ELSEVIER

Chemical Engineering Journal 71 (1998) 127–134

Chemical Engineering Journal

A technique for accurate collocation residual calculations

Raymond A. Adomaitis*, Yi-hung Lin

Department of Chemical Engineering and Institute for Systems Research, University of Maryland, College Park, MD 20742, USA

Received 17 March 1998; received in revised form 3 July 1998; accepted 24 July 1998

Abstract

A multiple-grid collocation method is presented that allows exact evaluation of residuals generated by truncated trial function expansion solutions to boundary-value problems with polynomial nonlinearities. The method is used to formulate a true, discrete analog to the Galerkin projection applicable to the same class of problems. The numerical techniques developed are used to study the convergence behavior of a nonlinear, reaction-diffusion problem as a function of Thiele modulus (ϕ) and trial function truncation number (N). The convergence problems encountered at high ϕ values are found to result from a second, physically meaningless solution to the modeling equations. This 'spurious' solution and the true solution are involved in a saddle-node bifurcation that limits the range of ϕ where solutions are found for most finite N; the solutions appear to asymptotically approach each other as ϕ, $N \to \infty$ regardless of the discretization method. The saddle–stable manifold of the spurious solution also defines the boundary of the set of initial conditions that diverge during dynamic simulations prior to the saddle-node bifurcation; all initial conditions are found to diverge after this bifurcation point.

Keywords: Orthogonal collocation; Galerkin projection; Method of weighted residual; Numerical analysis

1. Introduction

The orthogonal collocation method [1] was developed originally as a stable, predictable, and simple to implement pseudospectral technique. Because of its reliability, it has become a standard method for solving boundary-value problems by polynomial trial function expansions [2–5]. The interior formulation of this method [1] is based on choosing a set of trial functions $\{\eta_n\}_{n=1}^{N}$ from an orthogonal polynomial sequence, with the discretization points computed as the roots of the polynomial η_{N+1} next in the sequence. In most applications, this approximates the Galerkin procedure, because the residual is forced to have as its primary component the polynomial used to determine the collocation points. However, for some linear problems where the residual can be expressed exactly in terms of the chosen set of trial functions, the two methods give identical results.

In this paper, we develop a multiple-grid implementation of this collocation method that extends to some nonlinear problems the ability to calculate the residual exactly. Furthermore, the precise residual calculation capabilities of the multiple-grid collocation approach are used to modify the discretization procedure to obtain an exact analog to the Galerkin projection for the same class of problems. We use these discretization techniques to study the solutions to a nonlinear reaction–diffusion problem known to exhibit convergence problems for large Thiele modulus (ϕ) values. We use a combination of perturbation analysis and numerical continuation techniques to show that the problem stems from a second, physically meaningless solution generated during a saddle-node bifurcation at a value of ϕ that grows with increasing trial function truncation number N. This spurious, steady-state concentration profile, however, is not an artifact of the discretization procedure: the residual analysis techniques developed in this paper are used to show that it is a convergent solution to the boundary value problem. We find that the solutions asymptotically approach each other as ϕ, $N \to \infty$ regardless of the discretization method (cf. the parasitic solutions found in [7], which disappeared with increasing trial function truncation number). We demonstrate that the numerical problems encountered in previous simulation studies (e.g., [5,6]) are attributable to this spurious solution and its interaction with the 'true' solution. Further insight into the mechanisms that lead to divergence of dynamic simulations is obtained by the analysis of the global dynamic behavior of this system in the

*Corresponding author. Tel.: +1-301-405-2969; fax: +1-301-314-9920; e-mail: adomaiti@isr.umd.edu

1385-8947/98/$ – see front matter
PII: S1385-8947(98)00127-2

phase space of a two-point discretization of the reaction–diffusion problem. It will be shown that prior to the saddle-node bifurcation, the saddle-stable manifold of the spurious solution separates initial conditions that diverge from those that converge to the true solution. Because no N-mode steady-state solutions exist after the bifurcation points associated with the different N, all initial conditions will diverge when integrated forward in time for these Thiele modulus values.

2. The reaction–diffusion problem

We develop the multiple-grid collocation technique in the context of computing solutions to the nonlinear reaction–diffusion equation in a cylindrical physical domain

$$\frac{\partial c}{\partial t} = \frac{1}{x}\frac{\partial}{\partial x}\left(x\frac{\partial c}{\partial x}\right) - \phi^2 c^2$$

subject to boundary conditions $\partial c(0,t)/\partial x = 0$, $c(1,t) = 1$ and initial condition $c(x,0) = c_0(x)$. This problem can be converted into a nonhomogeneous partial differential equation subject to homogeneous boundary conditions with the transformation $c(x,t) = u(x,t) + 1$ to give

$$\frac{\partial u}{\partial t} = \frac{1}{x}\frac{\partial}{\partial x}\left(x\frac{\partial u}{\partial x}\right) - \phi^2(u+1)^2 \tag{1}$$

and boundary conditions $\partial u(0,t)/\partial x = 0$ and $u(1,t) = 0$. We represent $u(x,t)$ by the trial function expansion

$$u(x,t) = \sum_{n=1}^{\infty} b_n(t)(1-x^2)\eta_n(x^2) = \sum_{n=1}^{\infty} b_n(t)\psi_n(x) \tag{2}$$

