Linear Recurrences

Nikolai V. Ivanov

January 17, 2008

Let k be a natural number and let $a_1, a_2, \ldots, a_k$ be some complex numbers. We are interested in sequences $F(0), F(1), \ldots, F(m), \ldots$ of complex numbers satisfying the relation

$$\text{(R)} \qquad F(n) = a_1 F(n-1) + a_2 F(n-2) + \ldots + a_k F(n-k)$$

for all $n \geqslant k$. A relation of this form is called a *linear recurrence relation.* Without any loss of generality we may assume that $a_k \neq 0$. Slightly more generally, one can allow a non-zero coefficient in front of $F(n)$. Clearly, every relation of this more general form is equivalent to a relation of the form (R). A sequence $F(0), F(1), \ldots$ satisfying (R) for all $n \geqslant k$ will be called a *solution of* (R).

Our goal is to give a description of all solutions of (R). We will first approach this problem directly, by giving explicit formulas for solutions. Then we will describe the solutions in terms of their generating functions.*

Explicit formulas

Solutions of (R) form a vector space V over $\mathbf{C}$. Since every solution is determined by k numbers $F(0), F(1), \ldots, F(k-1)$, and these k numbers can be chosen arbitrarily, V has dimension k. Our goal in this section is to give explicit formulas for sequences in a basis of V. Since every solution of (R) is a linear combination of the elements of a basis of V, such formulas lead to an obvious description of all solutions of (R).

---

*Two aspects of our exposition appear to be non-standard. The first one is the notion of the *derived relation* (see the paragraph before Lemma 1), which leads to an elegant proof of the fact that sequences (1) from Theorem 2 are solutions of (R). The second point is the computation of a generalized Vandermonde determinant in Theorem 4, which definitely deserves to be more widely known.

Let us first look for solutions of the form $F(n) = \lambda^n$ for some non-zero $\lambda \in \mathbf{C}$. The relation (R) takes the form

$$\lambda^n = a_1\lambda^{n-1} + a_2\lambda^{n-2} + \ldots + a_k\lambda^{n-k},$$

which is equivalent to

$$\lambda^k - a_1\lambda^{k-1} - a_2\lambda^{k-2} - \ldots - a_k = 0$$

(because $\lambda \neq 0$). In other words, λ has to be a root of the polynomial

$$f(x) = x^k - a_1x^{k-1} - a_2x^{k-2} - \ldots - a_k,$$

which is called the *characteristic polynomial of* (R).

Theorem 1. *If the characteristic polynomial $f(x)$ has k distinct roots $\lambda_1, \ldots, \lambda_k$, then k sequences*

$$F_i(n) = \lambda_i^n,$$

$1 \leqslant i \leqslant k$, from a basis of V.

Proof. Clearly, every $F(n) = \lambda_i^n$ is a solution. Since the space of the solutions is k-dimensional, it is sufficient to prove that these k solutions are linearly independent. In order to prove that that sequences $F(n) = \lambda_i^n$ are linearly independent, it is sufficient to prove that their initial segments

$$1, \quad \lambda_i, \quad \lambda_i^2, \quad \ldots, \quad \lambda_i^{k-1}$$

are linearly independent. The linear independence of these segments is equivalent to the non-vanishing of the determinant

$$\begin{vmatrix} 1 & \lambda_1 & \lambda_1^2 & \ldots & \lambda_1^{k-1} \\ 1 & \lambda_2 & \lambda_2^2 & \ldots & \lambda_2^{k-1} \\ \multicolumn{5}{c}{\ldots\ldots\ldots\ldots\ldots\ldots\ldots} \\ 1 & \lambda_k & \lambda_k^2 & \ldots & \lambda_k^{k-1} \end{vmatrix}.$$

This is well known Vandermonde determinant, which is equal to

$$\prod_{i>j}(\lambda_i - \lambda_j).$$

Since λ_i are assumed to be distinct, this product is non-zero. The theorem follows. □

Remark. As is well known, k vectors in a k-dimensional vector space W from a basis if and only they are linearly independent. Similarly, k vestors in a k-dimensional vector space W from a basis if and only if they span W. The proof of Theorem 1 is based on the first of these two characterisations of bases. In what follows, we will use both characterizations several times.

