J. Numer. Math., Vol. 15, No. 3, pp. 181–208 (2007)
DOI 10.1515/ JNUM.2007.009

On the numerical solution of a semilinear elliptic eigenproblem of Lane–Emden type, I: Problem formulation and description of the algorithms

F. J. FOSS, II,* R. GLOWINSKI,* and R. H. W. HOPPE*

Received February 22, 2007
Received in revised form May 5, 2007

Abstract — In this first part of our two-part article, we present some theoretical background along with descriptions of some numerical techniques for solving a particular semilinear elliptic eigenproblem of Lane-Emden type on a triangular domain without any lines of symmetry. For solving the principal first eigenproblem, we describe an operator splitting method applied to the corresponding time-dependent problem. For solving higher eigenproblems, we describe an arclength continuation method applied to a particular perturbation of the original problem, which admits solution branches bifurcating from the trivial solution branch at eigenvalues of its linearization. We then solve the original eigenproblem by 'jumping' to a point on the unperturbed solution branch from a 'nearby' point on the corresponding continued perturbed branch, then normalizing the result. Finally, for comparison, we describe a particular implementation of Newton's method applied directly to the original constrained nonlinear eigenproblem.

Keywords: numerical method, Lane, Emden, semilinear, elliptic, eigenproblem, operator splitting, finite element, arclength continuation, least-squares, control, Newton's method

1. Introduction

Let Ω be a *bounded, Lipschitz domain* in $\mathbb{R}^d$ and denote its boundary by Γ. Consider the following model nonlinear eigenproblem:

$$-\Delta u = \lambda u^3 \quad \text{in } \Omega \tag{1.1}$$

$$u = 0 \quad \text{on } \Gamma \tag{1.2}$$

$$\int_\Omega u^4(x)\,\mathrm{d}x = c \tag{1.3}$$

where $c > 0$ is a normalization constant (we assume hereafter that $c \equiv 1$). The choice of the L^4 norm constraint (1.3) is natural and convenient, for if we multiply equation

*Department of Mathematics, University of Houston, 4800 Calhoun Rd, Houston, TX 77204-3008
This work was supported by NSF grant DMS 0412267.

(1.1) by any solution u (ignoring the natural existence question for the moment) and integrate, we immediately see that

$$\int_\Omega \left|\nabla u(x)\right|^2 \, \mathrm{d}x = \lambda \tag{1.4}$$

which, for one thing, shows that any eigenvalue λ corresponding to an eigenfunction u must be positive. It is worth mentioning that for $d > 4$, the unconstrained problem (1.1)–(1.2) has no nontrivial solution (cf. [9, Subsection 9.4.2]), and thus the constrained problem has no solution.

This problem falls into the class of nonlinear (more precisely, semilinear) elliptic eigenproblems, finding applications in, for example, the study of stellar equilibrium (e.g., the so-called *Lane–Emden model*, cf. [4]). Within the extensive literature on semilinear elliptic problems in general, some of the contributions on, or related to, such eigenproblems include [1–3,5,7,8,13–15,18,19,21], and further citations therein.

The most recent of these citations [21] is the first of three papers that, as of the final stages of this writing, are in various stages of prepublication. In their first paper, the authors summarize, rather well, the numerous and substantial difficulties encountered when attempting to characterize and solve constrained eigenproblems in a Banach space B arising as *Euler–Lagrange systems* of the form

$$F'(u) \;=\; \lambda G'(u) \tag{1.5}$$

$$G(u) \;=\; \alpha \tag{1.6}$$

obtained via differentiation of the associated *Lagrangian functional*

$$\mathcal{L}(u,\lambda) = F(u) - \lambda(G(u) - \alpha). \tag{1.7}$$

The first paper focuses on the case when the component functionals $F(\cdot)$ and $G(\cdot)$ possess what they refer to as the *iso-homogeneity property* defined by the existence of a positive integer $k = l$ such that

$$\begin{aligned} F'(tu) \;&=\; t^k F'(u) \\ G'(tu) \;&=\; t^l G'(u) \qquad \forall t > 0,\ u \in B. \end{aligned} \tag{1.8}$$

The authors show that this property is sufficient to characterize eigenpairs $\{u,\lambda\}$ solving (1.5)–(1.6) as critical point and value pairs $\{u,J(u)\}$ of the associated *Rayleigh quotient* functional

$$\begin{aligned} J(u) \;&:=\; \frac{F(u)}{G(u)}, \quad u \in B\backslash U \\ U \;&:=\; \{u \in B |\ G(u) = 0\}. \end{aligned} \tag{1.9}$$

The authors then present a so-called (*modified*) *Local MiniMax* (LMM) *method* for

finding multiple critical points of $J(\cdot)$, constrained to the unit sphere and ordered by their so-called (*local*) *MiniMax Index* (MMI) and show how the method relates to the established characterizations of Rayleigh–Ritz, Courant–Fischer, and Ljusternik–Schnirelman. Finally, they implement the modified LMM method and use it to solve a nonlinear p-Laplacian eigenproblem on a 2×2 square with some interesting and novel results. Although we have not seen their subsequent work, the authors evidently consider non iso-homogenous problems in their second paper, of which our model problem is a particular case as it satisfies a *bi-homogeneity property* with $k = 1$ and $l = 3$.

In the earlier paper [15], the author discusses and implements a Constrained Steepest Decent Method (CSDM) initializing a Constrained Mountain Pass Algorithm (CMPA) for solving constrained minimax problems arising as systems of variational functionals corresponding to various semilinear elliptic equations, including a particular case ($\lambda = 1$) of problem (1.1)–(1.3) on the unit square. The details of the methodology are rather intricate, but it is our basic understanding that the method first involves the finding of two suitable critical point solutions of the problem via the CSDM that satisfy the conditions of a constrained version of the classical mountain pass theorem. These two solutions are then used in the CMPA as endpoints of a path constructed (and possibly refined) in such a way as to traverse a so-called 'mountain pass', from the 'top' (i.e., local maximum point) of which the CSDM is used again to descend from this local maximum point along 'the ridge' of local maxima to the new mountain pass-type critical point solving the constrained minimax problem.

In the present work, we discuss and implement some alternative numerical methods and explore their shortcomings and merits. We restrict ourselves to the numerical investigation of problem (1.1)–(1.3) on a particular domain, looking for approximate variational solutions in a suitable Hilbert space.

In Section 2, we discuss the solution of problem (1.1)–(1.3) for the principal eigenpair (u_1, λ_1). Specifically, in Subsection 2.1, we prove that this problem is equivalent to energy minimization on the unit $L^4(\Omega)$ sphere in the Sobolev space $H_0^1(\Omega)$, and that the latter formulation (hence the former) has a solution. In Subsections 2.2 and 2.3, we present a computational algorithm for solving this problem based on the so-called *time-dependent approach* and *operator splitting*.

In Section 3, we discuss the solution of the unconstrained problem (1.1)–(1.2) in the setting of *arclength continuation theory* with the particular goal of finding higher eigenmodes, treating the problem with constraint (1.3) as a special case. In Subsection 3.1, we present the general and problem-specific arclength continuation framework. Within this framework, we discuss two local correction methodologies in Subsections 3.1.1 and 3.1.2.

Finally, for completeness and comparison purposes, we provide in Subsection 3.2 a direct approach to solving (1.1)–(1.3) based on an application of *affine covariant Newton's method w/wo damping* (à la P. Deuflhard).

2. The principal eigenproblem

2.1. Theoretical background

In this section, we present some of the supporting existence/uniqueness theory for problem (1.1)–(1.3), focussing on the *principal*, or *minimal*, *eigenproblem*. It is natural to look for *weak solutions* of this problem in the Sobolev space $H_0^1(\Omega) \times R$. The *weak formulation* of (1.1)–(1.3) is: Find $\{u, \lambda\} \in H_0^1(\Omega) \times R$ such that

$$\int_\Omega \nabla u(x) \cdot \nabla w(x)\, dx - \lambda \int_\Omega u^3(x) w(x)\, dx = 0 \quad \forall w \in H_0^1(\Omega) \tag{2.1}$$

$$\int_\Omega u^4(x)\, dx - 1 = 0. \tag{2.2}$$

Consider the following variational problem:

$$\text{Find } u \in \mathcal{E}_4 := H_0^1(\Omega) \cap \mathcal{S}_4 \text{ such that } J(u) \leqslant J(v) \ \forall v \in \mathcal{E}_4 \tag{2.3}$$

where $J(v) := \frac{1}{2}\int_\Omega |\nabla v(x)|^2\, dx$ and $\mathcal{S}_4 := \{v \in L^4(\Omega) \mid \int_\Omega v^4(x)\, dx = 1\}$. It is easy to see that, for any pair $\{u, \lambda\}$ solving (2.3), the weak formulation (2.1)–(2.2) comprises the so-called *first-order necessary optimality conditions* resulting from differentiation of the *Lagrangian functional* $\mathcal{L} : H_0^1(\Omega) \times R^+ \to R$ defined by

$$\begin{aligned} \mathcal{L}(v, \alpha) &:= J(v) - \frac{\alpha}{4}\left(\int_\Omega v^4(x)\, dx - 1\right) \\ &= \frac{1}{2}\int_\Omega |\nabla v(x)|^2\, dx - \frac{\alpha}{4}\left(\int_\Omega v^4(x)\, dx - 1\right). \end{aligned} \tag{2.4}$$

The first equation in the first-order necessary optimality system is the weak form of the *Euler–Lagrange equation* defined by

$$\langle \mathcal{L}'_v(u, \lambda), w\rangle := \langle J'(u), w\rangle - \lambda \int_\Omega u^3(x) w(x)\, dx = 0 \tag{2.5}$$

where we see that the eigenvalue λ is the *Lagrange multiplier* corresponding to the constraint defined by the second equation

$$\mathcal{L}'_\alpha(u, \lambda) := -\frac{1}{4}\left(\int_\Omega u^4(x)\, dx - 1\right) = 0. \tag{2.6}$$

Thus, we see that any solution u of (2.3) is an eigenfunction corresponding to the eigenvalue λ, the pair $\{u, \lambda\}$ necessarily solving (2.1)–(2.2). Since any $H_0^1(\Omega)$ function is continuously imbedded in $L^4(\Omega)$ (*Sobolev imbedding theorem*), we may normalize any nonzero $H_0^1(\Omega)$ function so that it lies in $\mathcal{S}_4$ (so $\mathcal{E}_4$ is nonempty) and then define the *principal eigenvalue* λ_1 as the *minimum value*

$$\lambda_1 := \inf_{v \in \mathcal{E}_4} \int_\Omega |\nabla v(x)|^2\, dx \equiv \inf_{v \in \mathcal{E}_4} 2J(v) \tag{2.7}$$

and a *principal eigenfunction* u_1 as a corresponding minimizer solving Problem (2.3), with the *principal eigenpair* $\{u_1, \lambda_1\}$ solving (2.1)–(2.2). We now show

