SOLVING LINEAR RECURSIONS OVER ALL FIELDS

KEITH CONRAD

1. Introduction

A sequence $\{a_n\} = (a_0, a_1, a_2, \dots)$ in a field K satisfies a *linear recursion* if there are $c_1, \dots, c_d \in K$ such that

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + \cdots + c_d a_{n-d} \tag{1.1}$$

for all $n \geq d$. For example, the Fibonacci sequence $\{F_n\} = (0, 1, 1, 2, 3, 5, 8, \dots)$ is defined by the linear recursion $F_n = F_{n-1} + F_{n-2}$ with initial values $F_0 = 0$ and $F_1 = 1$. (Often F_0 is ignored, but the values $F_1 = F_2 = 1$ and the recursion force $F_0 = 0$.) We will assume $c_d \neq 0$ and then say the recursion has *order* d; this is analogous to the degree of a polynomial. For instance, the recursion $a_n = a_{n-1} + a_{n-2}$ has order 2.

The sequences in K satisfying a common recursion (1.1) are a K-vector space under termwise addition. The initial terms $a_0, a_1, \dots, a_{d-1}$ determine the rest and if $c_d \neq 0$ then we can set the initial d terms arbitrarily[1], so solutions to (1.1) form a d-dimensional vector space in K. We seek an explicit basis for the solutions of (1.1) described by nice formulas.

Example 1.1. Solutions to $a_n = a_{n-1} + a_{n-2}$ in $\mathbf{R}$ are a 2-dimensional space. A power sequence λ^n with $\lambda \neq 0$ satisfies it when $\lambda^n = \lambda^{n-1} + \lambda^{n-2}$, which is equivalent to $\lambda^2 = \lambda + 1$. That makes λ a root of $x^2 - x - 1$, so $\lambda = \frac{1 \pm \sqrt{5}}{2}$. The sequences $(\frac{1+\sqrt{5}}{2})^n$ and $(\frac{1-\sqrt{5}}{2})^n$ are not scalar multiples, so they are a basis: in $\mathbf{R}$ if $a_n = a_{n-1} + a_{n-2}$ for all $n \geq 2$ then

$$a_n = \alpha \left(\frac{1+\sqrt{5}}{2} \right)^n + \beta \left(\frac{1-\sqrt{5}}{2} \right)^n$$

for unique α and β in $\mathbf{R}$. The Fibonacci sequence $\{F_n\}$ is the special case where $a_0 = 0$ and $a_1 = 1$. Solving $\alpha + \beta = 0$ and $\alpha(\frac{1+\sqrt{5}}{2}) + \beta(\frac{1-\sqrt{5}}{2}) = 1$ we get $\alpha = 1/\sqrt{5}$ and $\beta = -1/\sqrt{5}$ when $a_n = F_n$. The solution with $a_0 = 1$ and $a_1 = 0$ uses $\alpha = -\frac{1-\sqrt{5}}{2\sqrt{5}}$ and $\beta = \frac{1+\sqrt{5}}{2\sqrt{5}}$.

If we replace $\mathbf{R}$ by any field K in which $x^2 - x - 1$ has roots, even a field of characteristic p, the computations above still work if we replace $\frac{1 \pm \sqrt{5}}{2}$ by the roots in K unless K has characteristic 5: there $x^2 - x - 1 = x^2 + 4x + 4 = (x+2)^2$ so 3 is the only root and the only power sequence satisfying $a_n = a_{n-1} + a_{n-2}$ is $\{3^n\} = (1, 3, 4, 2, 1, 3, 4, 2 \dots)$.

Example 1.2. Let K have characteristic 5. What nice formula fits $a_n = a_{n-1} + a_{n-2}$ and is linearly independent of $\{3^n\} = (1, 3, 4, \dots)$? The sequence $a_n = n3^{n-1}$ works since

$$a_{n-1} + a_{n-2} = 3^{n-3}((n-1)3 + (n-2)) = 3^{n-3}(4n) = n3^n \frac{4}{27} = n3^{n-1} = a_n.$$

This starts out as $(0, 1, 1, 2, 3, 0, 3, 3, \dots)$, so it is F_n mod 5 and is not a multiple of $\{3^n\}$.

[1] What if $c_d = 0$? If $a_n = 2a_{n-1}$ then the first term determines the rest, so the solution space is 1-dimensional. Writing the recursion as $a_n = 2a_{n-1} + 0a_{n-2}$ doesn't make the solution space 2-dimensional unless we insist the recursion is for $n \geq 2$ rather than for $n \geq 1$. We will not address this option.

How is $\{n3^{n-1}\}$ found? If $\{\lambda^n\}$ and $\{\mu^n\}$ satisfy the same linear recursion, so does any linear combination. If $\lambda \neq \mu$ a linear combination is $(\lambda^n - \mu^n)/(\lambda - \mu)$, and as $\mu \to \lambda$ this "becomes" $\{n\lambda^{n-1}\}$, which up to scaling is $\{n\lambda^n\}$.[2] This suggests that since 3 is a double root of $x^2 - x - 1$ in characteristic 5, $\{n3^{n-1}\}$ should be a solution and we saw it really is. The formula $F_n \equiv n3^{n-1} \bmod 5$ goes back at least to Catalan [3, p. 86] in 1857.

