REDUCTION OF THE ORDINARY LINEAR DIFFERENTIAL EQUATION OF THE nTH ORDER WHOSE COEFFICIENTS ARE CERTAIN POLYNOMIALS IN A PARAMETER TO A SYSTEM OF n FIRST-ORDER EQUATIONS WHICH ARE LINEAR IN THE PARAMETER*

BY

CHARLES E. WILDER

The ordinary linear differential equation of the nth order herein considered,

$$\sum_{k=0}^{n} \frac{d^k y}{dx^k} P_k(x,\rho) = 0, \tag{1}$$

in which

$$\begin{aligned} P_k(x,\rho) &= \sum_{j=0}^{n-k} P_{n-k,j}(x)\rho^j, \quad k \neq n, \\ &= 1, \quad k = n, \end{aligned} \tag{2}$$

is a special case of the equation for which Birkhoff developed asymptotic solutions,† but more general than the one for which he solves the expansion problem.‡ In the form given above the equation with various types of boundary conditions has been extensively studied by Tamarkin.§ On the other hand Birkhoff and Langer|| have treated the boundary problem and developments associated with the system of ordinary linear differential equations of the first order

$$\frac{dy_i}{dx} = \sum_{j=1}^{n} A_{ij}(x,\rho) y_j \qquad (i = 1, 2, \cdots, n), \tag{3}$$

---

* Presented to the Society, September 9, 1926; received by the editors in April, 1926.

† Birkhoff, *On the asymptotic character of the solutions of certain linear differential equations containing a parameter*, these Transactions, vol. 9 (1908), pp. 219–231.

‡ Birkhoff, *Boundary value and expansion problems of ordinary linear differential equations*, these Transactions, vol. 9 (1908), pp. 373–395.

§ Tamarkin, *On certain general problems of the theory of ordinary linear differential equations and the expansion of an arbitrary function in series*, Petrograd, 1917 (Russian). Cf. also a paper under the same title which will appear shortly in the Mathematische Zeitschrift.

|| Birkhoff and Langer, *The boundary problems and developments associated with a system of ordinary linear differential equations of the first order*, Proceedings of the American Academy of Arts and Sciences, vol. 58 (1923), pp. 51–128.

in which

$$A_{ij}(x,\rho)=\rho a_{ij}(x)+b_{ij}(x) \qquad (i,j = 1,2, \cdots ,n). \tag{4}$$

It is the object of this paper to study the relation between the equation (1) and the system (3), and in particular to prove the following theorem:*

If the equation (1) *is such that the functions* $P_{i+k,i}(x)$ *possess* $n-k-1$ *continuous derivatives in the interval* $a\leqq x\leqq b$, *for* $i=1, 2, \cdots, n$, *and if the "characteristic equation"*

$$\sum_{t=0}^{n} P_{n-t,n-t}(x)a^t = 0 \tag{5}$$

has roots which are distinct for all values of x in the interval, then the equation (1) *may be reduced to a system of the form* (3), *in which the functions* $a_{ij}(x)$ *and* $b_{ij}(x)$ *are continuous in the same interval.*

In the system (3) set

$$y = \sum_{j=1}^{n} C_{0j}y_j. \tag{6}$$

From the equations obtained by differentiation,

$$y^{(k)} = \sum_{j=1}^{n} C_{kj}y_j \qquad (k = 1,2, \cdots ,n), \tag{7}$$

in which

$$C_{kj} = \sum_{r=1}^{n} C_{k-1,r}A_{rj} + C'_{k-1,j} \qquad (j,k = 1,2, \cdots ,n), \tag{8}$$

the y_i may be eliminated and the nth order equation

$$\begin{vmatrix} y & C_{01} & C_{02} & C_{03} & \cdots & C_{0n} \\ y' & C_{11} & C_{12} & C_{13} & \cdots & C_{1n} \\ y'' & C_{21} & C_{22} & C_{23} & \cdots & C_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ y^{(n)} & C_{n1} & C_{n2} & C_{n3} & \cdots & C_{nn} \end{vmatrix} = 0 \tag{9}$$

