STEPSIZE SELECTION IN CONTINUATION PROCEDURES AND DAMPED NEWTON'S METHOD

RANDOLPH E. BANK* AND HANS D. MITTELMANN†

Abstract. In this work we explore the relation of the problem of satisfying a sufficient decrease criterion in a damped Newton's method to the problem of stepsize selection for continuation methods. We show that there is a strong connection between the two, and that standard line search techniques used for computing damping parameters have direct application to the stepsize selection problem. The performance of the resulting continuation technique, also implemented in [2], is demonstrated for several standard example problems.

Key words. Continuation, Nonlinear Systems, Damped Newton's Method

1. Introduction. In this paper, we consider the solution of the system of nonlinear equations

$$G(u) = 0 \tag{1.1}$$

where $G : \mathcal{R}^{n+1} \to \mathcal{R}^n$ is continuously differentiable. The solution of (1.1) is formally a set of curves in $\mathcal{R}^n$; the structure of this set is often quite complicated, exhibiting *fold* or *limit points* (places where a curve turns back on itself) and *bifurcation points* (places where several curves intersect). In this paper we exclude the case of bifurcation points, and will assume that the solution of (1.1) is a single smooth curve Γ in $\mathcal{R}^n$ which does not intersect itself. In particular, we assume that there exists an $\epsilon > 0$ such that the cylindrical tube of radius ϵ about Γ does not intersect itself, and that the Jacobian matrix G_u has full rank (n) for all points lying within the tube.

A standard approach to the numerical solution of (1.1) is to use some form of continuation. In such a procedure one augments (1.1) with a normalization equation

$$N(u) - \sigma = 0 \tag{1.2}$$

where σ is the *steplength*. In this paper, we will consider two choices of $N(u)$:

$$N(u) = \dot{u}_0^t(u - u_0) \tag{1.3}$$

leading to the pseudo-arclength method of Keller [4], and

$$N(u) = \dot{u}_0^t e_k e_k^t(u - u_0) \tag{1.4}$$

where e_k is the k-th unit vector in $\mathcal{R}^n$ for some $1 \le k \le n$. This leads to a method similar to one proposed by Rheinboldt [8]. For a survey and bibliography of earlier works on continuation, see [1].

The vector u_0 is assumed to be a solution point, and $\dot{u}_0$ is the tangent vector defined by

$$G_u(u_0)\dot{u}_0 = 0 \tag{1.5}$$

*Department of Mathematics, University of California at San Diego, La Jolla, California 92093. The work of this author was supported by the Office of Naval Research under contract N00014-82K-0197.

†Department of Mathematics, Arizona State University, Tempe, Arizona 85287. The work of this author was supported by the Air Force Office of Scientific Research under contract AFOSR-84-0315.

$$\| \dot{u}_0 \| = 1 \tag{1.6}$$

Equations (1.5)-(1.6) uniquely define $\dot{u}_0$ up to the sign. The norm used in (1.6) is the standard ℓ^2 norm. The inclusion of the scalar $\dot{u}_0^t e_k$ in (1.4) emphasizes that k should be chosen such that $\dot{u}_0^t e_k \neq 0$; this can be insured, for example, by choosing k to correspond to the largest component of $\dot{u}_0$ in magnitude.

In any event, for either choice of N, the augmented system

$$H(u) = \begin{bmatrix} G(u) \\ N(u) - \sigma \end{bmatrix} = 0 \tag{1.7}$$

has a unique solution for sufficiently small stepsizes σ. A typical algorithm for tracing the solution manifold numerically consists of picking a sequence of steps σ_k, and solving (1.7) with u_0 being the current solution. Such a procedure is summarized below:

Procedure Continue

(C0) begin with initial solution u_0 and tangent vector $\dot{u}_0$
(C1) compute the step σ; predict $u \leftarrow u_0 + s\dot{u}_0$
(C2) correct u (i.e., solve (1.7))
(C3) set $u_0 \leftarrow u$; compute $\dot{u}_0$ using (1.5)-(1.6)
(C4) if done, then exit; else go to (C1)

We assume (C2) is carried out by a damped Newton iteration, using the predicted u as initial guess. Computing $\dot{u}_0$ on line (C3) is then simple since all the relevant machinery is immediately available. The predicted u on line (C1) is called the Euler predictor; for the pseudo arclength case one usually takes $s = \sigma$, while for $N(u)$ as in (1.4) one takes $s = \sigma/(\dot{u}_0^t e_k)^2$. These choices insure that the initial guess satisfies the normalization equation.

The main point of this paper is to show that there is a significant and strong connection between the stepsize selection/prediction procedures and damping strategies common to Newton's method. In particular, we suggest that widely studied and well understood damping strategies can have application to the problems of stepsize selection and prediction. We will explore this connection in a formal way in section 2. Once this connection is understood, the theory seems quite routine. In section 3 we will present some numerical illustrations.

