JOURNAL OF COMPUTATIONAL PHYSICS 14, 40–58 (1974)

An Invariant Imbedding Algorithm for the Solution of Inhomogeneous Linear Two-Point Boundary Value Problems

RICHARD C. ALLEN, JR. AND G. MILTON WING

Department of Mathematics and Statistics, The University of New Mexico, Albuquerque, New Mexico 87131

Received May 18, 1973

An algorithm is developed for the numerical solution of inhomogeneous linear two-point boundary value problems using the method of invariant imbedding. The method handles the case in which the standard Riccati equation of the imbedding method fails to have a solution over the entire interval of interest. Very general sets of boundary conditions are imposed. While the development is for scalar problems only, all steps generalize readily to the case of matrix equations. An efficient program implementing the algorithm has been written and several examples are solved by the method. Some of these are of such a nature that more standard integration schemes provide unsatisfactory results.

1. INTRODUCTION

The method of invariant imbedding has been under investigation for over thirty years, originally as a device for analyzing and solving problems in transport theory and radiative transfer, and more recently as a technique which has much more far reaching implications. In particular, it has recently been studied as a method of numerically solving linear two-point boundary problems. The structure of the basic equations involved in the application of the method is such as to sometimes provide faster, more accurate, and more stable integration of such problems than is possible with many classical algorithms.

However, the method itself suffers from the fact that one of the fundamental functions involved, a "reflection" function, satisfies a nonlinear Riccati-type equation. The solution to this equation may fail to exist over the interval of interest. A variety of attempts to avoid this difficulty have been made. Perhaps the most successful, in so far as homogeneous problems are concerned, is that based upon the use of certain recursive equations [1, 2].

The question immediately arises as to whether a similar device is available for the treatment of inhomogeneous problems, and, if so, if the resulting numerical method possesses any possible advantages. In this paper we shall develop such an algorithm and show by giving a number of numerical examples that it does indeed appear to be advantageous at times. It should be pointed out that many of the formulas obtained are similar to those found by Denman [3, 4, 5]. However, his approach might well be described as being "differential" in nature while ours is basically recursive.

In Section 2 we investigate the "well-behaved" case of the inhomogeneous problem. By "well-behaved" we mean that the troublesome reflection function is completely defined by the Riccati equation mentioned earlier. The results here are by no means new [6, 7]. However, it is valuable to go through this development briefly in order to realize the difficulties that can arise in the "ill-behaved" case, that in which the reflection function is not completely defined by its usual Riccati equation. The insight provided aids in finding a device which overcomes the problems.

Section 3 is somewhat technical. It proves that there is always a finite set of Riccati equations which, together with the recursive equations, do indeed completely define the troublesome reflection function. It is the existence of this set that makes our method viable.

In Section 4 we derive the additional recursion formulas which, together with the standard differential equations of invariant imbedding, provide the desired algorithm. Section 5 calls attention to the fact that in the case of periodic problems all calculations are considerably simplified. We leave the derivations of several equations to the reader.

Section 6 seems almost out of place in this paper, but it is rather difficult to find the results indicated written down in a convenient place. (Again, Denman [3] gives some of them, and a very through treatment, in a somewhat abstract setting, is provided by Golberg [8].) Historically, as was mentioned earlier, invariant imbedding was first extensively used in the study of transport phenomena. There the physics is such that the two-point boundary value problem arising has the value of one function assigned at one end of the system while the value of the second function is specified at the other end. Because of this background, most problems to which the method has been applied have had this feature. In Section 6 we point out that this structure is by no means necessary. We can examine problems in which the value of the same function is specified at both end points; we can also study initial value problems. Thus results of this paper, as well as those of many others, are far more widely applicable than is often supposed to be the case.

Numerical results are presented in Section 7, and comparisons are made with other methods of calculation. Two of the examples investigated are of the "unusual" sort mentioned in Section 6. Section 8 summarizes the results of the paper and suggests further directions for research.

In [9] it is concluded that the Riccati approach (basically equivalent to the invariant imbedding method) is often inferior to the very classical superposition method simply because the solutions to the Riccati equations become very large, eventually exceeding the capacity of the computer. Several examples are given in [9] which exhibit this property. Whether the solutions become large because the problem under study is truly "ill-behaved" or whether the functions involved become large despite the fact that the problem is actually "well-behaved" is not relevant. The methods of this paper are applicable in either situation. The conclusion reached in [9] is less valid than indicated there. While it is unwise to claim that any numerical algorithm is *always* superior to another, the technique we describe certainly extends the usefulness of the imbedding method greatly.

Throughout this section we have spoken of "functions." For ease of presentation we shall confine our study entirely to the scalar case. Extension to the case of matrix equations is fairly obvious (although frequently tedious). The possibility of using similar methods on more complicated functional equations also suggests itself.

2. An Algorithm for the Well-Behaved Case

We consider the two-point boundary value problem

$$(d/dz)\, u(z) = A(z)\, u(z) + B(z)\, v(z) + S^{+}(z), \tag{2.1a}$$

$$-(d/dz)\, v(z) = C(z)\, u(z) + D(z)\, v(z) + S^{-}(z), \tag{2.1b}$$

$$u(x) = u_x, \qquad v(y) = v_y, \qquad x \leqslant z \leqslant y. \tag{2.1c}$$

Here all functions are scalars, $A(z)$, $B(z)$,..., $S^{+}(z)$, and $S^{-}(z)$ are continuous on $x \leqslant z \leqslant y$, and u_x and v_y are given constants. Moreover, we suppose that all of the standard reflection and transmission functions of invariant imbedding exist and satisfy their respective differential equations over the entire interval $x' \leqslant z \leqslant y$ for any x', $x \leqslant x' < y$. For convenience we list one set of these differential equations [10]:

$$\begin{aligned}(d/dz)\, R_r(x', z) &= B(z) + [A(z) + D(z)]\, R_r(x', z) + C(z)\, R_r^2(x', z),\\ R_r(x', x') &= 0;\end{aligned} \tag{2.2a}$$

$$\begin{aligned}(d/dz)\, T_r(x', z) &= [D(z) + C(z)\, R_r(x', z)]\, T_r(x', z),\\ T_r(x', x') &= 1;\end{aligned} \tag{2.2b}$$

$$\begin{aligned}(d/dz)\, T_l(x', z) &= [A(z) + C(z)\, R_r(x', z)]\, T_l(x', z),\\ T_l(x', x') &= 1;\end{aligned} \tag{2.2c}$$

$$\begin{aligned}(d/dz)\, R_l(x', z) &= C(z)\, T_l(x', z)\, T_r(x', z),\\ R_l(x', x') &= 0, \qquad x \leqslant x' \leqslant z \leqslant y.\end{aligned} \tag{2.2d}$$

