

A New Type of Boundary Value Coupling for Second Order Sturm-Liouville Systems

Jon Geist

Institute for Basic Standards, National Bureau of Standards, Washington, D.C. 20234

(October 6, 1971)

A natural generalization of the familiar second order Sturm-Liouville system is presented. This generalization consists of considering a number of differential equations defined on different intervals, instead of just one equation on one interval. The self-adjoint character of the differential equations is retained in the generalization, but the boundary conditions are relaxed considerably. The most general boundary conditions which can be accommodated by this sort of generalization of Sturm-Liouville theory are discussed. The existence of eigenvalues is proved, and a generalized orthogonality and a weak eigenfunction expansion theorem are derived.

Key words: Boundary value problems; coupled equations; differential equations; self-adjoint equations; Sturm-Liouville systems.

1. Introduction

The solution to a large class of problems of interest in applied mathematics can be reduced to the solution of the second order Sturm-Liouville system,

$$(r(x)y'(x, \lambda))' = (q(x) - \lambda p(x))y(x, \lambda), \quad (1)$$

where $a \leq x \leq b$, and where

$$\alpha y(b, \lambda) + \beta y'(b, \lambda) = \gamma y(a, \lambda) + \delta y'(a, \lambda) = 0. \quad (2)$$

The primes indicate differentiation with respect to x , and the proper choice of $p(x) > 0$, $r(x) > 0$, $q(x)$, a , b , α , β , γ , and δ is dictated by the particular problem. The theory of such Sturm-Liouville systems is well known [1, 2]¹ and will not be reviewed here. Suffice it to say that it is a very powerful technique for handling the problems to which it is applicable.

Recently, generalizations of the second order Sturm-Liouville system have been developed which, while retaining the distinctive Sturm-Liouville character, are applicable to problems involving composite media [3, 4, 5, 6]. The purpose of this paper is to discuss a different generalization of the second order Sturm-Liouville system which retains the Sturm-Liouville character. In particular, an eigenvalue equation, an orthogonality theorem, and a weak eigenfunction expansion theorem will be derived. This eigenfunction expansion theorem is the first place where the results of this paper deviate from the previous work on Sturm-Liouville systems. It will be shown by counter-example, that the eigenfunctions of this Sturm-Liouville like system do not form a complete set in the sense of the previously described Sturm-Liouville systems.

The generalization of the second order Sturm-Liouville system discussed in this paper, and that which is applicable to composite media problems are similar in many ways. Each arises when

AMS Subject Classification: Primary 34B25; 42A64.

¹Figures in brackets indicate the literature references at the end of this paper.

attempting to solve time dependent heat transfer problems in complicated systems, and each consists of considering a number of differential equations defined on different intervals, instead of just one equation defined on one interval. In both generalizations the self-adjoint character of the differential equations is retained, but the conditions which the solutions must satisfy at the ends of the intervals are relaxed. In both cases, the solutions to all of the equations are coupled by relations involving their values and the values of their first derivatives at the ends of the intervals on which they are defined. The difference between the two generalizations is in the specific form of the coupling. In the case of the composite media generalization, the coupling connects the boundary values at the right end of the first interval with those at the left end of the second interval; the boundary values at the right end of the second interval with those at the left end of the third interval, and so forth. In the alternate generalization, introduced here, the coupling connects the boundary values at the left ends of all of the intervals. Figures 1a and 1b show examples of the two types of coupling schematically. It is interesting that the form of coupling introduced here results in a simpler eigenvalue equation than that which results from the composite media generalization. In the case where only two intervals are considered, the composite media generalization is a special case of the generalization introduced here. Except for this case, this generalization of the second order Sturm-Liouville system does not appear to have been studied previously. An example of an application of the theory developed in this paper, is the calculation of the temperature distribution as a function of time in a number of wires whose left ends are all welded together and whose right ends are all connected to a heat sink.

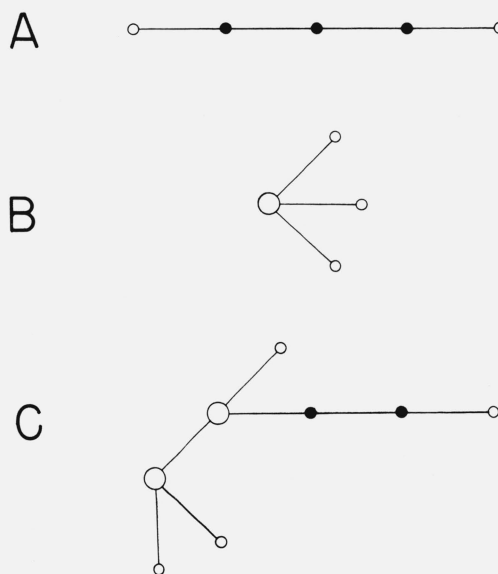


FIGURE 1. Schematic representation of different types of Boundary value coupling

- Interval on which self-adjoint differential equation is defined.
- End point of interval where Sturm-Liouville boundary conditions are imposed.
- End point of interval where boundary conditions appropriate to the composite media generalization are imposed.
- End point of interval where boundary conditions appropriate to the generalization introduced in this paper are imposed.

