

## ON COMPUTATION OF EIGENVALUES FOR DIFFERENTIAL EQUATIONS BY NEWTON'S METHOD

BY

ZEEV NEHARI

**ABSTRACT.** Professor Zeev Nehari, in the last months before his untimely death, had been analyzing and developing a new algorithm for computing eigenvalues of selfadjoint boundary value problems of arbitrary order. Apparently, his main goals were Theorems 3.1 and 3.2 below (which incidentally yield the eigenvalues of the given problem, as well as those of a related problem).

Unfortunately, as far as we know, Professor Nehari has not left a proof of the basic Lemma 3.1 which presumably was to be based on §2 or related results. It seems, however, worthwhile to publish this paper since the results of §§1 and 2 are complete and of independent interest and since the question of the validity of Lemma 3.1 is also of interest. While the presentation in §§1 and 2 is in most respects complete, there are a few places where some minor clarifications by selected footnotes were felt to be desirable. Also, Professor Nehari's original incomplete §3 has been replaced by a revised and abbreviated version.

1. Let  $L$  be an  $n$ th order selfadjoint linear differential operator of the form  $Ly = \sum_{\nu=0}^n q_{\nu}(x)y^{(\nu)}$ ,  $q_n > 0$ , where, to avoid irrelevant complications, it is assumed that the  $q_{\nu}$  are of class  $C^{(n)}$  on  $[0, 1]$ . If  $p > 0$  and is continuous on  $[0, 1]$ , we consider the eigenvalue problem

$$Lu + \lambda pu = 0, \quad B_{\nu}(u) = 0, \quad \nu = 1, \dots, n, \quad (1.1)$$

where the  $B_{\nu}$  are linearly independent linear selfadjoint homogeneous boundary conditions involving the values of  $u, u', \dots, u^{(n-1)}$  at 0 and 1, and it is assumed that the  $B_{\nu}$  do not depend on  $\lambda$ . For these boundary conditions we shall also use the abbreviated notation  $B(u) = 0$ . We note that it would be sufficient to treat the case  $p \equiv 1$ , but applied problems are usually of the form (1.1) with  $p \neq 1$ , and a preliminary transformation changing  $p$  into 1 can introduce considerable formal complications.

We shall further assume that the problem (1.1) is positive-definite, i.e., that, for any  $u \in C^n$  which satisfies the boundary conditions  $B(u) = 0$ , we have

$$-\int_0^1 uLu \, dx > 0. \quad (1.2)$$

Under these assumptions, we shall derive an effective algorithm, based on Newton's method, for the computation of the eigenvalues  $\lambda_1, \lambda_2, \dots$  (and the corresponding eigenfunctions) of (1.1).

---

Received by the editors February 12, 1979 and, in revised form, July 16, 1979.

AMS (MOS) subject classifications (1970). Primary 34B25, 65L15.

Key words and phrases. Associated selfadjoint boundary conditions, Schwarzian derivative.

© 1981 American Mathematical Society  
0002-9947/81/0000-0056/\$04.25

Many of our considerations do not depend on the positive-definiteness of the problem, and we therefore begin with those aspects of our problem which can be treated without the assumption (1.2).

If  $u, v \in C^n[0, 1]$  we have, by Lagrange's identity and the selfadjointness of  $L$ ,

$$\int_0^1 (vLu - uLv) dx = P(u, v) \quad (1.3)$$

where  $P(u, v)$  is a bilinear form in the variables  $u(0), u'(0), \dots, u^{(n-1)}(0), u(1), \dots, u^{(n-1)}(1)$  and  $v(0), \dots, v^{(n-1)}(0), v(1), \dots, v^{(n-1)}(1)$ . Using the linear independence of the linear forms  $B_\nu, \nu = 1, \dots, n$ , and the selfadjointness it is not difficult to show (see [1, Chapter 11, §3]) that (1.3) may be replaced by

$$\int_0^1 (vLu - uLv) dx = \sum_{\nu=1}^n [B_\nu(v)C_\nu(u) - B_\nu(u)C_\nu(v)] \quad (1.4)$$

where the  $C_\nu(u)$  are linear forms in the variables  $u(0), u'(0), \dots, u^{(n-1)}(1)$ . It may be noted that the  $C_\nu$  are not uniquely determined by the identity (1.4); clearly, the right-hand side of (1.4) will remain unchanged if  $C_\nu$  is replaced by  $C_\nu + \sum_{\mu=1}^n \alpha_{\nu\mu} B_\mu$ , where  $\alpha_{\nu\mu} = \alpha_{\mu\nu}$  and the  $\alpha_{\nu\mu}$  are otherwise arbitrary. It is not difficult to show that this is the only arbitrariness in the choice of the  $C_\nu$ . However, since we are not going to make use of this fact, we omit the proof.

We shall describe the relationship between the linear form  $B_\nu$  and the corresponding form  $C_\nu$  by saying that  $C_\nu$  is "associated" with  $B_\nu$ . Similarly, the boundary condition  $C_\nu = 0$  will be said to be associated with the boundary condition  $B_\nu = 0$ . We note here that the relation between  $B_\nu$  and  $C_\nu$  is symmetric except for a sign change: If  $C_\nu$  is associated with  $B_\nu$ , then  $-B_\nu$  is associated with  $C_\nu$ . This is an immediate consequence of the fact that the right-hand side of (1.4) remains unchanged under the substitution

$$B_\nu \rightarrow C_\nu, \quad C_\nu \rightarrow -B_\nu.$$

Note that given (1.4), it follows that  $B = 0$  is a set of selfadjoint boundary conditions for the operator  $L$ . Indeed, if  $B_\nu(u) = B_\nu(v) = 0, \nu = 1, \dots, n$ , we have

$$\int_0^1 vLu dx = \int_0^1 uLv dx \quad (1.5)$$

and the assertion follows. Since, as just pointed out, the substitution  $B_\nu \rightarrow C_\nu, C_\nu \rightarrow -B_\nu$  does not affect the right-hand side of (1.4), the relation (1.5) will remain valid if  $B_\nu$  is replaced by  $C_\nu$  for one or more subscripts  $\nu$ . We can thus obtain additional sets of selfadjoint boundary conditions for the equation  $Lu + \lambda pu = 0$ . For instance, if this substitution is carried out only for one subscript, say  $\nu = n$ , we obtain the selfadjoint boundary value problem

$$Lu + \lambda u = 0, \quad B_\nu(u) = 0, \quad \nu = 1, \dots, n-1, \quad C_n(u) = 0. \quad (1.6)$$

