MATHEMATICS OF COMPUTATION
VOLUME 64, NUMBER 210
APRIL 1995, PAGES 629–648

DEFECT CORRECTION FOR TWO-POINT BOUNDARY VALUE PROBLEMS ON NONEQUIDISTANT MESHES

J. C. BUTCHER, J. R. CASH, G. MOORE, AND R. D. RUSSELL

ABSTRACT. New finite difference formulae of arbitrary order are derived for the special class of second-order two-point boundary value problems $y'' = f(x, y(x))$, $a \leq x \leq b$. Variable mesh spacing is possible, and the required accuracy is achieved under a very mild mesh condition. A natural defect correction framework is set up to compute the higher-order approximations.

1. INTRODUCTION

In this paper we develop high-order finite difference formulae for solving the second-order two-point boundary value problem (TPBVP)

$$(1.1) \qquad y''(x) = f(x, y(x)), \quad g(y(a), y(b)) = 0, \quad y : [a, b] \to \mathbb{R}^N.$$

For the corresponding initial value problem

$$(1.2) \qquad y''(x) = f(x, y(x)), \quad y(0) = y_0, \quad y'(0) = y_0', \qquad 0 \leq x \leq x_F,$$

the most widely used numerical technique is to convert it to an 'equivalent' first-order system

$$(1.3) \qquad y'(x) = f(x, y(x)), \qquad y(0) = y_0$$

and then to use a standard software package. The conversion of (1.2) to (1.3) is appropriate for initial value problems first because storage is not normally a crucial factor, and so an increase in the dimension of the system is not a serious consideration, and second because the first derivative of y is needed from the beginning in order to advance the solution from step to step. However, for the TPBVP (1.1) the situation is quite different. If the solution is computed numerically, the storage space can become a major consideration. Treatment of the high-order equation directly, rather than conversion to the corresponding first-order system, can be considerably more efficient, e.g., as evidenced by the success of the collocation code COLSYS [2] for solving TPBVPs. Thus, it is natural to derive numerical methods tailored to this special form.

A major motivation for developing methods for second-order TPBVPs is that such problems arise when applying a method of lines procedure for solving hyperbolic and parabolic partial differential equations in one space variable, or

Received by the editor September 22, 1992 and, in revised form, May 14, 1993 and April 26, 1994.

1991 *Mathematics Subject Classification*. Primary 65L10.

elliptic equations in two space variables. In such contexts the development of efficient numerical methods on variable meshes is very important. A number of steps in this direction have been made, e.g., see Skeel and Berzins [27], where a finite element method is given, which is adapted for TPBVPs with an additional singularity arising by applying cylindrical or spherical symmetry to the PDE. While an eventual goal would be to extend our approach to handle general second-order TPBVPs for which $y'(x)$ appears explicitly, the form (1.1) does arise naturally, e.g., from discretizing first-order evolution equations $\partial u/\partial t + Au = F(u)$ in time t, where $Au \equiv \partial^2 u/\partial x^2$ and $F(u)$ is a nonlinearity in u (and not u_x), as for certain reaction-diffusion equations (see Jolly [18]) or for the complex Ginsburg-Landau equation [15].

Many approaches to solving (1.1) have appeared, viz. [7, 12, 13, 14, 17, 23], which are in some way connected to the approach here. However, many of these schemes are based on a uniform mesh, while an important property of the schemes we describe is that a variable mesh is allowed. Perhaps the simplest scheme for the solution of (1.1) would be to replace the derivative term in (1.1) by an appropriate finite difference approximation. Manteuffel and White [23] have recently analyzed this approach and have shown that centered difference schemes give second-order convergence even on nonuniform meshes. This important result has a bearing on the schemes we will derive in that we develop efficient high-order methods which appear to retain the high order of convergence for variable mesh spacing. Some high-order methods for the solution of (1.1) have been developed by Daniel and Martin [14]. They used a finite difference approach based on Numerov's method and increased the order of the basic method using iterated deferred corrections. Their approach can be regarded as an extension of Pereyra's method for first-order two-point boundary value problems [21, 22, 25] to the special second-order system (1.1). One of the present authors has recently proposed a different deferred correction approach to the solution of first-order systems of TPBVP [5, 6], based on mono-implicit Runge-Kutta formulae [9, 10]. Theoretical and numerical results indicate that this new approach is competitive with existing methods.

In what follows, we extend this basic approach to the special TPBVP (1.1). We again apply a deferred correction approach, using a natural generalization of mono-implicit Runge-Kutta formulae to second-order equations [8]. A change in methodology is introduced, however, since we adopt a Galerkin viewpoint as suggested in [3, 24]. Hence, we give a novel derivation of high-order finite difference methods, for nonuniform meshes, which is extremely natural and easy to understand. These high-order methods are developed using defect correction, which provides a very efficient way to solve the equations. The orders of convergence for the methods are shown using supraconvergence arguments. This term was introduced in [20, 23] and means that, under a mild mesh condition (cf. (2.8)), the truncation error is higher-order on average than pointwise. Hence, there is a strong connection with the idea of superconvergence in Galerkin/finite element methods.

In §2 we explain our underlying conditions on problem (1.1), and for notational simplicity we look at a single equation. Also, as discussed herein, homogeneous Dirichlet boundary conditions are assumed. Our basic framework for developing methods of arbitrary order is then presented and second/fourth-order approximations derived, the latter requiring a fundamental mesh condition necessary to exploit the supraconvergence phenomenon. In §3 the fourth-order accuracy is obtained by defect correction. Higher-order methods require extra function approximations to be obtained, and §4 explains how these can be generated, the supraconvergence again being crucial. In §5 the higher-order accuracy is achieved by defect correction. Finally, possible variations and generalizations within our defect correction framework are mentioned in §6, and §7 contains numerical results illustrating the orders of accuracy achieved.

2. The differential equation and its approximation

We consider the problem

$$y''(x) = f(x, y(x)), \qquad a < x < b, \tag{2.1}$$

with boundary conditions $y(a) = y(b) = 0$ and assume there exists a solution $y^* \in C^2[a, b]$. Later in this paper, y^* will be assumed to have more smoothness as required. As for the function $f(x, y)$, we assume once and for all that there exist constants α and $\delta > 0$ such that $\forall x \in [a, b]$

$$|g(x, u) - g(x, v)| \le \alpha |u - v| \tag{2.2}$$

if $|u - y^*(x)| < \delta$ and $|v - y^*(x)| < \delta$, where g is any of f, $\partial f/\partial x$, $\partial f/\partial y$, $\partial^2 f/\partial x^2$, $\partial^2 f/\partial x \partial y$, and $\partial^2 f/\partial y^2$. Finally, it is also assumed that the solution y^* is *isolated*, i.e., the linear problem

$$-y''(x) + q^*(x)y(x) = w(x), \qquad y(a) = y(b) = 0, \tag{2.3}$$

where $q^*(x) \equiv \partial f(x, y^*(x))/\partial y$, has a unique solution $z \in C^2[a, b]$ for each $w \in C[a, b]$, and there exists $\kappa > 0$ independent of w such that

$$\max_{x \in [a, b]} \{|z'(x)|\} \le \kappa \max_{x \in [a, b]} \left\{ \left| \int_a^x w(t)\, dt \right| \right\}.$$

We shall see later why it is appropriate to consider stability in this fashion.

