A Fundamental Theorem on Lambda-Matrices with Applications
— I. Ordinary Differential Equations with Constant Coefficients

Peter Lancaster
Department of Mathematics and Statistics, The University of Calgary
Calgary, Canada

Submitted by Hans Schneider

ABSTRACT

The fundamental theorem of the title refers to a spectral resolution for the inverse of a lambda-matrix \( L(\lambda) = \sum_{i=0}^{n} A_i \lambda^i \) where the \( A_i \) are \( n \times n \) complex matrices and \( \det A_i \neq 0 \). The idea of a Jordan normal form associated with such a lambda-matrix is developed in proving the theorem. Applications are made to the study of initial value problems and two-point boundary value problems for systems of constant-coefficient ordinary differential equations.

1. INTRODUCTION

It may be imagined that everything worth saying about the applications to be made in this paper has already been said. In spite of this, we present herein the development of a matrix formalism which admits a concise analysis and presentation of general solutions, as well as providing a new insight into their structure. For example, the existence of Green's function in the cases we consider is not surprising, but the explicit representation of these functions in spectral form, and in the generality we are able to cope with, seems to be new.

Let \( \mathbb{C}_n \) denote the complex vector space of column \( n \)-vectors of complex numbers, and let \( \mathbb{C}_{m \times n} \) be the space of \( m \times n \) complex matrices. If \( f \) is a continuous function of a real variable with values in \( \mathbb{C}_n \) and \( A_0, A_1, \ldots, A_l \subset \mathbb{C}_{n \times n} \), we consider the differential equation

\[
A_0 u + A_1 u^{(1)} + \cdots + A_l u^{(l)} = f,
\]

where the indices denote differentiation with respect to the independent real
variable \( t \). As is very well known, the general solution of the special case \( u^{(1)} = Au, A \in \mathbb{C}^{n \times n} \), can be written \( u(t) = e^{At}c \), where \( c \in \mathbb{C}^n \) is a vector of \( n \) arbitrary constants. If \( A \) has the Jordan normal form \( J \) and \( A = XJX^{-1} \), then the general solution can be written \( u(t) = Xe^{Jt}d \), where \( d \in \mathbb{C}^n \), and the columns of \( X \) can be interpreted as eigenvectors and generalized eigenvectors of \( A \). Furthermore, a solution of \( u^{(1)} = Au + f \) which is zero at \( t = a \) is given by

\[
u(t) = \int_a^t e^{A(t-\tau)} f(\tau) d\tau
\]
or, what is equivalent, the vector valued Green function for the initial value problem [2, Sec. II.7] is \( e^{A(t-\tau)} = Xe^{J(t-\tau)}X^{-1} \). We are going to show that these results generalize directly to the equation (1).

In order to do this, we develop in Secs. 3 and 4 the fundamental theorem of matrix theory referred to in the title. The line of proof and the conclusions are generalizations of those developed by the author in [7], or in Chapters 2, 3, and 4 of [8], where the results are confined to the case of linear elementary divisors for the eigenvalues. For the solution of the differential equations this meant that solutions with a "polynomial part" were not admitted; all solutions were linear combinations of exponential functions. By removing that restriction we bring the matrix theoretic approach to a more satisfactory conclusion. It should also be remarked that the homogeneous equation considered in the next section has been treated in some detail by Duffin [3].

2. THE HOMOGENEOUS EQUATION

With the differential operator of (1) there is associated the \( \lambda \)-matrix

\[
L(\lambda) = A_0 + A_1 \lambda + A_2 \lambda^2 + \cdots + A_k \lambda^k,
\]
and we assume throughout that \( \det L(\lambda) \neq 0 \). For brevity, we shall sometimes write \( L(D)u = f \) for Eq. (1). A number \( \lambda_i \) is called an eigenvalue of \( L \) if there is a nonzero \( x \in \mathbb{C}^n \) for which \( L(\lambda_i)x = 0 \), and \( x \) is called an eigenvector corresponding to \( \lambda_i \). The vectors \( x_1, x_2, \ldots, x_k \) form a Jordan chain (Keldys chain in the Russian literature) corresponding to an eigenvalue \( \lambda_i \) if \( x_1 \neq 0 \) and

\[
\sum_{p=0}^{j} \frac{1}{p!} L^{(p)}(\lambda_i)x_{j+1-p} = 0, \quad j = 0, 1, \ldots, k-1.
\]
In the case \( L(\lambda) = -\lambda I + A \) (associated with \( u^{(1)} = Au \)) the definition yields

\[
(A - \lambda I)x_1 = 0, \quad (A - \lambda I)x_j = x_{j-1}, \quad j = 2, \ldots, k - 1,
\]
so that the Jordan chain is just the more familiar set of generalized eigenvectors used to span a cyclic invariant subspace of \( A \) [4, p. 184] and, incidentally, to construct the matrix defining the reduction of \( A \) to Jordan normal form by a similarity transformation.

For the present analysis, the significance of the Jordan chain is that there correspond \( k \) linearly independent solutions of the homogeneous problem \( L(D)u = 0 \) to a chain of length \( k \), namely, the functions

\[
U_{i+1}(t) = e^{\lambda t} \left( \sum_{p=0}^{i} \frac{1}{p!} t^p x_{i+1-p} \right), \quad i = 0, 1, \ldots, k - 1
\]
(see [9] and p. 324 of [5]). If we now define the matrix \( X_i \in \mathbb{C}^{n \times k} \) with columns \( x_1, x_2, \ldots, x_k \) and define \( U_i(t) \) similarly to have columns \( u_1(t), \ldots, u_k(t) \), then it is important for our purposes to note that

\[
U_i(t) = X_i \left[
\begin{array}{cccc}
1 & t & \frac{t^2}{2} & \cdots & \frac{1}{(k-1)!} & t^{k-1} \\
0 & 1 & t & \cdots & \frac{1}{(k-2)!} & t^{k-2} \\
\vdots & \vdots & \ddots & \ddots & \vdots \\
0 & & & t & \\
0 & & & \cdots & 1 \\
\end{array}
\right] e^{\lambda t} = X_i e^{\lambda t},
\]
(3)

where \( I_i \) is the "Jordan block" \( I_i = \lambda I + H, \quad H = [\delta_{i+1, j}] \). Thus, the matrix \( U_i(t) \) may be viewed as an \( n \times k \) matrix solution of the differential equation. Furthermore, the algebraic relation corresponding to \( L(D) U_i(t) = 0 \) is

\[
A_0 X_i + A_1 X_i I_i + \cdots + A_k X_i I_i^{k-1} = 0.
\]
(4)