These boundary conditions—which we shall denote by the symbol  $B^*(u) = 0$ —and the corresponding solutions of the problem (1.6) depend, of course, on the numbering of the original boundary conditions. There are thus  $n$  different boundary conditions  $B^*(u) = 0$  of this type, all of which are equally suitable for our purposes (although some may be preferable from the point of view of computational convenience).

The proof of our main result will be greatly facilitated by the assumption that the problem

$$Lu + \lambda pu = 0, \quad B_\nu(u) = 0, \quad \nu = 1, \dots, n-1, \quad (1.7)$$

(for given real  $\lambda$ ) has a solution which is unique except for a multiplicative constant. (Later, it will be shown that it is sufficient to establish our result under this assumption.)<sup>1</sup> Because of this uniqueness property, the solution of (1.7) can be expressed in an elementary fashion in terms of the solution of the initial value problem (at  $x = 0$ ) for  $Lu + \lambda pu = 0$  and some algebraic manipulation. To determine the constant factor, which so far has remained arbitrary, we normalize this solution  $u$  by the condition

$$\int_0^1 pu^2 dx = 1 \quad (1.8)$$

and by some sign condition such as the requirement that  $u > 0$  in a small right neighborhood of 0.

If  $\lambda$  happens to be an eigenvalue of problem (1.1), we will clearly have  $B_n(u) = 0$ . Conversely, the equation  $B_n(u) = 0$  implies that  $\lambda$  is an eigenvalue. To emphasize the dependence of  $B_n(u)$  on the value  $\lambda$  for which the problem (1.7) has been solved (with the normalization (1.8)), we shall set

$$B_n(u) = \zeta(\lambda). \quad (1.9)$$

In this terminology, the problem of finding all the eigenvalues of (1.1) is identical with the problem of finding all the solutions of the equation  $\zeta(\lambda) = 0$ .

Analogously, the eigenvalues  $\lambda$  of problem (1.6) are characterized by the condition  $C_n(u) = 0$ , where  $u$  is the corresponding solution of (1.7). If the latter is normalized by (1.8), we set

$$C_n(u) = \eta(\lambda). \quad (1.10)$$

Clearly, the zeros of the function  $\eta(\lambda)$  coincide with the eigenvalues of problem (1.6). We note here that, as mentioned earlier, the linear form  $C_n$  is not uniquely determined by the forms  $B_\nu$ . Hence,  $\eta(\lambda)$  is not a uniquely determined function either. However, it will soon become apparent that all possible functions  $\eta$  are of the form  $A\eta_0 + B\zeta$ , where  $\eta_0$  is one of them and  $A, B$  are constants.

Finally, we introduce the function

$$\varphi(\lambda) = \zeta(\lambda)/\eta(\lambda). \quad (1.11)$$

Since, as will be shown presently, no zero of  $\zeta(\lambda)$  can coincide with a zero of  $\eta(\lambda)$ , the zeros of  $\varphi(\lambda)$  are the same as the zeros of  $\zeta(\lambda)$ . Thus, in view of the remark made above concerning the zeros of  $\zeta(\lambda)$ , the set of eigenvalues of problem (1.1) coincides with the set of zeros of the function  $\varphi(\lambda)$ . We shall attempt to find these zeros—and thus solve our eigenvalue problem—by applying Newton's method to the equation  $\varphi(\lambda) = 0$ . However, before we do so, a number of preparations are necessary.

---

<sup>1</sup>This argument was not given in the draft. Note that this assumption is *always* valid for  $\lambda$  not an eigenvalue.

We first remark that, by a standard argument, the solution  $u$  of any initial value problem of  $Lu + \lambda pu \equiv 0$  is an analytic function of  $\lambda$  for all finite  $\lambda$ . Since, as pointed out before, the conditions  $B_\nu(u) = 0$ ,  $\nu = 1, \dots, n-1$ , in conjunction with one normalization condition are equivalent to a set of initial conditions, it follows that  $B_n(u)$  and  $C_n(u)$ —i.e., by (1.9) and (1.10)  $\zeta(\lambda)$  and  $\eta(\lambda)$ —are analytic functions of  $\lambda$ , provided no extra singularities are introduced by the normalization. This, however, is not the case, at least for real values of  $\lambda$ , since we use the normalization (1.8), and the left-hand side of (1.8) cannot be zero if  $\lambda$  and  $u$  are real. We therefore conclude that  $\zeta(\lambda)$  and  $\eta(\lambda)$ —and also  $u$  if regarded as a function of  $\lambda$  for fixed  $x$ —are analytic functions of  $\lambda$  if  $\lambda$  is on the real axis.