which can be written in vector notation $u^N(x) = \boldsymbol{\psi}\boldsymbol{b}$ for truncated expansions. In this notation, the vector of trial functions is defined as $\boldsymbol{\psi}(x) = (1-x^2)[\eta_1(x^2), \ldots, \eta_N(x^2)]$ and the corresponding mode amplitude coefficients are $\boldsymbol{b} = [b_1, b_2, \ldots, b_N]^{\mathrm{T}}$. By using only even powers of x, the trial functions ψ_n satisfy the boundary condition at $x = 0$; the factor $(1-x^2)$ forces the trial functions to also satisfy the boundary condition at $x = 1$. The polynomials η_n are constructed as the normalized Jacobi polynomials defined by a sequence of $2n$-th degree polynomials orthonormal with respect to weighted inner product

$$\begin{aligned} &\int_0^1 \eta_i\eta_j(1-x^2)x\,\mathrm{d}x = 0 \quad \text{for} \quad i \neq j \\ &\int_0^1 \eta_i\eta_j(1-x^2)x\,\mathrm{d}x = 1 \quad \text{for} \quad i = j \end{aligned} \tag{3}$$

with $\eta_1 = 2$.

3. Interior collocation

The discrete-ordinate formulation of the interior collocation method is based on defining the discrete transformation array $\boldsymbol{Q}$:

$$\boldsymbol{u}^{N\times 1} = \boldsymbol{Q}^{N\times N}\boldsymbol{b}^{N\times 1} \text{ with } Q_{i,j} = \psi_j(x_i)$$

The collocation point locations x_n are determined as the roots of $\eta_{N+1}(x)$ (the interior roots of $\psi_{N+1}(x)$). The discretization array is defined as

$$\frac{1}{x}\frac{\mathrm{d}}{\mathrm{d}x}x\frac{\mathrm{d}\boldsymbol{u}^{N\times 1}}{\mathrm{d}x} = \boldsymbol{B}^{N\times N}\boldsymbol{u}^{N\times 1} \text{ with } \boldsymbol{B} = \left[\frac{1}{x}\frac{\mathrm{d}}{\mathrm{d}x}x\frac{\mathrm{d}\boldsymbol{Q}}{\mathrm{d}x}\right]\boldsymbol{Q}^{-1}$$

Therefore, Eq. (1) can be discretized to give the N ordinary differential equations in time

$$\frac{\mathrm{d}u_n}{\mathrm{d}t} = \boldsymbol{B}_n\boldsymbol{u} - \phi^2(u_n+1)^2 \quad n = 1, \ldots, N \tag{4}$$

with $\boldsymbol{B}_n$ defined as the n-th row of $\boldsymbol{B}$. This set of equations can be represented in matrix form by defining the column vector $\mathbf{1}$ as a $(N \times 1)$ vector with elements all equal to unity, to give

$$\frac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}t} = \boldsymbol{B}\boldsymbol{u} - \phi^2(\boldsymbol{u}+\mathbf{1})^2 \tag{5}$$

The discretized system Eq. (5) has more than one steady-state solution; this fact has not been discussed extensively in the literature on determining solutions to Eq. (1), but is critical to understand the problems associated with computing steady-state solutions and dynamic simulations for large ϕ. Consider Fig. 1 where the two coexisting steady-state solutions are plotted for $\phi = 4$. It is easy to visually distinguish between the 'true' solution and the physically meaningless (spurious) solution because the spurious solution crosses the $c = 0$ axis. To determine if this solution is an artifact of the collocation discretization of the original problem Eq. (1), the convergence of each solution is examined using the residual computation technique developed in the next section. Plotting the L^2 norm of each steady-state residual as a function of N confirms that each is a convergent solution (see Fig. 1, right-most plot).

3.1. Residual calculations

The residual function produced by the collocation procedure is defined as the function obtained by substituting the trial function expansion solution Eq. (2) into the original problem Eq. (1) with the values of the mode amplitude coefficients computed from Eq. (4). Solving Eq. (4) forces the residual to vanish at the collocation points; between the collocation points the residual function will either be zero or will be a function that, when decomposed in terms of the normalized trial functions $\{\eta_i\}_{i=1}^{\infty}$ by projecting the residual function onto each $\{\psi_i\}_{i=1}^{\infty}$, will normally have as its most significant component $d_{N+1}\eta_{N+1}$ with $d_{N+1} = d_{N+1}(t)$. If the trial function expansion Eq. (2) is truncated at $\psi_N(x)$, the trial function expansion approximating the solution will have polynomial terms of up to degree $2N$ and the residual produced by Eq. (1) will contain polynomials of up to degree $4N$. Because a polynomial function containing only terms of up to this degree can be expressed exactly by a