Now it can be shown (as in [9]) that to every elementary divisor \( (\lambda - \lambda_i)^k \) of \( L(\lambda) \) corresponds a Jordan chain of length \( k \). Furthermore, the union of all sets of solutions \( \{u_i(t)\}_{i=1}^k \), one set for each elementary divisor of \( L(\lambda) \), forms a linearly independent set which spans the solution space of \( L(D)u = 0 \). It
follows that the dimension of the solution space is just the degree of the polynomial $\det L(\lambda)$, say $\nu$. (This is known as Chrystal's theorem [9]. For elementary divisor theory see p. 142 of [4].) If there are $s$ elementary divisors of $L(\lambda)$ in all and we define

$$U(t) = \begin{bmatrix} U_1(t), \ldots, U_s(t) \end{bmatrix} \in C_{n \times r} \quad \text{for each } t,$$

$$X = \begin{bmatrix} X_1, \ldots, X_s \end{bmatrix} \in C_{n \times r}$$

and the block diagonal matrix of order $\nu \times \nu$,

$$J = \text{diag}\{J_1, \ldots, J_s\},$$

then

$$U(t) = X e^{Jt} = \begin{bmatrix} X_1 e^{J_1 t}, \ldots, X_s e^{J_s t} \end{bmatrix},$$

and $L(D)U(t) = 0$ implies

$$A_0 X + A_1 X J + \cdots + A_s X J^s = 0.$$ 

(7)

Note that $X$ is, in general, a rectangular matrix whose columns are members of Jordan chains for $L$. The matrix $J$ is to be seen as a Jordan matrix associated with $L$. We can now state

**Theorem 1.** Every solution of $L(D)u = 0$ can be written in the form $u(t) = X e^{Jt}c$ for some $c \in C_r$, where $\nu$ is the degree of $\det L(\lambda)$.

Consider now the initial value problem: to find a solution of $L(D)u = 0$ for which $u^{(r)}(a) = u_r, r = 0, 1, \ldots, l - 1$. Using the theorem, we have $u^{(r)}(a) = XJ^r e^{Ja}c$, so that an algebraic system

$$\begin{bmatrix} X \\ XJ \\ \vdots \\ XJ^{l-1} \end{bmatrix} e^{Ja}c = \begin{bmatrix} u_0 \\ u_1 \\ \vdots \\ u_{l-1} \end{bmatrix}$$

is to be solved and the matrix on the left is $ln \times \nu$. Thus the initial value problem can have a unique solution for any choice of initial values only if
Fundamental Theorem on Lambda-Matrices—I

\( v = \ln \). Since we propose to formulate a general solution of the initial value problem for \( L(D)u = f \), we now impose the condition that \( A_i \) must be nonsingular, in which case \( L(\lambda) \) is said to be a regular \( \lambda \)-matrix. In this case it is clear that \( v = \ln \) and it will transpire that (8) has a unique solution \( c \) for any initial conditions.

3. Preliminaries on Linear \( \lambda \)-Matrices

Let \( A, B \in \mathbb{C}^{p \times p} \) with \( B \) nonsingular, and consider the \( \lambda \)-matrix \( A + B\lambda \). Let \( J \) be the Jordan normal form for \( -B^{-1}A \), and suppose \( -B^{-1}A = QJQ^{-1} \). Then

\[
AQ + BQJ = 0. \tag{9}
\]

If \( I\lambda + B^{-1}A \) has \( s \) elementary divisors, then they coincide with those of \( A + B\lambda \), and we can write \( J \) in terms of Jordan blocks: \( J = \text{diag}\{J_1, \ldots, J_s\} \). Partition \( Q \) accordingly: \( Q = [Q_1, \ldots, Q_s] \), and we have \( AQ_i + BQ_iJ_i = 0 \), \( i = 1, 2, \ldots, s \). Now the relations (2) defining Jordan chains yield for \( A + B\lambda \)

\[
(A + B\lambda_i)q_1 = 0,
\]

\[
Aq_{i+1} + B(q_i + \lambda_i q_{i+1}) = 0, \quad j = 1, 2, \ldots, k - 1,
\]

which can be abbreviated to \( AQ_i + BQ_iJ_i = 0 \), where \( Q_i = [q_1, \ldots, q_k] \). Thus, the columns of the matrix \( Q \) in (9) may be identified with Jordan chains of \( A + B\lambda \) and the matrix \( Q \) is nonsingular.

Let \( A^T \) denote the transpose of any matrix \( A \), and define a nonsingular matrix \( R \in \mathbb{C}^{p \times p} \) by \( R^T = (BQ)^{-1} \). Then it follows from (9) that

\[
R^TA + JR^TB = 0. \tag{10}
\]

We claim that the columns of \( R \) can now be identified with the Jordan chains of the transposed \( \lambda \)-matrix, \( A^T + B^T\lambda \). To see this, observe that \( A^TR + B^TRJ^T = 0 \), and follow the same argument as applied to (9). The only difference arises from the presence of \( J^T \) in place of \( J \). This means that, if \( r_1, \ldots, r_k \) is a Jordan chain of \( A^T + B^T\lambda \) corresponding to the \( i \)th elementary divisor, we put \( R_i = [r_k, r_{k-1}, \ldots, r_1] \) and then \( R = [R_1, \ldots, R_s] \).
The constructions developed in this section now yield the following vital lemma.

