Separating variables in two-way diffusion equations

Nathaniel J. Fisch
Plasma Physics Laboratory, Princeton University, Princeton, New Jersey 08544

Martin D. Kruskal
Department of Astrophysical Sciences and Program in Applied Mathematics, Princeton University, Princeton, New Jersey 08544

(Received 10 October 1979; accepted for publication 4 January 1980)

It is shown that solutions to a class of diffusion equations of the two-way type may be found by a method akin to separation of variables. The difficulty with such equations is that the boundary data must be specified partly as initial and partly as final conditions. In contrast to the one-way diffusion equation, where the solution separates only into decaying eigenfunctions, the two-way equations separate into both decaying and growing eigenfunctions. Criteria are set forth for the existence of linear eigenfunctions, which may not be found directly by separating variables. A speculation with interesting ramifications is that the growing and decaying eigenfunctions are separately complete in an appropriate half of the problem domain. This conjecture is not proved, although it does enjoy some numerical support.

I. INTRODUCTION

Many physical systems may be described by what might be called a two-way diffusion equation, which we write in the form

$$h(\theta)\frac{\partial f(x,\theta)}{\partial x} = \frac{\partial}{\partial\theta}D(\theta)\frac{\partial}{\partial\theta}f(x,\theta) \tag{1}$$

in the domain $a<\theta<b$ and $0<x<L$, with $D(\theta)$ assumed positive. If $h(\theta)$ is also positive, then Eq. (1) represents the usual diffusion equation, which is well posed when initial conditions are given at $x=0$ and boundary conditions are given at $\theta=a$ and $\theta=b$. However, in the event that $h(\theta)$ changes sign in the interval (a,b), Eq. (1) then describes diffusion towards increasing x where h is positive and diffusion towards decreasing x where h is negative. Hence, we have the nomenclature "two-way" diffusion, which is also found in the literature as "forward–backward" diffusion. These equations are then well-posed only when initial conditions are given where h is positive and final conditions (i.e., at $x=L$) are given where h is negative. Consideration of these equations occurs in the literature as early as 1913.[1]

More complicated variations of Eq. (1) may be envisioned, for example, when h and D depend on x as well as θ. However, we restrict our consideration to h, D, and boundary conditions that are independent of x, so that Eq. (1) may be approached by the method of separation of variables. Also, for simplicity, we will assume, except in Sec. VIII, that h has only isolated zeros. The goal of this paper is, in part, to examine various subtleties arising in separating variables in the two-way diffusion equation.

In practice the restriction on h and D does not exclude most cases of interest arising in physics applications. For example, particles impinging with velocity $|\mathbf{v}|$ upon an infinite slab of randomly located, small-angle, elastic point-scatterers are governed by the diffusion equation

$$\left(\frac{\partial}{\partial t} + \mathbf{v}\cdot\frac{\partial}{\partial\mathbf{r}}\right)f = \alpha\Delta_v f, \tag{2}$$

where f is the particle phase space density, α is a constant, and Δ_v is the angular Laplacian operator in velocity space. The steady-state distribution of particles along the axis of the slab is described by[2]

$$\cos\theta\frac{\partial f}{\partial x} = \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\sin\theta\frac{\partial}{\partial\theta}f, \tag{3}$$

where θ is the angle the velocity vector makes with the x axis, x being the distance along the slab axis normalized to $|\mathbf{v}|/\alpha$. The range of θ is $(0,\pi)$, so that Eq. (3) is a special case of the two-way type considered in Eq. (1).

The two-dimensional analogue of Eq. (2) occurs when particles are scattered instead by randomly located rod scatterers, the rods being oriented in one direction and parallel to the face of the slab. Instead of Eq. (3), we then get

$$\sin\theta\frac{\partial f}{\partial x} = \frac{\partial^2}{\partial\theta^2}f, \quad -\pi<\theta<\pi, \tag{4}$$

where θ is now the angle the velocity vector makes with the direction perpendicular to both the slab axis and the rod orientation. The distance along the slab axis, x, is now normalized to $\alpha/v_\perp$, where $v_\perp$ is the particle velocity perpendicular to the rod orientation and is conserved now during scattering events. Equations (3) and (4) govern what is called diffuse reflection. Proper boundary conditions would specify the incident particle distribution [corresponding, say, in Eq. (4) to $\theta>0$] at $x=0$ and would specify that only outgoing particles are present at $x=L$. A recent problem of interest governed by Eq. (4) is the scattering of plasma waves in a tokamak by random density fluctuations aligned in rodlike fashion along the magnetic field.[3]

Much of the research on two-way diffusion equations has centered around a special case of Eq. (1), namely Eq. (3), which was derived by Bothe.[2] Bethe *et al.*[4] treated this equation by using separation of variables and finding the dominant behavior from the lowest eigenfunctions. A numerical check of those conclusions has been provided by Stein and Bernstein.[5] Beals[6] analytically proved the existence of a solution to Eq. (3) and, furthermore, proved that it could be represented by the eigenfunction expansion proposed by

 0022-2488/80/040740-11$1.00

Bethe *et al.* Considerations of other two-way equations occur in Refs. 7–11.

Our approach to the general equation, Eq. (1), is to use separation of variables as in Ref. 4. Here, our concern is that the resulting eigenvalue equation is not governed by the usual Sturm–Liouville theorems, a situation that we seek to remedy. The paper is organized as follows. In Sec. II we examine when, as Bethe *et al.* found in their case, the separation of variables solutions must be supplemented by an aditional singular eigenfunction. In Sec. III we show how this singular eigenfunction may be derived from a limiting case of completely separable (nonsingular) two-way equations. In Sec. IV we show that no other singular types of solutions are possible for this class of equations. In Secs. V and VI we prove a completeness theorem on the interval (a,b) for the eigenfunctions obtained by separating variables. In Sec. VII we conjecture a further completeness property of these eigenfunctions and we appeal to, among other things, a numerical computation that lends support to the conjecture. In Sec. VIII we show how to extend our considerations to the case that h vanishes over an interval. In Sec. IX we conclude with a summary of the salient findings, including ramifications of the proved completeness theorem and the conjectured completeness property.

Before concluding this introductory section, we wish to point out that the uniqueness of the solution to Eq. (1), if it exists, is an easy matter to show via the usual energy integral. Since Eq. (1) is linear, it is satisfied by the difference, $\phi = f_1 - f_2$, of any two supposed solutions. We multiply Eq. (1) by ϕ and integrate over x and θ. Upon integrating by parts in θ on the right-hand side, the surface terms vanish for suitable boundary conditions at $\theta = a$ and $\theta = b$ [see Eq. (7)], implying that ϕ vanishes, hence also uniqueness.