---

* The simple substitution $y=y_1$, $y_i'=\rho y_{i+1}$, will reduce the equation (1) to a system of first-order equations, but they will not all be of the form in the parameter required in (3).

obtained. If then the coefficient of $y^{(k)}$ is denoted by P_k, there are between the P's and the C's the relations

$$\sum_{k=0}^{n} P_k C_{kj} \equiv 0 \qquad (j = 1,2, \cdots, n). \tag{10}$$

It is now assumed that in the set (3)

$$\begin{aligned} a_{ij} &= 0, \qquad j>i, \\ b_{i,i+1} &= 1, \end{aligned} \tag{11}$$

and that all the other b_{ij} are zero except when $i=n$. It is to a system (3) of this sort that the equation (1) is reducible. From (11) follows immediately

$$\begin{aligned} A_{ij} &= 0, \quad j > i+1, \\ &= 1, \quad j = i+1, \\ &= \rho a_{ij}, \quad j < i+1,\ i \neq n, \\ &= \rho a_{nj} + b_{nj}, \quad i = n. \end{aligned} \tag{12}$$

Moreover in the substitution (6) it is now assumed that

$$\begin{aligned} C_{0j} &= 1, \quad j = 1, \\ &= 0, \quad j \neq 1. \end{aligned}$$

And it follows that

$$C_{kj} = \begin{cases} 0, j > k+1, \\ 1, j = k+1. \end{cases} \tag{13}$$

This is obviously true for $k=0$, while the recursion formula (8), modified by the restrictions (12) on the A_{ij}, furnishes the material for an easy proof by induction that it holds for any k.

In the expansion of (9) the coefficient of $y^{(n)}$ is unity because of (13), and a slight examination of the form of the C's in ρ, which will incidentally appear in what follows, shows that the equation (9) now actually has the form (1). That is, the system (3) if subject to the restrictions (11) can be transformed to an nth order equation of the form (1).

Under the restrictions (12) and (13) the recursion formula (8) becomes

$$\begin{aligned} C_{kj} &= C_{k-1,j-1} + \rho \sum_{r=j}^{k} C_{k-1,r} a_{rj} + C'_{k-1,j}, \quad k \neq n, \quad j \leqq k, \\ &= C_{n-1,j-1} + \rho \sum_{r=j}^{n} C_{n-1,r} a_{rj} + C'_{n-1,j} + b_{nj}, \quad k = n, \end{aligned}$$

in which of course a C with second subscript zero is identically zero. This is now a formula for $C_{kj}-C_{k-1,\,j-1}$. It can then be rewritten for $C_{k-1,j-1}-C_{k-2,j-2}$ and so on for diminishing values of the subscripts until the second subscript in the second C becomes zero. These formulas are then summed to obtain

$$C_{kj} = \rho \sum_{s=1}^{j} \sum_{r=s}^{k-j+s} C_{k-j+s-1,r}a_{rs} + \sum_{s=1}^{j} C'_{k-j+s-1,s}, \quad k \neq n,$$

in which for $k=n$ the term b_{nj} must be added. This is more convenient in a slightly altered form. Set $k=j+i$ and it becomes

$$(14) \qquad C_{j+i,j} = \rho \sum_{s=1}^{j} \sum_{r=s}^{s+i} C_{s+i-1,r}a_{rs} + \sum_{s=1}^{j} C'_{s+i-1,s}$$
$$(j+i \neq n\,;\; i = 0,1,\cdots,n-j-1),$$

in which for $j+i=n$ the term b_{nj} must be added.