**Lemma 1.** If $A, B \in C_{p \times p}$ with $B$ nonsingular, then there exist nonsingular matrices $Q, R$ whose columns are made up of Jordan chains of $A + B\lambda$ and $A^T + B^T\lambda$, respectively, and for which

$$R^T B Q = I \quad \text{and} \quad R^T A Q = -I,$$

where $J$ is a Jordan matrix associated with $A + B\lambda$. Furthermore, if $\lambda$ is not an eigenvalue, then

$$(A + B\lambda)^{-1} = Q (I\lambda - J)^{-1} R^T. \quad (12)$$

**Proof.** The existence of $Q$ and $R$ and the first of the relations (11) follow from the above construction and, in particular, the definition of $R$. The second relation follows on premultiplying (9) by $R^T$. From the relations (11) we deduce that $R^T (A + B\lambda) Q = I\lambda - J$, and (12) follows immediately.

Note that for a regular $\lambda$-matrix $A + B\lambda$ one is assured of the existence of *linearly independent* chains for each elementary divisor. This is not the case for $\lambda$-matrices of higher degree. Even though the columns of $U_i$ on the left of (3) are linearly independent functions of $t$, the columns of $X_i$ may well be dependent. Examples can be found in [9].

4. **THE FUNDAMENTAL THEOREM**

In the next section we shall formulate a Green's function for the initial value problem associated with $L(D)u = f$, and to do this, we shall need analogues of (11) and (12) for the regular $\lambda$-matrix problem of general degree. We obtain these results by analysis of a linear problem which is equivalent in the sense that it has the same associated Jordan matrix. This linearization can be achieved in several different ways. For our purposes we find a significant advantage in that followed by the author in [1] and [8]. (The nearest competitor seems to be that used by Baumgärtel [1, p. 325] and by Gohberg and Krein [5, p. 267].)
FUNDAMENTAL THEOREM ON LAMBDA-MATRICES—I

We define $n \times n$ matrices $A$ and $B$ by

\[
A = \begin{bmatrix}
A_0 & 0 & 0 & \cdots & 0 & 0 \\
0 & -A_2 & -A_3 & \cdots & -A_{l-1} & -A_l \\
0 & -A_3 & -A_4 & \cdots & -A_l & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & -A_{l-1} & -A_l & \cdots & 0 & 0 \\
0 & -A_l & 0 & \cdots & 0 & 0 \\
\end{bmatrix}
\]

\[
B = \begin{bmatrix}
A_1 & A_2 & \cdots & A_{l-1} & A_l \\
A_2 & A_3 & \cdots & A_l & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
A_{l-1} & A_l & \cdots & 0 & 0 \\
A_l & 0 & \cdots & 0 & 0 \\
\end{bmatrix}, \tag{13}
\]

If $z \in \mathbb{C}^n$, we will use component $z_i \in \mathbb{C}$, $i = 1, 2, \ldots, n$, for which $z^T = [z_1^T, \ldots, z_n^T]$. Then, using the fact that $A_l$ is nonsingular, it is easily verified that $\lambda_j$ is an eigenvalue of $L(\lambda)$ iff $\lambda_j$ is an eigenvalue of $A + B\lambda$. Furthermore, if $z$ is an eigenvector of $A + B\lambda$ corresponding to $\lambda_j$, then for $j = 1, 2, \ldots, l$, we have $z_j = \lambda_j^{j-1} z_1$, and $z_1$ is an eigenvector of $L(\lambda)$ corresponding to $\lambda_j$.

Let $Q_i \in \mathbb{C}^n \times k$ be built up from a Jordan chain of $A + B\lambda$ as described in Sec. 3; partition $Q_i$ into $n \times k$ blocks,

\[
Q_i = \begin{bmatrix}
X_{i,1} \\
\vdots \\
X_{i,k}
\end{bmatrix},
\]

and let $J_i$ be the associated Jordan block. Substituting for $Q_i$, $A$, and $B$ in the relation $AQ_i + BQ_i J_i = 0$, it is found that $X_{i,j} = X_{i,1} J_i^{j-1}$ for $j = 1, 2, \ldots, l$, and

\[
A_0 X_{i,1} + A_1 X_{i,1} J_i + \cdots + A_l X_{i,1} J_i^{l-1} = 0.
\]
It follows that

\[
Q_i = \begin{bmatrix}
X_{i,1} \\
X_{i,1}J_i \\
\vdots \\
X_{i,1}J_i^{l-1}
\end{bmatrix},
\]

and, as in (4), we find that the columns of \(X_{i,1}\) form a Jordan chain of \(L(\lambda)\). The construction can be reversed to show that to a Jordan chain of \(L(\lambda)\) there corresponds one of \(A + B\lambda\). It also follows that the elementary divisors of \(L(\lambda)\) and \(A + B\lambda\) are the same. If there are \(s\) such divisors, we obtain the Jordan matrix (6) for both \(L(\lambda)\) and \(A + B\lambda\), and if \(X = [X_{1,1}, X_{2,1}, \ldots, X_{s,1}] \in C_{n \times ln}\) is the matrix built up of the \(s\) Jordan chains for \(L(\lambda)\), then

\[
Q = [Q_1, \ldots, Q_s] = \begin{bmatrix}
X \\
XJ \\
\vdots \\
XJ^{l-1}
\end{bmatrix}
\]

is the nonsingular \(ln \times ln\) matrix of Jordan chains for \(A + B\lambda\) as constructed in Sec. 3. Incidentally, it is now clear that the Jordan matrix associated with \(L(\lambda)\) is just a Jordan normal form for \(-B^{-1}A\), and it is easily verified that \(B^{-1}A\) is the "block companion" matrix

\[
\begin{bmatrix}
0 & -I & 0 & \cdots & 0 \\
0 & 0 & -I & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & -I \\
A_l^{-1}A_0 & A_l^{-1}A_1 & A_l^{-1}A_2 & \cdots & A_l^{-1}A_{l-1}
\end{bmatrix}.
\]