Proposition 2.1. *Problem* (2.3) *has a solution.*

Proof. Since $H_0^1(\Omega)$ is, in fact, compactly imbedded in $L^4(\Omega)$ (by the *Rellich–Kondrachov imbedding theorem*), and since the functional J (being half the square of the equivalent energy norm $\|\cdot\| \equiv |\cdot|_{1,2,\Omega}$ on $H_0^1(\Omega)$) is continuous, coercive, and bounded below by zero on $H_0^1(\Omega)$, and so also on $\mathcal{E}_4$, there exists a *minimizing sequence* $\{v_k\}_{k\in\mathbb{N}}$ in $\mathcal{E}_4$ such that

$$\lim_{k\to\infty} J(v_k) = \inf_{v\in\mathcal{E}_4} J(v). \tag{2.8}$$

Since $\{J(v_k)\}_{k\in\mathbb{N}}$ is bounded in R^+, $\{v_k\}_{k\in\mathbb{N}}$ must be bounded in $H_0^1(\Omega)$ (by coercivity), and since $H_0^1(\Omega)$ is a Hilbert space, whence reflexive, it follows that there exists $\mathbb{N}' \subseteq \mathbb{N}$ and $u \in H_0^1(\Omega)$ such that the subsequence $\{v_{k'}\}_{k'\in\mathbb{N}'}$ converges *weakly* to u in $H_0^1(\Omega)$, that is, $\langle f, v_{k'}\rangle \to \langle f, u\rangle$ as $k' \to \infty$ for all $f \in H^{-1}(\Omega)$, or equivalently (by the *Riesz representation theorem*), $\int_\Omega \nabla w\cdot\nabla v_{k'}\,\mathrm{d}x \to \int_\Omega \nabla w\cdot\nabla u\,\mathrm{d}x$ for all $w \in H_0^1(\Omega)$. Now,

$$\begin{aligned} 0 \leqslant \frac{1}{2}\int_\Omega |\nabla(v_{k'}-u)(x)|^2\,\mathrm{d}x \;=\; & \frac{1}{2}\int_\Omega |\nabla v_{k'}(x)|^2\,\mathrm{d}x - \int_\Omega \nabla u(x)\cdot\nabla v_{k'}(x)\,\mathrm{d}x \\ & + \frac{1}{2}\int_\Omega |\nabla u(x)|^2\,\mathrm{d}x \quad \forall k' \in \mathbb{N}' \end{aligned}$$

and thus, upon taking the limit as $k' \to \infty$, we see that $J(u) \leqslant \inf_{v\in\mathcal{E}_4} J(v)$, which becomes an equality if we can show that $u \in \mathcal{E}_4$. But this follows from the compact imbedding of $H_0^1(\Omega)$ in $L^4(\Omega)$, since then the weak convergence of $\{v_{k'}\}_{k'\in\mathbb{N}'}$ in $H_0^1(\Omega)$ implies its strong convergence in $L^4(\Omega)$, and since $\|v_{k'}\|_{0,4,\Omega} = 1$ for all k', it follows that $\|u\|_{0,4,\Omega} = 1$. □

Remark 2.1. It is evident that if $\{u_1, \lambda_1\}$ is a principal eigenpair, then so is $\{-u_1, \lambda_1\}$. For higher eigenproblems, the nonuniqueness is less trivial than a sign change and depends, at least, on the geometry of Ω (cf. [22]). Although the nonuniqueness question is itself an interesting and important one, we do not explore it further herein.

2.2. Approximating the principal eigenproblem

Problem (1.1)–(1.3) is really a parameterized family of stationary nonlinear Dirichlet problems. Here, we are looking for the first (as a function of the parameter) such solution and the corresponding value of the parameter. For a general discussion of, and some additional references for, some methods used to solve stationary nonlinear

Dirichlet problems, see [11, Chapter VII, Section 3]. One such method discussed there, and which we employ here, is the so-called *time-dependent approach*. The general idea of this approach is to first introduce the *parabolic initial value problem* associated with stationary problem (1.1)–(1.3), namely

$$\frac{\partial u}{\partial t} - \Delta u = \lambda u^3 \quad \text{in } \Omega \times (0,+\infty) \tag{2.9}$$

$$u = 0 \qquad \text{on } \Gamma \times (0,+\infty) \tag{2.10}$$

$$\int_\Omega u^4(x,\cdot)\,\mathrm{d}x = 1, \qquad t \in (0,+\infty) \tag{2.11}$$

$$u(\cdot,0) = u_0 \qquad \text{in } \Omega\,. \tag{2.12}$$

For a particular choice of initial data u_0, we then discretize this problem in time and at each time step solve the (weak form of) the resulting semi-discrete problem in $H_0^1(\Omega)$. The only twist here is that we are solving not only for an update in u, but also in λ, at each time step, and therefore we need to initialize λ as well. With the proper time discretization and initialization (discussed below), the resulting approximating sequence of iterates $\{u^n\}_{n\in\mathbb{N}}$ will be a monotonically norm-decreasing, minimizing sequence converging to a steady state solving the principal eigenproblem for (1.1)–(1.3). To see that the sequence is monotonically norm-decreasing, multiply (2.9) by $\partial u/\partial t$ and integrate over Ω to obtain

$$\int_\Omega \left(\frac{\partial u(x,t)}{\partial t}\right)^2 \mathrm{d}x + \int_\Omega \frac{\partial}{\partial t}\left(\frac{1}{2}|\nabla u(x,t)|^2\right) \mathrm{d}x - \int_\Gamma \frac{\partial u(x,t)}{\partial n}\frac{\partial u(x,t)}{\partial t}\,\mathrm{d}s = \lambda \int_\Omega \frac{\partial}{\partial t}\left(\frac{1}{4}u^4(x,t)\right) \mathrm{d}x. \tag{2.13}$$

Now, the boundary integral vanishes since $\partial u/\partial t = 0$ on Γ (from (2.10)). Also, we may interchange the order of time differentiation and space integration to obtain (using (2.11))

$$\int_\Omega \left(\frac{\partial u(x,t)}{\partial t}\right)^2 \mathrm{d}x + \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega |\nabla u(x,t)|^2\,\mathrm{d}x = \frac{\lambda}{4}\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega u^4(x,t)\,\mathrm{d}x \equiv 0 \tag{2.14}$$

and thus

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega |\nabla u(x,t)|^2\,\mathrm{d}x = -\int_\Omega \left(\frac{\partial u(x,t)}{\partial t}\right)^2 \mathrm{d}x \leqslant 0. \tag{2.15}$$

This shows that the function $t \mapsto \|u(\cdot,t)\|$ is decreasing, and thus the approximating sequence of iterates $\{u^n\}$ will be monotonically norm-decreasing provided the discretization is consistent with this exact property of (2.9)–(2.12).

2.3. Numerical algorithm and discretization

For the time discretization of problem (2.9)–(2.12), we use the operator splitting theory of Lie as applied in the time-dependent PDE setting by Yanenko and Marchuk

(cf. [12, Chapters II and VII], and references therein), one possible implementation of which results in the following time-discrete system:

(1) $u^0 = u_0$ is given. (2.16)

For $n \geqslant 0$ until convergence, solve

(2) $$\frac{u^{n+1/2} - u^n}{\tau} = \lambda^{n+1}\left(u^{n+1/2}\right)^3 \quad \text{in } \Omega \tag{2.17}$$

$$\int_\Omega \left(u^{n+1/2}\right)^4 (x)\,\mathrm{d}x = 1. \tag{2.18}$$

(3) $$\frac{u^{n+1} - u^{n+1/2}}{\tau} - \Delta u^{n+1} = 0 \quad \text{in } \Omega \tag{2.19}$$

$$u^{n+1} = 0 \quad \text{on } \Gamma. \tag{2.20}$$

To transition from time step n to $n+1$, the nonlinear subproblem in Step (2) of this scheme requires the simultaneous solution of coupled cubic and integral equations defined on Ω to find the pair $\{u^{n+1/2}, \lambda^{n+1}\}$. The subproblem in Step (3) is a linear elliptic boundary value problem in u^{n+1} involving the solution found in Step (2). We discretize both subproblems in space using a standard piecewise linear finite element approximation of the variational forms in $H_0^1(\Omega)$ on a uniform, geometrically-conforming mesh and solve the linear subproblem using a direct method.

Although there is more than one way to choose the initialization in (2.16), we choose u_0 to be the principal eigenfunction w_1 satisfying the following constrained linear eigenproblem

$$-\Delta w = \propto w \quad \text{in } \Omega \tag{2.21}$$

$$w = 0 \quad \text{on } \Gamma \tag{2.22}$$

$$\int_\Omega w^4(x)\,\mathrm{d}x = 1 \tag{2.23}$$

which can be solved variationally in $H_0^1(\Omega)$ for the minimal eigenpair $\{w_1, \propto_1\}$ using the *inverse power method* (cf. [12, Chapter VII, Subsection 36.3]). Once we have its solution w_1, we simply renormalize via division by $\|w_1\|_{0,4}$ to satisfy the unit L^4 norm constraint (2.23) and take u_0 to be this result.

We solve the nonlinear subproblem in Step (2) of this scheme iteratively using two nested scalar implementations of Newton's method. Specifically, at each step k of the outer implementation (which solves the coupled cubic and integral equations), the inner implementation solves the set of scalar cubic equations

$$g_\tau(u_p) := \tau\lambda u_p^3 - u_p + u_p^n = 0, \quad p \in \Sigma_{0,h} \tag{2.24}$$

in u_p at every node p in the approximating interior finite element mesh $\Sigma_{0,h}$, assuming the current outer iterate of $\lambda = \lambda_k^{n+1}$ is given and u_p^n is known from the previous time step. The justification for solving $g_\tau(u) = 0$ pointwise is a combination of the

fact that we initialize the scheme with a smooth function u_0 (the linear eigenfunction w_1 solving (2.21)–(2.23)) and use the trapezoidal rule for approximating the integrals in the associated weak form of the equation, which diagonalizes the otherwise coupled set of nonlinear equations.

The integral constraint

$$H_\tau(\lambda) := \int_\Omega u_\lambda^4(x)\,\mathrm{d}x - 1 = 0 \tag{2.25}$$

is then used for the Newton update of the implicitly-defined λ using the newly computed iterate $u_\lambda = \sum_{p\in\Sigma_{0,h}} u_p\phi_p$ solving (2.24) pointwise on $\Sigma_{0,h}$, where the ϕ_p are the finite element nodal basis functions.