II. SEPARATION OF VARIABLES

To solve Eq. (1) by the method of separation of variables, we attempt an expansion

$$f = \sum_k c_k \phi_k(x) u_k(\theta)\,, \tag{5}$$

where c_k is a constant, $\phi_k(x) = \exp(kx)$, and u_k satisfies the eigenvalue equation

$$\mathscr{L}_k u_k(\theta) \equiv \left[\frac{d}{d\theta} D(\theta) \frac{d}{d\theta} - kh(\theta) \right] u_k(\theta) = 0\,. \tag{6}$$

We assume that the boundary conditions are given such that Eq. (6) is self-adjoint, i.e., the boundary conditions are such that

$$D(a) \left(\frac{du_j(a)}{d\theta} u_k(a) - \frac{du_k(a)}{d\theta} u_j(a) \right) = D(b) \left(\frac{du_j(b)}{d\theta} u_k(b) - \frac{du_k(b)}{d\theta} u_j(b) \right). \tag{7}$$

Self-adjointness assures the existence of an orthogonality relation between the u_k with weighting function h, i.e.,

$$\int_a^b h(\theta) u_k(\theta) u_l(\theta) = 0\,, \quad \text{if } k \neq l\,. \tag{8}$$

The question arises whether the u_k found from Eq. (6) comprise a complete set of eigenfunctions. Completeness is not assured by the usual Sturm–Liouville[12] theorems, which do not apply when h vanishes in the interval (a,b). In fact, the u_k are not, in general, complete. For example, suppose that, as in Eq. (4), we pick $D = 1$, $h = \sin\theta$, and we try to use the eigenfunctions calculated from Eq. (6) to describe $f(\theta)$ such that

$$\int_{-\pi}^{\pi} f(\theta) \sin\theta \, d\theta = J \neq 0\,. \tag{9}$$

Since all the u_k are orthogonal to $\sin\theta$, any function represented as a linear combination of the u_k must have $J = 0$. Hence, an f characterized by Eq. (9) is not representable, implying that the set $\{u_k\}$ is not complete. It turns out, however, that supplementing the set $\{u_k\}$ with $\sin\theta$ does produce a complete set of functions. This assertion will be proved in Sec. V.

More generally, suppose that there exists a function g obeying the boundary conditions of the eigenvalue Eq. (6), and satisfying

$$\frac{d}{d\theta} D(\theta) \frac{d}{d\theta} g(\theta) + h(\theta) = 0\,. \tag{10}$$

The conditions for the existence, which is not assured, of such a g are well known.[12] When g exists, there may be a solution to Eq. (1) of the form $x - g(\theta)$, which we refer to as the linear or diffusion solution. This solution, which is not obtained by means of product separation of variables like the other solutions, but by sum separation, may be used to complete the u_k, so that the union of the u_k and the linear solution can represent any function of θ at a given, i.e., constant, x.

Since the linear-in-x part of the diffusion solution must also satisfy the boundary conditions, in fact at every x, only a subset of the self-adjoint boundary conditions allow diffusion solutions. It may be seen that for diffusion solutions to exist in well-posed problems, the self-adjoint boundary conditions must be restricted such that for some constant η

$$D(a) \frac{\partial f(x,a)}{\partial\theta} = D(b) \frac{\partial f(x,b)}{\partial\theta} \equiv \eta \tag{11}$$

and either

$$f(x,a) = f(x,b)\,, \quad \text{if } \eta \neq 0\,, \tag{12}$$

or suitable conditions hold on allowable f if $\eta = 0$. In this latter category falls the first example of Sec. I, i.e., Eq. (3), where $D(a) = D(b) = 0$ and f is assumed to be nonsingular. In the former category, $\eta \neq 0$, falls Eq. (4), the example discussed in this section, where periodic boundary conditions of f are assumed. In this case we have $g = \sin\theta$, so that $x - \sin\theta$ is the diffusion solution that completes the u_k.

The rule is that when h is orthogonal to all the u_k, so that it cannot be expanded in the u_k, then the solution $x - g(\theta)$ exists and may be used to represent h. Furthermore, in such a case, g is orthogonal, with weighting function h, to all the u_k except u_0. This may be demonstrated by multiplying Eq. (10) by u_k, integrating twice by parts the left-hand side, and finally substituting from Eq. (6). Invoking the orthogonality of h and u_k on the right-hand side, which is obtained by virtue of Eq. (11), then gives the desired orthogonality property for g,

$$k\int_a^b h(\theta)\, g(\theta) u_k(\theta) = 0\,, \qquad (13)$$

which is nontrivial for $k \neq 0$.

III. DEGENERACY OF THE ZEROTH EIGENVALUE

In the previous section we noted that the diffusion solution is orthogonal to all but u_0, the $k = 0$ eigenfunction. This observation naturally leads us to suspect that there may be a particularly close connection between the diffusion solution and u_0. In this section we explore this connection and derive the diffusion eigenfunction in a natural way from the limiting form of a separable equation.

For convenience, we consider a specific example, although the conclusions are general. Therefore, considering the example in Sec. I, we attempt to break the degeneracy in Eq. (4) by posing instead

$$(\epsilon + \sin\theta)\frac{\partial f}{\partial x} = \frac{\partial^2 f}{\partial \theta^2}\,, \qquad (14)$$

again with periodic boundary conditions and in the interval $(-\pi,\pi)$. For $\epsilon \neq 0$, the boundary conditions are incompatible with the existence of a diffusion solution, i.e., no g can solve Eq. (10). It naturally follows, then, to ask how the diffusion solution can arise in the limit $\epsilon \to 0$.

For $\epsilon \neq 0$, we can solve Eq. (14) by separation of variables, i.e.,

$$f = \sum_k a_k \exp(kx)\phi_k(\theta)\,, \qquad (15)$$

where ϕ_k satisfies the eigenvalue equation

$$k(\epsilon + \sin\theta)\phi_k = \phi_k''\,, \qquad (16)$$

the prime denoting differentiation with respect to θ.

We wish to prove now that there is an eigenvalue k of order ϵ, for ϵ small. This eigenvalue is in addition to the eigenvalue $k = 0$. To find this eigenvalue, we assume $k \sim \epsilon$ and formally expand

$$\phi = \phi_0 + \phi_1 + \phi_2 + \cdots\,, \qquad (17)$$

where $\phi_n \sim \epsilon^n \sim k^n$. Inserting Eq. (17) into (16), we find to zeroth order

$$\phi_0'' = 0\,, \qquad (18)$$

whereupon invoking periodicity boundary conditions, we find $\phi_0 = C_0$, where C_0 is a constant. The order-ϵ equation now gives

$$\phi_1'' = C_0 k \sin\theta\,, \qquad (19)$$

and, again invoking periodicity boundary conditions, we find

$$\phi_1 = -C_0 k \sin\theta\,, \qquad (20)$$

which is of order ϵ, as supposed. So far k has not been determined. However, the order ϵ^2 equation may now be written as

$$\phi_2'' = k\,[(\sin\theta)\phi_1 + \epsilon\phi_0] = kC_0(\epsilon - k\sin^2\theta)\,, \qquad (21)$$

which has a solution for periodicity boundary conditions only if the consistency condition

$$\int_\pi^\pi (\epsilon - k\sin^2\theta)\, d\theta = 0 \qquad (22)$$

is satisfied. From Eq. (22), we determine $k = 2\epsilon$. This allows us to find

$$\phi_2 = -C_0(\tfrac{1}{2}\epsilon^2\cos 2\theta + C_2)\,, \qquad (23)$$

where $C_2 \sim \epsilon^2$ is a constant that may be lumped in with C_0.

Thus, we have found an eigenvalue of order ϵ to Eq. (16), and the corresponding term in Eq. (15) is

$$\begin{aligned} a_k \exp(kx)\phi_k &= a_k(1 + kx + \tfrac{1}{2}k^2x^2 + \cdots)(1 - k\sin\theta \\ &\quad - \tfrac{1}{8}k^2\cos 2\theta + \cdots) \\ &= a_k[1 + kx - k\sin\theta + O(k^2)]\,, \end{aligned} \qquad (24)$$

where we have expanded the exponential term in kx, assuming that x is $O(1)$ when retaining terms. The eigenfunction in Eq. (24) is additional to the eigenfunction corresponding to $k = 0$ exactly, which is a constant. Thus, we may use the $k = 0$ eigenfunction to substract out the constant part of Eq. (24), i.e., a_k, the remaining part also satisfying Eq. (14). We write this remaining part as $a_k k\,[x - \sin\theta + O(k)]$ and take the limit $k \to 0$ while taking $a_k \sim 1/k$. Thus, higher-order terms in k drop out and we are left with the diffusion eigenfunction, $x - \sin\theta$, as found in Sec. II.

