skolem-mahler-lech-theorem/)

He notes the following open problem: if \( a \) is an integer LRS, is the truth of the statement \( Z(a) = \emptyset \) decidable in finite time?

3. **Proof of the Theorem**

Let \( a \) be an integer LRS satisfying 
\[ a_n = c_1a_{n-1} + c_2a_{n-2} + \cdots + c_d a_{n-d} \]
for all \( n \geq d \). Without loss of generality we may assume \( c_d \neq 0 \). Define matrices:

\[
A = \begin{pmatrix}
0 & 1 & 0 & \cdots & 0 \\
0 & 0 & 1 & \cdots & 0 \\
\vdots \\
c_d & c_{d-1} & c_{d-2} & \cdots & c_1
\end{pmatrix}, \quad v = \begin{pmatrix}
a_0 \\
a_1 \\
\vdots \\
a_{d-1}
\end{pmatrix}, \quad e = \begin{pmatrix}
1 \\
0 \\
\vdots \\
0
\end{pmatrix},
\]

so that \( \langle A^n v, e \rangle = a_n \) for all \( n \geq 0 \).

Questions.

(1) Prove that \( A \) is invertible if and only if \( c_d \neq 0 \).
(2) Prove that there exists a finite integer \( m \geq 1 \) such that \( A^m \equiv 1 \pmod{p} \).
(3) Suppose that for such an integer \( m \) we can show that for each integer \( r = 0, \ldots, m-1 \), the set

\[
\{ n \in \mathbb{N} \mid a_{mn+r} = 0 \}
\]

is either finite or all of \( \mathbb{N} \). Prove that this implies Theorem 5.

It now suffices to prove (3). For this we fix integers \( r \) and \( m \) as in (3) and assume that \( S = \{ n \in \mathbb{N} \mid a_{mn+r} = 0 \} \) is infinite. We must prove that \( S = \mathbb{N} \). Let us define a function \( P: \mathbb{N} \to \mathbb{Z} \) by setting

\[
P(n) = a_{mn+r} = \langle (1+pB)^n A^r v, e \rangle,
\]

where \( B \) is a \( d \times d \)-integral matrix such that \( A^m = 1 + pB \). Note that \( P(n) \) is subsequence of \( a \) and \( S \) is its zero set.

Questions.

(1) Give an example of \( d \times d \)-matrices \( X \) and \( Y \) such that the binomial theorem does not hold for \( (X+Y)^k \) for some \( k \geq 0 \).
(2) Explain why one can expand \( (1+pB)^n \) using the binomial theorem.
(3) Prove that

\[
P(n) = \sum_{j=0}^{\infty} p^j P_j(n)
\]

for certain polynomials \( P_j(n) \in \mathbb{Z}[n] \).

To conclude the proof one can now use \( p \)-adic numbers. Briefly, these are infinite base-\( p \) expansions, where one does arithmetic in the usual way using “carries”. More precisely, the \( p \)-adic numbers \( \mathbb{Q}_p \) are obtained by completing the field \( \mathbb{Q} \) with respect to the metric obtained from the \( p \)-adic absolute value: write \( \alpha \in \mathbb{Q} \) in the form \( \alpha = p^r a/b \) where \( a, b \) are integers coprime to \( p \), and \( r \in \mathbb{Z} \). Then define

\[
|\alpha|_p := p^{-r}.
\]

The \( p \)-adic integers \( \mathbb{Z}_p \) are the closure of \( \mathbb{Z} \) inside of \( \mathbb{Q}_p \). Thanks to the powers of \( p^j \) in the expression for \( P(n) \) in (3), and thanks to funny properties of \( p \)-adic distances, one can prove that \( P(n) \) is a \( p \)-adically continuous function on \( \mathbb{Z} \), and hence it extends to a continuous function \( P: \mathbb{Z}_p \to \mathbb{Z}_p \). Moreover, (3) realizes \( P \) as a power series that converges on \( Z_p[1] \). But a nonzero convergent powers series can be shown to have a finite number of zeros in \( Z_p \) (which is a compact subset of \( \mathbb{Q}_p \)). Since we’re assuming that \( P \) has infinitely many zeros in \( \mathbb{N} \), it must in fact be the zero function on all of \( \mathbb{Z}_p \), and in particular \( S = \mathbb{N} \). This concludes the proof of Theorem 5.

\(^1\)To see this one needs to show that \( p^j \) is divisible by many more copies of \( p \) than \( j! \).