In particular for $i=0$, since by (13) $C_{s-1,s}=1$, $C'_{s-1,s}=0$, this gives

$$(15) \qquad \begin{aligned} C_{jj} &= \rho \sum_{s=1}^{j} a_{ss}, & j \neq n, \\ &= \rho \sum_{s=1}^{n} a_{ss} + b_{nn}, & j = n. \end{aligned}$$

Now $C_{j+i,j}$ is of degree $i+1$ in ρ. This is obviously true for $i=0$, and if it is assumed for $i=k$ it then follows at once from (14) for $i=k+1$. It is also clear that the term independent of ρ is missing except for $j+i=n$. Hence set

$$(16) \qquad \begin{aligned} C_{j+i,j} &= \sum_{k=1}^{i+1} E_{j,i,k}\rho^k, & i+j \neq n, \\ &= \sum_{k=1}^{i+1} E_{j,i,k}\rho^k + b_{nj}, & i+j = n, \end{aligned}$$

and substitute in (14):

$$\begin{aligned} \sum_{k=1}^{i+1} E_{j,i,k}\rho^k &= \rho \sum_{s=1}^{j} \sum_{r=s}^{s+i-1} \sum_{k=1}^{s+i-1} E_{r,s+i-1-r,k}\rho^k a_{rs} \\ &\quad + \sum_{s=1}^{j} \sum_{k=1}^{i} E'_{s,i-1,k}\,\rho^k + \rho \sum_{s=1}^{j} a_{s+i,s} \\ &= \rho \sum_{s=1}^{j} \sum_{k=1}^{i} \sum_{r=s}^{s+i-k} E_{r,s+i-1-r,k}a_{rs}\rho^k \\ &\quad + \sum_{s=1}^{j} \sum_{k=1}^{i} E'_{s,i-1,k}\rho^k + \rho \sum_{s=1}^{j} a_{s+i,s}, \end{aligned}$$

so that the recursion formulas for the E's are

$$(17)\qquad E_{j,i,k} = \sum_{s=1}^{j}\left\{\sum_{r=s}^{s+i-k+1} E_{r,s+i-1-r,k-1}a_{rs} + E'_{s,i-1,k}\right\},\quad k \neq 1,$$

in which $E_{s,i,i+2}=0$,

$$E_{j,i,1} = \sum_{s=1}^{j}\left\{E'_{s,i-1,1} + a_{s+i,s}\right\}.$$

In particular for $k=i+1$ these become

$$E_{j,i,i+1} = \sum_{s=1}^{j}\left\{E_{s,i-1,i}a_{ss} + E'_{s,i-1,i+1}\right\},$$

but since in an E the last subscript can be at most one greater than the second, the last term in the above is zero, so that

$$(17a)\qquad E_{j,i,i+1} = \sum_{s=1}^{j} E_{s,i-1,i}a_{ss}.$$

But this may be rewritten with the first subscript on the E equal to $j-1$ instead of j, and the result combined with the above gives

$$(18)\qquad E_{j,i,i+1} = E_{j-1,i,i+1} + a_{jj}E_{j,i-1,i}.$$

This again may be rewritten for descending values of i, and from the system of equations all E's with first subscript j and second subscript less than i may be eliminated. Thus is obtained the formula

$$(19)\qquad E_{j,i,i+1} = \sum_{k=0}^{i+1} a_{jj}^{k}E_{j-1,i-k,i-k+1}$$

in which $E_{j,-1,0}=1$.

It is now convenient to deduce an identity between these E's which will be of use in what follows. From (19)

$$E_{j,i-1,i} = \sum_{k=0}^{i} a_{jj}^{k}E_{j-1,i-k-1,i-k}.$$

Set $k+1=l$,

$$E_{j,i-1,i} = \sum_{l=1}^{i+1} a_{jj}^{l-1}E_{j-1,i-l,i-l+1},$$

and then multiply this by a_{jj}, obtaining

$$(20)\qquad a_{jj}E_{j,i-1,i} - \sum_{l=1}^{i+1} a_{jj}^{l}E_{j-1,i-l,i-l+1} = 0.$$

