JOURNAL OF COMPUTATIONAL PHYSICS 66, 356–390 (1986)

Galerkin Method for Differential Equations with Regular Singular Points

A. D. MILLER

Centre for Mathematical Analysis,
The Australian National University, Canberra A.C.T., Australia 2601

AND

R. L. DEWAR

Department of Theoretical Physics, Research School of Physical Sciences,
The Australian National University, Canberra A.C.T. Australia 2601

Received October 11, 1985; revised February 4, 1986

DEDICATED TO THE MEMORY OF RAYMOND C. GRIMM

A Galerkin method is presented for calculating the general weak solution of self-adjoint differential equations with regular singular points, such as the ideal MHD equation for zero-frequency displacements about a finite beta, cylindrical plasma equilibrium (the Newcomb equation). In this case such solutions could be used for constructing the eigenfunctions of resistive instabilities by the method of matched asymptotic expansions. A Galerkin method using finite elements, singular if necessary, is used to calculate the finite energy part of the solution, with the infinite energy part acting as a forcing term. The asymptotic behaviour near the singular point is accurately estimated by postprocessing the Galerkin solution using a generalized Green's function method. The effectiveness of the method is demonstrated on some simple test cases. © 1986 Academic Press, Inc.

1. INTRODUCTION

Since the work of Furth, Killeen, and Rosenbluth (FKR) [1], a standard way to treat linear resistive instabilities in weakly resistive plasmas has been to use the method of matched asymptotic expansions. In this approach the plasma is separated into thin inner layers containing the mode rational surfaces, where resistivity and frequency are effectively of order unity, and large outer regions away from mode rational surfaces where the eigenfunctions, to lowest order, satisfy the linearized ideal (zero resistivity) magnetohydrodynamic equation of motion at zero frequency (henceforth called the ideal marginal equation).

The FKR approach has been generalized to compressible, cylindrical, finite

0021-9991/86 $3.00

pressure plasmas by Coppi, Greene, and Johnson [2], and this work has in turn been generalized to toroidal geometry by Glasser, Greene, and Johnson [3]. Recently [4] a similar inner layer method has been developed to treat nonlinear, but thin, magnetic islands. In principle, the physics of the inner layer can be modelled to almost any level of complexity provided only that the width is small compared with typical equilibrium scale lengths. This is because, even in a toroidal plasma, the equations which must be solved are one dimensional. Although it may well be impossible to solve these equations analytically, the reduction in dimension provided by the thin layer approximation greatly reduces the computational time for accurate numerical solution of the equations containing the complicated physics.

In this context we note the exhaustive comparison of three numerical techniques for solving the linear resistive inner layer problem recently carried out by Glasser, Jardin, and Tesauro [5]. Of particular interest to us here is their general dispersion relation, Eq. (94), which shows that in general two exterior quantities, Δ_L and Δ_R are needed to determine the growth rate. The quantities Δ_L and Δ_R are, respectively, the ratios of the leading coefficients of the nonanalytic Frobenius solutions of the ideal marginal equation to the left and right of the mode rational surface. These ratios are determined by solving the ideal marginal equation everywhere, except at the mode rational surface (where the ideal marginal equation is singular), taking proper account of the boundary conditions and other geometric complications such as the toroidal nature of the plasma.

It is the fact that the full geometrical complexity of the problem need only be faced using the relatively simple ideal marginal equations which represents the power of the matching method. Unfortunately, even with this reduction, the computational problem is highly nontrivial in realistic geometries. This is due to the difficulty of accurately extracting the asymptotic behaviour of the solution in the neighbourhood of the singular surface from approximate numerical solutions typically based on spatial discretization (e.g., the toroidal stability codes [6, 7]).

This difficulty has led to an attempt [8, 9] to adapt the finite element method of [6] so as to represent the singular behaviour more accurately by the use of singular elements based on the Frobenius expansion. Mesh refinement near the singular surface was also used. As well as simply allowing a more accurate representation, the singular finite elements were used to perform two important tasks: First, the "big" elements (carrying the dominant asymptotic behaviour) were assigned unit amplitudes, thus acting as normalized forcing terms for the remaining part of the solution. Second, the amplitudes of the "small" elements were interpreted as the coefficients of the Frobenius solutions corresponding to the more positive solution of the indicial equation, thus giving Δ_L and Δ_R.

The method has proved quite effective at low plasma pressure, where the exponents of the big and small solutions differ by unity, approximately. However, convergence problems were encountered for higher pressure equilibria. This led [10] to the examination of a simple one-dimensional test problem to gain an understanding of the fundamental basis of the method. Numerical analysis [10] revealed that the singular element procedure as outlined above was very sensitive to

small errors near the singular point. It is the purpose of the present paper to present an alternative approach which, because it obtains Δ_L or Δ_R by an integration over the whole domain of the solution, is much less sensitive to local errors. Also, because it separates the representational problem from the problem of extracting the asymptotic behaviour, it is not inherently tied to the use of singular elements. This makes the method easier to analyze and should make generalization to the toroidal case relatively straightforward. To simplify the exposition of the method, however, we shall in this paper confine our attention to the one-dimensional problem. The results are thus only directly applicable to the determination of the resistive stability of a cylindrical plasma of circular cross section.

The new method was suggested by analogy with the (in some ways similar) problem of calculating the coefficients of stress singularities about a crack in an elastic solid [11]. It may also have application in other areas, such as in calculating the stability of a shear flow in a weakly viscous fluid [12].

2. Formulation of the Problem

As our model problem we consider the Newcomb equation [13],

$$F(y) \equiv \frac{d}{dx}\left(f(x)\frac{d}{dx}y\right) - g(x)\,y = 0, \qquad 0 < x < 1, \tag{1a}$$

with the boundary condition

$$y(1) = 0. \tag{1b}$$

Here $f(x)$ and $g(x)$ are smooth functions which are analytic at $x = 0$. In addition, f does not vanish for $0 < x \leqslant 1$, though it has a zero of order 2 at $x = 0$. Without any loss of generality we may suppose that f is nonnegative, and that it has the Taylor expansion

$$f(x) = x^2 + \sum_{m=3}^{\infty} f_m x^m, \qquad f_m \text{ real constant},$$

valid in some neighbourhood of $x = 0$. Likewise, let the Taylor expansion of g be

$$g(x) = \sum_{m=0}^{\infty} g_m x^m, \qquad g_m \text{ real constant},$$

again converging in some neighbourhood of $x = 0$. In the plasma stability setting, y is the radial component of the plasma displacement, and F is proportional to the radial component of the force density (the other two components having been assumed to vanish). The point $x = 0$ corresponds to a mode rational surface with the boundary condition (1b) describing a perfectly conducting wall. The boundary

value problem (1) models the displacement on one side of a mode rational surface, corresponding to the calculation of Δ_R in [5]. The calculation of Δ_L would be completely analagous.

By virtue of our assumptions on f and g, (1a) has a regular singular point at $x=0$. Formally at least, a complete set of solutions of (1a) can be readily determined. We may write the most general solution as

$$y(x) = a_b\, y^{(b)}(x) + a_s\, y^{(s)}(x), \tag{2}$$

where a_b and a_s are arbitrary real constants, and $y^{(b)}$, $y^{(s)}$ is a suitable pair of linearly independent solutions obtained by a Frobenius expansion technique about $x=0$. The indicial equation in this case is

$$\alpha(\alpha+1) - g_0 = 0,$$

which has roots

$$\alpha = \alpha_\mu^{(b)} \equiv -1/2 - \mu \qquad \text{and} \qquad \alpha = \alpha_\mu^{(s)} \equiv -1/2 + \mu,$$

where

$$\mu = (1/4 + g_0)^{1/2}.$$

We shall suppose that

$$g_0 > -\tfrac{1}{4}. \tag{3}$$

This ensures that μ is real and positive, and thus

$$\alpha_\mu^{(b)} < -\tfrac{1}{2} < \alpha_\mu^{(s)}. \tag{4}$$

In the notation of [3], $\mu = (-D_I)^{1/2}$, and (3) corresponds to the Suydam criterion for stability of the plasma against ideal interchange modes: if the plasma were ideally unstable, resistive stability would only be of academic interest.

Corresponding to the larger exponent $\alpha_\mu^{(s)}$ is a solution $y_\mu^{(s)}$, which has the following expansion valid in some deleted neighbourhood of $x=0$,

$$y_\mu^{(s)}(x) = x^{\alpha_\mu^{(s)}} \sum_{m=0}^{\infty} p_m^{(s)} x^m, \tag{5}$$

where

$$p_0^{(s)} = 1,$$

$$p_m^{(s)} = -\sum_{l=0}^{m-1} \frac{[(\alpha_\mu^{(s)}+l)(\alpha_\mu^{(s)}+m+1) f_{m+2-l} - g_{m-l}]\, p_l^{(s)}}{m(m+2\mu)}, \qquad m = 1, 2, 3, \ldots . \tag{6}$$

Provided $2\mu \neq 1, 2, 3,...$, the other component $\alpha_\mu^{(b)}$ gives rise to a solution $y_\mu^{(b)}$ with the following expansion, valid in some deleted neighbourhood of $x = 0$,

$$y_\mu^{(b)}(x) = x^{\alpha_\mu^{(b)}} \sum_{m=0}^{\infty} p_m^{(b)} x^m, \tag{7}$$

where

$$p_0^{(b)} = 1,$$

$$p_m^{(b)} = -\sum_{l=0}^{m-1} \frac{[(\alpha_\mu^{(b)} + l)(\alpha_\mu^{(b)} + m + 1) f_{m+2+l} - g_{m-l}]\, p_l^{(b)}}{m(m - 2\mu)}, \qquad m = 1, 2, 3,.... \tag{8}$$

If 2μ is a positive integer, the recurrence relation (8) in general breaks down at $m = 2\mu$. This, however, is just a technical difficulty. It reflects a growing linear dependence between $y_\mu^{(b)}$ and $y_\mu^{(s)}$ as μ approaches a half integer. It is particularly important to be able to treat the case $\mu = \frac{1}{2}$, since this corresponds to a plasma with zero pressure gradient at the singular surface. The classical Frobenius technique is to define a second solution $y_\mu^{(b)}$ by other means in these exceptional cases. (This typically introduces $\log x$ terms into the expansion.) However, this is not entirely satisfactory, as it does not do anything about the near linear dependence of $y_\mu^{(b)}$ and $y_\mu^{(s)}$ for μ near, but not at, a half integer. To overcome this difficulty we need to redefine $y_\mu^{(b)}$ for a full range of μ's by adding appropriate multiples of $y_\mu^{(s)}$. This can be done in many ways, with one possibility being as follows: Set

$$\gamma_m = \operatorname*{lmt}_{\mu \to m/2} \sin 2\mu\pi\, p_m^{(b)}, \qquad m = 1, 2, 3,...,$$

which is certainly finite. Let $\gamma(\cdot)$ be some sufficiently smooth function defined on the positive real numbers which satisfies

$$\gamma\left(\frac{m}{2}\right) = \gamma_m, \qquad m = 1, 2, 3,... .$$

Now in place of (7), let us redefine $y_\mu^{(b)}$ by

$$y_\mu^{(b)}(x) = x^{\alpha_\mu^{(b)}} \sum_{m=0}^{\infty} p_m^{(b)} x^m - \frac{\gamma(\mu)}{\sin 2\mu\pi}\, y_\mu^{(s)}(x) \tag{9a}$$

which is certainly well defined for $2\mu \neq 1, 2, 3,...$. While for the half integers we define

$$y_{m/2}^{(b)}(x) = \operatorname*{lmt}_{\mu \to m/2} y_\mu^{(b)}(x), \qquad m = 1, 2, 3,... . \tag{9b}$$

The validity of the limiting process in (9b) is readily seen from the explici expression for $y_\mu^{(b)}(x)$,

$$y_\mu^{(b)}(x) = \sum_{n=0}^{N_\mu - 1} p_n^{(b)} x^{\alpha_\mu^{(b)} + n} + \sum_{n=N_\mu}^{\infty} \left[\left\{ p_n^{(b)} - \frac{\gamma(\mu)}{\sin 2\mu\pi} p_{n-N_\mu}^{(s)} \right\} x^{\alpha_\mu^{(s)} + n - N_\mu} - \{(2\mu - N_\mu)\, p_n^{(b)}\} \left\{ \frac{x^{2\mu - N_\mu} - 1}{2\mu - N_\mu} \right\} x^{\alpha_\mu^{(b)} + n} \right], \tag{10}$$

where N_μ denotes the nearest positive integer to 2μ, and for $\mu = m/2$ $(m = 1, 2, 3, \ldots)$ the expressions in $\{\ \}$ take their limiting values as $\mu \to m/2$. The first expression in $\{\ \}$ is finite as $\mu \to m/2$, since for $n \geqslant m$,

$$p_n^{(b)} \sim \frac{\gamma_m\, p_{n-m}^{(s)}}{\sin 2\mu\pi}.$$

Note that the third term in $\{\ \}$ will generate the characteristic $\log x$ term at the half integers.