As in Sec. 3, Jordan chains for \(A^T + B^T\lambda\) are determined by columns of \(R\), where \(R^T = (BQ)^{-1}\). If we write \(R^T = [Y_1^T, \ldots, Y_s^T]\), the partitions being determined by the degrees of the \(s\) elementary divisors [as in (14)], and substitute for \(R\), \(A^T\), and \(B^T\) [from (13)] in \(A^TR + B^TR^T = 0\) [Eq. (10)], then the argument used above to bring out the structure of \(Q\) can be applied to investigate the structure of \(R\). It is found that \(Y_i = Y_1(J^T)^{i-1}\) for \(j=1,2,\ldots,l\),
and

\[ A_0^T Y_1 + A_1^T Y_1 J^T + \cdots + A_l^T Y_1 (J^T)^l = 0. \] (15)

Thus, the partition \( Y_1 \) of \( R \), which we subsequently call \( Y \in C_{n \times bn} \), is made up of Jordan chains of \( L^T(\lambda) \), and we have

\[
R = [R_1, \ldots, R_s] = \begin{bmatrix} Y \\ YJ^T \\ \vdots \\ Y(J^T)^{l-1} \end{bmatrix}.
\] (16)

We now reexamine the first of the conditions (11), from which, as we have seen, the second condition and Eq. (12) follow very easily. Using (13), (14), and (16) we have

\[
\begin{bmatrix} Y^T & JY^T & \cdots & J^{l-1}Y^T \end{bmatrix} \begin{bmatrix} A_1 & A_2 & \cdots & A_l \\ A_2 & A_3 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ A_l & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} X \\ XJ \\ \vdots \\ X(J^T)^{l-1} \end{bmatrix} = I,
\]

which can be written

\[
\sum_{k=1}^l \sum_{s=1}^k J^{k-s} (Y^T A_k X) J^{s-1} = I.
\] (17)

Now with any unilateral function \( P: C_{n \times n} \to C_{n \times n} \) defined by \( P(U) = \sum_{i=0}^l P_i U^i \) we can associate a bilateral function

\[
\mathcal{P}(U) = P_0 + \frac{1}{2} (P_1 U + UP_1) + \frac{1}{3} (P_2 U^2 + UP_2 U + U^2 P_2) + \cdots 
= \sum_{k=1}^{l+1} \sum_{s=1}^k k^{-1} U^{k-s} P_{k-1} U^{s-1},
\]
and the Frechet derivative of $g^p$ at $I$ in the direction of $J$ is

$$
g^{(1)}(J) = \sum_{k=1}^{l} \sum_{s=1}^{k} J^{k-s} P_{s} J^{s-1}.
$$

This suggests that (17) can be phrased in terms of the Frechet derivative of $\ell$ where

$$
\ell(U) = \sum_{k=1}^{l+1} \sum_{s=1}^{k} k^{-1} U^{k-s} (Y A_{k-1} X) U^{s-1}.
$$

Indeed, (17) becomes simply

$$
\ell^{(1)}(J) = I.
$$

This is an $ln \times ln$ matrix equation, and it may be more convenient in practice to see the implications for separate elementary divisors. Thus, suppose elementary divisors $i$ and $j$ have degrees $k_{1}$ and $k_{2}$ respectively, and we define $\ell_{ij}: C_{k_{1} \times k_{1}} \times C_{k_{2} \times k_{2}} \rightarrow C_{k_{2} \times k_{1}}$ by

$$
\ell_{ij}(U, V) = \sum_{k=1}^{l+1} \sum_{s=1}^{k} k^{-1} V^{k-s} (Y A_{k-1} X) U^{s-1}.
$$

Then (17), or (18), can be written

$$
\ell_{ij}^{(1)}(J_{i}, J_{j}) = \begin{cases} 0 & \text{if } i \neq j, \\ I & \text{if } i = j \end{cases}
$$

for $i, j = 1, 2, \ldots, s$, where the zero is a $k_{2} \times k_{1}$ matrix and the identity is $k_{1} \times k_{1}$.

These expressions simplify if the elementary divisors are linear, for in this case $X_{i}$ and $Y_{j}$ are column vectors and $J_{i}, J_{j}$ are $1 \times 1$ matrices. In particular, if $i, j$ refer to linear elementary divisors associated with the same eigenvalue, then the biorthogonality and normalization conditions (19) become

$$
y_{i}^{T} L^{(1)}(\lambda_{i}) x_{1} = \delta_{ij}
$$

(note that $x_{1}, y_{1}$ are still associated with the $i$th and $j$th divisors, respectively),
where $L^{(1)}$ denotes the derivative of $L$ with respect to $\lambda$, as asserted in Theorem 4.5 of [8].

We are now in a position to state our central algebraic result giving a spectral resolution for the matrix $L(\lambda)^{-1}$.

**Theorem 2.** Let $L(\lambda)$ be a regular $n \times n \lambda$-matrix of degree 1 with $s$ elementary divisors, let $J \in \mathbb{C}^{n \times n}$ be the associated Jordan matrix, and let $X, Y$ be $n \times \ln$ matrices of Jordan chains. Then $X, Y$ can be defined in such a way that (17) [or (18), or (19)] is true, in which case,

$$\lambda^{-1}L(\lambda)^{-1} = XJ^{-1}(I\lambda - J)^{-1}Y^T, \quad r = 1, 2, \ldots, l,$$

and

$$\lambda^r L(\lambda)^{-1} = XJ^r (I\lambda - J)^{-1}Y^T + A_1^{-1}. \quad (21)$$

**Remarks.** (1) To see the nature of this spectral resolution, observe that if $X, Y, J$ are partitioned according to elementary divisors, as in (5) and (6), then (20) may be written

$$\lambda^{-1}L(\lambda)^{-1} = \sum_{i=1}^{s} X_i J_i^{-1} (I\lambda - J_i)^{-1}Y_i^T, \quad r = 1, 2, \ldots, l.$$