Explicitly, then, we have the following algorithm for solving the problem in Step (2) for each time step:

(2_1) $\quad \lambda_0^{n+1} = 0$ or $\lambda_0^{n+1} = \infty_1$. (2.26)

For $k \geqslant 0$ until convergence,

(2_2) at every mesh node $p \in \Sigma_{0,h}$,

(2_{2_1}) $\quad$ take $\left|u_{p,k,0}^{n+1/2}\right| \in \left[0, \frac{2}{3}\frac{1}{\sqrt{3\tau\lambda_k^{n+1}}}\right)$. (2.27)

For $l \geqslant 0$ until convergence,

(2_{2_2}) $\quad u_{p,k,l+1}^{n+1/2} = u_{p,k,l}^{n+1/2} - \dfrac{g_\tau(u_{p,k,l}^{n+1/2})}{g'_\tau(u_{p,k,l}^{n+1/2})}$. (2.28)

(2_3) $\quad \lambda_{k+1}^{n+1} = \lambda_k^{n+1} - \dfrac{H_\tau(\lambda_k^{n+1})}{H'_\tau(\lambda_k^{n+1})}$. (2.29)

Concerning the choice of τ, we notice immediately that $g'_\tau(u_p) = 3\tau\lambda u_p^2 - 1$, showing that critical points of g_τ occur at $u_p^\pm = \pm\frac{1}{\sqrt{3\tau\lambda}}$ with corresponding critical values $g_\tau\left(\pm\frac{1}{\sqrt{3\tau\lambda}}\right) = u_p^n \mp \frac{2}{3}\frac{1}{\sqrt{3\tau\lambda}}$. Since $g''_\tau(u_p) = 6\tau\lambda u_p$, we see that $u_p = 0$ is an inflection point with corresponding inflection value $g_\tau(0) = u_p^n$. For a given outer Newton iterate $\lambda = \lambda_k^{n+1}$, in order for there to be a root between the two critical points near the inflection point and most recently computed solution u_p^n, the critical values must have opposite signs (or one must itself be zero, in which case the critical point is a double root, a situation that we would like to avoid). This means that we must have $|u_p^n| < \frac{2}{3}\frac{1}{\sqrt{3\tau\lambda_k^{n+1}}}$ (from which we deduce the upper bound in (2.27)) or $\tau < \tau_{p,k,n} := \frac{4}{27\lambda_k^{n+1}\left(u_p^n\right)^2}$ for all $p \in \Sigma_{0,h}$, for all k and n. Equivalently, we have the following necessary constraint on the time step τ:

$$\tau < \tau_n := \frac{4}{27\max\limits_k\{\lambda_k^{n+1}\}\left(\max\limits_{p\in\Sigma_{0,h}}\{|u_p^n|\}\right)^2} \qquad \forall n. \tag{2.30}$$

Unfortunately, this constraint is implicit in τ since λ_k^{n+1} and u_p^n depend on τ in a rather complicated way through the two Newton iteration processes involving $g_\tau(\cdot)$, $H_\tau(\cdot)$ and their derivatives. Thus, it is only useful as an *a posteriori* monitor of whether or not the chosen τ is satisfactory with respect to this condition.

The calculation of $H'_\tau(\cdot)$ is straightforward but more involved. First, since u_λ is a function of λ through equation (2.24), we have upon differentiation with respect to λ that $\tau u_\lambda^3 + 3\tau\lambda u_\lambda^2 u'_\lambda - u'_\lambda = 0$ so that $u'_\lambda = \tau u_\lambda^3/(1 - 3\tau\lambda u_\lambda^2)$. Using this result we find that

$$H'_\tau(\lambda) = \int_\Omega 4u_\lambda^3(x)u'_\lambda(x)\,\mathrm{d}x = \int_\Omega \frac{4\tau u_\lambda^6(x)}{1 - 3\tau\lambda u_\lambda^2(x)}\,\mathrm{d}x. \tag{2.31}$$

From this expression for $H'_\tau(\cdot)$, with $\lambda = \lambda_k^{n+1}$ and $u_\lambda = u_k^{n+1/2}$ we see that another *a posteriori* necessary condition on τ is that

$$\tau < \tau_n := \frac{1}{3\max\limits_k \left\{ \lambda_k^{n+1} \left(\max\limits_{p\in\Sigma_{0,h}} \left\{ \left| u_{p,k}^{n+1/2} \right| \right\} \right)^2 \right\}} \quad \forall n. \tag{2.32}$$

From numerical experiments, this condition appears to be consistently less restrictive than condition (2.30), and thus one would use the latter to monitor the choice of τ.

It is well known (cf. [12, Chapter VI]) that the Marchuk–Yanenko scheme is at most first-order accurate, and its stability and convergence properties depend heavily on the operators appearing in each subproblem and the choice of τ. It is important to note that the necessary constraints on τ derived above are by no means sufficient for overall convergence of scheme (2.16)–(2.20). Indeed, these constraints on τ only guarantee the solvability of equation (2.24) and well-posedness of the $H'_\tau(\cdot)$ integral (2.31). The final choice of τ must also be consistent with overall stability and convergence of the operator splitting scheme. Finally, from numerical experiments, it is our experience that extreme care must be used when attempting to adaptively modify τ in this case.

For the results of the numerical experiments with our implementation of this method, we refer the reader to Part II, Section 2, of our article.

3. Higher eigenproblems

Attempts to adapt the methodology used to solve the principal eigenproblem for use in solving even the second eigenproblem (let alone higher ones) were not entirely successful for a variety of reasons. Although we implemented four methods that were successful at solving the first two eigenproblems (cf. [10]), only one of these proved robust enough (without further fine-tuning) to solve the third and higher eigenproblems. All of the implementations that failed to solve eigenproblems beyond the second were based on solving an approximating linear formulation of the

original semilinear problem (although this fact alone doesn't account for the failures of these methods). The one method robust enough to solve the higher eigenproblems preserves the original semilinear structure of the problem and incorporates it into the solution strategy together with a particular perturbation term that gives rise to a natural initialization of the numerical scheme. The approach uses the machinery of the classical technique of *arclength continuation* (cf. [16,17]) and that of its subsequent application to the efficient numerical solution of least-squares formulations of some nonlinear boundary value problems (cf. [13]).

In the sequel, we focus our discussion on the implementation of the arclength continuation method. For completeness and comparison purposes, however, we also offer some results obtained from the implementations of so-called *error-oriented*, or *affine covariant*, *undamped* and *damped Newton iterations*, discussed in a general setting by P. Deuflhard in [8]. In contrast to the methods previously discussed, these Newton methods are applied directly to the original constrained semilinear eigenproblem (1.1)–(1.3).

3.1. The arclength continuation framework

For a fairly detailed account of the theory of arclength continuation applied to the least squares formulation of general, and some specific, nonlinear boundary value problems, we refer the reader to Glowinski, *et al.* [13]. In this section, we summarize the presentation found there in the context of the current problem.

The general idea behind the use of arclength continuation for solving a nonlinear problem, say $S(u,\lambda) = 0$ with u in a (real, in this case) Hilbert space $(V,(\cdot,\cdot))$ and $\lambda \in R$, is to adjoin a so-called *arclength constraint* $l(u,\lambda,s) = 0$ that parameterizes solution branches $\{\{u(s),\lambda(s)\}\}$ in terms of an arclength parameter s. Recall that any parameterized solution branch $\{\{u(s),\lambda(s)\}\} \subset V \times R$ is said to be *parameterized by arclength* provided $\|\dot{u}(s)\|^2 + |\dot{\lambda}(s)|^2 - 1 = 0$ for all s, that is, the tangent vector $\{\dot{u}(s),\dot{\lambda}(s)\}$ has unit length for all s, and is the natural candidate for the arclength constraint l. We then employ the *implicit function theorem* and *bifurcation theory* in order to assert, depending on the behavior of the respective partial derivatives of S and l with respect to the variables u and λ, the local existence and uniqueness of solution branches in the neighborhood of a known solution $\{u_0,\lambda_0\} := \{u(s_0),\lambda(s_0)\}$. Note that along any branch of solutions, the derivatives with respect to arclength must vanish since the functions are identically zero there. This leads to the so-called *Davidenko equations* for the tangent vector $\{\dot{u},\dot{\lambda}\}$ along the branch. If we know or can solve for a corresponding tangent vector $\{\dot{u}_0,\dot{\lambda}_0\} := \{\dot{u}(s_0),\dot{\lambda}(s_0)\}$ at s_0, then we may predict to first order the location of the next iterate along the branch and use it to solve for another nearby solution on the same branch using an appropriate nonlinear solver, and thus (theoretically anyway) produce the entire branch via iteration (cf. [16]).

More concretely, to solve the system

$$S(u,\lambda) \;=\; 0 \tag{3.1}$$

$$l(u,\lambda,s) \;=\; 0 \tag{3.2}$$

along a branch of solutions $\{(u(s),\lambda(s))\}$ in $V \times R$ parameterized by arclength s, one particular arclength continuation process is the following *predictor–corrector method*:

Step 0: Initialization

Assume a regular point $\{u_0,\lambda_0\} := \{u(s_0),\lambda(s_0)\}$ on, and a tangent $\{\dot{u}_0,\dot{\lambda}_0\} := \{\dot{u}(s_0),\dot{\lambda}(s_0)\}$ to, a solution branch (3.1)–(3.2) are known.

Step 1: Continuation

Step 1.1: Tangent line prediction

Set

$$\{u_1^0,\lambda_1^0\} = \{u_0,\lambda_0\} + \{\dot{u}_0,\dot{\lambda}_0\}\Delta s_0 \tag{3.3}$$

for a suitably chosen arclength step $\Delta s_0 := s_1 - s_0$.

Step 1.2: Correction

Solve for $\{u_1,\lambda_1\} := \{u(s_1),\lambda(s_1)\}$ on the solution branch via Newton's method

$$\begin{pmatrix} S_u(u_1^k,\lambda_1^k) & S_\lambda(u_1^k,\lambda_1^k) \\ l_u(u_1^k,\lambda_1^k,s_1) & l_\lambda(u_1^k,\lambda_1^k,s_1) \end{pmatrix} \begin{pmatrix} \Delta u_1^k \\ \Delta\lambda_1^k \end{pmatrix} = \begin{pmatrix} -S(u_1^k,\lambda_1^k) \\ -l(u_1^k,\lambda_1^k,s_1) \end{pmatrix} \tag{3.4}$$

$$\{u_1^{k+1},\lambda_1^{k+1}\} = \{u_1^k,\lambda_1^k\} + \{\Delta u_1^k,\Delta\lambda_1^k\} \tag{3.5}$$

for $k = 0,1,\ldots$.