Consider now the solution  $u$  of

$$Lu + \lambda pu = 0, \quad (1.12)$$

subject to the boundary conditions  $B_\nu(u) = 0$ ,  $\nu = 1, \dots, n-1$ , and the normalization (1.8). This solution is unique and, as just seen, analytic in  $\lambda$ . It is thus permissible to differentiate (1.12) with respect to  $\lambda$ . Setting

$$v = \partial u / \partial \lambda, \quad (1.13)$$

we find that  $v$  is a solution of the differential equation

$$Lv + \lambda pv + pu = 0. \quad (1.14)$$

Since, regardless of the value of  $\lambda$ ,  $B_\nu(u) = 0$ ,  $\nu = 1, \dots, n-1$ , we also have  $B_\nu(v) = 0$ ,  $\nu = 1, \dots, n-1$ . Hence, by (1.4),

$$\int_0^1 (vLu - uLv) dx = B_n(v)C_n(u) - B_n(u)C_n(v).$$

By (1.9), (1.10) and (1.13), the right-hand side is equivalent to  $\zeta'(\lambda)\eta(\lambda) - \zeta(\lambda)\eta'(\lambda)$ . Since, by (1.12), (1.14) and (1.8), the left-hand side is equal to 1, we obtain the relation

$$\zeta'(\lambda)\eta(\lambda) - \eta'(\lambda)\zeta(\lambda) = 1. \quad (1.15)$$

Since both  $\zeta(\lambda)$  and  $\eta(\lambda)$  are analytic for real  $\lambda$  this proves, incidentally, our earlier assertion that  $\zeta$  and  $\eta$  cannot have common real zeros.

Differentiating this with respect to  $\lambda$ , we have  $\zeta''\eta - \eta''\zeta = 0$ , i.e.,  $\zeta''/\zeta = \eta''/\eta$ . If we denote this common value by  $-q(\lambda)$ , we thus have

$$\zeta''(\lambda) + q(\lambda)\zeta(\lambda) = 0, \quad \eta''(\lambda) + q(\lambda)\eta(\lambda) = 0. \quad (1.16)$$

Both  $\zeta(\lambda)$  and  $\eta(\lambda)$  are thus found to be solutions of the same linear second-order differential equation. This, incidentally, proves the assertion made earlier that all possible functions  $\eta(\lambda)$  associated with a given  $\zeta(\lambda)$  are of the form  $\eta(\lambda) = A\eta_0(\lambda) + B\zeta(\lambda)$ , where  $\eta_0(\lambda)$  is one of them and  $A, B$  are constants.

We now prove that the coefficient  $q(\lambda)$  of equation (1.16) must be positive, a fact which will later be found of crucial importance. Differentiating (1.14) with respect to  $\lambda$  and setting

$$\partial v / \partial \lambda = w \quad (1.17)$$

we find that  $w$  satisfies the equation

$$Lw + \lambda pw + 2pv = 0 \quad (1.18)$$

and, of course, the boundary conditions  $B_\nu(w) = 0, \nu = 1, \dots, n - 1$ . Thus, by (1.4), (1.14) and (1.18),

$$B_n(w)C_n(v) - B_n(v)C_n(w) = -\int_0^1 puw \, dx + 2\int_0^1 pv^2 \, dx.$$

By (1.9), (1.10), (1.13) and (1.17), the left-hand side is equal to  $\zeta''(\lambda)\eta'(\lambda) - \eta''(\lambda)\zeta'(\lambda)$ . Since, by twice differentiating (1.8) and observing (1.13) and (1.17), we obtain

$$0 = 2\int_0^1 puw \, dx + 2\int_0^1 pv^2 \, dx,$$

we thus arrive at the identity

$$\zeta''(\lambda)\eta'(\lambda) - \eta''(\lambda)\zeta'(\lambda) = 3\int_0^1 pv^2 \, dx.$$

By (1.16) and (1.15),

$$\zeta''(\lambda)\eta'(\lambda) - \eta''(\lambda)\zeta'(\lambda) = q(\lambda)[\zeta'(\lambda)\eta(\lambda) - \eta'(\lambda)\zeta(\lambda)] = q(\lambda),$$

and our identity simplifies to

$$q(\lambda) = 3\int_0^1 pv^2 \, dx. \tag{1.19}$$

The coefficient  $q(\lambda)$  in the differential equation (1.16) is thus indeed found to be positive, as asserted.

We are now ready to begin our attempt to solve the equation  $\varphi(\lambda) = 0$  (where  $\varphi$  is the function in (1.11)) by means of Newton's method. If  $\lambda$  is a tentative value for a zero of  $\varphi(\lambda)$ , we replace it by a modified value  $\lambda_1$  defined by

$$\lambda_1 = \lambda - \varphi(\lambda)/\varphi'(\lambda). \tag{1.20}$$

Iterating this procedure, we obtain a sequence  $\lambda, \lambda_1, \lambda_2, \dots$  which will rapidly converge to a zero  $\bar{\lambda}$  of  $\varphi(\lambda) = 0$  if the initial guess  $\lambda$  was "close enough" to  $\bar{\lambda}$ . This last remark is of course of no practical value, since we do not know even approximately where the zeros of  $\varphi(\lambda)$  are located. We shall, however, see that considerable information concerning the behavior of such sequences  $\lambda, \lambda_1, \lambda_2, \dots$  can be extracted from the definition (1.20) of the basic step, without having any prior information regarding the location of the zeros.

The definition (1.20) of  $\lambda_1$ , can be cast into a simpler and more suggestive form. Since, by (1.11) and (1.15),

$$\varphi'(\lambda) = 1/\eta^2(\lambda), \tag{1.21}$$

it follows, by another use of (1.11), that (1.20) is equivalent to

$$\lambda_1 = \lambda - \zeta(\lambda)\eta(\lambda). \tag{1.22}$$

We note that, by (1.21),  $\varphi(\lambda)$  is increasing from  $-\infty$  to  $+\infty$  if  $\lambda$  increases from one zero of  $\eta(\lambda)$  to the adjacent zero on the right. Thus, between two adjacent zeros of  $\eta(\lambda)$  there will be precisely one zero of  $\varphi(\lambda)$  (or, equivalently of  $\zeta(\lambda)$ ). This, of course, also follows from the Sturm separation theorem if it is observed that both  $\zeta(\lambda)$  and  $\eta(\lambda)$  are solutions of the equation (1.16).