One may deduce from this that $L(\lambda)^{-1}$ has a pole at each eigenvalue $\lambda_i$ whose order is the maximal degree of the elementary divisors associated with $\lambda_i$. If all elementary divisors have degree one (the case considered in [8]) then the $X_i, Y_i$ reduce to single column vectors (eigenvectors) and $J_i = [\lambda_i]$. Thus

$$\lambda^{-1}L(\lambda)^{-1} = \sum_{i=1}^{s} \frac{\lambda_i^{r-1}}{\lambda - \lambda_i} (x_i y_i^T), \quad r = 1, 2, \ldots, l.$$

(2) The precise nature of the poles of $L(\lambda)^{-1}$ and coefficients in the principal part of the Laurent expansion about an eigenvalue are often of interest. All this information is contained in our expansion. It has also been developed in [9], but in that case the columns of $X$ and $Y$ are constructed by examination of the Smith normal form for $L(\lambda)$. In that work the hypothesis that $L(\lambda)$ is regular was not required.
Proof of the theorem. We apply our lemma to the $\lambda$-matrix $A + B\lambda$ with $A, B$ defined by (13). Suppose that $C = (A + B\lambda)^{-1}$ has $n \times n$ partitions $C_{ij}$ for $i, j = 1, 2, \ldots, l$. Then examining the first column of the product $(A + B\lambda)C = I$ yields

$$C_{r1} = \lambda^{-1}C_{11}, \quad r = 1, 2, \ldots, l,$$

(22)

and

$$(A_0 + A_1\lambda)C_{11} + A_2C_{21}\lambda + \cdots + A_lC_{l1}\lambda = I,$$

whence $L(\lambda)C_{11} = I$ and, using (22) again,

$$\lambda^{-1}L(\lambda)^{-1} = C_{r1}, \quad r = 1, 2, \ldots, l.$$

Using the lemma we also have $C = (A + B\lambda)^{-1} = Q(I\lambda - J)^{-1}R^T$. Substitute from (14) for $Q$ and (16) for $R$ and evaluate the element $C_{r1}$; then (20) is obtained.

Now write

$$(A_0 + A_1\lambda + \cdots + A_l\lambda^l)L(\lambda)^{-1} = I$$

and apply (20) for each $r$ to obtain

$$(A_0X + A_1XJ + \cdots + A_lXJ^l)(I\lambda - J)^{-1}Y^T$$

$$- A_lXJ^l(I\lambda - J)^{-1}Y^T + \lambda^lA_lL(\lambda)^{-1} = I;$$

then using (7) we obtain (21).

\[\square\]

Corollary. The matrices $X, Y,$ and $J$ of Theorem 2 satisfy

$$XJ^TY^T = \begin{cases} 0, & r = 0, 1, \ldots, l - 2, \\ A_l^{-1}, & r = l - 1. \end{cases}$$

Proof. Put $L(\lambda)^{-1} = X(I\lambda - J)^{-1}Y^T$ in (21) and deduce that

$$X(I\lambda - J^l)(I\lambda - J)^{-1}Y^T = A_l^{-1}.$$
from which it follows that
\[ X(I\lambda^{l-1} + J\lambda^{l-2} + \cdots + J^{l-1})Y^T = A_l^{-1} \]
and hence the conclusion. \(\blacksquare\)

4. THE NONHOMOGENEOUS EQUATION

We first make a natural extension of the theory for one scalar equation (Birkhoff and Rota, [2], p. 89, for example) and claim the following result:

**Theorem 3.** If the function \(G(t, \tau)\) satisfies \(G(t, \tau) = 0\) when \(a < t < \tau\), \(L(G) = 0\) (for the independent variable \(t\)) when \(t > \tau \geq a\), and
\[
G^{(r)}|_{t=\tau} = \begin{cases} 
0, & r = 0, 1, \ldots, l-2, \\
A_l^{-1}, & r = l-1 
\end{cases}
\]
(the derivatives being with respect to \(t\)), then \(G\) is the Green's function for the initial value problem \(L(D)u = f\), \(u^{(r)}(a) = 0\) for \(r = 0, 1, \ldots, l-1\).

It is a simple matter of verification to see that, once \(G\) is known, \(L(D)u = f\) has a solution
\[
u(t) = \int_a^t G(t, \tau) f(\tau) d\tau
\]
which, together with \(u^{(r)}(a) = 0, r = 0, 1, \ldots, l-1\), is the defining property of \(G\).

**Theorem 4.** The Green's function for the initial value problem of Eq. (1) is given by \(G(t, \tau) = Xe^{J(t-\tau)}Y^T\) for \(t > \tau\), where \(X, Y\) are the matrices of Jordan chains of \(L(\lambda), L(\lambda)^T\) used in Theorem 2, and \(J\) is the corresponding Jordan matrix.

**Proof.** Differentiating with respect to \(t\) we have
\[
L(G) = (A_0X + A_1XJ + \cdots + A_lXJ^l)e^{J(t-\tau)}Y^T = 0,
\]
using Eq. (7). Then it is found that
\[ G^{(r)}|_{t=\tau} = XJ^r Y^T, \]
and the required initial conditions follow from the Corollary to Theorem 2.

Combining Theorems 1 and 4, it follows that when \( L(\lambda) \) is regular, a general solution to (1) can be written
\[ u(t) = X e^{Jc} + X \int_{a}^{t} e^{J(t-\tau)} Y^T f(\tau) d\tau, \quad (23) \]
where \( c \in C_{ln} \) and can be determined by imposing the usual initial conditions of standard type, for example.

5. **EXPLICIT SOLUTION OF THE INITIAL VALUE PROBLEM**

We have seen that if \( L(\lambda) \) is a regular \( \lambda \)-matrix, then the general solution of \( L(D)u = 0 \) is \( u(t) = X e^{Jc} \) for some \( ln \)-vector \( c \). We have also noted that in the case of the standard initial value problem, there is a unique solution for \( c \) in terms of the initial vectors \( u_0, u_1, \ldots, u_{l-1} \) [the matrix \( Q \) in (8) and (14) is nonsingular]. We are now to solve explicitly for \( c \) in terms of the \( u_i \). We treat the homogeneous case, but in view of (23), this also yields the solution to the nonhomogeneous problem.