Our goal is to prove the following theorem about a basis for solutions to a linear recursion over a general field K.

Theorem 1.3. *Let $a_n = c_1a_{n-1} + c_2a_{n-2} + \cdots + c_da_{n-d}$ be a linear recursion of order d with $c_i \in K$. Assume $1 - c_1x - c_2x^2 - \cdots - c_dx^d$ factors in $K[x]$ over its reciprocal roots – the λ such that $1/\lambda$ is a root – as*

$$(1 - \lambda_1 x)^{e_1} \cdots (1 - \lambda_r x)^{e_r},$$

where the λ_i are distinct and $e_i \geq 1$. A basis for the solutions of the linear recursion in K is given by the e_i sequences $\{\lambda_i^n\}$, $\{n\lambda_i^n\}$, $\{\binom{n}{2}\lambda_i^n\}, \ldots, \{\binom{n}{e_i-1}\lambda_i^n\}$ for $i = 1, \ldots, r$.

Example 1.4. The simplest case of Theorem 1.3 is when each reciprocal root has multiplicity 1: if $1 - c_1x - c_2x^2 - \cdots - c_dx^d = (1 - \lambda_1 x) \cdots (1 - \lambda_d x)$ has d distinct reciprocal roots λ_i then the solutions of (1.1) are unique linear combinations of the λ_i^n: $a_n = \alpha_1\lambda_1^n + \cdots + \alpha_d\lambda_d^n$.

Example 1.5. The recursion $a_n = 8a_{n-1} - 24a_{n-2} + 32a_{n-3} - 16a_{n-4}$ has order 4 if K does not have characteristic 2. Since $1 - 8x + 24x^2 - 32x^3 + 16x^4 = (1 - 2x)^4$, the solutions have basis $\{2^n\}$, $\{n2^n\}$, $\{\binom{n}{2}2^n\}$, and $\{\binom{n}{3}2^n\}$, so every solution is $(b_0 + b_1n + b_2\binom{n}{2} + b_3\binom{n}{3})2^n$ for unique $b_i \in K$.

The classical version of Theorem 1.3 in $\mathbf{C}$, or more generally characteristic 0, is due to Lagrange and says a basis of the solutions is the sequences $\{n^k\lambda_i^n\}$ for $k < e_i$ and $i = 1, \ldots, r$. This works in characteristic p if all $e_i \leq p$ but breaks down if any $e_i > p$, so it's essential to use $\{\binom{n}{k}\lambda_i^n\}$ to have a result valid in all fields. While $n^k \bmod p$ has period p in n, $\binom{n}{k} \bmod p$ has a longer period if $k \geq p$. For instance, in characteristic 2 the sequence $\binom{n}{2} \bmod 2$ has period 4 with repeating values 0, 0, 1, 1 when n runs through the nonnegative integers.

We will prove Theorem 1.3 in two ways: by generating functions and by an analogy with differential equations. Anna Medvedovsky brought this problem in characteristic p to my attention and the second proof is a variation on hers. After writing this up I found a result equivalent to Theorem 1.3 in a paper of Fillmore and Marx [4, Thm. 1, 2] and the case of finite K in McEliece's Ph.D. thesis [5, p. 19]. The earliest paper I found mentioning the basis $\{\binom{n}{k}\lambda_i^n\}$ in characteristic p is by Engstrom [2, p. 215] in 1931 when $\max e_i = p$, but $\{n^k\lambda_i^n\}$ still works in that case. In 1933, Milne-Thomson [7, p. 388] in characteristic 0 gave the basis $\{n^k\lambda_i^n\}$ and remarked that the alternative $\{\binom{n-1}{k}\lambda_i^n\}$ "is sometimes convenient."

2. First Proof: Generating Functions

It is easy to show the sequence $\{\lambda^n\}$ fits (1.1) if λ is a reciprocal root of $1 - c_1x - \cdots - c_dx^d$:

$$\begin{aligned} \lambda^n = c_1\lambda^{n-1} + c_2\lambda^{n-2} + \cdots + c_d\lambda^{n-d} \text{ for all } n \geq 0 \quad &\iff \quad \lambda^d = c_1\lambda^{d-1} + c_2\lambda^{d-2} + \cdots + c_d \\ &\iff \quad \lambda^d - c_1\lambda^{d-1} - \cdots - c_d = 0, \\ &\iff \quad 1 - \frac{c_1}{\lambda} - \cdots - \frac{c_d}{\lambda^d} = 0. \end{aligned}$$

[2] There is an analogous result in differential equations: the solution space to $y''(t) + ay'(t) + by(t) = 0$ has basis $\{e^{\lambda t}, e^{\mu t}\}$ if λ and μ are different roots of $x^2 + ax + b$. If $x^2 + ax + b$ has a double root λ then a basis of the solution space is $\{e^{\lambda t}, te^{\lambda t}\}$. So $\lambda^n \leftrightarrow e^{\lambda t}$ and $n\lambda^n \leftrightarrow te^{\lambda t}$. We'll return to this analogy in Section 3.

It's more difficult to show $\binom{n}{k}\lambda^n$ for a $k \geq 1$ satisfies (1.1) if λ is a reciprocal root of $1-c_1x-\cdots-c_dx^d$ with multiplicity greater than k. To do this, we will rely on the following theorem characterizing linearly recursive sequences in terms of their generating functions.