The pair of solutions $y_\mu^{(s)}$ and $y_\mu^{(b)}$ given by (5) and (9) constitute an acceptable pair of linearly independent solutions of (1a) for all μ bounded away from zero. They will though, degenerate as μ approaches zero, but this will not concern us. The above construction bears an obvious similarity to the method used to define Bessel functions of the second kind [15]. Note that $y_\mu^{(b)}$ and $y_\mu^{(s)}$ have been normalized such that

$$y_\mu^{(b)} \sim x^{\alpha_\mu^{(b)}} \qquad \text{and} \qquad y_\mu^{(s)} \sim x^{\alpha_\mu^{(s)}} \quad (x \to 0+). \tag{11}$$

In particular, $y_\mu^{(b)}$ is dominant, or in the terminology of [13], big, at $x = 0$; while $y_\mu^{(s)}$ is recessive, or small.

The solution of (1) is only determined up to an arbitrary multiplicative factor. For definiteness, we normalize (2) by requiring that

$$a_b = 1, \tag{12a}$$

or, what is equivalent,

$$y \sim x^{\alpha_\mu^{(b)}} \qquad (x \to 0+). \tag{12b}$$

This normalization requires that $y_\mu^{(s)}(1) \neq 0$, but this will be a consequence of an assumption that we shall introduce later (see Lemma A.3 of Appendix A).

Our interest in (1) lies in the asymptotic behaviour of y near $x = 0$. As outlined in the Introduction, this information may be used to match y to some inner solution which is valid in a narrow region about $x = 0$ where the outer equation (1a) may no longer be physically reasonable. Given the expansions (5) and (9) for $y_\mu^{(s)}$ and $y_\mu^{(b)}$, it

is immediate from (2) and (12a) that a_s completely determines the asymptotic behaviour of y. The task we set ourselves then, is to find the coefficient a_s. Let us remark that the numerical value of a_s is of course relative to the particular choice of the dominant solution $y_\mu^{(b)}$, since we can include arbitrary multiples of $y_\mu^{(s)}$ in the definition of $y_\mu^{(b)}$. In particular, a_s is not in general equal to Δ_R of [5], but it contains the same physical information. In fact, $\Delta_R = a_s - (\gamma(\mu)/\sin 2\mu\pi)$. Note also that for the model problem (1) we have in fact the explicit result

$$a_s = -\frac{y_\mu^{(b)}(1)}{y_\mu^{(s)}(1)}, \tag{13}$$

provided $y_\mu^{(s)}(1) \neq 0$. This, however, may not be a very practical method for calculating a_s. First, the Frobenius expansions for $y_\mu^{(b)}$ and $y_\mu^{(s)}$ may not converge at $x = 1$. This could occur if the coefficient functions f and g were not analytic throughout all of the interval $0 \leqslant x \leqslant 1$. Second, even if the expansions converge, the calculation of a sufficient number of coefficients $p_m^{(s)}$, $p_m^{(b)}$ may be prohibitive; especially, if there is no simple means of obtaining the Taylor series coefficients f_m, g_m of f and g. The main point is, however, that (1) is only a model for problems that have a more complex coupling between a "far" boundary condition, such as (1b), and the asymptotic behaviour at a regular singular point (see [8, 9]). For such problems, explicit expressions such as (13) are likely to be even less useful. We therefore seek a numerical method for calculating a_s, which should have the potential to extend to a wider class of problems than just (1).

We shall proceed in two stages. First, an approximate solution of (1a), (1b), (12b) will be found using a finite element approach, and second, an approximation to a_s will be extracted from this approximate solution by a generalized Green's function technique.

3. Finite Element Approximation

3.1. *The Finite Element Approximation of y*

Our finite element approximation will be based on a Galerkin formulation of (1). To see what shape such a formulation should take, multiply the left-hand side of (1a) by an arbitrary test function v, and formally integrate by parts to obtain

$$\int_0^1 F(y)\, v = f(x)\, y'v\,|_{x=0}^{x=1} - \int_0^1 (f(x)\, y'v' + g(x)\, yv).$$

Therefore it seems natural to consider the bilinear form

$$W(u, v) \equiv \int_0^1 (fu'v' + guv). \tag{14}$$

We start by describing some assumptions on $W(\cdot, \cdot)$.

Introduce the norm and inner product

$$\|u\| = \left[\int_0^1 (u^2 + x^2(u')^2\right]^{1/2},$$

$$(u, v) = \int_0^1 (uv + x^2 u'v'),$$

which are defined, at least to start with, for functions belonging to

$$\mathscr{D} = \{v \in C^\infty([0, 1])\}.$$

Denote by $\mathscr{H}$ the formal completion of $\mathscr{D}$ in the norm $\|\cdot\|$. $\mathscr{H}$ is a Hilbert space. Both the norm $\|\cdot\|$ and inner product $(\cdot, \cdot)$ extend naturally to $\mathscr{H}$, and we may regard $W(u, v)$ as defined for all $u, v \in \mathscr{H}$; in fact, for some constant $C > 0$,

$$|W(u, v)| \leqslant C \|u\| \|v\|, \qquad \forall u, v \in \mathscr{H}. \tag{15}$$

Define the subspace $\mathscr{H}_0 \subseteq \mathscr{H}$ by

$$\mathscr{H}_0 = \{u \in \mathscr{H}; u(1) = 0\}.$$

$\mathscr{H}_0$ is also a Hilbert space. We shall assume that $W(\cdot, \cdot)$ is positive definite over $\mathscr{H}_0$, that is, for some constant $\eta > 0$,

$$W(u, u) \geqslant \eta \|u\|^2, \qquad \forall u \in \mathscr{H}_0. \tag{16}$$

We remark that the validity of our earlier assumption (3) on g_0, in fact, follows from (16) (see Theorem A.1 of Appendix A).

The assumption (16) will obviously hold if for some real constant $G > 0$, $g(x) \geqslant G$. If g becomes negative though, the situation is more delicate. For our model problem the Newcomb condition [13] provides a sufficient criterion for (16) to be valid.

Because of the obvious mechanical analogy we shall sometimes refer to $W(u, u)$ as the "energy" of u. In a plasma context it is proportional to the quadratic part δW of the plasma potential energy after minimization over components of the plasma displacement lying within the magnetic surfaces [13]. Thus (16) is simply the assumption of ideal stability. The kinetic energy does not enter as we are solving the ideal marginal equation.

Functions in $\mathscr{H}$ are well behaved away from $x = 0$. However, as $x \to 0+$, $x^{-\beta}$–like growth is permitted as long as $\beta < 1/2$. In particular, the $x^{-1/2+\mu}$–like growth of the recessive solution $y_\mu^{(s)}$ is admissible, while the $x^{-1/2-\mu}$–like behaviour of the dominant solution $y_\mu^{(b)}$ is not (see Lemma A.2). It would seem then that a straightforward Galerkin formulation of (1) based on the bilinear form $W(\cdot, \cdot)$ is not possible, as the energy $W(y, y)$ of the solution itself would be infinite. To over-

come this difficulty we need to decompose the solution y into a finite energy part and an infinite energy part. To do this, recall (12a) and rewrite (2) as

$$y(x) = y_0(x) + y_\infty(x), \tag{17}$$

where y_∞ is the first $N_\mu + 1$ terms of (10),

$$y_\infty(x) = \sum_{m=0}^{N_\mu - 1} p_m^{(b)} x^{\alpha_\mu^{(b)} + m} + \left[\left\{ p_{N_\mu}^{(b)} - \frac{\gamma(\mu)}{\sin 2\mu\pi} \right\} x^{\alpha_\mu^{(s)}} - \{(2\mu - N_\mu)\, p_{N_\mu}^{(b)}\} \left\{ \frac{x^{2\mu - N_\mu} - 1}{2\mu - N_\mu} \right\} x^{\alpha_\mu^{(b)} + N_\mu} \right], \tag{18a}$$

and

$$y_0(x) = a_s\, y_\mu^{(s)}(x) + (y_\mu^{(b)}(x) - y_\infty(x)). \tag{18b}$$

Note that the asymptotic behaviour of $y_0(x)$ as $x \to 0+$ is

$$y_0(x) \sim \begin{cases} a_s x^{\alpha_\mu^{(s)}} + O(x^{\alpha_\mu^{(s)} + 1 + \min(0, N_\mu - 2\mu)}) & \text{if } 2\mu \neq N_\mu \\ a_s x^{\alpha_\mu^{(s)}} + O(x^{\alpha_\mu^{(s)} + 1} \log x) & \text{if } 2\mu = N_\mu. \end{cases} \tag{19}$$

Thus y_0 has admissible behaviour at $x = 0$, and moreover its leading term is just $a_s x^{\alpha_\mu^{(s)}}$. Note also that for all μ the difference between the order of the leading term and the order of the remainder of y_0 is uniformally bounded away from 0.

Using the decomposition (17) we may rewrite (1) as

$$-F(y_0) = F(y_\infty), \qquad 0 < x < 1, \tag{20a}$$

$$y_0(1) = -y_\infty(1). \tag{20b}$$

This can be regarded as a boundary value problem for the unknown y_0, with boundary condition and forcing term derived from the known y_∞. Multiplying (20a) by an arbitrary test function $v \in \mathscr{H}_0$ and integrating by parts gives

$$W(y_0, v) - f(x)\, y_0' v \big|_{x=0}^{x=1} = \int_0^1 F(y_\infty)\, v.$$

The endpoint terms on the left-hand side will both vanish, since $v = 0$ at $x = 1$, while near $x = 0$, $f(x)\, y_0'$ has an $x^{1/2+\mu}$-type behaviour with v growing no faster than $x^{-1/2}$. Thus y_0 satisfies

$$y_0 \in \mathscr{H}_1 \equiv \{u\colon u(x) = v(x) - y_\infty(1),\ v \in \mathscr{H}_0\}, \tag{21a}$$

$$W(y_0, v) = \int_0^1 F(y_\infty)\, v, \qquad \forall v \in \mathscr{H}_0, \tag{21b}$$

where the trial space $\mathscr{H}_1$ is simply a translation of the test space $\mathscr{H}_0$ to accom-

modate the boundary condition (20b). On the other hand, standard Hilbert space theory [14] shows that (21) has a unique solution. Thus the formulation (21) completely characterizes y_0. It will be the basis of our finite element method. Note that implicit in (21) is the extra "boundary condition" of (20) at $x=0$,

$$f(x)\, y_0' = O(x^{1/2}), \qquad x \to 0+.$$

Let $\tilde{\mathscr{H}}_0$ be some finite element subspace of $\mathscr{H}_0$. We shall be more specific concerning $\tilde{\mathscr{H}}_0$ in Section 4; for the moment we just want to outline the general approach. We seek a function $\tilde{y}_0$ which satisfies

$$\tilde{y}_0 \in \tilde{\mathscr{H}}_1 \equiv \{u\colon u(x) = v(x) - y_\infty(1),\ v \in \tilde{\mathscr{H}}_0\}, \tag{22a}$$

$$W(\tilde{y}_0, v) = \int_0^1 F(y_\infty)\, v, \qquad \forall v \in \tilde{\mathscr{H}}_0. \tag{22b}$$

Standard theory [14] guarantees that such a $\tilde{y}_0$ exists and is unique. We also have the usual finite element orthogonality relation

$$W(y_0 - \tilde{y}_0, v) = 0, \qquad \forall v \in \tilde{\mathscr{H}}_0, \tag{23a}$$

and the best approximation result

$$\|y_0 - \tilde{y}_0\| \leqslant C \inf_{v \in \tilde{\mathscr{H}}_1} \|y_0 - v\| \tag{23b}$$

for some constant $C > 0$, independent of y_0 and $\tilde{\mathscr{H}}_1$.