2. Postulates and Principal Results

Consider the N self-adjoint differential equations obtained by letting $i = 1, 2, \dots, N$ in the equation,

$$(r_i(x_i)y_i'(x_i, \lambda))' = (q_i(x_i) - \lambda p_i(x_i))y_i(x_i, \lambda) \quad (3)$$

which is to hold for $a_i \leq x_i \leq b_i$, where $r_i(x_i) > 0$ and $p_i(x_i) > 0$. For the present, retain the usual Sturm-Liouville boundary condition,

$$\alpha_i y_i(b_i, \lambda) + \beta_i y_i'(b_i, \lambda) = 0, \quad (4)$$

at the right end of each interval, but at the left end of the intervals employ a generalized boundary condition, consisting of the $N-1$ equations obtained by letting $i=1, 2, \dots, N-1$ in the equation,

$$\gamma_i y_i(a_i, \lambda) + \delta_i y_i'(a_i, \lambda) = \gamma_{i+1} y_{i+1}(a_{i+1}, \lambda) + \delta_{i+1} y_{i+1}'(a_{i+1}, \lambda) \neq 0. \quad (5)$$

One more coupling condition is needed to give a total of $2N$ boundary conditions, so introduce the general relation,

$$\sum_{i=1}^N C_i y_i(a_i, \lambda) + D_i y_i'(a_i, \lambda) = 0, \quad (6)$$

where

$$D_i \gamma_i - C_i \delta_i > 0, \quad (7)$$

for $i=1, 2, \dots, N$.

Equations 3, 4, 5, 6 and 7 define the system. Now it is necessary to determine under what conditions a set of functions, $(y_1(x_1, \lambda), y_2(x_2, \lambda), \dots, y_N(x_N, \lambda))$, exists such that these five equations are satisfied.

First concentrate attention on eqs (3) and (4). For any positive integer value of $i \leq N$, it can be shown, following Ince [7], that for every value of λ , eqs (3) and (4) determine, except for the specification of a multiplicative constant, one and only one solution which is continuous and has a continuous first derivative on the interval, (a_i, b_i) . Keeping this in mind, divide eq (6) by $\lambda_1 y_1(a_1, \lambda) + \delta_1 y_1'(a_1, \lambda)$ and substitute the $N-1$ equations represented by eq (5) into the result, to obtain the equation,

$$\sum_{i=1}^N \frac{C_i y_i(a_i, \lambda) + D_i y_i'(a_i, \lambda)}{\gamma_i y_i(a_i, \lambda) + \delta_i y_i'(a_i, \lambda)} = 0, \quad (8)$$

which is generally a transcendental equation in λ , and which is independent of the multiplicative constant which is as yet unspecified in $y_i(x_i, \lambda)$. Equation (8) is called the characteristic or eigenvalue equation of the system, since its roots are the eigenvalues of the problem; that is, the only values of λ for which eqs (3), (4), (5), and (6) are mutually consistent. The existence of a countably infinite number of such roots is proved in appendix A of this paper. Let λ_m be the m th root of eq (8), then it is called the m th eigenvalue of the system, and $y_i(x_i, \lambda_m)$ will be called the m th eigenfunction on the i th interval of the system. This completes the discussion of the eigenvalue equation.

Now attention will be turned to exhibiting the Sturm-Liouville like properties of the eigenfunctions. Other than the mere existence of eigenvalues, the most distinctive characteristic of Sturm-Liouville theory is the theorem expressing the orthogonality of the different eigenfunctions and the related eigenfunction expansion theorem. For the system described in this paper, the orthogonality theorem, which is proved in appendix B, takes the form,

$$\sum_{i=1}^N W_i \int_{a_i}^{b_i} p_i(x_i) y_i(x_i, \lambda_m) y_i(x_i, \lambda_n) dx_i = 0, \quad (9)$$

for $m \neq n$, where

$$W_i = \frac{D_i \gamma_i - C_i \delta_i}{r_i(a_i)} \quad (10)$$

Now the problem of an eigenfunction expansion theorem will be considered. Since the functions, $y_i(x_i, \lambda_m)$, are determined only to within a multiplicative constant, it is convenient to define

$$d_{im} = \left[\int_{a_i}^{b_i} p_i(x_i) y_i(x_i, \lambda_m) y_i(x_i, \lambda_m) dx_i \right]^{1/2}, \quad (11)$$

and

$$\phi_{im}(x_i) = y_i(x_i, \lambda_m) / d_{im}. \quad (12)$$

Thus each of the functions, $\phi_{im}(x_i)$, is square normalized to unity on the i th interval with respect to the weighting functions, $p_i(x_i)$, and this property uniquely determines the function, $\phi_{im}(x_i)$. Of course, the d_{im} are as yet undetermined, since the $y_i(x_i, \lambda_m)$ are determined only to within a multiplicative constant.