To get a more accurate idea of the behavior of  $\varphi(\lambda)$  between two adjacent zeros of  $\eta(\lambda)$ , we compute the Schwarzian derivative

$$\{\varphi, \lambda\} = \left(\frac{\varphi''}{\varphi'}\right)' - \frac{1}{2}\left(\frac{\varphi''}{\varphi'}\right)^2 = \frac{\varphi'''}{\varphi'} - \frac{3}{2}\left(\frac{\varphi''}{\varphi'}\right)^2. \tag{1.23}$$

Since, by (1.21),  $\varphi''/\varphi' = -2\eta'/\eta$  we easily obtain

$$\{\varphi, \lambda\} = -2\eta''/\eta = 2q(\lambda) \tag{1.23'}$$

(the last equality following from (1.16)). As shown before (cf. (1.19)),  $q(\lambda) > 0$ . Hence,  $\{\varphi, \lambda\} > 0$ , and we can conclude from (1.23) and (1.21) that  $\varphi''' > 0$  i.e.,  $\varphi''(\lambda)$  is increasing in an interval  $I$  between two adjacent zeros of  $\eta(\lambda)$ . Since, by (1.15),  $\eta$  and  $\eta'$  cannot be zero simultaneously, it follows from  $\varphi'' = -2\eta'/\eta^3$  that  $\varphi''$  increases from  $-\infty$  to  $+\infty$  in this interval. As a result, the interval  $I$  contains a point  $\lambda_0$  such that  $\varphi''(\lambda_0) = 0$ ,  $\varphi''(\lambda) > 0$ , for  $\lambda > \lambda_0$ ,  $\lambda \in I$ , and  $\varphi''(\lambda) < 0$  for  $\lambda < \lambda_0$ ,  $\lambda \in I$ .

We saw earlier that the function  $\varphi(\lambda)$  is likewise increasing from  $-\infty$  to  $+\infty$  as  $\lambda$  traverses the interval  $I$  and that, as a result, there exists a unique point in  $I$ , say  $\alpha$ , such that  $\varphi(\alpha) = 0$ . Our aim is to find the point  $\alpha$ . It would therefore be very satisfactory if we could show that a Newton algorithm based on the iteration of the step (1.22) and starting with a value  $\lambda \in I$  will invariably converge to  $\alpha$ . While this does not seem to be true, the following, less general result is easily established.

**THEOREM A.** *Let  $I$  be the open interval between two consecutive zeros of  $\eta(\lambda)$ , and denote by  $\alpha$  the unique zero of  $\zeta(\lambda)$  in  $I$ . Then one of the two open intervals—say  $I_0$ —into which  $I$  is divided by  $\alpha$  has the following property. If  $\lambda_1$  is any point of  $I_0$ , and  $\lambda_2, \lambda_3, \dots$  are defined by the iteration*

$$\lambda_{k+1} = \lambda_k - \zeta(\lambda_k)\eta(\lambda_k), \tag{1.24}$$

*then  $\lambda_k$  converges monotonically (and quadratically) to  $\alpha$ .*

Before we prove this, we point out that, by (1.8), (1.9), and (1.10), (1.24) is equivalent to

$$\lambda_{k+1} = \lambda_k - \frac{B_n(u_k)C_n(u_k)}{\int_0^1 pu_k^2 dx}, \tag{1.25}$$

where  $u_k$  is a solution of  $Lu_k + \lambda_k pu_k = 0$  which satisfies the boundary conditions  $B_\nu(u_k) = 0$ ,  $\nu = 1, \dots, n - 1$ .

The interval  $I_0$  in which the iteration (1.24) converges can be characterized in terms of the point  $\lambda_0 \in I$  at which  $\varphi''(\lambda_0) = 0$ :  $I_0$  is the subinterval of  $I$  bounded by  $\alpha$ , which does *not* contain  $\lambda_0$  as an interior point. To prove that  $I_0$  indeed has the property stated above, suppose first that  $\lambda_0 \leq \alpha$  and that, as a result,  $I_0$  is at the right of  $\alpha$ . By the definition of  $\lambda_0$  and  $\alpha$  we then have  $\varphi(\lambda) > 0$  and  $\varphi''(\lambda) > 0$  if  $\lambda \in I_0$ . Hence, if  $\lambda_1 \in I_0$  and  $\lambda_2$  is defined by

$$\lambda_2 = \lambda_1 - \varphi(\lambda_1)/\varphi'(\lambda_1) = \lambda_1 - \zeta(\lambda_1)\eta(\lambda_1), \tag{1.26}$$

we have

$$\varphi(\lambda_2) \geq \varphi(\lambda_1) + \varphi'(\lambda_1)(\lambda_2 - \lambda_1) = \varphi'(\lambda_1)[\lambda_2 - \lambda_1 + \varphi(\lambda_1)/\varphi'(\lambda_1)]$$

i.e., by (1.26),  $\varphi(\lambda_2) \geq 0$ , and this implies  $\lambda_2 \geq \alpha$ . On the other hand, since  $\varphi(\lambda_1) > 0$  and (by (1.21)),  $\varphi'(\lambda_1) > 0$ , (1.26) shows that  $\lambda_2 < \lambda_1$ . Thus,  $\lambda_1 \in I_0$  implies  $\lambda_2 \in I_0$ , and  $\lambda_2$  is closer to  $\alpha$  than  $\lambda_1$ . Hence, the indefinite iteration of the step (1.26) will converge to a point  $\geq \alpha$ . It is obvious from (1.26) that this point must be a zero of either  $\zeta(\lambda)$  or  $\eta(\lambda)$ . Since  $\eta(\lambda) \neq 0$  in  $I$  and the only zero of  $\zeta(\lambda)$  in  $I$  is  $\alpha$ , it follows that the iteration converges to  $\alpha$ , as asserted. If  $\lambda_0 < \alpha$ , the result follows by a trivial modification of the same argument.