**Theorem 5.** If \( L(\lambda) \) is a regular \( \lambda \)-matrix, there is a unique solution of \( L(D)u = 0 \) for which \( u^{(r)}(a) = u_r \), for \( r = 0, 1, \ldots, l-1 \), and this solution is given by putting
\[ c = \begin{bmatrix} \quad Y^T & JY^T & \cdots & J^{l-1}Y^T \end{bmatrix} \begin{bmatrix} A_1 & A_2 & \cdots & A_l \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \\ \vdots \\ u_{l-1} \end{bmatrix} \quad (24) \]
in the solution of Theorem 1.

**Proof.** The existence of a unique solution has already been established.
We verify that (24) yields the appropriate solution. The vector $c$ must satisfy

$$
\begin{bmatrix}
  u_0 \\
  u_1 \\
  \vdots \\
  u_{l-1}
\end{bmatrix}
= 
\begin{bmatrix}
  X \\
  XJ \\
  \vdots \\
  XJ^{l-1}
\end{bmatrix} 
Q c.
$$

But, for a regular $\lambda$-matrix, $Q^{-1}$ exists and is given by Eq. (11) in the form $Q^{-1} = R^T B$. The result is then obtained from (13) and (16).

Equation (24) can also be written in the form

$$
c = \sum_{k=1}^{l} J^{k-1} Y^T \sum_{r=1}^{l-k+1} A_{r+k-1} u_{r-1},
$$

so that

$$
u(t) = X e^{J(t-a)} \sum_{k=1}^{l} J^{k-1} Y^T \sum_{r=1}^{l-k+1} A_{r+k-1} u_{r-1} = X e^{J(t-a)} c,
$$

or, examining the contributions of the $s$ elementary divisors to the solution,

$$
u(t) = \sum_{i=1}^{s} X_i e^{h_i(t-a)} c_i,
$$

where

$$
c_i = \sum_{k=1}^{l} J^{k-1} Y_i^T \sum_{r=1}^{l-k+1} A_{r+k-1} u_{r-1}.
$$

(25)

Note that if the $i$th divisor has degree $k$, then $c_i \in C_k$.

The columns of $U_i(t) = X_i e^{h_i(t-a)}$ are fundamental solutions of Eq. (2), and the elements of the vectors $c_i$ give the coefficients of that linear combination of these fundamental solutions having the prescribed initial values. Indeed, the relations

$$
u_r = \sum_{i=1}^{s} X_i J^r c_i, \quad r = 0, 1, \ldots, l-1,
$$
together with (25) show explicitly how each of the initial vectors is expressed as a linear combination of the $ln$ vectors in the union of the $s$ Jordan chains. The existence of these expansions is described by Keldys [6] in a more general context, as the $l$-fold completeness of the generalized eigenvectors.

In the case that all elementary divisors of $L(\lambda)$ are linear, the above expansions take the relatively simple forms

$$u(t) = \sum_{i=1}^{ln} c_i x_i e^{\lambda_i t}, \quad u_r = \sum_{i=1}^{ln} \lambda_i^r c_i x_i, \quad r = 0, 1, \ldots, l-1.$$ 

The coefficients $c_i$ being given by

$$c_i = \sum_{i=1}^{l} \lambda_i^{k-1} \left( y_i^T \sum_{r=1}^{l-k+1} A_{r+k-1} u_{r-1} \right).$$

**Example.** We consider the (otherwise trivial) problem in which $l=2$,

$$A_0 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad A_1 = 0, \quad \text{and} \quad A_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

—in other words, the pair of simultaneous equations $x_1^{(2)} + x_2 = 0$, $x_2^{(2)} = 0$. In this case

$$L(\lambda) = \begin{bmatrix} \lambda^2 & 1 \\ 0 & \lambda^2 \end{bmatrix},$$

and zero is the only eigenvalue with only one elementary divisor and that has degree four. A Jordan chain for this divisor is $x_1, x_2, x_3, x_4$, where

$$x_1 = x_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad x_3 = x_4 = \begin{bmatrix} 0 \\ -1 \end{bmatrix}.$$ 

A Jordan chain for $L(\lambda)^T$ is equally easy to determine, but care has to be taken to see that one is chosen so that (17) is satisfied. It is found that

$$y_1 = \begin{bmatrix} 0 \\ -1 \end{bmatrix}, \quad y_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad y_3 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad y_4 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$
is appropriate. Then we have

\[ X = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & -1 & -1 \end{bmatrix}, \quad Y = \begin{bmatrix} -1 & 1 & 0 & 0 \\ 1 & -1 & 1 & -1 \end{bmatrix}. \]

The Green’s function for the initial value problem is then determined by

\[ G(t) = \begin{bmatrix} t & -\frac{1}{6}t^3 \\ 0 & t \end{bmatrix}, \]

and the vector \( c \) of (25) which determines the solution in terms of initial vectors \( u_0, u_1 \in \mathbb{C}_2 \) is given by

\[
c = Y^T (A_1u_0 + A_2u_1) + JY^T (A_2u_0) = \begin{bmatrix} -1 & 1 \\ 1 & -1 \\ 0 & 1 \\ 0 & -1 \end{bmatrix} u_1 + \begin{bmatrix} 1 & -1 \\ 0 & 1 \\ 0 & -1 \\ 0 & 0 \end{bmatrix} u_0.
\]

6. TWO POINT BOUNDARY VALUE PROBLEMS

To facilitate the study of two point boundary value problems we begin with the construction of what might be called the pre-Green’s function. First, let \( J_1 \) and \( J_2 \) be Jordan matrices with the property that \( J = J_1 \oplus J_2 \) is a Jordan matrix associated with the regular \( \lambda \)-matrix \( L(\lambda) \). Thus, each elementary divisor of \( L(\lambda) \) is associated with either a block of \( J_1 \), or \( J_2 \), but not with both. Let \( J_1 \) be \( p \times p \) and \( J_2 \) be \( q \times q \), so that \( p + q = \ln \). Let \( X, Y \) be the \( n \times \ln \) matrices of Jordan chains as introduced earlier and defined to satisfy (17). Then there are blocks \( X_1, X_2 \) of \( X \) and \( Y_1, Y_2 \) of \( Y \) compatible with the partition of \( J \). With these understandings we have