Now let a be a parameter, and in (18) set $i=\mu-k$,

$$E_{j,\mu-k,\mu-k+1} - E_{j-1,\mu-k,\mu-k+1} - a_{jj}E_{j,\mu-k-1,\mu-k} = 0,$$

and multiply this by a^k and sum with respect to k:

$$\sum_{k=1}^{\mu} a^k(E_{j,\mu-k,\mu-k+1} - E_{j-1,\mu-k,\mu-k+1}) - a_{jj}\sum_{k=0}^{\mu} a^k E_{j,\mu-k-1,\mu-k} = 0.$$

To this add the negative of (20) with μ substituted for i, and then multiplying the identity

$$E_{j,-1,0} - E_{j-1,-1,0} = 0$$

by $a^{\mu+1}$ add it in also, thus obtaining

$$\sum_{k=1}^{\mu+1} a^k(E_{j,\mu-k,\mu-k+1} - E_{j-1,\mu-k,\mu-k+1}) - a_{jj}\sum_{k=0}^{\mu} a^k E_{j,\mu-k-1,\mu-k} + \sum_{k=1}^{\mu+1} a_{jj}^k E_{j-1,\mu-k,\mu-k+1} = 0.$$

In the second summation set $k'=k+1$; then since k' is a variable of summation the prime may be dropped and the terms rearranged in the form

$$\sum_{k=1}^{\mu+1} \left[(a^k - a_{jj}a^{k-1})E_{j,\mu-k,\mu-k+1} - (a^k - a_{jj}^k)E_{j-1,\mu-k,\mu-k+1}\right] = 0.$$

Separate this into two sums, and in the first replace $k-1$ by k:

$$\sum_{k=0}^{\mu} (a^{k+1} - a_{jj}a^k)E_{j,\mu-k-1,\mu-k} - \sum_{k=1}^{\mu+1} (a^k - a_{jj}^k)E_{j-1,\mu-k,\mu-k+1} = 0\,;$$

then divide this by $a-a_{jj}$ and set $\mu=t-j$, obtaining finally

$$\sum_{k=0}^{t-j} a^k E_{j,t-j-k-1,t-j-k} = \sum_{k=1}^{t-j+1} \left(\frac{a^k - a_{jj}^k}{a - a_{jj}}\right) E_{j-1,t-j-k,t-j-k+1}, \tag{21}$$

which is the identity that will be needed later.

In particular note from (17a) that

$$E_{1,i,i+1} = E_{1,i-1,i}a_{11}\,;$$

but $E_{1,0,1}=a_{11}$, whence it follows at once that

$$E_{1,i-1,i} = a_{11}^i. \tag{22}$$

In order to study the manner in which the a's enter the E's it is convenient to say that $a_{s+q,s}$ is earlier than $a_{t+r,t}$ if $q<r$, and also that $a_{s+q,s}$ is earlier than $a_{r+q,r}$ if $s<r$. Thus the a's are arranged in order.

From the work above it is clear that $E_{j,\sigma,\sigma+1}$ contains of the a's only $a_{11}, a_{22}, \cdots, a_{jj}$. Consider next

$$E_{j,i,i} = \sum_{s=1}^{j} \left\{ \sum_{r=s}^{s+1} E_{r,s+i-r-1,i-1}a_{rs} + E'_{s,i-1,i} \right\}$$

$$(23) \qquad = \sum_{s=1}^{j} \left\{ E_{s,i-1,i-1}a_{ss} + E_{s+1,i-2,i-1}a_{s+1,s} + E'_{s,i-1,i} \right\},$$

in which the derivative term is of the form that contains only the a_{jj}. In particular

$$E_{j,1,1} = \sum_{s=1}^{j} \left\{ E'_{s,0,1} + a_{s+1,s} \right\},$$

and so contains only the a_{ii} and $a_{i+1,i}$ for values of i in the range $1 \leqq i \leqq j$. Hence by (23) $E_{j,i,i}$ contains the a_{ii} and their derivatives and $a_{i+1,i}$ for the same range of values of i.