Finally, we note that before removing the constant term in Eq. (24) and taking the limit $\epsilon \to 0$, all eigenfunctions are orthogonal with respect to the weighting function, $\epsilon + \sin\theta$. Taking the limit $\epsilon \to 0$ may be viewed as the merging of two eigenvalues or a degeneracy in the zeroth eigenvalue. Thus, the two eigenfunctions remain orthogonal to the remaining eigenfunctions. They may also be made orthogonal to each other at any x, but not simultaneously at all x.

These arguments may be applied to the general case, where h is altered perturbatively in such a manner that g no longer exists. This removes the degeneracy in the lowest eigenvalue, and taking the limit of zero perturbation recovers the diffusion solution in a manner entirely analogous to the case presented. That the resulting set of eigenfunctions, the u_k plus the diffusion solution, is complete is still not assured, although this property is now somewhat motivated by the observations on the degeneracy as a limiting case. In the next section we show that other degeneracies, not in the zeroth eigenvalue, are impossible. The proof of the completeness property is reserved for Sec. V.

IV. SIMPLICITY OF THE NONZERO EIGENVALUES

The results of the previous section concerning the degeneracy of the zeroth eigenvalue point naturally to the possibility of degeneracy in the nonzero eigenvalues also. If such a degeneracy were to occur, then by analogy to the degeneracy already studied, we may expect solutions of the form

$$f = g_1(\theta)x\exp(kx) + g_2(\theta)\exp(kx)\,. \qquad (25)$$

Substituting into Eq. (1), we see that g_1 and g_2 must satisfy

$$\mathscr{L}_k g_1(\theta) = 0\,, \qquad (26a)$$

$$\mathscr{L}_k g_2(\theta) = h(\theta) g_1(\theta)\,. \qquad (26b)$$

We assume boundary conditions that allow diffusion solutions, i.e.,

$$f(a) = f(b)\,, \qquad (27a)$$

$$D(a)\frac{\partial f(a)}{\partial \theta} = D(b)\frac{\partial f(b)}{\partial \theta}, \tag{27b}$$

although we prove now that such solutions cannot occur for $k \neq 0$.

Note that g_1 satisfies the usual eigenvalue equation, i.e., Eq. (6), and thus has the usual property

$$k\int_a^b hg_1^2\, d\theta = -\int_a^b D\left(\frac{dg_1}{d\theta}\right)^2 d\theta. \tag{28}$$

Now we also find

$$\int_a^b (g_2\mathscr{L}_k g_1 - g_1\mathscr{L}_k g_2)\, d\theta = -\int_a^b hg_1^2\, d\theta = 0, \tag{29}$$

where the second equality is obtained because the left-hand side of Eq. (29) vanishes upon application of the boundary conditions after the obvious integration by parts. Equation (29), which is a necessary condition for any degenerate solution, implies that h must change sign. Furthermore, Eqs. (28) and (29) together imply that

$$\int_a^b D\left(\frac{dg_1}{d\theta}\right)^2 d\theta = 0, \tag{30}$$

which, in turn, implies that g_1 must be constant, since D is not zero at any interior points. However, if g_1 is constant, then k must be zero to satisfy Eq. (26a). Hence, all nonzero eigenvalues are nondegenerate or simple, and degenerate solutions can occur only for $k = 0$, i.e., they are of the type considered in the previous section. The necessary and sufficient conditions for their occurrence are that the boundary conditions permit the constant solution to be an allowable eigenfunction and that Eq. (29) hold, which may now be written simply as

$$\int_a^b h(\theta)\, d\theta = 0. \tag{31}$$

Although we have shown that merging of roots may occur only at $k = 0$, we have not yet limited the number of roots which may merge, i.e., the order of the degeneracy. For example, if n roots were to merge at $k = 0$, then a solution of the form

$$f = x^n + x^{n-1}g_1(\theta) + \cdots + xg_{n-1}(\theta) + g_n(\theta) \tag{32}$$

could exist, where the coefficient of x^n must obviously be independent of θ as taken. We show now that multiple degeneracies of the form of Eq. (32) cannot exist for $n > 1$. For Eq. (32) to satisfy Eq. (1), the g_n must satisfy

$$nh = \frac{d}{d\theta}\frac{Ddg_1}{d\theta}, \tag{33a}$$

$$(n-l)hg_l = \frac{d}{d\theta}\frac{Ddg_{l+1}}{d\theta}, \quad 1 \leqslant l \leqslant n, \tag{33b}$$

and obey boundary conditions of the type given by Eqs. (27). From Eq. (33a) we find that

$$\int_a^b D\left(\frac{dg_1}{d\theta}\right)^2 d\theta = -n\int_a^b hg_1\, d\theta = 0, \tag{34}$$

where the second equality arises from use of Eq. (33b) when $n \neq 1$ and implies that $dg_1/d\theta = 0$, which is impossible in view of Eq. (33a). Hence there cannot exist any solutions of the type given by Eq. (32) for $n > 1$.

In summary the rule is that there is only one possible degeneracy, namely, the merging of two roots at $k = 0$. Corresponding to this degeneracy, there can be at most one independent diffusion solution of the form $x - g(\theta)$. [Any other diffusion solution, say $x - y(\theta)$, is a linear combination of the above solution and the u_0 eigenfunction, since $g - y$ satisfies Eq. (6) with $k = 0$.]

V. PROOF OF CLOSEDNESS

In the next two sections, we prove that any function defined on the interval (a,b) and obeying the boundary conditions at $\theta = a$ and $\theta = b$ may be written as a linear combination of the u_k and g, which are found from

$$\mathscr{L}_k u_k(\theta) = 0, \tag{35a}$$

$$\mathscr{L}_0 g(\theta) = h(\theta), \tag{35b}$$

where $\mathscr{L}_k$ is defined in Eq. (6) and the boundary conditions are given by Eq. (11) and either Eq. (12) or a suitable replacement as discussed in Sec. II. The completeness proof that we offer is motivated by the method of Kneser,[13] which was useful in proving completeness for proper Sturm–Liouville problems. First we prove a "closedness" property, i.e., if any function is orthogonal to all of the eigenfunctions, then it must be zero. The completeness property follows from the closedness property. Kneser's idea, which may be found within the context of subsidiary theorems in Ref. 12, is to construct a series solution to a related inhomogeneous problem. Information concerning the closedness of the eigenfunctions of the homogeneous problem then follows from the convergence properties of the series.

Motivated in this manner, we first consider the inhomogeneous equation

$$\frac{d}{d\theta}D(\theta)\frac{d}{d\theta}v(\theta) + kh(\theta)v(\theta) + p(\theta) = 0, \tag{36}$$

with boundary conditions imposed on v that allow diffusion solutions to the related homogeneous equation, i.e., Eq. (36) with $p = 0$. We construct a solution to Eq. (36) by means of the expansion

$$v(\theta) = v_0 + kv_1 + \cdots + k^n v_n + \cdots, \tag{37a}$$

where the terms of the series are found from

$$\frac{d}{d\theta}D\frac{d}{d\theta}v_0 + p = 0, \tag{37b}$$

$$\frac{d}{d\theta}D\frac{d}{d\theta}v_n + hv_{n-1} = 0, \quad n \geqslant 1, \tag{37c}$$

with the v_n obeying the same boundary conditions that v obeys.