It will be observed that only the Riccati equation (2.2a) can be a source of difficulty; all the other equations are linear equations sequentially dependent upon the solution to (2.2a). Our hypothesis of good behavior implies that this Riccati equation is soluble for $R_r(x', z)$ for all $x \leqslant x' \leqslant z \leqslant y$.

The set of equations (2.2) completely defines the reflection and transmission functions. Another set (see [11]) which is sometimes useful allows integration in the reverse direction, i.e., defines the quantities $R_r(z, y')$, $T_r(z, y')$, etc. One may think of these as "backwards differential equations."

In the homogeneous case, $S^+(z) = S^-(z) \equiv 0$, the basic functional relationship which makes the imbedding approach so useful is (see, e.g., [12])

$$u(z) = u_x T_l(x, z) + v(z) R_r(x, z), \tag{2.3a}$$

$$v(z) = u(z) R_l(z, y) + v_y T_r(z, y). \tag{2.3b}$$

To take account of the S-functions in the more general inhomogeneous case we amend these relationships and attempt to find functions g and h such that

$$u(z) = u_x T_l(x, z) + v(z) R_r(x, z) + g(x, z), \tag{2.4a}$$

$$v(z) = u(z) R_l(z, y) + v_y T_r(z, y) + h(z, y). \tag{2.4b}$$

Obviously this pair of equations can be solved for $u(z)$ and $v(z)$. Since the R and T functions are defined by (2.2), our problem is solved completely if the functions g and h can be found.

Let us differentiate (2.4a) with respect to z. Using a somewhat condensed notation we get

$$u' = u_x T_l' + v' R_r + v R_r' + g'. \tag{2.5}$$

After use of Eqs. (2.1) and considerable manipulation, we have

$$\begin{aligned} &u_x(T_l' - CT_l R_r - AT_l) + v(-CR_r^2 - DR_r + R_r' - AR_r - B) \\ &\quad = -g' + CgR_r + S^- R_r + Ag + S^+. \end{aligned} \tag{2.6}$$

The coefficients of u_x and of v are identically zero by virtue of Eqs. (2.2). Thus we must require of g that it satisfies

$$(d/dz)\, g(x, z) = [A(z) + C(z) R_r(x, z)]\, g(x, z) + S^-(z) R_r(x, z) + S^+(z). \tag{2.7}$$

From Eq. (2.4a) we obtain $u(x) = u_x + g(x, x)$ and so

$$g(x, x) = 0. \tag{2.7a}$$

That Eqs. (2.7, 2.7a) have a unique solution follows from the linearity and the properties of the coefficient functions involved.

Equation (2.4b) may now be treated similarly to obtain a differential equation satisfied by $h(z, y)$. Since we have listed only the equations for the R and T functions with arguments (x', z), it may be a bit more consistent to find an expression for $h(x', z)$. To obtain such an expression we rewrite (2.4b) in the form

$$v(x') = u(x')\, R_l(x', z) + v(z)\, T_r(x', z) + h(x', z) \tag{2.8}$$

and differentiate with respect to z:

$$0 = u(x')\, R_l'(x', z) + v'(z)\, T_r(x', z) + v(z)\, T_r'(x', z) + h'(x', z). \tag{2.9}$$

Further use of Eqs. (2.1) and considerable manipulation lead eventually to

$$(d/dz)\, h(x', z) = [S^-(z) + C(z)\, g(x', z)]\, T_r(x', z) \tag{2.10}$$

and

$$h(x', x') = 0. \tag{2.10a}$$

We observe that the equation for h involves initial argument x' where x' is quite arbitrary, $x \leqslant x' \leqslant z \leqslant y$. The equation for g involves the initial argument x. A little thought shows that (2.7) may also be stated for $g(x', z)$. It is this ability to write equations involving functions with relatively arbitrary arguments which makes the imbedding method useful.

We have now obtained a complete set of equations for all the functions involved in the determination of $u(z)$ and $v(z)$ (Eq. (2.4)). Our assumption that this is the well-behaved case insures that all these equations may be integrated. As with the R and T functions, one can derive differential equations for g and h with arguments (z, y'). It is important to note that none of the R, T, g, or h functions is in any way dependent upon the boundary conditions (2.1c).

3. The Ill-Behaved Case. A Theoretical Analysis

As mentioned in the Introduction, we shall refer to the Problem (2.1) as "ill-behaved" when the Riccati equation for $R_r(x', z)$ fails to have a solution over the entire interval $x \leqslant x' \leqslant z \leqslant y$. Clearly, when this happens, none of the equations for the other vital functions is meaningful and the method outlined in Section 2 ceases to be valid. However, it is known that the function R_r can be extended by means of the recursive equations, as can R_l and the T functions (see [1, 2] and Eqs. 4.2 of this paper). It seems reasonable to suppose that similar recursion formulas should be obtainable for the functions g and h.