We were led to Theorem A in an attempt to solve the eigenvalue problem (1.1) by applying Newton's method to the equation  $\varphi(\lambda) = 0$ , where

$$\varphi(\lambda) = \frac{\zeta(\lambda)}{\eta(\lambda)} = \frac{B_n(u)}{C_n(u)}$$

(cf. (1.9), (1.10), (1.11)). The same procedure may be applied to the eigenvalue problem (1.6), which is obtained from (1.1) if the boundary condition  $B_n = 0$  is replaced by  $C_n = 0$ . Since, as shown before, the linear form "associated with"  $C_n$  is  $-B_n$  (if  $C_n$  was associated with  $B_n$ ) the only change necessary to adapt our analysis to the present case is the substitution of  $\eta(\lambda)$  for  $\zeta(\lambda)$ , and of  $-\zeta(\lambda)$  for  $\eta(\lambda)$ . In accordance with (1.11), the function  $\varphi(\lambda)$  will then be replaced by

$$\psi(\lambda) = -\frac{\eta(\lambda)}{\zeta(\lambda)} = -\frac{1}{\varphi(\lambda)}, \tag{1.27}$$

and, by (1.15), its derivative is found to be

$$\psi'(\lambda) = \frac{1}{\zeta^2(\lambda)}. \tag{1.28}$$

Thus  $\psi(\lambda)$  is increasing from  $-\infty$  to  $+\infty$  if  $\lambda$  traverses an interval  $K$  between two consecutive zeros of  $\zeta(\lambda)$  and  $\psi$  is found to have precisely one zero, say  $\beta$ , in  $K$ . Since, as noted before,  $\zeta(\lambda)$  and  $\eta(\lambda)$  have no zeros in common, this is also the only zero of  $\eta(\lambda)$  in  $K$ .

The Schwarzian derivative (1.23) remains unchanged if  $\varphi(\lambda)$  is replaced by  $-1/\varphi(\lambda)$ . Thus, by (1.23') and (1.16),  $\{\psi, \lambda\} = 2q(\lambda) > 0$ , and we have (cf. (1.23))

$$\frac{\psi'''}{\psi'} - \frac{3}{2} \left( \frac{\psi''}{\psi'} \right)^2 > 0.$$

Hence, by (1.28),  $\psi''' > 0$  and—as in the case of the function  $\varphi$ —we conclude from this that  $\psi''$  grows from  $-\infty$  to  $+\infty$  as  $\lambda$  traverses the interval  $K$ . A comparison with the proof of Theorem A shows that the properties of  $\psi$  are identical with the properties of  $\varphi$  used in proving Theorem A. If we note that, by (1.27) and (1.28),

$$\lambda - \frac{\psi(\lambda)}{\psi'(\lambda)} = \lambda + \zeta(\lambda)\eta(\lambda),$$

we have established the following companion result to Theorem A.

**THEOREM B.** Let  $K$  be the open interval between two consecutive zeros of  $\zeta(\lambda)$ , and denote by  $\beta$  the unique zero of  $\eta(\lambda)$  in  $K$ . Then one of the two open intervals—say  $K_0$ —into which  $K$  is divided by  $\beta$  has the following property: If  $\lambda_1$  is any point of  $K_0$ , and  $\lambda_2, \lambda_3, \dots$  are defined by the iteration

$$\lambda_{k+1} = \lambda_k + \zeta(\lambda_k)\eta(\lambda_k), \quad (1.29)$$

then  $\lambda_n$  converges monotonically (and quadratically) to  $\beta$ .

2. As mentioned earlier, Lagrange's identity can be written in the form

$$\int_0^1 (vLu - uLv) dx = \sum_{\nu=1}^n [B_\nu(v)C_n(u) - B_\nu(u)C_n(v)].$$

If the functions  $u$  and  $v$  satisfy the  $n - 1$  boundary conditions  $B_1 = \dots = B_{n-1} = 0$ —or, as we shall occasionally say,  $u, v \in B_{(n-1)}$ —this reduces to

$$\int_0^1 (vLu - uLv) dx = B_n(v)C_n(u) - B_n(u)C_n(v), \quad (2.1)$$

or

$$\mathcal{F}(u, v) = \mathcal{F}(v, u) \quad (2.1')$$

where

$$\mathcal{F}(u, v) = B_n(u)C_n(v) - \int_0^1 uLv dx. \quad (2.2)$$

As follows from (2.2), the symmetric bilinear form  $\mathcal{F}(u, v)$  can be written—after some preliminary integrations by parts—as an integral of a bilinear form in  $u, u', \dots, u^{(n/2)}, v, v', \dots, v^{(n/2)}$ , and, possibly, a number of boundary terms (in many of the classical problems these boundary terms are absent).

From now on, we shall restrict our considerations to operators  $L$  and boundary conditions  $B_\nu$  for which the quadratic form  $\mathcal{F}(u, u)$  is positive-definite for  $u \in B_{(n-1)}$ . We thus assume

ASSUMPTION  $A_3$ .

$$\mathcal{F}(u, u) \equiv B_n(u)C_n(u) - \int_0^1 uLu dx > 0 \quad (2.3)$$

if  $u$  satisfies the boundary conditions  $B_1(u) = \dots = B_{n-1}(u) = 0$  and does not vanish identically.