The radius of convergence of the series solution is bounded by $\rho = |l|$, where l is the smallest (in absolute value) eigenvalue of the homogeneous system for which

$$\int_a^b p(\theta)u_l(\theta)d = 0 \tag{38}$$

does not hold. If no such l exists, then $\rho = \infty$. For $|k| < \rho$, the series will converge (e.g., Sec. 11.3 in Ref. 12 with minor modification assures this) so long as successive v_n may be found unambiguously. The condition on the existence of the series' coefficients v_n imposes restrictions on p in addition to those already imposed by Eq. (38). The analysis now departs somewhat from that for a proper Sturm–Liouville problem, for which the existence of the v_n is assured without further restrictions on p.

First, we note that upon integrating Eqs. (37b) and (37c) over the interval (a,b) and applying the boundary condition in Eq. (11), we obtain

$$\int_a^b p(\theta)\, d\theta = 0\,, \tag{39a}$$

$$\int_a^b h(\theta) v_n(\theta)\, d\theta = 0\,, \tag{39b}$$

the latter relation holding for all n. The restriction on p given by Eq. (39a) is necessary for the convergence of the series anywhere and is equivalent to Eq. (38) with $l = 0$, where then u_l is the zeroth eigenfunction, which is a constant when the boundary conditions allow diffusion solutions. Nevertheless, the necessity of satisfying Eq. (39a) means that v_0 cannot be fully determined from Eq. (37b) alone. Similarly, the compatibility condition, Eq. (39b), means that Eq. (37c) alone is insufficient to determine v_n unambiguously. It may be suspected, however, that the compatibility conditions provide sufficient additional restrictions to determine the v_n. We now set out to prove that, in fact, this is so.

Multiplying Eq. (37c) by g and integrating yields

$$\begin{aligned}\int_a^b g h v_{n-1}\, d\theta &= -\int_a^b g\left(\frac{d}{d\theta} D \frac{d}{d\theta} v_n\right) d\theta \\ &= \int_a^b \left(\frac{dv_n}{d\theta}\right) D \frac{dg}{d\theta}\, d\theta \\ &= -\int_a^b v_n \left(\frac{d}{d\theta} D \frac{dg}{d\theta}\right) d\theta \\ &= -\int_a^b v_n h\, d\theta\,,\end{aligned} \tag{40}$$

where, upon each integration by parts, the boundary terms vanished by virtue of the assumed boundary conditions. The last equality was written on the basis of substitution from Eq. (35b). Now by virtue of Eq. (39b), the right-hand side of Eq. (40) vanishes, from which we obtain the orthogonality of g and the v_n with weighting function h.

Although g is orthogonal to the v_n, it is not orthogonal to the constant function, as may be shown by multiplying Eq. (35b) by g and integrating once by parts to obtain

$$\int_a^b hg\, d\theta = -\int_a^b D\left(\frac{dg}{d\theta}\right)^2 d\theta \neq 0\,. \tag{41}$$

The inequality above obtains because D does not pass through zero and g cannot be constant.

It may be seen that each v_n is determined from Eq. (37c) up to an arbitrary additive constant, say of the form

$$v_n = C + R(\theta)\,, \tag{42}$$

where $R(\theta)$ is a known function and C is an unknown constant. We show now that Eqs. (40) and (41) are sufficient to determine C and resolve the ambiguity in the v_n. Substituting for v_n from Eq. (42) into Eq. (40) and using Eqs. (39b) and (41), we construct

$$C = \int_b^a Rhg\, d\theta \left[\int_a^b D\left(\frac{dg}{d\theta}\right)^2 d\theta\right]^{-1}. \tag{43}$$

The point is that C is always determined because the denominator cannot vanish. It may be noted that simply substituting v_n from Eq. (42) directly into the compatibility condition, Eq. (39b), would not determine C since h has zero area.

It should be noted that, in the above, nonzero k has been tacitly assumed in the application of the orthogonality relations, i.e., Eqs. (39b) and (40). When k vanishes, these relations, and hence Eq. (43) also, do not necessarily hold and C cannot be uniquely determined. In fact, when $k = 0$, the solution to Eq. (36) is not unique and can be determined only up to an arbitrary constant. Similarly, an additional contraint is needed to uniquely determine the solution to Eq. (37b), which is identical to Eq. (36) with $k = 0$. Here, nearly any additional constraint will do; we could, for example, assume that Eq. (43) holds also for $k = 0$ in which case a unique solution to Eq. (36) exists, even for $k = 0$, and may be constructed by means of the series for $|k| < \rho$.

Finally, multiplying Eq. (37b) by g, integrating, and manipulating the subsequent expression in a manner analogous to Eq. (40), we obtain

$$\int_a^b gp\, d\theta = 0\,, \tag{44}$$

which represents an additional constraint on p if the series is to have a nonzero radius of convergence. It should be appreciated that this condition is essentially anticipated by the resemblance of Eq. (44) to Eq. (38) in view of the discussion given in Sec. III.

To recapitulate, we have so far shown that if Eqs. (39a) and (44) hold, then the series solution converges in a finite interval about $k = 0$, namely for $|k| < \rho$. The series coefficients may be determined uniquely order by order by means of Eqs. (37) and (43). We now proceed to exploit the fact that if $\rho = \infty$, i.e., if Eq. (38) holds for all l, then v is an entire function of k.

We consider two cases of Eq. (37c), say

$$\frac{d}{d\theta} D \frac{d}{d\theta} v_n + h v_{n-1} = 0\,, \tag{45a}$$

$$\frac{d}{d\theta} D \frac{d}{d\theta} v_{m+1} + h v_m = 0\,. \tag{45b}$$

Multiplying Eq. (45a) by v_{m+1} and Eq. (45b) by v_n, subtracting, and integrating, we obtain

$$\int_a^b (v_{m+1} v_{n-1} - v_n v_m) h\, d\theta = 0\,, \tag{46}$$

where we made use of the boundary condition, Eq. (11). We may note that since n and m were arbitrarily chosen, the definition

$$W_q \equiv \int_a^b v_n v_m h\, d\theta\,, \quad n + m = q\,, \tag{47}$$

is unambiguous.

For proper Sturm–Liouville problems, the proof proceeds somewhat more simply as W_q is always non-negative for q even, since h is non-negative. This simplification does not occur in our case, where h does, in fact, pass through zero. We can, however, show that W_q is always non-negative for q odd. We do so by multiplying Eq. (45a) by v_n and integrating to obtain

$$0 \leqslant \int_a^b D\left(\frac{dv_n}{d\theta}\right)^2 d\theta = \int_a^b h v_n v_{n-1}\ d\theta \equiv W_{2n-1}\ . \tag{48}$$

We may also consider the quantity

$$\delta \equiv \int_a^b \left[\alpha\left(\frac{dv_n}{d\theta}\right) + \beta\left(\frac{dv_m}{d\theta}\right)\right]^2 d\theta$$
$$= \alpha^2 W_{2n-1} + 2\alpha\beta W_{m+n-1} + \beta^2 W_{2m-1}\ , \tag{49}$$

where α and β are arbitrary constants. Since δ is non-negative for any choice of α and β, we must have

$$W^2_{m+n-1} \leqslant W_{2n-1}\ W_{2m-1}\ . \tag{50}$$

Let q be an odd integer. From Eq. (50) we see that if $W_q = 0$ for any q, then all the W_q must be zero, except possibly for W_1. We now, in fact, show that W_1 must then also be zero, although this is not implied by Eq. (50). Using instead Eq. (48) with $n = 2$, we see that $W_3 = 0$ implies that $dv_2/d\theta = 0$, which, in turn, implies from Eq. (45a) with $n = 2$ that $hv_1 = 0$, which requires that W_1 vanish too. Hence, either all the W_q vanish or none do. Suppose first that none do. Since by Eq. (48) $W_q \geqslant 0$, Eq. (50) implies the inequalities

$$\frac{W_3}{W_1} \leqslant \frac{W_5}{W_3} \leqslant \frac{W_7}{W_5} \leqslant \cdots \leqslant \frac{W_{2n+1}}{W_{2n-1}} \leqslant \cdots\ , \tag{51}$$

which in turn implies that

$$W_{2n+1} \geqslant T^n W_1\ , \tag{52}$$

where we have defined

$$T \equiv W_3/W_1 > 0\ . \tag{53}$$

Now when v is an entire function of k [i.e., when p is orthogonal to all the u_l and g, with a constant as the weighting function, as in Eqs. (38) and (44)], then so is the quantity

$$\int_a^b v v_0 h\ d\theta = W_0 + kW_1 + \cdots + k^n W_n + \cdots\ . \tag{54}$$