When the characteristic polynomial has multiple roots, there is not enough solutions of the form $F(n) = \lambda^n$ to form a basis. A multiple root of a polynomial is also a root of its derivative. This suggests to look for the relation having the derivative $f'(x)$ of $f(x)$ as its characteristic polynomial. Let us call the relation

$$kF(n) = a_1(k-1)F(n-1) + a_2(k-2)F(n-2) + \ldots + a_{k-1}F(n-k+1)$$

the *derived relation of* (R). Strictly speaking, we should divide it by k in order to bring it to the form (R). Clearly, the characteristic polynomial of the derived relation of (R) is equal to the derivative $f'(x)$. By analogy with the usual derivatives, we may consider the *higher derived relations.* We will denote by (R$'$) the derived relation of (R), by (R$''$) the derived relation of (R$'$), and, in general, by (R$^{(d)}$) the derived relation of (R$^{(d-1)}$).

Lemma 1. *If the sequence $F(n)$ is a solution of both* (R) *and* (R$'$)*, then the sequence $nF(n)$ is a solution of* (R)*.*

Proof. (R) implies that

$$(n-k)F(n) = a_1(n-k)F(n-1)+a_2(n-k)F(n-2)+\ldots+a_k(n-k)F(n-k).$$

By taking the sum of this relation with with (R$'$), we get

$$nF(n) = (n-1)a_1F(n-1) + (n-2)a_2F(n-2) + \ldots + (n-k)a_kF(n-k).$$

This means that $nF(n)$ is a solution of (R). $\square$

Corollary 1. *If the sequence $F(n)$ is a solution of* (R)*,* (R$'$)*, ...,* (R$^{(d)}$)*, then the sequences $nF(n)$, $n^2F(n)$,..., $n^dF(n)$ are solutions of* (R)*.*

Proof. By the lemma, the corollary is true for $d = 1$. Let us assume that we already proved the corollary for for $d = D$, and prove it for $d = D + 1$. Let $F(n)$ be a solution of (R), (R$'$), ... (R$^{(D+1)}$). If we apply the corollary for $d = D$ to $F(n)$ and (R), we conclude that $n^DF(n)$ is a solution of (R). If we apply the corollary for $d = D$ to $F(n)$ and (R′), we conclude that $n^D F(n)$ is a solution of (R′). Now, the lemma implies that $n^{D+1}F(n)$ is a solution of (R). The induction completes the proof. □

Suppose λ is a root of $f(x)$ of the multiplicity d. Then λ is also a root of derivatives $f'(x), f''(x), \ldots, f^{(d-1)}$ of $f(x)$ of order $\leqslant d-1$. Therefore, the sequence $F(n) = \lambda^n$ is a solution of (R) and of the derived relations $(\mathrm{R}), (\mathrm{R}'), \ldots, (\mathrm{R}^{(d-1)})$. Now, the corollary implies that d sequences $\lambda^n, n\lambda^n, \ldots, n^{d-1}\lambda^n$ are solutions of (R). So, for a root of $f(x)$ of the multiplicity d we get d solutions of our recurrence relation. Since $f(x)$ has exactly k roots counted with their multiplicity, altogether we get k solutions.

Theorem 2. *Let $\lambda_1, \ldots, \lambda_m$ be the roots of the characteristic polynomial $f(x)$ of multiplicity $d_1, \ldots, d_m$ respectively. The k solutions*

$$F(n) = n^a \lambda_i^n, \tag{1}$$

where $1 \leqslant i \leqslant m$ and $0 \leqslant a \leqslant d_i - 1$, form a basis of V.