Step 2: Update

Solve the Davidenko equations (which arise from differentiation with respect to s along the solution branch)

$$\begin{pmatrix} S_u(u_1,\lambda_1) & S_\lambda(u_1,\lambda_1) \\ l_u(u_1,\lambda_1,s_1) & l_\lambda(u_1,\lambda_1,s_1) \end{pmatrix} \begin{pmatrix} \dot{u}_1 \\ \dot{\lambda}_1 \end{pmatrix} = \begin{pmatrix} 0 \\ -l_s(u_1,\lambda_1,s_1) \end{pmatrix} \tag{3.6}$$

for $\{\dot{u}_1,\dot{\lambda}_1\} := \{\dot{u}(s_1),\dot{\lambda}(s_1)\}$.

Set $s_0 = s_1$, $\{u_0,\lambda_0\} = \{u_1,\lambda_1\}$, $\{\dot{u}_0,\dot{\lambda}_0\} = \{\dot{u}_1,\dot{\lambda}_1\}$, and return to Step 1.

Remark 3.1. As a practical matter, the system in the correction step is solved via the particular equivalent *Schur complement* system

$$\left(\begin{array}{cc} S_u & S_\lambda \\ 0 & l_\lambda - l_u S_u^{-1} S_\lambda \end{array} \right) \Bigg|_{\{u_1^k, \lambda_1^k, s_1\}} \left(\begin{array}{c} \Delta u_1^k \\ \Delta \lambda_1^k \end{array} \right) = \left(\begin{array}{c} -S \\ -l + l_u S_u^{-1} S \end{array} \right) \Bigg|_{\{u_1^k, \lambda_1^k, s_1\}} \tag{3.7}$$

while from the Davidenko equations in the update step we have that

$$\dot{u}_1 = \dot{\lambda}_1 \hat{u}, \text{ where } \hat{u} \text{ solves } S_u(u_1, \lambda_1)\hat{u} = -S_\lambda(u_1, \lambda_1) \tag{3.8}$$

and depending on the form of the second equation in (3.6), $\dot{\lambda}_1$ is found either from that equation as

$$\dot{\lambda}_1 = \frac{-l_s(u_1, \lambda_1, s_1)}{l_u(u_1, \lambda_1, s_1)\hat{u} + l_\lambda(u_1, \lambda_1, s_1)} \tag{3.9}$$

or from the arclength constraint (3.2) via the solution of $l(u_1, \lambda_1, s_1) = 0$, for example the natural arclength constraint (3.13) gives

$$\dot{\lambda}_1 = \pm \frac{1}{\sqrt{1 + \|\hat{u}\|^2}}. \tag{3.10}$$

Finally, it may only be necessary to solve the Davidenko equations in the update step periodically through the continuation process to 'renormalize' the tangent. Otherwise, it is sufficient to approximate the tangent via

$$\{\dot{u}_1, \dot{\lambda}_1\} = \left\{ \frac{u_1 - u_0}{s_1 - s_0}, \frac{\lambda_1 - \lambda_0}{s_1 - s_0} \right\}. \tag{3.11}$$

In [16], it is mentioned that imposing the arclength constraint (3.2) periodically is good policy so that more uniform steps are taken during the continuation process. In fact, we shall see later that failure to renormalize the tangent via the Davidenko equations can result in the arclength step becoming too small or too large, leading to the failure of the method to continue the desired nontrivial branch of solutions.

Let us now apply this general arclength continuation framework to our particular problem. Casting the original (unconstrained) eigenproblem (1.1)–(1.2) in this framework, the augmented problem we wish to solve is

$$S(u, \lambda) := -\Delta u - \lambda u^3 = 0 \tag{3.12}$$

$$l(u, \lambda) := \|\dot{u}\|^2 + |\dot{\lambda}|^2 - 1 = 0 \tag{3.13}$$

in $H_0^1(\Omega) \times R$ (note that l does not depend explicitly on s).

From equation (3.12), it is clear that $\{\{0, \lambda\}\}$ is a trivial branch of solutions. Concerning the existence of nontrivial branches of solutions, it is natural to wonder if there are any bifurcating from the trivial branch. We can determine whether or

not this is the case by examining the linearization of system (3.12)–(3.13). Upon differentiating with respect to s, we have that

$$\begin{pmatrix} S_u(u,\lambda) & S_\lambda(u,\lambda) \\ l_u(u,\lambda,s) & l_\lambda(u,\lambda,s) \end{pmatrix} = \begin{pmatrix} -\Delta - 3\lambda u^2 & -u^3 \\ 2(\dot{u}, \frac{\mathrm{d}}{\mathrm{d}s}\cdot) & 2\dot{\lambda}\frac{\mathrm{d}}{\mathrm{d}s} \end{pmatrix}, \quad l_s(u,\lambda,s) = 0 \tag{3.14}$$

so along any branch of solutions $\{\{u(s),\lambda(s)\}\}$, the Davidenko equations (3.6) must be satisfied and therefore we must have

$$\begin{pmatrix} -\Delta - 3\lambda u^2 & -u^3 \\ 2(\dot{u}, \frac{\mathrm{d}}{\mathrm{d}s}\cdot) & 2\dot{\lambda}\frac{\mathrm{d}}{\mathrm{d}s} \end{pmatrix} \begin{pmatrix} \dot{u} \\ \dot{\lambda} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{3.15}$$

Along the trivial branch of solutions this reduces to

$$\begin{pmatrix} -\Delta & 0 \\ 2(\dot{u}, \frac{\mathrm{d}}{\mathrm{d}s}\cdot) & 2\dot{\lambda}\frac{\mathrm{d}}{\mathrm{d}s} \end{pmatrix} \begin{pmatrix} \dot{u} \\ \dot{\lambda} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{3.16}$$

Since $S_u(0,\lambda) = -\Delta : H_0^1(\Omega) \to H^{-1}(\Omega)$ is an (isometric) isomorphism (cf. [6]), we see from the first equation in system (3.16) that the trivial branch of solutions $\{\{0,\lambda\}\}$ is isolated, i.e., for no λ along the trivial branch can $\dot{u} \neq 0$, so we cannot possibly have a bifurcation from this branch. Thus, we cannot hope to continue along a nontrivial solution branch starting from a trivial solution. Notice that the second equation in system (3.16) (or (3.15)) is equivalent to $(\dot{u},\ddot{u}) + \dot{\lambda}\ddot{\lambda} = 0$, which is simply a statement of the fact that the (*unit*) *tangent vector* $\{\dot{u},\dot{\lambda}\}^t$ is orthogonal to the (*principal*) *normal vector* $\{\ddot{u},\ddot{\lambda}\}^t$ in $H_0^1(\Omega) \times R$ along the solution branch, which is always true.

Since there is no nontrivial solution branch bifurcating from the trivial solution branch, we still need an initializing solution for the arclength continuation process along a nontrivial branch. In this case, however, even if we knew a solution of (3.12) on a nontrivial branch, then we would have the solution satisfying the unit L^4 norm constraint (1.3) as well and there would be no need to continue further along this branch. To see this, suppose $\{u,\lambda\}$ is a known nontrivial solution of (3.12). Take $\alpha = \|u\|_{0,4}$ and define $\{u_\alpha,\lambda_\alpha\} := \{u/\alpha, \alpha^2\lambda\}$. Then it is easy to verify that $\{u_\alpha,\lambda_\alpha\}$ satisfies the original L^4 norm constrained eigenproblem (1.1)–(1.3).

Remark 3.2. For problem (3.12) (and similar problems), we can obtain some qualitative information about the behavior of the solution set simply by looking at a one dimensional analog having the same differential behavior in the state variable as the infinite dimensional problem. Since $-\Delta u$ is linear in u, we can model this term in the one dimensional case with a linear term and therefore consider the solution sets for

$$x - \lambda x^3 = 0. \tag{3.17}$$

Other than the trivial solution branch $\{\{0,\lambda\}\}$, we see that nontrivial solution branches satisfy $x^2 = 1/\lambda$, and therefore there are no bifurcations from the trivial branch except at infinity.

At this point, it would seem that we have a major dilemma when it comes to using arclength continuation, as it is currently formulated, for solving problem (3.12)–(3.13) directly. On the one hand, we need to have a known solution $\{u_0, \lambda_0\}$ on a nontrivial branch of solutions to initialize the continuation process, but if we had such a solution, no continuation would be necessary because, modulo an appropriate normalization, the original problem would be solved.

To overcome this dilemma and salvage the technique, instead of pursuing the one-step strategy:

1. Continue along a nontrivial solution branch starting from a known nontrivial solution,

we pursue the two-step strategy:

1. Formulate and solve a perturbation of problem (3.12) that admits perturbed nontrivial solution branches bifurcating from the trivial branch and which are asymptotic to the corresponding unperturbed nontrivial solution branches.
2. On any of these perturbed solution branches, continue to a point 'close enough' to the corresponding unperturbed solution branch that it becomes possible to 'jump' from this point to a point on the unperturbed branch.

Note that in the second step of the two-step strategy, 'close enough' means inside the radius of convergence of the nonlinear solver applied to the unperturbed problem, and 'jump' means convergence to a point on the correct unperturbed branch in a single step using the 'close enough' perturbed branch point as an initial guess in the nonlinear solver.

With these ideas in mind, consider the following alternative to the system (3.12)–(3.13):

$$\tilde{S}(u, \lambda, \delta) := -\Delta u - \lambda(u^3 + \delta u) = 0 \tag{3.18}$$

$$\tilde{l}(u, \lambda, s) := (\dot{u}_0, u - u_0) + \dot{\lambda}_0(\lambda - \lambda_0) - (s - s_0) = 0 \tag{3.19}$$

where δ is a perturbation parameter (which we henceforth suppress in the notation) and the form of the perturbation was inspired by a third-order approximation to a model problem posed in the NETLIB software package `PLTMG` (see Part II, Section 3). Notice that this perturbed system has the trivial branch in common with the original system, and we have replaced the natural arclength constraint l with a *pseudo-arclength constraint* $\tilde{l}$ that depends explicitly on s and is based on a first-order approximation of l at s_0. Specifically, $\tilde{l}$ defines the length $s - s_0$ of the $\{\dot{u}_0, \dot{\lambda}_0\}$–projection (i.e. tangent projection) of the first-order difference $\{u - u_0, \lambda - \lambda_0\}$. A nice explanation of this choice can be found in [17].