Theorem 2.1. *If the linear recursion (1.1) has $c_d \neq 0$ then a sequence $\{a_n\}$ in K satisfies (1.1) if and only if the generating function $\sum_{n\geq 0} a_nx^n$ is a rational function of the form $N(x)/(1-c_1x-c_2x^2-\cdots-c_dx^d)$ where $N(x)=0$ or $\deg N(x) < d$.*

Proof. Set $F(x) = \sum_{n\geq 0} a_nx^n$. Using (1.1),

$$\begin{aligned}
F(x) &= a_0 + a_1x + \cdots + a_{d-1}x^{d-1} + \sum_{n\geq d} a_nx^n \\
&= a_0 + a_1x + \cdots + a_{d-1}x^{d-1} + \sum_{n\geq d}(c_1a_{n-1} + c_2a_{n-2} + \cdots + c_da_{n-d})x^n \\
&= a_0 + a_1x + \cdots + a_{d-1}x^{d-1} + \sum_{i=1}^{d}\sum_{n\geq d} c_ia_{n-i}x^n \\
&= a_0 + a_1x + \cdots + a_{d-1}x^{d-1} + \sum_{i=1}^{d} c_ix^i\left(\sum_{n\geq d} a_{n-i}x^{n-i}\right) \\
&= a_0 + a_1x + \cdots + a_{d-1}x^{d-1} + \sum_{i=1}^{d} c_ix^i\left(\sum_{n\geq d-i} a_nx^n\right) \\
&= a_0 + a_1x + \cdots + a_{d-1}x^{d-1} + \sum_{i=1}^{d} c_ix^i\left(F(x) - \sum_{n=0}^{d-i-1} a_nx^n\right).
\end{aligned}$$

The term in the sum at $i = d$ is just $c_dx^dF(x)$; the inner sum from $n = 0$ to $n = -1$ in this case is 0. Bringing $\sum_{i=1}^{d} c_ix^iF(x)$ over to the left side, we can solve for $F(x)$ as a rational function:

$$F(x) = \frac{N(x)}{1-c_1x-c_2x^2-\cdots-c_dx^d}$$

where $N(x)$, if not identically 0, is a polynomial of degree at most $d-1$.

Conversely, assume for $\{a_n\}$ in K that $\sum_{n\geq 0} a_nx^n = N(x)/(1-c_1x-\cdots-c_dx^d)$ where $N(x) = 0$ or $\deg N(x) < d$. Then

$$N(x) = \left(\sum_{n\geq 0} a_nx^n\right)(1-c_1x-c_2x^2-\cdots-c_dx^d).$$

Equating the coefficient of x^n on both sides for $n \geq d$,

$$0 = a_n - c_1a_{n-1} - c_2a_{n-2} - \cdots - c_da_{n-d},$$

which is the linear recursion (1.1). □

Corollary 2.2. *In the linear recursion (1.1) suppose $c_d \neq 0$. For $\lambda \in K^\times$, if $1-\lambda x$ is a factor of $1-c_1x-\cdots-c_dx^d$ with multiplicity $e \geq 1$ then for $0 \leq k \leq e-1$ the sequence $\{\binom{n}{k}\lambda^n\}$ satisfies (1.1).*

Proof. Theorem 2.1 tells us that our task is equivalent to showing the generating function $\sum_{n\geq 0}\binom{n}{k}\lambda^n x^n$ can be written in the form $N(x)/(1-c_1x-\cdots-c_dx^d)$ where $N(x)=0$ or $\deg N(x)<d$. We'll do this with an $N(x)$ of degree $d-1$.

In $\mathbf{Z}[[x]]$, differentiating the geometric series formula $\sum_{n\geq 0}x^n = 1/(1-x)$ a total of k times and then dividing both sides by $k!$ gives us the formal power series identity

$$\sum_{n\geq k}\binom{n}{k}x^{n-k} = \frac{1}{(1-x)^{k+1}}. \tag{2.1}$$

Since $\mathbf{Z}$ has a (unique) homomorphism to any commutative ring, (2.1) is true in $K[[x]]$.[3] Multiply both sides by x^k:

$$\sum_{n\geq 0}\binom{n}{k}x^{n} = \frac{x^k}{(1-x)^{k+1}}. \tag{2.2}$$

We changed the sum on the left to run over $n\geq 0$ instead of $n\geq k$, which is okay since $\binom{n}{k}=0$ for $0\leq n\leq k-1$. Replacing x with λx in (2.2),

$$\sum_{n\geq 0}\binom{n}{k}\lambda^n x^{n} = \frac{\lambda^k x^k}{(1-\lambda x)^{k+1}}. \tag{2.3}$$