Before proving this theorem, let us introduce another family of solutions of (R). Recall that the *falling factorials*† are defined as

$$n^{\underline{a}} = n(n-1)\ldots(n-a+1)$$

for $a > 0$. We set $n^{\underline{0}} = 1$.

Theorem 3. *Let $\lambda_1, \ldots, \lambda_m$ be the roots of the characteristic polynomial $f(x)$ of multiplicity $d_1, \ldots, d_m$ respectively. The k sequences*

$$F(n) = n^{\underline{a}} \lambda_i^n, \tag{2}$$

where $1 \leqslant i \leqslant m$ and $0 \leqslant a \leqslant d_i - 1$, form a basis of V.

Proof of Theorems 2 and 3. Clearly, $n^{\underline{a}} = n^a +$terms of lower degree in n. This implies that the sequences (2) are linear combinations of the solutions (1), and, therefore, are indeed solutions and do belong to V. Notice that our solutions (1) or (2) form a basis of V if and only if they span V (because we have exactly $\dim V$ solutions). Clearly, if solution (1) do not span V, their linear combinations (2) do not span V either. Therefore, both theorems will

†It would be better to call them *falling powers*.

follow once we prove that solutions (2) span V. This fact, in turn, will follow once we prove that solutions (2) are linearly independent (again, because we have exactly $\dim V$ solutions).

Since we assume that our relation (R) has length k, the free term a_k of $f(x)$ is non-zero, and therefore none of the roots λ_i of $f(x)$ is equal to zero. It follows that it is sufficient to prove that the sequences

$$F(n) = n^{\underline{a}}\lambda_i^{n-a}, \tag{3}$$

where $1 \leqslant i \leqslant m$ and $0 \leqslant a \leqslant d_i - 1$ are linearly independent (obviously, they are solutions of (R)).

As in the proof of Theorem 1, it is sufficient to prove that the initial segments of length k of sequences (3) are linearly independent, and this amounts to proving that the determinant having these initial segments as rows is non-zero. Such a determinant with rows arranged in a natural order is computed in the next theorem. See (5). Since λ_i's are all different, the product in (5) is non-zero. This proves Theorems 2 and 3 modulo the next theorem. □

Theorem 4. *Let $\lambda_1, \dots, \lambda_m$ be different non-zero complex numbers. Let $d_1, \dots, d_m$ be some positive integers, and let $k = d_1 + \dots + d_m$. Let A be a $k \times k$ matrix having the sequences*

$$\{n^{\underline{a}}\lambda_i^{n-a}\}_{0 \leqslant n \leqslant k-1}, \tag{4}$$

where $1 \leqslant i \leqslant m$ and $0 \leqslant a \leqslant d_i - 1$, as rows. Suppose that the rows are ordered in the increasing order of i, and for a fixed i in the increasing order of a. Then

$$\det(A) = \left(\prod_{j=1}^{m} \prod_{j=0}^{d_i - 1} j! \right) \prod_{p>q} (\lambda_p - \lambda_q)^{d_p d_q}. \tag{5}$$

Proof. We will compute this determinant following an idea of Gel'fond; see [Ge], Chapter I, Section 4.2.

Let us start with the Vandermonde determinant

$$\begin{vmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{k-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{k-1} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_k & x_k^2 & \dots & x_k^{k-1} \end{vmatrix}.$$

Denote by D_j the partial derivative $\partial/\partial x_j$. We claim that applying the iterated partial derivative D_j^a to the Vandermonde determinant amounts to applying D_j^a to its j-th row. Indeed, let

$$1 \cdot A_0 - x_j A_1 + \ldots \pm x_j^{k-1} A_{k-1}$$

be the expansion of the Vandermonde determinant by j-th row. Since A_i's do not depend on x_j, the derivatives $D_j^a A_i$ are all equal to 0. The claim follows.