Differentiating with respect to s, we have that

$$\begin{pmatrix} -\Delta - \lambda(3u^2 + \delta) & -(u^3 + \delta u) \\ (\dot{u}_0, \cdot) & \dot{\lambda}_0 \end{pmatrix} \begin{pmatrix} \dot{u} \\ \dot{\lambda} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{3.20}$$

which along the trivial branch reduces to

$$\begin{pmatrix} -\Delta - \lambda\delta & 0 \\ (\dot{u}_0, \cdot) & \dot{\lambda}_0 \end{pmatrix} \begin{pmatrix} \dot{u} \\ \dot{\lambda} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{3.21}$$

Since $\tilde{S}_u(0,\lambda) = -\Delta - \lambda\delta$ is singular whenever $\propto := \lambda\delta$ is an eigenvalue of the linear eigenproblem (2.21)–(2.22), we have bifurcation at points $\{0, \propto/\delta\}$ along the trivial branch. Let $\{w_n, \propto_n\}$ be the nth eigenpair solving the linear eigenproblem (2.21)–(2.22), where w_n is normalized to have unit L^2 norm (recall that these eigenpairs form an orthonormal basis of $L^2(\Omega)$). Then, choosing the point $\{u_0, \lambda_0\} = \{0, \propto_n/\delta\}$ along the trivial branch to initialize the continuation process along the nth bifurcating nontrivial branch and, assuming we have simple bifurcation at this point (our tacit assumption here because of the choice of our symmetry-breaking domain), we have from the first equation in (3.21) that $\dot{u}_0 = c_n w_n$, where c_n is a constant, and from the second equation that c_n and $\dot{\lambda}_0$ satisfy $c_n^2 \propto_n + \dot{\lambda}_0^2 = 1$, which is the equation of a $\{\dot{\lambda}_0, c_n\}$-ellipse on which the choice of $\dot{\lambda}_0$ determines c_n and conversely. Taking $\dot{\lambda}_0 = 0$ gives $c_n = \pm 1/\sqrt{\propto_n}$, which is the theoretically-consistent choice for initializing the continuation of the nontrivial solution branch (see next paragraph). Alternatively, we could take $\dot{\lambda}_0 = \pm 1/\sqrt{1 + \propto_n}$ (à la equation (3.10)), which gives $c_n = \pm\dot{\lambda}_0$. If we take $\dot{\lambda}_0 = \pm 1$, then $c_n = 0$, which results in an initial step along the trivial branch (not a very good start if we are trying to produce the nontrivial branch).

From an implementational point of view, our ability to, and the accuracy with which we, resolve the beginning portion of the nontrivial solution branch depends on the initial tangent choice. With this in mind, it is somewhat disconcerting that the specification of the initial tangent $\{c_n w_n, \dot{\lambda}_0\}$ can only be narrowed down to the parameterizing ellipse defined by $c_n^2 \propto_n + \dot{\lambda}_0^2 = 1$. In theory, this fact can be resolved by restricting ourselves to the *distinct* roots of the quadratic bifurcation equation defining the pair $\{\dot{\lambda}_0, c_n\}$, which in this case (following the development in [16]) can be shown to reduce to the purely bilinear equation

$$-2\delta \propto_n c_n \dot{\lambda}_0 = 0. \tag{3.22}$$

The two canonical distinct roots of this equation are $\{\dot{\lambda}_0, c_n\} = \{1, 0\}$ and $\{\dot{\lambda}_0, c_n\} = \{0, 1\}$, which define, respectively, tangents parallel and orthogonal to the trivial solution branch. This shows that the bifurcating nontrivial solution branch is orthogonal to the trivial solution branch and tangent to the linear eigenmanifold at the trivial solution point $\{0, \propto_n/\delta\}$, so in this case, tangent line prediction from this point in the orthogonal direction produces a point on the linear eigenmanifold from which we correct to the nonlinear solution branch.

Substituting these quantities computed for our specific problem into the previously stated general arclength continuation process, we obtain

Step 0: Initialization

Take

$$\{u_0, \lambda_0\} = \left\{0, \frac{\alpha_n}{\delta}\right\} \tag{3.23}$$

$$\text{and } \{\dot{u}_0, \dot{\lambda}_0\} = \{c_n(\dot{\lambda}_0) w_n, \dot{\lambda}_0\} \tag{3.24}$$

where $\{w_n, \alpha_n\}$ is the n^{th} eigenpair solving the linear eigenproblem (2.21)–(2.22).

Step 1: Continuation

Step 1.1: Tangent line prediction

Set

$$\{u_1^0, \lambda_1^0\} = \{u_0, \lambda_0\} + \{\dot{u}_0, \dot{\lambda}_0\} \Delta s_0 \tag{3.25}$$

for a suitably chosen arclength step $\Delta s_0 := s_1 - s_0$.

Step 1.2: Correction

Solve for $\{u_1, \lambda_1\} := \{u(s_1), \lambda(s_1)\}$ on the solution branch via Newton's method

$$\begin{pmatrix} -\Delta - \lambda_1^k (3(u_1^k)^2 + \delta) & -((u_1^k)^3 + \delta u_1^k) \\ (\dot{u}_0, \cdot) & \dot{\lambda}_0 \end{pmatrix} \begin{pmatrix} v_1^k \\ \alpha_1^k \end{pmatrix}$$

$$= \begin{pmatrix} -(-\Delta u_1^k - \lambda_1^k ((u_1^k)^3 + \delta u_1^k) \\ -((\dot{u}_0, u_1^k - u_0) + \dot{\lambda}_0 (\lambda_1^k - \lambda_0) - (s_1 - s_0)) \end{pmatrix} \tag{3.26}$$

$$\{u_1^{k+1}, \lambda_1^{k+1}\} = \{u_1^k, \lambda_1^k\} + \{v_1^k, \alpha_1^k\} \tag{3.27}$$

for $k = 0, 1, \ldots$.

Step 2: Update

Solve the Davidenko equations

$$\begin{pmatrix} -\Delta - \lambda_1 (3u_1^2 + \delta) & -(u_1^3 + \delta u_1) \\ (\dot{u}_0, \cdot) & \dot{\lambda}_0 \end{pmatrix} \begin{pmatrix} \dot{u}_1 \\ \dot{\lambda}_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{3.28}$$

for $\{\dot{u}_1, \dot{\lambda}_1\} := \{\dot{u}(s_1), \dot{\lambda}(s_1)\}$.

Set $s_0 = s_1$, $\{u_0, \lambda_0\} = \{u_1, \lambda_1\}$, $\{\dot{u}_0, \dot{\lambda}_0\} = \{\dot{u}_1, \dot{\lambda}_1\}$, and return to Step 1.

The continuation process proceeds along the perturbed solution branch until an attempt is made to 'jump' to the unperturbed solution branch, which entails setting $\delta = 0$ in (3.26) and attempting to correct to the unperturbed branch instead of the perturbed branch in the correction step. If the 'jump' to the unperturbed branch is successful, we normalize the solution as indicated previously so that the L^4 norm constraint is satisfied, and we are done. Otherwise, we restore δ to its previous value and proceed with the continuation process as before.

3.1.1. Newton's method correction. The particular Schur complement system of interest corresponding to system (3.26) in Step 1.2 of the continuation process is

$$\begin{pmatrix} -\Delta - \lambda_1^k(3(u_1^k)^2 + \delta) & -((u_1^k)^3 + \delta u_1^k) \\ 0 & \dot{\lambda}_0 - (\dot{u}_0, z_1^k) \end{pmatrix} \begin{pmatrix} v_1^k \\ \alpha_1^k \end{pmatrix} = -\begin{pmatrix} -\Delta u_1^k - \lambda_1^k((u_1^k)^3 + \delta u_1^k) \\ (\dot{u}_0, u_1^k - u_0) + \dot{\lambda}_0(\lambda_1^k - \lambda_0) - (s_1 - s_0)) - (\dot{u}_0, y_1^k + \lambda_1^k z_1^k) \end{pmatrix} \tag{3.29}$$

where y_1^k and z_1^k solve the system

$$-\Delta y_1^k - \lambda_1^k\big(3(u_1^k)^2 + \delta\big)y_1^k = -\Delta u_1^k \tag{3.30}$$

$$-\Delta z_1^k - \lambda_1^k\big(3(u_1^k)^2 + \delta\big)z_1^k = -\big((u_1^k)^3 + \delta u_1^k\big). \tag{3.31}$$

The solution of this system is readily seen to be

$$\alpha_1^k = \frac{(\dot{u}_0, u_1^k - (u_0 + y_1^k + \lambda_1^k z_1^k)) + \dot{\lambda}_0(\lambda_1^k - \lambda_0) - (s_1 - s_0)}{(\dot{u}_0, z_1^k) - \dot{\lambda}_0} \tag{3.32}$$

$$v_1^k = -(y_1^k + \lambda_1^k z_1^k + \alpha_1^k z_1^k) \tag{3.33}$$

provided the solutions y_1^k and z_1^k of the two elliptic problems (3.30) and (3.31) exist, and $\dot{\lambda}_0 \neq (\dot{u}_0, z_1^k)$. Because the elliptic operators in (3.30) and (3.31) are singular and indefinite, an iterative method that can handle such systems (e.g. a preconditioned minimum residual method) must be used to solve them. As an alternative to this correction methodology, we elect to use a different approach that we now describe.

3.1.2. Least-squares conjugate gradient correction. As an alternative to using Newton's method in the correction step to solve the system (3.12)–(3.13) (or in this case, the perturbed system (3.18)–(3.19)) directly, it is possible to correct via the solution of an equivalent least-squares problem to which we can apply the *conjugate gradient method*.

To begin, we note that for each $\{u, \lambda\} \in H_0^1(\Omega) \times R$, $\{\tilde{S}(u, \lambda), \tilde{l}(u, \lambda, s)\}$ is in $H^{-1}(\Omega) \times R$ (thanks again, in part, to an appropriate Sobolev imbedding result). Thus, the least-squares formulation of problem (3.18)–(3.19) is:

Find $\{u, \lambda\} \in H_0^1(\Omega) \times R$ such that

$$\tilde{J}_s(u,\lambda) \leqslant \tilde{J}_s(v,\propto) \quad \forall\{v,\propto\} \in H_0^1(\Omega)\times R \tag{3.34}$$

where the (homogeneous, in this case) *least-squares functional* $\tilde{J}_s$ is defined by

$$\tilde{J}_s(v,\propto) := \frac{1}{2}\|\tilde{S}(v,\propto)\|_{-1}^2 + \frac{1}{2}|\tilde{l}(v,\propto,s)|^2 \tag{3.35}$$

where the dual norm in $H^{-1}(\Omega)$ is defined by $\|f\|_{-1} := \sup_{w\in H_0^1(\Omega)\setminus\{0\}} |\langle f,w\rangle|/\|w\|$, and the primal norm in $H_0^1(\Omega)$ is induced by the inner product in $H_0^1(\Omega)$ defined by $(u,v) := \int_\Omega \nabla u\cdot\nabla v\,\mathrm{d}x$. Henceforth, in our discussion of the functional $\tilde{J}_s$, we suppress any explicit dependence on the arclength parameter s for notational clarity.