Since $k\leq e-1$, $(1-\lambda x)^{k+1}$ is a factor of $(1-\lambda x)^e$, which is a factor of $1-c_1x-\cdots-c_dx^d$. Set $1-c_1x-\cdots-c_dx^d=(1-\lambda x)^e g(x)$. If we multiply the top and bottom of the right side of (2.3) by $(1-\lambda x)^{e-(k+1)}g(x)$, which has degree $d-(k+1)$ because $c_d\neq 0$, we get

$$\sum_{n\geq 0}\binom{n}{k}\lambda^n x^{n} = \frac{N(x)}{1-c_1x-c_2x^2-\cdots-c_dx^d}$$

where $\deg N(x) = k+(d-(k+1)) = d-1<d$. □

We proved in Corollary 2.2 that the sequences $\{\binom{n}{k}\lambda_i^n\}$ for $1\leq i\leq r$ and $0\leq k\leq e_i-1$ satisfy (1.1) when $c_d\neq 0$. The number of these sequences is[4] $\sum_{i=1}^r e_i = d$, which is the dimension of the solution space, so to finish the proof of Theorem 1.3 we will show these d sequences are linearly independent: if $b_{ik}\in K$ satisfy

$$\sum_{i=1}^{r}\sum_{k=0}^{e_i-1} b_{ik}\binom{n}{k}\lambda_i^n = 0 \text{ for all } n\geq 0 \tag{2.4}$$

then we want to show each b_{ik} is 0. The sequence $\{\sum_{i=1}^r\sum_{k=0}^{e_i-1} b_{ik}\binom{n}{k}\lambda_i^n\}$ for $n\geq 0$ has generating function

$$\sum_{n\geq 0}\left(\sum_i\sum_k b_{ik}\binom{n}{k}\lambda_i^n\right)x^n = \sum_i\sum_k b_{ik}\left(\sum_{n\geq 0}\binom{n}{k}\lambda_i^n x^n\right) \overset{(2.3)}{=} \sum_i\sum_k \frac{b_{ik}\lambda_i^k x^k}{(1-\lambda_i x)^{k+1}},$$

[3]We can't prove (2.1) in $K[[x]]$ directly for all fields K using repeated differentiation, since in fields of characteristic p the pth and higher-order derivatives are identically 0. It could be proved directly in $K[[x]]$ if K has characteristic p by using Hasse derivatives.

[4]Here we require that the linear recursion has order d, or equivalently that $c_d\neq 0$.

so this double sum is 0, since it's the generating function of the zero sequence. For each i, the inner sum over k is

$$\frac{b_{i0}}{1-\lambda_i x} + \frac{b_{i1}\lambda_i x}{(1-\lambda_i x)^2} + \frac{b_{i2}\lambda_i^2 x^2}{(1-\lambda_i x)^3} + \cdots + \frac{b_{i\,e_i-1}\lambda_i^{e_i-1}x^{e_i-1}}{(1-\lambda_i x)^{e_i}}. \tag{2.5}$$

Putting these terms over a common denominator, the sum is $q_i(x)/(1-\lambda_i x)^{e_i}$ for a polynomial $q_i(x)$ and the vanishing generating function for (2.4) becomes

$$\frac{q_1(x)}{(1-\lambda_1 x)^{e_1}} + \cdots + \frac{q_r(x)}{(1-\lambda_r x)^{e_r}} = 0. \tag{2.6}$$

By construction, each $q_i(x)$ is 0 or $\deg q_i(x) < e_i$. What can we say about each $q_i(x)$?

Lemma 2.3. *Let $\lambda_1, \ldots, \lambda_r$ in $K^\times$ be distinct such that (2.6) is satisfied, where $e_1, \ldots, e_r$ are positive integers and $q_1(x), \ldots, q_r(x)$ are in $K[x]$ with $q_i(x) = 0$ or $\deg q_i(x) < e_i$ for all i. Then every $q_i(x)$ is 0.*

Proof. We argue by induction on r. The case $r = 1$ is obvious. If $r \geq 2$ and the result is true for $r-1$ then multiply (2.6) through by the product $(1-\lambda_1 x)^{e_1} \cdots (1-\lambda_r x)^{e_r}$:

$$\sum_{i=1}^{r} q_i(x)(1-\lambda_1 x)^{e_1} \cdots \widehat{(1-\lambda_i x)^{e_i}} \cdots (1-\lambda_r x)^{e_r} = 0,$$

where the hat indicates an omitted factor in the ith term, for every i. Each term in this sum is a polynomial, and all the terms besides the one for $i = r$ have $(1-\lambda_r x)^{e_r}$ as a factor. Thus the term at $i = r$ is divisible by $(1-\lambda_r x)^{e_r}$. That term is $q_r(x)(1-\lambda_1 x)^{e_1} \cdots (1-\lambda_{r-1} x)^{e_{r-1}}$. Since $\lambda_1, \ldots, \lambda_{r-1}$ are distinct from λ_r, $(1-\lambda_r x)^{e_r}$ must divide $q_r(x)$. But $q_r(x)$, if not 0, has degree less than e_r by hypothesis. Therefore $q_r(x) = 0$, so the rth term in (2.6) is 0, which makes every other $q_i(x)$ equal to 0 by induction. □

Remark 2.4. This lemma becomes obvious if a term in (2.6) is moved to the other side, say $q_r(x)/(1-\lambda_r x)^{e_r}$. If $q_r(x) \neq 0$ then the right side blows up at $x = 1/\lambda_r$ since the numerator can't completely cancel the denominator (because $\deg q_r(x) < e_r$), but the left side without the term $q_r(x)/(1-\lambda_r x)^{e_r}$ has a finite value at $x = 1/\lambda_r$. Thus $q_r(x) = 0$.

