Commun. Comput. Phys.
doi: 10.4208/cicp.290709.120210a

Vol. **8**, No. 4, pp. 845-865
October 2010

Adaptivity and A Posteriori Error Control for Bifurcation Problems I: The Bratu Problem

K. Andrew Cliffe[1], Edward J. C. Hall[1], Paul Houston[1,*], Eric T. Phipps[2] and Andrew G. Salinger[2]

[1] *School of Mathematical Sciences, University of Nottingham, University Park, Nottingham NG7 2RD, UK.*
[2] *Computer Science Research Institute, Sandia National Laboratories, Albuquerque, New Mexico, USA.*

Received 29 July 2009; Accepted (in revised version) 12 February 2010

Available online 17 May 2010

Abstract. This article is concerned with the numerical detection of bifurcation points of nonlinear partial differential equations as some parameter of interest is varied. In particular, we study in detail the numerical approximation of the Bratu problem, based on exploiting the symmetric version of the interior penalty discontinuous Galerkin finite element method. A framework for *a posteriori* control of the discretization error in the computed critical parameter value is developed based upon the application of the dual weighted residual (DWR) approach. Numerical experiments are presented to highlight the practical performance of the proposed *a posteriori* error estimator.

AMS subject classifications: 37G10, 37M20, 65N12, 65N15, 65N30

Key words: Bifurcation theory, Bratu problem, *a posteriori* error estimation, adaptivity, discontinuous Galerkin methods.

1 Introduction

Understanding the nature of solutions to nonlinear partial differential equations (PDEs) remains one of the greatest challenges in modern scientific computing. Some fundamental questions include: "How many solutions exist as some parameter of interest is varied?"; "Are the solutions linearly stable?"; and "At what critical parameter value does a bifurcation occur?". In this article we consider the latter question and in particular address the issue concerning the accuracy of the computed critical value by means of

*Corresponding author. *Email addresses:* Andrew.Cliffe@nottingham.ac.uk (K. A. Cliffe), Edward.Hall@nottingham.ac.uk (E. J. C. Hall), Paul.Houston@nottingham.ac.uk (P. Houston), etphipp@sandia.gov (E. T. Phipps), agsalin@sandia.gov (A. G. Salinger)

a posteriori error estimation. For this purpose we investigate the Bratu problem, see, for example, Wazwaz [35], which can be viewed as a model of some phenomenon exhibiting diffusion with exponential growth. Although the Bratu problem is essentially of academic interest, it serves as an excellent model situation in which to demonstrate the computational approach developed in this article, as it contains many of the key features inherent in the study of more general nonlinear PDEs of practical interest.

In the numerical study of nonlinear PDEs, the estimation of the critical parameter at which a bifurcation may occur can be performed by discretizing a suitable extended system of PDEs; see, for example, Seydel et al. [30, 31] and Moore and Spence [28]. In essence, this process involves determining the parameter value and associated solution at which the Jacobian of the underlying nonlinear PDE has a zero eigenvalue. For the discretization of the extended system we propose to exploit the symmetric version of the interior penalty discontinuous Galerkin (DG) method; see, for example, [2] where a unified analysis of a number of DG methods is presented. Our use of a DG method is primarily due to the benefits in terms of the ease of implementation of automatic mesh adaptation procedures.

Over the past few decades, tremendous progress has been made in the area of *a posteriori* error estimation and adaptive finite element approximation of partial differential equations; for a review of some of the main developments in the subject we refer to the recent monographs [1, 32, 34], and the articles [5, 15]. Despite a number of significant advances in the field, much of the research to date has focused on source problems. In the context of eigenvalue error estimation for determining whether a solution to a PDE is linearly stable or not, we mention the recent articles [13,14,25,29] for the finite element approximation of second-order self-adjoint elliptic eigenvalue problems. For related work, based on considering the eigenvalue problem as a parameter-dependent nonlinear equation, see Verfürth [33, 34], for example, while convergent adaptive algorithms for eigenvalue problems have been analysed in [7,17]. More recently, in the article [11], we considered the *a posteriori* estimation of the error in the leading eigenvalue for the hydrodynamic stability problem. In particular, we employed a dual weighted residual (DWR) *a posteriori* error estimator, see [4, 16, 21], for example, specifically tailored to assess the accuracy of the computed leading eigenvalue. Here, the discretization error stemming from both the numerical approximation of the steady incompressible Navier-Stokes equations, as well as the error arising from the approximation of the corresponding eigenvalue problem itself was controlled. The purpose of this article is to consider the natural extension of these ideas to bifurcation problems. More precisely, we derive computable *a posteriori* bounds on the error in the DG approximation of the critical parameter value for the Bratu problem, based on exploiting the general DWR methodology. Additionally, we extend the ideas presented in Moore & Spence [28] to develop an efficient solution algorithm for both the underlying primal and dual problems to the DG setting.

Rigorous proofs of the existence of bifurcation points in continuous systems, such as the Bratu problem in 2 or more dimensions, are extremely difficult. Indeed one of the primary motivations for developing numerical methods for such problems is that they

are beyond the reach of current analytical approaches. What is usually done theoretically is to assume that a particular bifurcation point exists in the continuous problem and then show that the discretised problem has the same type of bifurcation for a nearby parameter value, which converges to the exact value as the mesh is refined. To our knowledge, this kind of analysis has not yet been carried out for DG methods applied to bifurcation problems and is the topic of reference [10]. However, what one would really like is a theorem that says if a bifurcation is present in the discretised problem and certain additional (and computable) conditions are satisfied then the continuous problems also has such a bifurcation. Such results are not available at present and are likely to be very difficult to obtain. Current practice in this area is to compute bifurcations for the discretised problem and then infer their presence in the continuous problem. In applying a posteriori error estimation to bifurcation problems, for the first time, we are taking a significant step towards the goal of establishing rigorously the existence of bifurcation points for the continuous problem, but we fully acknowledge that much still remains to be done in terms of rigorous analysis.

The article is structured as follows. In the next section we discuss the calculation of simple fold points, specifically we shall be interested in quadratic fold points. In Section 3 we then recall the DWR error estimation technique applied to a general Galerkin finite element method and propose its application for the control of the error in the computed critical parameter. Computation of critical parameters involves the solution of an extended system for the base solution, null-function and the critical parameter; similarly, the error estimation involves the computation of an associated dual solution which satisfies a corresponding adjoint extended system. In Section 4 we therefore discuss how the solution of these extended systems may be computed in an efficient manner. The Bratu problem and its DG discretization are presented in Section 5 and an error representation formula for the error in the computed critical parameter is developed. Numerical experiments for the Bratu problem in one- and two-dimensions are then carried out in Section 6 before we draw some conclusions in Section 7.

2 Calculation of simple fold points

Following the discussion presented in [12], we consider a nonlinear problem of the form

$$F(u,\lambda)=0, \tag{2.1}$$

where F is a map from $V\times\mathbb{R}\to V$, for some Banach space V, with norm $\|\cdot\|$. For the purpose of this article, we shall primarily be concerned with the case when F is a partial differential operator defined over a given computational domain Ω, subject to appropriate boundary/initial conditions. We assume that F is smooth, that is

$$F:V\times\mathbb{R}\to V,$$

is a C^3 mapping. In applications, it is often of interest to compute paths or branches of solutions of (2.1), where λ is some distinguished parameter, e.g., the flow rate or Reynolds

number, and u is a state variable, e.g., the temperature or velocity field. We denote the Fréchet derivative of F with respect to u at a fixed point $(w,\chi)\in V\times\mathbb{R}$, by $F'_u(w,\chi;\cdot)$ and similarly the derivative with respect to λ by $F'_\lambda(w,\chi)$. Here and throughout the paper, we use the convention that in semi-linear forms such as $F'_u(\cdot,\cdot;\cdot)$, the form is linear with respect to all arguments to the right of the semicolon. We will assume that

$$F'_u(u,\lambda;\cdot):V\to V,$$

is Fredholm of index 0 for all $(u,\lambda)\in V\times\mathbb{R}$. For convenience, at a given point (u^0,λ^0), we define

$$F^0=F(u^0,\lambda^0),\quad F^0_u(\cdot)=F'_u(u^0,\lambda^0;\cdot),\quad\text{and}\quad F^0_\lambda=F'_\lambda(u^0,\lambda^0).$$

Higher order Fréchet derivatives are expressed in much the same way, for example, the Fréchet derivative of $F'_u(w,\chi;\cdot)$ with respect to u at a fixed point v is denoted by $F''_{uu}(w,\chi;\cdot,v)$, and similarly, at a given point (u^0,ϕ^0,λ^0), we define

$$F^0_{uu}\phi^0(\cdot)=F''_{uu}(u^0,\lambda^0;\cdot,\phi^0),\quad\text{and}\quad F^0_{u\lambda}\phi^0=F''_{u\lambda}(u^0,\lambda^0;\phi^0).$$