3.2. *Extraction of the Coefficients a_s*

Once we have the finite element solution $\tilde{y}_0$, we must then somehow extract an approximate value for a_s, the coefficient of the leading term of y_0. We shall employ what can be termed a generalized Green's function technique [11]. This technique relies upon deriving a suitable integral representation for a_s in terms of y_0 and y_∞.

To describe the method, let ψ be a function which satisfies

$$\text{(i)}\quad \psi(x) \sim x^{\alpha_\mu^{(b)}}$$
$$\text{(ii)}\quad d\psi/dx \sim \alpha_\mu^{(b)} x^{\alpha_\mu^{(b)} - 1}, \qquad \text{as } x \to 0+, \tag{24a}$$

$$\text{(iii)}\quad \psi(1) = 0, \tag{24b}$$

(ψ is not related to the magnetic flux, often also denoted by ψ in plasma applications).

An obvious choice would be

$$\psi(x) = \psi_0(x) \equiv y_\infty(x) - x y_\infty(1), \tag{25}$$

but for the moment we shall work with a general ψ satisfying (24). Now multiply (20a) by ψ and integrate by parts twice over the interval $\varepsilon \leqslant x \leqslant 1$, where $\varepsilon > 0$ is arbitrary. This gives

$$-\int_\varepsilon^1 F(y_\infty)\,\psi = \int_\varepsilon^1 F(y_0)\,\psi$$

$$= f(x)\,\frac{dy_0}{dx}(x)\,\psi(x)\Big|_{x=\varepsilon}^{x=1}$$

$$- f(x)\,y_0(x)\,\frac{d\psi}{dx}(x)\Big|_{x=\varepsilon}^{x=1} + \int_\varepsilon^1 F(\psi)\,y_0. \tag{26}$$

Consider the endpoint terms in more detail. At $x = 1$, they reduce to

$$-f(1)\,y_0(1)\,\frac{d\psi}{dx}(1) = f(1)\,y_\infty(1)\,\frac{d\psi}{dx}(1)$$

by virtue of (24b) and (20b). Using the asymptotic properties (19) and (24a) of y_0 and ψ, respectively, and the fact that $f(x) \sim x^2$, it follows that in the limit as $\varepsilon \to 0+$ the $x = \varepsilon$ endpoint terms yields

$$\lim_{\varepsilon \to 0+} \left[f(\varepsilon)\,\frac{dy_0}{dx}(\varepsilon)\,\psi(\varepsilon) - f(\varepsilon)\,y_0(\varepsilon)\,\frac{d\psi}{dx}(\varepsilon) \right] = 2\mu a_s.$$

So taking the limit $\varepsilon \to 0+$ in (26) and rearranging we find

$$2\mu a_s = \lim_{\varepsilon \to 0+} \left(\int_\varepsilon^1 F(\psi)\,y_0 + \int_\varepsilon^1 F(y_\infty)\,\psi \right) + f(1)\,y_\infty(1)\,\frac{d\psi}{dx}(1). \tag{27}$$

This is the basic integral representation that we shall exploit. Note that if we were able to choose ψ so as to satisfy

$$F(\psi) = 0 \tag{28}$$

as well as (24), then (27) would provide a representation for a_s solely in terms of the known right-hand sides of (20). We could think of ψ as a Green's function for a_s in this case. Usually, however, we cannot readily satisfy (28), and must be content with (27) in its general form which includes the weighted integral of y_0. In this case we choose to call ψ a generalized Green's function. Its usefulness lies in the fact that if we replace y_0 in (27) by $\tilde{y}_0$, all other terms of (27) being known, we will obtain some approximation for a_s. Intuitively we might expect this approximation to be reasonably accurate, since it is obtained by a weighted averaging of $\tilde{y}_0$. We shall now show that this is the case.

From now on we only consider the particular choice $\psi = \psi_0$ given in (25). Note in this case that

$$F(\psi_0) = \begin{cases} O(x^{\mu+1/2+\min(0, N_\mu - 2\mu)}) + O(x) & \text{if } 2\mu \neq N_\mu \\ O(x^{\mu+1/2}\log x) + O(x) & \text{if } 2\mu = N_\mu. \end{cases}$$

Since $y_0 = O(x^{-1/2+\mu})$, it follows that the first integral of (27) is proper. We may therefore rewrite (27) as

$$2\mu a_s = \int_0^1 F(\psi_0)\, y_0 + \int_0^1 F(y_\infty)\, \psi_0 + f(1)\, y_\infty(1)\, \frac{d\psi_0}{dx}\,(1), \tag{29}$$

where the second integral may be improper. Denote by $\tilde{a}_s$ the approximation to a_s obtained from

$$2\mu \tilde{a}_s = \int_0^1 F(\psi_0)\, \tilde{y}_0 + \int_0^1 F(y_\infty)\, \psi_0 + f(1)\, y_\infty(1)\, \frac{d\psi_0}{dx}\,(1). \tag{30}$$

Subtracting (30) from (29) gives

$$2\mu(a_s - \tilde{a}_s) = \int_0^1 F(\psi_0)(y_0 - \tilde{y}_0), \tag{31}$$

showing that the error in $\tilde{a}_s$ is some weighted average of the error in the approximate solution $\tilde{y}_0$ over the entire interval. To gain some further theoretical understanding of $(a_s - \tilde{a}_s)$ let us introduce (just for the purpose of the error analysis) the auxiliary problem

$$\begin{aligned} &-F(\xi) = F(\psi_0), \\ &\xi(1) = 0 \\ &f(x)\,\xi' = O(x^{1/2}), \qquad x \to 0+. \end{aligned} \tag{32}$$

Just as for (20), this can be cast in a weak Galerkin form

$$\xi \in \mathscr{H}_0, \qquad W(\xi, v) = \int_0^1 F(\psi_0)\, v, \qquad \forall v \in \mathscr{H}_0. \tag{33}$$

In particular, making the choice $v = (y_0 - \tilde{y}_0)$ we obtain from (31),

$$\begin{aligned} 2\mu(a_s - \tilde{a}_s) &= \int_0^1 F(\psi_0)(y_0 - \tilde{y}_0) \\ &= W(\xi, y_0 - \tilde{y}_0) \\ &= W(\xi - u, y_0 - \tilde{y}_0) \end{aligned}$$

for any $u \in \tilde{\mathscr{H}}_0$ by (23a) (note that $W(\cdot, \cdot)$ is symmetric). Thus by (15),

$$\begin{aligned} |a_s - \tilde{a}_s| &= \frac{1}{2\mu} \inf_{u \in \tilde{\mathscr{H}}_0} |W(\xi - u, y_0 - \tilde{y}_0)| \\ &\leqslant C \inf_{u \in \tilde{\mathscr{H}}_0} \|\xi - u\| \, \|y_0 - \tilde{y}_0\| \\ &\leqslant C \inf_{u \in \tilde{\mathscr{H}}_0} \|\xi - u\| \inf_{v \in \tilde{\mathscr{H}}_1} \|y_0 - v\| \end{aligned} \tag{34}$$

using (23b). What we see here, is that the error in $\tilde{a}_s$ can be directly related to how well the solutions y_0 and ξ of (20) and (32) can be approximated by finite element functions from $\tilde{\mathscr{H}}_1$ and $\tilde{\mathscr{H}}_0$, respectively. This clearly will have a bearing on how we should choose the finite element subspace $\tilde{\mathscr{H}}_0$. We take this matter up in the next section.

4. Choice of Finite Element Subspace $\tilde{\mathscr{H}}$

4.1. *"Singular" Behaviour of y_0 and ξ*

The obvious conclusion to be drawn from the estimate (34) is that $\tilde{\mathscr{H}}_0$ should be chosen so as to optimize, in some sense, the accuracy of the approximation of y_0 and ξ by functions from $\tilde{\mathscr{H}}_1$ and $\tilde{\mathscr{H}}_0$ as measured in the norm $\|\cdot\|$. This accuracy will, to a large extent, be influenced by any "singular" behaviour that y_0 or ξ may display. Let us gather together some information on this aspect of y_0 and ξ. First, we know that y_0 is relatively well behaved away from $x = 0$, but near $x = 0$ it behaves as

$$\begin{aligned} y_0(x) &= O(x^{-1/2+\mu}), \\ y_0'(x) &= \begin{cases} O(x^{-3/2+\mu}) & \text{if } \mu \neq \frac{1}{2} \\ O(\log x) & \text{if } \mu = \frac{1}{2}, \end{cases} \\ y_0''(x) &= \begin{cases} O(x^{-5/2+\mu}) & \text{if } \mu \neq \frac{1}{2}, \frac{3}{2} \\ O(x^{-1}) & \text{if } \mu = \frac{1}{2} \\ O(\log x) & \text{if } \mu = \frac{3}{2}. \end{cases} \end{aligned} \tag{35}$$

A similar result holds for ξ. To see this, note that by (32) $F(\xi + \psi_0) = 0$, and so using (2) we must have $\xi = A_s y_\mu^{(s)} + A_b y_\mu^{(b)} - \psi_0$ for certain real constants A_s, A_b. Applying the boundary condition of (32) at $x = 1$ allows us to eliminate A_s, and thereby obtain

$$\xi(x) = -A_b \frac{y_\mu^{(b)}(1)}{y_\mu^{(s)}(1)} y_\mu^{(s)}(x) + A_b y_\mu^{(b)}(x) - \psi_0(x)$$

assuming that $y_\mu^{(s)}(1) \neq 0$ (see Appendix A). Requiring that $f\xi' = O(x^{1/2})$ implies, in

particular, that $A_b = 1$ to ensure that the $x^{-1/2-\mu}$ term of $\psi_0(x)$ is cancelled. In fact, the first $N_\mu + 1$ terms of $y_\mu^{(b)}$ are eliminated by this choice. Thus by (13) and (25),

$$\begin{aligned}\xi(x) &= a_s y_\mu^{(s)}(x) + (y_\mu^{(b)}(x) - y_\infty(x)) + xy'_\infty(1) \\ &= y_0(x) + xy'_\infty(1) \end{aligned} \tag{36}$$

using (18b).