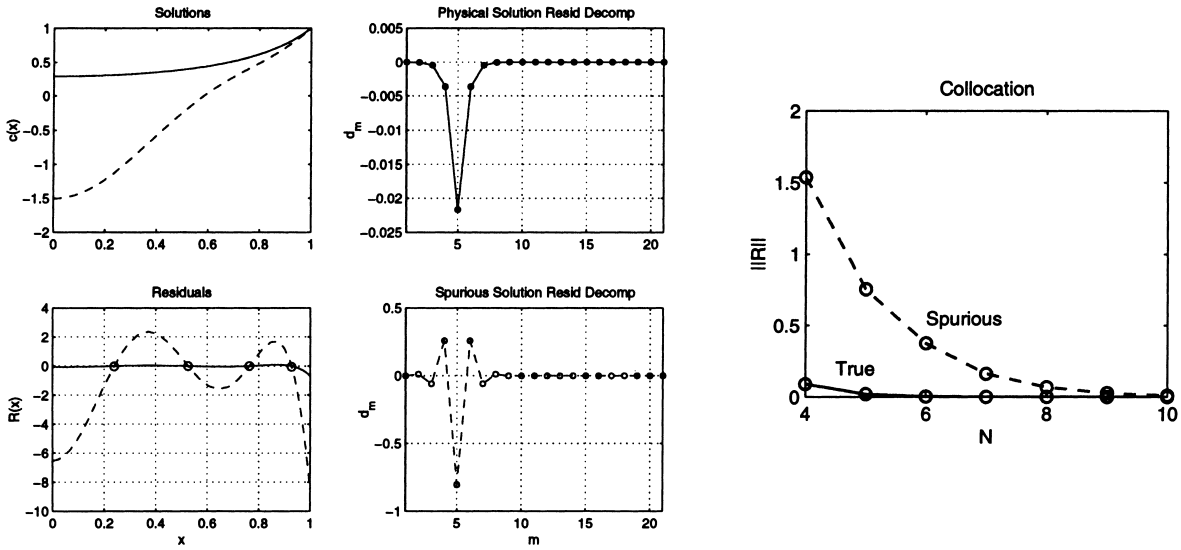

Fig. 1. Steady-state solutions and convergence behavior produced by the orthogonal collocation discretization technique for $\phi = 4$ and $N = 4$. Spurious solution is denoted with the dashed curve.

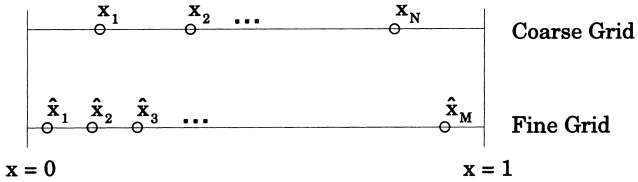

Fig. 2. Coarse- and fine-scale spatial discretization grids.

combination of the first $M = (4N/2) + 1$ functions η_m, we can use a finer scale of collocation points $\hat{x}_m$, $m = 1, \ldots, M$ (shown in Fig. 2) to define the discrete transformation and discretization arrays necessary to compute the mode amplitude coefficients d_m associated with the trial functions η_m for the residual function $R(x,t)$ evaluated at these points (see Fig. 1, center plots, for the residual decomposition results).

The $\hat{x}_m$ are computed as the roots of η_{M+1}. To compute the fine-grid residual values corresponding to a solution defined on the coarse grid, we define the interpolation array

$$\hat{\boldsymbol{u}}^{M\times 1} = \hat{\boldsymbol{Q}}^{M\times N}\boldsymbol{b}^{N\times 1} = \hat{\boldsymbol{Q}}\boldsymbol{Q}^{-1}\boldsymbol{u}$$

with $\hat{Q}_{m,n} = \psi_n(\hat{x}_m)$, $n = 1, \ldots, N$, and $m = 1, \ldots, M$. This defines the transformation array $\hat{\boldsymbol{Q}}\boldsymbol{Q}^{-1}$ that takes the function $u(x)$ defined by the coarse-discretization grid ($\boldsymbol{x}$) values of u_n and exactly interpolates the function to the fine grid. Similarly, we can generate the non-square discretization array necessary for computing the residual by

$$\frac{1}{x}\frac{\mathrm{d}}{\mathrm{d}x}x\frac{\mathrm{d}\hat{\boldsymbol{u}}^{M\times 1}}{\mathrm{d}x} = \hat{\boldsymbol{B}}^{M\times N}\boldsymbol{u}^{N\times 1} \text{ with } \hat{\boldsymbol{B}} = \left[\frac{1}{x}\frac{\mathrm{d}}{\mathrm{d}x}x\frac{\mathrm{d}\hat{\boldsymbol{Q}}}{\mathrm{d}x}\right]\boldsymbol{Q}^{-1}$$

Therefore, the residual on the fine grid can be computed directly from values of u_n on the coarse grid by

$$\hat{R}_m = \left[\hat{\boldsymbol{Q}}\boldsymbol{Q}^{-1}\right]_m\frac{\mathrm{d}u}{\mathrm{d}t} - \hat{\boldsymbol{B}}_m\boldsymbol{u} + \phi^2\left(\left[\hat{\boldsymbol{Q}}\boldsymbol{Q}^{-1}\right]_m\boldsymbol{u} + 1\right)^2$$

$m = 1, \ldots, M$

or

$$\hat{\boldsymbol{R}} = \hat{\boldsymbol{Q}}\boldsymbol{Q}^{-1}\frac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}t} - \hat{\boldsymbol{B}}u + \phi^2(\hat{\boldsymbol{Q}}\boldsymbol{Q}^{-1}\boldsymbol{u} + \boldsymbol{1})^2 \quad (6)$$

where $[\hat{\boldsymbol{Q}}\boldsymbol{Q}^{-1}]_m$ denotes the m-th row of the product $\hat{\boldsymbol{Q}}\boldsymbol{Q}^{-1}$.