We define the set S by

$$S:=\Big\{(u,\lambda)\in V\times\mathbb{R}:F(u,\lambda)=0\Big\}.$$

If $(u^0,\lambda^0)\in S$ with F^0_u an isomorphism on V, then the Implicit Function Theorem (IFT) ensures the existence of a unique smooth path of solutions $u(\lambda)\in C^3$, satisfying

$$F\big(u(\lambda),\lambda\big)=0,$$

for λ in a neighbourhood of λ^0, with $F'_u(u,\lambda;\cdot)$ an isomorphism. In this article we consider only the case of simple singular points, i.e., where (u^0,λ^0) satisfies

$$F^0=0,\quad\text{and}\quad\dim\ker(F^0_u)=1. \tag{2.2}$$

Furthermore, these singular points will be *quadratic* fold points and thus, denoting by V' the dual space of V and by $\langle\cdot,\cdot\rangle$ the duality pairing between the spaces V and V', the additional side constraints

$$\langle F^0_\lambda,\psi^0\rangle\neq0,\quad\text{and}\quad\langle F^0_{uu}\phi^0(\phi^0),\psi^0\rangle\neq0, \tag{2.3}$$

will also hold, for any $\psi^0\in\ker((F^0_u)')$. For notational simplicity, in the sequel we suppress the superscript "0", when it is clear from the context that the solution under consideration is indeed a singular (quadratic fold) point of (2.1). With this in mind, to determine the quadratic fold point of (2.1), we seek to compute a solution of the following extended system: find

$$\mathbf{u}=(u,\phi,\lambda)\in\mathbf{V}=V\times V\times\mathbb{R},$$

such that

$$T(\mathbf{u}) \equiv \begin{pmatrix} F(u,\lambda) \\ F'_u(u,\lambda;\phi) \\ \langle \phi, c\rangle - 1 \end{pmatrix} = 0, \tag{2.4}$$

where $c \in V'$ is some chosen functional, which satisfies $\langle \phi, c\rangle \neq 0$, for $\phi \in \ker(F'_u(u,\lambda;\cdot))$, see [28, 30, 31]. The equation $\langle \phi, c\rangle - 1 = 0$ acts to normalise the nullfunction ϕ, thus ensuring that, if a solution to (2.4) exists at some λ, the solution is unique.

The following lemma will prove useful.

Lemma 2.1 ("ABCD" Lemma)**.** Keller, [24, Lemma 2.8]. *Let V be a Banach space and consider the linear operator $M : V \times \mathbb{R} \to V \times \mathbb{R}$ of the form*

$$M = \begin{pmatrix} A & b \\ \langle \cdot, c\rangle & d \end{pmatrix}, \tag{2.5}$$

where $A : V \to V$, $b \in V \setminus \{0\}$, $c \in V' \setminus \{0\}$, $d \in \mathbb{R}$. Then

1. If A is an isomorphism on V, then M is an isomorphism on $V \times \mathbb{R}$ if and only if

$$d - \langle A^{-1}b, c\rangle \neq 0.$$

2. If $\dim\ker(A) = \mathrm{codim}\,\mathrm{Range}(A) = 1$*, then M is an isomorphism if and only if*

$$\begin{aligned} &(a) \quad \langle b, \psi\rangle \neq 0, \qquad \forall \psi \in \ker(A') \setminus \{0\}, \\ &(b) \quad \langle \phi, c\rangle \neq 0, \qquad \forall \phi \in \ker(A) \setminus \{0\}. \end{aligned}$$

3. If $\dim\ker(A) \geq 2$*, then M is singular.*

3 *A posteriori* error estimation

In this section we develop a general theoretical framework for the derivation of computable *a posteriori* estimates for the error in the computed bifurcation point when the extended system (2.4) is numerically approximated by a general Galerkin finite element method. To this end, we exploit the duality-based *a posteriori* error estimation techniques developed by C. Johnson and R. Rannacher and their collaborators. For a detailed discussion, we refer to the series of articles [5, 15, 23, 26], and the references cited therein.

We begin by first introducing a suitable finite element approximation of the bifurcation problem (2.4). To this end, we consider a sequence of finite element spaces $S_{h,p}$ consisting of piecewise polynomial functions of degree p on a partition $\mathcal{T}_h$, of granularity h. The Galerkin finite element approximation consists of finding the triple

$$\mathbf{u}_h = (u_h, \phi_h, \lambda_h) \in \mathbf{S}_{h,p} = S_{h,p} \times S_{h,p} \times \mathbb{R},$$

such that

$$\mathcal{N}(\mathbf{u}_h; \mathbf{v}_h) \equiv \hat{\mathcal{N}}(u_h, \lambda_h; v_h) + \hat{\mathcal{N}}'_u(u_h, \lambda_h; \phi_h, \varphi_h) + \chi_h\big((c, \phi_h) - 1\big) = 0, \quad \forall \mathbf{v}_h \in \mathbf{S}_{h,p}, \tag{3.1}$$

where $\mathbf{v}_h=(v_h,\varphi_h,\chi_h)$, $\hat{\mathcal{N}}(\cdot;\cdot)$ is the semi-linear form associated with the discretization of the underlying partial differential equation (2.1) and $\hat{\mathcal{N}}'_u(\cdot,\cdot;\cdot,\cdot)$ is the Jacobian of $\hat{\mathcal{N}}(\cdot,\cdot;\cdot)$ with respect to u. Further, we assume that (u_h,ϕ_h,λ_h) also satisfies the properties of a quadratic fold point, i.e.,

$$\hat{\mathcal{N}}'_\lambda(u_h,\lambda_h;\psi_h)\neq 0, \qquad \hat{\mathcal{N}}''_{uu}(u_h,\lambda_h;\phi_h,\phi_h,\psi_h)\neq 0, \tag{3.2}$$

where $\psi_h\in\ker\big(\hat{\mathcal{N}}'_u(u_h,\lambda_h;\cdot,\varphi_h)\big)$ for all $\varphi_h\in S_{h,p}$. Finally, we also assume that (3.1) is a consistent discretization of (2.4); namely that the analytical solution $\mathbf{u}=(u,\phi,\lambda)$ to (2.4) satisfies

$$\mathcal{N}(\mathbf{u};\mathbf{v}_h)=0, \quad \forall\mathbf{v}_h\in\mathbf{S}_{h,p}, \tag{3.3}$$

and moreover, we assume that, as the mesh is refined, $\mathbf{u}_h$ converges to $\mathbf{u}$ with respect to some appropriate norm. These assumptions are very reasonable; indeed, for a continuous problem exhibiting a simple singular point $(u^0,\lambda^0)\in V\times\mathbb{R}$, Brezzi et al. [6] showed that a numerical discretization of the problem (such as the standard conforming Galerkin finite element method), which approximates functions in V with functions in a finite dimensional subspace $V_h\subset V$, will also possess a simple singular point $(u_h^0,\lambda_h^0)\in V_h\times\mathbb{R}$ in a neighbourhood of (u^0,λ^0), and furthermore $(u_h^0,\lambda_h^0)\to(u^0,\lambda^0)$, as $h\to 0$. For analogous results for problems with simple singular points in the context of discontinuous Galerkin methods (where $V_h\not\subset V$) we refer the reader to [10].

Remark 3.1. We remark that, in a slight variation to the standard approach of the location of critical parameters, we have recast the equation $(c,\phi_h)-1=0$ in the weak form $\chi_h\big((c,\phi_h)-1\big)=0$, for all $\chi_h\in\mathbb{R}$. As $\mathbb{R}=\text{span}\{1\}$, this has no effect when calculating the approximate critical parameter, but this formulation is required for the error estimation which follows.

3.1 DWR approach for functionals

For a linear target functional of practical interest $J:\mathbf{V}\to\mathbb{R}$, we briefly outline the key steps involved in estimating the approximation error $J(\mathbf{u})-J(\mathbf{u}_h)$ employing the DWR technique. We write $\mathcal{M}(\cdot,\cdot;\cdot,\cdot)$ to denote the mean value linearization of $\mathcal{N}(\cdot;\cdot)$, defined by

$$\begin{aligned}\mathcal{M}(\mathbf{u},\mathbf{u}_h;\mathbf{u}-\mathbf{u}_h,\mathbf{w})&=\mathcal{N}(\mathbf{u};\mathbf{w})-\mathcal{N}(\mathbf{u}_h;\mathbf{w})\\&=\int_0^1\mathcal{N}'_{\mathbf{u}}\big(\theta\mathbf{u}+(1-\theta)\mathbf{u}_h;\mathbf{u}-\mathbf{u}_h,\mathbf{w}\big)\,\mathrm{d}\theta,\end{aligned} \tag{3.4}$$

for some $\mathbf{w}\in\mathbf{V}$. We now introduce the following (formal) *dual problem*: find $\mathbf{z}\in\mathbf{V}$, such that

$$\mathcal{M}(\mathbf{u},\mathbf{u}_h;\mathbf{w},\mathbf{z})=J(\mathbf{w}), \qquad \forall\mathbf{w}\in\mathbf{V}. \tag{3.5}$$

We assume that (3.5) possesses a unique solution. This assumption is, of course, dependent on both the definition of $\mathcal{M}(\mathbf{u},\mathbf{u}_h;\cdot,\cdot)$ and the target functional under consideration. We point out that well-posedness of the underlying *primal* problem does not automatically imply the well-posedness of the corresponding dual problem. Indeed, this must be verified for the problem at hand; this is a particularly pertinent issue when considering first-order hyperbolic conservation laws, see, for example, the discussion in [20]. For the proceeding error analysis, we assume that (3.5) is well-posed. By using the linearity of $J(\cdot)$, combining (3.4) and (3.5) and using the consistency condition (3.3) we arrive at the following error representation formula

$$\begin{aligned} J(\mathbf{u})-J(\mathbf{u}_h) &= J(\mathbf{u}-\mathbf{u}_h) = \mathcal{M}(\mathbf{u},\mathbf{u}_h;\mathbf{u}-\mathbf{u}_h,\mathbf{z}) \\ &= \mathcal{M}(\mathbf{u},\mathbf{u}_h;\mathbf{u}-\mathbf{u}_h,\mathbf{z}-\mathbf{z}_h) \\ &= -\mathcal{N}(\mathbf{u}_h,\mathbf{z}-\mathbf{z}_h), \qquad \forall \mathbf{z}_h \in \mathbf{S}_{h,p}. \end{aligned} \tag{3.6}$$