To quantify the "singular" nature of the kinds of functions we are handling it is convenient to introduce a class of weighted norms. For any real number $0 \leqslant \beta < 1$ and suitable function u, define the norm

$$|||u|||_\beta = \left[\int_0^1 u^2 + x^{2\beta}(u')^2 + x^{2(1+\beta)}(u'')^2\right]^{1/2}.$$

Note that if u is a well-behaved function away from $x = 0$, and near $x = 0$ it satisfies $u = O(x^\alpha)$, $u' = O(x^{\alpha-1})$, and $u'' = O(x^{\alpha-2})$ with $\alpha > -\frac{1}{2}$ then $|||u|||_\beta < \infty$ as long as $\beta > 1/2 - \alpha$. In particular, by (35) and (36)

$$|||y_0|||_\beta,\ |||\xi|||_\beta < \infty \qquad (1-\mu < \beta < 1), \tag{37}$$

with the exceptional case of $\mu = \frac{1}{2}$,

$$|||y_0|||_\beta,\ |||\xi|||_\beta < \infty, \qquad 0 \leqslant \beta < 1. \tag{38}$$

4.2. *The Finite Element Mesh*

Consider a partition of the interval [0, 1],

$$0 = x_0 < x_1 < \cdots < x_M = 1, \tag{39}$$

say. For the moment we shall only consider linear elements; that is, we consider functions which are linear on each element (x_i, x_{i+1}) $(i = 0,..., M-1)$ individually, and which are continuous at each internal nodal point x_i $(i = 1,..., M-1)$. Any such function can be expanded as

$$v(x) = \sum_{i=0}^{M} v(x_i)\,\varphi_i(x),$$

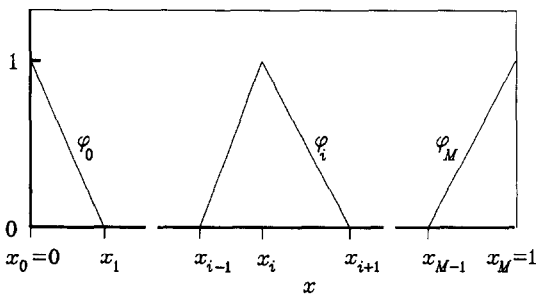

FIG. 1. The linear shape functions φ_i.

where the φ_i are local "shape" functions (see Fig. 1) given by

$$\varphi_0 = \begin{cases} \dfrac{x_1 - x}{x_1}, & 0 \leqslant x \leqslant x_1 \\ 0, & x > x_1, \end{cases}$$

$$\varphi_i = \begin{cases} \left(\dfrac{x - x_{i-1}}{x_i - x_{i-1}}\right), & x_{i-1} \leqslant x \leqslant x_i \\ \left(\dfrac{x_{i+1} - x}{x_{i+1} - x_i}\right), & x_i \leqslant x \leqslant x_{i+1} \\ 0, & x < x_{i-1} \text{ or } x > x_{i+1} \qquad (i = 1,..., M-1), \end{cases} \tag{40}$$

$$\varphi_M = \begin{cases} 0, & x < x_{M-1} \\ \left(\dfrac{x - x_{M-1}}{1 - x_{M-1}}\right), & x_{M-1} \leqslant x \leqslant 1. \end{cases}$$

We choose $\tilde{\mathscr{H}}_0$ to be the linear span of φ_0, φ_1,..., φ_{M-1}; that is, the space of all functions of the form

$$v(x) = \sum_{i=0}^{M-1} v_i \varphi_i(x), \qquad v_i \text{ real constant.} \tag{41}$$

Clearly $\tilde{\mathscr{H}}_0$ is a subspace of $\mathscr{H}_0$, and the corresponding $\tilde{\mathscr{H}}_1$ is the set of all functions of the form

$$v(x) = \sum_{i=0}^{M-1} v_i \varphi_i(x) - y_\infty(1)\, \varphi_M(x), \qquad v_i \text{ real constant.} \tag{42}$$

In terms of the φ_i the finite element equations (22b) may be written as a system of linear equations for the unknown nodal values $\tilde{y}_0(x_i)$ $(i = 0,..., M-1)$,

$$\sum_{i=0}^{M-1} W(\varphi_i, \varphi_j)\, \tilde{y}_0(x_i) = \begin{cases} \displaystyle\int_0^1 F(y_\infty)\, \varphi_j, & j = 0,..., M-2 \\ \displaystyle\int_0^1 F(y_\infty)\, \varphi_{M-1} + W(\varphi_M, \varphi_{M-1})\, y_\infty(1), & j = M-1. \end{cases} \tag{43}$$

Because of the local support of the shape functions φ_i, the matrix $W(\varphi_i, \varphi_j)$ is tridiagonal. This permits very efficient solution procedures.

To completely specify $\tilde{\mathscr{H}}_0$ we need to decide upon a particular partition (39). The simplest choice would be a uniform partition. As we shall see in Section 5 where we discuss some numerical examples. Such a choice often proves quite satisfactory. Indeed, we would expect a uniform mesh to be reasonable away from $x = 0$, for, as

we saw in Section 4.1, both y_0 and ξ are well behaved there. It is around $x=0$, where y_0 and ξ have some form of "singular" behaviour, that a nonuniform mesh may be called for. For a given positive integer M, we would ideally like a prescription for a partition (39) that leads to optimal approximations for y_0 and ξ in the norm $\|\cdot\|_\infty$. This of course is asking too much; however, we can at least say something in this respect in an asymptotic sense (i.e., as $M \to \infty$). First, let us define some classes of nonuniform meshes.

A convenient method of describing nonuniform meshes is by means of a mesh grading function. If Δ is a strictly increasing function from $[0, 1]$ to $[0, 1]$ such that $\Delta(0)=0$ and $\Delta(1)=1$, we shall call Δ a mesh grading function. For any positive integer M we can generate a partition of $[0, 1]$ from Δ by

$$x_i = \Delta\left(\frac{i}{M}\right), \qquad i=0,\ldots, M. \tag{44}$$

For instance, if $\Delta(t)=t$ then (44) generates a uniform mesh; while if $\Delta(t)=t^\gamma$, $\gamma>1$, then the partition is nonuniform, with mesh points being concentrated near $x=0$ (see Fig. 2). We can think of such mesh grading functions as relating a nonuniform mesh in the original variable $x=\Delta(t)$ to a uniform mesh in the "stretched" variable $t=x^{1/\gamma}$.

Let $D>0$ and $0\leqslant\lambda<1$. We shall say that the partition (39) is of class (D, λ) if

$$\text{(i)} \quad x_{i+1}-x_i \leqslant \frac{D}{M}\, x_i^\lambda, \quad i=1,\ldots, M-1 \tag{45a}$$

$$\text{(ii)} \quad D^{-1}\,\frac{1}{M^{1/(1-\lambda)}} \leqslant x_1 \leqslant \frac{D}{M^{1/(1-\lambda)}}. \tag{45b}$$

Meshes of class (D, λ) are relatively easy to construct. For instance,

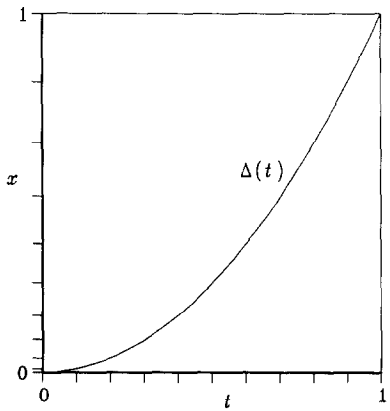

FIG. 2. Mesh grading function $\Delta(t)$. A uniform mesh in the t-variable generates a nonuniform mesh in the x-variable.

LEMMA 4.1. *All meshes generated by* (44) *using the mesh grading function* $\Delta(t)=t^{\gamma}$, $\gamma \geqslant 1$, *are of class* $(D(\gamma),\ 1-(1/\gamma))$ *where* $D(\gamma)>0$ *is some number depending on* γ.

Proof. See Appendix B. ∎

We shall say that $\tilde{\mathscr{H}}_0$, $\tilde{\mathscr{H}}_1$ are of class (D, λ) if the corresponding mesh (39) is of class (D, λ).

THEOREM 4.2. (i) *Suppose* $\tilde{\mathscr{H}}_1$ *is of class* (D, λ), *and that* $|||\, y_0 \,|||_\lambda < \infty$, *then*

$$\inf_{v \in \tilde{\mathscr{H}}_1} \| y_0 - v \| \leqslant \frac{C}{M} |||\, y_0 \,|||_\lambda,$$

where $C = C(D, \lambda) > 0$ *is a constant depending only on* D *and* λ.

(ii) *Suppose* $\tilde{\mathscr{H}}_0$ *is a class* (D, λ), *and that* $|||\, \xi \,|||_\lambda < \infty$, *then*

$$\inf_{u \in \tilde{\mathscr{H}}_0} \| \xi - u \| \leqslant \frac{C}{M} |||\, \xi \,|||_\lambda,$$

where $C = C(D, \lambda) > 0$ *is a constant depending only on* D *and* λ.

Proof. See Appendix B. ∎

An immediate corollary is

COROLLARY 4.3. *Suppose* $\tilde{\mathscr{H}}_0$, $\tilde{\mathscr{H}}_1$ *are of class* (D, λ) *where* $\lambda > 1-\mu$, $\lambda \geqslant 0$ *then both*

$$\inf_{v \in \tilde{\mathscr{H}}_1} \| y_0 - v \|, \qquad \inf_{u \in \tilde{\mathscr{H}}_0} \| \xi - u \| = O(M^{-1}). \tag{46}$$

In particular, (46) *will hold if the partition* (39) *is generated by the mesh grading function*

$$\Delta(t) = t^{\gamma}, \tag{47}$$

where

$$\gamma > \frac{1}{\mu}, \qquad \gamma \geqslant 1.$$

For the exceptional case $\mu = \frac{1}{2}$, *the choice* $\lambda = 0$, $\gamma = 1$ (*uniform mesh*) *will give* (46).

Proof. The result (46) follows from Theorem 4.2 upon recalling (37) (or (38) if $\mu = \frac{1}{2}$). Lemma 4.1 then implies the special case covered by (47). ∎

The practical conclusion to be drawn from this corollary is that by a suitable mesh grading we can always ensure that y_0 and ξ are approximated by $\tilde{\mathscr{H}}_0$ and $\tilde{\mathscr{H}}_1$ with an error in the norm $\|\cdot\|$ of $O(M^{-1})$ as $M \to \infty$, where M is the number of

mesh points. If $\mu > 1$, then a uniform mesh ($\Delta(t) = t$) will achieve this asymptotic rate, whereas if $0 < \mu < 1$ a graded mesh, such as that generated by (47), will in general be necessary. The asymptotic rate of $O(M^{-1})$ predicted by the corollary is actually optimal, in the sense that no matter what partition (39) is used, $\mathscr{H}_0$ cannot approximate in the norm $\|\cdot\|$ any better than $O(M^{-1})$, except in some special cases (e.g., linear functions on $[0, 1]$).

Returning now to our primary objective of approximating a_s, the estimate (34) shows that if the partition (39) is chosen as above then

$$|a_s - \tilde{a}_s| = O(M^{-2}). \tag{48}$$

4.3. *Singular Elements*

In the last section we utilized mesh gradings in an attempt to overcome the problems caused by the singular behaviour of y_0 and ξ near $x = 0$ when $\mu \leqslant 1$. From one point of view, however, this is not a particularly efficient remedy, especially as $\mu \to 0$. To see why, consider the mesh grading function (47), say. For any mesh generated by this grading function the proportion θ of mesh points located in the subinterval $(0, \delta)$ of $(0, 1)$ is $\theta = \delta^{1/\mu} > \delta^{\mu}$. Note that $\theta \to 1$ as $\mu \to 0+$; that is, for sufficiently small values of μ, the mesh grading function (47) is packing the majority of mesh points into the (arbitrarily) small neighbourhood $(0, \delta)$. All these mesh points will have a minimal influence on how well the "smooth" parts of y_0 and ξ can be approximated; their only role is to approximate the singular behaviour in $(0, \delta)$. But the form of this singularity is precisely known from (19); and so it would seem to make sense to somehow use this information directly in the finite element approximation, rather than rely upon the finite element equations themselves to discover it at the cost of a large number of extra degrees-of-freedom.