There is first a theoretical question to be answered. Suppose that the function $R_r(x, z)$, obtained by integration of (2.2a), has a singular point at $\zeta_1 < y$. One may restart the integration with $R_r(\zeta_1^-, \zeta_1^-) = 0$, $\zeta_1^- < \zeta_1$ and continue until another singular point $\zeta_2 < y$ is encountered. Again one may restart with $R_r(\zeta_2^-, \zeta_2^-) = 0$, $\zeta_2^- < \zeta_2$, and continue the process to ζ_3, etc. It is completely possible that the points $\zeta_1, \zeta_2, \zeta_3, \ldots$, have a point of accumulation $z^* < y$. If this should happen the scheme must stop and there is no hope, even with the recursive equations, of proceeding beyond this point of accumulation. In this event the effort to use the invariant imbedding analysis would be doomed to failure. One might argue that in such a case one could turn to the "backward equations," those whose functions have arguments (z, y') and work to the left from y. But the same phenomenon might occur, and if the accumulation point were to the right of z^* there would be an interval of z points inaccessible to the method.

We shall now prove that such a point of accumulation does not exist. To do so, we must recall to mind what the function $R_r(z', z'')$ *really* is. It is the value of $u(z'')$, the solution of (2.1a, b) with $S^+ = S^- \equiv 0$, when these equations are subject to the conditions $u(z') = 0$ and $v(z'') = 1$. It is always possible that this problem has no solution, and this is reflected in the fact that $R_r(z', z'')$ fails to exist. This can *only* occur if the problem (2.1a, b) with $S^+ = S^- \equiv 0$, $u(z') = 0$ and $v(z') = 1$ has a zero of v at z''. This can be seen as a relatively easy consequence of the fundamental existence and uniqueness theorem for initial value problems. Thus our analysis is thrown back to the study of initial value problems, a much more comfortable domain.

It is important at this point not to try to dispense with the whole question with a simple reference to this fundamental theorem. A little thought shows that if the R_r function does cease to exist at a point ζ_1 and one starts a new integration for $R_r(\zeta_1^-, z)$, then he is looking at a *completely new* problem. Thus the study must involve a *sequence* of initial value problems, not just one. We first prove a fundamental lemma.

LEMMA. *Consider the problem*

$$u' = A(z)\, u + B(z)\, v, \tag{3.1a}$$

$$-v' = C(z)\, u + D(z)\, v, \tag{3.1b}$$

where $A, B, \ldots$, are continuous in the interval $x \leqslant z \leqslant y$. Let z' be in this interval, $z' < y$. Then the problem (3.1a, b) *subject to the initial conditions*

$$u(z') = 0, \qquad v(z') = 1, \tag{3.1c}$$

has no zeros of v in the interval $z' \leqslant z \leqslant z''$ where $z'' - z' = \delta$, and δ is a number independent of z'.

Remark. It is important to understand the meaning of the lemma. It really says that there is a *constant* δ with the property that the initial value problem (3.1) cannot have a zero of v in the interval $(z', z' + \delta)$ regardless of how we choose z'. Thus the corresponding R_r problem over this interval can be resolved by simply solving its Riccati equation. Therefore, the accumulation point z^* alluded to earlier in this section cannot exist.

Proof of the lemma. Denote by A_0 the maximum value of $|A(z)|$ in $x \leqslant z \leqslant y$, with similar meanings for B_0, etc. Suppose there is no such δ. Then we can choose z', $x \leqslant z' < y$ such that the first zero of v is arbitrarily close to z'. Call this zero $\tilde{z}$, $z' < \tilde{z}$.

Let $v(z)$ achieve its maximum value in $[z', \tilde{z}]$ at $\hat{z}$; call this value $\hat{v}$. By the mean value theorem,

$$v'(\bar{z}) = (0 - v(\hat{z}))/(\tilde{z} - \hat{z}) = \hat{v}/(\hat{z} - \tilde{z}) < 0, \tag{3.2}$$

where $\hat{z} < \bar{z} < \tilde{z}$.

From Equation (3.1b)

$$|v'(z)| \leqslant C_0 |u| + D_0 |v| . \tag{3.3}$$

Thus at the point $\bar{z}$

$$\hat{v}/(\tilde{z} - z') \leqslant \hat{v}/(\tilde{z} - \hat{z}) = |v'(\bar{z})| \leqslant C_0\hat{u} + D_0\hat{v} \tag{3.4}$$

where $\hat{u}$ is the maximum value of $|u|$ in $z' \leqslant z \leqslant \tilde{z}$.

Now turn to Eq. (3.1a) subject to $u(z') = 0$. It is readily solved in terms of v:

$$u(z) = \int_{z'}^{z} B(s)\, v(s) \exp\left[-\int_{z'}^{s} A(t)\, dt\right] ds. \tag{3.5}$$

Thus in the interval $z' \leqslant z \leqslant \tilde{z}$,

$$\max |u(z)| = \hat{u} \leqslant (\tilde{z} - z')\, B_0\hat{v}e^{(\tilde{z}-z')A_0} . \tag{3.6}$$

From Eq. (3.4),

$$\hat{v}/(\tilde{z} - z') \leqslant C_0(\tilde{z} - z')\, B_0\hat{v}e^{(\tilde{z}-z')A_0} + D_0\hat{v}. \tag{3.7}$$

However, $\hat{v} > 0$. Therefore,

$$1/(\tilde{z} - z') \leqslant C_0(y - x)\, B_0e^{(y-x)A_0} + D_0 . \tag{3.8}$$

The right side of Eq. (3.8) is completely independent of z' and $\tilde{z}$. But we assumed that z' was chosen so that $\tilde{z}$ was arbitrarily close to it, thus making the left side of

(3.8) arbitrarily large. This is an obvious contradiction and the Lemma is proved. It is interesting to note that the right side of (3.8) provides a bound on δ.

THEOREM. *Let the problem* (3.1) *subject to the conditions* $u(x) = 0, v(y) = 1$ *be soluble. Then there exist a finite number of points,* z_i, $x = z_0 < z_1 < \cdots < z_n = y$, *such that the functions* $R_r(x, z_i)$, $i = 1, 2,..., n$ *and the functions* $R_r(z_i, y)$, $i = n - 1, n - 2,..., 0$ *all* exist *and such that the functions* $R_r(z_i, z_{i+1})$, $i = 0, 1,..., n - 1$ *all exist and can be found by integrating the Riccati equations for* $R_r(z_i, z)$ *from* z_i *to* z_{i+1}.