Now consider the possibility of choosing these multiplicative constants (and hence the d_{im}) in such a way as to express the N arbitrary functions, $g_i(x_i)$, $i=1, 2, \dots, N$, as a sum of eigenfunctions. For any given value of i , the constants, d_{im} , which are obtained by varying m are independent of each other. However, when m is held constant and i is varied, the resulting d_{im} are not independent. Once one, say d_{jm} , is specified all of the others can be obtained from the equation,

$$\frac{d_{im}}{d_{jm}} = \frac{\gamma_j \phi_{jm}(a_j) + \delta_j \phi'_{jm}(a_j)}{\gamma_i \phi_{im}(a_i) + \delta_i \phi'_{im}(a_i)}, \quad (13)$$

which is obtained by substituting eq (12) into eq (5). Thus in the eigenfunction expansions,

$$f_i(x_i) = \sum_{m=1}^{\infty} d_{im} \phi_{im}(x_i), \quad (14)$$

the N functions, $f_i(x_i)$, cannot be chosen arbitrarily. Once one of them is specified, all of the rest of them are uniquely determined. This is a strong restriction, but it does not necessarily prevent the theory from being useful. However, it does weaken the eigenvalue expansion theorem considerably. The theorem stated below and proved in appendix C is the strongest that can be proved when the functions to be expanded are left completely arbitrary.

For each value of $j=1, 2, \dots, N$, the function,

$$f_j(x_j) = \sum_{m=1}^{\infty} d_{jm} \phi_{jm}(x_j) = \sum_{m=1}^{\infty} y_j(x_j, \lambda_m), \quad (15)$$

will be a "best" approximation to the function, $g_j(x_j)$, in the sense that the expression,

$$I = \sum_{i=1}^N W_i \int_{a_i}^{b_i} p_i(x_i) [g_i(x_i) - f_i(x_i)]^2 dx_i, \quad (16)$$

is a minimum, if and only if for each $m=1, 2, \dots$,

$$d_{jm} = \frac{\sum_{i=1}^N R_{mji} W_i \int_{a_i}^{b_i} p_i(x_i) g_i(x_i) \phi_{im}(x_i) dx_i}{\sum_{i=1}^N (R_{mji})^2 W_i} \quad (17)$$

where

$$R_{mji} = \frac{d_{im}}{d_{jm}} = \frac{\gamma_j \phi_{jm}(a_j) + \delta_j \phi'_{jm}(a_j)}{\gamma_i \phi_{im}(a_i) + \delta_i \phi'_{im}(a_i)}, \quad (18)$$

Since the above eigenfunction expansion theorem does not require that each function, $f_i(x_i)$, be identical to the function, $g_i(x_i)$, a criterion for goodness of fit is useful. It is shown in appendix C that the ratio,

$$R = \sum_{n=1}^{\infty} \sum_{i=1}^N W_i (d_{in})^2 \bigg/ \sum_{i=1}^N W_i \int_{a_i}^{b_i} p_i(x_i) [g_i(x_i)]^2 dx_i \quad (19)$$

has the property that, $0 \leq R \leq 1$, and $R = 1$ if and only if $f_i(x_i) = g_i(x_i)$, for $i = 1, 2, \dots, N$. Thus R can be used as a criterion for goodness of fit in a qualitative manner. Furthermore the set of functions for which the eigenfunctions of the problem form a complete set are characterized by $R = 1$.

3. An Example

As an example, let $i = 1, 2$, and 3 , in eqs (3), (4), (5), and (6), and let $a_i = -1$, $b_i = 0$, $\epsilon = \lambda^{1/2}$, $r_i(x_i) = 1$, $q_i(x_i) = 0$, $p_i(x_i) = (i)^2$, $\alpha_i = 1$, $\beta_i = 0$, $\gamma_i = 1$, $\delta_i = 0$, $C_i = 0$, and $D_i = 1$. The resulting system arises when attempting to determine the temperature distribution, $T_i(x_i, t)$, as a function of time in three straight wires of identical thermal conductivity and area, whose specific heats per unit volume are in the ratio 1, 4, 9, and whose right ends are attached to a block whose thermal conductivity and heat capacity are infinite, and whose left ends are attached to a block whose thermal conductivity is infinite, but whose heat capacity is zero. If the initial condition for this problem is that $T_i(x_i, 0) = -x_i$, for $i = 1, 2, 3$, then $T_i(x_i, t)$ can be written in the form

$$T_i(x_i, t) = \sum_{m=1}^{\infty} d_{im} \phi_{im}(x_i) \exp [- (\epsilon_m)^2 t], \quad (20)$$

where $\epsilon_m^2 = \lambda_m$ is the m th root of eq (8), if and only if $f_i(x_i) = -x_i$, for $i = 1, 2, 3$, where $f_i(x_i)$ is defined in eq (15). However, this is not the case. The functions,

$$Z_i(x_i) = \sum_{m=1}^{150} y_i(x_i, \lambda_m), \quad (21)$$

are shown in figure 2. It is clear from the figure, that these functions are not very good approximations to $-x_i$. For this particular case, $R = 0.990159$. There is no significant difference between these results and the results when only the first fifteen terms are included in the sum, and presumably there is no significant difference between these results and the results when all of the terms are included in the sum. Thus in order to obtain $T_i(x_i, t)$ it is necessary to add to the right hand side of eq (20), a nonseparable function of x_i and t , which is a solution to the corresponding time dependent heat transfer equation for the problem. The properties of such non-separable solutions remain to be investigated.