One technique for doing this is to enrich the finite element spaces $\mathscr{H}_0$, $\mathscr{H}_1$ by the addition of a so-called singular shape function which has the required $x^{\alpha_\mu^{(s)}}$-type singular behaviour built into it. However, there are conflicting considerations surrounding the use of such functions. On one hand, computational efficiency would tend to favour singular shape functions with only local support: that is, they should overlap with as few as possible of the "regular" shape functions φ_i of Section 4.2. The system of linear algebraic equations arising from (22b) would then remain relatively sparse. On the other hand, the singular nature of $x^{\alpha_\mu^{(s)}}$ is felt on some fixed (i.e., mesh independently) subinterval of $(0, 1)$. The support of the singular shape function should therefore spread over this entire subinterval; but as the mesh size decreases, this leads to overlap with more and more regular shape functions, and a consequent loss of sparsity in the resulting algebraic system.

We shall adopt a combined mesh grading and singular element approach, which attempts to minimize the "wastage" of mesh points implicit in the mesh gradings of Section 4.3, yet maintains the tridiagonal character of the resulting algebraic system. We restrict ourselves to the case $\mu < \frac{1}{2}$. This is not an unreasonable restriction, since the difficulties with the mesh grading approach mentioned above only became significant as $\mu \to 0$.

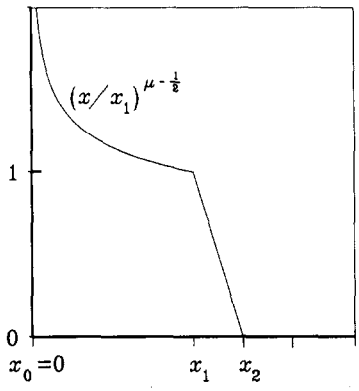

FIG. 3. The singular shape function Φ_1, in the case $\mu = 0.3$.

Given the partition (39), define the singular shape function (see Fig. 3),

$$\Phi_1(x) \equiv \begin{cases} \left(\dfrac{x}{x_1}\right)^{\alpha_\mu^{(s)}}, & 0 < x \leqslant x_1 \\ \left(\dfrac{x_2 - x}{x_2 - x_1}\right), & x_1 < x \leqslant x_2 \\ 0, & x > x_2 \end{cases} \tag{49a}$$

and the regular shape functions

$$\Phi_i(x) \equiv \varphi_i(x), \qquad i = 2, \ldots, M. \tag{49b}$$

Note that the support of Φ_1 extends only over the interval $0 \leqslant x \leqslant x_2$. Observe also that any function v in the linear span of the Φ_i's can be written

$$v(x) = \sum_{i=1}^{M} v(x_i)\, \Phi_i(x).$$

We define the finite element subspace $\tilde{\mathscr{H}}_0$ by

$$\tilde{\mathscr{H}}_0 = \left\{ v \colon v(x) = \sum_{i=1}^{M-1} v_i \Phi_i(x),\ v_i \text{ real constant} \right\}.$$

Clearly $\tilde{\mathscr{H}}_0 \subseteq \mathscr{H}_0$, and the corresponding translated space $\tilde{\mathscr{H}}_1$ is simply the set of all functions of the form

$$v(x) = \sum_{i=1}^{M-1} v_i \Phi_i(x) - y_\infty(1)\, \Phi_M(x), \qquad v_i \text{ real constant.}$$

In place of the linear system (43) we now have

$$\sum_{i=1}^{M-1} W(\Phi_i, \Phi_j,) \tilde{y}_0(x_i) = \begin{cases} \int_0^1 F(y_\infty)\, \Phi_j, & j = 1,..., M-2 \\ \int_0^1 F(y_\infty)\, \Phi_{M-1} + W(\Phi_M, \Phi_{M-1})\, y_\infty(1), & j = M-1. \end{cases}$$

The local support of both the regular and singular shape functions guarantees that the matrix $W(\Phi_i, \Phi_j)$ will be tridiagonal.

Again let $D > 0$ and $0 \leqslant \lambda < 1$. We shall say that the partition (39) is of class $(D, \lambda, *)$ if it has the following properties

$$\text{(i)} \quad \frac{D^{-1}}{M} \leqslant x_1 \leqslant \frac{D}{M}, \tag{50a}$$

$$\text{(ii)} \quad x_{i+1} - x_i \leqslant \frac{D}{M} x_i^\lambda, \qquad i = 1,..., M-1. \tag{50b}$$

Partitions satisfying (50) are relatively easy to generate. For instance,

LEMMA 4.4. *Let* $L > 0$, $\gamma \geqslant 1$. *Define the partition* (39) *by*

$$\begin{aligned} x_0 &= 0, \\ x_1 &= \frac{L}{M}, \\ x_i &= \left[\left(\frac{L}{M}\right)^{1/\gamma} + \left(1 - \left(\frac{L}{M}\right)^{1/\gamma}\right) \frac{i-1}{M-1}\right]^\gamma, \qquad i = 2,..., M-1, \\ x_M &= 1, \end{aligned} \tag{51}$$

then, provided $M > \max(2, L)$, *the partition is of class* $(D(\gamma, L), 1 - (1/\gamma), *)$, *where* $D(\gamma, L) > 0$ *is some number depending on* γ *and* L, *but not* M.

Proof. See Appendix C. ∎

Note that in the limit as $\gamma \to \infty$ (with M fixed) the partition (51) becomes

$$\begin{aligned} x_0 &= 0, \\ x_1 &= \frac{L}{M}, \\ x_i &= \left(\frac{L}{M}\right)^{(M-i)/(M-1)}, \qquad i = 2,..., M-1, \\ x_M &= 1. \end{aligned}$$

In particular, the mesh points do not accumulate around a single point as occurred for the corresponding limiting case of (47).

We shall say that $\tilde{\mathscr{H}}_0$, $\tilde{\mathscr{H}}_1$ are of class (D, λ, μ) if the partition (39) is of class $(D, \lambda, *)$. The third parameter μ is included to clearly indicate that the definition of the singular shape function Φ_1 by (49a) depends on μ.

THEOREM 4.5. *Suppose* $\tilde{\mathscr{H}}_0$, $\tilde{\mathscr{H}}_1$ *are of class* (D, λ, μ) *where* $\lambda > 1 - \mu$, $\lambda \geqslant 0$ *then both*

$$\inf_{v \in \tilde{\mathscr{H}}_1} \|y_0 - v\|, \qquad \inf_{u \in \tilde{\mathscr{H}}_0} \|\xi - u\| = O(M^{-1}). \tag{52}$$

In particular, (52) *will hold if the partition* (39) *is generated by* (51) *with* $\gamma > 1/\mu$, $\gamma \geqslant 1$.

Proof. See Appendix C. ∎

Thus, by the combined mesh grading, singular element approach we are again able to approximate y_0 and ξ to within $O(M^{-1})$ in $\|\cdot\|$. The estimate (48) will also follow just as in Section 4.2.

In Section 4.2 and 4.3 we have really only taken account of the behaviour of y_0 and ξ near $x = 0$. Away from $x = 0$ both functions were assumed to be relatively well behaved. Our mesh grading strategy was based on the premise that the singular behaviour near $x = 0$ was the principle source of numerical error. While this will certainly be true in the limit as $M \to \infty$, it need not be the case in practice when, of course, we are dealing with some finite M. Thus, although Theorems 4.2, 4.3, 4.5, etc., will, strictly speaking, remain valid in such circumstances, the conclusions drawn from them may not yield particularly efficient meshes. For instance, according to the strategy outlined above, uniform meshes should be used whenever $\mu > 1$; however, it would seem reasonable that for "large" values of μ some form of mesh grading towards $x = 1$ may be justified. A more careful analysis is needed to cover such cases than we can go into here.

5. NUMERICAL EXAMPLES

In this section we shall present a variety of numerical examples to illustrate some of the points made in our earlier theoretical discussion.

5.1. *A Class of Test Problems*

Our test problems will all be of the simple form

$$\frac{d}{dx}\left(x^2 \frac{dy}{dx}\right) - (a + bx + cx^2)\, y = 0, \qquad 0 < x < 1,$$

$$y(1) = 0, \tag{53}$$

where $a > -\frac{1}{4}$, b and c are constants. The parameter μ is then given by

$$\mu = (\tfrac{1}{4} + a)^{1/2} > 0. \tag{54}$$

After some manipulation, the Frobenius solutions (5) and (7) can be recognized as (assuming $c \neq 0$)

$$y_\mu^{(s)}(x) = x^{-(1/2)+\mu} e^{-\theta x/2} M(\tfrac{1}{2} + \mu + \kappa,\ 1 + 2\mu,\ \theta x) \tag{55}$$

and provided $2\mu \neq 1, 2, 3, \ldots$,

$$y_\mu^{(b)}(x) = x^{-(1/2)-\mu} e^{-\theta x/2} M(\tfrac{1}{2} - \mu + \kappa,\ 1 - 2\mu,\ \theta x), \tag{56}$$

where $\theta = 2c^{1/2}$, $\kappa = b/\theta$ and $M(\cdot, \cdot, \cdot)$ is the confluent hypergeometric Kummer function [15]. In the case $\kappa = 0$ (i.e., $b = 0$), (55) and (56) become expressible in terms of either Bessel functions or modified Bessel functions depending on the sign of c (see [15]).

As discussed in Section 2, we need to redefine the second solution $y_\mu^{(b)}$ given by (56) if we are to have an acceptable pair of linearly independent solutions for a full range of μ's. Proceeding as in Section 2, we choose to define

$$\begin{aligned} \gamma(\mu) &= \sin 2\mu\pi\ p_{N_\mu}^{(b)}(\mu), \qquad && 2\mu \neq 1, 2, 3, \ldots, \\ \gamma\left(\frac{m}{2}\right) &= \operatorname*{lmt}_{\mu \to m/2} \gamma(\mu), && m = 1, 2, 3, \ldots, \end{aligned} \tag{57}$$

where we have written $p_{N_\mu}^{(b)}(\mu)$ to clearly indicate that $p_{N_\mu}^{(b)}$ depends on μ other than just through the index N_μ (see (8)). We may then redefine $y_\mu^{(b)}$ by (9). Admittedly, γ will in general have discontinuities at the odd quarter-integers where the index N_μ jumps; however, this causes us no problem. Note that the customary second solution of a confluent hypergeometric equation, in terms on either the Kummer function U or the Whittaker function W (see [15]), is not entirely satisfactory for our purpose, since both U and W become increasingly linearly dependent on M as certain combinations of their parameters are approached.

For the special case $c = 0$, the solutions corresponding to (55) and (56) are

$$y_\mu^{(s)}(x) = x^{-(1/2)+\mu} \sum_{n=0}^{\infty} \frac{b^n}{(1+2\mu)_n} \frac{x^n}{n!} \tag{58}$$

and

$$y_\mu^{(b)}(x) = x^{-(1/2)-\mu} \sum_{n=0}^{\infty} \frac{b^n}{(1-2\mu)_n} \frac{x^n}{n!}, \tag{59}$$

where $(\cdot)_n$ is Pochhammer's symbol, defined by $(z)_0 = 1$, $(z)_n = (z)_{n-1}(z+n-1)$. Again $y_\mu^{(b)}$ may be redefined using (57) and (9).

The various choices of a, b, and c that we shall consider are listed in Table I.

TABLE I

Example	a, b, c (zero unless otherwise stated)	μ	$\alpha_\mu^{(s)}$	$\alpha_\mu^{(b)}$
1	$c=-3$	0.5	0	-1
2	$a=20, c=1$	4.5	4	-5
3	$a=0.8525, b=-1.86, c=0.36$	1.05	0.55	-1.55
4	$a=-0.0475, b=2$	0.45	-0.05	-0.95
5	$a=-0.1875, c=1$	0.25	-0.25	-0.75
6	$a=-0.24, c=1$	0.1	-0.4	-0.6

These will give values for the small exponent $\alpha_\mu^{(s)}$ ranging from the badly singular $\alpha_\mu^{(s)}=-0.4$, to the well behaved $\alpha_\mu^{(s)}=4$. Because of the relative simplicity of the Frobenius expansions (55), (56), (58), and (59) it is not difficult to find exact values for a_s using (13); these can then be used to judge the accuracy of our approximations $\tilde{a}_s$.