As it stands, the error representation formula (3.6) is still non-computable, since $\mathbf{z}$ is unknown. Instead, we must seek a finite dimensional approximation $\hat{\mathbf{z}}_h$ to $\mathbf{z}$. Unfortunately it is not possible to seek $\hat{\mathbf{z}}_h \in \mathbf{S}_{h,p}$, otherwise the resulting error representation would be identically zero due to (3.1). A number of possible alternatives exist. The first involves keeping the degree p of the approximating polynomial the same as that for $\mathbf{u}_h$, but computing $\hat{\mathbf{z}}_h$ on a sequence of dual finite element meshes $\hat{\mathcal{T}}_{\hat{h}}$ which, in general, differ from the "primal meshes" $\mathcal{T}_h$. Alternatively $\hat{\mathbf{z}}_h \in \mathbf{S}_{h,\hat{p}}$ may be computed using polynomials of degree $\hat{p} > p$ on the same finite element mesh $\mathcal{T}_h$ employed for the primal problem. A variant of this second approach is to compute the approximate dual solution using the same polynomial degree p as used for the primal problem and to extrapolate the resulting approximate dual solution $\hat{\mathbf{z}}_h$. Although this latter approach is the cheapest of the three methods, and is still capable of producing adaptively refined meshes specifically tailored to the selected target functional, the quality of the resulting approximate error representation formula may be poor, cf. [19], for example. On the basis of numerical experimentation ([19]), we favour the second approach due to its computational simplicity of implementation.

In our case, we are interested in controlling the error in the critical bifurcation parameter and hence the target functional of interest is simply $J(\mathbf{u}) = \lambda$. The linearization performed in (3.4) is carried out at a convex combination of $\mathbf{u}$ and $\mathbf{u}_h$, however, as $\mathbf{u}$ is not available we linearize only about our approximate solution $\mathbf{u}_h$ and thereby the integral in (3.4) is redundant. Hence, the (approximate) dual problem, attained by linearizing about $\mathbf{u}_h$, for the estimation of the error in the computed critical parameter is defined by: find $\hat{\mathbf{z}}_h = (z_u, z_\phi, z_\lambda) \in \mathbf{S}_{h,\hat{p}}$, such that

$$\begin{aligned} &\hat{\mathcal{N}}'_u(u_h,\lambda_h;v_h,z_u)+\hat{\mathcal{N}}'_\lambda(u_h,\lambda_h;z_u)\chi_h+\hat{\mathcal{N}}''_{uu}(u_h,\lambda_h;v_h,\phi_h,z_\phi) \\ &\quad +\hat{\mathcal{N}}'_u(u_h,\lambda_h;\varphi_h,z_\phi)+\hat{\mathcal{N}}''_{u\lambda}(u_h,\lambda_h;\phi_h,z_\phi)\chi_h+z_\lambda(c,\varphi_h)=1, \qquad \forall \mathbf{v}_h \in \mathbf{S}_{h,\hat{p}}, \end{aligned} \tag{3.7}$$

where $\mathbf{v}_h = (v_h, \varphi_h, \chi_h)$.

4 Solution procedure

In this section we discuss how to solve the primal and dual problems in an efficient manner by reducing the extended problems to a succession of smaller ones, cf. [28].

4.1 Primal problem

To determine the numerical solution $\mathbf{u}_h$ to the nonlinear system of equations (3.1), we employ a damped Newton method. This nonlinear iteration generates a sequence of approximations $\mathbf{u}_h^n$, $n=1,2,\cdots$, to the actual numerical solution $\mathbf{u}_h$ using the following algorithm. Given an iterate $\mathbf{u}_h^n$, the update $\mathbf{d}_h^n=(du_h^n,d\phi_h^n,d\lambda_h^n)$, for $\mathbf{u}_h^n$ to get to the next iterate

$$\mathbf{u}_h^{n+1}=\mathbf{u}_h^n+\omega^n\mathbf{d}_h^n,\qquad 0<\omega^n\leq 1,$$

is defined by: find $\mathbf{d}_h^n$, such that

$$\hat{\mathcal{N}}_u'(u_h^n,\lambda_h^n;du_h^n,v_h)+\hat{\mathcal{N}}_\lambda'(u_h^n,\lambda_h^n;v_h)d\lambda_h^n=r_1^n(v_h), \tag{4.1a}$$

$$\hat{\mathcal{N}}_{uu}''(u_h^n,\lambda_h^n;du_h^n,\phi_h^n,\varphi_h)+\hat{\mathcal{N}}_u'(u_h^n,\lambda_h^n;d\phi_h^n,\varphi_h)+\hat{\mathcal{N}}_{u\lambda}''(u_h^n,\lambda_h^n;\phi_h^n,\varphi_h)d\lambda_h^n=r_2^n(\varphi_h), \tag{4.1b}$$

$$\chi_h(d\phi_h^n,c)=r_3^n(\chi_h), \tag{4.1c}$$

for all $\mathbf{v}_h=(v_h,\varphi_h,\chi_h)\in\mathbf{S}_{h,p}$. Here, $r_1^n(\cdot)$, $r_2^n(\cdot)$ and $r_3^n(\cdot)$ are residuals given, respectively, by

$$r_1^n(v_h)=-\hat{\mathcal{N}}(u_h^n,\lambda_h^n;v_h),\quad r_2^n(\varphi_h)=-\hat{\mathcal{N}}_u'(u_h^n,\lambda_h^n;\phi_h^n,\varphi_h),\quad r_3^n(\chi_h)=-\chi_h\big((\phi_h^n,c)-1\big).$$

Furthermore, the step length ω^n is automatically selected to ensure that the l_2-norm of the residual of the underlying approximation is reduced at each Newton step. If the finite element space $S_{h,p}$ is of dimension N, then the system defined in (4.1) is of size $2N+1$, which may be extremely large for problems of engineering interest. Instead, we would like to reduce it to a collection of smaller problems, though, we point out that, a block LU-decomposition is not applicable since it will lead to the inversion of $\hat{\mathcal{N}}_u'(u_h^n,\lambda_h^n;\cdot,\cdot)$, which is singular at the bifurcation point. Instead, we follow the proceeding steps: we assume a Galerkin type approximation of $\mathbf{u}_h$, in which case

$$u_h^n=\sum_{i=1}^N U_i^n\varphi_i,\quad \phi_h^n=\sum_{i=1}^N\Phi_i^n\varphi_i,\quad du_h^n=\sum_{i=1}^N dU_i^n\varphi_i,\quad d\phi_h^n=\sum_{i=1}^N d\Phi_i^n\varphi_i,$$

where $\{\varphi_i\}_{i=1}^N$ is the set of linearly independent finite element basis functions which span $S_{h,p}$. We let

$$\boldsymbol{\phi}_h^n=\{\Phi_i\}_{i=1}^N,\quad \boldsymbol{du}_h^n=\{dU_i^n\}_{i=1}^N,\quad \boldsymbol{d\phi}_h^n=\{d\Phi_i^n\}_{i=1}^N,$$

and in an abuse of notation, we can rewrite (4.1) as

$$\begin{bmatrix} \mathrm{F}_u^n & 0 & \boldsymbol{F}_\lambda^n \\ \mathrm{F}_{uu}^n & \mathrm{F}_u^n & \boldsymbol{F}_{u\lambda}^n \\ \mathbf{0}^\top & \boldsymbol{l}^\top & 0 \end{bmatrix}\begin{bmatrix} \boldsymbol{du}_h^n \\ \boldsymbol{d\phi}_h^n \\ d\lambda_h^n \end{bmatrix}=\begin{bmatrix} \boldsymbol{r}_1^n \\ \boldsymbol{r}_2^n \\ r_3^n \end{bmatrix}, \tag{4.2}$$

where the matrices F_u^n and F_{uu}^n are given, respectively, by

$$\{\mathrm{F}_u^n\}_{i,j=1}^N = \{\hat{\mathcal{N}}_u'(u_h^n,\lambda_h^n;\varphi_i,\varphi_j)\}_{i,j=1}^N, \qquad \{\mathrm{F}_{uu}^n\}_{i,j=1}^N = \{\hat{\mathcal{N}}_{uu}''(u_h^n,\lambda_h^n;\varphi_i,\phi_h^n,\varphi_j)\}_{i,j=1}^N,$$

and the vectors $\boldsymbol{F}_\lambda^n$ and $\boldsymbol{F}_{u\lambda}^n$ are given, respectively, by

$$\{\boldsymbol{F}_\lambda^n\}_{i=1}^N = \{\hat{\mathcal{N}}_\lambda'(u_h^n,\lambda_h^n;\varphi_i)\}_{i=1}^N, \qquad \{\boldsymbol{F}_{u\lambda}^n\}_{i=1}^N = \{\hat{\mathcal{N}}_{u\lambda}''(u_h^n,\lambda_h^n;\phi_h^n,\varphi_i)\}_{i=1}^N.$$