(EXAMPLES.<sup>2</sup> If  $Lu \equiv u''$ ,  $B_1(u) = u(0)$ , then  $\mathcal{F}(u, u) = \int_0^1 (u')^2 dx$ . If  $Lu \equiv -u^{iv}$ ,  $B_1(u) = u(0)$ ,  $B_2(u) = u'(0)$ ,  $B_3(u) = u(1)$  then  $\mathcal{F}(u, u) = \int_0^1 (u'')^2 dx$ .)

Our final restriction on  $L$ —and this is a very mild one—is the following:<sup>3</sup> We

<sup>2</sup>In the first example, the expression  $\mathcal{F}(u, u)$  will be as given if one selects, say,  $B_2(u) = u(1)$  and correspondingly,  $C_2(u) = u'(1)$ .

In the second example, one may take  $B_4(u) = u'(1)$ ,  $C_4(u) = u''(1) + \beta u(1)$ , to obtain the given formula for  $\mathcal{F}(u, u)$ .

<sup>3</sup>Under the present hypotheses on the coefficients of  $L$ , this appears to follow from Assumption  $A_3$  and the hypothesis described in (1.7).

assume that, within the class of functions  $u$  satisfying  $B_\nu(u) = 0, \nu = 1, \dots, n - 1, C_n(u) + \mu B_n(u) = 0$  ( $\mu > 0$  is fixed but arbitrary), the equation  $Lu = w$  ( $w$  continuous) has a unique inverse  $u = Gw$ , where  $G$  is a compact operator. (Equivalently, this property may be expressed in terms of the existence of a Green's function of  $Lu = 0$  with the above boundary conditions.)

Under all these assumptions we can prove the following result:

Let  $u \in C^n[0, 1]$ , and let  $u$  be subject to the boundary conditions  $B_\nu(u) = 0, \nu = 1, \dots, n - 1$ . Then

$$\min_u [\mathcal{J}(u, u) + \mu B_n^2(u)] = \mathcal{J}(w, w) + \mu B_n^2(w) = \lambda, \tag{2.4}$$

where  $\mu > 0$  and  $w$  is a solution of the differential equation

$$Lw + \lambda pw = 0 \tag{2.5}$$

which satisfies, in addition to the boundary conditions imposed on  $u$ , the condition  $C_n(w) + \mu B_n(w) = 0$  and  $\lambda$  is the smallest eigenvalue of (2.5) associated with these boundary conditions.

We shall further show that the higher eigenvalues of (2.5) for these boundary conditions can be defined by the classical max-min characterization associated with the minimum problem (2.4).

Suppose now that  $u$  is a function subject to the boundary conditions  $B_\nu(u) = 0, \nu = 1, \dots, n - 1$ , and the normalization

$$\int_0^1 pu^2 dx = 1. \tag{2.6}$$

We define a function  $w$  for which

$$Lw = -\lambda pu \tag{2.7}$$

and for which  $B_\nu(w) = 0, \nu = 1, \dots, n - 1, C_n(w) + \mu B_n(w) = 0$ . By our assumptions, there exists such a function  $w$  for any  $\lambda$ . We determine  $\lambda$ , and thus determine the function  $w$  completely, by the normalization

$$\int_0^1 pw^2 dx = 1. \tag{2.8}$$

From (2.7), we have

$$-\int_0^1 wLw dx = \lambda \int_0^1 puw dx,$$

and thus, by (2.6), (2.8) and (2.3),

$$\mathcal{J}(w) - B_n(w)C_n(w) \leq \lambda \quad (\mathcal{J}(w) = \mathcal{J}(w, w)). \tag{2.9}$$

Again by (2.7) and (2.6),

$$-\int_0^1 uLw dx = \lambda$$

and therefore, by (2.2),

$$\lambda = \mathcal{J}(u, w) - B_n(u)C_n(w).$$

Since the quadratic form  $\mathcal{J}(u) \equiv \mathcal{J}(u, u)$  is positive-definite we have

$$\mathcal{J}^2(u, w) \leq \mathcal{J}(u)\mathcal{J}(w),$$

and it follows from the last identity that, for arbitrary positive  $\mu$ ,

$$\lambda \leq \sqrt{\mathcal{F}(u) + \mu B_n^2(u)} \sqrt{\mathcal{F}(w) + \frac{1}{\mu} C_n^2(w)}. \quad (2.10)$$

The function  $w$  is subject to the condition  $C_n(w) + \mu B_n(w) = 0$ . If we use this number  $\mu$  in (2.10), we have

$$\mathcal{F}(w) + \frac{1}{\mu} C_n^2(w) = \mathcal{F}(w) - B_n(w)C_n(w) \leq \lambda$$

(the last inequality following from (2.9)), and (2.10) is seen to imply the right half of the inequality

$$\mathcal{F}(w) + \mu B_n^2(w) \leq \lambda \leq \mathcal{F}(u) + \mu B_n^2(u). \quad (2.11)$$

The left half of this inequality follows from (2.9) and the fact that  $C_n(w) + \mu B_n(w) = 0$ .