The solution y^* will be approximated by a mesh function in the following way. Consider a mesh on $I \equiv [a, b]$, i.e.,

$$a = x_0 < x_1 < \cdots < x_{N-1} < x_N = b,$$

and introduce the notation

$$I_{j-\frac{1}{2}} = [x_{j-1}, x_j], \quad I_j = I_{j-\frac{1}{2}} \cup I_{j+\frac{1}{2}}, \quad h_{j-\frac{1}{2}} \equiv x_j - x_{j-1},$$

$$h_j \equiv \frac{h_{j-\frac{1}{2}} + h_{j+\frac{1}{2}}}{2}, \quad \theta_j \equiv \frac{h_{j-\frac{1}{2}} - h_{j+\frac{1}{2}}}{h_{j-\frac{1}{2}} + h_{j+\frac{1}{2}}} \quad \text{and} \quad h \equiv \max_j \{h_{j-\frac{1}{2}}\}.$$

Define the basis functions $\{\varphi_j\}$ according to

$$\varphi_j(x) \equiv \begin{cases} (x - x_{j-1})/h_{j-\frac{1}{2}}, & x \in I_{j-\frac{1}{2}}, \\ (x_{j+1} - x)/h_{j+\frac{1}{2}}, & x \in I_{j+\frac{1}{2}}, \\ 0 & x \notin I_j, \end{cases} \qquad j = 1, \dots, N-1,$$

so that $\int \varphi_j = h_j$. Now inserting y^* into (2.1), multiplying both sides by φ_j / h_j, and integrating leads to

$$(D^2 Y^*)_j + \frac{1}{h_j} \int f(x, y^*(x)) \varphi_j\, dx = 0, \qquad j = 1, \dots, N-1, \tag{2.4}$$

where Y^* is the mesh function with values $y^*(x_j)$ and

$$(D^2Y)_j \equiv \frac{(Y_j - Y_{j-1})/h_{j-\frac{1}{2}} - (Y_{j+1} - Y_j)/h_{j+\frac{1}{2}}}{h_j}.$$

The whole of this paper is concerned with obtaining difference equations of increasing accuracy for (2.1) by using weighted quadrature rules of increasing order to approximate the integral in (2.4). Of course, these quadrature rules are only allowed to use information based on the mesh values Y^*_{j-1}, Y^*_j and Y^*_{j+1}.

The simplest quadrature rule is perhaps the "generalized weighted midpoint" rule, i.e.,

$$\int_{-1}^{1} z(t)\varphi_\theta \, dt \approx z(\theta),$$

where $\theta \in (-1, 1)$ and

$$\varphi_\theta(t) \equiv \begin{cases} (t+1)/(\theta+1), & t \in [-1, \theta], \\ (1-t)/(1-\theta), & t \in [\theta, 1], \end{cases}$$

derived by integrating the constant polynomial which interpolates z at $t = \theta$ exactly against φ_θ. The error in this integration method is given by

$$\left| \int_{-1}^{1} z(t)\varphi_\theta \, dt - z(\theta) + \frac{2\theta}{3} z'(\theta) \right| \leq C_0 \max_{t \in [-1, 1]} \{|z''(t)|\}.$$

Hence, applying this quadrature rule, by a change-of-variable, to the integrals in (2.4) leads to the difference equation

$$(2.5) \qquad (D^2Y)_j + f(x_j, Y_j) = 0, \qquad j = 1, \ldots, N-1,$$

with $Y_0 = Y_N = 0$, and the local truncation error

$$\tau^0_j \equiv \frac{1}{h_j} \int f(x, y^*(x))\varphi_j \, dx - f(x_j, Y^*_j)$$

satisfies

$$\left| \tau^0_j + \frac{2}{3} h_j \theta_j y^{*'''}(x_j) \right| \leq C_0 h_j^2 \|y^{*(\mathrm{iv})}\|_{I_j},$$

where $\|\cdot\|_A$ denotes the maximum norm over an interval A. Hence, although $|\tau^0_j| = O(h_j)$ in general,

$$\left| \frac{2}{3} \sum_{i=1}^{j} h_i^2 \theta_i y^{*'''}(x_i) \right| \leq \frac{1}{6} \Big\{ 2 \max_j \{h^2_{j-\frac{1}{2}} \|y^{*'''}\|_{I_{j-\frac{1}{2}}}\} + (b-a) \max_j \{h^2_{j-\frac{1}{2}} \|y^{*(\mathrm{iv})}\|_{I_{j-\frac{1}{2}}}\} \Big\}$$

and thus $\|\tau^0\|_{-1}$ (where this norm is defined in (2.6)) is second order in terms of the size of the local mesh and local derivatives. It can then be shown that, providing the local mesh size mirrors sufficiently closely the behavior of $y^*(x)$, there is a unique solution of (2.5) near Y^*. We state the following results without proof:

(i) if

$$\kappa \max_{j=1, \ldots, N} \{(b-a) h_{j-\frac{1}{2}} [\|q\|_{I_{j-\frac{1}{2}}} + \|q'\|_{I_{j-\frac{1}{2}}}]\} \leq \frac{1}{2},$$

then the linear difference equation

$$(D^2Y)_j + q^*(x_j)Y_j = \beta_j, \qquad j = 1, \dots, N-1,$$

with $Y_0 = Y_N = 0$ has a unique solution Z for each β and

$$\|Z\|_1 \leq 2\kappa\|\beta\|_{-1},$$

where

$$\|Z\|_1 \equiv \max_{j=1,\dots,N}\{|(Z_j - Z_{j-1})/h_{j-\frac{1}{2}}|\},$$

$$\tag{2.6} \|\beta\|_{-1} \equiv \max_{j=1,\dots,N-1}\left\{\left|\sum_{i=1}^{j} h_i\beta_i\right|\right\};$$

(ii) if

$$(b-a)^2 2\kappa\alpha d_0 \leq \frac{1}{2},$$

where $d_0 \equiv 4\kappa\|\tau_0\|_{-1}$, a locally unique solution Y^0 of (2.5) exists in $\overline{B}(Y^*, d_0)$ [the closed ball of radius d_0 in $\|\cdot\|_1$ centered on Y^*], and the linear difference equation

$$(L_h^0 Y)_j \equiv (D^2Y)_j + \frac{\partial f}{\partial y}(x_j, Y_j^0)Y_j = \beta_j, \qquad j = 1, \dots, N-1,$$

with $Y_0 = Y_N = 0$ has a unique solution Z for each β with

$$\|Z\|_1 \leq \kappa_0\|\beta\|_{-1},$$

where $\kappa_0 \equiv 4\kappa$.

A more accurate difference scheme is obtained by using the "generalized weighted Simpson" rule to approximate the integral in (2.4), i.e.,

$$\int_{-1}^{1} z(t)\varphi_\theta\, dt \approx \alpha(\theta)z(1) + \beta(\theta)z(\theta) + \alpha(-\theta)z(-1),$$

where

$$\text{(i)}\quad \alpha(\theta) \equiv \frac{1 - 4\theta - \theta^2}{12(1-\theta)} \qquad \text{(ii)}\quad \beta(\theta) \equiv \frac{5 - \theta^2}{6(1-\theta^2)},$$

derived by integrating the quadratic polynomial which interpolates z at $t = \pm 1, \theta$ exactly against φ_θ. The error in this integration method is given by

$$\left|\int_{-1}^{1} z(t)\varphi_\theta\, dt - \{\alpha(\theta)z(1) + \beta(\theta)z(\theta) + \alpha(-\theta)z(-1)\} - \frac{(21\theta + \theta^3)}{90}z'''(\theta)\right|$$
$$\leq C_1 \max_{t\in[-1,1]}\{|z^{(\text{iv})}(t)|\}.$$

Hence, applying this quadrature rule, by a change-of-variable, to the integrals in (2.4) leads to the difference equation

$$\tag{2.7} (D^2Y)_j + \alpha(\theta_j)f(x_{j+1}, Y_{j+1}) + \beta(\theta_j)f(x_j, Y_j) + \alpha(-\theta_j)f(x_{j-1}, Y_{j-1}) = 0$$

with $Y_0 = Y_N = 0$. We shall use this equation to obtain a fourth-order accurate mesh function in the next section. Here we merely point out that the truncation error τ^1, defined by

$$\tau_j^1 \equiv \frac{1}{h_j}\int f(x, y^*(x))\varphi_j\, dx$$
$$- \{\alpha(\theta_j)f(x_{j+1}, Y_{j+1}^*) + \beta(\theta_j)f(x_j, Y_j^*) + \alpha(-\theta_j)f(x_{j-1}, Y_{j-1}^*)\},$$

satisfies

$$\left|\tau_j^1 - h_j^3 \frac{(21\theta_j + \theta_j^3)}{90} y^{*(\mathrm{v})}(x_j)\right| \le C_1 h_j^4 \|y^{*(\mathrm{vi})}\|_{I_j}$$

and thus, in general, we only have $|\tau_j^1| = O(h_j^3)$. Moreover, it is not true that $\|\tau^1\|_{-1} = O(h^4)$ for arbitrary meshes [20]. If, however, our meshes satisfy a supraconvergence condition [20, 23], it is possible to show that $\|\tau^1\|_{-1}$ is fourth-order accurate. For our purposes we define this condition as follows.