A similar argument shows that applying $D_1^{a_1} D_2^{a_2} \ldots D_k^{a_k}$ to the Vandermonde determinant amounts to the simultaneous application of the derivative $D_j^{a_j}$ to j-th row for all j, $1 \leqslant j \leqslant k$.

By differentiating a times x_j^n, we get $D_j^a(x_j^n) = n^{\underline{a}} x_j^{n-a}$. Therefore, if apply D_j^a to j-th row of the Vandermonde determinant, we will get the sequence (a row)

$$\{n^{\underline{a}} x_j^{n-a}\}_{0 \leqslant n \leqslant k-1}.$$

If we replace x_j by λ_i, this will be exactly the sequence (4).

Now we see that we can get our determinant in the following manner. Apply $D_1^0 = \mathrm{id}$ to the first row, D_2^1 to the second row, and continue in this way till the d_1-th row, to which we apply $D_{d_1}^{d_1-1}$. Apply

$$D_{d_1+1}^0, D_{d_2+2}^1, \ldots D_{d_1+d_2}^{d_2-1}$$

to the next d_2 rows respectively,

$$D_{d_1+d_2+1}^0, D_{d_1+d_2+2}^1, \ldots D_{d_1+d_2+d_3}^{d_3-1}$$

to the next d_3 rows, etc. Next, set

$$x_1 = x_2 = \ldots = x_{d_1} = \lambda_1,$$

$$x_{d_1+1} = x_{d_1+2} = \ldots = x_{d_1+d_2} = \lambda_2,$$

etc. The result will be our determinant $\det(A)$.

Let us call *clusters* the sets of variables x_i which are set to be equal at the last step. So,

$$\{x_1, x_2, \ldots, x_{d_1}\}$$

is the first cluster,

$$\{x_{d_1+1}, x_{d_1+2}, \ldots, x_{d_1+d_2}\}$$

is the second cluster, etc.

Now, we will apply the same derivatives to the standard formula

$$\prod_{i>j}(x_i - x_j) \tag{6}$$

for the Vandermonde determinant, and then substitute λ_p's for x_i's in the same clusters in the above manner. This will evaluate $\det(A)$.

Notice that any derivative of order $\geqslant 2$ of any factor $(x_i - x_j)$ is equal to 0. Therefore, Leibniz formula implies that the result of application of our derivatives to (6) is a sum of terms of the form

$$\prod_{i>j} D_{i,j}(x_i - x_j), \tag{7}$$

where each $D_{i,j}$ is equal either to some D_r, or to id. Such a term can be non-zero only if each $D_{i,j}$ is equal to either D_i, or D_j, or id.

After computing the derivatives, we will set the variables in each cluster to be equal. In order to get a non-zero term after this substitution, we need to apply D_i or D_j to every factor $(x_i - x_j)$ such that x_i, x_j belong to the same cluster. Consider, for example, the first cluster. D_1 is not available (since we are applying only $D_1^0 = \mathrm{id}$), so, D_2 should be applied to $(x_2 - x_1)$. After D_2 is used, we need two D_3's to apply them to $(x_3 - x_2)$ and $(x_3 - x_1)$. These two D_3 are available, since we are applying D_3^2 to our expression. Continuing in this way, we see that the available $j-1$ derivatives D_j's need to be applied to the factors $(x_j - x_{j-1}), \ldots, (x_j - x_1)$. We see that for a term (7) to remain non-zero after setting equal the variables in clusters, it should have the form

$$\left(\prod_{C} \prod_{\substack{i>j \\ i,j \in C}} D_i(x_i - x_j)\right) \prod_{r>s}(x_r - x_s), \tag{8}$$

where the first product is taken over all clusters C, and the last product is taken over $r > s$ such that x_r, x_s belong to different clusters.