Now suppose that we eliminate the third interval in the above problem, restricting i to be 1 or 2. Then the problem is equivalent to a composite media problem, and interestingly enough the eigenfunctions form a complete set as can be seen from figure 3. In this case, $R = 0.999989$, for just the first fifteen terms of the sum.

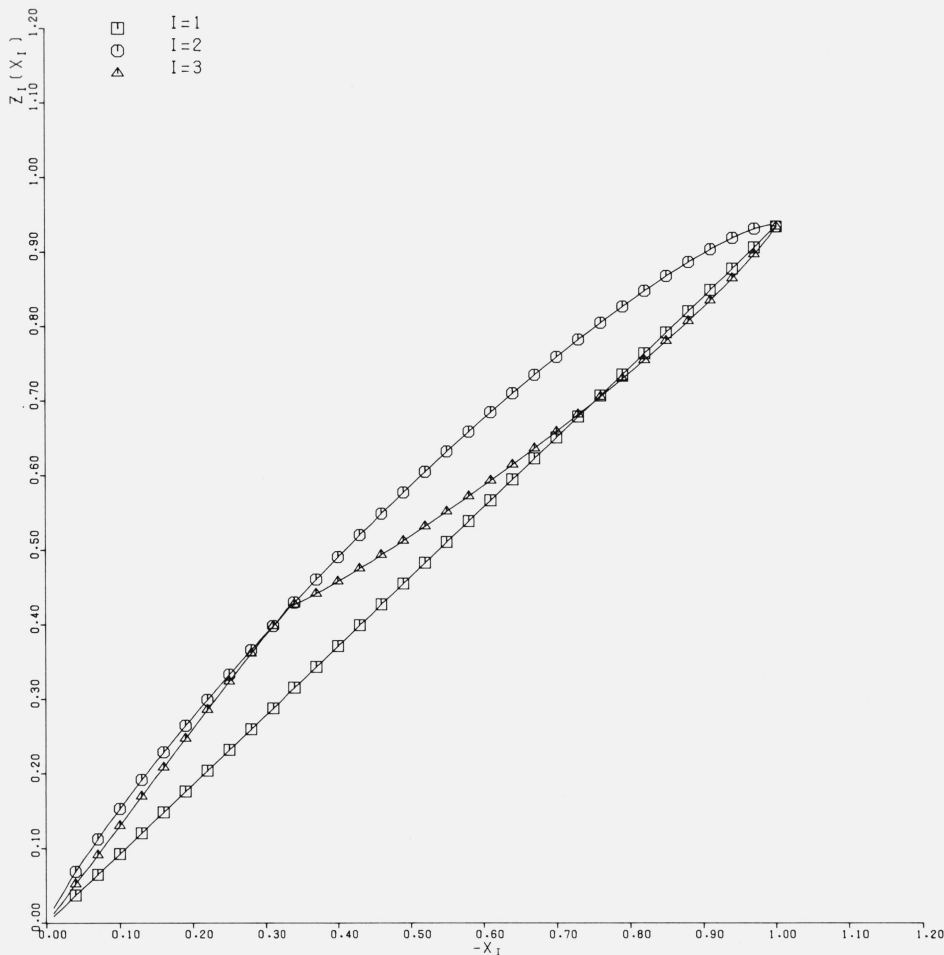


FIGURE 2. Graphs of $Z_i(x_i)$ for $i=1, 2,$ and $3,$ showing that $Z_i(x_i)$ is not a very good approximation to $-x_i$ in the three interval example considered in the paper.

4. Discussion

The fact that nonseparable solutions appear in a prominent way in the generalization of Sturm-Liouville theory discussed here certainly limits the usefulness of the theory to some extent. How serious this limitation is in various applications will depend upon the general properties of these non-separable solutions. Thus an investigation of such non-separable functions would appear to be quite worthwhile.

Nowhere in the derivation of the results presented in this paper does the assumption that $a_i < x_i < b_i$ play a fundamental role. All of the results would have been identical if for some of the values of $i,$ $b_i < x_i < a_i,$ but the discussion would have been somewhat more complicated. To further generalize the second order Sturm-Liouville system, the boundary conditions at the points, $b_i,$ could be relaxed. It is possible to replace them by generalized boundary conditions identical in form to those stated earlier in this paper, by introducing a number of new intervals on which more self-adjoint differential equations are defined. The type of coupling envisioned here should be obvious after glancing at figure 1c. The detailed development of this further generalization of the second order Sturm-Liouville system introduces no new ideas, and it is fraught with notational difficulties. Consequently it will not be pursued here. However, a few remarks are in order.