Theorem 2.5. *For $\lambda_1, \ldots, \lambda_r \in K^\times$ and positive integers $e_1, \ldots, e_r$, the sequences $\{\binom{n}{k}\lambda_i^n\}$ for $i = 1, \ldots, r$ and $k = 0, \ldots, e_i - 1$ are linearly independent over K.*

Proof. If the sequences satisfy a K-linear relation (2.4) then applying Lemma 2.3 to (2.6) shows each $q_i(x)$ vanishes, so (2.5) vanishes for each i. Since (2.5) is the generating function of the sequence with nth term $\sum_{k=0}^{e_i-1} b_{ik}\binom{n}{k}\lambda_i^n$, we get

$$\sum_{k=0}^{e_i-1} b_{ik}\binom{n}{k}\lambda_i^n = 0 \text{ for each } i \text{ and all } n \geq 0. \tag{2.7}$$

We passed from a linear relation (2.4) involving several λ_i's to a linear relation (2.7) that involves just a single λ_i that is one of the inner sums in (2.4). In (2.7) we can cancel the common nonzero factor λ_i^n:

$$\sum_{k=0}^{e_i-1} b_{ik}\binom{n}{k} = 0 \text{ for all } n \geq 0.$$

Let's write this out as a system of linear equations at $n = 0, 1, \ldots, e_i - 1$:

$$\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ 1 & 2 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & n & \binom{n}{2} & \cdots & 1 \end{pmatrix} \begin{pmatrix} b_{i0} \\ b_{i1} \\ b_{i2} \\ \vdots \\ b_{i\,e_i-1} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{2.8}$$

The matrix is invertible in K, so $b_{ik} = 0$ for all i and k. □

Remark 2.6. Linear recursions over finite fields are of interest to coding theorists because of their close relation to cyclic codes, a special type of linear code. The important constructions of cyclic codes, like Reed–Solomon and BCH codes, are related to linear recursions whose characteristic polynomial[5] has distinct roots. Cyclic codes where the characteristic polynomial has repeated roots have been studied [1], and for a number of reasons they are not competitive with the standard "distinct root" cyclic codes.

3. Interlude: Analogy with Differential Equations

Linear recursions are analogous to linear differential equations, and our second proof of Theorem 1.3 will be motivated by this analogy, which we set up in this section.

A sequence $\{a_n\}$ satisfying $a_n = c_1 a_{n-1} + c_2 a_{n-2} + \cdots + c_d a_{n-d}$ can be compared with a function $y(t)$ satisfying

$$y^{(d)}(t) = c_1 y^{(d-1)}(t) + c_2 y^{(d-2)}(t) + \cdots + c_d y(t), \tag{3.1}$$

which is a dth-order linear ODE with constant coefficients. The solution space to such an ODE is d-dimensional. How similar are solutions to the recursion and the ODE?

Example 3.1. A first-order linear recursion $a_n = ca_{n-1}$ has general solution $a_n = a_0 c^n$, while a first-order ODE of the form $y'(t) = cy(t)$ has general solution $y(t) = y(0)e^{ct}$. The geometric progression c^n is analogous to the exponential function e^{ct}.

Example 3.2. A second-order linear recursion $a_n = ba_{n-1} + ca_{n-2}$ has a general solution that depends on whether or not the factorization $1 - bx - cx^2 = (1 - \lambda x)(1 - \mu x)$ has distinct or repeated reciprocal roots:

$$a_n = \begin{cases} \alpha\lambda^n + \beta\mu^n, & \text{if } \lambda \neq \mu, \\ \alpha\lambda^n + \beta n\lambda^n, & \text{if } \lambda = \mu. \end{cases}$$

A second-order ODE of the form $y''(t) = by'(t) + cy(t)$ has a general solution that depends on whether or not the factorization $x^2 - bx - c = (x - \lambda)(x - \mu)$ has distinct or repeated roots:

$$y(t) = \begin{cases} \alpha e^{\lambda t} + \beta e^{\mu t}, & \text{if } \lambda \neq \mu, \\ \alpha e^{\lambda t} + \beta t e^{\lambda t}, & \text{if } \lambda = \mu. \end{cases}$$

Letting $D = d/dt$, the differential equation (3.1) can be written as

$$(D^d - c_1 D^{d-1} - \cdots - c_d)(y(t)) = 0, \tag{3.2}$$

so solutions of (3.1) are the nullspace of the differential operator

$$D^d - c_1 D^{d-1} - \cdots - c_d, \tag{3.3}$$

[5] This is $x^d - c_1 x^{d-1} - \cdots - c_d$ in our notation.

which acts on the real vector space of smooth functions $\mathbf{R} \to \mathbf{R}$. On sequences, the analogue of D is the left-shift operator L: if $\mathbf{a} = (a_0, a_1, a_2, \dots)$ then $L(\mathbf{a}) = (a_1, a_2, a_3, \dots)$, or equivalently $L(\{a_n\}) = \{a_{n+1}\}$. This is a linear operator on the K-vector space $\mathrm{Seq}(K)$ of sequences with coordinates in K. Here is the analogue of (3.2) that can serve as a characterization of sequences satisfying a linear recursion in place of Theorem 2.1.