Our statement is now a direct consequence of the inequalities (2.11). If  $\{u_k\}$  is a sequence of admissible functions for which

$$\lim_{k \rightarrow \infty} \mathcal{F}(u_k) + \mu B_n^2(u_k) = \lambda_0 = \inf[\mathcal{F}(u) + \mu B_n^2(u)], \quad (2.12)$$

then, by (2.11), the sequence of functions  $w_k$  determined from the  $u_k$  by means of (2.7) and the prescribed boundary conditions also has the minimum property expressed by (2.12). By Assumption  $A_3$ , the  $w_k$  belong to a compact set  $S$ , and it is easy to see that there exists a subsequence of the sequence  $w_k$  which converges to a function  $w_0 \in S$  for which

$$\mathcal{F}(w_0) + \mu B_n^2(w_0) = \lambda_0. \quad (2.13)$$

Identifying  $w_0$  with the function  $u$  in (2.7) and defining a new function  $w$  by (2.7), we find that the minimum property (2.13) of  $w_0$  will be contradicted unless  $w$  and  $w_0$  are proportional. From this it follows immediately that  $w_0$  must be a solution of (2.5), and that the eigenvalue  $\lambda$  appearing in (2.5) has the meaning expressed in (2.4).

To establish the statement concerning the max-min characterization of the higher eigenvalues, it is sufficient to show that the minimum problem (2.4) will yield the  $k$ th eigenvalue, say  $\lambda_k$ , of the differential equation (2.5) if  $u$  is restricted by the conditions

$$\int_0^1 p u w_\nu dx = 0, \quad \nu = 1, \dots, k-1, \quad (2.14)$$

where  $w_\nu$  is the  $\nu$ th eigenfunction of (2.5). This, in turn, will follow by a simple argument if we can show that the transformation  $u \mapsto w$  given by (2.7) and the appropriate boundary conditions does not affect the orthogonality conditions (2.14). To do so, we note that, by (2.7) and (2.14),

$$\int_0^1 w_\nu Lw dx = 0.$$

Hence, by (2.1),

$$\int_0^1 w Lw_\nu dx = B_n(w)C_n(w_\nu) - C_n(w)B_n(w_\nu).$$

Because of  $C_n(w) + \mu B_n(w) = 0$  and  $C_n(w_\nu) + \mu B_n(w_\nu) = 0$ , this reduces to

$$\int_0^1 wLw_\nu dx = 0.$$

Finally, since  $w_\nu$  satisfies the equation (2.5) (with  $\lambda = \lambda_\nu$ ), we obtain

$$\int_0^1 pww_\nu dx = 0, \quad \nu = 1, \dots, k - 1.$$

Thus,  $w$  is found to be subject to the same orthogonality relations as  $u$ . The rest of the argument establishing our assertion is elementary.

3. [This is a completely revised and abbreviated version of Professor Nehari's original §3.]

The assertion in Theorem B has an interesting extension when the following lemma is valid.

LEMMA 3.1. *Let  $\alpha_m, \alpha_{m+1}$  be two consecutive zeros of  $\zeta(\lambda)$  ( $\alpha_m < \alpha_{m+1}$ ) and let  $\beta_{m+1}$  be the unique zero of  $\eta(\lambda)$  in  $(\alpha_m, \alpha_{m+1})$ . If  $\lambda_0 \in (\beta_{m+1}, \alpha_{m+1})$ , then*

$$\lambda_0 + \zeta(\lambda_0)\eta(\lambda_0) > \alpha_m. \tag{3.1}$$

For notation see (1.9) and (1.10) and recall that Assumption  $A_3$  is in effect (see (2.3)).

It is not difficult to see that if Lemma 3.1 holds then Theorem B implies

THEOREM 3.1. *Let  $\lambda_1$  be a positive number larger than the smallest zero of  $\eta(t)$ . If  $\zeta(\lambda_1)\eta(\lambda_1) < 0$ , then the iteration*

$$\lambda_{k+1} = \lambda_k + \zeta(\lambda_k)\eta(\lambda_k), \quad k = 1, 2, \dots, \tag{3.2}$$

*converges (quadratically) to the largest zero of  $\eta(t)$  at the left of  $\lambda_1$ .*

Relative to Theorem A the analogous assertion would be

LEMMA 3.1\*. *If  $\lambda_0 \in (\alpha_m, \beta_{m+1})$  ( $m > 1$ ) then*

$$\lambda_0 - \zeta(\lambda_0)\eta(\lambda_0) > \beta_m.$$

If this is valid then Theorem A implies the following

THEOREM 3.2. *Let  $\lambda_1$  be a positive number larger than the smallest zero of  $\zeta(t)$ . If  $\zeta(\lambda_1)\eta(\lambda_1) > 0$ , then the iteration*

$$\lambda_{k+1} = \lambda_k - \zeta(\lambda_k)\eta(\lambda_k), \quad k = 1, 2, \dots, \tag{3.3}$$

*converges (quadratically) to the largest zero of  $\zeta(t)$  at the left of  $\lambda_1$ .*

Thus, these algorithms always converge to the nearest zero of  $\zeta(\lambda)$  or  $\eta(\lambda)$ , respectively, at the left of the starting value. Since the zeros of  $\zeta(\lambda)$  and  $\eta(\lambda)$  alternate, it is possible to combine these procedures into one algorithm which gives, in descending order, all the zeros of  $\zeta(\lambda)$  and  $\eta(\lambda)$  which are smaller than the starting value  $\lambda_1$ . If, say, the algorithm (3.1) converges to a zero  $\alpha$  of  $\zeta(\lambda)$ , we start the algorithm (3.2) with a value  $\alpha - \epsilon$ , where  $\epsilon$  is a small positive number (replacing  $\alpha - \epsilon$  by  $\alpha - \epsilon/2$  if it should be found that  $\zeta(\alpha - \epsilon)\eta(\alpha - \epsilon) > 0$ ). A similar step is taken at a zero  $\beta$  of  $\eta(\lambda)$  to which the iteration (3.2) converges. This

method may, of course, miss some eigenvalues whenever a zero of  $\zeta(\lambda)$  is very close to a zero of  $\eta(\lambda)$  or, equivalently, if  $\epsilon$  is not taken small enough.