It is clear that solving the least-squares formulation is equivalent to solving the original problem. The only difficulty with solving it as stated lies with the explicit presence of the dual norm in the functional expression. Fortunately, we can overcome this difficulty with some powerful theory that admits a reformulation in terms of the primal norm. In particular, from the *Riesz representation theorem*, for each $f\in H^{-1}(\Omega)$, there exists a unique $v_f\in H_0^1(\Omega)$ such that $\langle f,w\rangle = (v_f,w)$ for all $w\in H_0^1(\Omega)$, and furthermore, $\|f\|_{-1} = \|v_f\|$. On the other hand, we know that $-\Delta$ is an isometric isomorphism of $H_0^1(\Omega)$ onto $H^{-1}(\Omega)$ so that we may identify f with $-\Delta v_f$. Therefore, we see that for each $f\in H^{-1}(\Omega)$, there exists a unique $v_f\in H_0^1(\Omega)$ such that $\langle -\Delta v_f, w\rangle = (v_f,w) = \langle f,w\rangle$ for all $w\in H_0^1(\Omega)$, and $\|-\Delta v_f\|_{-1} = \|v_f\| = \|f\|_{-1}$. Replacing each of $H_0^1(\Omega)$ and $H^{-1}(\Omega)$ by R, f by α, v by a, w by b, and $-\Delta$ by 1, we have the same (rather pedantic and unnecessary) argument for the scalar component.

Applying this general theory to the current setting, we take

$$\{f,\alpha\} = \{\tilde{S}(v,\propto),\tilde{l}(v,\propto,s)\},\quad \{v_f,a_\alpha\} = \{\tilde{v},\tilde{\propto}\},\quad \{w,b\} = \{\tilde{w},\tilde{\nu}\}.$$

Then we have the following reformulation of (3.35):

$$\tilde{J}(v,\propto) := \frac{1}{2}\|\tilde{v}\|^2 + \frac{1}{2}|\tilde{\propto}|^2 \tag{3.36}$$

where each of $\tilde{v}$ and $\tilde{\propto}$ is a function of $\{v,\propto\}$ through

$$(\tilde{v},\tilde{w}) = \langle \tilde{S}(v,\propto),\tilde{w}\rangle \;=\; \langle -\Delta v - \propto(v^3+\delta v),\tilde{w}\rangle \quad \forall\tilde{w}\in H_0^1(\Omega) \tag{3.37}$$

$$\tilde{\propto} = \tilde{l}(v,\propto,s) \;=\; (\dot{u}_0, v-u_0) + \dot{\lambda}_0(\propto-\lambda_0) - (s-s_0). \tag{3.38}$$

The least-squares problem, posed in the primal norm, may be solved using a very efficient quadratic solver, namely the *conjugate gradient method*, which gives the following alternative correction step for the arclength continuation process:

Step 1.2: Correction

Step 1.2.0: Initialize the conjugate gradient direction

$$\text{Solve } (g_u^0, w) = \langle \tilde{J}'_u(u_1^0, \lambda_1^0), w \rangle \quad \forall\, w \in H_0^1(\Omega) \tag{3.39}$$

$$\text{set } g_\lambda^0 = \tilde{J}'_\lambda(u_1^0, \lambda_1^0) \tag{3.40}$$

$$\text{and take } \{v_1^0, \propto_1^0\} = \{g_u^0, g_\lambda^0\} \tag{3.41}$$

where $\tilde{J}'_u$ and $\tilde{J}'_\lambda$ are the partial derivatives of $\tilde{J}(u, \lambda)$ with respect to u and λ, respectively.

Compute $\{u_1^{k+1}, \lambda_1^{k+1}\}$, $\{g_u^{k+1}, g_\lambda^{k+1}\}$, $(v_1^{k+1}, \propto_1^{k+1})$ from $\{u_1^k, \lambda_1^k\}$, $\{g_u^k, g_\lambda^k\}$, $\{v_1^k, \propto_1^k\}$ via

Step 1.2.1: Compute optimal step size for descent

Find ρ_k such that

$$\tilde{J}(u_1^k - \rho_k v_1^k, \lambda_1^k - \rho_k \propto_1^k) \leqslant \tilde{J}(u_1^k - \rho v_1^k, \lambda_1^k - \rho \propto_1^k) \quad \forall \rho \in R. \tag{3.42}$$

Step 1.2.2: Update and test for convergence

$$\{u_1^{k+1}, \lambda_1^{k+1}\} = \{u_1^k - \rho_k v_1^k, \lambda_1^k - \rho_k \propto_1^k\} \tag{3.43}$$

If $\tilde{J}(u_1^{k+1}, \lambda_1^{k+1}) \leqslant \varepsilon$, take $\{u_1, \lambda_1\} = \{u_1^{k+1}, \lambda_1^{k+1}\}$ and stop; else

Step 1.2.3: Update conjugate gradient direction

$$\text{Solve } (g_u^{k+1}, w) = \langle \tilde{J}'_u(u_1^{k+1}, \lambda_1^{k+1}), w \rangle \quad \forall\, w \in H_0^1(\Omega) \tag{3.44}$$

$$\text{set } g_\lambda^{k+1} = \tilde{J}'_\lambda(u_1^{k+1}, \lambda_1^{k+1}) \tag{3.45}$$

$$\text{compute } \gamma_k = \frac{(g_u^{k+1} - g_u^k, g_u^{k+1}) + (g_\lambda^{k+1} - g_\lambda^k) g_\lambda^{k+1}}{\|g_u^k\|^2 + |g_\lambda^k|^2} \tag{3.46}$$

$$\text{and take } \{v_1^{k+1}, \propto_1^{k+1}\} = \{g_u^{k+1}, g_\lambda^{k+1}\} + \gamma_k \{v_1^k, \propto_1^k\}. \tag{3.47}$$

$k \leftarrow k+1$ and return to Step 1.2.1.

For the implementation of this method, we must elaborate on two details. First, we need to compute the *Fréchet derivative* of the least squares functional $\tilde{J}(u, \lambda)$. Differentiating (3.36) (noting that, for each $\{v, \propto\}$ in $H_0^1(\Omega) \times R$, $\tilde{J}'(v, \propto) \in \mathcal{L}(H_0^1(\Omega) \times R, R)$), we have

$$\langle \tilde{J}'(v, \propto), \{w, \nu\} \rangle = \big(\tilde{v}'(v, \propto)\{w, \nu\}, \tilde{v}(v, \propto)\big) + \langle \tilde{\propto}'(v, \propto), \{w, \nu\} \rangle \tilde{\propto}(v, \propto) \tag{3.48}$$

where, from (3.37)–(3.38), we find that $\tilde{v}'(v,\alpha) \in \mathcal{L}(H_0^1(\Omega)\times R, H_0^1(\Omega))$ and $\tilde{\alpha}'(v,\alpha) \in \mathcal{L}(H_0^1(\Omega)\times R, R)$ are defined by

$$\begin{aligned}\big(\tilde{v}'(v,\alpha)\{w,\nu\},\tilde{w}\big) &= \big\langle \tilde{S}'(v,\alpha)\{w,\nu\},\tilde{w}\big\rangle \\ &= \big\langle -\Delta w - \alpha(3v^2+\delta)w - (v^3+\delta v)\nu, \tilde{w}\big\rangle \end{aligned} \tag{3.49}$$

for all $\tilde{w} \in H_0^1(\Omega)$, and

$$\big\langle \tilde{\alpha}'(v,\alpha),\{w,\nu\}\big\rangle = \big\langle \tilde{l}'_u(v,\alpha,s),w\big\rangle + \tilde{l}'_\lambda(v,\alpha,s)\nu = (\dot{u}_0,w) + \dot{\lambda}_0\nu. \tag{3.50}$$

On the other hand,

$$\big\langle \tilde{J}'(v,\alpha),(w,\nu)\big\rangle = \big\langle \tilde{J}'_u(v,\alpha),w\big\rangle + \tilde{J}'_\lambda(v,\alpha)\nu \tag{3.51}$$

for all $\{w,\nu\} \in H_0^1(\Omega)\times R$, so from (3.49)–(3.51) we deduce that the partial derivatives of $\tilde{J}$ satisfy

$$\big\langle \tilde{J}'_u(v,\alpha),w\big\rangle = \big\langle -\Delta w - \alpha(3v^2+\delta)w,\tilde{v}\big\rangle + (\dot{u}_0,w)\tilde{\alpha} \tag{3.52}$$

$$\tilde{J}'_\lambda(v,\alpha) = \big\langle -(v^3+\delta v),\tilde{v}\big\rangle + \dot{\lambda}_0\tilde{\alpha} \tag{3.53}$$

for all $\{w,\nu\} \in H_0^1(\Omega)\times R$. We use these expressions in the implementation.

Next, we need to solve the one-dimensional minimization problem in Step 1.2.1 for the optimal step size ρ_k for descent. Although there is more than one method that can be used for this, we have chosen *Newton's method*, for which we give the details now. Define $\mathbf{r} : R \to H_0^1(\Omega)\times R : \rho \mapsto \{v-\rho w, \alpha-\rho\nu\}$ and take $\phi(\rho) := \tilde{J}(\mathbf{r}(\rho))$. Taking $\{v,\alpha\} = \{u_1^k,\lambda_1^k\}$ and $\{w,\nu\} = \{v_1^k,\alpha_1^k\}$, we solve (3.42) for the optimal step size ρ_k by applying Newton's method to the derivative ϕ' in order to find the root corresponding to the (unique in this case) minimizer of ϕ, giving

$$\rho_k^{n+1} = \rho_k^n - \frac{\phi'(\rho_k^n)}{\phi''(\rho_k^n)} \tag{3.54}$$

for $n = 0,1,\ldots$ until convergence, where calculation gives

$$\phi'(\rho) = \big\langle \tilde{J}'(\mathbf{r}(\rho)),\mathbf{r}'(\rho)\big\rangle = -\big\langle \tilde{J}'(v-\rho w,\alpha-\rho\nu),\{w,\nu\}\big\rangle \tag{3.55}$$

and

$$\phi''(\rho) = \big\langle \tilde{J}''(\mathbf{r}(\rho))\mathbf{r}'(\rho),\mathbf{r}'(\rho)\big\rangle = \big\langle \tilde{J}''(v-\rho w,\alpha-\rho\nu)\{w,\nu\},\{w,\nu\}\big\rangle. \tag{3.56}$$

For initialization, we take ρ_k^0 to be the optimal descent step size found in the kth CG iteration, if it exists, during the arclength continuation process for the most recently found solution along the solution branch. If there was no kth CG iteration required for the previously found solution, we set $\rho_k^0 = 1$ (a full step in the descent direction).