Supraconvergent mesh condition. There are real constants, $\overline{U}$, $\overline{C}$ and an integer constant $\overline{N}$ such that all meshes satisfy

$$\text{(i)} \quad |\theta_j| \le \overline{U} < 1$$

and at least *one* of

$$\text{(QUANTITATIVE)} \quad \text{(ii)} \quad \sum_{j=1}^{N-1} |\theta_j| \le \overline{C}$$

(QUALITATIVE) (iii) θ_j changes sign no more than $\overline{N}$ times.

Hence, if (ii) holds, we can deduce that

$$\left|\sum_{i=1}^{j} h_i^4 \frac{(21\theta_i + \theta_i^3)}{90} y^{*(\mathrm{v})}(x_i)\right| \le \frac{11}{45}\overline{C} \max_j \{h_{j-\frac{1}{2}}^4 \|y^{*(\mathrm{v})}\|_{I_{j-\frac{1}{2}}}\},$$

while, on the other hand, if (iii) holds, then

$$\left|\sum_{i=1}^{j} h_i^4 \frac{(21\theta_i + \theta_i^3)}{90} y^{*(\mathrm{v})}(x_i)\right| \\ \le \frac{11(\overline{N}+1)}{360}\left[2 \max_j \{h_{j-\frac{1}{2}}^4 \|y^{*(\mathrm{v})}\|_{I_{j-\frac{1}{2}}}\} + (b-a) \max_j \{h_{j-\frac{1}{2}}^4 \|y^{*(\mathrm{vi})}\|_{I_{j-\frac{1}{2}}}\}\right].$$

In either case, $\|\tau^1\|_{-1}$ is fourth-order accurate in terms of the local mesh size and local derivatives. The significance of the quantitative/qualitative assumptions is that either implies

$$\sum_{j=1}^{N-1} h_j |\theta_j| = O(h), \tag{2.8}$$

and this ensures the supraconvergence.

Note that this mesh condition will be satisfied by all *practical* meshes. (Part (i) will be required later for higher-order methods and to justify the defect correction procedure.) It can be seen that this gain in accuracy, i.e., smaller $\|\cdot\|_{-1}$ than $\|\cdot\|_0$, where the latter is just the simple maximum of the moduli of the components, will occur whenever the leading error term is a polynomial in θ which is zero at $\theta = 0$. More generally, it will also occur when the leading error term is a function $g(\theta)$ which satisfies $|g(\theta)| \le C|\theta|$. This result will be used repeatedly later for our higher-order approximations.

3. Fourth-order accuracy by defect correction

We will compute a fourth-order approximation Y^1 to Y^* which satisfies

$$(3.1)\quad \begin{aligned}&(D^2Y^1)_j + \alpha(\theta_j)f(x_{j+1}, Y^1_{j+1}) + \beta(\theta_j)(x_j, Y^1_j)\\ &\qquad + \alpha(-\theta_j)f(x_{j-1}, Y^1_{j-1}) = 0, \qquad j = 1, \dots, N-1,\end{aligned}$$

with $Y^1_0 = Y^1_N = 0$. Its existence, local uniqueness and construction is based on the fixed point equation

$$(3.2)\qquad Z = \mathscr{G}_1(Z),$$

where

$$[\mathscr{G}_1(Z)]_j \equiv Z_j - W_j$$

with W satisfying

$$\begin{aligned}(L^0_hW)_j = (D^2Z)_j &+ \alpha(\theta_j)f(x_{j+1}, Z_{j+1}) + \beta(\theta_j)f(x_j, Z_j)\\ &+ \alpha(-\theta_j)f(x_{j-1}, Z_{j-1}), \qquad j = 1, \dots, N-1,\end{aligned}$$

and $W_0 = W_N = 0$. We will show next that $\mathscr{G}_1$, for sufficiently small h, is a contraction mapping in $\overline{B}(Y^0, d_1)$, the closed ball of radius

$$d_1 \equiv 2(\|\mathscr{G}_1(Y^*) - Y^*\|_1 + \|Y^0 - Y^*\|_1)$$

in $\|\cdot\|_1$ centered on Y^0. [Note that $(L^0_h[\mathscr{G}_1(Y^*) - Y^*])_j = \tau^1_j$, and so d_1 just depends on $\|\tau^1\|_{-1}$ and $\|\tau^0\|_{-1}$.] Hence, the defect correction iteration

$$(3.3)\qquad Z^{(m+1)} = \mathscr{G}_1(Z^{(m)}),$$

starting from $Z^{(0)} = Y^0$, will converge to a locally unique Y^1.

If $U, V \in \overline{B}(Y^0, d_1)$, then

$$(3.4)\qquad (L^0_h[\mathscr{G}_1(U) - \mathscr{G}_1(V)])_j = -\delta^0_j - \lambda^0_j,$$

where

(i)

$$\begin{aligned}\delta^0_j \equiv\ &\alpha(\theta_j)[f(x_{j+1}, U_{j+1}) - f(x_{j+1}, V_{j+1})]\\ &+ \beta(\theta_j)[f(x_j, U_j) - f(x_j, V_j)]\\ &+ \alpha(-\theta_j)[f(x_{j-1}, U_{j-1}) - f(x_{j-1}, V_{j-1})]\\ &- [f(x_j, U_j) - f(x_j, V_j)],\end{aligned}$$

and

(ii)

$$\lambda^0_j \equiv f(x_j, U_j) - f(x_j, V_j) - \frac{\partial f}{\partial y}(x_j, Y^0_j)[U_j - V_j].$$

Here, δ^0 is the key term in our analysis of defect correction. It compares the difference in the higher-order formula at U and V with the difference in the basic formula at these two mesh functions. This depends on the smoothness of the difference $U - V$. In the present case we have

$$\begin{aligned}\delta^0_j = \frac{h_j}{12(1-\theta_j^2)}\{&(1 - 5\theta_j^2)[(DE)_{j+\frac12} - (DE)_{j-\frac12}]\\ &- (3\theta_j + \theta_j^3)[(DE)_{j+\frac12} + (DE)_{j-\frac12}]\},\end{aligned}$$

where $E_j \equiv f(x_j, U_j) - f(x_j, V_j)$ and $(DY)_{j-\frac{1}{2}} \equiv (Y_j - Y_{j-1})/h_{j-\frac{1}{2}}$. Writing this as

$$\begin{aligned}\delta_j^0 = \frac{h_j}{12}[(DE)_{j+\frac{1}{2}} - (DE)_{j-\frac{1}{2}}] \\ - \frac{h_j}{12(1-\theta_j^2)}\{4\theta_j^2[(DE)_{j+\frac{1}{2}} - (DE)_{j-\frac{1}{2}}] \\ + (3\theta_j + \theta_j^3)[(DE)_{j+\frac{1}{2}} + (DE)_{j-\frac{1}{2}}]\},\end{aligned}$$

we see that the θ_j factors will ensure supraconvergence for the latter term. Summing the former, however, gives

$$\begin{aligned}&\left|\sum_{i=1}^{j} \frac{h_i^2}{12}[(DE)_{i+\frac{1}{2}} - (DE)_{i-\frac{1}{2}}]\right| \\ &= \left|-\frac{h_i^2}{12}(DE)_{\frac{1}{2}} + \sum_{i=2}^{j} \frac{h_{i-1}^2 - h_i^2}{12}(DE)_{i-\frac{1}{2}} + \frac{h_j^2}{12}(DE)_{j+\frac{1}{2}}\right| \\ &\le \left(\frac{h^2}{6} + \frac{h}{3}\sum_{i=1}^{N-1} h_i|\theta_i|\right)\|E\|_1 .\end{aligned}$$

Consequently, on a supraconvergent mesh we have

$$\|\delta^0\|_{-1} \le \text{const}(\overline{U}, \overline{C}, \overline{N})h^2\|E\|_1,$$

and thus

$$\|\delta^0\|_{-1} \le \text{const}(\overline{U}, \overline{C}, \overline{N})h^2\|U - V\|_1 .$$

We emphasize that this constant only depends on the supraconvergence properties of the mesh. The other term in (3.4), λ^0, is just a simple linearization error and satisfies

$$\|\lambda^0\|_0 \le \alpha(b-a)d_1\|U - V\|_0 .$$

Consequently, $\mathscr{G}_1$ is a contraction on $\overline{B}(Y^0, d_1)$ for sufficiently small h and d_1, and we may assume that the contraction constant $\gamma_1 \le 1/4$.