Finally, $\boldsymbol{l}$ is the vector given by

$$\{\boldsymbol{l}\}_{i=1}^N = (\varphi_i, c), \qquad \{\boldsymbol{r}_1^n\}_{i=1}^N = r_1^n(\varphi_i),$$

and so on. We introduce the auxiliary variable $\mu = \boldsymbol{l}^\top \boldsymbol{du}_h^n$, and consider the following equation

$$\begin{bmatrix} \mathrm{F}_u^n & \boldsymbol{F}_\lambda^n \\ \boldsymbol{l}^\top & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{du}_h^n \\ d\lambda_h^n \end{bmatrix} = \begin{bmatrix} \boldsymbol{r}_1^n \\ \mu \end{bmatrix} \equiv \begin{bmatrix} \boldsymbol{r}_1^n \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \mu. \tag{4.3}$$

Using Lemma 2.1 and the conditions of a quadratic fold point (2.2) and (2.3), we see that, even at the fold point, the matrix in (4.3) is non-singular. Hence, the following holds

$$\begin{bmatrix} \boldsymbol{du}_h^n \\ d\lambda_h^n \end{bmatrix} = \begin{bmatrix} \boldsymbol{a} \\ \alpha \end{bmatrix} + \begin{bmatrix} \boldsymbol{b} \\ \beta \end{bmatrix} \mu, \tag{4.4}$$

where

$$\begin{bmatrix} \boldsymbol{a} \\ \alpha \end{bmatrix} = \begin{bmatrix} \mathrm{F}_u^n & \boldsymbol{F}_\lambda^n \\ \boldsymbol{l}^\top & 0 \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{r}_1^n \\ 0 \end{bmatrix}, \qquad \begin{bmatrix} \boldsymbol{b} \\ \beta \end{bmatrix} = \begin{bmatrix} \mathrm{F}_u^n & \boldsymbol{F}_\lambda^n \\ \boldsymbol{l}^\top & 0 \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{0} \\ 1 \end{bmatrix}.$$

Using (4.4), the second and third equations of (4.2) can then be rewritten as

$$\begin{aligned} \begin{bmatrix} \mathrm{F}_u^n & \boldsymbol{F}_\lambda^n \\ \boldsymbol{l}^\top & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{d\phi}_h^n \\ \mu \end{bmatrix} &= \begin{bmatrix} \boldsymbol{r}_2^n + \boldsymbol{F}_\lambda^n \mu - \boldsymbol{F}_{u\lambda}^n d\lambda_h^n - \mathrm{F}_{uu}^n \boldsymbol{du}_h^n \\ r_3^n \end{bmatrix} \\ &\equiv \begin{bmatrix} \boldsymbol{r}_2^n - \boldsymbol{F}_{u\lambda}^n \alpha - \mathrm{F}_{uu}^n \boldsymbol{a} \\ r_3^n \end{bmatrix} + \begin{bmatrix} \boldsymbol{F}_\lambda^n - \boldsymbol{F}_{u\lambda}^n \beta - \mathrm{F}_{uu}^n \boldsymbol{b} \\ 0 \end{bmatrix} \mu, \end{aligned}$$

which in turn implies

$$\begin{bmatrix} \boldsymbol{d\phi}_h^n \\ \mu \end{bmatrix} = \begin{bmatrix} \boldsymbol{c} \\ \gamma \end{bmatrix} + \begin{bmatrix} \boldsymbol{d} \\ \delta \end{bmatrix} \mu, \tag{4.5}$$

where

$$\begin{bmatrix} \boldsymbol{c} \\ \gamma \end{bmatrix} = \begin{bmatrix} \mathrm{F}_u^n & \boldsymbol{F}_\lambda^n \\ \boldsymbol{l}^\top & 0 \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{r}_2^n - \boldsymbol{F}_{u\lambda}^n \alpha - \mathrm{F}_{uu}^n \boldsymbol{a} \\ r_3^n \end{bmatrix},$$

and

$$\begin{bmatrix} \boldsymbol{d} \\ \delta \end{bmatrix} = \begin{bmatrix} \mathrm{F}_u^n & \boldsymbol{F}_\lambda^n \\ \boldsymbol{l}^\top & 0 \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{F}_\lambda^n - \boldsymbol{F}_{u\lambda}^n \beta - \mathrm{F}_{uu}^n \boldsymbol{b} \\ 0 \end{bmatrix}.$$

Hence, μ is given in closed form by

$$\mu = \frac{\gamma}{1-\delta},$$

which can then be used in (4.4) and (4.5) to compute $\boldsymbol{du}_h^n$, $d\lambda_h^n$ and $\boldsymbol{d\phi}_h^n$. It then remains to show that $\delta \neq 1$.

Lemma 4.1. *Consider δ as given in* (4.5). *At a quadratic fold bifurcation point (u_h,ϕ_h,λ_h), we have that $\delta \neq 1$.*

Proof. We have that

$$\begin{bmatrix} \mathrm{F}_u & \boldsymbol{F}_\lambda \\ \boldsymbol{l}^\top & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{b} \\ \beta \end{bmatrix} = \begin{bmatrix} \boldsymbol{0} \\ 1 \end{bmatrix}, \tag{4.6}$$

where the superscript "n" have been dropped to indicate evaluation at the bifurcation point. We premultiply the above equation by $(\boldsymbol{\psi}_h^\top,0)$, where $\boldsymbol{\psi}_h \in \ker(\mathrm{F}_u)^\top$ to obtain

$$\beta = 0.$$

Hence, (4.6) becomes

$$\begin{bmatrix} \mathrm{F}_u & \boldsymbol{F}_\lambda \\ \boldsymbol{l}^\top & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{b} \\ 0 \end{bmatrix} = \begin{bmatrix} \boldsymbol{0} \\ 1 \end{bmatrix}, \tag{4.7}$$

from which we deduce that $\boldsymbol{b} = \boldsymbol{\phi}_h$. Furthermore,

$$\begin{bmatrix} \mathrm{F}_u & \boldsymbol{F}_\lambda \\ \boldsymbol{l}^\top & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{d} \\ \delta \end{bmatrix} = \begin{bmatrix} \boldsymbol{F}_\lambda - \mathrm{F}_{uu}\boldsymbol{\phi}_h \\ 0 \end{bmatrix}. \tag{4.8}$$

Premultiplying this equation by $(\boldsymbol{\psi}_h^\top,0)$, then gives

$$\boldsymbol{\psi}_h^\top \boldsymbol{F}_\lambda \delta = \boldsymbol{\psi}_h^\top \boldsymbol{F}_\lambda - \boldsymbol{\psi}_h^\top \mathrm{F}_{uu}\boldsymbol{\phi}_h.$$

Using the side constraints (3.2), we have

$$\boldsymbol{\psi}_h^\top \boldsymbol{F}_\lambda \neq 0, \qquad \boldsymbol{\psi}_h^\top \mathrm{F}_{uu}\boldsymbol{\phi}_h \neq 0,$$

thus we can be sure that δ is well defined and $\delta \neq 1$. □

Remark 4.1. A continuity argument shows that in a neighbourhood of (u_h,ϕ_h,λ_h), Newton's method can be used in the manner proposed above without the matrices present in the inner (linear) iteration becoming singular. The solution of the primal problem thus requires four solves with the same matrix for each Newton iteration.

Remark 4.2. Although we do not do this in our numerical experiments, for simplicity, it is possible to approximate the second derivatives $\mathrm{F}_{uu}^n\boldsymbol{a}$ and $\mathrm{F}_{uu}^n\boldsymbol{b}$ using a directional finite differencing technique. For example,

$$\mathrm{F}_{uu}^n\boldsymbol{a} \approx \frac{\mathrm{F}_u(u_h^n + \epsilon\phi_h^n)\boldsymbol{a} - \mathrm{F}_u^n\boldsymbol{a}}{\epsilon},$$

where $\epsilon = \varepsilon\big(\varepsilon + \|u_h^n\|/\|\phi_h^n\|\big)$, for $\varepsilon = 10^{-6}$, and $\mathrm{F}_u(\cdot)$ is the matrix, such that

$$\big\{\mathrm{F}_u(\cdot)\big\}_{i,j=1}^N = \hat{\mathcal{N}}_u'(\cdot,\lambda_h^n;\varphi_i,\varphi_j).$$

4.2 Dual problem

In this section we outline the numerical procedure employed to compute the solution of the (approximate) dual problem defined in (3.7). To this end, we first write

$$z_u = \sum_{i=1}^{\hat{N}} Z_{u,i}\hat{\varphi}_i, \qquad z_\phi = \sum_{i=1}^{\hat{N}} Z_{\phi,i}\hat{\varphi}_i,$$

where $\{\hat{\varphi}_i\}_{i=1}^{\hat{N}}$ is the set of linearly independent finite element basis functions, which span $\hat{\mathbf{S}}_{h,\hat{p}}$. Defining

$$\boldsymbol{z}_u = \{Z_{u,i}\}_{i=1}^{\hat{N}}, \qquad \boldsymbol{z}_\phi = \{Z_{\phi,i}\}_{i=1}^{\hat{N}},$$

we rewrite the dual problem (3.7) as: find the triple $(\boldsymbol{z}_u, \boldsymbol{z}_\phi, z_\lambda)$, such that

$$\begin{bmatrix} (\hat{\mathrm{F}}_u)^\top & (\hat{\mathrm{F}}_{uu})^\top & \mathbf{0} \\ 0 & (\hat{\mathrm{F}}_u)^\top & \hat{\boldsymbol{l}} \\ (\hat{\boldsymbol{F}}_\lambda)^\top & (\hat{\boldsymbol{F}}_{u\lambda})^\top & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{z}_u \\ \boldsymbol{z}_\phi \\ z_\lambda \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ 1 \end{bmatrix}. \tag{4.9}$$