Proof. We begin by calling attention to the fact that no claim is made that the $R_r(x, z_i)$ (nor $R_r(z_i, y)$) can be obtained by solving Riccati equations. To see that these functions exist, consider the initial value problem defined by (3.1) and $u(x) = 0, v(x) = 1$. The solution of this problem certainly exists over $[x, y]$. Moreover, the function v can have only a finite number of zeroes on this interval; call them $x < t_1 < t_2 < \cdots < t_k < y$. At any other point z' in the interval the two-point boundary value problem with $u(x) = 0$, $v(z') = 1$ must then be soluble. Hence the function $R_r(x, z')$ exists. Similar reasoning shows that $R_r(z'', y)$ exists for any z'' not in any exceptional finite set $x < w_1 < w_2 < \cdots < w_j < y$.

Now we need only choose a set of points z_i such that the distance from z_i to z_{i+1} does not exceed δ and such that no z_i coincides with a t or w point. Clearly, this can be done. The Lemma then asserts that the functions R_r have the requisite properties.

Without stating a formal corollary we observe that all of the functions $R_r(x, z_i)$, $R_r(z_i, z_{i+1})$, $T_r(z_i, y)$, $T_r(z_i, z_{i+1})$, etc., are obtainable either through the integration of differential equations or through the use of the recursive equations.

4. THE BASIC RECURSION FORMULAS

We are now in a position to derive the complete set of recursion formulas that constitute the computing algorithm we are seeking. In the light of the previous section we may assume that the various R and T functions are known through integration over any interval $[z_i, z_{i+1}]$, $i = 0, 1, 2,..., n - 1$. By virtue of Eqs. (2.7) and (2.10) the g and h functions may then also be assumed known over these intervals. For the time being we suppose that we are interested in the values of u and v at a point z which is itself a z_i point. (Recall from the work of the previous section that only a finite number of z points cannot be z_i points.) Later, even this requirement will be dropped.

Using Eqs. (2.4), but evaluating all functions at more convenient points, we obtain

$$u(z_i) = [u(z_0)\, T_l(z_0, z_i) + v(z_n)\, R_r(z_0, z_i)\, T_r(z_i, z_n) + R_r(z_0, z_i)\, h(z_i, z_n) + g(z_0, z_i)][1 - R_r(z_0, z_i)\, R_l(z_i, z_n)]^{-1}, \quad (4.1a)$$

$$v(z_i) = [v(z_n)\, T_r(z_i, z_n) + u(z_0)\, R_l(z_i, z_n)\, T_l(z_0, z_i) + R_l(z_i, z_n)\, g(z_0, z_i) + h(z_i, z_n)][1 - R_r(z_0, z_i)\, R_l(z_i, z_n)]^{-1}. \quad (4.1b)$$

Here we know that all functions exist except perhaps the g and h functions. While we could discuss their existence now, we prefer to defer the matter briefly. That the expression $1 - R_r(z_0, z_i)\, R_l(z_i, z_n) \neq 0$ is shown in [2]. Clearly to obtain $u(z_i)$ and $v(z_i)$ we need only find the quantities $R_r(z_0, z_i)$, $R_l(z_i, z_n)$, $T_r(z_i, z_n)$, $T_l(z_0, z_i)$, $g(z_0, z_i)$, and $h(z_i, z_n)$. The first four functions may be found by adroit use of the classical recursive equations. We list these below for reference (see [1, 2, 3]).

$$R_r(\alpha_1, \alpha_3) = \frac{T_l(\alpha_2, \alpha_3)\, T_r(\alpha_2, \alpha_3)\, R_r(\alpha_1, \alpha_2)}{1 - R_r(\alpha_1, \alpha_2)\, R_l(\alpha_2, \alpha_3)} + R_r(\alpha_2, \alpha_3), \quad (4.2a)$$

$$T_r(\alpha_1, \alpha_3) = \frac{T_r(\alpha_1, \alpha_2)\, T_r(\alpha_2, \alpha_3)}{1 - R_r(\alpha_1, \alpha_2)\, R_l(\alpha_2, \alpha_3)}, \quad (4.2b)$$

$$R_l(\alpha_1, \alpha_3) = \frac{T_l(\alpha_1, \alpha_2)\, T_r(\alpha_1, \alpha_2)\, R_l(\alpha_2, \alpha_3)}{1 - R_r(\alpha_1, \alpha_2)\, R_l(\alpha_2, \alpha_3)} + R_l(\alpha_1, \alpha_2), \quad (4.2c)$$

$$T_l(\alpha_1, \alpha_3) = \frac{T_l(\alpha_1, \alpha_2)\, T_l(\alpha_2, \alpha_3)}{1 - R_r(\alpha_1, \alpha_2)\, R_l(\alpha_2, \alpha_3)}. \quad (4.2d)$$

The choice $\alpha_1 = z_0$, $\alpha_2 = z_{k-1}$, $\alpha_3 = z_k$ yields

$$R_r(z_0, z_k) = \frac{T_l(z_{k-1}, z_k)\, T_r(z_{k-1}, z_k)\, R_r(z_0, z_{k-1})}{1 - R_r(z_0, z_{k-1})\, R_l(z_{k-1}, z_k)} + R_r(z_{k-1}, z_k), \quad (4.3a)$$

$$T_r(z_0, z_k) = \frac{T_r(z_0, z_{k-1})\, T_r(z_{k-1}, z_k)}{1 - R_r(z_0, z_{k-1})\, R_l(z_{k-1}, z_k)}, \quad (4.3b)$$

$$R_l(z_0, z_k) = \frac{T_l(z_0, z_{k-1})\, R_r(z_0, z_{k-1})\, R_l(z_{k-1}, z_k)}{1 - R_r(z_0, z_{k-1})\, R_l(z_{k-1}, z_k)} + R_l(z_{k-1}, z_k), \quad (4.3c)$$

$$T_l(z_0, z_k) = \frac{T_l(z_0, z_{k-1})\, T_l(z_{k-1}, z_k)}{1 - R_r(z_0, z_{k-1})\, R_l(z_{k-1}, z_k)}, \quad (4.3d)$$

for $k = 2, 3, \ldots, n$. This set of coupled first-order difference equations may now be iterated to obtain $R_r(z_0, z_i)$ and $T_l(z_0, z_i)$.