In addition, $\mathscr{G}_1$ maps $\overline{B}(Y^0, d_1)$ onto itself since $Z \in \overline{B}(Y^0, d_1)$ implies

$$\|\mathscr{G}_1(Z) - Y^0\|_1 \le \|\mathscr{G}_1(Z) - \mathscr{G}_1(Y^*)\|_1 + \|\mathscr{G}_1(Y^*) - Y^*\|_1 + \|Y^* - Y^0\|_1 .$$

Hence, our conclusion is that the defect correction iteration (3.3) converges to the locally unique Y^1 satisfying (3.1) and that, as usual,

$$Y^* - Y^1 = Y^* - \mathscr{G}_1(Y^*) + \mathscr{G}_1(Y^*) - Y^1$$

implies

$$\|Y^1 - Y^*\|_1 \le \frac{\kappa_0}{1-\gamma_1}\|\tau^1\|_{-1} . \tag{3.5}$$

We underline the important fact that the global mesh size h only appears in the contraction constant γ_1, while the error $Y^1 - Y^*$ just depends on the product of local mesh size and local solution derivative size.

4. Function values for higher-order quadrature

We wish to approximate the integral in (2.4) more accurately than is possible with the generalized weighted Simpson rule, but still only using the mesh values

Y_{j-1}^*, Y_j^* and Y_{j+1}^*. To do this, we must make use of the differential equation (2.1), which y^* satisfies, in order to generate accurate approximations to y^* at other points besides x_{j-1}, x_j and x_{j+1}. This is exactly what Runge-Kutta methods do for first-order systems. The values of $y^{*\prime\prime}$, at x_{j-1}, x_j and x_{j+1}, i.e., $f(x_{j-1}, Y_{j-1}^*)$, $f(x_j, Y_j^*)$ and $f(x_{j+1}, Y_{j+1}^*)$, are immediately available and naively one would suppose that it would be possible to fit a quintic polynomial to these six pieces of data and thus obtain an $O(h_j^6)$ approximation to y^* on I_j. It is, however, easily checked that for $\theta_j = 0$, i.e., $h_{j+\frac{1}{2}} = h_{j-\frac{1}{2}}$, the data is incompatible in general. Hence, we shall define $p_2(Y_j^*; x)$ to be the quartic polynomial which satisfies

$$(4.1)\qquad \begin{array}{lll} \text{(i)} & p_2(Y_j^*; x_i) = Y_i^*, & i = j-1, j+1, \\ \text{(ii)} & p_2''(Y_j^*; x_i) = f(x_i, Y_i^*), & i = j-1, j, j+1, \end{array}$$

i.e., we are no longer trying to collocate with y^* at x_j. In general, of course, $p_2(Y_j^*; x)$ can only be an $O(h_j^5)$ approximation to y^* on I_j, but we shall make use of the supraconvergence property to obtain $O(h^6)$ in a negative norm, which has proved to be an important tool in the analysis of supraconvergence [26].

To obtain more accurate approximations to y^*, we generate higher-degree polynomials inductively. In particular, for $k = 3, 4, 5, \ldots$, we define $p_k(Y_j^*; x)$ to be the $(2k)$th-degree polynomial which satisfies

$$(4.2)\qquad \begin{array}{lll} \text{(i)} & p_k(Y_j^*; x_i) = Y_i^*, & i = j-1, j+1, \\ \text{(ii)} & p_k''(Y_j^*; x_{jl}) = f(x_{jl}, p_{k-1}(Y_j^*; x_{jl})), & l = 0, \pm 1, \ldots, \pm(k-1), \end{array}$$

where

$$(4.3)\qquad x_{jl} \equiv \begin{cases} x_j + \frac{l}{k-1} h_{j+\frac{1}{2}}, & l \geq 0, \\ x_j + \frac{l}{k-1} h_{j-\frac{1}{2}}, & l \leq 0. \end{cases}$$

Although $p_k(Y_j^*; x)$ will be only an $O(h_j^{2k+1})$ approximation to $y^*(x)$ on I_j in general, the following lemma shows that the leading error term has an important property.

Lemma 1. *For $x \in I_j$ we have*

$$(4.4)\qquad \begin{aligned} y^*(x) - p_k(Y_j^*; x) = h_j^{2k+1} \sum_{p=2}^{k} & [f_y(x_j, Y_j^*)]^{k-p} y^{*(2p+1)}(x_j) \\ & \cdot P_{k,p}\left(\theta_j; \frac{x - x_j}{h_j} + \theta_j\right) + O(h_j^{2k+2}), \end{aligned}$$

where

(i) *$P_{k,k}(\theta; t)$ is a polynomial of degree $2k+1$ in t with coefficients that are polynomials in θ and $P_{k,k}(0; t)$ is odd,*

(ii) *$P_{k,i}(\theta; t)$, $i = 2, \ldots, k-1$, are polynomials of degree $2k$ in t with coefficients that are rational in θ (the divisors being powers of $m - \theta$ for nonzero integers m) and $P_{k,i}(0; t)$, $i = 2, \ldots, k-1$, are odd.*

Proof. First consider p_2 so that

$$y^*(x) - p_2(Y_j^*; x) = h_j^5 P_{2,2}\left(\theta_j; \frac{x - x_j}{h_j} + \theta_j\right) y^{*(5)}(x_j) + O(h_j^6),$$

where $P_{2,2}(\theta; t)$ is a quintic polynomial in t with zeros at ± 1 and whose second derivative is zero at θ and ± 1. Hence,

$$P''_{2,2}(\theta; t) \equiv \text{const}(t^2 - 1)(t - \theta)$$

and so is odd for $\theta = 0$, and this property is inherited by $P_{2,2}(0; t)$.

For $k \geq 3$ we argue by induction. Thus suppose the lemma is true for $y^*(x) - p_{k-1}(Y_j^*; x)$. Then

$$y^*(x) - p_k(Y_j^*; x) = y^*(x) - q_k(Y_j^*; x) + q_k(Y_j^*; x) - p_k(Y_j^*; x),$$

where $q_k(Y_j^*; x)$ is the $(2k)$th-degree polynomial which satisfies

$$\begin{aligned} &\text{(i)} \quad q_k(Y_j^*; x_i) = Y_i^*, && i = j-1, j+1, \\ &\text{(ii)} \quad q_k''(Y_j^*; x_{jl}) = y^{*\prime\prime}(x_{jl}), && l = 0, \pm 1, \dots, \pm(k-1). \end{aligned}$$

Consequently, for $x \in I_j$,

$$y^*(x) - q_k(Y_j^*; x) = h_j^{2k+1} P_{k,k}\left(\theta_j; \frac{x - x_j}{h_j} + \theta_j\right) y^{*(2k+1)}(x_j) + O(h_j^{2k+2}),$$

where $P_{k,k}(\theta; t)$ is a $(2k+1)$th-degree polynomial in t with zeros at ± 1 and whose second derivative has zeros at $t_l(\theta_j)$, $l = 0, \pm 1, \dots, \pm(k-1)$, where

$$t_l(\theta) \equiv \frac{l}{k-1} + \theta\left(1 - \frac{|l|}{k-1}\right). \tag{4.5}$$

Hence,

$$P''_{k,k}(\theta; t) \equiv \text{const} \prod_{l=-(k-1)}^{k-1} (t - t_l(\theta))$$

and so is odd for $\theta = 0$, and this property is inherited by $P_{k,k}(0; t)$.