3.2. The residual function and norm

A subtle but important point is that the residual function does not necessarily vanish at the catalyst pellet outer boundary even if the trial functions representing the solution vanish at $x = 1$. Therefore, the residual function $R(x,t)$ must be represented in terms of the η_m, not the ψ_n. This gives the truncated trial function expansion $R(x,t) = \boldsymbol{\eta}d$. We stress that this is the exact residual corresponding to the truncated trial function expansion $u^N(x,t)$ solution for this problem and similar, nonlinear problems. By defining the discrete transformation array $\hat{\boldsymbol{P}}$ with $\hat{P}_{i,j} = \eta_j(\hat{x}_i)$ with $i, j = 1, \ldots, M$, the discretized residual function can be written as $\hat{\boldsymbol{R}} = \hat{\boldsymbol{P}}d$.

The residual function decomposition mode amplitudes are computed from $\boldsymbol{d} = \hat{\boldsymbol{P}}^{-1}\hat{\boldsymbol{R}}$. Plotting these coefficients $d_m, m = 1, \ldots, M$ as a function of m (Fig. 1), we see that the residual of the steady-state solution contains components corresponding to mode numbers less than $N + 1$. This demonstrates one difference between solutions obtained by the collocation and Galerkin methods: that the collocation solution residual is not truly orthogonal to the η_n for $n = 1, \ldots, N$ with respect to weight function $x(1 - x^2)$. However, because the maximum magnitude mode

Table 1
Collocation constants

N	x	$\mathbf{B}$	$\sum_j B_{nj}$
1	[0.5774]	[−6]	[−6]
2	$\begin{bmatrix}0.3938\\0.8031\end{bmatrix}$	$\begin{bmatrix}-9.9024 & 12.2997\\9.0337 & -32.7643\end{bmatrix}$	$\begin{bmatrix}2.3973\\-23.7306\end{bmatrix}$
3	$\begin{bmatrix}0.2976\\0.6399\\0.8875\end{bmatrix}$	$\begin{bmatrix}-15.8814 & 19.6364 & -5.2812\\11.1519 & -34.4974 & 29.2357\\-3.5406 & 34.5121 & -99.6212\end{bmatrix}$	$\begin{bmatrix}-1.5262\\5.8902\\-68.6496\end{bmatrix}$

corresponds to η_{N+1} (in Fig. 1 the peak is located at $m = 5$), this discretization procedure constitutes a good approximation to the Galerkin method. We also note that the residual contains no mode components with mode number greater than $2N + 1$ (i.e., polynomials of degree greater than $4N$), which is consistent with our previous analysis of the residual generated by this type of nonlinearity.

The square of the residual norm is computed by the inner (dot) product $||R(x,t)||^2 = \hat{\boldsymbol{w}} \cdot \hat{\boldsymbol{R}}^2$. The quadrature weights used to compute the residual norm are found from

$$\hat{\boldsymbol{w}}^{M\times 1} = \left[\hat{\boldsymbol{P}}^{-1}\right]^{\mathrm{T}} \int_{x=0}^{1} \boldsymbol{\eta}^{\mathrm{T}} x\,\mathrm{d}x \text{ or } \left[\hat{\boldsymbol{P}}^{-1}\right]^{\mathrm{T}} \int_{x=0}^{1} \boldsymbol{\psi}^{\mathrm{T}} x\,\mathrm{d}x$$

depending on which norm weight function is desired[1]. The choice of collocation point locations guarantees that the computed value of the residual norm is exact if the residual function is exact. Plotting the residual norm as a function of truncation number N shows that both solutions converge (Fig. 1); since both solutions also satisfy the problem boundary conditions, both are true solutions to the mathematical model of the problem. Only one, however, is a true, physical solution.

4. Solution behavior as a function of ϕ and N

To understand the role the spurious solution plays in the difficulties associated with computing the true solution at large ϕ values and a fixed number of discretization points N, consider steady-state solutions to the discretized problem Eq. (5). When ϕ is large, the solution must be approximately $\boldsymbol{u} \approx -\mathbf{1}$, otherwise the product $\phi^2(u_n + 1)^2$ would result in numbers too large to be balanced by the remaining terms of the discretized equation. Therefore, by using the original reactant concentration variable c in discretized form, $c = \boldsymbol{u} + \mathbf{1}$, if ϕ is large and if each term $|c_n| \ll 1$, we can compute an approximate solution as

$$c_n = \pm\frac{1}{\phi}\sqrt{-\sum_{j=1}^{N} B_{n,j}} \tag{7}$$

[1]The quadrature weights computed by the first method are used to compute the residual norm and the second for computing mean concentration values.

for the discretized equations. The conclusion of this perturbation analysis is that if $\sum_{j=1}^{N} B_{n,j} > 0$ for any value of n, no solutions can exist for large ϕ and fixed N. On the other hand, solutions will exist if the sums of the row elements of $\boldsymbol{B}$ are all negative numbers. For example, consider Table 1: solutions will exist for large ϕ for the one-point collocation solution, but will not for $N = 2$ and $N = 3$. In fact, numerical tests show that there are no values of $N \leq 20$ (besides $N = 1$) where the sums of the row elements of $\boldsymbol{B}$ are all negative. This property depends on the collocation point locations and trial functions – we can design discretization arrays $\boldsymbol{B}$ so that solutions exist for large ϕ, but have found that the solutions for smaller values of ϕ can be inaccurate.