The assumption $W_q \neq 0$, however, leads to a contradictory statement, since the sum of the subseries of all the odd terms of the series in Eq. (54),

$$\sum_{m=0}^{\infty} k^{2m+1} W_{2m+1} \geqslant kW_1 \sum_0^{\infty} (k^2 T)^m\ , \tag{55}$$

clearly diverges for some k. Hence, we have proved by contradiction that $W_q = 0$ and, in particular, that $W_1 = 0$, which implies that

$$0 = W_1 \equiv \int_a^b v_1 v_0 h\ d\theta = \int_a^b D\left(\frac{dv_1}{d\theta}\right)^2 d\theta\ , \tag{56}$$

which means that v_1 must be a constant. This latter statement, in turn, implies from Eq. (37c) that $hv_0 = 0$, which finally implies from Eq. (37b) that $p = 0$.

What we have proved is that if p is orthogonal to all the u_k and g, then $p = 0$. This closedness property is shown in the next section to imply completeness, i.e., that any suitable p may be constructed as a linear combination of the u_k and g in the manner

$$p(\theta) = \sum_k c_k u_k(\theta) + c_g g(\theta)\ , \tag{57}$$

which is the completeness relation that we seek. The constants c_k and c_g may then be determined easily from the orthogonality properties of the eigenfunctions. We multiply Eq. (57) by either h, hu_l, or hg and integrate to obtain, respectively,

$$c_g \int_a^b hg\ d\theta = \int_a^b hp\ d\theta\ , \tag{58a}$$

$$c_l \int_a^b hu_l^2\ d\theta = \int_a^b hpu_l\ d\theta\ , \quad l \neq 0\ , \tag{58b}$$

$$c_0 \int_a^b hg\ d\theta = \int_a^b p\, gh\ d\theta - c_g \int_a^b g^2 h\ d\theta\ . \tag{58c}$$

The constants are all determined from Eqs. (58) since all the integrals on the left-hand sides have been shown not to vanish. The case $l = 0$ is excluded in Eq. (58b) for just that reason, since h has zero area, and instead use is made of Eq. (58c) to determine c_0.

VI. PROOF OF COMPLETENESS

In order to prove that the closedness of the eigenfunctions implies their completeness, we begin by pointing out that the asymptotic behavior of large eigenvalues is given by

$$k_n \sim n|n|\ , \quad n = 0\ , \pm 1, \pm 2, \cdots\ , \tag{59}$$

where n indexes the k_n and spans $(-\infty, \infty)$. In contrast, for Sturm–Liouville problems, $k_n \sim n^2$ where n spans only $(0, \infty)$. The validity of Eq. (59) can be demonstrated by means of matched asymptotic expansions. For example, suppose that $H \equiv h/D$ passes through zero only once in the interval (a,b), say at $\theta = 0$, with finite positive slope $C > 0$. We can then solve Eq. (6) away from $\theta \approx 0$ by means of a WKB expansion, which matches onto Airy functions near $\theta = 0$. Doing so for $k \to +\infty$, we find $u_k = \alpha u_A(\theta) + \beta u_B(\theta)$, where α and β are constants to be determined from the boundary conditions, and

$$u_A(\theta) = \begin{cases} \dfrac{1}{2\sqrt{\pi}} H^{-1/4} \exp\left(-k^{1/2}\displaystyle\int_0^\theta [H(\theta')]^{1/2} d\theta'\right), & \theta > 0, \\ \mathrm{Ai}(k^{1/3} C^{1/2} \theta), & \theta \approx 0, \\ \dfrac{1}{\sqrt{\pi}} [-H]^{-1/4} \sin\left(k^{1/2}\displaystyle\int_\theta^0 [-H(\theta')]^{1/2} d\theta' + \frac{\pi}{4}\right), & \theta < 0, \end{cases} \tag{60a}$$

$$u_B(\theta) = \begin{cases} \dfrac{1}{\sqrt{\pi}} H^{-1/4} \exp\left(k^{1/2}\displaystyle\int_b^\theta [H(\theta')]^{1/2} d\theta'\right), & \theta > 0, \\ 0, & \theta \lesssim 0. \end{cases} \tag{60b}$$

The point is that in order to satisfy self-adjoint boundary conditions, $k^{1/2}$ is determined only up to multiples of $2\pi\Gamma$, where Γ is a finite constant, whence Eq. (59) follows. For example, for periodicity boundary conditions, we have

$$\frac{\beta}{\alpha} \simeq \left[\frac{-H(b)}{H(a)} \right]^{1/4} \sin\left(k^{1/2} \int_a^0 (-H)^{1/2}\, d\theta + \frac{\pi}{4} \right), \tag{61}$$

and k is determined from

$$\tan\left(k^{1/2} \int_a^0 (-H)^{1/2}\, d\theta + \frac{\pi}{4} \right) = \left[\frac{-H(b)}{H(a)} \right]^{1/2}. \tag{62}$$

We now turn to the question of uniform convergence. Consider the functions which represent the decomposition of f into orthogonal modes, i.e.,

$$F_n(\theta) = u_n(\theta) \int_a^b f h u_n \, d\theta , \tag{63}$$

where u_n is the eigenfunction corresponding to k_n and is normalized according to

$$\int_a^b h u_n^2 = 1 . \tag{64}$$

Note that from the asymptotic representation of the u_n for large n, we see that $H^{1/4}u_n$ is uniformly bounded for all n and θ.

Suppose that f obeys the same boundary conditions as the u_n and is twice differentiable. From Eq. (6) we have

$$\begin{aligned} F_n &= k_n^{-1} u_n \int_a^b f \frac{d}{d\theta} D \frac{d}{d\theta} u_n \, d\theta \\ &= k_n^{-1} u_n \int_a^b u_n \frac{d}{d\theta} D \frac{d}{d\theta} f \, d\theta . \end{aligned} \tag{65}$$

By virtue of the boundedness of $H^{1/4}u_n$ and the integrability of the weak singularity (i.e., since $u_n \sim H^{-1/4}$) in the integrand on the right side of Eq. (65), it is seen that

$$H^{1/4}F_n \sim 1/k_n \sim 1/n^2 , \quad \text{as } n \to \pm\infty , \tag{66}$$

where we now made use of Eq. (59). Hence, the series of partial sums of $H^{1/4}F_n$ is absolutely and uniformly convergent in the interval (a,b).