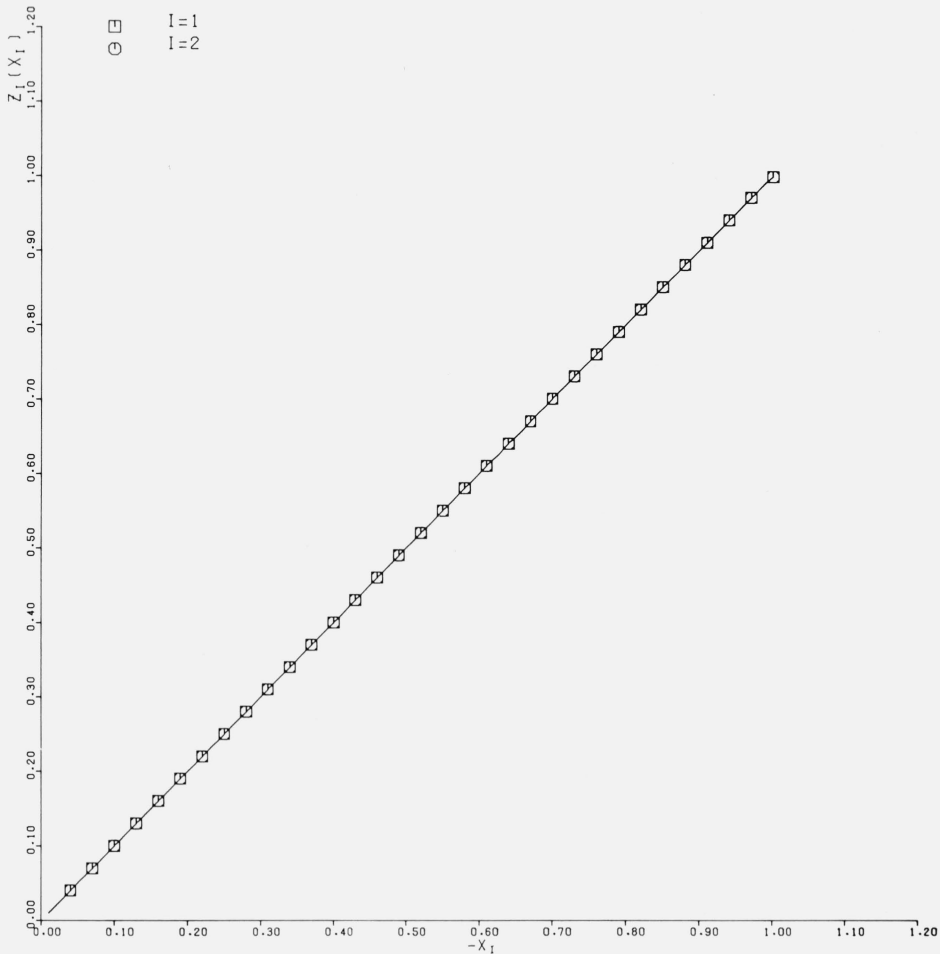


FIGURE 3. Graphs of $Z_i(x_i)$ for $i=1$, and 2, showing that $Z_i(x_i)$ is an excellent approximation to $-x_i$ in the two interval (composite media) example considered in the paper.

First, cross-linking among the branching intervals will not generally allow a nontrivial solution. But it is possible to cross-link any two intervals for which $r(a) = r(b)$ for both intervals. Second, both the composite media generalization and the generalization described earlier in this paper are special cases of this generalization of the second order Sturm-Liouville system. Third, in a detailed development of this theory, it will not be possible to derive a simple eigenvalue equation as was done here. A determinantal equation, similar in principle to, but more complicated in actual detail than that in reference [4] will have to be derived.

5. Appendix A

In this appendix, eq (8) will be proved to have an infinite number of roots. In discussing eq (8), it is convenient to define,

$$F_i(\lambda) = \frac{C_i y_i(a_i, \lambda) + D_i y_i'(a_i, \lambda)}{\gamma_i y_i(a_i, \lambda) + \delta_i y_i'(a_i, \lambda)}, \quad (1A)$$

and,

$$F(\lambda) = \sum_{i=1}^N F_i(\lambda). \quad (2A)$$

First it will be proved that for each $i = 1$ to N , $F_i(\lambda)$ has an infinite number of singularities, and that it is positive in some interval immediately to the right of each singularity, and negative in some interval immediately to the left of each singularity. In order to prove this, it is necessary to consider each of the four cases, $C_i = \delta_i = 0$, $D_i = \gamma_i = 0$, $C_i \gamma_i \neq 0$, and $D_i \delta_i \neq 0$, separately. Every possible choice of C_i , γ_i , D_i , and δ_i which is consistent with eq (7) is covered by at least one of these cases. Before looking into each of these cases separately, it is necessary to review some of the properties of $y_i(a_i, \lambda)$. Ince [8] has shown that if $y_i(x_i, \lambda)$ satisfies eqs (3) and (4), then the equation, $y_i(a_i, \lambda) = 0$, has a countably infinite number of roots. These will be denoted by μ_{im} , where the μ_{im} are labeled so that $\mu_{im} < \mu_{im+1}$. Ince [9] has shown that as λ is increased from μ_{im} to μ_{im+1} , the expression,

$$\frac{y'_i(a_i, \lambda)}{y_i(a_i, \lambda)},$$

decreases monotonically from $+\infty$ to $-\infty$. Thus $y'_i(a_i, \lambda)$ also has an infinite number of roots, denoted by ω_{im} , and as λ is increased from ω_{im} to ω_{im+1} , the expression,

$$\frac{y_i(a_i, \lambda)}{y'_i(a_i, \lambda)},$$

increases monotonically from $-\infty$ to $+\infty$.