We shall also be interested in comparing our post-processed approximation for a_s with that obtained by other, more "direct" methods. Given the asymptotic relation (19), perhaps the most straightforward approach to finding a_s is simply to take the limit

$$a_s = \operatorname*{lmt}_{x\to 0+} \frac{y_0(x)}{x^{\alpha_\mu^{(s)}}}. \tag{60}$$

This suggests the approximation

$$a_s^* = \frac{\tilde{y}_0(x^*)}{(x^*)^{\alpha_\mu^{(s)}}}, \tag{61}$$

where x^* is some mesh point "close" to $x=0$. (Of course, in the case $\alpha_\mu^{(s)}=0$, we could use $x^*=0$.) This method may be thought of as a pointwise "fitting" of $\tilde{y}_0$ to the known asymptotic form of y_0. There is, of course, no reason why this approach should only utilize the first term of (19); more sophisticated versions of (61) could fit $\tilde{y}_0$ to more and more terms of (19) at x^*. We shall, however, only consider the basic fitting method (61).

We may estimate the error in a_s^* by writing

$$(a_s - a_s^*) = \frac{1}{(x^*)^{\alpha_\mu^{(s)}}} (y_0(x^*) - \tilde{y}_0(x^*)) + \left(a_s - \frac{y_0(x^*)}{(x^*)^{\alpha_\mu^{(s)}}}\right). \tag{62}$$

Using (19) we find that the second term of (62) behaves like

$$\begin{aligned} &O((x^*)^{1+\min(0, N_\mu - 2\mu)}) && \text{if} \quad 2\mu \neq N_\mu \\ &O(x^* \log x^*) && \text{if} \quad 2\mu = N_\mu. \end{aligned} \tag{63}$$

Note that if a more sophisticated fitting method were used, the analog of this term could be made higher order in x^*. However, such modifications of (61) would not appreciably alter the first term of (62); it would retain the form

$$\frac{y_0(x^*) - \tilde{y}_0(x^*)}{(x^*)^{\alpha_\mu^{(s)}}(1 + o(1))}. \tag{64}$$

Thus, there are two considerations influencing the accuracy of a_s^*: first, x^* needs to be close to $x = 0$ so that (63) is sufficiently small, and second, $\tilde{y}_0(x^*)$ must be accurate enough for (64) also to be small. As we shall see in some of the examples, these are sometimes conflicting considerations.

5.2. *Discussion of Examples (see Figs. 4 and 5)*

EXAMPLE 1. Here a sequence of uniform meshes was used, with the spaces $\mathscr{H}_0$, $\mathscr{H}_1$ being as described in Section 4.2. As can be seen, both $\tilde{a}_s$ and a_s^* (with $x^* = 0$) appear to converge as M^{-2} (curves (a) and (b)). On the other hand, if the fitting point x^* is chosen to be x_1, the nearest mesh point to $x = 0$, then a_s^* only converges at a rate of M^{-1}. This is consistent with (61); for if we suppose on the basis of curve (b), that (64) converges as M^{-2}, then the major contribution to the error in a_s^* will come from (63), which in this case behaves as $O(x^*) = O(M^{-1})$. (The extra $\log x^*$ factor of (63) not being present in this particular case.) Note also that the error for (c) is some two orders of magnitude greater than that for either (a) or (b).

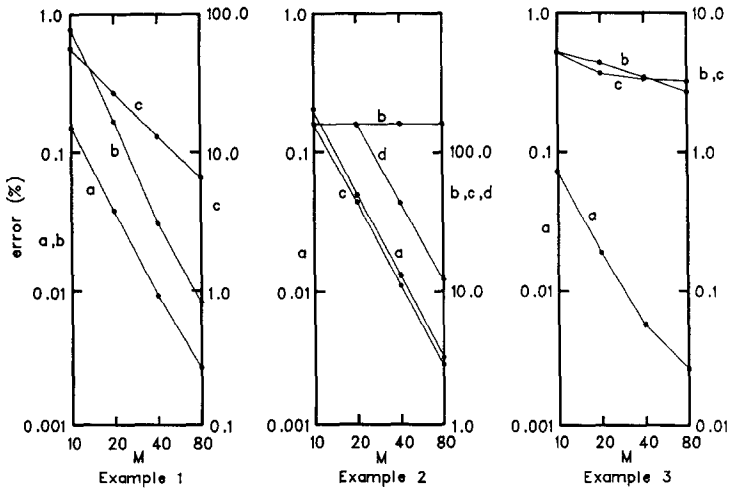

FIG. 4. Relative error $|a_s - a_s^{\text{est}}|/|a_s|$ in estimates for the coefficient a_s of the small solution, using (a) the generalized Green's function method $[a_s^{\text{est}} \equiv \tilde{a}_s$, see Eq. (30)], and (b)–(d) the generalized derivative method $[a_s^{\text{est}} \equiv a_s^*$, see Eq. (61)], versus the number of mesh points M (uniform mesh). Example 1 $\mu = 0.5$, (b) $x^* = 0$, (c) $x^* = x_1$. Example 2 $\mu = 4.5$, (b) $x^* = x_1$, (c) $x^* = 0.1$, (d) $x^* = 0.05$. Example 3 $\mu = 1.05$, (b) $x^* = x_1$, (c) $x^* = 0.1$.

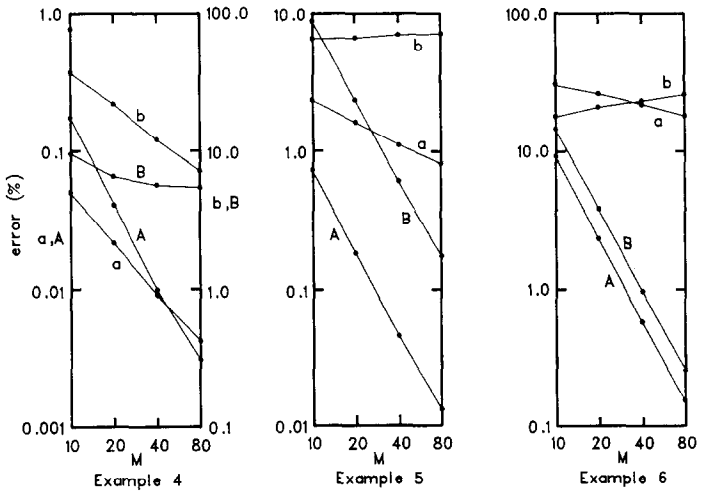

FIG. 5. Relative errors, as in Fig. 4: (a) and (A) using generalized Green's function method, and (b) and (B) using generalized derivative method [(b) $x^* = x_1$, (B) $x^* = 0.01$] versus M, the number of mesh points. In cases (a) and (b) a uniform mesh is used, while a graded mesh is used in cases (A) and (B). Example 4 $\mu = 0.45$. Example 5, $\mu = 0.25$. Example 6, $\mu = 0.1$. See text for further details.

EXAMPLE 2. In this example, the same sequence of uniform meshes and spaces $\mathscr{H}_0$, $\mathscr{H}_1$ as in Example 1 was used. We see again that $\tilde{a}_s$ converges as M^{-2} (curve (a)). The behaviour of a_s^*, however, is now quite different. Choosing $x^* = x_1$, the closest mesh point to $x = 0$, gives an approximation a_s^* which does not seem to converge (curve (b)), whereas placing x^* at some fixed point (i.e., the same point for all meshes) seems to give an M^{-2} rate of convergence (curves (c) and (d)). To see what this means, again consider (62). For the choice $x^* = x_1$, the term (63) will obviously tend to 0 as $x^* \to 0$. So the obvious conclusion is that the term (64) does not converge as $M \to \infty$. Note that this implies that no generalization of (61) based only on increasing the order of (63) will be successful. This sort of behaviour is perhaps not unexpected, since (64) corresponds in some sense to taking an "$\alpha_\mu^{(s)}$th derivative" of $\tilde{y}_0$ at $x = 0$. On the other hand, if we fix x^*, the term (64) should converge as $M \to \infty$, but now (63) will not. The apparent M^{-2} rate of convergence of (c) and (d) presumably means that for the range of M under consideration, the contribution of (64) overwhelms that of (63). However, as $M \to \infty$, the (fixed) term (63) must ultimately dominate and the curves (c) and (d) will flatten out. Note that the curves (c) and (d) represent errors some three orders of magnitude greater than that of curve (a).

EXAMPLE 3. In this example the value of μ, and consequently that of $\alpha_\mu^{(s)}$, has been decreased in comparison to Example 2. The uniform meshes and associated spaces $\mathscr{H}_0$, $\mathscr{H}_1$ of the previous examples have been retained. Again we see that $\tilde{a}_s$ appears to converge, at least initially, as M^{-2}. (The slight flattening of curve (a) from $M = 40$ to $M = 80$ is conjectured to be caused by accumulated numerical

roundoff. Bear in mind that for $M = 80$, a_s and $\tilde{a}_s$ differ only in the fifth significant figure.) Turning now to a_s^*, we see that a choice of $x^* = x_1$ gives an apparent rate of convergence of $M^{-0.3}$ (curve (b)); whereas letting x^* be fixed (curve (c)) seems to lead to the flattening out of the error predicted in the discussion of Example 2.

EXAMPLE 4. For this example $\mu = 0.45$. Our theory of Section 4 predicted that uniform meshes were not "optimal" in this case. In (a) and (b) the errors in $\tilde{a}_s$ and a_s^* (for $x^* = x_1$) are plotted for a sequence of uniform meshes. These appear to converge as $M^{-1.2}$ and $M^{-0.8}$, respectively. Next, a combined mesh grading and singular element approach was taken. In particular the meshes were constructed by taking

$$\begin{aligned} x_0 &= 0, \\ x_1 &= 0.01, \\ x_i &= \left[(.01)^{1/\gamma} + (1 - (.01)^{1/\gamma}) \frac{i-1}{M-1} \right]^{\gamma}, \qquad i = 2,..., M-1, \\ x_M &= 1, \end{aligned} \tag{65}$$

with $\gamma = 2$ ($\approx 1/\mu$) and $M = 10, 20, 40, 80$. The spaces $\tilde{\mathscr{H}}_0$, $\tilde{\mathscr{H}}_1$ were then constructed as described in Section 4.3. The corresponding errors in $\tilde{a}_s$ and a_s^* (with $x^* = 0.01$) are shown in curves (A) and (B), respectively. Observe that we have regained the M^{-2} rate of convergence for $\tilde{a}_s$, and that we again see the flattening out of the error in a_s^* that we have come to expect for cases where x^* is fixed. For the range of M considered here, there is little difference between the uniform and graded meshes as far as the accuracy of $\tilde{a}_s$ is concerned. Indeed, the uniform mesh leads to a more accurate value to start with ($M = 10, 20$).

The mesh (65) differs slightly from that described, say, by (51), in that x_1 is fixed. Of course we must imagine x_1 as having an M^{-1}-type behaviour as $M \to \infty$. We chose the value x_1 on the basis that it should approximately equal M^{-1} for the largest M being considered ($M = 80$).

EXAMPLE 5. This example with $\mu = 0.25$ ($\alpha_\mu^{(s)} = -0.25$) represents a more singular case than the previous one. With a uniform mesh, $\tilde{a}_s$ now only converges as $M^{-0.5}$ whereas a_s^* (with $x^* = x_1$) does not appear to converge at all. With the combined graded mesh, singular element approach described for the previous example, but now with $\gamma = 4$ ($= 1/\mu$), we obtain an apparent M^{-2} rate of convergence for both $\tilde{a}_s$ and a_s^* (with $x^* = 0.01$).