Here, $\hat{\mathrm{F}}_u$ is the Jacobi matrix defined on the space $S_{h,\hat{p}}$ evaluated at u_h, and so on. In analogy to the solution of the primal problem, we reduce (4.9) to a collection of smaller matrix problems. First, we introduce an auxiliary variable $z_\mu = (\hat{\boldsymbol{F}}_\lambda)^\top \boldsymbol{z}_\phi$ and consider the set of equations

$$\begin{bmatrix} (\hat{\mathrm{F}}_u)^\top & \hat{\boldsymbol{l}} \\ (\hat{\boldsymbol{F}}_\lambda)^\top & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{z}_\phi \\ z_\lambda \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix} z_\mu. \tag{4.10}$$

Once again, Lemma 2.1 can be used to show that the matrix on the left-hand side of (4.10) is non-singular at a quadratic fold point. We point out that in this case, as the dual solution belongs to a finite element space consisting of higher order polynomials than that used for the numerical approximation of the primal solution, $\hat{F}_u$ may not necessarily be singular, though it is expected to be highly ill-conditioned, particularly as the finite element mesh is enriched. Hence, we first write

$$\begin{bmatrix} \boldsymbol{z}_\phi \\ z_\lambda \end{bmatrix} = \begin{bmatrix} \boldsymbol{a}_z \\ \alpha_z \end{bmatrix} z_\mu, \tag{4.11}$$

where

$$\begin{bmatrix} \boldsymbol{a}_z \\ \alpha_z \end{bmatrix} = \begin{bmatrix} (\hat{\mathrm{F}}_u)^\top & \hat{\boldsymbol{l}} \\ (\hat{\boldsymbol{F}}_\lambda)^\top & 0 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Thus, the first and third equations of (4.9) can be rewritten as

$$\begin{bmatrix} (\hat{\mathrm{F}}_u)^\top & \hat{\boldsymbol{l}} \\ (\hat{\boldsymbol{F}}_\lambda)^\top & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{z}_u \\ z_\mu \end{bmatrix} = \begin{bmatrix} z_\mu \hat{\boldsymbol{l}} - (\hat{\mathrm{F}}_{uu})^\top \boldsymbol{z}_\phi \\ 1 - (\hat{\boldsymbol{F}}_{u\lambda})^\top \boldsymbol{z}_\phi \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix} + \begin{bmatrix} \hat{\boldsymbol{l}} - (\hat{\mathrm{F}}_{uu})^\top \boldsymbol{a}_z \\ -(\hat{\boldsymbol{F}}_{u\lambda})^\top \boldsymbol{a}_z \end{bmatrix} z_\mu.$$

Hence,

$$\begin{bmatrix} \boldsymbol{z}_u \\ z_\mu \end{bmatrix} = \begin{bmatrix} \boldsymbol{a}_z \\ \alpha_z \end{bmatrix} + \begin{bmatrix} \boldsymbol{b}_z \\ \beta_z \end{bmatrix} z_\mu, \tag{4.12}$$

where

$$\begin{bmatrix} \boldsymbol{b}_z \\ \beta_z \end{bmatrix} = \begin{bmatrix} (\hat{\mathrm{F}}_u)^\top & \hat{\boldsymbol{l}} \\ (\hat{\boldsymbol{F}}_\lambda)^\top & 0 \end{bmatrix}^{-1} \begin{bmatrix} \hat{\boldsymbol{l}} - (\hat{\mathrm{F}}_{uu})^\top \boldsymbol{a}_z \\ -(\hat{\boldsymbol{F}}_{u\lambda})^\top \boldsymbol{a}_z \end{bmatrix}.$$

Therefore,

$$z_\mu = \frac{\alpha_z}{1-\beta_z},$$

and (4.11) and (4.12) can be used to calculate z_u, z_ϕ and z_λ. We now seek to show that $\beta_z \neq 1$.

Lemma 4.2. *Consider β_z as defined in* (4.12)*, but with $S_{h,\hat{p}} = S_{h,p}$, then $\beta_z \neq 1$.*

Proof. We have

$$\begin{bmatrix} (\hat{\mathrm{F}}_u)^\top & \hat{\boldsymbol{l}} \\ (\hat{\boldsymbol{F}}_\lambda)^\top & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{b}_z \\ \beta_z \end{bmatrix} = \begin{bmatrix} \hat{\boldsymbol{l}} - (\hat{\mathrm{F}}_{uu})^\top \boldsymbol{a}_z \\ -(\hat{\boldsymbol{F}}_{u\lambda})^\top \boldsymbol{a}_z \end{bmatrix}, \tag{4.13}$$

and

$$\begin{bmatrix} (\hat{\mathrm{F}}_u)^\top & \hat{\boldsymbol{l}} \\ (\hat{\boldsymbol{F}}_\lambda)^\top & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{a}_z \\ \alpha_z \end{bmatrix} = \begin{bmatrix} \boldsymbol{0} \\ 1 \end{bmatrix}. \tag{4.14}$$

We premultiply (4.13) by $((\boldsymbol{\phi}_h)^\top, 0)$, with $\boldsymbol{\phi}_h$ the null-function of F_u (and therefore also of $\hat{\mathrm{F}}_u$, as $S_{h,\hat{p}} = S_{h,p}$), to obtain

$$\beta_z = 1 - (\boldsymbol{\phi}_h)^\top (\hat{\mathrm{F}}_{uu})^\top a_z = 1 - (\hat{\mathrm{F}}_{uu}\boldsymbol{\phi}_h)^\top a_z.$$

Hence, we must now show that $(\hat{\mathrm{F}}_{uu}\boldsymbol{\phi}_h)^\top \boldsymbol{a}_z \neq 0$. We premultiply (4.14) by $((\boldsymbol{\phi}_h)^\top, 0)$ to obtain $\alpha_z = 0$, and hence $\boldsymbol{a}_z \neq 0$, but we must have

$$(\hat{\mathrm{F}}_u)^\top \boldsymbol{a}_z = 0,$$

or in other words, $\boldsymbol{a}_z$ is in the null-space of the operator $(\hat{\mathrm{F}}_u)^\top$, and hence, by the constraint (3.2), $(\hat{\mathrm{F}}_{uu}\boldsymbol{\phi}_h)^\top \boldsymbol{a}_z \neq 0$. □

Remark 4.3. We notice that if $\phi_h \to \phi$, then $z_\lambda \to 0$, which will be witnessed in the proceeding numerical examples. Although the dual problem requires the solution on an enriched finite element space, only two solves with the same matrix is required, as opposed to four for each Newton iteration of the primal problem.

Remark 4.4. As with the primal problem, although we do not use this technique in our numerical experiments, finite differencing can be used for the calculation of the second order derivatives. For example,

$$(\hat{\mathrm{F}}_{uu})^{\top}\boldsymbol{a}_z=\left(\boldsymbol{a}_z^{\top}(\hat{\mathrm{F}}_{uu})\right)^{\top}\approx\left(\frac{\boldsymbol{a}_z^{\top}\left(\hat{\mathrm{F}}_u(u_h+\epsilon\phi_h)-\hat{\mathrm{F}}_u\right)}{\epsilon}\right)^{\top}, \tag{4.15}$$

where $\epsilon=\varepsilon\left(\varepsilon+\|u_h\|/\|\phi_h\|\right)$, for $\varepsilon=10^{-6}$.

5 Bratu problem and DG discretization

The Bratu problem on an open bounded domain $\Omega\in\mathbb{R}^d$, $d\geq 1$, with boundary $\Gamma=\partial\Omega$, is defined by

$$\Delta u+\lambda \mathrm{e}^u=0,\quad \mathbf{x}\in\Omega, \tag{5.1}$$

subject to homogeneous boundary conditions

$$u=0,\quad \mathbf{x}\in\Gamma. \tag{5.2}$$

When $d\leq 2$ the nonlinear operator on the left hand side of (5.1) maps $H_0^1(\Omega)$ to $H^{-1}(\Omega)$ (see Section 17 of reference [18]). We remark that Eq. (5.1) may be posed within the abstract setting outlined in Section 2, based on first applying the inverse Laplacian operator to (5.1). This approach was first developed by Brezzi and co-workers (see, e.g., [6]); a survey of this technique, together with the extension to the incompressible Navier-Stokes equations is presented in [12].