Now consider the $q_k(Y_j^*; x) - p_k(Y_j^*; x)$ term, which is the $(2k)$th-degree polynomial with zeros at $x_{j\pm 1}$ and whose second derivative satisfies

$$q_k''(Y_j^*; x_{jl}) - p_k''(Y_j^*; x_{jl}) = f(x_{jl}, y^*(x_{jl})) - f(x_{jl}, p_{k-1}(Y_j^*; x_{jl})),$$
$$l = 0, \pm 1, \dots, \pm(k-1),$$

and thus is zero at $x_{j\pm 1}$. By the induction hypothesis,

$$\begin{aligned} q_k''&(Y_j^*; x) - p_k''(Y_j^*; x) \\ &= h_j^{2k-1} \sum_{l=-(k-2)}^{(k-2)} \left\{ \sum_{p=2}^{k-1} [f_y(x_j, Y_j^*)]^{k-p} y^{*(2p+1)}(x_j) P_{k-1,p}(\theta_j; t_l(\theta_j)) \right\} \\ &\qquad \cdot l_{k,l}\left(\theta_j; \frac{x - x_j}{h_j} + \theta_j\right) + O(h_j^{2k}), \end{aligned}$$

where the $l_{k,l}(\theta; t)$ are the Lagrange polynomials for the $2k-1$ points $t_l(\theta)$, i.e.,

$$l_{k,l}(\theta; t) \equiv \prod_{\substack{i=-(k-1) \\ i \neq l}}^{k-1} (t - t_i(\theta))/(t_l(\theta) - t_i(\theta)). \tag{4.6}$$

Hence, the induction will be complete if we can show that

$$\begin{aligned}P''_{k,p}(\theta\,;\,t) \equiv{}& P_{k-1,p}(\theta\,;\,t_0(\theta))l_{k,0}(\theta\,;\,t)\\&+\sum_{l=1}^{k-2} P_{k-1,p}(\theta\,;\,t_l(\theta))l_{k,l}(\theta\,;\,t)+P_{k-1,p}(\theta\,;\,t_{-l}(\theta))l_{k,-l}(\theta\,;\,t)\end{aligned}$$

is odd when $\theta = 0$. Since it is easily checked that

$$P_{k-1,p}(0\,;\,t_{-l}(0)) = -P_{k-1,p}(0\,;\,t_l(0)),$$

through $P_{k-1,p}(0\,;\,t)$ being odd and $t_{-l}(0) = -t_l(0)$, and

$$l_{k,l}(0\,;\,t) = \text{const}(|l|)t\left(t-\frac{l}{k-1}\right)\prod_{\substack{i=1\\ i\neq l}}^{k-1}\left(t^2-\left[\frac{i}{k-2}\right]^2\right),$$

we are finished. □

We shall also require these polynomials evaluated at other mesh functions near Y^*, and so we use the obvious definition for $p_k(Y_j\,;\,x)$, $j = 1, \dots, N$; $k = 2, 3, 4, \dots$. The following lemma shows how $p_k(Y_j^*\,;\,x) - p_k(Y_j\,;\,x)$ and its derivatives are bounded in terms of the mesh function $Y^* - Y$.

Lemma 2. *There holds*

$$\begin{aligned}&\text{(i)}\quad \|p''_k(Y_j^*\,;\,\cdot) - p''_k(Y_j\,;\,\cdot)\|_{I_j} \le H_k\|Y^* - Y\|_0,\\ (4.7)\quad &\text{(ii)}\quad \|p'_k(Y_j^*\,;\,\cdot) - p'_k(Y_j\,;\,\cdot)\|_{I_j} \le 2\|Y^* - Y\|_1 + \frac{h_j}{2}H_k\|Y^* - Y\|_0,\\ &\text{(iii)}\quad \|p_k(Y_j^*\,;\,\cdot) - p_k(Y_j\,;\,\cdot)\|_{I_j} \le (1 + H_k h_j^2/6)\|Y^* - Y\|_0,\end{aligned}$$

where $H_2 \equiv \alpha L_2(\theta_j)$ *and* $H_{k+1} \equiv \alpha L_k(\theta_j)(1 + H_k h_j^2/6)$, $k \ge 2$, *with* $L_k(\theta_j)$ *being the norm of the Lagrange interpolation operator on* $[-1, 1]$ *for the* $2k-1$ *points* $t_l(\theta_j)$, $l = 0, \pm 1, \dots, \pm(k-1)$.

Proof. A simple induction based on

$$\|p''_k(Y_j^*\,;\,\cdot) - p''_k(Y_j\,;\,\cdot)\|_{I_j} \le L_k(\theta_j)\alpha\|p_{k-1}(Y_j^*\,;\,\cdot) - p_{k-1}(Y_j\,;\,\cdot)\|_{I_j}. \qquad \square$$

We have not emphasized the point, but note that the bounds in (4.7) are just in terms of $Y^* - Y$ at points $j-1, j, j+1$.

5. Higher-order accuracy by defect correction

We compute a sequence $\{Y^k\}$, $k = 2, 3, 4, \dots$, of mesh functions which will be shown to be $(2k+2)$th-order approximations to Y^*. Their defining equations are

$$(5.1)\qquad (D^2Y^k)_j + Q_j^k(\theta_j)\{f(x, p_k(Y_j^k\,;\,x))\} = 0, \qquad j = 1, \dots, N-1,$$

with $Y_0^k = Y_N^k = 0$. Here, $Q^k(\theta)$ is the weighted Gauss-Lobatto rule with k interior points for approximating

$$\int_{-1}^{1} z(t)\varphi_\theta\,dt$$

and, by a change-of-variable, $Q_j^k(\theta_j)\{f(x, p_k(Z_j; x))\}$ approximates

$$\frac{1}{h_j}\int f(x, p_k(Z_j; x))\varphi_j\, dx.$$

Hence, Q^k integrates $(2k+1)$th-degree polynomials exactly, and Q_j^k is $O(h_j^{2k+2})$ accurate for sufficiently smooth integrands. Existence, local uniqueness and construction of Y^k is based on the fixed point equation

$$Z = \mathscr{G}_k(Z), \tag{5.2}$$

where

$$[\mathscr{G}_k(Z)]_j \equiv Z_j - W_j$$

with W satisfying

$$(L_h^0 W)_j = (D^2 Z)_j + Q_j^k(\theta_j)\{f(x, p_k(Z_j; x))\}, \qquad j = 1, \ldots, N-1,$$

and $W_0 = W_N = 0$. We will show that $\mathscr{G}_k$, for sufficiently small h, is a contraction mapping on $\overline{B}(Y^{k-1}, d_k)$, where

$$d_k \equiv 2(\|\mathscr{G}_k(Y^*) - Y^*\|_1 + \|Y^{k-1} - Y^*\|_1),$$

and hence that the defect correction iteration

$$Z^{(m+1)} = \mathscr{G}_k(Z^{(m)}), \tag{5.3}$$

starting from $Z^{(0)} = Y^{k-1}$, will converge to a locally unique Y^k. First, however, we analyze $\mathscr{G}_k(Y^*) - Y^*$, i.e., the accuracy of the higher-order methods.

Note that

$$(L_h^0[\mathscr{G}_k(Y^*) - Y^*])_j = \tau_j^k + \pi_j^k, \tag{5.4}$$

where

$$\text{(i)}\quad \tau_j^k \equiv \frac{1}{h_j}\int f(x, y^*(x))\varphi_j\, dx - Q_j^k(\theta_j)\{f(x, y^*(x))\},$$

$$\text{(ii)}\quad \pi_j^k \equiv Q_j^k(\theta_j)\{f(x, y^*(x)) - f(x, p_k(Y_j^*; x))\}.$$

We consider these two terms separately.

(a) The first, τ^k, as in the defect correction error analysis for Y^1, is just a quadrature error. By our choice of $Q^k(\theta)$ we have

$$|\tau_j^k| \le C_k h_j^{2k+2}\|y^{*(2k+2)}\|_{I_j}.$$

(b) The second term, π^k, was not present in the defect correction error analysis for Y^1 and appears here because, for $k \ge 2$, we need to use the polynomials p_k to generate extra approximations for the quadrature. Using Lemma 1, we may write

$$\pi_j^k \equiv h_j^{2k+1}\sum_{p=2}^{k}[f_y(x_j, Y_j^*)]^{k-p+1} y^{*(2p+1)}(x_j) Q_j^k(\theta_j)$$

$$\cdot\left\{P_{k,p}\left(\theta_j; \frac{x - x_j}{h_j} + \theta_j\right)\right\} + O(h_j^{2k+2}),$$

and so the key terms are

$$Q_j^k(\theta_j)\left\{P_{k,p}\left(\theta_j;\frac{x-x_j}{h_j}+\theta_j\right)\right\}\equiv\int_{-1}^1 P_{k,p}(\theta_j;t)\varphi_{\theta_j}\,dt,\qquad p=2,\dots,k.$$

However, $P_{k,p}(0;t)$ is an odd polynomial. In addition, the formula

$$\int_{-1}^1 t^m\varphi_\theta\,dt=\frac{1+(-1)^{m+2}+\theta[1-(-1)^{m+2}]-2\theta^{m+2}}{(m+1)(m+2)(1-\theta^2)}$$

shows that integrating an odd polynomial in t against φ_θ gives an odd polynomial in θ. Hence,

$$g(\theta)\equiv\int_{-1}^1 P_{k,p}(\theta;t)\varphi_\theta\,dt$$

is a rational function of θ with $g(0)=0$, and this ensures that

$$\|\pi^k\|_{-1}=\mathrm{const}_k(\alpha,\overline{U},\overline{C},\overline{N})\max_j\left\{h_j^{2k+2}\sum_{p=5}^{2k+2}\|y^{*(p)}\|_{I_j}\right\}.$$

Hence, d_2 just depends on $\|\tau^2\|_{-1}$, $\|\pi^2\|_{-1}$ and $\|\tau^1\|_{-1}$; and we shall show by induction that, for $k>2$, d_k only depends on $\|\tau^k\|_{-1}$, $\|\pi^k\|_{-1}$, $\|\tau^{k-1}\|_{-1}$ and $\|\pi^{k-1}\|_{-1}$.