A similar argument can be applied to show that $E_{j,i,k}$ contains $a_{j+i+1-k,j}$ and earlier a's and their derivatives, and the order of the derivative of $a_{r+s,r}$ occurring is not greater than $i+1-k-s$.

It is now necessary to determine the coefficient of this latest a, namely $a_{j+i+1-k,j}$, in the expression for $E_{j,i,k}$. By reference to (17) it will be seen that this a occurs only in the two terms

$$a_{j+i-k+1,j}E_{j+i-k+1,k-2,k-1} + a_{jj}E_{j,i-1,k-1}.$$

The first E is of the form that contains only the a_{ii}. By use of the same formula the terms in $E_{j,i-1,k-1}$ that contain $a_{j+i-k+1,j}$ may then be found. Each repetition of this procedure lowers the second and last subscripts of the last E by unity, so that finally the coefficient of the latest a, namely $a_{j+i-k+1,j}$, in $E_{j,i,k}$ is

$$(24) \qquad \sum_{l=0}^{k-1} a_{jj}^{\,l} E_{j+i-k+1,k-l-2,k-l-1}.$$

It has already been noted that the differential equation (9) by reason of (13) has coefficients of the form in ρ prescribed for (1). In order to prove the theorem of this paper it is then necessary to show how the a's and b's of the first-order system may be determined in terms of the P's of the nth order equation. This is done by means of the relation (10), which because of (13) may be written

$$(25) \qquad P_{j-1} + \sum_{t=j}^{n} P_t C_{tj} = 0 \qquad (j = 1, 2, \cdots, n).$$

By substitution from (2) and (16) in (25),

$$\sum_{\mu=0}^{n-j+1} P_{n-j+1,\mu}\rho^{\mu} = \sum_{t=j}^{n}\left(\sum_{\mu=0}^{n-t} P_{n-t,\mu}\rho^{\mu}\right)\left(\sum_{k=1}^{t-j+1} E_{j,t-j,k}\rho^{k}\right) + b_{nj} = 0,$$

from which it follows that

$$P_{n-j+1,0} = -b_{nj} \qquad (j = 1,2,\cdots,n),$$

so that the b's are determined.

The coefficient of ρ^{ν} above gives then

$$(26)\qquad P_{n-j+1,\nu} + \sum_{t=j}^{n}\sum_{\mu=0}^{n-t} P_{n-t,\mu}E_{j,t-j,\nu-\mu} = 0 \qquad (j = 1,2,\cdots,n;\ \nu = 1,\cdots,n-j+1).$$

Consider first the equation in which ν has its largest value, namely

$$(27)\qquad P_{n-j+1,n-j+1} + \sum_{t=j}^{n} P_{n-t,n-t}E_{j,t-j,t-j+1} = 0.$$

In particular for $j=1$ this gives

$$P_{nn} + \sum_{t=1}^{n} P_{n-t,n-t}E_{1,t-1,t} = 0.$$

But by (22) this becomes

$$P_{nn} + \sum_{t=1}^{n} P_{n-t,n-t}a_{11}^{t} = 0,$$

which shows that a_{11} is a root of the characteristic equation. Denote by $f_j(a_{jj})$ the expression (27), and by $f_j(a)$ the expression formed from this by substituting a for a_{jj}. By (27) and (19),

$$f_j(a_{jj}) = P_{n-j+1,n-j+1} + \sum_{t=j}^{n} P_{n-t,n-t}\sum_{k=0}^{t-j+1} a_{jj}^{k}E_{j-1,t-j-k,t-j-k+1},$$

in which the E's now occurring do not contain a_{jj}. Hence

$$f_j(a) = P_{n-j+1,n-j+1} + \sum_{t=j}^{n} P_{n-t,n-t}\sum_{k=0}^{t-j+1} a^{k}E_{j-1,t-j-k,t-j-k+1}.$$

and from these two expressions is obtained

$$
(28)\qquad
\begin{aligned}
\frac{f_j(a)-f_j(a_{jj})}{a-a_{jj}} &= \sum_{t=j}^{n} P_{n-t,n-t}\sum_{k=0}^{t-j+1}\frac{a^k-a_{jj}^k}{a-a_{jj}}E_{j-1,t-j-k,t-j-k+1}\\
&= P_{n-j,n-j}+\sum_{t=j+1}^{n} P_{n-t,n-t}\sum_{k=1}^{t-j+1}\left(\frac{a^k-a_{jj}^k}{a-a_{jj}}\right)E_{j-1,t-j-k,t-j-k+1}.
\end{aligned}
$$