Next, the selection $\alpha_1 = z_i$, $\alpha_2 = z_{k-1}$, $\alpha_3 = z_k$ produces

$$R_r(z_i, z_k) = \frac{T_l(z_{k-1}, z_k)\, T_r(z_{k-1}, z_k)\, R_r(z_i, z_{k-1})}{1 - R_r(z_i, z_{k-1})\, R_l(z_{k-1}, z_k)} + R_r(z_{k-1}, z_k), \quad (4.4)$$

$$g(z_0, z_n) = \frac{\ldots}{1 - R_r(z_0, z_{n-1})\, R_l(z_{n-1}, z_n)} [g(z_0, z_{n-1}) + R_r(z_0, z_{n-1})\, h(z_{n-1}, z_n)] + g(z_{n-1}, z_n). \quad (4.7)$$

Finally we recognize that the index n was chosen only as a matter of notational convenience. The point z_n in Eq. (4.7) can be replaced by any z_k point. Hence,

$$g(z_0, z_k) = \frac{T_l(z_{k-1}, z_k)}{1 - R_r(z_0, z_{k-1})\, R_l(z_{k-1}, z_k)} [g(z_0, z_{k-1}) + R_r(z_0, z_{k-1})\, h(z_{k-1}, z_k)] + g(z_{k-1}, z_k) \quad (4.8)$$

for $k = 2, 3, \ldots, n$. The structure of this simple difference equation shows that $g(z_0, z_i)$ does indeed exist.

We leave the derivation of the equation for h to the reader:

$$h(z_i, z_k) = \frac{T_r(z_i, z_{k-1})}{1 - R_r(z_i, z_{k-1}) R_l(z_{k-1}, z_k)} [h(z_{k-1}, z_k) + R_l(z_{k-1}, z_k) g(z_i, z_{k-1})] + h(z_i, z_{k-1}), \tag{4.9}$$

for $k = i + 2, i + 3,..., n$. It should be noted that Eq. (4.8) must be solved before (4.9) since the latter equation depends upon $g(z_i, z_{k-1})$.

We have now derived equations for all the functions required to calculate $u(z_i)$ and $v(z_i)$. Moreover, the recursive nature of the formulas is most appropriate for machine calculation. In practice, it is convenient to have still another set of equations. Those we have derived might be called "forward recursion formulas." Just as in the case with the differential equations of the method (see Section 2), it is possible to derive "backwards recursion formulas." We omit the details, but allude to the use of these in the description of our numerical algorithm (see Section 7).

Finally, we must dispense with the case in which the values of u and v are desired at a point z which cannot be a z_i point. Such a z must lie between two z_i points, say $z_j < z < z_{j+1}$. The method derived in this section may be used to find the value of u at z_j and the value of v at z_{j+1}. Now we may consider a new problem simply over the interval $[z_j, z_{j+1}]$. But this interval is such that we are in the well-behaved case, completely resolved in Section 2. Thus $u(z)$ and $v(z)$ may be readily calculated.

5. Periodic Problems

It is known (see [11, 13]) that when the coefficients in Eq. (3.1) are periodic with the same period the invariant imbedding method is especially successful in providing numerical results. It is reasonable to ask if similar simplifications occur in the inhomogeneous case, Eq. (2.1), when S^+ and S^- as well as A, B, etc., are all periodic with the same period. We address this problem briefly, leaving many of the details to the reader.

We first note that a closed form solution for (4.8) is obtainable even without any assumptions of periodicity. We chose not to examine this solution in the previous section since it is not particularly interesting in the general case. In the periodic problem it is valuable. We sketch the derivation.

First, define

$$G_k = g(z_0, z_k) - g(z_{k-1}, z_k) \tag{5.1}$$

so that (4.8) becomes

$$G_k = \frac{T_l(z_{k-1}, z_k)}{1 - R_r(z_0, z_{k-1}) R_l(z_{k-1}, z_k)} \{G_{k-1} + g(z_{k-2}, z_{k-1}) + R_r(z_0, z_{k-1}) h(z_{k-1}, z_k)\}. \tag{5.2}$$

But Eq. (4.3b) allows this to be put in the form

$$\frac{G_k}{T_l(z_0, z_k)} - \frac{G_{k-1}}{T_l(z_0, z_{k-1})} = \frac{g(z_{k-2}, z_{k-1})}{T_l(z_0, z_{k-1})} + \frac{R_r(z_0, z_{k-1})}{T_l(z_0, z_{k-1})} h(z_{k-1}, z_k). \tag{5.3}$$

Summing from $k = 2$ to i we obtain

$$\frac{G_i}{T_l(z_0, z_i)} = \frac{g(z_0, z_i) - g(z_{i-1}, z_i)}{T_l(z_0, z_i)}$$
$$= \sum_{k=2}^{i} \left[\frac{g(z_{k-2}, z_{k-1})}{T_l(z_0, z_{k-1})} + \frac{R_r(z_0, z_{k-1})}{T_l(z_0, z_{k-1})} h(z_{k-1}, z_k)\right].$$

Thus, with no assumption of periodicity,

$$g(z_0, z_i) = T_l(z_0, z_i)\left[\sum_{k=1}^{i} \frac{g(z_{k-1}, z_k)}{T_l(z_0, z_{k-1})} + \sum_{k=1}^{i-1} \frac{R_r(z_0, z_k)}{T_l(z_0, z_k)} h(z_k, z_{k+1})\right]. \tag{5.4}$$