Now we will show that $\mathcal{G}_k$, for sufficiently small h, is a contraction mapping on $\overline{B}(Y^{k-1},d_k)$. If $U, V\in\overline{B}(Y^{k-1},d_k)$, then

$$(L_h^0[\mathcal{G}_k(U)-\mathcal{G}_k(V)])_j=-\delta_j^{k-1}-\lambda_j^{k-1},\tag{5.5}$$

where

(i) $\delta_j^{k-1}\equiv Q_j^k(\theta_j)\{f(x,p_k(U_j;x))-f(x,p_k(V_j;x))\}$
$-[f(x,U_j)-f(x,V_j)]$,

(ii) $\lambda_j^{k-1}\equiv f(x_j,U_j)-f(x_j,V_j)-\dfrac{\partial f}{\partial y}(x_j,Y_j^0)[U_j-V_j]$.

We consider these two terms separately.

(a) The first, δ^{k-1}, as in the case of Y^1, is the key term in the defect correction error analysis and relies on the smoothness of $U-V$. Since $Q^k(\theta)$ integrates linear functions exactly against φ_θ, we may write

$$\begin{aligned}\delta_j^{k-1}&=Q_j^k(\theta_j)\{f(x,p_k(U_j;x))-f(x,p_k(V_j;x))-e_1^k(x)\}\\&\quad+\frac{1}{2}\{h_{j+\frac12}(DE)_{j+\frac12}-h_{j-\frac12}(DE)_{j-\frac12}\}\\&\quad+\frac{\theta_j}{6}\{h_{j+\frac12}(DE)_{j+\frac12}+h_{j-\frac12}(DE)_{j-\frac12}\},\end{aligned}$$

where e_1^k is the linear polynomial interpolating

$$f(x,p_k(U_j;x))-f(x,p_k(V_j;x))$$

at $x_{j\pm1}$ and $E_j\equiv f(x_j,U_j)-f(x_j,V_j)$. The first term on the right-hand side will be $O(h_j^2\|U_j-V_j\|_1)$ if we use standard interpolation theory, Lemma 2 and the conditions on f given by (2.2). [Note that it is only here that Lipschitz continuity of the second derivatives of f is needed; cf. §6.] The other two terms we write as

$$\frac{h_j}{2}[(DE)_{j+\frac{1}{2}} - (DE)_{j-\frac{1}{2}}] \\ - h_j\left\{\frac{\theta_j^2}{6}[(DE)_{j+\frac{1}{2}} - (DE)_{j-\frac{1}{2}}] + \frac{\theta_j}{3}[(DE)_{j+\frac{1}{2}} + (DE)_{j-\frac{1}{2}}]\right\}.$$

The θ_j factors will ensure supraconvergence of the latter term, while for the former we have

$$\begin{aligned}&\left|\sum_{i=1}^{j}\frac{h_i^2}{2}[(DE)_{i+\frac{1}{2}} - (DE)_{i-\frac{1}{2}}]\right| \\ &= \left|-\frac{h_1^2}{2}(DE)_{\frac{1}{2}} + \sum_{i=2}^{j}\frac{h_{i-1}^2 - h_i^2}{2}(DE)_{i-\frac{1}{2}} + \frac{h_j^2}{2}(DE)_{j+\frac{1}{2}}\right| \\ &\le \left(h^2 + 2h\sum_{i=1}^{N-1} h_i|\theta_i|\right)\|E\|_1 .\end{aligned}$$

Consequently, on a supraconvergent mesh we have

$$\|\delta^{k-1}\|_{-1} \le \operatorname{const}(\overline{U}, \overline{C}, \overline{N})h^2\|E\|_1$$

and thus

$$\|\delta^{k-1}\|_{-1} \le \alpha\operatorname{const}(\overline{U}, \overline{C}, \overline{N})h^2\|U - V\|_1 .$$

We emphasize again that this constant only depends on the supraconvergence properties of the mesh.

(b) As before, λ_j^{k-1} is just a simple linearization error satisfying

$$\|\lambda^{k-1}\|_0 \le \alpha(b-a)d_1\|U - V\|_0 .$$

Hence $\mathcal{G}_k$, for sufficiently small h and d_k, is a contraction on $\overline{B}(Y^{k-1}, d_k)$, and we may assume that the contraction constant $\gamma_k \le \frac{1}{4}$.

In addition, $\mathcal{G}_k$ maps $\overline{B}(Y^{k-1}, d_k)$ onto itself since $Z \in \overline{B}(Y^{k-1}, d_k)$ implies

$$\|\mathcal{G}_k(Z) - Y^{k-1}\|_1 \le \|\mathcal{G}_k(Z) - \mathcal{G}_k(Y^*)\|_1 + \|\mathcal{G}_k(Y^*) - Y^*\|_1 + \|Y^* - Y^{k-1}\|_1 .$$

So our final conclusion is that the sequence of defect correction iterations (5.3) converges to locally unique $\{Y^k\}$ satisfying (5.1) and that, as usual,

$$Y^* - Y^k = Y^* - \mathcal{G}_k(Y^*) + \mathcal{G}_k(Y^*) - Y^k$$

implies

$$\|Y^k - Y^*\|_1 \le \frac{\kappa_0}{1-\gamma_k}(\|\tau^k\|_{-1} + \|\pi^k\|_{-1}). \tag{5.6}$$

We finish by repeating the crucial fact that the global mesh size h only appears in the contraction constants $\{\gamma_k\}$, while the errors $\{Y^k - Y^*\}$ just depend on the product of local mesh size and local derivative size.

6. Generalization and variations

(a) *Systems of differential equations.* For simplicity we have described the application of defect correction to a single second-order differential equation, but there is no difficulty in extending our results to systems.

(b) *Derivative boundary conditions.* Although we are specifically considering second-order differential equations without a first derivative term, there is no difficulty in catering for a derivative in the boundary conditions. These may be imposed variationally in the usual finite element way. For example, if

$$\varphi_0(x) \equiv \begin{cases} (x_1 - x)/h_{\frac{1}{2}}, & x \in I_{\frac{1}{2}}, \\ 0, & x \notin I_{\frac{1}{2}}, \end{cases}$$

then integrating (1) against φ_0 gives

$$y^{*\prime}(0) = \frac{y^*(x_1) - y^*(x_0)}{h_{\frac{1}{2}}} - \int f(x, y^*(x))\varphi_0\, dx,$$

and this expression can replace $y^{*\prime}(0)$ in any boundary condition. Of course, the integral must be approximated by increasingly accurate weighted quadrature rules, this time with weight φ_0, using only Y_0^*, Y_1^* and the differential equation.