Theorem 3.3. *A sequence* $\mathbf{a} = \{a_n\}$ *in* $\mathrm{Seq}(K)$ *satisfies the linear recursion* (1.1) *if and only if* $(L^d - c_1 L^{d-1} - \cdots - c_d I)(\mathbf{a}) = \mathbf{0}$*, where* I *is the identity operator on* $\mathrm{Seq}(K)$ *and* $\mathbf{0} = (0, 0, 0, \dots)$.

Proof. For $i \geq 0$, the sequence $L^i(\mathbf{a})$ has nth component a_{n+i}, so the sequence $c_i L^i(\mathbf{a})$ has nth component $c_i a_{n+i}$. The nth component of $(L^d - c_1 L^{d-1} - \cdots - c_d I)(\mathbf{a})$ is $a_{n+d} - c_1 a_{n+d-1} - \cdots - c_d a_n$, which is 0 for all n if and only if $\mathbf{a}$ satisfies (1.1). □

Example 3.4. If a sequence $\mathbf{a}$ satisfies $a_n = a_{n-1} + a_{n-2}$ then $(L^2 - L - I)(\mathbf{a})$ has nth component $a_{n+2} - a_{n+1} - a_n$, which is 0 for all n, so $(L^2 - L - I)(\mathbf{a}) = \mathbf{0}$.

To solve the differential equation (3.2), factor the polynomial $x^d - c_1 x^{d-1} - \cdots - c_d$ over $\mathbf{C}$ to get a factorization of the differential operator (3.3):

$$x^d - c_1 x^{d-1} - \cdots - c_d = \prod_{i=1}^{r} (x - \lambda_i)^{e_i} \Longrightarrow D^d - c_1 D^{d-1} - \cdots - c_d = \prod_{i=1}^{r} (D - \lambda_i)^{e_i},$$

where $\lambda_1, \dots, \lambda_r$ are distinct and $e_i \geq 1$. We have to allow $\lambda_i \in \mathbf{C}$, so for compatibility let $D = d/dt$ act on the smooth functions $\mathbf{R} \to \mathbf{C}$ (functions whose real and imaginary parts are ordinary smooth functions $\mathbf{R} \to \mathbf{R}$). The operators $(D - \lambda_i)^{e_i}$ for $i = 1, 2, \dots, r$ commute, so $\mathbf{C}$-valued solutions $y(t)$ to the differential equation $(D - \lambda_i)^{e_i}(y(t)) = 0$ are solutions to (3.1). This is enough to describe all solutions of (3.1):

Theorem 3.5. *Using the above notation, a* $\mathbf{C}$*-basis of solutions to* $(D - \lambda_i)^{e_i}(y(t)) = 0$ *is* $e^{\lambda_i t}, te^{\lambda_i t}, \dots, t^{e_i - 1} e^{\lambda_i t}$*, and putting these together for* $i = 1, \dots, r$ *gives a* $\mathbf{C}$*-basis of solutions to* (3.1).

We omit a proof of this theorem, which for us serves only as motivation. In the simplest case that each e_i is 1, so $x^d - c_1 x^{d-1} - \cdots - c_d = \prod_{i=1}^{d} (x - \lambda_i)$, a basis of the solution space to (3.1) is $e^{\lambda_1 t}, \dots, e^{\lambda_d t}$, which is analogous to Example 1.4.

Remark 3.6. Since the λ_i in Theorem 3.5 are in $\mathbf{C}$, the solution space in the theorem is the complex-valued solutions of (3.1). If the coefficients c_i in (3.1) are all real, then even if some λ_i in Theorem 3.5 is not real and therefore some $t^k e^{\lambda_i t}$ is not a real-valued function, it can be proved that the $\mathbf{R}$-valued solution space to (3.1) is d-dimensional over $\mathbf{R}$.

4. Second Proof: Linear Operators

We will apply the ideas from Section 3 to the linear operator $L^d - c_1 L^{d-1} - \cdots - c_d I$ in Theorem 3.3 to reprove Theorem 1.3 using an argument of Anna Medvedovsky [6, App. B].

Our first approach to proving Theorem 1.3 involved the polynomial $1 - c_1 x - \cdots - c_d x^d$ and its reciprocal roots (and their multiplicities). By analogy with the method of solving differential equations, we will now use the polynomial $x^d - c_1 x^{d-1} - \cdots - c_d$ instead. These two polynomials are reciprocal in the sense that

$$x^d - c_1 x^{d-1} - \cdots - c_d = x^d \left(1 - \frac{c_1}{x} - \cdots - \frac{c_d}{x^d}\right).$$

Therefore

$$x^d - c_1x^{d-1} - \cdots - c_d = \prod_{i=1}^{r}(x-\lambda_i)^{e_i} \iff 1 - c_1x - \cdots - c_dx^d = \prod_{i=1}^{r}(1-\lambda_i x)^{e_i},$$

where $\lambda_i \neq \lambda_j$ for $i \neq j$, so reciprocal roots of $1 - c_1x - c_2x^2 - \cdots - c_dx^d$ are ordinary roots of $x^d - c_1x^{d-1} - \cdots - c_d$, with matching multiplicities. Since $c_d \neq 0$, no λ_i is 0.