Computing the Fréchet derivative of (5.1) with respect to u in the direction ϕ, we deduce that at a singular point (u^0,ϕ^0,λ^0) the following holds

$$\mathcal{L}^u(u^0,\lambda^0)\equiv\Delta u^0+\lambda^0\mathrm{e}^{u^0}=0, \qquad \mathbf{x}\in\Omega, \tag{5.3a}$$

$$\mathcal{L}^{\phi}(u^0,\lambda^0;\phi_h^0)\equiv\Delta\phi^0+\lambda^0\mathrm{e}^{u^0}\phi^0=0, \qquad \mathbf{x}\in\Omega, \tag{5.3b}$$

subject to the homogeneous boundary conditions

$$u^0=0, \qquad \mathbf{x}\in\Gamma, \tag{5.4a}$$

$$\phi^0=0, \qquad \mathbf{x}\in\Gamma, \tag{5.4b}$$

and the normalisation condition

$$(\phi^0,c)=1,$$

for some $c\in L^2(\Omega)$. The precise choice of c is not unique; clearly, any function which is not orthogonal to ϕ^0 may be employed. For the case of the Bratu problem, numerical experiments indicate that simply selecting $c=1$ is indeed a suitable choice.

5.1 Meshes and traces

In this section we introduce the notation needed to define the symmetric interior penalty DG discretization of the primal problem (5.3)-(5.4). Specifically, we consider $\Omega\subset\mathbb{R}^d$, $d=2$, with the definition for $d=1$ following in a natural manner.

To this end, we assume that Ω can be subdivided into shape-regular meshes $\mathcal{T}_h=\{\kappa\}$ (with possible hanging nodes) consisting of tensor-product elements κ (quadrilaterals, when $d=2$). For the sake of simplicity, we shall suppose that the mesh is 1-regular in the sense that there is at most one hanging node per element-face, which we assume to be the barycenter of the face. We denote by h the piecewise constant mesh function with

$$h(\mathbf{x})\equiv h_\kappa=\mathrm{diam}(\kappa),$$

when $\mathbf{x}$ is in element κ. An interior face of $\mathcal{T}_h$ is defined as the (non-empty) $(d-1)$-dimensional interior of $\partial\kappa_i\cap\partial\kappa_j$, where κ_i and κ_j are two adjacent elements of $\mathcal{T}_h$, not necessarily matching. A boundary face of $\mathcal{T}_h$ is defined as the (non-empty) $(d-1)$-dimensional interior of $\partial\kappa\cap\Gamma$, where κ is a boundary element of $\mathcal{T}_h$. We denote by Γ_{int} the union of all interior faces of $\mathcal{T}_h$. Given a face $f\subset\Gamma_{\text{int}}$, shared by the two elements κ_i and κ_j, where the indices i and j satisfy $i>j$, we write $\mathbf{n}_f$ to denote the (numbering-dependent) unit normal vector which points from κ_i to κ_j; on boundary faces, we put $\mathbf{n}_f=\mathbf{n}$. Further, for v sufficiently smooth, we define the jump of v across f and the mean value of v on f, respectively, by

$$[v]=v|_{\partial\kappa_i\cap f}-v|_{\partial\kappa_j\cap f},\qquad \langle v\rangle=\frac{1}{2}\Big(v|_{\partial\kappa_i\cap f}+v|_{\partial\kappa_j\cap f}\Big).$$

On a boundary edge $f\subset\partial\kappa$, we set

$$[v]=v|_{\partial\kappa\cap f},\qquad \langle v\rangle=v|_{\partial\kappa\cap f}.$$

Finally, given a smooth function v and an element $\kappa\in\mathcal{T}_h$, we denote by v_κ^+ (respectively, v_κ^-) the interior (respectively, exterior) trace of v defined on $\partial\kappa$ (respectively, $\partial\kappa\backslash\Gamma$). Since below it will always be clear from the context which element κ in the subdivision $\mathcal{T}_h$ the quantities v_κ^+ and v_κ^- correspond to, for the sake of notational simplicity, we shall suppress the letter κ in the subscript and write, respectively, v^+ and v^- instead.

Given that κ is an element in the subdivision $\mathcal{T}_h$, we denote by $\partial\kappa$ the union of $(d-1)$-dimensional open faces of κ. Let $\mathbf{x}\in\partial\kappa$ and suppose that $\mathbf{n}_\kappa(\mathbf{x})$ denotes the unit outward normal vector to $\partial\kappa$ at $\mathbf{x}$.

For a given mesh $\mathcal{T}_h$ and polynomial degree $p\geq 1$, we introduce the following finite element space

$$S_{h,p}=\big\{v\in L^2(\Omega):v|_\kappa\in\mathcal{Q}^p(\kappa),\ \ \forall\kappa\in\mathcal{T}_h\big\}.$$

Here, $\mathcal{Q}^p(\kappa)$ denotes the space of tensor product polynomials on κ of degree at most p in each coordinate direction. We then define the space $\mathbf{S}_{h,p}$ by

$$\mathbf{S}_{h,p}=S_{h,p}\times S_{h,p}\times\mathbb{R},$$

with which we shall approximate the base solution, the null-function and the critical parameter value.

5.2 Symmetric interior penalty DG method

The symmetric interior penalty DG approximation of (5.3), (5.4) is defined as follows, where again for notational simplicity we have suppressed the superscript "0": find $\mathbf{u}_h = (u_h,\phi_h,\lambda_h)$ in $\mathbf{S}_{h,p}$, such that

$$\mathcal{N}(\mathbf{u}_h;\mathbf{v}_h) = 0, \tag{5.5}$$

for all $\mathbf{v}_h = (v_h,\psi_h,\chi_h) \in \mathbf{S}_{h,p}$, where

$$\begin{aligned}
\mathcal{N}(\mathbf{u}_h;\mathbf{v}_h) = &-B_a(u_h,v_h)+B_f(v_h,u_h)+B_f(u_h,v_h)-B_\vartheta(u_h,v_h)\\
&+\lambda_h\int_\Omega \mathrm{e}^{u_h}v_h\,\mathrm{d}x+\lambda_h\int_\Omega \mathrm{e}^{u_h}\phi_h\psi_h\,\mathrm{d}x+\chi_h\big((\phi_h,g)-1\big)\\
&-B_a(\phi_h,\psi_h)+B_f(\psi_h,\phi_h)+B_f(\phi_h,\psi_h)-B_\vartheta(\phi_h,\psi_h),\\
B_a(w,v) = &\sum_{\kappa\in\mathcal{T}_h}\int_\kappa \nabla w\cdot\nabla v\,\mathrm{d}x,\\
B_f(w,v) = &\int_{\Gamma_{\mathrm{int}}\cup\Gamma}\langle(\nabla w)\cdot\mathbf{n}_f\rangle[v]\,\mathrm{d}s,\\
B_\vartheta(w,v) = &\int_{\Gamma_{\mathrm{int}}\cup\Gamma}\vartheta[w][v]\,\mathrm{d}s.
\end{aligned}$$

Here, ϑ is called the *discontinuity-penalization* parameter and is defined by $\vartheta|_f = \vartheta_f$, for $f\subset\Gamma_{\mathrm{int}}\cup\Gamma$, where ϑ_f is a non-negative constant on face f. We select ϑ_f as follows: writing $\mathrm{h}\in L^\infty(\Gamma_{\mathrm{int}}\cup\Gamma)$ to denote the mesh function defined by

$$h(\mathbf{x}) = \begin{cases} \min\{h_\kappa,h_{\kappa'}\}, & \mathbf{x}\in f=\partial\kappa\cap\partial\kappa'\subset\Gamma_{\mathrm{int}},\\ h_\kappa, & \mathbf{x}\in f=\partial\kappa\cap\Gamma, \end{cases}$$

we set

$$\vartheta_f = C_\vartheta\frac{p^2}{h}.$$

Here, C_ϑ is a positive constant which is independent of the mesh size and polynomial degree p. Selecting C_ϑ to be sufficiently large guarantees the well-posedness of the interior penalty DG method (5.5). For details concerning the construction of the DG method (5.5), we refer the reader to the article [22], for example.

5.3 *A posteriori* error estimation

We are now in a position to apply the DWR *a posteriori* error estimation technique outlined in Section 3 to the DG method proposed in the previous section. To this end, we have the following result.

Proposition 5.1 (Error Representation Formula)**.** Let $\mathbf{u}$ and $\mathbf{u}_h$ denote the solutions of (5.3)-(5.4) and (5.5), respectively, and suppose that the corresponding dual problem (3.5) is well posed, with solution denoted by $\mathbf{z}=(z'_u,z'_\phi,z'_\lambda)$. Then

$$\lambda-\lambda_h=\varepsilon_\Omega(\mathbf{u},\mathbf{u}_h;\mathbf{z}-\mathbf{z}_h)=\sum_{\kappa\in\mathcal{T}_h}\eta_\kappa, \tag{5.6}$$

for all $\mathbf{z}_h=(z_{u,h},z_{\phi,h},z_{\lambda,h})\in\mathbf{S}_{h,p}$. Here, $\eta_\kappa=\eta_\kappa^u+\eta_\kappa^\phi$,

$$\begin{aligned}\eta_\kappa^u=&\sum_{\kappa\in\mathcal{T}_h}\int_\kappa\mathcal{L}^u(u_h,\lambda_h)w_h\,\mathrm{d}x+\frac{1}{2}\int_{\partial\kappa\setminus\Gamma}\left\{[u_h]\nabla w_h^+\cdot\mathbf{n}_\kappa-w_h^+[\nabla u_h\cdot\mathbf{n}_\kappa]\right\}\mathrm{d}s\\&-\int_{\partial\kappa\setminus\Gamma}\vartheta[u_h]w_h^+\,\mathrm{d}s+\int_{\partial\kappa\cap\Gamma}R_D^u(u_h)\nabla w_h^+\cdot\mathbf{n}\mathrm{d}s-\int_{\partial\kappa\cap\Gamma}\vartheta R_D^u(u_h)w_h^+\,\mathrm{d}s,\end{aligned} \tag{5.7}$$

$$\begin{aligned}\eta_\kappa^\phi=&\sum_{\kappa\in\mathcal{T}_h}\int_\kappa\mathcal{L}^\phi(u_h,\lambda_h;\phi_h)\omega_h\,\mathrm{d}x+\frac{1}{2}\int_{\partial\kappa\setminus\Gamma}\left\{[\phi_h]\nabla\omega_h^+\cdot\mathbf{n}_\kappa-\omega_h^+[\nabla\phi_h\cdot\mathbf{n}_\kappa]\right\}\mathrm{d}s\\&-\int_{\partial\kappa\setminus\Gamma}\vartheta[\phi_h]\omega_h^+\,\mathrm{d}s+\int_{\partial\kappa\cap\Gamma}R_D^\phi(\phi_h)\nabla\omega_h^+\cdot\mathbf{n}\mathrm{d}s-\int_{\partial\kappa\cap\Gamma}\vartheta R_D^\phi(\phi_h)\omega_h^+\,\mathrm{d}s.\end{aligned} \tag{5.8}$$