The explicit form of $\phi'(\rho)$ for our problem may be found from equations (3.48)–(3.50) by replacing $\{v, \propto\}$ by $\{v-\rho w, \propto-\rho\nu\}$, which gives

$$\begin{aligned}\phi'(\rho) \;=\; &\big\langle -\Delta w-(\propto-\rho\nu)(3(v-\rho w)^2+\delta)w\\ &\qquad -((v-\rho w)^3+\delta(v-\rho w))\nu, \tilde{v}(v-\rho w, \propto-\rho\nu)\big\rangle\\ &+\big(\dot{u}_0, w\big)+\dot{\lambda}_0\nu\big)\,\tilde{\propto}(v-\rho w, \propto-\rho\nu)\end{aligned} \tag{3.57}$$

where from (3.37)–(3.38)

$$(\tilde{v}(v-\rho w, \propto-\rho\nu), \tilde{w}) = \big\langle -\Delta(v-\rho w)-(\propto-\rho\nu)((v-\rho w)^3+\delta(v-\rho w)), \tilde{w}\big\rangle \tag{3.58}$$

for all $\tilde{w}\in H_0^1(\Omega)$, and

$$\tilde{\propto}(v-\rho w, \propto-\rho\nu) = (\dot{u}_0, v-\rho w-u_0)+\dot{\lambda}_0(\propto-\rho\nu-\lambda_0)-(s-s_0). \tag{3.59}$$

To find the explicit form of $\phi''(\rho)$ for our problem, we need the second derivative mapping $\tilde{J}''$ of $\tilde{J}$ (more precisely, its action). Differentiating (3.48), we obtain

$$\begin{aligned}\big\langle \tilde{J}''(v, \propto)\{w_2, \nu_2\}, \{w_1, \nu_1\}\big\rangle \;=\; &\big(\tilde{v}''(v, \propto)\{w_2, \nu_2\}\{w_1, \nu_1\}, \tilde{v}(v, \propto)\big)\\ &+\big(\tilde{v}'(v, \propto)\{w_1, \nu_1\}, \tilde{v}'(v, \propto)\{w_2, \nu_2\}\big)\\ &+\big\langle \tilde{\propto}''(v, \propto)\{w_2, \nu_2\}, \{w_1, \nu_1\}\big\rangle\tilde{\propto}(v, \propto)\\ &+\big\langle \tilde{\propto}'(v, \propto), \{w_1, \nu_1\}\big\rangle\big\langle \tilde{\propto}'(v, \propto), \{w_2, \nu_2\}\big\rangle\end{aligned} \tag{3.60}$$

where, from (3.49)–(3.50), $\tilde{v}''(v, \propto)\in \mathcal{L}\big(H_0^1(\Omega)\times R, \mathcal{L}\big(H_0^1(\Omega)\times R, H_0^1(\Omega)\big)\big)$ and $\tilde{\propto}''(v, \propto)\in \mathcal{L}\big(H_0^1(\Omega)\times R, \mathcal{L}\big(H_0^1(\Omega)\times R, R\big)\big)$ are defined by

$$\begin{aligned}\big(\tilde{v}''(v, \propto)\{w_2, \nu_2\}\{w_1, \nu_1\}, \tilde{w}\big) \;&=\; \big\langle \tilde{S}''(v, \propto)\{w_2, \nu_2\}\{w_1, \nu_1\}, \tilde{w}\big\rangle\\ &= \big\langle -6\propto v w_2 w_1-(3v^2+\delta)w_2\nu_1-(3v^2+\delta)\nu_2 w_1, \tilde{w}\big\rangle\end{aligned} \tag{3.61}$$

for all $\tilde{w}\in H_0^1(\Omega)$ (note there is no $\nu_1\nu_2$ term because $S(v, \propto)$ is linear in $\propto$), and

$$\begin{aligned}\big\langle \tilde{\propto}''(v, \propto)\{w_2,\nu_2\}, \{w_1,\nu_1\}\big\rangle \;=\; &\big\langle \tilde{l}''_{uu}(v, \propto, s)w_2, w_1\big\rangle+\big\langle \tilde{l}''_{u\lambda}(v, \propto, s)\nu_2, w_1\big\rangle\\ &+\tilde{l}''_{\lambda u}(v, \propto, s)w_2\nu_1+\tilde{l}''_{\lambda\lambda}(v, \propto, s)\nu_2\nu_1\equiv 0.\end{aligned} \tag{3.62}$$

Taking $\{w_2, \nu_2\}=\{w_1, \nu_1\}=\{w, \nu\}$, replacing $\{v, \propto\}$ with $\{v-\rho w, \propto-\rho\nu\}$ in (3.60), and using (3.49)–(3.50) and (3.61)–(3.62), we finally obtain that

$$\begin{aligned}\phi''(\rho) \;=\; &\big\langle -6(\propto-\rho\nu)(v-\rho w)w^2-2(3(v-\rho w)^2+\delta)w\nu, \tilde{v}(v-\rho w, \propto-\rho\nu)\big\rangle\\ &+\|\tilde{v}'(v-\rho w, \propto-\rho\nu)(w, \nu)\|^2+|(\dot{u}_0, w)+\dot{\lambda}_0\nu|^2\end{aligned} \tag{3.63}$$

where $\tilde{v}(v-\rho w, \propto-\rho\nu)$ and $\tilde{\propto}(v-\rho w, \propto-\rho\nu)$ are again defined by (3.58)–(3.59).

For the results of the numerical experiments with our implementation of this method, we refer the reader to Part II, Section 3, of our article.

3.2. Newton's method applied directly to the original eigenproblem

Before presenting the formulation for our specific problem, we summarize the general framework for the damped Newton method. Let $F : X \mapsto Y$ be a C^1 map between Banach spaces such that $F(x) = 0$ has at least one solution. Under suitable regularity restrictions on F and for a suitable initial guess x^0, it can be shown that *ordinary Newton's method*

$$F'(x^k)\Delta x^k = -F(x^k) \tag{3.64}$$

$$x^{k+1} = x^k + \Delta x^k \tag{3.65}$$

converges to a solution x^* of $F(x) = 0$. Note that this method is *local* in the sense that its convergence depends on having a suitable initial guess x^0. The rationale for damping the *ordinary Newton increments* Δx^k is to remove any restrictions on x^0 and in this sense *globalize* the method. Such damping typically utilizes second order information available in the problem to restrict the sizes of the steps taken in the sequential *Newton directions* $\Delta x^k / \|\Delta x^k\|$, which are initial tangent directions to the sequential *Newton paths* defined by the sequential *Davidenko IVPs* (cf. [20, Subsection 7.5] for a summary of Davidenko's work and references)

$$F'(x(\sigma))\dot{x}(\sigma) + F(x(0)) = 0 \tag{3.66}$$

$$x(0) = x^k, \qquad x(1) = x^* \tag{3.67}$$

which in turn are derived by differentiating each link in the sequential *homotopy chain*

$$\Phi_k(x, \sigma) := F(x) - (1 - \sigma)F(x^k) \equiv 0, \quad k = 0, 1, 2, \ldots. \tag{3.68}$$

Damping the ordinary Newton increments simply involves multiplying them by corresponding *damping factors* σ_k in the interval $(0, 1]$. The derivation of theoretically-optimal damping factors, and their computationally-available estimates, is technical and for which we refer the curious reader to [8]. We simply invoke such estimates in the following general error-oriented damped Newton algorithm adapted from Deuflhard (cf. [8, Algorithm NLEQ-ERR]):

Step 0: Initialization

Guess x^0.

For $k = 0, 1, \ldots$, until convergence, compute x^{k+1} from x^k via

Step 1: Natural level function descent

Step 1.1: Compute ordinary Newton increment.

Solve

$$F'(x^k)\Delta x^k = -F(x^k). \tag{3.69}$$

Step 1.2: Test for convergence.

If $\|\Delta x^k\| \leqslant \varepsilon$, take $x^* = x^k + \Delta x^k$ and stop; else

Step 1.3: Predict Newton increment damping factor.

If $k = 0$, set $\sigma_0 \leqslant 1$;

else

Step 1.3.1: Compute a priori local trial Lipschitz constant estimate.

Define

$$[\overline{\omega}_k] := \frac{\|\overline{\Delta x}^k - \Delta x^k\|}{\sigma_{k-1}\|\Delta x^{k-1}\| \cdot \|\overline{\Delta x}^k\|}. \tag{3.70}$$

Step 1.3.2: Compute predicted damping factor from a priori local trial Kantorovich quantity estimate.

Set

$$\sigma_k = \min\left\{1, \frac{1}{[\overline{\omega}_k]\,\|\Delta x^k\|}\right\}. \tag{3.71}$$

Step 1.4: Regularity test

If $\sigma_k < \sigma_{\min}$, stop (no convergence);

else

Step 1.5: Update damped Newton iterate and compute trial simplified Newton increment.

Set

$$x^{k+1} = x^k + \sigma_k \Delta x^k \tag{3.72}$$

then solve

$$F'(x^k)\overline{\Delta x}^{k+1} = -F(x^{k+1}). \tag{3.73}$$

Step 1.6: Correct Newton increment damping factor.

Step 1.6.1: Compute a posteriori local trial Kantorovich quantity estimate.

Define

$$[h_k] := \frac{2\|\overline{\Delta x}^{k+1} - (1-\sigma_k)\Delta x^k\|}{(\sigma_k)^2\,\|\Delta x^k\|}. \tag{3.74}$$

Step 1.6.2: Restricted natural monotonicity test

If

$$\|\overline{\Delta x}^{k+1}\| > \left(1 - \frac{\sigma_k}{4}\right)\|\Delta x^k\| \tag{3.75}$$

(failed restricted monotonicity test), set

$$\overline{\sigma}_k = \min\left\{\frac{1}{2}\sigma_k, \frac{1}{[h_k]}\right\} \tag{3.76}$$

correct: set $\sigma_k = \overline{\sigma}_k$,

and return to Step 1.4;

else

(passed restricted monotonicity test) set

$$\overline{\sigma}_k = \min\left\{1, \frac{1}{[h_k]}\right\}. \tag{3.77}$$

If $\overline{\sigma}_k \geqslant 4\sigma_k$

correct: set $\sigma_k = \overline{\sigma}_k$,

and return to Step 1.5;

else

If $\overline{\sigma}_k = \sigma_k = 1$

If $\|\overline{\Delta x}^{k+1}\| \leqslant \varepsilon$

take $x^* = x^{k+1} + \overline{\Delta x}^{k+1}$ and stop.

set $x^k = x^{k+1}$, $k \leftarrow k+1$ and return.