A predictor–corrector pseudo-arclength continuation technique was used to verify the predictions of the asymptotic solution Eq. (7). Results are presented in Fig. 3 where we see that solutions do indeed exist for all values of ϕ in the range of this plot[2] for $N = 1$. As one might expect, the range of ϕ where solutions exist for $N > 1$ increases with N. In fact, it appears that doubling the number of collocation points N more than doubles the value of ϕ where the saddle-node bifurcation takes place. The saddle-node bifurcation takes place when the spurious and true solutions coalesce and subsequently disappear (see Fig. 4 and [8]). At this point, the determinant of the Jacobian array of the linearized system vanishes ($\det(\boldsymbol{J}) = 0$); the nearly singular Jacobian array can contribute to numerical difficulties experienced when solutions are sought in the neighborhood of the bifurcation points. Saddle-node bifurcation point locations were computed directly with Newton–Raphson iterations of the discretized steady-state equations augmented with the $\det(\boldsymbol{J}) = 0$ condition. Results are reported in Fig. 3 which includes plots of the exact residual norms computed using Eq. (6).

5. Discrete Galerkin projection

We can take advantage of the precise residual computations made possible by the multiple-grid technique to

[2]We plot the reactant concentration at $x = 0$ in this Figure, resulting in an asymptotic value different from that predicted by (7) which is valid only at the collocation points. We also note that the residual norm becomes unbounded for $\phi \to \infty$ for the one-point collocation-discretized system.

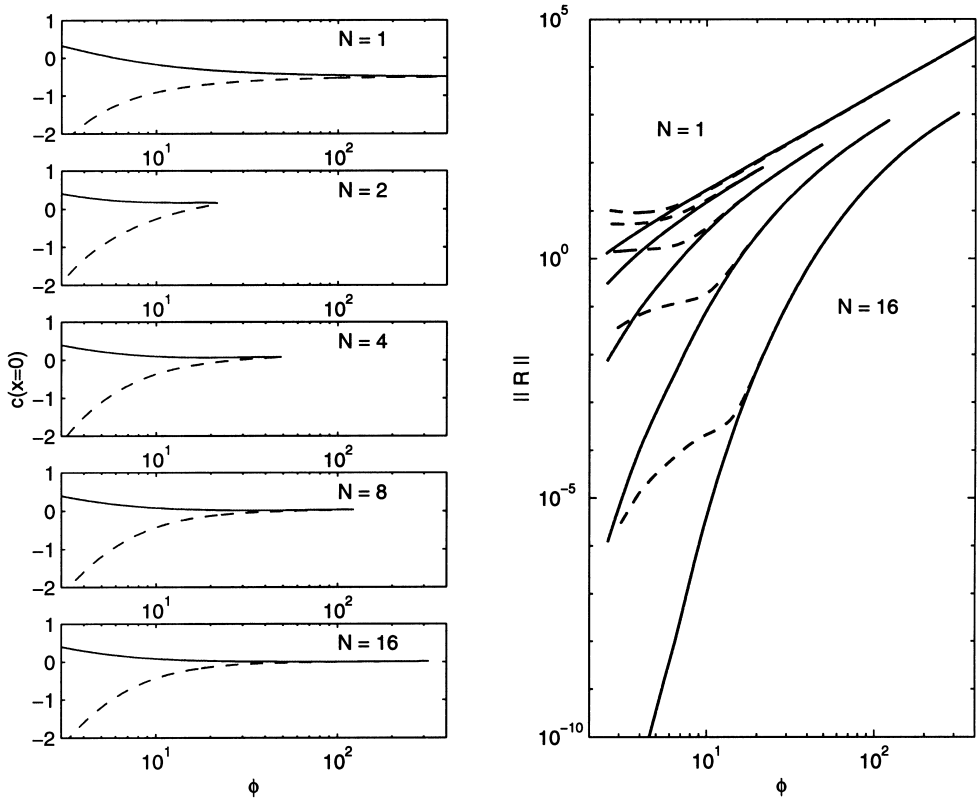

Fig. 3. Continuation results for the collocation discretization method showing the reactant concentration at the catalyst pellet center (left) and both solution residual norms (right) as a function of ϕ. Dashed curves represent spurious solutions and solid curves indicate true solutions. Saddle-node bifurcation point locations are $\phi \to \infty$ for $N = 1$, $\phi = 21.8634$ for $N = 2$, $\phi = 49.5011$ for $N = 4$, $\phi = 122.7844$ for $N = 8$, and $\phi = 318.2965$ for $N = 16$.