Theorem 3.3 tells us that a sequence $\mathbf{a} \in \mathrm{Seq}(K)$ satisfies (1.1) precisely when $\mathbf{a}$ is in the kernel of $L^d - c_1L^{d-1} - \cdots - c_dI$. Since

$$x^d - c_1x^{d-1} - \cdots - c_d = \prod_{i=1}^{r}(x-\lambda_i)^{e_i} \Longrightarrow L^d - c_1L^{d-1} - \cdots - c_dI = \prod_{i=1}^{r}(L-\lambda_i I)^{e_i} \tag{4.1}$$

and the operators $(L-\lambda_i I)^{e_i}$ for different λ_i commute, any $\mathbf{a}$ in the kernel of some $(L-\lambda_i I)^{e_i}$ is a solution of (1.1). Solutions to $(L-\lambda_i I)^{e_i}(\mathbf{a}) = \mathbf{0}$ belong to the *generalized λ_i-eigenspace* of L, which is the set of $\mathbf{a}$ killed by some positive integer power of $L - \lambda_i I$. If $e_i = 1$, such $\mathbf{a}$ form the λ_i-eigenspace of L: $(L - \lambda_i I)(\mathbf{a}) = \mathbf{0}$ if and only if $L(\mathbf{a}) = \lambda_i\mathbf{a}$. The λ_i-eigenvectors are the nonzero vectors in the λ_i-eigenspace, and the nonzero vectors in a generalized eigenspace are called *generalized eigenvectors*.

Our second proof of Theorem 1.3, like the first, is established in two steps by proving results like Corollary 2.2 and Theorem 2.5.

Theorem 4.1. *If $\lambda \in K^\times$ is a root of $x^d - c_1x^{d-1} - \cdots - c_d$ with multiplicity $e \geq 1$ then the sequence $\{\binom{n}{k}\lambda^n\}$ for $k = 0, 1, \ldots, e-1$ satisfies (1.1).*

Proof. Since $(x-\lambda)^e$ is a factor of $x^d - c_{d-1}x^{d-1} - \cdots - c_d$, it suffices by Theorem 3.3 to show the sequence $\{\binom{n}{k}\lambda^n\}$ for $0 \leq k \leq e-1$ is killed by $(L-\lambda I)^e$ to make it satisfy (1.1).

First we treat $k = 0$. Since $L(\{\lambda^n\}) = \{\lambda^{n+1}\} = \lambda\{\lambda^n\}$, we get $(L - \lambda I)(\{\lambda^n\}) = \mathbf{0}$. Therefore $(L - \lambda I)^e(\{\lambda^n\}) = \mathbf{0}$.

Let $k \geq 1$. Applying $L - \lambda I$ to $\{\binom{n}{k}\lambda^n\}$, we get the sequence

$$(L-\lambda I)\left\{\binom{n}{k}\lambda^n\right\} = L\left\{\binom{n}{k}\lambda^n\right\} - \lambda\left\{\binom{n}{k}\lambda^n\right\} = \left\{\binom{n+1}{k}\lambda^{n+1} - \binom{n}{k}\lambda^{n+1}\right\}.$$

For $k \geq 1$, $\binom{n+1}{k} = \binom{n}{k-1} + \binom{n}{k}$, so

$$(L-\lambda I)\left\{\binom{n}{k}\lambda^n\right\} = \left\{\binom{n}{k-1}\lambda^{n+1}\right\}.$$

By induction,

$$(L-\lambda I)^i\left\{\binom{n}{k}\lambda^n\right\} = \left\{\binom{n}{k-i}\lambda^{n+i}\right\}$$

for $1 \leq i \leq k$. Thus $(L-\lambda I)^k(\{\binom{n}{k}\lambda^n\}) = \{\lambda^{n+k}\} = \lambda^k\{\lambda^n\}$. The sequence $\{\lambda^n\}$ is a λ-eigenvector of L, and hence is in the kernel of $L - \lambda I$, so applying $L - \lambda I$ more than k times to the sequence $\{\binom{n}{k}\lambda^n\}$ kills it. Thus $(L-\lambda I)^e(\{\binom{n}{k}\lambda^n\}) = \mathbf{0}$ for $e > k$. □

Second proof of Theorem 2.5. Suppose a K-linear combination of these sequences vanishes, say

$$\sum_{i=1}^{r}\sum_{k=0}^{e_i-1} b_{ik}\left\{\binom{n}{k}\lambda_i^n\right\} = \mathbf{0}. \tag{4.2}$$

with $b_{ik} \in K$. We want to show each b_{ik} is 0.

For each i, the sequences $\{\binom{n}{k}\lambda_i^n\}$ for $0 \le k \le e_i - 1$ are all killed by $(L - \lambda_i I)^{e_i}$ by Theorem 4.1, so the inner sum $\mathbf{v}_i := \sum_{k=0}^{e_i-1} b_{ik}\{\binom{n}{k}\lambda_i^n\}$ in (4.2) belongs to the generalized λ_i-eigenspace of L. A standard theorem in linear algebra says that eigenvectors of a linear operator associated to different eigenvalues are linearly independent, and this extends to *generalized* eigenvectors of a linear operator associated to different eigenvalues; a proof of that is in the appendix and serves as an analogue of Lemma 2.3. Since $\mathbf{v}_1, \dots, \mathbf{v}_r$ belong to generalized eigenspaces associated to different eigenvalues of L, and $\mathbf{v}_1 + \cdots + \mathbf{v}_r = \mathbf{0}$, each $\mathbf{v}_i$ must be $\mathbf{0}$; if any $\mathbf{v}_i$ were not $\mathbf{0}$ then the vanishing sum over the nonzero $\mathbf{v}_i$ would be a linear dependence relation among generalized eigenvectors associated to distinct eigenvalues.