We are now in a position to demonstrate that closedness implies at least a mild type of completeness, i.e., completeness in that any continuous twice-differentiable function f may be uniformly approximated. This restrictive or narrow property is needed before we relax the conditions on f to include all continuous functions. Proceeding in this vein, we define the function

$$\begin{aligned} \phi(\theta) &\equiv f(\theta) - \sum_k u_k(\theta) \int_a^b u_k(\theta')h(\theta')f(\theta')\, d\theta' \\ &\equiv f(\theta) - \sum_k u_k(\theta) a_k . \end{aligned} \tag{67}$$

Consider the quantity

$$\begin{aligned} \int_a^b h u_l \sum_k u_k a_k &= \int_a^b h u_l |h|^{-1/4} \sum_k |h|^{1/4} u_k a_k \\ &= \sum_k a_k \int_a^b h u_l u_k = a_l , \end{aligned} \tag{68}$$

where reversing the order of integration and summation to obtain the second equality is clearly allowable since the series is uniformly convergent by virtue of Eq. (66). Thus, we have for every u_l

$$\int_a^b \phi h u_l \, d\theta = \int_a^b f h u_l - a_l = 0 , \tag{69}$$

which implies from the closedness of the u_l that $\phi h = 0$. Hence, from Eq. (67) we see that the u_k can represent f everywhere except possibly at the isolated point $\theta = 0$, which is the mild completeness property we sought to prove. It is now possible to follow relatively standard but tedious procedures to show that the above property implies that, in fact, any continuous f, not necessarily twice-differentiable, may be uniformly approximated by the u_n (except where $\theta = 0$). This last step of the completeness proof is given in Appendix A.

Note that this proof relied on a finite first derivative of h at $\theta = 0$ in order to write down the asymptotic expansion in terms of Airy functions. Nevertheless, it is easy to show that similar conclusions may be drawn when this derivative vanishes or when more general (but self-adjoint) boundary conditions are used. The case of h vanishing over an interval is discussed in Sec. VIII.

We now show that the proved completeness property on the interval (a,b) guarantees that if a solution to Eq. (1) exists, then it can be expanded in the separation-of-variables eigenfunctions. For f may then be put in the form

$$f = \sum_k A_k(x) u_k(\theta) + B(x)[x - g(\theta)] \tag{70}$$

and it suffices to show that the A_k and B are in fact given by

$$A_k(x) = c_k e^{kx} \tag{71a}$$

$$B = -c_g , \tag{71b}$$

where the c_k and c_g are constants found, say, from evaluating Eqs. (58) at $x = 0$. Equations (58) may now be used to establish Eqs. (71). For example, using Eq. (58b) and assuming the u_l have been normalized, we have for $k \neq 0$

$$A_k(x) = \int_a^b f h u_k \, d\theta , \tag{72}$$

and differentiating by x gives

$$\begin{aligned} \frac{d}{dx} A_k(x) &= \int_a^b h u_k \frac{df}{dx} \, d\theta \\ &= \int_a^b u_k \frac{d}{d\theta} D \frac{d}{d\theta} f \, d\theta \\ &= \int_a^b f \frac{d}{d\theta} D \frac{d}{d\theta} u_k \, d\theta \\ &= \int_a^b k f h u_k \, d\theta \\ &= k A_k(x) , \end{aligned} \tag{73}$$

where in deriving the third equality we integrated twice by parts. Thus, Eq. (71a) follows for $k \neq 0$. In a similar manner, $A_0(x)$ and $B(x)$ may be shown to be constant as required.

VII. COMPLETENESS ON THE HALF-INTERVAL

In Sec. VI, we showed that the u_k are complete on the interval (a,b). This implied that the expansion by separation of variables could be used to represent f everywhere, when f exists.

However, the existence of f for arbitrary initial and final conditions indicates that the u_k possess a far stronger property than completeness on the interval (a,b). By taking the limit $d \to \infty$, it can be seen that only the non-growing-in-x eigenfunctions can contribute at $x = 0$, while only the non-decaying eigenfunctions can contribute at $x = d$. Thus, assuming that h vanishes only once, say, $h(0) = 0$, we see that the exponentially decaying eigenfunctions, supplemented by the constant and linear eigenfunctions, should be complete on the interval where h is positive, say in $(0,b)$. Similarly, the growing eigenfunctions, also supplemented by the constant and linear eigenfunctions, should be complete in the interval $(a,0)$. (There is no problem in switching the order of taking the limit $d \to \infty$ with taking the limit of the number of eigenfunctions becoming infinite, since the higher order eigenfunctions, decaying or growing most rapidly, certainly cannot contribute at both boundaries.)

Thus, it may be seen that existence of the solution and its representability by the eigenfunction expansion should imply the further completeness property of the u_k, that "half" the u_k are complete in the interval $(0,b)$ while the other half are complete in the interval $(a,0)$. Furthermore, it may presumably be shown that the converse is also true, i.e., if the u_k possess this completeness property, then f must exist. This should follow from a construction of f by means of a convergent, iterative scheme, where the boundary conditions are alternatively satisfied by the decaying set of u_k at $x = 0$ and the growing set at $x = d$.

Unfortunately, we have not been able to prove the completeness of the eigenfunctions on the half-interval, which would have, from the above argument, presumably provided a general and independent proof of the existence of solutions to Eq. (1). We must therefore content ourselves, at present, with the reverse argument. Thus, in special cases, where we may rely on other proofs for existence, then we can infer that the associated eigenfunctions possess a completeness property on the half-inteval.

To demonstrate how the eigenfunctions may be used to construct the solution, and to provide additional support for the conjecture that they are complete on the half-interval, we numerically consider one example of Eq. (1), namely

$$\sigma(\theta)\frac{\partial f}{\partial x} = \frac{\partial^2 f}{\partial \theta^2}, \quad -\pi<\theta<\pi,\ 0<x<d, \qquad (74)$$

where

$$\sigma(\theta) = \begin{cases} 1, & \theta>0, \\ -1, & \theta<0, \end{cases} \qquad (75)$$

and periodicity boundary conditions are assumed. This particular example was chosen for numerical analysis because the eigenfunctions are particularly simple. Since $\sigma(\theta)$ has zero area, we expect a linear eigenfunction of the form $x - u_{00}(\theta)$, where

$$u_{00}(\theta) = \begin{cases} \theta(\pi-\theta), & \theta>0, \\ \theta(\pi+\theta), & \theta<0. \end{cases} \qquad (76)$$

The eigenfunctions that are even about $\pi/2$ in the interval $(0,\pi)$ are of the form, for $n>0$,

$$u_n(\theta) = \begin{cases} \cos[\lambda_n(\theta-\pi/2)], & \theta>0, \\ C_n\cosh[\lambda_n(\theta+\pi/2)], & \theta<0, \end{cases} \qquad (77)$$

where

$$C_n = [\cos(\lambda_n\pi/2)][\cosh(\lambda_n\pi/2)]^{-1} \qquad (78)$$

and λ_n solves

$$\cos\lambda_n\pi = \operatorname{sech}\lambda_n\pi. \qquad (79)$$

For $n<0$, i.e., for the decaying eigenfunctions of the form $u_n(\theta)\exp(-\lambda_n x)$, we have $u_n(\theta) = u_{-n}(-\theta)$. Similarly, there are eigenfunctions that are odd about $\theta = \pi/2$ in the interval $(0,\pi)$. Since we will assume even boundary conditions, the odd functions need not be considered here.

Part of the reason for giving these eigenfunctions in detail is that as $n \to \infty$ the u_n exponentially fast approach $\cos[(2n+\frac{1}{2})(\theta-\pi/2)]$. If the u_n are complete on $(0,\pi)$, which we will demonstrate numerically, it is expected that the $\cos[(2n+\frac{1}{2})(\theta-\pi/2)]$, supplemented by a constant or roughly constant function, must also be complete on $(0,\pi)$. Nevertheless, despite its simple form, we have been unable to analytically demonstrate that the set $\cos[(2n+\frac{1}{2})(\theta-\pi/2)]$ is complete, which presumably might be an easier task than to demonstrate that property for the u_n.