Case 1: $C_i = \delta_i = 0$

In this case,

$$F_i(\lambda) = \frac{D_i y'_i(a_i, \lambda)}{\gamma_i y_i(a_i, \lambda)}, \quad (3A)$$

and dividing eq (7) by $(\gamma_i)^2$ shows that D_i/γ_i is a positive number. Thus $F_i(\lambda)$ decreases monotonically from $+\infty$ to $-\infty$, as λ is increased from μ_{im} to μ_{im+1} , so it has an infinite number of singularities and it is positive immediately to the right of each singularity and it is negative immediately to the left of each singularity.

Case 2: $D_i = \gamma_i = 0$

In this case,

$$F_i(\lambda) = \frac{C_i y_i(a_i, \lambda)}{\delta_i y'_i(a_i, \lambda)}, \quad (4A)$$

but dividing eq (7) by $(\delta_i)^2$ shows that C_i/δ_i is a negative number. Thus $F_i(\lambda)$ decreases monotonically from $+\infty$ to $-\infty$ as λ is increased from ω_{im} to ω_{im+1} . So again, $F_i(\lambda)$ has an infinite number of singularities, is positive to the right of each singularity, and is negative to the left of each singularity.

Case 3: $D_i \delta_i \neq 0$

$$\text{Let } f_i(\lambda) = \left(\frac{C_i}{D_i} + \frac{y'_i(a_i, \lambda)}{y_i(a_i, \lambda)} \right) \left(\frac{\gamma_i}{\delta_i} + \frac{y'_i(a_i, \lambda)}{y_i(a_i, \lambda)} \right)^{-1}, \quad (5A)$$

then $F_i(\lambda) = (D_i/\delta_i) f_i(\lambda)$. Now the equations,

$$\frac{y'_i(a_i, \lambda)}{y_i(a_i, \lambda)} + \frac{C_i}{D_i} = 0, \quad (6A)$$

and

$$\frac{y'_i(a_i, \lambda)}{y_i(a_i, \lambda)} + \frac{\gamma_i}{\delta_i} = 0, \quad (7A)$$

each have one and only one root between μ_{im} and μ_{im+1} . This root of eq (7A) will be denoted by ξ_{im} and the equivalent root of eq (6A) will be denoted by η_{im} . First supposing that $0 < D_i \delta_i$, use eq (7) to show that $-\gamma_i/\delta_i < -C_i/D_i$, so that $\mu_{im} < \eta_{im} < \xi_{im} < \mu_{im+1}$. Thus

$$\frac{-\gamma_i}{\delta_i} < \frac{y'_i(a_i, \lambda)}{y_i(a_i, \lambda)} < -\frac{C_i}{D_i} \text{ for } \eta_{im} < \lambda < \xi_{im}, \quad (8A)$$

and

$$\frac{y'_i(a_i, \lambda)}{y_i(a_i, \lambda)} < -\frac{\gamma_i}{\delta_i} \text{ for } \xi_{im} < \lambda < \mu_{im+1}. \quad (9A)$$

Since ξ_{im} is a root of eq (7A), $f_i(\lambda)$ has a singularity at ξ_{im} , and eq (8A) can be used to show that $f_i(\lambda)$ is negative when $\eta_{im} < \lambda < \xi_{im}$, and eq (9A) can be used to show that $f_i(\lambda)$ is positive when $\xi_{im} < \lambda < \mu_{im+1}$. Letting $m=1, 2, \dots$ in the above discussion shows that $f_i(\lambda)$ has an infinite number of singularities and that it is positive immediately to the right of each singularity and that it is negative immediately to the left of each singularity. Now, supposing that $D_i \delta_i < 0$, use eq (7) to show that

$$\frac{-C_i}{D_i} < -\frac{\gamma_i}{\delta_i},$$

so that $\mu_{im} < \xi_{im} < \eta_{im} < \mu_{im+1}$.

Thus

$$-\frac{C_i}{D_i} < \frac{y'_i(a_i, \lambda)}{y_i(a_i, \lambda)} < -\frac{\gamma_i}{\delta_i} \text{ for } \xi_{im} < \lambda < \eta_{im}, \quad (10A)$$

and

$$-\frac{\gamma_i}{\delta_i} < \frac{y'_i(a_i, \lambda)}{y_i(a_i, \lambda)}, \text{ for } \mu_{im} < \lambda < \xi_{im} \quad (11A)$$

Since ξ_{im} is a root of eq (7A), $f_i(\lambda)$ has a singularity at ξ_{im} , and eq (10A) implies that $f_i(\lambda)$ is negative when $\xi_{im} < \lambda < \eta_{im}$, and eq (11A) implies that $f_i(\lambda)$ is positive when $\mu_{im} < \lambda < \xi_{im}$. Letting $m=1, 2, \dots$ in the above discussion shows that $f_i(\lambda)$ has an infinite number of singularities and it is negative in some interval immediately to the right of each singularity and that is positive in some interval immediately to the left of each singularity. So in both the case where $D_i \delta_i > 0$ and in the case where $D_i \delta_i < 0$, $F_i(\lambda) = (D_i/\delta_i) f_i(\lambda)$ has an infinite number of singularities and it is positive immediately to the right of each singularity and negative immediately to the left of each singularity.