Applying this general framework to our constrained nonlinear elliptic eigenproblem, we define

$$x = \begin{pmatrix} u \\ \lambda \end{pmatrix}, \qquad F(u,\lambda) = \begin{pmatrix} -\Delta u - \lambda u^3 \\ \int_\Omega u^4(x)\,\mathrm{d}x - 1 \end{pmatrix} \tag{3.78}$$

and a quick calculation gives

$$F'(u,\lambda)=\begin{pmatrix} -\Delta-3\lambda u^2 & -u^3 \\ 4\int_\Omega u^3(x)\,\mathrm{d}x & 0 \end{pmatrix} \tag{3.79}$$

where the iteration takes place in $H_0^1(\Omega)\times R$ using the corresponding product norm. Steps 1.1 and 1.5 take the respective forms

$$\begin{pmatrix} -\Delta-3\lambda^k\left(u^k\right)^2 & -\left(u^k\right)^3 \\ 4\int_\Omega\left(u^k(x)\right)^3\,\mathrm{d}x & 0 \end{pmatrix}\begin{pmatrix} v^k \\ \propto^k \end{pmatrix}=-\begin{pmatrix} -\Delta u^k-\lambda^k\left(u^k\right)^3 \\ \int_\Omega\left(u^k(x)\right)^4\,\mathrm{d}x-1 \end{pmatrix} \tag{3.80}$$

and

$$\begin{pmatrix} -\Delta-3\lambda^k\left(u^k\right)^2 & -\left(u^k\right)^3 \\ 4\int_\Omega\left(u^k(x)\right)^3\,\mathrm{d}x & 0 \end{pmatrix}\begin{pmatrix} \overline{v}^{k+1} \\ \overline{\propto}^{k+1} \end{pmatrix} =-\begin{pmatrix} -\Delta u^{k+1}-\lambda^{k+1}\left(u^{k+1}\right)^3 \\ \int_\Omega\left(u^{k+1}(x)\right)^4\,\mathrm{d}x-1 \end{pmatrix} \tag{3.81}$$

where

$$\begin{pmatrix} u^{k+1} \\ \lambda^{k+1} \end{pmatrix}=\begin{pmatrix} u^k \\ \lambda^k \end{pmatrix}+\sigma_k\begin{pmatrix} v^k \\ \propto^k \end{pmatrix}. \tag{3.82}$$

The solutions of systems (3.80) and (3.81) may be found easily from the corresponding *Schur complement* systems

$$\begin{pmatrix} -\Delta-3\lambda^k\left(u^k\right)^2 & -\left(u^k\right)^3 \\ 0 & -4\int_\Omega\left(u^k(x)\right)^3 z^k(x)\,\mathrm{d}x \end{pmatrix}\begin{pmatrix} v^k \\ \propto^k \end{pmatrix} =-\begin{pmatrix} -\Delta u^k-\lambda^k\left(u^k\right)^3 \\ \int_\Omega\left(u^k(x)\right)^4\,\mathrm{d}x-1-4\int_\Omega\left(u^k(x)\right)^3\left(y^k+\lambda^k z^k\right)(x)\,\mathrm{d}x \end{pmatrix} \tag{3.83}$$

and

$$\begin{pmatrix} -\Delta-3\lambda^k\left(u^k\right)^2 & -\left(u^k\right)^3 \\ 0 & -4\int_\Omega\left(u^k(x)\right)^3 z^k(x)\,\mathrm{d}x \end{pmatrix}\begin{pmatrix} \overline{v}^{k+1} \\ \overline{\propto}^{k+1} \end{pmatrix}$$

$$=-\begin{pmatrix} -\Delta u^{k+1}-\lambda^{k+1}\left(u^{k+1}\right)^3 \\ \int_\Omega\left(u^{k+1}(x)\right)^4\,\mathrm{d}x-1-4\int_\Omega\left(u^k(x)\right)^3\left(y^{k+1}+\lambda^{k+1}z^{k+1}\right)(x)\,\mathrm{d}x \end{pmatrix} \tag{3.84}$$

where, for $j = k, k+1$, y^j solves

$$-\Delta y - 3\lambda^k \left(u^k\right)^2 y = -\Delta u^j \quad \text{in } \Omega \tag{3.85}$$

$$y = 0 \quad \text{on } \Gamma \tag{3.86}$$

and z^j solves

$$-\Delta z - 3\lambda^k \left(u^k\right)^2 z = -\left(u^j\right)^3 \quad \text{in } \Omega \tag{3.87}$$

$$z = 0 \quad \text{on } \Gamma. \tag{3.88}$$

The solutions of (3.83) and (3.84) are readily seen to be

$$v^k = -\left(y^k + \lambda^k z^k + \alpha^k z^k\right) \tag{3.89}$$

and

$$\overline{v}^{k+1} = -\left(y^{k+1} + \lambda^{k+1} z^{k+1} + \overline{\alpha}^{k+1} z^k\right) \tag{3.90}$$

where

$$\alpha^k = \frac{\displaystyle\int_\Omega \left(u^k(x)\right)^4 \mathrm{d}x - 1 - 4\int_\Omega \left(u^k(x)\right)^3 \left(y^k + \lambda^k z^k\right)(x)\,\mathrm{d}x}{\displaystyle 4\int_\Omega \left(u^k(x)\right)^3 z^k(x)\,\mathrm{d}x} \tag{3.91}$$

and

$$\overline{\alpha}^{k+1} = \frac{\displaystyle\int_\Omega \left(u^{k+1}(x)\right)^4 \mathrm{d}x - 1 - 4\int_\Omega \left(u^k(x)\right)^3 \left(y^{k+1} + \lambda^{k+1} z^{k+1}\right)(x)\,\mathrm{d}x}{\displaystyle 4\int_\Omega \left(u^k(x)\right)^3 z^k(x)\,\mathrm{d}x}. \tag{3.92}$$

From this discussion, we see that the solvability of the systems (3.80) and (3.81) boils down to the solvability of the four (two each for $j = k$ and $j = k+1$) elliptic boundary value problems (3.85)–(3.86) and (3.87)–(3.88), for all k. In turn, as discussed before in Subsection 3.1.1 in the context of similar problems, the solvability of these problems hinges on the consistency of the systems as specified by the Fredholm alternative and the application of an appropriate solver.

For the results of the numerical experiments with our implementation of this method, we refer the reader to Part II, Section 4, of our article.

Acknowledgments. For inspiring this work, we would like to thank Mónica Clapp. For their interest in, and comments on, various aspects of this work, we would like to thank Herb Keller and Peter Deuflhard. Finally, for funding this work, we would like to thank NSF.

References

1. A. Aftalion and F. Pacella, Qualitative properties of nodal solutions of semilinear elliptic equations in radially symmetric domains. *Tech. Report No. 1*, C. R. Acad. Sci. Paris, 2004.
2. G. Bognár, Finite element approximation of the first eigenvalue of a nonlinear problem for some special domains. In: *Proceedings of the 6th Colloquium on QTDE*, Electronic Journal of the Qualitative Theory of Differential Equations, 2000, No. 1, pp. 1 – 11.
3. A. Castro and M. Clapp, Upper estimates for the energy of solutions of nonhomogeneous boundary value problems. In: *Proceedings of the American Mathematical Society*, American Mathematical Society, 2005, Vol. 134, No. 1, pp. 167 – 175.
4. S. Chandrasekhar, *An Introduction to the Study of Stellar Structure*. Dover Publications, Inc., 1957, corrected republication of 1st (1939) University of Chicago Press edition.
5. S-N. Chow and J. K. Hale, *Methods of Bifurcation Theory*, Vol. 251, Springer-Verlag, New York Inc., 1982.
6. R. Dautray and J.-L. Lions, Functional and variational methods. In: *Mathematical Analysis and Numerical Methods for Science and Technology*, Springer-Verlag, Berlin-Heidelberg, 1988, Vol. 2.
7. D. G. de Figueiredo and P. L. Lions, On pairs of positive solutions for a class of semilinear elliptic problems. *Indiana University Math. J.* (1985) **34**, No. 3, 591 – 605.
8. P. Deuflhard, *Newton Methods for Nonlinear Problems: Affine Invariance and Adaptive Algorithms*. Springer Series in Computational Mathematics, Springer-Verlag, Berlin-Heidelberg, 2004, No. 35.
9. L. C. Evans, Partial differential equations. *Graduate Studies in Mathematics. American Mathematical Society*, Providence, Rhode Island, 1998, No. 19.
10. F. J. Foss, II, On the numerical exact pointwise interior controllability of the scalar wave equation and solution of nonlinear elliptic eigenproblems. *PhD thesis*, University of Houston, Houston, TX, 2006.
11. R. Glowinski, *Numerical Methods for Nonlinear Variational Problems*. Springer Series in Computational Physics, Springer-Verlag, New York Inc., 1984.
12. R. Glowinski, Finite element methods for incompressible viscous flow. In: *Numerical Methods for Fluids, Part III* (Eds. P. G. Ciarlet and J.-L. Lions), North-Holland, Amsterdam, 2003.
13. R Glowinski, H. B. Keller, and L. Reinhart, Continuation-conjugate gradient methods for the least squares solution of nonlinear boundary value problems. *SIAM J. Sci. Stat. Comp.* (1985) **6**, No. 4, 793 – 832.
14. J. W. He and R. Glowinski, Neumann control of unstable parabolic systems: Numerical approach. *J. Opt. Theory Appl.* (1998) **96**, No. 1, 1 – 55.
15. J. Horák, Constrained mountain pass algorithm for the numerical solution of semilinear elliptic problems. *Numerische Mathematik* (2004) **98**, 251 – 276.
16. H. B. Keller, Numerical solution of bifurcation and nonlinear eigenvalue problems. In: *Applications of Bifurcation Theory* (Ed. P. H. Rabinowitz), Publications of the Mathematics Research Center, The University of Wisconsin at Madison, Academic Press, Inc., 1977, No. 38, pp. 359 – 384.
17. H. B. Keller, Global homotopies and newton methods. In: *Recent Advances in Numerical Analysis* (Eds. C. de Boor and G. H. Golub), Publications of the Mathematics Research Center, The University of Wisconsin at Madison, Academic Press, Inc., 1978, No. 41, pp. 73 – 94.
18. J.-M. Morel and L. Oswald, Remarks on the equation $-\triangle u = \lambda f(u)$ with f nondecreasing. In:

Contributions to Nonlinear Partial Differential Equations (Eds. J. I. D´iaz and P. L. Lions), Pitman Research Notes in Mathematics, Longman Scientific & Technical, Longman Group UK Ltd, 1987, Vol. II, pp. 184 – 192.

19. J. M. Neuberger and J. W. Swift, Newton's method and Morse index for semilinear elliptic PDEs. *Int. J. Bifurcation and Chaos* (2001) **11**, No. 3, 801 – 820. World Scientific Publishing Company.
20. J. M. Ortega and W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*. Classics in Applied Mathematics. Society for Industrial and Applied Mathematics, Philadelphia, 2000, No. 30.
21. X. Yao and J. Zhou, Numerical methods for computing nonlinear eigenpairs: Part I. Isohomogeneous cases. *SIAM J. Sci. Comp.* (2007) **29**, No. 4, 1355 – 1374.
22. H. Zou, On the effect of the domain geometry on uniqueness of positive solutions of $\triangle u + u^p = 0$. *Annali della Scuola Normale Superiore di Pisa* (1994) **XXI**, No. 3, 343 – 356.