The application of Leibniz formula will lead to many terms equal to (8). Let $D = D_j$ for some j. If we compute $D^j(f_1 f_2 \cdots f_j)$ by Leibniz formula, and then discard terms involving derivatives of order $\geqslant 2$ (which are equal to 0 in our case), we get a sum of terms equal to $D(f_1)D(f_2)\cdots D(f_j)$. To get such a term, we need to pick up a f_a to apply the first D, then pick up a f_b with $b \neq a$ to apply the second D, etc. Clearly, this can be done in $j!$ ways, and, therefore, there are $j!$ terms equal to $D(f_1)D(f_2)\cdots D(f_j)$ in our sum.

By applying this observation to our situation, and noticing that there are m clusters of sizes are $d_1, d_2, \ldots, d_m$ respectively, we see that every term (8) appears in our calculation

$$\prod_{i=1}^{m} \prod_{j=0}^{d_i-1} j! \tag{9}$$

times.

Now, it is time to observe that every derivative in (8) is equal to 1, and, therefore, the product over clusters in (8) is equal to 1. Therefore, (8) is equal to

$$\prod_{r>s} (x_r - x_s),$$

where the product is taken over $r > s$ such that x_r, x_s belong to different clusters. We need to take sum of (9) such equal terms. The result is

$$\left(\prod_{i=1}^{m} \prod_{j=0}^{d_i-1} j! \right) \prod_{r>s} (x_r - x_s).$$

Finally, we have to substitute λ_p's for x_r, x_s in this expression. Recalling that d_p is the size of the cluster of variables set to be equal to λ_p, we see that the result is equal to the right hand side of (5). This completes the proof. □

Generating functions

We need a version of Theorem 3. Recall that the *rising factorials*[‡] are defined as

$$n^{\overline{a}} = n(n+1)\ldots(n+a-1)$$

for $a > 0$. We set $n^{\overline{0}} = 1$.

Theorem 5. *Let $\lambda_1, \ldots, \lambda_m$ be the roots of the characteristic polynomial $f(x)$ of multiplicity $d_1, \ldots, d_m$ respectively. The k sequences*

$$F(n) = (n+1)^{\overline{a}} \lambda_i^n, \tag{10}$$

where $1 \leqslant i \leqslant m$ and $0 \leqslant a \leqslant d_i - 1$, form a basis of V.

[‡]It would be better to call them *rising powers*.

Proof. Clearly, $(n+1)^{\overline{a}} = n^a +$ terms of lower degree in n. This implies (by using induction on a, for example) that, vice versa, $n^a = (n+1)^{\overline{a}} +$ a linear combination of $(n+1)^{\overline{b}}$ with $b < a$. Therefore, solutions (1) are linear combinations of solutions (10). Since the solutions (1) form a basis of V, the solutions (10) span V. Since there are k of them, and $\dim V = k$, this implies that they form a basis of V. □

The generating functions of sequences (10) are provided by the following lemma.

Lemma 2. *Suppose that $\lambda \neq 0$. The generating function of the sequence*

$$F(n) = (n+1)^{\overline{a}} \lambda^n$$

is equal to $\lambda^{-a} G^{(a)}(z)$, where $G(z) = (1-\lambda z)^{-1}$, and $G^{(a)}(z)$ is, as usual, the a-th derivative of $G(z)$.

Proof. Recall that

$$G(z) = (1-\lambda z)^{-1} = \sum_{m \geqslant 0} \lambda^m z^m.$$

By differentiating this formula a times, we get

$$G^{(a)}(z) = \sum_{m \geqslant a} m(m-1)\dots(m-a+1)\lambda^m z^{m-a} =$$

$$= \lambda^a \sum_{n \geqslant 0} (n+a)(n+a-1)\dots(n+1)\lambda^{n+a} z^n =$$

$$= \lambda^a \sum_{n \geqslant 0} (n+1)^{\overline{a}} \lambda^n z^n$$

(on the second line we made a change of variables $n = m - a$). The lemma follows. □

The derivatives $G^{(a)}(x)$ can be easily computed.