(c) *Alternatives to p_k.* There is a natural alternative to our definition of p_k in §4, which should be somewhat more accurate. We define $\hat{p}_2(Y_j^*; x)$ to be the continuous piecewise quartic polynomial which satisfies

$$\text{(6.1a)} \qquad \begin{aligned} &\text{(i)} \quad \hat{p}_2(Y_j^*; x_i) = Y_i^*, && i = j-1, j, \\ &\text{(ii)} \quad \hat{p}_2''(Y_j^*; x_i) = f(x_i, Y_i^*), && i = j-1, j, j+1, \end{aligned}$$

for $x \in I_{j-\frac{1}{2}}$, and

$$\text{(6.1b)} \qquad \begin{aligned} &\text{(i)} \quad \hat{p}_2(Y_j^*; x_i) = Y_i^*, && i = j, j+1, \\ &\text{(ii)} \quad \hat{p}_2''(Y_j^*; x_i) = f(x_i, Y_i^*), && i = j-1, j, j+1, \end{aligned}$$

for $x \in I_{j+\frac{1}{2}}$. Similarly, for $k = 3, 4, 5, \ldots$ we define $\hat{p}_k(Y_j^*; x)$ inductively to be the continuous piecewise $(2k)$th-degree polynomial which satisfies

(6.2a)

$$\begin{aligned} &\text{(i)} \quad \hat{p}_k(Y_j^*; x_i) = Y_i^*, && i = j-1, j, \\ &\text{(ii)} \quad \hat{p}_k''(Y_j^*; x_{jl}) = f(x_{jl}, \hat{p}_{k-1}(Y_j^*; x_{jl})), && l = 0, \pm 1, \ldots, \pm(k-1), \end{aligned}$$

for $x \in I_{j-\frac{1}{2}}$, and

(6.2b)

$$\begin{aligned} &\text{(i)} \quad \hat{p}_k(Y_j^*; x_i) = Y_i^*, && i = j, j+1, \\ &\text{(ii)} \quad \hat{p}_k''(Y_j^*; x_{jl}) = f(x_{jl}, \hat{p}_{k-1}(Y_j^*; x_{jl})), && l = 0, \pm 1, \ldots, \pm(k-1), \end{aligned}$$

for $x \in I_{j+\frac{1}{2}}$, where x_{jl} is defined in (4.3).

It is easy to develop analogues of Lemmas 1 and 2 for the $\hat{p}_k$, but a complete defect correction error analysis is complicated by their piecewise polynomial nature. It is, however, expected to hold and the numerical results in §7 provide verification.

We further note that it is possible to vary the x_{jl} in the definition of p_k and $\hat{p}_k$, always enforcing the condition that they are symmetric with respect to the center of I_j when $\theta_j = 0$. The appearance of norms of Lagrange interpolation operators in the proof of Lemma 2 indicates that an analogue of Chebyshev points would be advantageous, but this has yet to be investigated.

(d) *Alternative* $Q^k(\theta)$. There is no special reason for our choice of weighted Gauss-Lobatto quadrature, apart from the fact that it provides the highest accuracy with the minimum number of function evaluations. If the $\hat{p}_k$ piecewise polynomials are used, so that $f(x_j, \hat{p}(Y_j; x_j)) \equiv f(x_j, Y_j)$ is an extra free integrand value, there are other possibilities which use an equal number or sometimes even fewer function evaluations.

(i) For odd k we look at a generalization of the $k = 1$ approach. Thus, we require a "generalized" weighted Gauss-Lobatto rule with k interior points, one of which is fixed at x_j. For $\theta_j = 0$, these are ordinary weighted Gauss-Lobatto rules and integrate $(2k + 1)$th-degree polynomials exactly. For $\theta \neq 0$ the schemes will only have abscissae in $(-1, 1)$ for $|\theta|$ sufficiently small (bounds given in the table below), and the supraconvergence property must be relied on to obtain the required accuracy, as with $k = 1$.

k	1	3	5
$\lvert\theta\rvert <$	1	.27	.16

Note that one integrand evaluation is saved, compared with the weighted Gauss-Lobatto rules.

(ii) For even k we may look at a generalization of the $k = 0$ approach, i.e., "generalized" weighted Gauss rules with one point fixed at x_j. For $\theta_j = 0$ these are ordinary weighted Gauss rules and integrate $(2k + 1)$th-degree polynomials exactly. For $\theta \neq 0$ the schemes will only have abscissae in $(-1, 1)$ for $|\theta|$ sufficiently small (bounds given in the table below), and the supraconvergence property must again be relied on to achieve the required accuracy, as with $k = 0$.

k	0	2	4
$\lvert\theta\rvert <$	1	.30	.17

Note that these schemes use the same number of extra integrand evaluations as the weighted Gauss-Lobatto rules.

(e) *Alternative* $k = 0$ *formulae.* There is no difficulty in using a different difference equation as the basic second-order accurate method on which our defect correction procedure relies. Those derived by integrating linear functions (weighted trapezoidal rule) or continuous piecewise linear functions exactly against φ_j are respectively:

$$\text{(i)} \quad (D^2Y)_j + \left(\frac{1}{2} + \frac{\theta_j}{6}\right) f(x_{j+1}, Y_{j+1}) + \left(\frac{1}{2} - \frac{\theta_j}{6}\right) f(x_{j-1}, Y_{j-1}) = 0,$$

$$\text{(ii)} \quad (D^2Y)_j + \frac{1-\theta_j}{6} f(x_{j+1}, Y_{j+1}) + \frac{2}{3} f(x_j, Y_j) + \frac{1+\theta_j}{6} f(x_{j-1}, Y_{j-1}) = 0.$$

It is easily checked that the defect correction analysis in §5 still carries through.

More interesting is the choice of the fourth-order accurate method (2.7) as our basic $k = 0$ formula. Our new L_h^0 will still only involve the solution of a tridiagonal system, so this is a serious possibility. It is straightforward to adapt the theory to show that a similar improvement of $O(h^2)$ per correction is still achieved, and this is verified numerically in the next section.

7. Numerical results

In this section we present some numerical results for the integration of two challenging second-order equations. The particular problems we consider are:

$$(1) \qquad \varepsilon^2 y'' = y - x, \qquad 0 \le x \le 1, \; y(0) = 1, \; y(1) = 2,$$

which has the solution

$$y(x) = x + \frac{e^{(x-1)/\varepsilon}}{1 + e^{-1/\varepsilon}} + \frac{e^{-x/\varepsilon} - e^{-(x+1)/\varepsilon}}{1 - e^{-2/\varepsilon}}.$$

For small positive ε, this solution has a boundary layer at both ends of the integration range and is smooth away from these layer regions. For the purpose of our numerical experiments we take $\varepsilon = 10^{-2}$.

$$(2) \qquad y'' = \mu \sinh \mu y, \qquad 0 \le x \le 1, \; y(0) = 0, \; y(1) = 1.$$

This is a very well-known problem due to Troesch. This equation does not have an analytic solution and becomes increasingly more difficult to solve numerically as the parameter μ increases. In particular, it presents a very difficult problem for $\mu = 20$, owing to the presence of a sharp boundary layer at $x = 1$.

For both of these examples, the behavior of the solution varies significantly over the interval of x, and so it makes good sense to use a nonuniform grid. The purpose of using such a grid is two-fold. First, by concentrating grid points in regions where the solution is varying rapidly (and putting relatively few grid points in smooth regions) we hope to achieve good accuracy using relatively few grid points, certainly less than if a uniform grid were to be used. Secondly, for nonlinear problems, we expect that a nonuniform grid will facilitate the convergence of the Newton iteration scheme used to solve the nonlinear algebraic equations which define the numerical solution.

The deferred correction algorithm on which we base our integration scheme is

$$\begin{aligned} &\varphi_4(\eta) = 0, \\ &\varphi_4(\overline{\eta}) = -\varphi_6(\eta), \\ &\varphi_4(\overline{\overline{\eta}}) = \varphi_4(\overline{\eta}) - \varphi_8(\overline{\eta}), \end{aligned}$$

where φ_i denotes a formula which is of order i on a uniform grid. It is not difficult to derive these formulae explicitly using the theory presented in previous sections, and they are available from the authors on request. The precise way in which these formulae are used in a deferred correction framework is described in [6]. Although our theory applies to very general meshes, we simplify our implementation by allowing only mesh halving/doubling. This corresponds to the case $\theta = 0$ or $\pm 1/3$ in §2. It is straightforward to store quadrature formulae for these three cases and also to store coefficients which generate $p_k(Y_j\,;\,\cdot)$ from $p_{k-1}(Y_j\,;\,\cdot)$. However, in future work we wish to examine the possibility of using more general meshes in our numerical implementation.