Now turning to the periodic case we make the assumption that the z_k points may be so chosen that $z_k - z_{k-1} = P$, where P is the period and $k = 1, 2,..., n$. Then (see [13]) for all such k define

$$\begin{aligned} R_r(z_{k-1}, z_k) &= \rho_r, \qquad T_r(z_{k-1}, z_k) = \tau_r, \\ R_l(z_{k-1}, z_k) &= \rho_l, \qquad R_l(z_{k-1}, z_k) = \tau_l, \end{aligned} \tag{5.5}$$

where the ρ's and τ's are constants. Thus, for example,

$$R_r(z_0, z_k) = \frac{\tau_l \tau_r R_r(z_0, z_{k-1})}{1 - \rho_l R_r(z_0, z_{k-1})} + \rho_r, \tag{5.6a}$$

$$T_l(z_0, z_k) = \frac{\tau_l T_l(z_0, z_{k-1})}{1 - \rho_l R_r(z_0, z_{k-1})}, \tag{5.6b}$$

result immediately from Eq. (4.3). These recursion formulas are especially fast and stable in computation. The derivation just given is considerably simpler than the original [13].

As yet, we have made no use of the assumed periodicity of the S functions. Let us turn to the basic defining equation for the g function over the interval $[z_{k-1}, z_k]$ (Eq. (2.7)). We write it in the form

$$\begin{aligned} &(d/dw)\, g(z_{k-1}, z_{k-1} + w) \\ &\quad = [A(z_{k-1} + w) + C(z_{k-1} + w)\, R_r(z_{k-1}, z_{k-1} + w)]\, g(z_{k-1}, z_{k-1} + w) \\ &\quad\quad + S^-(z_{k-1} + w)\, R_r(z_{k-1}, z_{k-1} + w) + S^+(z_{k-1} + w) \\ &\quad = [A(z_0 + w) + C(z_0 + w)\, R_r(z_0, z_0 + w)]\, g(z_{k-1}, z_{k-1} + w) \\ &\quad\quad + S^-(z_0 + w)\, R_r(z_0, z_0 + w) + S^+(z_0 + w). \end{aligned} \tag{5.7}$$

Here we have made strong use of the periodicity. Now, if we select $k = 1$ we find that $g(z_0, z_0 + w)$ and $g(z_{k-1}, z_{k-1} + w)$, as functions of w, both satisfy the same equation. Moreover, they both vanish at $w = 0$. Therefore, they are identical, and so

$$g(z_{k-1}, z_k) = g(z_0, z_1) = \hat{g}, \tag{5.8}$$

where $\hat{g}$ is a constant. Similarly,

$$h(z_{k-1}, z_k) = h(z_0, z_1) = \hat{h}. \tag{5.9}$$

Finally, from (5.4),

$$g(z_0, z_i) = T_l(z_0, z_i)\left[\hat{g} \sum_{k=1}^{i} \frac{1}{T_l(z_0, z_{k-1})} + \hat{h} \sum_{k=1}^{i-1} \frac{R_r(z_0, z_k)}{T_l(z_0, z_k)}\right]. \tag{5.10}$$

A similar closed-form expression for $h(z_i, z_n)$ may be found. Computations of periodic problems have been made using these simplified equations and have been compared with calculations which used the results of Section 4 in a straightforward manner. Very considerable improvements in speed and accuracy are obtained when the methods of this section are utilized.

6. Other Boundary Conditions

Up to this point we have been using the standard boundary conditions found in most invariant imbedding treatments, i.e., $u(x)$ and $v(y)$ specified. As noted in the Introduction, these conditions are imposed largely as a matter of historical tradition. Actually a wide variety of boundary value problems can be handled.

Consider, for example, the case in which u is given at both x and y. We rewrite Eq. (2.4) so as to avoid any notational bias concerning which quantities are known and which are unknown.

$$u(z) = u(x)\, T_l(x, z) + v(z)\, R_r(x, z) + g(x, z), \tag{6.1a}$$

$$v(z) = u(z)\, R_l(z, y) + v(y)\, T_r(z, y) + h(z, y). \tag{6.1b}$$

If in (6.1a) we set $z = y$ and then solve for $v(y)$ we find

$$v(y) = [u(y) - u(x)\, T_l(x, y) - g(x, y)]/R_r(x, y). \tag{6.2}$$

Provided $R_r(x, y)$ is not zero (a matter into which we shall not delve since this section is purely expository) we now have $v(y)$ given in terms of quantities which may be presumed known. Thus the problem is reduced to the standard one with $u(x)$ and $v(y)$ prescribed.

It is obvious that cases in which u and v are both given at x (or at y), in which v is specified at both x and y, and in which various linear combinations of u and v are given at the two ends may be similarly treated. The interested reader is referred to [8] for a more thorough treatment. We have chosen to discuss the case in which u is prescribed at both x and y since two of our numerical examples are in that category.

7. Some Numerical Examples

In this section we shall illustrate the use of the algorithm we have derived by applying it to four examples. No direct effort has been made to compute exactly the points referred to in Section 3 as z_i. We simply integrate the R_r equation forward until either we experience numerical difficulty or sense that the integrator is working too hard. (We measure this by counting the number of function evaluations the integrator requires.) A z_i point is then chosen as a point at which R_r can still be computed to within the imposed accuracy criteria or the maximum function count is reached. Integration then automatically begins again with initial data at z_i.