Lemma 3. *If $G(z) = (1-\lambda z)^{-1}$, then $G^{(a)}(z) = a!\lambda^a (1-\lambda z)^{-a-1}$.*

Proof. Use the induction by a. □

Corollary 2. *Let $\lambda_1, \ldots, \lambda_m$ be the roots of the characteristic polynomial $f(x)$ of multiplicity $d_1, \ldots, d_m$ respectively. A function is the generating function of a solution of (R) if and only if it is a linear combination (with complex coefficients) of functions*

$$(1-\lambda_i z)^{-b},$$

where $1 \leqslant i \leqslant m$ and $1 \leqslant b \leqslant d_i$.

Proof. By Theorem 5 and Lemmas 2 and 3, a function is the generating function of a solution if and only if it is a linear combination of of functions $\lambda^{-a}(a!\lambda^a(1-\lambda z)^{-a-1}) = a!(1-\lambda_i z)^{-a-1}$, where $1 \leqslant i \leqslant m$ and $0 \leqslant a \leqslant d_i - 1$. Clearly, we can discard the multiplicative constants $a!$ in this statement. The corollary follows (after we set $b = a+1$). □

Corollary 3. *Let $\lambda_1, \ldots, \lambda_m$ be the roots of the characteristic polynomial $f(x)$ of multiplicity $d_1, \ldots, d_m$ respectively. Let*

$$\widehat{f}(z) = (1-\lambda_1 z)^{d_1}(1-\lambda_2 z)^{d_2} \ldots (1-\lambda_m z)^{d_m}.$$

The generating function of every solution of (R) has the form $g(z)/\widehat{f}(z)$, where $g(z)$ is a polynomial of degree $\leqslant \deg \widehat{f}(z) - 1 = k-1$.

Proof. Let $\lambda \neq 0$ and let d be a positive integer. Any linear combination of functions $(1-\lambda z)^{-b}$ with $1 \leqslant b \leqslant d$ has the form

$$\frac{a_1}{1-\lambda z} + \frac{a_2}{(1-\lambda z)^2} + \ldots + \frac{a_d}{(1-\lambda z)^d},$$

which can be rewritten as

$$\frac{a_1(1-\lambda z)^{d-1} + a_2(1-\lambda z)^{d-2} + \ldots + a_d}{(1-\lambda z)^d} = \frac{h(z)}{(1-\lambda z)^d},$$

where $h(z)$ is a polynomial of degree $\leqslant d-1$. It follows that the generating function of every solution of (R) has the form

$$\frac{h_1(z)}{(1-\lambda_1 z)^{d_1}} + \frac{h_2(z)}{(1-\lambda_2 z)^{d_2}} + \ldots + \frac{h_m(z)}{(1-\lambda_m z)^{d_m}}, \tag{11}$$

where $h_i(z)$ is a polynomial of degree $\leqslant d_i - 1$, $1 \leqslant i \leqslant m$. Clearly, the expression (11) is equal to a fraction $g(z)/\widehat{f}(z)$ as stated. □.

In fact, the converse is also true.

Theorem 6. *A function is a generating function of a solution of (R) if and only if it is equal to a fraction $g(z)/\widehat{f}(z)$, where $\widehat{f}(z)$ is defined in Corollary 3, and $g(z)$ is a polynomial of degree $\leqslant \deg \widehat{f}(z) - 1 = k - 1$.*

Proof. The *"only if"* part of the theorem is Corollary 3. The proof of the *"if"* part is based on a well known algebraic theorem about presentation of rational functions as sums of *partial fractions*. Namely, $g(z)/\widehat{f}(z)$ is equal to an expression of the form (11), with $\deg h_i(z) \leqslant d_i - 1$, $1 \leqslant i \leqslant m$ (the condition $\deg g(z) \leqslant \deg \widehat{f}(z) - 1 = k - 1$ is needed to exclude a potential polynomial summand). See [La], Section IV.5, Theorem 5.2, for example.