EXAMPLE 6. In this example we take μ even closer to the critical case $\mu = 0$. For uniform meshes $\tilde{a}_s$ now converges as $M^{-0.25}$, and again a_s^* (with $x^* = x_1$) does not seem to converge at all. The combined graded mesh, singular shape function approach with $\gamma = 10$ ($= 1/\mu$) in (65) suffices to give an apparent M^{-2} rate of convergence for both $\tilde{a}_s$ and a_s^* (with $x^* = 0.01$).

Observe that although in Examples 4–6 we have always been able to achieve an M^{-2} rate of convergence for $\tilde{a}_s$ by suitable mesh grading, etc., the general level of error has tended to increase as μ approaches 0. This is consistent with the theory outlined in Section 4, where a more careful analysis would show that the generic constants appearing in the various estimates become unbounded as $\mu \to 0+$ ($\gamma \to \infty$, $\lambda \to 1-$, etc.).

6. Conclusions

We have presented an efficient and accurate finite element method for computing the coefficient of the leading term of the recessive solution of an ordinary differential equation at a regular singular point, given an appropriate normalization of the dominant solution and a boundary condition at a finite distance from the singular point.

Such singular differential equations arise in the theory of resistive modes in a slab or cylindrical plasma with high magnetic Reynold's (Lundquist) number, [1]. This was the motivation for the development of this method. The advantage of the Galerkin formulation is its ready generalization to partial differential equations with singular surfaces, as are encountered in a toroidal or helical plasma. The Frobenius expansion of the ideal marginal equation near a singular surface is more complicated in two dimensions, but can be carried out compactly in operator notation [17], or in Fourier representation when straight-magnetic-field-line coordinates [18] are used. The weak solutions do not, as in the one-dimensional case, separate into those with support to one side or the other of the singular layer, but instead are global. The Galerkin method handles such global problems easily.

It is anticipated that the present method, when generalized to two dimensions, and implemented in the PEST 3 program, will represent a significant improvement over the singular element method used previously [8, 9]. It should be noted that the present method shares with all boundary layer methods the problem that, when the exponent $-\frac{1}{2}+\mu$ of the recessive (small) solution differs from the exponent $-\frac{1}{2}-\mu$ of the dominant (big) solution by a large number, then many terms of the Frobenius expansion of the big solutions will need to be computed (otherwise, the remainder would overwhelm the small solution). Thus the boundary layer approach would not be very appropriate for such cases, unless an analysis of the physics of the inner layer showed that the big solutions are negligible at such singular layers. If this were the case, then ideal MHD would apply in the neighbourhood of singular layers for which $\mu \gg 1$, and so the standard Galerkin approach [6] could be used.

Another area where our analysis of the outer problem suggests a re-examination of the inner, boundary layer theory is in the appropriate definitions of the "big" and "small" solutions. To construct a theory uniformly valid for all μ, whether near a half integer or not, we were led to redefine the big solution to be a linear com-

bination of the big and small solutions, as conventionally defined [2, 3]. This should be reflected also in the asymptotic analysis of the inner layer equations as $|x| \to \infty$, where x is the scaled distance from the singular surface.

THEOREM A.1. *If* (16) *holds, then* $g_0 > -\frac{1}{4}$.

Proof. For any $\alpha > -\frac{1}{2}$ and $1 > \delta > 0$, define

$$v = \begin{cases} x^\alpha - \delta^\alpha, & 0 < x \leqslant \delta \\ 0, & \delta \leqslant x \leqslant 1. \end{cases}$$

It is readily verified that $v \in \mathscr{H}_0$. Let us evaluate

$$\begin{aligned} W(v, v) &= \int_0^\delta f(v')^2 + gv^2 \\ &= \int_0^\delta (x^2 + O(x^3))(v')^2 + (g_0 + O(x))\, v^2 \\ &= \left(\frac{\alpha^2}{2\alpha+1} + g_0\left(\frac{1}{2\alpha+1} - \frac{2}{\alpha+1} + 1\right)\right)\delta^{2\alpha+1} + O(\delta^{2\alpha+2}) \end{aligned}$$

by our assumptions on f and g. Likewise

$$\|v\|^2 = \left(\frac{\alpha^2}{2\alpha+1} + \frac{1}{2\alpha+1} - \frac{2}{\alpha+1} + 1\right)\delta^{2\alpha+1}.$$

By (16), $W(v, v) \geqslant \eta \|v\|^2$ and so we have after dividing by $\delta^{2\alpha+1}$,

$$\begin{aligned} &\frac{\alpha^2}{2\alpha+1} + g_0\left(\frac{1}{2\alpha+1} - \frac{2}{\alpha+1} + 1\right) + O(\delta) \\ &\qquad \geqslant \eta\left(\frac{\alpha^2}{2\alpha+1} + \frac{1}{2\alpha+1} - \frac{2}{\alpha+1} + 1\right). \end{aligned}$$

Letting $\delta \to 0+$, multiplying through by $2\alpha + 1$, and then letting $\alpha \to -\frac{1}{2}+$ we find

$$\tfrac{1}{4} + g_0 \geqslant \eta$$

which proves the result. ∎

LEMMA A.2. *If* $u \in \mathscr{H}$, *then*

$$\lim_{x \to 0+} x^{1/2} u(x) = 0.$$

Proof. To begin with suppose $u \in \mathscr{D}$, and let $\phi \in C^\infty[0, 1]$ be a cutoff function which is unity in some neighborhood $[0, \delta]$, and which vanishes at $x = 1$. Writing $U = u\phi$, we clearly have for any $x > 0$,

$$U(x) = -\int_x^1 U'(t)\, dt.$$

So

$$\begin{aligned} |x^{1/2}U(x)| &= x^{1/2}\left|\int_x^1 U'(t)\, dt\right| \\ &\leqslant x^{1/2}\left(\int_x^1 |U'(t)|^2\, t^2\, dt\right)^{1/2}\left(\int_x^1 \frac{dt}{t^2}\right)^{1/2} \\ &\leqslant x^{1/2}\left(\int_0^1 |U'(t)|^2\, t^2\, dt\right)^{1/2}\left(\frac{1}{x} - 1\right)^{1/2} \\ &\leqslant \int_0^1 |U'(t)|^2\, t^2\, dt\ (1-x)^{1/2} \\ &\leqslant C\left(\int_0^1 u^2 + t^2(u')^2\, dt\right)^{1/2}, \end{aligned}$$

for some constant $C > 0$. In particular, if $0 < x < \delta$

$$|x^{1/2}u(x)| \leqslant C\, \|u\|. \tag{A.1}$$

Now if $v \in \mathscr{H}$, there must be some sequence $u_n \in \mathscr{D}$ such that $u_n \to v$ in $\mathscr{H}$. In particular it follows from (A.1) that for any $x \in (0, \delta)$,

$$|x^{1/2}u_n(x) - x^{1/2}u_m(x)| \leqslant C\, \|u_n - u_m\|.$$

Since this convergence is uniform on $(0, \delta)$, it follows that $x^{1/2}u_n(x)$ must converge pointwise to a continuous function on $[0, \delta]$. Thus

$$\lim_{x \to 0+} x^{1/2}v(x) = \lim_{n \to \infty} \lim_{x \to 0+} x^{1/2}u_n(x) = 0 \quad \blacksquare$$

LEMMA A.3. *If* (16) *holds, then the coefficient* a_b *of* $y_\mu^{(b)}$ *in* (2) *will vanish if and only if* y *is identically zero on* $(0, 1)$. *In particular,* $y_\mu^{(s)}(1) \neq 0$.

Proof. (a) Suppose y is identically zero on $(0, 1)$, that is

$$a_b\, y_\mu^{(b)}(x) + a_s\, y_\mu^{(s)}(x) = 0, \qquad 0 < x < 1.$$

So multiplying through by $x^{(1/2)+\mu}$, and then letting $x \to 0+$, shows that $a_b = 0$.

(b) Conversely, suppose that $a_b = 0$. Then $y = a_s y_\mu^{(s)} \in \mathscr{H}$; moreover, because of (1b), $y(1) = 0$, and so $y \in \mathscr{H}_0$. It follows then, by the same sort of considerations used to obtain (21), that

$$W(y, v) = 0, \qquad \forall v \in \mathscr{H}_0.$$

In particular, $W(y, y) = 0$, and so by (16), $y = 0$.

(c) Suppose by way of contradiction that $y_\mu^{(s)}(1) = 0$; then, $y = y_\mu^{(s)}(x)$ would satisfy (1). Thus by (b) above, $y_\mu^{(s)}(x)$ must vanish identically on (0,1). But this is clearly impossible, since $y_\mu^{(s)}(x) \sim x^{-(1/2)+\mu}$ as $x \to 0+$. ∎

APPENDIX B

Proof of Lemma 4.1. First, we show (45a). For $i = 1, 2, .., M-1$,

$$\begin{aligned} \Delta\left(\frac{i+1}{M}\right) - \Delta\left(\frac{i}{M}\right) &= \left(\frac{i+1}{M}\right)^\gamma - \left(\frac{i}{M}\right)^\gamma \\ &= \frac{i^\gamma}{M^\gamma}\left(\left(1+\frac{1}{i}\right)^\gamma - 1\right) \\ &= \frac{i^\gamma}{M^\gamma}\left(\frac{1}{i}\,\gamma(1+\xi)^{\gamma-1}\right) \end{aligned}$$

by the mean value theorem, where $0 < \xi < 1/i \leqslant 1$. Thus

$$\begin{aligned} \Delta\left(\frac{i+1}{M}\right) - \Delta\left(\frac{i}{M}\right) &\leqslant \frac{\gamma 2^{\gamma-1}}{M}\left(\frac{i}{M}\right)^{\gamma-1} \\ &\leqslant \frac{\gamma 2^{\gamma-1}}{M}\left[\left(\frac{i}{M}\right)^\gamma\right]^{1-(1/\gamma)} \end{aligned}$$

which is precisely (45a) with $D = \gamma 2^{\gamma-1}$, $\lambda = 1 - (1/\gamma)$.

For (45b), we need only note that with the above choice of λ,

$$\Delta\left(\frac{1}{M}\right) = M^{-\gamma} = M^{-(1/(1-\lambda))}. \quad ∎$$

Proof of Theorem 4.2. The proofs of parts (i) and (ii) of the theorem are almost identical. We shall only deal with (ii) here.