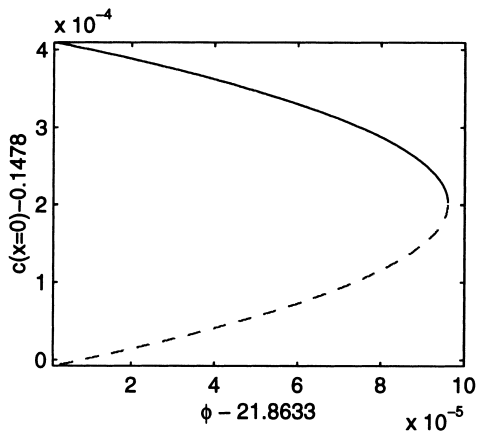

Fig. 4. A magnified view of the solutions in the neighborhood of the saddle-node bifurcation point for the case $N = 2$.

compute values of u_n on the coarse grid that force the residual, weighted by the function $(1 - x^2)x$, to be orthogonal to the first N trial functions η_n. This gives a true discrete analog to the Galerkin projection when the residual can be expressed exactly in terms of a finite number of trial functions η_n (cf. [3]). The motivation for developing this numerical technique was to provide results equivalent to the exact Galerkin projection while retaining the simplicity of implementation of the original orthogonal collocation method. If we define the nonsquare array $\bar{\boldsymbol{P}}^{N\times M}$ by the first N rows of $\hat{\boldsymbol{P}}^{-1}$, i.e.,

$$\bar{P}_{n,m} = [\hat{\boldsymbol{P}}^{-1}]_{n,m} \; n = 1, \ldots, N, \; m = 1, \ldots, M,$$

the first $n = 1, \ldots, N$ mode amplitude coefficients d_n of the collocation-discretized solution residual can be computed from

$$\boldsymbol{d} = \bar{\boldsymbol{P}}\left[\hat{\boldsymbol{Q}}Q^{-1}\frac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}t} - \hat{\boldsymbol{B}}u + \phi^2(\hat{\boldsymbol{Q}}Q^{-1}\boldsymbol{u} + \boldsymbol{1})^2\right] \tag{8}$$

Setting these coefficients to zero forces the residual function to be orthogonal (in terms of inner product Eq. (3)) to the first $n = 1, \ldots, N$ trial functions η_n; rearranging Eq. (8) gives the N ordinary differential equations in time

$$\frac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}t} = \left[\bar{\boldsymbol{P}}\hat{\boldsymbol{Q}}Q^{-1}\right]^{-1}\bar{\boldsymbol{P}}\left[\hat{\boldsymbol{B}}u - \phi^2\left(\hat{\boldsymbol{Q}}Q^{-1}\boldsymbol{u} + \boldsymbol{1}\right)^2\right]$$

to find the values of u_n on the coarse scale. Representative solutions and corresponding residual functions of the true and spurious steady-state solutions obtained with the Galerkin method are shown in Fig. 5. We can observe all of the characteristics of a true Galerkin procedure in Fig. 5: the residual is orthogonal to the first N trial functions since $d_m = 0$, $m = 1, \ldots, N$; the residual mode with greatest magnitude is $d_{N+1}\eta_{N+1}$; and the residual contains no modal

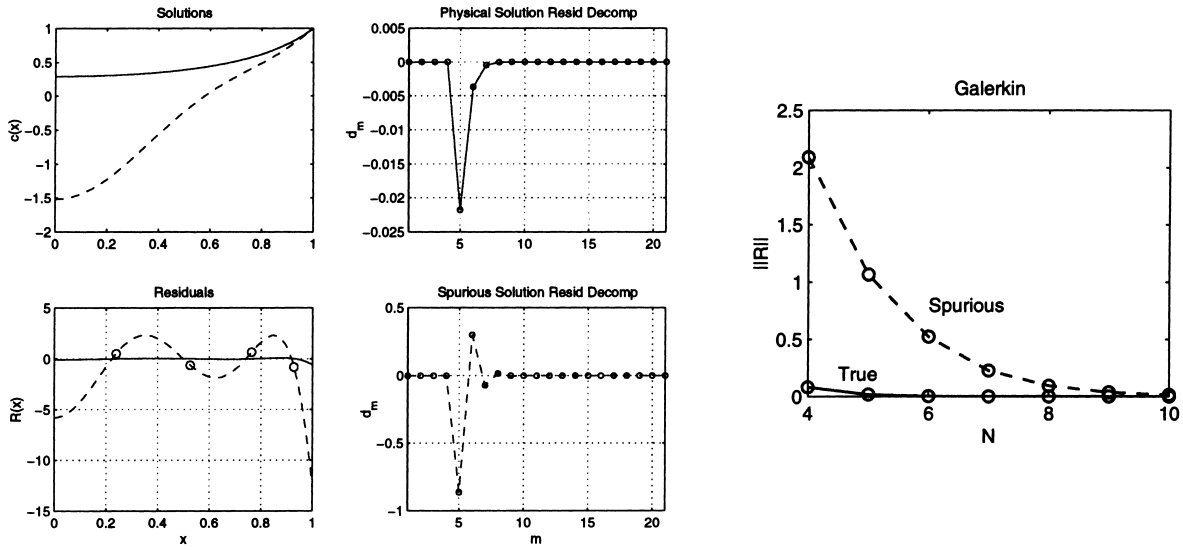

Fig. 5. Steady-state solutions and convergence behavior produced by the Galerkin discretization technique for $\phi = 4$ and $N = 4$. Spurious solution is denoted with the dashed curve. Note that the residual does not necessarily vanish at the collocation point locations (marked in the residual function plot).

contribution from modes $m \geq 2N + 1$, which is an observation consistent with the assumptions upon which the residual calculation techniques are based (see Section 3.1). We note that because the solution procedure uses the values of the residual evaluated on the fine grid, the accuracy assessment of a solution essentially comes at no additional computational cost when using this numerical technique.