Now, let $\lambda \neq 0$ and let d be a positive integer. Let $h(z)$ be a polynomial of degree $\leqslant d-1$. By substituting $z = -\lambda^{-1}(1-\lambda z)+\lambda^{-1}$ into $h(z)$, we can write $h(z)$ as a polynomial in $(1-\lambda z)$ of the same degree as $h(z)$. Therefore,

$$h(z) = a_1(1-\lambda z)^{d-1} + a_2(1-\lambda z)^{d-2} + \ldots + a_d,$$

and

$$\frac{h(z)}{(1-\lambda z)^d} = \frac{a_1}{1-\lambda z} + \frac{a_2}{(1-\lambda z)^2} + \ldots + \frac{a_d}{(1-\lambda z)^d},$$

for some $a_1, a_2, \ldots, a_d$. In fact, this observation is also a part of the theory of partial fractions (see [La], Section IV.5, Theorem 5.3 and the paragraph preceding this theorem, for the general result).

By combining the results of the last two paragraphs, we see that $g(z)/\widehat{f}(z)$ is a linear combination of functions $(1-\lambda_i z)^{-b}$, $1 \leqslant i \leqslant m$, $1 \leqslant b \leqslant d_i$, from Corollary 2. The *"if"* part follows. $\square$.

Remark. $\widehat{f}(z)$ can be expressed directly in terms of $f(x)$. Indeed, since

$$f(x) = (x-\lambda_1)^{d_1}(x-\lambda_2)^{d_2}\ldots(x-\lambda_m)^{d_m},$$

and $k = d_1 + d_2 + \ldots + d_m$, we have

$$\widehat{f}(z) = z^k f(1/z),$$

and, therefore,

$$\widehat{f}(z) = 1 - a_1 z - a_2 z^2 - \ldots - a_k z^k.$$

Theorem 6 can be also proved directly, without using an explicit basis of the space of solutions V.

A generating functions proof of Theorem 6. Let $F(n)$ be a solution of (R). Let

$$G(z) = \sum_{n \geqslant 0} F(n) z^n$$

be the generating function of the sequence $F(n)$. Let us multiply (R) by z^n and rewrite the result as

$$F(n)z^n = a_1 z F(n-1) z^{n-1} + a_2 z^2 F(n-2) z^{n-2} + \ldots + a_k z^k F(n-k) z^{n-k}.$$

Now, let us sum these identities over all $n \geqslant k$:

$$\sum_{n \geqslant k} F(n) z^n = a_1 z \sum_{n \geqslant k} F(n-1) z^{n-1} + a_2 z^2 \sum_{n \geqslant k} F(n-2) z^{n-2} + \ldots$$

$$+ a_k z^k \sum_{n \geqslant k} F(n-k) z^{n-k}.$$

It follows that

$$G(z) - g_k(z) = a_1 z (G(z) - g_{k-1}(z)) + a_2 z^2 (G(z) - g_{k-2}(z)) + \ldots + a_k z^k G(z),$$

where $g_i(z)$ is the sum of the first i terms of $G(z)$,

$$g_i(z) = \sum_{n=0}^{i-1} F(n) z^n.$$

Therefore,

$$G(z) - a_1 z G(t) - a_2 z^2 G(z) - \ldots - a_k z^k G(z) = g(z), \tag{12}$$

where

$$g(z) = g_k(z) - a_1 z g_{k-1}(z) - a_2 z^2 g_{k-2}(z) - \ldots - a_k z^k$$

is a polynomial of degree $\leqslant k-1$ (because $\deg g_i(z) \leqslant i-1$ for all i).

We can rewrite (12) as

$$G(a)(1 - a_1 z - a_2 z^2 - \ldots - a_k z^k) = g(z),$$

i.e. as $G(z)\widehat{f}(z) = g(z)$. It follows that $G(z) = g(z)/\widehat{f}(z)$. This proves that the generating function of every solution of (R) has the form $g(z)/\widehat{f}(z)$.