Using the equations we have developed our computational algorithm proceeds as follows. Suppose we want the solution to the problem (2.1) at the point z^*. Assume that z^* is not an exceptional point of the kind discussed in the last paragraph of Section 4. Then z^* may be included in the set of z_i points computed above. If it is not we simply augment that set by z^* and reindex. We determine the values $R_r(z_{i-1}, z_i)$, $T_l(z_{i-1}, z_i)$, $R_l(z_{i-1}, z_i)$, $T_r(z_{i-1}, z_i)$, $g(z_{i-1}, z_i)$ and $h(z_{i-1}, z_i)$ for $i = 1, 2, \ldots, n$, by integrating the differential equations (2.2a–d), (2.7), and (2.9). The recursion formulas (4.2a-d) and (4.8) are then used to find the quantities $R_r(x, z^*)$, $T_l(x, z^*)$ and $g(x, z^*)$. Equations of the type (4.4) together with (4.8) and (4.9) yield $R_l(z^*, y)$, $T_r(z^*, y)$, and $h(z^*, y)$. The solution to (2.1) is then obtained by solving the trivial linear algebraic system (2.4).

The reader will note that while conceptually it is somewhat easier to find the z_i points (other than z^*) before implementing the algorithm, in practice this can be done during the integration of the differential equations, thus economizing on computing time.

When the solution of the problem (2.1) is desired at several points z^*, the above approach is rather wasteful. The values $R_l(z^*, y)$, $T_r(z^*, y)$, and $h(z^*, y)$ can be calculated more easily by employing "backward recursion fomulas" of the kind mentioned at the end of Section 4. We do not pursue this and other nuances of the method here.

In our examples all differential equations were integrated using a fourth-order Runge–Kutta–Fehlberg scheme designed to estimate the local error and to control the step size to insure user input accuracy requirements [14]. We used an absolute and relative error of 10^{-6} per unit step in the independent variable and a maximum function count of 120. All calculations were done on an IBM 360/67 in extended precision arithmetic. (Some early problems were run in single precision and satisfactory results, compatible with the arithmetic, were obtained.)

EXAMPLE 1. The first example to be given was one described in [15] as presenting some difficulties when various versions of invariant imbedding, different from the one we have developed, were used for its solution. It is simply

$$u'(z) = v(z) + 1.0 \tag{7.1a}$$

$$-v'(z) = -u(z) \tag{7.1b}$$

$$u(0) = 0,\ v(20) = 1.0. \tag{7.1c}$$

The analytical solution is readily found to be

$$u(z) = 2\operatorname{sech}(20)\sinh(z);\ v(z) = 2\operatorname{sech}(20)\cosh(z) - 1.0. \tag{7.2}$$

For this system the R_r equation is

$$R_r' = 1.0 - R_r^2 \tag{7.3}$$

which has a solution for all z. Therefore, the example under consideration is actually a well-behaved case. Nevertheless, it seems valuable to verify that the method is effective in this simpler situation. This problem was solved by use of our general program and the solution was found accurate to five significant digits. In particular, those values of z which produced difficulties in [15] were handled with ease. It might be added that computations in well-behaved cases ordinarily proceed very rapidly

EXAMPLE 2. This example was also chosen to be a very easy one:

$$y'' + y(z) = z, \tag{7.4a}$$

$$y(0) = 1.0,\ y'(10.0) = \cos(10.0) - \sin(10.0) + 1.0. \tag{7.4b}$$

This is equivalent to

$$u'(z) = v(z), \tag{7.5a}$$

$$-v'(z) = u(z) - z, \tag{7.5b}$$

$$u(0) = 1.0,\ v(10.0) = \cos(10.0) - \sin(10.0) + 1.0. \tag{7.5c}$$

Again the analytical solution may be found

$$u(z) = \sin(z) + \cos(z) + z,\ v(z) = \cos(z) - \sin(z) + 1.0. \tag{7.6}$$

The R_r equation here is

$$R_r' = 1.0 + R_r^2. \tag{7.7}$$

Clearly this equation cannot be integrated over any interval of length greater than or equal to $\pi/2$. Hence the example is in the ill-behaved category and provides a genuine test of our algorithm and program since there are several "break points" (points at which the integration is automatically started over) on the interval $0 \leqslant z \leqslant 10.0$. Table 1 gives the results of our calculations as well as the absolute errors in the computed quantities. The results are compatible with the accuracy criteria used.

TABLE I

z	computed $u(z)$	computed $v(z)$	error $u(z)$	error $v(z)$
0.0	1.0000000	1.9999967	0.0	3.3(−6)
2.0	2.4931471	−3.2544501(−1)	3.5(−6)	7.5(−6)
4.0	2.5895545	1.1031621	5.9(−7)	3.2(−6)
6.0	6.8075711	2.2395829	2.3(−6)	2.8(−6)
8.0	8.8438561	−1.3485984	2.1(−6)	1.6(−6)
10.0	8.6169074	7.0495446	4.9(−6)	0.0

EXAMPLE 3. We now consider an equation which arises in a practical setting. It describes the stress distribution in a spherical membrane with normal and tangential loads [16].

$$y'' + [3\cot(z) + 2\tan(z)]\, y' + 0.7y(z) = 0, \tag{7.8a}$$

$$y(30°) = 0,\ y(60°) = 5.0. \tag{7.8b}$$

Previous numerical studies have been made of the solution, many of them rather unsuccessfully. The value of y actually rises from zero to about 283 as z changes from 30° to 30.7°. The matter is further complicated by the fact that in the physical application the value of y' is also of interest.

The reader will note that the problem is homogeneous and hence does not really fit the spirit of this paper. Of course, the algorithm we have derived can be used when $S^+ = S^- = 0$. We actually carried out this calculation with good results. However, in order to show the effectiveness of our full algorithm we converted the given problem into an inhomogeneous one by writing

$$y(z) = u(z) - 1.0 \tag{7.9}$$

and using the system

$$u'(z) = v(z), \tag{7.10a}$$

$$-v'(z) = 0.7u(z) + [3\cot(z) + 2\tan(z)]\, v(z) - 0.7, \tag{7.10b}$$

$$u(30°) = 1.0,\ u(60°) = 6.0. \tag{7.10c}$$

The R_r equation for this problem is very badly behaved and there were a total of 57 break points in the integration, more or less equally spaced. Table 2 provides our results. They are consistent to six significant digits with other methods (see, e.g., [16]).