In Fig. 1 we show the results of numerically fitting the eigenfunction expansion to boundary data using the method of least squares. The boundary conditions are that $f(\theta<0, x=d) = 0$ and $f(\theta>0, x=0)$ has the Gaussian type of dependence shown in Fig. 1(a). Here, we have taken $d = 10$. We find that the root-mean-square difference between the eigenfunction expansion and the given bundary data converges to zero as $1/N$, where N is the number of eigenfunctions employed. This type of convergence is indicative of the presence of Gibb's phenomenon in this problem. Indeed, some such phenomenon is expected since the solution on the boundary cannot be analytic over the whole interval $(-\pi,\pi)$. For example, at $x = d$, there cannot exist a smooth fit of any non-zero solution in the region $\theta>0$ to the zero data given in the region $\theta<0$. (Parenthetically, we remark that in analogy with the examples given in Sec. I, the zero data correspond to a condition of only outgoing particles

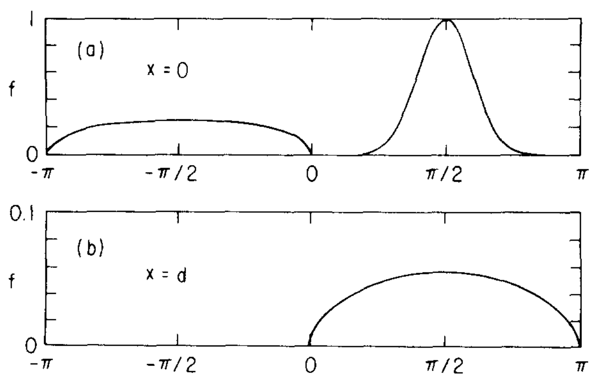

FIG. 1. Data on the boundaries. (a) $f(\theta)$ vs θ on the boundary at $x = 0$; for $\theta>0$, the data is given, whereas for $\theta<0$ the data is computed, using 250 eigenfunctions. (b) $f(\theta)$ vs θ on the boundary at $x = 10$; for $\theta<0$, the data is given, whereas for $\theta>0$ the data is computed. Note change of scale.

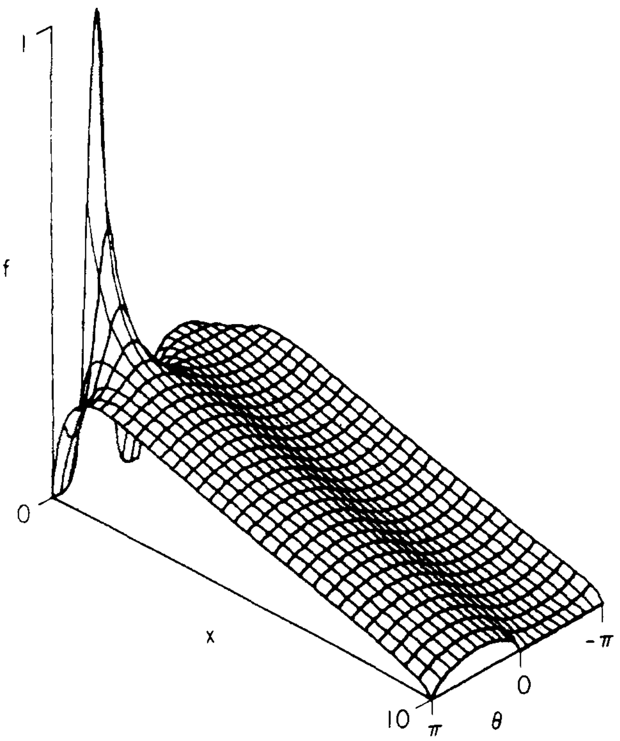

FIG. 2. Surface of $f(\theta,x)$ using the boundary data given in Fig. 1.

present at one end of the collision region.) In Fig. 2 we show, using a large number of eigenfunctions, f as a function of both x and θ. Note that, except near the boundaries, f is described mainly by the linear solution.

Numerical results for large d are similar. For very large d, only the decaying eigenfunctions remain finite. Thus, there is numerical evidence, at least in this case, for the conjectured completeness property on the half-interval.

A remark, appropriate before closing this section, is that the eigenfunctions u_n, which are supposed complete on $(0,\pi)$, pass through zero n times in that interval and not at all in $(-\pi,0)$. The reverse holds true for those eigenfunctions complete on the other half-interval.

VIII. WHEN $h(\theta)$ VANISHES OVER AN INTERVAL

When $h(\theta)$ vanishes over an interval, it turns out that the closedness and completeness properties of the u_k do not extend over the full interval (a,b). Instead, it may be shown that these properties extend over the interval (a,b) with the exclusion of the subintervals over which h vanishes. For example, in deducing the closedness property following Eq. (56), we relied on the vanishing of hv_0. Were h to vanish over an interval, then we could conclude only that p vanishes where h does not. Similarly, the conjectures regarding properties on the half-interval would then apply only where h is either positive or negative.

That the completeness property fails to apply where $h = 0$ presents no difficulty in describing the solution to Eq. (1) by means of an eigenfunction expansion. In fact, it can be shown via an energy integral (in the manner described at the end of Sec. I) that specifying boundary conditions at $x = 0$ where $h > 0$ and at $x = d$ where $h < 0$ implies uniqueness of the solution. Hence, nothing can be specified where $h = 0$ anyway, which is eminently consistent with an eigenfunction expansion wherein the eigenfunctions are complete on the interval that excludes the set where $h = 0$.

IX. SUMMARY AND CONCLUSIONS

In a manner of summary, we remark that the important features in this work are the proof of completeness of the $u_k(\theta)$ on (a,b), the criteria for the existence of a diffusion solution, and the conjecture regarding the completeness of half of the $u_k(\theta)$ on the interval $(a,0)$ or $(0,b)$.

The completeness of the eigenfunctions $u_k(\theta)$ on the interval (a,b), which was proved in Secs. V and VI, guarantees that when a solution to Eq. (1) exists it may be expanded in the form given by Eq. (5). It should be pointed out, however, that it may not always be practical to employ the expansion in numerically solving Eq. (1), since the u_k themselves may be hard to compute and, once computed, do not enjoy useful orthogonality properties for properly posed boundary value problems. Nevertheles, we have seen that when the u_k are easily found, as in the example given in Sec. VII, the eigenfunction expansion is certainly convenient.

The criteria for the existence of the diffusion solution allow the gathering of partial information about the solution without obtaining it completely. For example, the existence of a diffusion solution to Eq. (3) implies the well-known fact that the amount of sunlight reaching earth through a nonabsorbing cloud layer drops off only as the reciprocal of the layer thickness rather than with an exponential dependence on it.

Finally, the conjecture regarding the half-interval problem perhaps has the farthest reaching implications of all and its verification is worthy of future investigation. The correctness of the conjecture should imply an independent proof of the existence of a solution to the related two-way diffusion equation. There is the added academic interest in that the conjecture relates to functions which do not satisfy a Sturm–Liouville equation on the interval upon which the completeness property is supposed. This is in contrast to the other findings here, which may be viewed somewhat as a supplement or extension to the standard Sturm–Liouville theory.

The proof of the conjecture regarding completeness on the half-interval has now been provided by R. Beals.[14]

ACKNOWLEDGMENT

The work of one of the authors (NJF) was supported by the United States Department of Energy Contract No. EY-76-C-02-3073.