In order to prove the converse statement, suppose that $G(z) = g(z)/\widehat{f}(z)$ for a polynomial $g(z)$ of degree $\leqslant k-1$. Then $G(z)\widehat{f}(z) = g(z)$, i.e.

$$G(z)(1 - a_1 z - a_2 z^2 - \ldots - a_k z^k) = g(z).$$

By equating the terms of degree $n \geqslant k$ at both sides, we get the relation

$$F(n) - a_1 F(n-1) - a_2 F(n-2) - \ldots - a_k F(n-k) = 0$$

for all $n \geqslant k$ (notice that $g(z)$ has no terms of degree $\geqslant k$ by the assumption). Clearly, this is equivalent to (R). This completes the proof. $\square$

Remarks. This generating function proof of Theorem 6 can be used as a starting point of another approach to the whole theory.

Namely, after we proved that that the generating function of every solution of (R) has the form $g(z)/\widehat{f}(z)$, we can use the theory of the partial fractions as in the first proof of the *"if"* part of Theorem 6. It implies that the generating function of every solution of (R) is a linear combination of functions $(1-\lambda_i z)^{-b}$, $1 \leqslant i \leqslant m$, $1 \leqslant b \leqslant d_i$, from Corollary 2. Using Lemmas 2 and 3, we see that every solution of (R) is a linear combination of solutions $F(n) = (n+1)^{\overline{a}} \lambda_i^n$ from Theorem 5. Since we have exactly k these solutions, where $k = \dim V$ is the dimension of the space V of all solution, they form a basis of V. This gives a new proof of Theorem 5.

Next, since $(n+1)^{\overline{a}}$ is a polynomial in n of degree a, we see that every solution of (R) is a linear combination of sequences $F(n) = n^a \lambda_i^n$ from Theorem 2. This implies that these sequences (1) form a basis of V (because we have exactly $k = \dim V$ sequences). This gives a new proof of Theorem 2.

Now we can deduce Theorem 3 from Theorem 2. Recall that $n^{\underline{a}} = n^a +$ terms of lower degree in n. This implies (by using induction on a, for example) that, vice versa, $n^a = n^{\underline{a}} +$ a linear combination of $n^{\underline{b}}$ with $b < a$ (cf. the proof of Theorem 5). Therefore, solutions $F(n) = n^a \lambda_i^n$ are linear combinations of solutions $F(n) = n^{\underline{a}} \lambda_i^n$. Therefore, the latter solutions span V and, hence, form a basis of V. This gives a new proof of Theorem 3.

Theorem 3 implies that the determinant from Theorem 4 is non-zero, but, of course, is not sufficient to compute its value.

Theorems 2, 3, and 5 should be considered as the main results of the theory. As we saw, they can be easily deduced one from another, using relations between the powers n^a and the falling and rising factorials $n^{\underline{a}}$, $(n+1)^{\overline{a}}$. The generating functions allow us to compress Theorem 5 into a concise form of Theorem 6 (but with a loss of information). They also can be used as a tool. This tool is powerful enough to deduce the main theorems without the computation of the determinant from Theorem 4. However, instead of this computation, the generating functions approach requires another non-trivial algebraic tool: the theory of partial fractions. So, there is a sort of balance between two approaches. Of course, the partial fractions expansion is widely known (although, probably, the relevant proofs are not). Nevertheless, Theorem 4 is beautiful in its own right.

References

[Ge] A.O. Gel'fond, The calculus of finite differences (in Russian), Moskva, Gos. izd-vo tekhniko-teoret. lit-ry, 1952. 479 pp.

[La] S. Lang, Algebra, 3rd rev. ed., Springer, 2002. XV, 914 pp.

Michigan State University, Department of Mathematics, Wells Hall, East Lansing, MI 48824-1027.

ivanov@math.msu.edu