In a practical algorithm it would be necessary to derive an automatic mesh selection procedure. However, this has proved to be difficult and is certainly beyond the scope of the present paper. In deriving our numerical results, we have used a somewhat crude method for deriving an adaptive grid. As in the case of the deferred correction algorithm described in [5, 6], we seek to choose the grid so as to equidistribute the eighth-order deferred correction $|\varphi_4(\bar{\eta}) - \varphi_8(\bar{\eta})|$. However, in the second-order case considered in this paper we have imposed the constraint that successive grid spacings should be in the ratio $1:1$, $1:2$ or $2:1$. Although this does simplify the mesh selection process, it is still nontrivial to satisfy this constraint. In deriving our numerical results, what we actually did was as follows. Given an initial grid, we used our deferred correction program to compute 4th, 6th and 8th-order numerical solutions and to derive the eighth-order deferred correction $\varphi_4(\bar{\eta}) - \varphi_8(\bar{\eta})$ associated with each mesh. If the error criterion was not satisfied, we worked out a new grid by hand on the basis of approximately equidistributing the eighth-order deferred correction, fed this new grid into our deferred correction algorithm and computed one more loop of the iteration. This process was continued until either the required accuracy was obtained or more than a maximum number of grid points was used.

Using the procedure just described, we solved problem (1) with a requested absolute accuracy of 10^{-10}. We ended up solving this problem on a nonuniform grid of 146 points, of which 51 were in $[0, 0.1]$ and 47 in $[0.9, 1]$. On this mesh, the maximum errors in η, $\bar{\eta}$ and $\bar{\bar{\eta}}$ were $.203 \times 10^{-5}$, $.553 \times 10^{-8}$ and $.551 \times 10^{-10}$, respectively. On an equally spaced grid we found that about 500 points were needed to obtain the same accuracy. Finally, we halved the nonuniform grid to obtain a grid with 291 points to see how these maximum errors behaved. On this new halved grid, the maximum errors in η, $\bar{\eta}$ and $\bar{\bar{\eta}}$ were $.129 \times 10^{-6}$, $.877 \times 10^{-10}$ and $.169 \times 10^{-12}$, with the ratio between the new and old maximum errors being 15.7, 63 and 311, respectively.

For problem (2) the situation is more complicated. Not only do we need to choose an adaptive grid to satisfy the accuracy requirements, but we also need to ensure that the Newton iteration scheme used to solve for the numerical solution will converge. So, as not to introduce additional complications, we used a straightforward (undamped) Newton scheme with initial approximation $y = 0$. In what follows, we describe the solution of problem (2) with an accuracy requirement of 10^{-8}. We started with $\mu = 2$ and 10 equally spaced grid points, and performed continuation in μ with increments of 2 to get a reasonable grid for $\mu = 20$. This process produced a grid of 59 points with 46 of these being in $[0.9, 1]$. We solved problem (2) on this grid (using Newton with initial guess $y = 0$) and found that the maximum errors in η, $\bar{\eta}$ and $\bar{\bar{\eta}}$ were $.287 \times 10^{-5}$, $.368 \times 10^{-7}$ and $.700 \times 10^{-8}$. Halving this grid, we found the maximum errors to be $.190 \times 10^{-6}$, $.628 \times 10^{-9}$, $.290 \times 10^{-10}$, with the respective ratios being $15.1, 59, 276$. Finally, we attempted to solve problem (2) using our deferred correction scheme on a uniform grid. However, we found that we were unable to obtain convergence of the Newton iteration scheme, even with a grid of 2000 points.

Bibliography

1. U. Ascher, R. Mattheij, and R. D. Russell, *Numerical solution of boundary value problems for ODEs*, Prentice-Hall, Englewood Cliffs, NJ, 1988.

→ U. Ascher, J. Christiansen, and R. D. Russell, *Collocation software for boundary value ordinary differential equations*, ACM Trans. Math. Software 7 (1981), 209–222.

3. J. W. Barrett, G. Moore, and K. W. Morton, *Optimal recovery in the finite element method, Part* 2: *Defect correction for ordinary differential equations*, IMA J. Numer. Anal. **8** (1988), 527–540.

4. K. Bohmer and H. J. Stetter (eds.), *Defect correction methods: Theory and applications*, Springer-Verlag, Wien, 1984.

5. J. R. Cash, *Numerical integration of nonlinear two-point boundary value problems using iterated deferred correction*–I. *A survey and comparison of some one-step formulae*, Comput. Math. Appl. **12** (1986), 1029–1048.

6. ——, *Numerical integration of nonlinear two-point boundary value problems using iterated deferred correction*–II. *The development and analysis of highly stable deferred correction formulae*, SIAM J. Numer. Anal. **25** (1988), 862–882.

7. ——, *Adaptive Runge-Kutta methods for nonlinear two-point boundary value problems with mild boundary layers*, Comput. Math. Appl. **12** (1985), 605–620.

8. —→—, *High order P-stable formulae for the numerical integration of periodic initial value problems*, Numer. Math. **37** (1981), 355–370.

9. J. R. Cash and A. Singhal, *Mono-implicit Runge-Kutta formulae for the numerical integration of stiff differential equations*, IMA J. Numer. Anal. **2** (1982), 211–227.

10. ——, *High order methods for the numerical solution of two-point boundary value problems*, BIT **22** (1986), 184–199.

11. J. R. Cash and M. H. Wright, *A deferred correction method for nonlinear two-point boundary value problems: implementation and numerical evaluation*, AT&T Bell Laboratories, Comput. Sci. Report No. 146, Murray Hills, NJ, 1989.

12. M. M. Chawla, *A sixth order tri-diagonal finite difference method for nonlinear two-point boundary value problems*, BIT **17** (1977), 128–133.

13. ——, *An eighth order tri-diagonal finite difference method for nonlinear two-point boundary value problems*, BIT **17** (1977), 281–285.

14. J. W. Daniel and A. J. Martin, *Numerov's method with deferred corrections for two-point boundary value problems*, SIAM J. Numer. Anal. **14** (1977), 1033–1050.

→ C. R. Doering, J. D. Gibbon, D. D. Holm, and B. Nicolaenko, *Low dimensional behaviour in the complex Ginzburg-Landau equation*, Nonlinearity **1** (1988), 279–309.

16. R. D. Grigorieff, *On stability constants and condition number of discretization methods*, Report No. 149, Technische Universität Berlin, 1986.

17. P. Henrici, *Discrete variable methods in ordinary differential equations*, Wiley, New York, 1962.

18. M. S. Jolly, *Explicit construction of an inertial manifold for a reaction diffusion equation*, J. Differential Equations **78** (1989), 220–261.

19. H. O. Kreiss, *Difference approximations for boundary and eigenvalue problems for ODEs*, Math. Comp. **26** (1972), 605–624.

20. H. O. Kreiss, T. A. Manteuffel, B. Swartz, B. Wendroff, and A. B. White, *Supra-convergent schemes on irregular grids*, Math. Comp. **47** (1986), 537–554.

21. M. Lentini and V. Pereyra, *A variable order finite difference method for nonlinear multi-point boundary value problems*, Math. Comp. **28** (1974), 981–1004.

22. ——, *An adaptive finite difference solver for nonlinear two-point boundary value problems with mild boundary layers*, SIAM J. Numer. Anal. **14** (1977), 91–111.

23. T. A. Manteuffel and A. B. White, *The numerical solution of second-order boundary value problems on nonuniform meshes*, Math. Comp. **47** (1986), 511–535.

24. G. Moore, *Defect correction from a Galerkin viewpoint*, Numer. Math. **52** (1988), 565–582.

25. V. Pereyra, PASVA3: *An adaptive finite difference FORTRAN program for first order nonlinear, ordinary boundary value problems*, Codes for Boundary Value Problems in Ordinary Differential Equations (B. Childs, M. Scott, J. E. Daniel, E. Denman, and P. Nelson, eds.), Springer-Verlag, Berlin, 1979, pp. 67–88.

26. R. D. Skeel, *A theoretical framework for proving accuracy results for deferred corrections*, SIAM J. Numer. Anal. **19** (1982), 171–196.

27. R. D. Skeel and M. Berzins, *A method for the spatial discretization of parabolic equations in one space variable*, SIAM J. Sci. Statist. Comput. **11** (1990), 1–32.

(Butcher) Department of Mathematics and Statistics, University of Auckland, Private Bag 92019, Auckland, New Zealand
E-mail address, J. C. Butcher: `butcher@mat.auckland.ac.nz`

(Cash and Moore) Department of Mathematics, Imperial College, South Kensington, London SW7 2BZ, England
E-mail address, J. R. Cash: `J.Cash@ic.ac.uk`
E-mail address, G. Moore: `G.Moore@ic.ac.uk`

(Russell) Department of Mathematics, Simon Fraser University, Burnaby, British Columbia, Canada V5A 1S6
E-mail address, R. D. Russell: `russell@cs.sfu.ca`