Comparing the Galerkin results to collocation in terms of both convergence behavior and the range of ϕ where solutions exist, we find the differences to be minor for this problem, giving further evidence to support the conclusion that the spurious solution is not an artifact of the discretization procedure. The Galerkin procedure appears to give slightly more accurate results for larger values of N; one interesting difference, however, is that the single mode Galerkin-discretized problem has no solutions for $\phi > 7.03$.

6. Dynamic behavior

Linearizing both the Galerkin and collocation-discretized systems at each steady-state solution and computing the eigenvalues and associated eigenfunctions, we find that the physically relevant solution is a stable node and that the spurious solution has saddle-type stability, with one unstable mode. Therefore, the spurious solution will, for all practical purposes, never be directly observed during dynamic simulations when using the discretized equations. However, because the saddle-type solution's stable manifold W^s defines the boundary separating the basin of attraction of the physically relevant solution from the initial conditions that diverge to infinity, the spurious solution is actually responsible for determining the global stability characteristics of the dynamic simulations (see Fig. 6). In these phase portraits, the saddle-stable manifold W^s is approximated by the curves obtained from integrating the discretized equations in reverse time, starting from initial conditions located near and to each side of the spurious solution fixed point along the saddle-stable eigenvector [9].

For $N = 2$, the spurious and physically relevant solutions meet during a saddle-node bifurcation at $\phi = \phi_{sn} = 21.8634$. In the phase portraits shown in Fig. 6, we see that the solutions approach each other as ϕ is increased from $\phi = 4$ to $\phi = 14$. This results in a growing portion of phase space consisting of those initial conditions that diverge to negative infinity, a region that increases in size until it suddenly fills the entire phase space after the saddle-node bifurcation. The result is that all N-mode simulations diverge for the values of $\phi > \phi_{sn}$ corresponding to each N, regardless of the initial conditions. This explains the difficulties associated with numerical simulations of transient behavior for large ϕ; the continuation results presented in Fig. 3 demonstrate the relationship between the number of collocation points used and the values of ϕ after which divergence of the dynamic simulations is guaranteed. In Fig. 7, we see the time-history of the concentration values at the collocation points for a value of $\phi > \phi_{sn}$; we see the potentially deceptive behavior exhibited by the simulation (which would occur when using any integrator): the slow drift in the neighborhood where the steady-state solution existed for $\phi < \phi_{sn}$, followed by the quick divergence to infinity.

7. Conclusions

In this paper, we developed a numerical technique that allows exact computation of the error generated by collocation discretization techniques applied to a nonlinear boundary-value problem. These capabilities were used to analyze

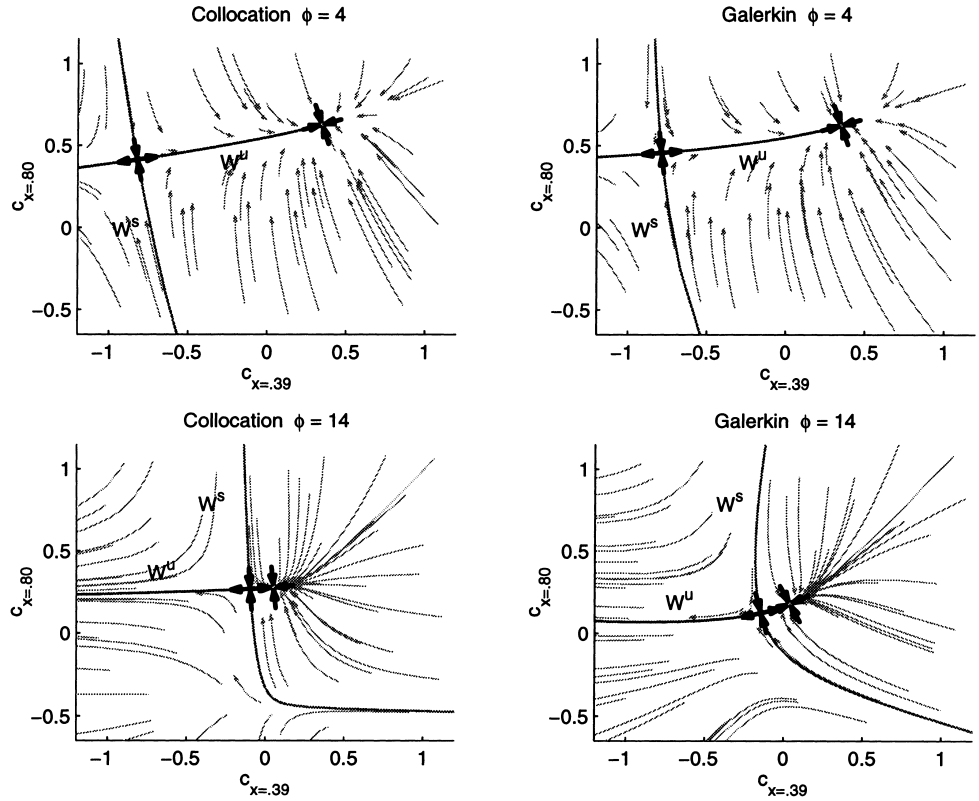

Fig. 6. Phase planes for the collocation- and Galerkin-discretized problem solutions with $N = 2$ illustrating how the true (right-most) and spurious (left-most) solutions approach each other with increasing ϕ. The axes denote concentration values at the two collocation points. The basin of attraction of the true solution and the set of initial conditions that lead to divergence of the dynamic solution are separated by the saddle-stable manifold W^S of the spurious solution. Thicker arrows denote eigenvectors associated with the linearized solutions – the arrow lengths indicate the eigenvalue magnitudes.