Moreover, $w_h=z'_u-z_{u,h}$, $\omega_h=z'_\phi-z_{\phi,h}$, $R_D^u(u_h)$ and $R_D^u(\phi_h)$, represent the boundary residuals for u and ϕ, respectively. Since homogeneous Dirichlet boundary conditions have been employed, we have that

$$R_D^u(u_h)\big|_{\partial\kappa\cap\Gamma}=u_h^+\big|_{\partial\kappa\cap\Gamma},\quad\text{and}\quad R_D^\phi(\phi_h)\big|_{\partial\kappa\cap\Gamma}=\phi_h^+\big|_{\partial\kappa\cap\Gamma}.$$

Proof. The error representation formula follows after an application of (3.6) and performing integration by parts. □

6 Numerical experiments

In this section, we present numerical examples to highlight the practical performance of our proposed *a posteriori* error indicator on adaptively refined computational meshes.

6.1 Example 1

In this first example we consider the Bratu problem in one-dimension on the domain $\Omega=(0,1)$. In this case it can be shown that the Bratu problem has zero, one, or two solutions when $\lambda>\lambda^0$, $\lambda=\lambda^0$, and $\lambda<\lambda^0$, respectively, where the critical value λ^0 satisfies the equations

$$1=\frac{1}{4}\sqrt{2\lambda^0}\sinh\left(\frac{\theta^0}{4}\right),\quad\text{and}\quad\theta^0=\sqrt{2\lambda^0}\cosh\left(\frac{\theta^0}{4}\right),$$

see [3]. A simple iterative solution procedure reveals that $\lambda^0 = 3.5138307191$ to 10 decimal places.

We begin with a uniform starting grid which divides $[0,1]$ into 16 elements and carry out an adaptive mesh refinement strategy based on the *a posteriori* error estimate derived in the previous section. For the primal problem a polynomial degree of $p=1$ is used for the numerical approximation of both the base solution and the null-function; on the other hand, the dual problem is approximated with discontinuous piecewise polynomials of degree $\hat{p}=2$. Elements are marked for refinement/derefinement using a fixed fraction strategy according to the size of the (approximate) error indicators $|\hat{\eta}_\kappa|$, with refinement and derefinement fractions set to 20% and 10%, respectively. Here, the approximate error indicator $\hat{\eta}_\kappa$ is defined in an analogous fashion to η_κ in Proposition 5.1 with $\mathbf{z}$ replaced by $\hat{\mathbf{z}}_h \in \mathbf{S}_{h,2}$.

Table 1 shows the number of elements and the number of degrees of freedom employed in the finite element space $S_{h,p}$, the computed critical parameter λ_h^0, the dual critical parameter z_λ, the true error $|\lambda^0 - \lambda_h^0|$, the predicted error $|\sum_\kappa \hat{\eta}_k|$ and the effectivity index $\tau = |\sum_\kappa \hat{\eta}_k| / |\lambda^0 - \lambda_h^0|$, as the mesh $\mathcal{T}_h$ is refined. We first notice that, even on very coarse meshes, the error indicator is performing extremely well, with effectivity indices of 1.00 on all but the first two meshes. We remark that effectivities less than unity are possible since the equality (5.6) only holds with the analytical dual solution; note here, that the dual solution has been numerically computed as part of the *a posteriori* error estimation procedure. Secondly, we note that as the mesh is refined z_λ does indeed appear to be tending to 0.

Table 1: Convergence and effectivity indices for the 1D Bratu problem.

No. Elements	DOF	λ_h^0	z_λ	$\lvert\lambda^0 - \lambda_h^0\rvert$	$\lvert\sum_\kappa \hat{\eta}_k\rvert$	τ
16	32	3.5249864	3.63E-05	1.116E-02	1.115E-02	0.99
21	42	3.5204068	1.04E-05	6.576E-03	6.571E-03	0.99
28	56	3.5169520	2.50E-06	3.121E-03	3.120E-03	1.00
36	72	3.5161228	1.14E-06	2.292E-03	2.291E-03	1.00
46	92	3.5151761	4.42E-07	1.345E-03	1.345E-03	1.00
59	118	3.5147078	2.06E-07	8.771E-04	8.770E-04	1.00
75	150	3.5143572	5.94E-08	5.265E-04	5.264E-04	1.00
96	192	3.5141461	2.45E-08	3.154E-04	3.154E-04	1.00
123	246	3.5140308	1.01E-08	2.001E-04	2.001E-04	1.00
157	314	3.5139494	3.40E-09	1.187E-04	1.187E-04	1.00

6.2 Example 2

In this second example we consider the Bratu problem in two-dimensions on the domain $\Omega = (0,1)^2$. As in the one-dimensional setting, there exists a critical parameter value λ^0, such that for $\lambda > \lambda^0$ the problem has no solution, for $\lambda = \lambda^0$ there exists exactly one

solution, and for $\lambda < \lambda^0$ there are two solutions. To the authors' knowledge there is no analytical expression for the value λ^0 in this case, but calculations have revealed that $\lambda^0 = 6.808124423$ to 9 decimal places, see [27].

Once again we carry out a fixed fraction adaptive strategy using the *a posteriori* error estimator developed in the previous section starting from a uniform grid consisting of 256 elements. As before, we assign a polynomial degree of $p=1$ on each element for the numerical approximation of the primal problem, and employ bi-quadratic elements for the numerical solution of the dual problem.

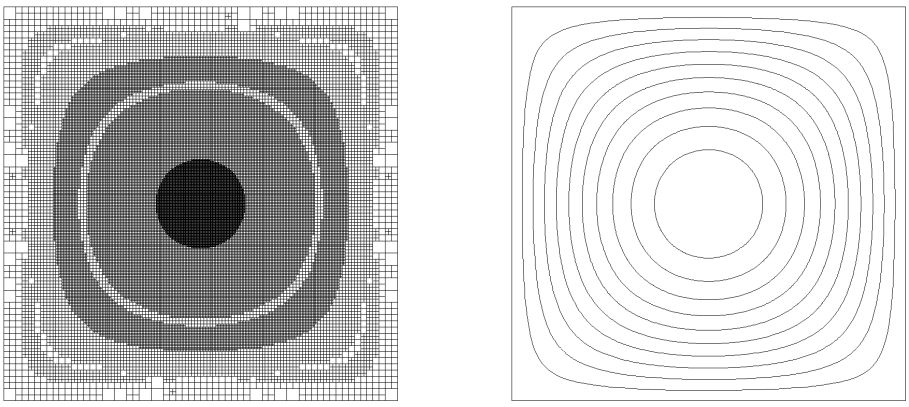
Table 2 shows the number of elements and the number of degrees of freedom employed in the finite element space $S_{h,p}$, the computed critical parameter λ_h^0, the dual critical parameter z_λ, the true error $|\lambda^0 - \lambda_h^0|$, the predicted error $|\sum_\kappa \hat{\eta}_k|$ and the effectivity index $\tau = |\sum_\kappa \hat{\eta}_k| / |\lambda^0 - \lambda_h^0|$, as the mesh is refined. As with the one-dimensional case we witness extremely good error predictions on all meshes, even the very coarse ones. Indeed, except for the first two grids the effectivity index $\tau \approx 1.00$. As the mesh is refined we again see an indication that the dual critical parameter is tending to zero.

Table 2: Convergence and effectivity indices for the 2D Bratu problem.