Geometrically speaking, the relation (3.1) refers to the amount by which the Newtonian iterate calculated at  $\lambda_1$  for the nearby zero  $\beta_{m+1}$  of  $\psi(\lambda)$  can overshoot this zero. Of course no overshooting at all is possible when the inflection point of the curve  $\psi(\lambda)$  lies in the subinterval  $(\alpha_m, \beta_{m+1})$  since in that case  $\psi$  is convex throughout  $(\beta_{m+1}, \alpha_{m+1})$ . This is the context of Theorem B.

In seeking how to prove Lemma 3.1 it may be worth noting the following variational description of relation (3.1). First of all, a minimax comparison permits one to deduce that under the present hypotheses the family of quadratic forms  $\{\mathcal{J}^\mu\}_{\mu \geq 0}$  defined on the subspace  $B_{(n-1)}$  (notation introduced in §2) by

$$\mathcal{J}^\mu(u) = \mu B_n^2(u) + B_n(u)C_u(u) - \int uLu \, dx, \quad u \in B_{(n-1)},$$

all possess eigenvalues which interweave one another. (Note that  $\mathcal{J}^0 = \mathcal{J}$ .) That is, denoting eigenvalues of  $\mathcal{J}^\mu$  by  $\lambda_i^\mu, i \geq 1$ , we have in addition to the obvious relation

$$0 < \mu_1 < \mu_2 \Rightarrow \lambda_i^0 \leq \lambda_i^{\mu_1} \leq \lambda_i^{\mu_2}, \quad i = 1, 2, \dots, \tag{1}$$

the further relation

$$\lambda_i^\mu < \lambda_{i+1}^0, \quad 0 \leq \mu. \tag{2}$$

Thus by reference to (2.4) and (2.5) we see that the eigenvalues of the selfadjoint operator  $L$  relative to the sets of boundary conditions

$$C_n(u) + \mu B_n(u), \quad \mu \geq 0, \quad B_k(u) = 0, \quad 1 \leq k \leq n - 1, \tag{3}$$

fill out disjoint intervals on the positive real axis. In fact, if we use  $\lambda_j^\infty$  to refer to the eigenvalues associated with what is formally obtained from (3) when  $\mu = \infty$ :

$$B_n(u) = 0, \quad B_k(u) = 0, \quad 1 \leq k \leq n - 1, \tag{3'}$$

then it follows that the nonoverlapping intervals of the positive  $\lambda$  axis denoted by

$$I_j = [\lambda_j^0, \lambda_j^\infty], \quad j = 1, 2, \dots,$$

consist entirely of eigenvalues of the forms  $\{\mathcal{J}^\mu\}_{0 < \mu < \infty}$ . (Similarly, the gaps between the intervals  $I_j$  are intervals consisting of eigenvalues of the quadratic forms  $\{\mathcal{J}^\mu\}_{\mu < 0}$ .)

Utilizing this information we see that the point  $\lambda_1$  occurring in Lemma 3.1 can be described as the eigenvalue  $\lambda_{m+1}^\mu$  for some  $\mu > 0$  and is thus associated with a normalized eigenfunction  $u_{m+1}^\mu$  of the quadratic form  $\mathcal{J}^\mu$  [and hence of the operator  $L$ ]. In this description the iteration (3.1) can be expressed by noting that

$$\lambda_0 + \zeta(\lambda_0)\eta(\lambda_0) = - \int u_{m+1}^\mu Lu_{m+1}^\mu \, dx + B_n(u_{m+1}^\mu)C_n(u_{m+1}^\mu) = \mathcal{J}(u_{m+1}^\mu).$$

Thus we have the alternate "variational" formulation

$$\mathcal{J}^0(u_{m+1}^\mu) > \alpha_m = \mathcal{J}^0(u_m^\infty). \tag{3.1'}$$

A point that was contributed by the referee is the observation that the variational Assumption  $A_3$  is crucial to the possible validity of Lemma 3.1. Otherwise as he indicated one could simply select  $\eta$  as any solution of  $\zeta'' + q\zeta = 0$  subject to the

condition

$$\zeta' \eta - \zeta \eta' = 1$$

and arrange for its zero  $\beta \in (\alpha_m, \alpha_{m+1})$  to be located so close to  $\alpha_m$  that  $\zeta(\lambda_0)\eta(\lambda_0) = -\zeta^2(\lambda_0) \int_{\beta}^{\lambda_0} d\lambda / \zeta^2(\lambda)$  would be a very large negative number.

Finally, we note that under the transition  $B_n \rightarrow C_n$ ,  $C_n \rightarrow -B_n$  the form  $\mathcal{G}(u, v)$  is replaced by

$$\mathcal{G}_*(u, v) = -B_n(v)C_n(u) - \int_0^1 uLv \, dx,$$

which is also selfadjoint on  $B_{(n-1)}$ . Setting

$$\mathcal{G}_*^\mu = C_n^2(u)/\mu + \mathcal{G}_*(u), \quad \mu > 0,$$

one finds that eigenfunctions of  $\mathcal{G}_*^\mu$  are eigenfunctions of  $L$  which satisfy the (previously omitted) boundary conditions

$$C_n(u) - \mu B_n(u) = 0, \quad \mu > 0.$$

It should be noted that the quadratic form

$$\mathcal{G}_*(u) = -B_n(u)C_n(u) - \int_0^1 uLu \, dx$$

cannot be positive-definite on  $B_{(n-1)}$ , if  $A_3$  holds.

#### REFERENCES

1. E. A. Coddington and N. Levinson, *Theory of ordinary differential equations*, McGraw-Hill, New York, 1955.

DEPARTMENT OF MATHEMATICS, CARNEGIE-MELLON UNIVERSITY, PITTSBURGH, PENNSYLVANIA 15213