The equation $\mathbf{v}_i = \mathbf{0}$ says

$$\sum_{k=0}^{e_i-1} b_{ik}\left\{\binom{n}{k}\lambda_i^n\right\} = \mathbf{0}. \tag{4.3}$$

The passage from (4.2) to (4.3) is an analogue of the passage from (2.4) to (2.7), and (4.3) for $i = 1, \dots, r$ is exactly the same as (2.7), so we can finish off this proof in the same way that we did before: equating the coordinates on both sides of (4.3) for $n = 0, \dots, e_i - 1$ and dividing by λ_i^n leads to the matrix equation (2.8) so all b_{ik} are 0.

□

Appendix A. Linear Independence of Generalized Eigenvectors

Theorem A.1. *Let V be a K-vector space, $A\colon V \to V$ be a linear operator, and $v_1, \dots, v_r$ in V be generalized eigenvectors of A associated to distinct respective eigenvalues $\lambda_1, \dots, \lambda_r$. Then $v_1, \dots, v_r$ are linearly independent over K.*

Proof. Since v_i is a generalized eigenvector of A associated to the eigenvalue λ_i, $v_i \neq 0$ and $(A - \lambda_i I)^{e_i}(v_i) = 0$ for some $e_i \ge 1$. Suppose there is a linear relation

$$b_1 v_1 + \cdots + b_r v_r = 0$$

for some $b_1, \dots, b_r \in K$. We want to prove each b_i is 0, and will argue by induction on r. The result is clear if $r = 1$, since $v_1 \neq 0$, so suppose $r \ge 2$ and the lemma is proved for $r - 1$ generalized eigenvectors associated to distinct eigenvalues.

The operators $(A - \lambda_i I)^{e_i}$ commute, so applying the product $(A - \lambda_1 I)^{e_1} \cdots (A - \lambda_{r-1})^{e_{r-1}}$ to the linear relation kills the first $r - 1$ terms and leaves us with

$$b_r (A - \lambda_1 I)^{e_1} \cdots (A - \lambda_{r-1})^{e_{r-1}}(v_r) = 0.$$

If $b_r \neq 0$ then v_r is killed by $(A - \lambda_1 I)^{e_1} \cdots (A - \lambda_{r-1})^{e_{r-1}}$. It is also killed by $(A - \lambda_r I)^{e_r}$. In $K[x]$ the polynomials $(x - \lambda_1)^{e_1} \cdots (x - \lambda_{r-1})^{e_{r-1}}$ and $(x - \lambda_r)^{e_r}$ are relatively prime (since $\lambda_r \neq \lambda_1, \dots, \lambda_{r-1}$), so there's a polynomial identity

$$g(x)(x - \lambda_1)^{e_1} \cdots (x - \lambda_{r-1})^{e_{r-1}} + h(x)(x - \lambda_r)^{e_r} = 1$$

for some $g(x)$ and $h(x)$ in $K[x]$. Thus

$$g(A)(A - \lambda_1 I)^{e_1} \cdots (A - \lambda_{r-1} I)^{e_{r-1}} + h(A)(A - \lambda_r I)^{e_r} = I,$$

and applying both sides to v_r implies $0 = v_r$, which is a contradiction. Thus $b_r = 0$. The linear relation among the v_i simplifies to $b_1 v_1 + \cdots + b_{r-1} v_{r-1} = 0$, so by induction all the remaining b_i equal 0. □

References

[1] G. Castagnoli, J. L. Massey, P. A. Schoeller, and N. von Seeman, "On Repeated-Root Cyclic Codes," *IEEE Trans. Inform. Theory* **37** (1991), 337–342.

[2] H. T. Engstrom, "On sequences defined by linear recurrence relations," *Trans. AMS* **33** (1931), 210–218. URL http://www.ams.org/journals/tran/1931-033-01/S0002-9947-1931-1501585-5/S0002-9947-1931-1501585-5.pdf.

[3] E. Catalan, *Manuel des Candidats à l'École Polytechnique*, Tome 1, 1857. URL https://archive.org/details/manueldescandid00catagoog/mode/2up.

[4] J. P. Fillmore and M. L. Marx, "Linear Recursive Sequences," *SIAM Review* **10** (1968), 342–353.

[5] R. J. McEliece, "Linear Recurring Sequences over Finite Fields," Ph.D. thesis, Caltech, 1967. URL http://thesis.library.caltech.edu/3856/1/McEliece_rj_1967.pdf.

[6] A. Medvedovsky, "Lower bounds on dimensions of mod-p Hecke algebras: The nilpotence method," Ph.D. thesis, Brandeis, 2015.

[7] L. M. Milne-Thomson, *The Calculus of Finite Differences*, MacMillan and Co., London, 1933. URL https://archive.org/details/calculusoffinite032017mbp/mode/2up?view=theater.