No. Elements	DOF	λ_h	z_λ	$\lvert\lambda^0 - \lambda_h^0\rvert$	$\lvert\sum_\kappa \hat{\eta}_k\rvert$	τ
256	1024	6.8290830	7.83E-05	2.096E-02	2.093E-01	0.99
448	1792	6.8169639	1.12E-05	8.839E-03	8.833E-02	0.99
784	3136	6.8130504	3.73E-06	4.926E-03	4.924E-02	1.00
1342	5368	6.8110225	1.09E-06	2.898E-03	2.897E-02	1.00
2167	8668	6.8102161	7.09E-07	2.092E-03	2.091E-02	1.00
3583	14332	6.8092367	1.85E-07	1.112E-03	1.112E-02	1.00
5902	23608	6.8087960	6.67E-08	6.715E-04	6.715E-03	1.00
9691	38764	6.8085714	3.30E-08	4.469E-04	4.469E-03	1.00
15922	63688	6.8083832	1.08E-08	2.587E-04	2.588E-03	1.00
26449	105796	6.8082700	3.41E-09	1.455E-04	1.455E-03	1.00

Fig. 1(a) shows a plot of the resultant grid after 9 refinement steps; Fig. 1(b) shows the numerical approximation of the primal base solution computed on that grid. We notice immediately that the mesh has been refined to resolve the features present in the base solution. We remark that the primal null-function and both components of the dual solution exhibit the same features as the primal base solution and thus plots of these have been omitted for brevity. The DWR technique only leads to refinement of those regions in the computational domain where the residual weighted with the dual solution is large. For this problem, the primal and dual solutions are so similar that refinement has occurred to resolve all features present in the primal solution; this will not be the case for more general problems.

Finally, for this two-dimensional problem we discuss the overall cost of our proposed computational algorithm. Given the simplicity of the underlying problem, the Newton iteration for the primal problem converges in around 3 steps, excluding the initial first



(a) (b)

Figure 1: (a) Grid after 9 refinement steps and (b) Primal Base Solution.

mesh, where 6 iterations are necessary; here, we have reduced the l_2-norm of the underlying residual to be below a tolerance of 10^{-11}. In terms of computational time, the cost of computing the dual solution is approximately 1.4 times that of the primal solution. However, we point out that for more complicated problems, the cost of the evaluating the dual solution in comparison to that of computing the primal is significantly reduced.

7 Conclusions

In this article we have developed a framework for *a posteriori* error estimation targeted at numerically estimating critical parameters for nonlinear problems exhibiting quadratic fold points. To this end, we employed the DWR approach, originally developed for the numerical approximation of target functionals of the solution. This general approach was then applied to the symmetric interior penalty DG approximation of the Bratu problem. Numerical experiments presented in both one- and two-dimensions clearly highlight the practical performance of the proposed *a posteriori* error indicator within an automatic adaptive mesh refinement strategy. The extension of these ideas to more complex problems involving incompressible fluid flows in open systems will be considered in the companion articles [8, 9]. Moreover, the rigorous analysis of discontinuous Galerkin methods for the numerical approximation of simple singular points will be undertaken in the forthcoming article [10].

Acknowledgments

KAC, PH, and EJCH gratefully acknowledge the financial support of the EPSRC under the grant EP/E013724. In addition, all of the authors acknowledge the support of the EPSRC under the grant EP/F01340X.

References

[1] M. Ainsworth and J. T. Oden, A Posteriori Error Estimation in Finite Element Analysis, Series in Computational and Applied Mathematics, Elsevier, 1996.

[2] D. N. Arnold, F. Brezzi, B. Cockburn, and L. D. Marini, Unified analysis of discontinuous Galerkin methods for elliptic problems, SIAM. J. Numer. Anal., 39 (2001), 1749–1779.

[3] U. M. Ascher, M. M. Mattheij, and R. D. Russell, Numerical Solution of Boundary Value Problems for Ordinary Differential Equation, SIAM, Philadelphia, PA, 1995.

[4] W. Bangerth and R. Rannacher, Adaptive Finite Element Methods for Differential Equations (Lectures in Mathematics, ETH Zurich), Birkhauser Verlag AG, 2003.

[5] R. Becker and R. Rannacher, An Optimal Control Approach to A-Posteriori Error Estimation in Finite Element Methods, In A. Iserles, editor, Acta Numerica, Cambridge University Press, 2001.

[6] F. Brezzi, J. Rappaz, and P. A. Raviart, Finite dimensional approximation of non-linear problems .2. limit points, Numer. Math., 37 (1) (1981), 1–28.

[7] C. Carstensen and J. Gedicke, An oscillation–free adaptive FEM for symmetric eigenvalue problems, Technical Report 489, DFG Research Center MATHEON, 2008.

[8] K. A. Cliffe, E. Hall, and P. Houston, Adaptivity and a posteriori error control for bifurcation problems II: incompressible fluid flow in open systems with Z_2 symmetry, submitted.

[9] K. A. Cliffe, E. Hall, and P. Houston, Adaptivity and a posteriori error control for bifurcation problems III: incompressible fluid flow in open systems with O(2) symmetry, in preparation.

[10] K. A. Cliffe, E. Hall, and P. Houston, Discontinuous Galerkin methods for bifurcation problems, in preparation.

[11] K. A. Cliffe, E. Hall, and P. Houston, Adaptive discontinuous Galerkin methods for eigenvalue problems arising in incompressible fluid flows, SIAM. J. Sci. Comput., 31(6) (2010), 4607–4632.

[12] K. A. Cliffe, A. Spence, and S. J. Tavener, The Numerical Analysis of Bifurcation Problems with Application to Fluid Mechanics, In A. Iserles, editor, Acta Numerica, Cambridge University Press, 2000.

[13] R.G. Durán, L. Gastaldi, and C. Padra, A posteriori error estimators for mixed approximations of eigenvalue problems, Math. Models. Methods. Appl. Sci., 9 (1999), 1165–1178.

[14] R. G. Durán, C. Padra, and R. Rodriguez, A posteriori error estimates for the finite element approximation of eigenvalue problems, Math. Models. Methods. Appl. Sci., 13(8) (2003), 1219–1229.

[15] K. Eriksson, D. Estep, P. Hansbo, and C. Johnson, Introduction to Adaptive Methods for Differential Equations, In A. Iserles, editor, Acta Numerica, Cambridge University Press, 1995.

[16] J. Gedicke and C. Carstensen, A posteriori error estimators for non-symmetric eigenvalue problems, Technical Report 659, DFG Research Center MATHEON, 2009.

[17] S. Giani and I. Graha, A convergent adaptive method for elliptic eigenvalue problems, SIAM. J. Numer. Anal., 47 (2009), 1067–1091.

[18] R. Glowinski, Numerical Methods for Fluids, In P.G. Ciarlet and J.L. Lions, editors, Handbook of Numerical Analysis, North-Holland, Amsterdam, 2003.

[19] R. Hartmann, Adaptive Finite Element Methods for the Compressible Euler Equations, PhD thesis, University of Heidelberg, 2002.

[20] R. Hartmann and P. Houston, Adaptive discontinuous Galerkin finite element methods for nonlinear hyperbolic conservation laws, SIAM. J. Sci. Comput., 24 (2002), 979–1004.

[21] V. Heuveline and R. Rannacher, A posteriori error control for finite element approximations of elliptic eigenvalue problems, Adv. Comp. Math., 15 (2001), 107–138.
[22] P. Houston, C. Schwab, and E. Süli, Discontinuous hp-finite element methods for advection–diffusion–reaction problems, SIAM. J. Numer. Anal., 39 (2002), 2133–2163.
[23] P. Houston and E. Süli, Adaptive Finite Element Approximation of Hyperbolic Problems, In T. Barth and H. Deconinck, editors, Error Estimation and Adaptive Discretization Methods in Computational Fluid Dynamics, Lect. Notes Comput. Sci. Engrg., 25, pages 269–344, Springer, 2002.
[24] H. B. Keller, Numerical solution of bifurcation and nonlinear eigenvalue problems, In P.H. Rabinowitz, editor, Applications of Bifurcation Theory, pages 359–384, Academic Press, New York, 1977.
[25] M. G. Larson, A posteriori and a priori error analysis for finite element approximations of self–adjoint elliptic eigenvalue problems, SIAM. J. Numer. Anal., 38 (2000), 608–625.
[26] M. G. Larson and T. J. Barth, A Posteriori Error Estimation for Discontinuous Galerkin Approximations of Hyperbolic Systems, In B. Cockburn, G. E. Karniadakis, and C.-W. Shu, editors, Discontinuous Galerkin Methods: Theory, Computation and Applications, Lecture Notes in Computational Science and Engineering, Vol. 11. Springer, 2000.
[27] A. Mohsen, L. F. Sedeek, and S.A. Mohamed, New smoother to enhance multigrid–based methods for bratu problem, Appl. Math. Comput., 204 (2008), 325–339.
[28] G. Moore and A. Spence, The calculation of turning points of nonlinear equations, SIAM. J. Numer. Anal., 17 (1980), 567–576.
[29] C. Nystedt, A priori and a posteriori error estimates and adaptive finite element methods for a model eigenvalue problem, Technical Report 1995-05, Chalmers Finite Element Center, Chalmers University, 1995.
[30] R. Seydel, Numerical computation of branch points in nonlinear equations, Numer. Math., 32 (1979), 339–352.
[31] R. Seydel, Numerical computation of branch points in ordinary differential equations, Numer. Math., 32 (1979), 51–68.
[32] B. Szabó and I. Babuška, Finite Element Analysis, J. Wiley & Sons, New York, 1991.
[33] R. Verfürth, A posteriori error estimates for nonlinear problems, Math. Comp., 62 (1989), 445–475.
[34] R. Verfürth, A Review of a Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques, B. G. Teubner, Stuttgart, 1996.
[35] A.M. Wazwaz, Adomian decomposition method for a reliable treatment of a Bratu-type equation, Appl. Math. Comput., 166 (2005), 652–663.