**Lemma 2.** The matrix-valued function defined on \([a, b] \times [a, b]\) by

\[
G_0(t, \tau) = \begin{cases} 
-X_1e^{J_1(t-\tau)}Y_1^T & \text{if } a < t < \tau, \\
X_2e^{J_2(t-\tau)}Y_2^T & \text{if } \tau < t < b 
\end{cases}
\] (26)
satisfies the following conditions:

(a) Differentiating with respect to \( t \), \( L(D)G_0 = 0 \), for \( (t, \tau) \in [a, b] \times [a, b] \), as long as \( t \neq \tau \).

(b) Differentiating with respect to \( t \),

\[
G_0^{(r)}|_{t=\tau^+} - G_0^{(r)}|_{t=\tau^-} = \begin{cases} 
0 & \text{for } r=0,1,\ldots,l-2, \\
A_t^{-1} & \text{for } r=l-1.
\end{cases}
\]

(c) The function

\[
u(t) = \int_a^b G_0(t, \tau) f(\tau) d\tau
\]

is a solution of \( L(D)u = f \).

Proof.

(a) We have

\[
L(D)G_0 = \begin{cases} 
- \left( \sum_{r=0}^{l} A_r X_1 J_1^r \right) e^{J_1^l(t-\tau)} Y_1^T & \text{if } a < t < \tau, \\
\left( \sum_{r=0}^{l} A_r X_2 J_2^r \right) e^{J_2^{l}(t-\tau)} Y_2^T & \text{if } \tau < t < b.
\end{cases}
\]

The relation (4) implies that the summation applied to each elementary divisor associated with \( J_1 \) or \( J_2 \) is zero, and it follows that \( L(D)G_0 = 0 \) if \( t \neq \tau \).

(b) We have

\[
G_0^{(r)}|_{t=\tau^+} - G_0^{(r)}|_{t=\tau^-} = X_2 J_2^l Y_2^T + X_1 J_1^l Y_1^T = XJY^T,
\]

and the conclusion follows from the Corollary to Theorem 2.

(c) This part is proved by verification. For brevity we treat the case \( l=2 \). The same technique applies more generally. Write

\[
u(t) = -X_1 e^{J_1 t} \int_t^b e^{-J_1 \tau} Y_1^T f(\tau) d\tau + X_2 e^{J_2 t} \int_a^t e^{-J_2 \tau} Y_2^T f(\tau) d\tau.
\]
Differentiating and using the fact that $X_1 Y_1^T + X_2 Y_2^T = XY^T = 0$, we obtain

$$u^{(1)}(t) = -X_1 e^{i t} \int_{t}^{b} e^{-i\tau Y_1^T} f(\tau) \, d\tau + X_2 e^{i t} \int_{a}^{t} e^{-i\tau Y_2^T} f(\tau) \, d\tau.$$  

Differentiate once more and use $XY^T = 0$, $XJY^T = A_{2}^{-1}$ to obtain

$$u^{(2)}(t) = -X_1 e^{i t} \int_{t}^{b} e^{-i\tau Y_1^T} f(\tau) \, d\tau + X_2 e^{i t} \int_{a}^{t} e^{-i\tau Y_2^T} f(\tau) \, d\tau + A_{2}^{-1} f(t).$$

Then, since $\sum_{\tau=0}^{2} A_{\tau} X_{\tau}$ for $i = 1$ and 2, we obtain

$$L(D)u = \sum_{\tau=0}^{2} A_{\tau} u^{(\tau)} = f.$$  

We now seek to modify $G_0$ in order to produce a Green's function $G$ which will retain properties (a), (b), and (c) of the lemma but will, in addition, satisfy two point boundary conditions. To formulate general conditions of this kind, let $\hat{y}_c$ denote the ln-vector determined by $\hat{y}_c^T = [y(c)^T, \ldots, y^{(l-1)}(c)^T]$ for any $y \in C^{(l-1)}[a, b]$ and any $c \in [a, b]$. Then let $M, N$ be ln $\times$ ln constant matrices, and consider homogeneous boundary conditions of the form

$$M\hat{y}_a + N\hat{y}_b = 0.$$  

For brevity, write $V(y) = M\hat{y}_a + N\hat{y}_b$, and our problem is the boundary value problem

$$L(D)u = f, \quad V(u) = 0,$$  

for the case in which the homogeneous problem, $L(D)u = 0$, $V(u) = 0$ has only the trivial solution.

We have seen that every solution of $L(D)u = 0$ is expressible in the form $u(t) = Xe^{Jc}$ for some $c \in C_{ln}$, so the boundary condition $V(u) = 0$ implies

$$M\hat{u}_a + N\hat{u}_b = (MQe^{J_0} + NQe^{J_0})c = 0,$$  

where $Q$ is given by (14). Since this equation is to have only the trivial solution, we may assume that the ln $\times$ ln matrix $MQe^{J_0} + NQe^{J_0}$ is nonsingular.
Theorem 6. If the homogeneous problem $L(D)u = 0$, $V(u) = 0$ has only the trivial solution, then there is a unique $m \times n$ matrix $K(\tau)$ independent of $t$ and depending analytically on $\tau$, in $[a, b]$, for which the function

$$G(t, \tau) = G_0(t, \tau) + Xe^J K(\tau)$$

satisfies conditions (a) and (b) of Lemma 2 (applied to $G$) as well as

(d) $V(G) = 0$, where $V$ acts on $b$ as a function of $t$.