On the other hand

$$
f_{j+1}(a_{j+1,j+1}) = P_{n-j,n-j}+\sum_{t=j+1}^{n} P_{n-t,n-t}\sum_{k=0}^{t-j} a_{j+1,j+1}^k E_{j,t-j-k-1,t-j-k},
$$

and so

$$
(29)\qquad f_{j+1}(a) = P_{n-j,n-j}+\sum_{t=j+1}^{n} P_{n-t,n-t}\sum_{k=0}^{t-j} a^k E_{j,t-j-k-1,t-j-k}.
$$

But by (21) the right hand sides of the equations (28) and (29) are equal, so that

$$
\frac{f_j(a)-f_j(a_{jj})}{a-a_{jj}} = f_{j+1}(a),
$$

and it follows at once that the a_{jj} are the roots of the characteristic equation $f_1(a)=0$.

The earliest of the a's are thus determined. The rest of the proof of the theorem then consists in showing how any a may be determined in terms of P's and earlier a's. By reference to (26) it is seen that the latest a is contained in the E's that have the greatest difference between the last two subscripts, that is, in those for which $t-j-\nu+\mu$ is a maximum, which it is for $\mu=n-t$. It is then contained in the terms

$$
\sum_{t=n+1-\nu}^{n} P_{n-t,n-t}\,E_{j,t-j,\nu+t-n},
$$

and by using (24) it is seen that the coefficient of this latest a, namely $a_{n+1-\nu,j}$, is

$$
(30)\qquad \sum_{t=n+1-\nu}^{n} P_{n-t,n-t}\sum_{l=0}^{k-1} a_{jj}^l E_{n+1-\nu,k-l-2,k-l-1}
$$

where $k=\nu+t-n$.

Thus the equation (26) is linear in the latest a occurring in it, with the coefficient of this a given by (30). If then it can be shown that this coefficient is not zero the proof of the theorem will be complete. Now set $n+1-\nu=i$, and (30) becomes

$$\sum_{t=i}^{n} P_{n-t,n-t} \sum_{l=0}^{t-i} a_{jj}^{l} E_{i,\,t-i-l-1,\,t-i-l},$$

which may be written

$$P_{n-i,n-i} + \sum_{t=i+1}^{n} P_{n-t,n-t} \sum_{l=0}^{t-i} a_{jj}^{k} E_{i,\,t-i-l-1,\,t-i-l},$$

and by (29) this is $f_{i+1}(a_{jj})$, and so the coefficient (30) is

$$f_{n-\nu+2}(a_{jj}) = \prod_{i=n-\nu+2}^{n} (a_{ii} - a_{jj}),$$

which cannot be zero for any value of x in the interval (a, b). Thus the proof is complete.

To find the number of derivatives that any $P_{i,j}$ must possess, apply to formula (26) the known facts concerning the manner in which the E's involve the a's. It is thus found that the determination of $a_{j+k,j}$ involves derivatives of $a_{r+s,r}$ of order not greater than $k-s$. Again by referring to (26) it will be seen that $P_{t+s,t}$ is not involved in the determination of any a earlier than $a_{r+s,r}$, so that in the determination of $a_{j+k,j}$ its derivatives will not occur to an order greater than $k-s$. Since the maximum number of derivatives of any $P_{t+s,t}$ will occur in the determination of the latest of the a's, namely a_{n1}, the function $P_{t+s,t}$ will not need to be differentiated more than $n-1-s$ times.

Dartmouth College,
Hanover, N. H.