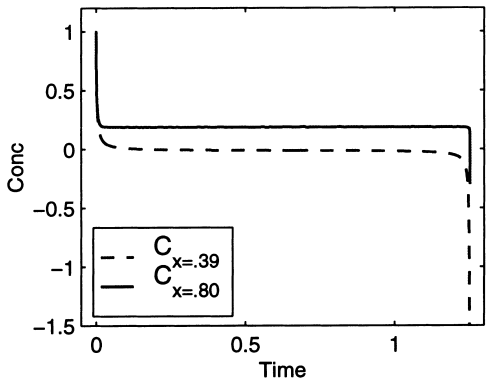

Fig. 7. Simulation results illustrating the divergence of initial conditions corresponding to the case $N = 2$ and $\phi = 22$. The simulation diverges regardless of the integration method used.

a nonlinear reaction-diffusion problem known to exhibit convergence problems. Numerical bifurcation analysis techniques were used to study the role a second, physically meaningless solution played in both the dynamic and steady-state solution behavior. An interesting conclusion of this study is that while the second solution has no physical meaning, it plays a central role in the solution dynamics and convergence of methods designed to determine the true, steady-state solution, and so cannot be dismissed in this context. The precise residual calculations made possible by the multiple–grid method were used to develop a discrete form of the Galerkin projection for a class of nonlinear systems, a goal much sought after in previous collocation studies (see, e.g., [3]). We are currently studying the applicability of this method to implementing other, more advanced weighted residual methods, such as the least squares projection and even nonlinear Galerkin methods [10–12].

In this paper, we consider a problem with a quadratic nonlinearity. Increasing the number of fine-grid discretization points allows analysis of higher-degree nonlinearities; for example, a boundary value problem with a polynomial nonlinearity of degree q would generally require $qN + 1$ fine-scale discretization points to exactly evaluate the residual corresponding to a trial function expansion composed of polynomials with maximum degree N. Numerical techniques are available for accurately performing computations based on high-degree ($M > 500$) polynomial discretizations [13]. Furthermore, preliminary numerical studies of systems with exponential and other 'harder' nonlinearities indicate that for these problems, where the residual still cannot be exactly determined, the new method more closely approximates the Galerkin procedure than the original, orthogonal collocation method. We believe the numerical methods

developed in this paper are a first step towards developing more powerful tools applicable to the analysis of convergence and simulation difficulties for a wide range of more detailed modeling and simulation problems.

Acknowledgements

This work was supported by the National Science Foundation through grant NSF EEC 95-27 576 and the Petroleum Research Fund of the American Chemical Society through grant 31391-G9.

References

[1] J.V. Villadsen, W.E. Stewart, Solution of boundary-value problems by orthogonal collocation, Chem. Eng. Sci. 22 (1967) 1483–1501.
[2] B.A., Finlayson, Nonlinear Analysis in Chemical Engineering, Chem. Eng. Sci. Series, McGraw-Hill, New York, 1980.
[3] M.L. Michelsen, J. Villadsen, Polynomial solution of differential equations. R.S.H. Mah, W.D. Seider (Eds.), Foundations of Computer-Aided Chemical Process Design, Vol. I, 1981, pp. 341–368.
[4] R.G. Rice, D.D. Do, Applied Mathematics and Modeling for Chemical Engineers, Chem. Eng. Series, Wiley, New York, 1995.
[5] J. Villadsen, M.L. Michelsen, Solution of Differential Equation Models by Polynomial Approximation, Int. Series in Phys. and Chem. Eng. Sci, Prentice-Hall, Englewood Cliffs, NJ, 1978.
[6] K.S. Denison, C.E. Hamrin Jr, G. Fairweather, Solution of boundary value problems using software packages DD04AD and COLSYS, Chem. Eng. Commun. 22 (1983) 1–9.
[7] J.P. Sorensen, E.W. Guertin, W.E. Stewart, Computational models for cylindrical catalyst particles, AIChE J. 19 (1973) 969–975.
[8] J. Guckenheimer, P. Holmes, Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields, Appl. Math. Sci. Series, Springer, New York, 42, 1983.
[9] I.G. Kevrekidis, M.S. Jolly, On the use of interactive graphics in the numerical study of chemical dynamics. In: Paper No. 22c, 1987 AIChE Annual Meeting, New York, 1987.
[10] A.E. Deane, I.G. Kevrekidis, G.E. Karniadakis, S.A. Orszag, Low-dimensional models for complex geometry flows: Application to grooved channels and circular cylinders, Phys. Fluids A 10 (1991) 2237–2354.
[11] L. Dettori, D. Gottlieb, R. Temam, A nonlinear Galerkin method: The two-level Fourier collocation case, J. Sci. Comp. 10 (1995) 371–389.
[12] M.D. Graham, I.G. Kevrekidis, Alternative approaches to the Karhunen–Loéve decomposition for model reduction and data analysis, Comp. Chem. Eng. 20 (1996) 495–506.
[13] Y.-h. Lin, H.-Y. Chang, R.A. Adomaitis, MWRtools: A library for weighted residual method calculations. ISR TR (1998) 24–98.