In this case, the unique solution of (27) is given by

$$u(t) = \int_a^b G(t, \tau)f(\tau)d\tau.$$  

(30)

Proof. It is easily seen that $G$ inherits conditions (a) and (b) from $G_0$ for any $K(\tau)$. For condition (d) observe that $V(G) = 0$ if and only if $V(Xe^J K(\tau)) = - V(G_0)$. As in (28), this may be written

$$(MQu^a + NQu^b)K(\tau) = - V(G_0).$$

(31)

We have seen that our assumption on the homogeneous problem implies that the matrix on the left is nonsingular so that a unique $K$ is obtained. The analytic dependence of $K$ on $\tau$ [through $V(G_0)$] is clear.

Finally, with $u$ defined by (30) we have $V(u) = 0$ from condition (d), and

$$L(D)u = L(D)\left\{\int_a^h G_0(t, \tau)f(\tau)d\tau + Xe^J \int_a^h K(\tau)f(\tau)d\tau \right\}.$$  

The first integral reduces to $f$ by virtue of condition (c) in the lemma, and $L(D)Xe^J = (\sum_{\tau=0}^{l} A_{\tau}XJ^\tau)e^J$ is zero because of (7). This completes the proof.

7. A SECOND ORDER PROBLEM

As an important special case of the foregoing analysis we consider second order problems:

$$L(D)u = A_0 + A_1 u^{(1)} + A_2 u^{(2)} = f$$
with the boundary conditions \( u(a) = u(b) = 0 \). These boundary conditions are obtained by defining the \( 2n \times 2n \) matrices

\[
M = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}, \quad N = \begin{bmatrix} 0 & 0 \\ I & 0 \end{bmatrix},
\]

where the blocks are all \( n \times n \). Insisting that the homogeneous problem have only the trivial solution then implies that the matrix

\[
MQe^{Ja} + NQe^{Jb} = \begin{bmatrix} Xe^{Ja} \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ Xe^{Jb} \end{bmatrix} = \begin{bmatrix} Xe^{Ja} \\ Xe^{Jb} \end{bmatrix}
\]

appearing in (28) is nonsingular.

In this case we have

\[
V(G_0) = -M \begin{bmatrix} X_1 \\ X_1J_1 \end{bmatrix} e^{J_1(a-\tau)}Y_1^T + N \begin{bmatrix} X_2 \\ X_2J_2 \end{bmatrix} e^{J_2(b-\tau)}Y_2^T
\]

\[
= \begin{bmatrix} -X_1e^{J_1(a-\tau)}Y_1^T \\ X_2e^{J_2(b-\tau)}Y_2^T \end{bmatrix},
\]

and so the solution of (31) is

\[
K(\tau) = \begin{bmatrix} Xe^{Ja} \\ Xe^{Jb} \end{bmatrix}^{-1} \begin{bmatrix} X_1e^{J_1(a-\tau)}Y_1^T \\ X_2e^{J_2(b-\tau)}Y_2^T \end{bmatrix},
\]

yielding a more explicit representation of the Green’s function in (29).

The condition on \( a, b \) and the Jordan chains of \( L(\lambda) \) which is necessary and sufficient for the existence of a unique solution to our two point boundary value problem is, therefore, that the matrix (32) be nonsingular.

Sufficient conditions for this to be the case will now be formulated. First we make hypothesis H1: There exist \( n \times n \) matrices \( J_1 \) and \( J_2 \) for which \( J = J_1 \oplus J_2 \) and for which the associated \( n \times n \) matrices of Jordan chains, \( X_1 \) and \( X_2 \), are both nonsingular. This hypothesis is equivalent to asserting the existence of matrices

\[
S_i = X_iJ_iX_i^{-1}, \quad i = 1, 2,
\]
which are solutions of the unilateral equation in matrices:

\[ A_0 + A_1 S + A_2 S^2 = 0. \]  \hspace{1cm} (34)

Then it is found that matrices \( e^{S_1 t} \) are solutions of the matricial differential equation \( A_0 U + A_1 U^{(1)} + A_2 U^{(2)} = 0. \)

Suppose in addition to hypothesis \( H_1 \), we make hypothesis \( H_2 \): The matrix \( e^{S_1 (b-a)} - e^{S_2 (b-a)} \) is nonsingular. Then \( (32) \) is nonsingular. If we use the same partitioning of \( J \) in the construction of the Green’s function, then \( K(\tau) \) can be computed explicitly from \( (33) \) in the form

\[ K(\tau) = \begin{bmatrix}
-X_1^{-1} e^{-S_1 a} (e^{S_1 (b-a)} - e^{S_2 (b-a)})^{-1} e^{S_2 (b-a)} (X e^{I (a-\tau) Y T})
\end{bmatrix}.
\]  \hspace{1cm} (35)

Note also that in this case the pre-Green’s function can be written

\[ G_0(t, \tau) = \begin{cases}
-e^{S_1 (t-\tau)} & \text{if } a \leq t < \tau, \\
e^{S_2 (t-\tau)} & \text{if } \tau \leq t \leq b.
\end{cases} \]

This discussion makes an interesting connection between the purely algebraic problem of finding solutions for (34)—a difficult and interesting problem in its own right (see, for example, \([9], [10] \), and p. 296 of \([5]\))—and the boundary value problem considered here. Our result may be summarized as follows:

**Theorem 7.** If there exist \( n \times n \) matrices \( S_1, S_2 \) which are solutions of (34), if the union of their sets of elementary divisors is the set of elementary divisors of \( L(\lambda) \), and if \( e^{S_1 (b-a)} - e^{S_2 (b-a)} \) is nonsingular, then the second order boundary value problem

\[ L(D) u = f, \quad u(a) = u(b) = 0 \]

has a unique solution. This solution has an explicit representation given by combining equations \((26), (29), (30), \) and \((35)\).

This research was supported by the Science Research Council of Great Britain during the tenure of a Senior Visiting Fellowship at the University of Dundee, and also by the National Research Council of Canada.
The author is grateful to P. Rozsa for re-awakening his interest in this problem area.

REFERENCES


Received 29 March 1976; revised 1 September 1976