Case 4: $C_i \gamma_i \neq 0$

The proof is similar at each step to that of Case 3 and will not be presented here.

If $F_i(\lambda)$ has a singularity at κ_{im} , then eq (2A) implies that $F(\lambda)$ also has a singularity there, and $F(\lambda)$ behaves like $F_i(\lambda)$ in some neighborhood of κ_{im} . Thus, if κ_{im} and κ_{jm} are two consecutive

singularities of $F(\lambda)$, and $\kappa_{im} < \kappa_{jn}$, then $F(\lambda)$ is positive immediately to the right of κ_{im} and negative immediately to the left of κ_{jn} . Now since $F(\lambda)$ is continuous for $\kappa_{im} < \lambda < \kappa_{jn}$, it must have a zero in this interval. Thus eq (8) has at least one root between any two consecutive singularities of $F(\lambda)$. Finally, since $F(\lambda)$ has an infinite number of singularities, it has an infinite number of roots, as was to be proved.

6. Appendix B

In this appendix, eq (9) will be proved. The proof requires a lemma which will be proved separately, but in somewhat more general form than needed.

Lemma

Since this lemma consists almost entirely of algebraic manipulations, it will prove convenient to use the shorthand notation,

$$M_i = M_i(a_i, \xi), \quad (1B)$$

$$N_i = N_i(a_i, \eta), \quad (2B)$$

and

$$r_i = r_i(a_i), \quad (3B)$$

where $M_i(x_i, \xi)$ and $N_i(x_i, \eta)$ satisfy eqs (3), (4), and (5), with $\lambda = \xi$, and $\lambda = \eta$, respectively. Now let

$$S = (\gamma_i N_i + \delta_i N'_i) \sum_{j=1}^N (C_j M_j + D_j M'_j) - (\gamma_i M_i + \delta_i M'_i) \sum_{j=1}^N (C_j N_j + D_j N'_j), \quad (4B)$$

and notice that if M_j and N_j both satisfy eq (6), then $S=0$, but if only one of them satisfies eq (6), then $S \neq 0$.

Since both M_i and N_i satisfy eq (5),

$$S = \sum_{j=1}^N [(C_j M_j + D_j M'_j) (\gamma_j N_j + \delta_j N'_j) - (C_j N_j + D_j N'_j) (\gamma_j M_j + \delta_j M'_j)]. \quad (5B)$$

When the indicated multiplications in eq (5B) are carried out, some of the terms cancel one another, and the remaining terms can be combined to yield,

$$S = \sum_{j=1}^N (D_j \gamma_j - C_j \delta_j) (M'_j N_j - N'_j M_j). \quad (6B)$$

Thus by eq (10),

$$S = \sum_{j=1}^N W_j r_j (M'_j N_j - N'_j M_j) \quad (7B)$$

Churchill [10] has shown that eqs (3) and (4) are sufficient conditions for

$$(\xi - \eta) \int_{a_j}^{b_j} p_j(x_j) M_j(x_j, \xi) N_j(x_j, \eta) dx_j = r_j (M'_j N_j - N'_j M_j). \quad (8B)$$

Thus

$$(\xi - \eta) \sum_{j=1}^N \mathcal{W}_j \int_{a_j}^{b_j} p_j(x_j) M_j(x_j, \xi) N_j(x_j, \eta) dx_j = S, \quad (9B)$$

where S is given by eq (4B). This is the significant result of the lemma.

Now let $\xi = \lambda_m$ and $\eta = \lambda_n$. Then $M_j(x_j, \xi) = y_j(x_j, \lambda_m)$ and $N_j(x_j, \eta) = y_j(x_j, \lambda_n)$, so M_j and N_j both satisfy eq (6), so $S=0$. And if $m \neq n$, eq (9B) becomes eq (9). This completes the proof.

7. Appendix C

In this appendix, it will be shown that eq (17) is a necessary and sufficient condition for the expression, I , defined in eq (16), to have a minimum with respect to variation of the d_{jm} subject to the constraints imposed by the $N-1$ equations represented by eq (13).