Define $U \in \mathscr{H}_0$ by requiring

$$\begin{aligned} U(0) &= \xi(x_1), \\ U(x_i) &= \xi(x_i), \qquad i = 1, ..., M-1. \end{aligned}$$

In addition, we have since $U \in \tilde{\mathscr{H}}_0$ and $\xi \in \mathscr{H}_0$,

$$U(x_M) = \xi(x_M) = 0.$$

Consider the interval $J_i = (x_i, x_{i+1})$, $i = 1,..., M-1$. U is linear on J_i, and it interpolates ξ at the endpoints of J_i. It is a standard result (see, e.g., [16]) that for some generic constant $C > 0$, independent of the mesh and of ξ,

$$\int_{J_i} (\xi(x) - U(x))^2 \leqslant C\,|J_i|^4 \int_{J_i} (\xi''(x))^2\,dx, \tag{B.1}$$

where $|J_i| = (x_{i+1} - x_i)$ denotes the length of the interval J_i. Turning now to

$$\begin{aligned}\int_{J_i} x^2(\xi'(x) - U'(x))^2 &\leqslant (x_{i+1})^2 \int_{J_i} (\xi'(x) - U'(x))^2 \\ &\leqslant C(x_{i+1})^2\,|J_i|^2 \int_{J_i} (\xi''(x))^2\end{aligned} \tag{B.2}$$

again using a standard result. Now if the mesh is of class (D, λ), then dividing (45a) through by x_i gives

$$\begin{aligned}\frac{x_{i+1}}{x_i} - 1 &\leqslant \frac{D}{M}\, x_i^{\lambda - 1} \\ &\leqslant \frac{D}{M}\, x_1^{\lambda - 1} \\ &\leqslant \frac{D}{M}\left(\frac{D^{-1}}{M^{1/(1-\lambda)}}\right)^{\lambda - 1} \leqslant D^{2-\lambda}\end{aligned} \tag{B.3}$$

using (45b). Thus for some constant C, depending only on D and λ,

$$\frac{x_{i+1}}{x_i} \leqslant C.$$

Using this along with (45a) we obtain the estimates

$$|J_i|^4 \leqslant x_{i+1}^2\,|J_i|^2 \leqslant C x_i^2\,\frac{x_i^{2\lambda}}{M^2}.$$

Combining this with (B.1) and (B.2) we find

$$\begin{aligned}\int_{J_i} (\xi - U)^2 + x^2(\xi' - U')^2 &\leqslant \frac{C}{M^2}\, x_i^{2+2\lambda} \int_{J_i} (\xi'')^2 \\ &\leqslant \frac{C}{M^2} \int_{J_i} x^{2+2\lambda} (\xi'')^2.\end{aligned} \tag{B.4}$$

Now, consider the first interval $J_0 = (0, x_1)$. Recall that $U(x) = \xi(x_1)$ on J_0, in particular $U' = 0$. Therefore

$$\int_{J_0} x^2(\xi' - U')^2 = \int_{J_0} x^2(\xi')^2 \leqslant x_1^{2-2\lambda} \int_{J_0} x^{2\lambda}(\xi')^2$$

$$\leqslant \frac{C}{M^2} \int x^{2\lambda}(\xi')^2 \tag{B.5}$$

using (45b). Now clearly, for $x \in J_0$,

$$|\xi(x) - U(x)| = |\xi(x) - \xi(x_1)|$$

$$= \left| \int_x^{x_1} \xi'(t)\, dt \right|$$

$$\leqslant \left(\int_x^{x_1} t^{2\lambda}(\xi'(t))^2 \right)^{1/2} \left(\int_x^{x_1} \frac{dt}{t^{2\lambda}} \right)^{1/2}$$

using Schwartz's inequality. Thus

$$\int_{J_0} |\xi(x) - U(x)|^2 \leqslant \left(\int_{J_0} x^{2\lambda}(\xi')^2 \right) \left(\int_0^{x_1} \int_x^{x_1} t^{-2\lambda}\, dt\, dx \right)$$

$$\leqslant \left(\int_{J_0} x^{2\lambda}(\xi')^2 \right) \left(\int_0^{x_1} t^{1-2\lambda}\, dt \right)$$

$$\leqslant C \left(\int_{J_0} x^{2\lambda}(\xi')^2 \right) x_1^{2-2\lambda}$$

$$\leqslant \frac{C}{M^2} \int_{J_0} x^{2\lambda}(\xi')^2 \tag{B.6}$$

again using (45b).

Thus, summing (B.4) over $i = 1, \ldots, M-1$ and then adding (B.5) and (B.6) we obtain

$$\|\xi - U\| \leqslant \frac{C}{M} |||\xi|||_\lambda$$

which suffices to prove part (ii) of the theorem. ∎

APPENDIX C

Proof of Lemma 4.4. Condition (50a) is clear, while (50b) follows by the same sort of argument used in Lemma 4.1: Write $\delta = (1 - (L/M)^{1/\gamma})/(M-1)$, then

$$\begin{aligned} x_{i+1} - x_i &= (x_i^{1/\gamma} + \delta)^\gamma - x_i \\ &= x_i((1 + \delta x_i^{-1/\gamma})^\gamma - 1) \\ &= x_i\, \delta x_i^{-1/\gamma} \gamma (1+\xi)^{\gamma-1} \end{aligned}$$

by the mean value theorem, where

$$\begin{aligned} 0 < \xi < \delta x_i^{-1/\gamma} &\leqslant \delta x_1^{-1/\gamma} \\ &\leqslant \frac{1}{M-1} \left(\frac{L}{M}\right)^{-1/\gamma}. \end{aligned}$$

This is bounded independently of M as $M \to \infty$ since $\gamma \geqslant 1$. Thus

$$\begin{aligned} x_{i+1} - x_i &\leqslant C x_i^{1-(1/\gamma)} \delta \\ &\leqslant \frac{C}{M} x_i^{1-(1/\gamma)}. \quad \blacksquare \end{aligned}$$

Proof of Theorem 4.5. The proofs of the two estimates of (52) are almost identical. We shall only deal with the second here. By (37) we know that $\|\!| \xi |\!\|_\lambda < \infty$.

Define $U \in \tilde{\mathscr{H}}_0$ by requiring

$$U(x_i) = \xi(x_i), \qquad i = 1, \ldots, M-1.$$

In addition we have since $U \in \tilde{\mathscr{H}}_0$ and $\xi \in \mathscr{H}_0$,

$$U(x_M) = \xi(x_M) = 0.$$

For the intervals $J_i = (x_i, x_{i+1})$, $i = 1, \ldots, M-1$, we obtain the same estimates (B.1) and (B.2) as in the proof of Theorem 4.2. Instead of (B.3), we now have by (50b),

$$\frac{x_{i+1}}{x_i} - 1 \leqslant \frac{D}{M} x_1^{\lambda - 1} \leqslant D^{2-\lambda} M^{-\lambda}.$$

Since $\lambda \geqslant 0$ we still have $x_{i+1}/x_i \leqslant C$, with $C > 0$ depending only on D and λ. Thus the estimate (B.4) remains valid.

Turn now to the interval $J_0 = (0, x_1)$. By (36) and (19) we may write

$$\xi(x) = a_s x^\alpha + \xi^*(x),$$

where $\|\!| \xi^* |\!\|_0 < \infty$ and

$$\xi^*(x) = O(x^{\alpha+1}) \qquad \text{as} \quad x \to 0+, \tag{C.1}$$

and for convenience we have set $\alpha = \alpha_\mu^{(s)}$. So on J_0 we have

$$\begin{aligned}\xi(x) - U(x) &= \xi(x) - \xi(x_1)\left(\frac{x}{x_1}\right)^\alpha \\ &= (a_s x^\alpha + \xi^*(x)) - (a_s x_1^\alpha + \xi^*(x_1))\left(\frac{x}{x_1}\right)^\alpha \\ &= \xi^*(x) - \xi^*(x_1)\left(\frac{x}{x_1}\right)^\alpha \\ &= [\xi^*(x) - \xi^*(x_1)] + \left[\xi^*(x_1)\left(1 - \left(\frac{x}{x_1}\right)^\alpha\right)\right] \\ &= e_1 + e_2 \qquad \text{say.} \end{aligned} \tag{C.2}$$

Just as in the proof of Theorem 4.2 (in the case $\lambda = 0$) we obtain

$$\int_{J_0} e_1^2 + x^2(e_1')^2 \leqslant \frac{C}{M^2}\int_{J_0} (\xi^{*\prime})^2. \tag{C.3}$$

For the remaining term of (C.2) note that

$$\int_0^{x_1} \left(1 - \left(\frac{x}{x_1}\right)^\alpha\right)^2 + x^2\left(\frac{d}{dx}\left(\frac{x}{x_1}\right)^\alpha\right)^2 \leqslant C x_1$$

with C depending only on α. Introduce the notation

$$|\xi^*|_{\alpha+1} = \sup_{x\in(0,1)} \left|\frac{\xi^*(x)}{x^{\alpha+1}}\right| < \infty$$

by (C.1). We therefore have

$$\begin{aligned}\int_{J_0} e_2^2 + x^2(e_2')^2 &\leqslant C(\xi^*(x_1))^2 x_1 \\ &\leqslant C x_1^{3+2\alpha} |\xi^*|_{\alpha+1}^2 \\ &\leqslant C M^{-(3+2\alpha)} |\xi^*|_{\alpha+1}^2 \\ &\leqslant C M^{-2} |\xi^*|_{\alpha+1}^2 \end{aligned} \tag{C.4}$$

by (50a) and the fact that $\alpha = \alpha_\mu^{(s)} > -\frac{1}{2}$.

Now bringing together (C.2), (C.3), (C.4), and (B.4) we have

$$\|\xi - U\| \leqslant \frac{C}{M}(|||\xi|||_\lambda + |||\xi^*|||_0 + |\xi^*|_{\alpha_\mu^{(s)}+1})$$

which suffices to prove the second estimate of (52). ∎

ACKNOWLEDGMENT

One of us (R.L.D.) wishes to acknowledge his profound debt to the late Dr. Raymond Grimm for his inspiration of this work. His enthusiastic readiness to implement and experiment with new numerical approaches to this (and other) problems provided the impetus to seek a better and more fundamental understanding of the finite element approach to the asymptotic matching method.

REFERENCES

1. H. P. FURTH, J. KILLEEN, AND M. N. ROSENBLUTH, *Phys. Fluids* **6**, 459 (1963).
2. B. COPPI, J. M. GREENE, AND J. L. JOHNSON, *Nucl. Fusion* **6**, 101 (1966).
3. A. H. GLASSER, J. M. GREENE, AND J. L. JOHNSON, *Phys. Fluids* **18**, 875 (1975).
4. M. KOTSHENREUTHER, A. Y. AYDEMIR, D. C. BARNES, J. R. CARY, J. D. HANSON, R. D. HAZELTINE, AND P. J. MORRISON, in *Plasma Physics and Controlled Nuclear Fusion Research 1984*, (IAEA, Vienna, 1985), Vol. 2, p. 223.
5. A. H. GLASSER, S. C. JARDIN, AND G. TESAURO, *Phys. Fluids* **27**, 1225 (1984).
6. R. C. GRIMM, R. L. DEWAR, AND J. MANICKAM, *J. Comput. Phys.* **49**, 94 (1983).
7. R. GRUBER, F. TROYON, D. BERGER, L. C. BERNARD, S. ROUSSET, R. SCHREIBER, W. KERNER, W. SCHNEIDER, AND K. V. ROBERTS, *Comput. Phys. Commun.* **21**, 323 (1981).
8. R. C. GRIMM, R. L. DEWAR, J. MANICKAM, S. C. JARDIN, A. H. GLASSER, AND M. S. CHANCE, in *Plasma Physics and Controlled Nuclear Fusion Research* 1982, (IAEA, Vienna, 1983), Vol. III, p. 35.
9. J. MANICKAM, R. C. GRIMM, AND R. L. DEWAR, in *Energy Modeling and Simulation*, edited by A. S. Kydes *et al.* (North-Holland, Amsterdam, 1983), p. 355.
10. R. L. DEWAR AND R. C. GRIMM, in *Computational Techniques and Applications: CTAC-83*, edited by J. Noye and C. Fletcher (North-Holland, Amsterdam, 1984), p. 730.
11. I. BABUŠKA AND A. D. MILLER, *Int. J. Numer. Methods Eng.* **20**, 1085 (1984).
12. S. CHANDRASEKHAR, *Hydrodynamic and Hydromagnetic Stability* (Oxford Univ. Press, London, 1961), p. 359.
13. W. A. NEWCOMB, *Ann. Phys.* **10**, 232 (1960).
14. I. BABUŠKA AND A. K. AZIZ, in *The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations*, edited by A. K. Aziz (Academic Press, New York/London, 1972).
15. M. ABRAMOWITZ AND I. STEGUN, *Handbook of Mathematical Functions*, (Dover, New York, 1972).
16. P. G. CIARLET, *The Finite Element Method for Elliptic Problems* (North-Holland, Amsterdam, 1980).
17. R. L. DEWAR, "Annual Controlled Fusion Theory Conference," Sherwood Theory Meeting, Arlington, Virginia, 1983, Paper 1R24.
18. R. L. DEWAR, D. A. MONTICELLO, AND W. N-C. SY, *Phys. Fluids* **27**, 1723 (1984).