APPENDIX A

In this appendix we show that the capability of the eigenfunction expansion to represent continuous twice-differentiable f implies the capability to also represent any f that is merely continuous. The proof of this step, which supplements the completeness proof in Sec. VI, follows Ref. 12, Sec. 11.52. The idea is to compare the expansion in the u_n with a uniformly convergent expansion in a known set of complete orthogonal functions, χ_n. For simplicity, we consider the intervals $(a,0)$ and $(0,b)$ separately. Assume that the χ_n are defined on $(0,b)$ and are orthogonal with respect to the weighting function $\nu(\theta)>0$. We define the χ_n to be zero for $\theta<0$. Then we can define the partial sums

$$S_N(\theta)\equiv\int_a^b f(t)\sum_{-N}^{N} u_n(\theta)u_n(t)h(t)\,dt, \tag{A1a}$$

$$\sigma_N(\theta)\equiv\int_a^b f(t)\sum_{0}^{N}\chi_n(\theta)\chi_n(t)\nu(t)\,dt, \tag{A1b}$$

where $\sigma_N(\theta)$ is known to converge uniformly to $f(\theta)$ in the interval $(0,b)$, and we would like to prove the same for $S_N(\theta)$. Thus, we will try to show that

$$S_N(\theta)-\sigma_N(\theta)\to 0 \tag{A2}$$

uniformly in $(0,b)$ as $N\to\infty$.

To expedite matters, we define the function

$$\Phi_N(\theta,t)\equiv\sum_{n=0}^{N}[u_n(\theta)u_n(t)+u_{-n}(\theta)u_{-n}(t) - \chi_n(\theta)\chi_n(t)], \tag{A3}$$

where, for simplicity of notation, the diffusion solution, if it exists, is not written explicitly but is understood to be included in the summation. Now if $G(\theta)$ is continuous and twice differentiable, then

$$\int_a^b \Phi_N(\theta,t)G(t)\to 0 \tag{A4}$$

uniformly in $(0,b)$ as $N\to\infty$. This is a consequence of the uniform convergence of S_N to such G (proved in Sec. VI). The same property holds for the σ_N by assumption. Equation (A4) is obtained since their difference must then uniformly converge to zero.

Consider the sequence of functions

$$G_1, G_2, \ldots, G_n,$$

such that the G_n are continuous, twice-differentiable functions that uniformly converge to f in $(0,b)$. The G_n could be, for example, n-term polynomial or Fourier approximations to f. We may then write

$$S_N(\theta)-\sigma_N(\theta)=\int_a^b\Phi_N(\theta,t)[f(t)-G_m(t)]\,dt + \int_a^b\Phi_N(\theta,t)G_m(t)\,dt. \tag{A5}$$

If $\Phi_N(\theta,t)$ is uniformly bounded for all N,θ, and t, then the first integral in Eq. (A5) may be made arbitrarily small by taking m large enough. Then, by Eq. (A4), the second integral can be made arbitrarily small by taking N large enough. Hence, Eq. (A2) follows.

It remains to show that $\Phi_N(\theta,t)$ is, in fact, uniformly bounded. For simplicity, we begin with a specific example, namely that considered in Eq. (74). We take the χ_n to be $\cos[2n(\theta-\pi/2)]$. Since only functions even about $\pi/2$ are considered here, it suffices to examine the interval $(0,\pi/2)$. Consider one of the contributions to Φ_N, which may be written as

$$A\equiv\sum_{N_0}^{N}\cos[(2n+\tfrac12)\theta]\cos[(2n+\tfrac12)t] - \cos(2n\theta)\cos(2nt)$$

$$=\sum_{N_0}^{N}\tfrac12\{\cos[(2n+\tfrac12)\psi]-\cos(2n\psi)\} + \tfrac12[\cos(2n+\tfrac12)z-\cos(2nz)], \tag{A6}$$

where $\psi\equiv\theta+t$ and $z\equiv\theta-t$. It is assumed that N_0 is a sufficiently large eigenvalue, so that the asymptotic representation of the u_n as $\cos(2n+\frac12)\theta$ is valid. We note, for example, that

$$\sum_{N_0}^{N}\cos(2n+\tfrac12)\psi-\cos2n\psi=\mathrm{Re}\left(\frac{e^{i\psi/2}-1}{e^{2N_0i\psi}-1}(e^{2Ni\psi}-1)\right), \tag{A7}$$

which is obviously uniformly bounded. Thus A is bounded.

The other contribution to $\Phi_N(\theta,t)$ arises from the decaying eigenfunctions in the interval $(0,b)$, and may be written as

$$B\equiv\sum_{N_0}^{N}u_{-n}(\theta)u_{-n}(t)\sim\sum_{N_0}^{N}\exp[-(2n+\tfrac12)(\theta+t)], \tag{A8}$$

which is uniformly bounded for θ in any *closed* interval in $(0,\pi/2)$ that does not include zero. Since $\Phi_N(\theta,t)$ $=A+B+$ (a finite number of terms), it follows that $\Phi_N(\theta,t)$ is similarly uniformly bounded in any closed interval not including zero.

This property of $\Phi_N(\theta,t)$ implies, through Eq. (A5), that $S_N(\theta)$ uniformly converges to any continuous $f(\theta)$ in any closed interval in $(0,\pi/2)$ not including the origin. The exclusion of the origin [i.e., where $h(\theta)=0$] from the completeness proof is expected, as in other aspects of this problem. In a similar manner, it can be shown that $S_n(\theta)$ converges to $f(\theta)$ in the rest of the interval $(-\pi,\pi)$. Note that for periodicity boundary conditions, the endpoints, $\theta=\pm\pi$, must be excluded in the same manner and for the same reason that the point $\theta=0$ is excluded from the interval over which the completeness holds.

The general case in which h passes through zero at $\theta=0$ is handled similarly. From the asymptotic expansion of the u_n, given by Eqs. (60), it can be seen that the u_n asymptotically approach sinusoidal or decaying functions. The Airy function behavior occurs only near the origin, so that for any closed interval not including the origin, it is possible to begin the summation at a large enough N_0, as in Eqs. (A7) and (A8), that the asymptotic behavior is valid. From the example given above, it is clear that the Φ_N would be similarly bounded. For one may, for example, pick the χ_n such that

$$\frac{d}{d\theta}D\frac{d}{d\theta}\chi_n+k_n|h|\chi_n=0, \tag{A9}$$

with Sturm–Liouville-type boundary conditions at the endpoints of either interval, $(0,b)$ or $(a,0)$. The eigenfunctions χ_n would then be asymptotically sinusoidal, but shifted from the u_n, a case known from Eq. (A7) to have the requisite properties.

[1] M. Gevrey, J. Math. Sér. **6**, 9 (1913) and 10 (1914).
[2] W. Bothe, Zeits f. Phys. **54**, 161 (1929).
[3] A. Sen and N.J. Fisch, Massachusetts Institute of Technology MIT—PRR 78/16 (1978).

[4]H.A. Bethe, M.E. Rose, and L.P. Smith, Proc. Amer, Philos. Soc. **78**, 573 (1938).
[5]D. Stein and I.B. Bernstein, Phys. Fluids **19**, 811 (1976).
[6]R. Beals, J. Math. Anal. Appl. **58**, 32 (1977).
[7]R. Beals, J. Funct. Anal. (to be published).
[8]C.D. Pagani, Rend. Inst. Lombardo **103**, 618 (1969).
[9]W. Fleming, U.S. Army Math. Res. Center, Univ. of Wisc., M.R. C. Rep. (1963).
[10]M.S. Baovendi and P. Grisvard, J. Funct. Anal. **2**, 352 (1968).
[11]C.D. Pagani and G. Talenti, Ann. Mat. Pura Appl. **90**, 1 (1971).
[12]E.L. Ince, *Ordinary Differential Equations* (Dover, New York, 1956).
[13]A. Kneser, Mathematische Ann. **58**, 136 (1904).
[14]R. Beals, J. Math. Phys. (to be published).