First notice that by the restrictions stated after eq (3), by eq (7) and eq (10), $I \geq 0$, so it has a minimum. Also notice that $I=0$, if and only if $f_i(x_i) = g_i(x_i)$ for $i=1, 2, \dots, N$. Next take the variation of I with respect to d_{jm} where $m=1, 2, \dots$, and j is fixed. Using eq (13) to justify replacing the partial derivative of d_{im} with respect to d_{jm} by d_{im}/d_{jm} , rearrange the result to obtain,

$$\delta(I) = \sum_{m=1}^{\infty} \delta(d_{jm}) \sum_{i=1}^N \left(\frac{d_{im}}{d_{jm}} \right) \mathcal{W}_i \int_{a_i}^{b_i} p_i(x_i) (g_i(x_i) - f_i(x_i)) \phi_{im}(x_i) dx_i = 0. \quad (1C)$$

Now since each d_{jm} is independent of d_{jn} for $n \neq m$, the coefficient of every $\delta(d_{jm})$ must be zero. Combining eqs (15), and (18) with this result yields

$$\sum_{i=1}^N R_{mji} \mathcal{W}_i \int_{a_i}^{b_i} p_i(x_i) g_i(x_i) \phi_{im}(x_i) dx_i = \frac{1}{d_{jm}} \sum_{i=1}^N \mathcal{W}_i \int_{a_i}^{b_i} p_i(x_i) f_i(x_i) y_i(x_i, \lambda_m) dx_i, \quad (2C)$$

for $m=1, 2, \dots$. Finally applying eqs (15), (9), and (11) to the right-hand side of eq (2C), and again using eq (18) one sees that the simultaneous equations are diagonalized and that they are equivalent to eq (17). All of the steps in this proof are reversible, so eq (17) is a necessary and sufficient condition for I to have a minimum value as was to be shown.

In order to obtain a criterion for goodness of fit, expand I to obtain,

$$I = \sum_{i=1}^N \mathcal{W}_i \int_{a_i}^{b_i} p_i(x_i) (g_i(x_i))^2 dx_i + \sum_{i=1}^N \mathcal{W}_i \int_{a_i}^{b_i} p_i(x_i) (f_i(x_i))^2 dx_i - 2 \sum_{i=1}^N \mathcal{W}_i \int_{a_i}^{b_i} p_i(x_i) g_i(x_i) f_i(x_i) dx_i. \quad (3C)$$

Notice that if eq (2C) is multiplied by d_{jm} and summed over m , the result is

$$\sum_{i=1}^N \mathcal{W}_i \int_{a_i}^{b_i} p_i(x_i) g_i(x_i) f_i(x_i) dx_i = \sum_{i=1}^N \mathcal{W}_i \int_{a_i}^{b_i} p_i(x_i) (f_i(x_i))^2 dx_i. \quad (4C)$$

Substitute this result into eq (3C) to obtain

$$I = \sum_{i=1}^N \mathcal{W}_i \int_{a_i}^{b_i} p_i(x_i) (g_i(x_i))^2 dx_i - \sum_{i=1}^N \mathcal{W}_i \int_{a_i}^{b_i} p_i(x_i) (f_i(x_i))^2 dx_i. \quad (5C)$$

Finally substitute eq (15) into eq (5C), and use the orthogonality theorem, eq (9) to obtain,

$$I = \sum_{i=1}^N \mathcal{W}_i \int_{a_i}^{b_i} p_i(x_i) (g_i(x_i))^2 dx_i - \sum_{m=1}^{\infty} \sum_{i=1}^N \mathcal{W}_i (d_{im})^2. \quad (6C)$$

Since $I \geq 0$, $0 \leq R \leq 1$, where R was defined in eq (19). The second property of R follows from the property of I that it is zero if and only if $g_i(x_i) = f_i(x_i)$ for $a_i \leq x_i \leq b_i$, and $i = 1, 2, \dots, N$.

I thank R. C. Kraft and W. P. Reid for constructive criticism on various aspects of this work, and J. S. Gallagher for preparing some of the figures.

8. References

- [1] Ince, E. L., Ordinary Differential Equations (Dover Publications, New York, 1944), 1st American ed., p. 217 and Ch. 10.
- [2] Churchill, R. V., Operational Mathematics (McGraw-Hill, New York, 1958, 2nd ed., Ch. 9.
- [3] Reid, W. P., J. Franklin Inst. **274**, 352 (1962).
- [4] Bulavin, P. E. and Kashcheev, V. M., Int. Chem. Eng., **5**, 112 (1965).
- [5] Tittle, C. W., J. Appl. Phys. **36**, 1486 (1965).
- [6] Formerly such problems were solved by using integral transforms, and the evaluation of a complex inversion integral was required. See Carslaw, H. C. and Jaeger, J. C., Conduction of Heat in Solids, (Oxford University Press, London, 1959), 2d ed., p. 319, and references listed there. Actually, the two techniques are not unrelated. An orthogonality theorem identical to that employed in the direct eigenvalue expansion is often used to study the properties of solutions derived by transform techniques, and the same eigenvalue equation occurs in both cases.
- [7] Ince, E. L., Ordinary Differential Equations (Dover Publications, New York, 1944), 1st American ed., pp. 73-75.
- [8] Ince, E. L., Ordinary Differential Equations (Dover Publications, New York, 1944), 1st American ed., pp. 231-232.
- [9] Ince, E. L., Ordinary Differential Equations (Dover Publications, New York, 1944), 1st American ed., pp. 229-230, 232.
- [10] Churchill, R. V., Operational Mathematics (McGraw-Hill, Inc., New York, 1958), 2d ed., pp. 264-265. (Note that Churchill denotes by $-\lambda$, what Ince and the author of the present paper denote by λ .)

(Paper 75B3&4-352)