TABLE 2

	Computed	
z	$u(z)$	$v(z)$
30.0	1.00000	1896.44
40.0	90.07069	−12.1522
50.0	22.26790	−3.13099
60.0	6.00000	−0.693964

EXAMPLE 4. As our final example we choose an equation for which standard superposition algorithms are quite inadequate. The equation is

$$y'' - (1 + z^2)\, y(z) = 0, \tag{7.11a}$$

$$y(0) = 1.0,\ y(10.2) = 0. \tag{7.11b}$$

This was converted into the form

$$u'(z) = v(z), \tag{7.12a}$$

$$-v'(z) = -(1.0 + z^2)\, u(z), \tag{7.12b}$$

$$u(0) = 1.0,\ u(10.2) = 0. \tag{7.12c}$$

Equation (7.12) is admittedly homogeneous. However, we feel that the effectiveness of our method on inhomogeneous problems has been demonstrated by the first three examples and that the interest here lies in the ability of our algorithm to handle problems with very badly behaved solutions.

Again the R_r equation produces difficulty, resulting in a total of eight break points in the calculation. Table 3 gives the results. There we also present results obtained by S. Pruess using a method of approximating coefficients [17]. (See also results by M. R. Osborne [18].)

TABLE 3

	Imbedding		Pruess	
z	u	v	u	v
0.0	1.000	−1.128	1.000	−1.128
2.0	3.456(−2)	−8.358(−2)	3.456(−2)	−8.358(−2)
4.0	4.595(−5)	−1.946(−4)	4.596(−5)	−1.947(−4)
6.0	1.412(−9)	−8.700(−9)	1.413(−9)	−8.707(−9)
8.0	8.846(−16)	−7.186(−15)	8.863(−16)	−7.199(−15)
10.0	1.061(−23)	−1.109(−22)	1.064(−23)	−1.112(−22)
10.2	0.0	−2.883(−23)	0.0	−2.887(−23)

8. Summary and Final Remarks

In this paper we have developed an algorithm to solve inhomogeneous linear differential equations by the method of invariant imbedding for cases in which the solution to the equation for the reflection function is ill-behaved. For purposes of ease of exposition we have confined our investigation to the case of scalar equations. However, the methods generalize completely to matrix systems. The algorithm has been tested numerically on numerous scalar problems. Four are given as examples. The scheme seems to provide much greater accuracy in some cases than more standard methods, although it can be somewhat costly in time, especially when there are many break points.

The use of the recursive equations in generating the reflection and transmission functions appears quite powerful. In homogeneous scalar problems these equations may often be avoided by the simple device of replacing the reflection function by its reciprocal in the vicinity of a singularity. An analagous trick is probably available for the inhomogeneous problem, but it seems quite likely that for matrix problems an approach analogous to ours may be superior numerically to any based on the analog of taking reciprocals of reflection functions. It is hoped that this whole area may be investigated.

One basic difficulty with the invariant imbedding procedure is that it depends strongly on the linearity of the system under study. (It is true, of course, that imbedding equations have been derived for nonlinear problems, but they have not been found to be of any great numerical value.) However, one way of approaching nonlinear problems is by the use of skillfully devised sequences of linear problems. The algorithm we have constructed is of potential great use here since the solution of the approximating linear problem at the n-th stage usually contributes to the inhomogeneous term of the approximating problem at the $(n + 1)$ stage. Our method can easily take advantage of this phenomenon.

In conclusion, we might remark that the continuity condition on the various functions $A(z)$, $B(z)$, etc., is considerably stronger than necessary. One can easily allow piecewise continuity by simply making small modifications in Section 3. It is likely that even this restriction can be significantly relaxed.

REFERENCES

1. R. M. REDHEFFER, *J. Math. Phys.* **41** (1962), 1.
2. R. C. ALLEN, JR. AND G. M. WING, *J. Math. Anal. Appl.* **42** (1973), 397.
3. E. DENMAN, Invariant imbedding and linear systems, in "Invariant Imbedding," Springer-Verlag, Berlin-Heidelberg-New York, 1971.
4. E. DENMAN, *Inter. J. Numer. Meth. Eng.* **4** (1972), 587.
5. E. DENMAN, "Coupled Modes in Plasma, Elastic Media, and Parametric Amplifiers," American Elsevier, New York, 1971.
6. G. M. WING, *J. Math. Anal. Appl.* **13** (1966), 361.
7. P. NELSON AND M. SCOTT, *J. Math. Anal. Appl.* **37** (1972), 501.
8. M. GOLBERG, *J. Math. Anal. Appl.* **33** (1971), 518.
9. B. D. TAPLEY AND W. E. WILLIAMSON, *J. Amer. Inst. Aeron. Astron.* **10** (1972), 1154.
10. W. T. REID, *J. Math. Mech.* **8** (1959), 221.
11. R. C. ALLEN, JR., J. W. BURGMEIER, P. MUNDORFF, AND G. M. WING, *J. Math. Anal. Appl.* **37** (1972), 725.
12. R. C. ALLEN, JR., *J. Math. Anal. Appl.* **30** (1970), 48.
13. R. C. ALLEN JR. AND G. M. WING, *J. Math. Anal. Appl.* **29** (1970), 141.
14. L. SHAMPINE AND R. C. ALLEN, JR., "Numerical Computing: An Introduction," Saunders, Philadelphia, Penn., 1973.
15. P. NELSON, *J. Computational Phys.* **10** (1972), 374.
16. A. L. ROARK AND L. F. SHAMPINE, On the numerical solution of a linear two-point boundary value problem, Sandia Laboratories Report SC-TM-67-588, September, 1967.
17. S. PRUESS, *Math. Comput.* **27** (1973), 551.
18. M. R. OSBORNE, *J. Math. Anal. Appl.* **27** (1969), 417.