A version of the continuation method using damped Newton's method as discussed below has been implemented in [2]. The nonlinear systems arise from finite element discretizations of elliptic boundary value problems that may depend, in a general way, on several parameters. Simple fold and birfurcation points can be handled by this package. Continuation is done through a pseudo-arclength type method in (ρ, λ)-space, where ρ is a functional of the solution and λ is one of the parameters. For more details on this and an earlier method see [6].

For section 3, we have used a simplified version of the continuation procedure employed in [2], adapted to problems of the form (1.1)-(1.4). In its present form, this simple package can handle only fold points.

2. The Relation of Damping to Stepsize Selection. Suppose that $G(u_0) = 0$ and consider the Newton linearization of $H(u) = 0$ about u_0:

$$\begin{bmatrix} G_u \\ N_u \end{bmatrix} \Delta = \begin{bmatrix} 0 \\ \sigma \end{bmatrix} \tag{2.1}$$

From $G_u\Delta = 0$ and (1.5)-(1.6), we see that Δ is a scalar multiple of $\dot{u}_0$; the second block equation determines the constant. Thus

$$\Delta = \sigma\dot{u}_0/\dot{N}_0 \tag{2.2}$$

where $\dot{N}_0 = 1$ for $N(u)$ given by (1.3) and $\dot{N}_0 = (\dot{u}_0^t e_k)^2$ for $N(u)$ given by (1.4).

We define

$$\|H\|^2 = \|G\|^2 + |N - \sigma|^2 \tag{2.3}$$

where $\|G\|$ is the ℓ^2 norm. We assume here that $G(u)$ is properly scaled relative to $N(u)$; in practice, one often must explicitly rescale G or N or use a weighted norm in place of (2.3). From (2.3) we see that

$$\|H(u_0)\| = |\sigma| \tag{2.4}$$

If we apply a damped Newton iteration to (2.1), we would set

$$u = u_0 + t\Delta \tag{2.5}$$

for some $t\in(0, 1]$. The damping parameter t is chosen such that a sufficient decrease criterion like

$$\|H(u)\|\leq(1 - t\delta)\|H(u_0)\| \tag{2.6}$$

holds for some fixed $0 < \delta < 1$. From (2.2) we see that in this case

$$u = u_0 + t\sigma\dot{u}_0/\dot{N}_0 \tag{2.7}$$

From (2.7) we see that for this first Newton step, the stepsize σ and the damping parameter t enter in exactly the same way. Within this framework, choosing σ is in some sense "predamping" the system.

Let us now assume Lipschitz continuity for G_u.

$$\|G_u(v) - G_u(w)\|\leq L\|v - w\| \tag{2.8}$$

Then

$$\|H(u)\|^2\leq(L\|\Delta\|^2/2)^2 + (\sigma(1 - t))^2 \tag{2.9}$$

and the sufficient decrease criterion would be satisfied by choosing $|\sigma| > 0$ and $t\in(0, 1]$ such that

$$(C\sigma t^2)^2 + (1 - t)^2\leq(1 - t\delta)^2 \tag{2.10}$$

where

$$C = L/(2\dot{N}_0^2)$$

Suppose that we fix t, say, let $t = 1$; then if we pick the stepsize such that

$$0 < |\sigma|\leq(1 - \delta)/C$$

we would know that the standard Euler predictor would be a good initial approximation in the sense that the sufficient decrease criterion

$$\|H(u_0 + \sigma\dot{u}_0/\dot{N}_0)\|\leq(1 - \delta)\|H(u_0)\|$$

will be satisfied. This will result in a fairly conservative step-picking strategy. Suppose we fix $t < 1$, say $t = \bar{t}$. We then pick σ such that

$$0 < |\sigma|^2 \leq ((1 - \bar{t}\delta)^2 - (1 - \bar{t})^2)/(C\bar{t}^2)^2$$

and we would have

$$\|H(u_0 + \bar{t}\sigma\dot{u}_0/\dot{N}_0)\| \leq (1 - \bar{t}\delta)\|H(u_0)\| \tag{2.11}$$

Fixing $0 < \bar{t} << 1$ generally would allow for larger steps, but also suggests that the subsequent damped Newton corrector might require more damping and possibly more iterations for convergence.

We summarize this discussion with

THEOREM 2.1. *Let $\delta \in (0, 1)$ be given. Then for any fixed $t \in (0, 1]$, there exists a $|\sigma| > 0$ such that the predicted solution*

$$u = u_0 + \alpha\dot{u}_0$$

$$\alpha = t\sigma/\dot{N}_0$$

satisfies the sufficient decrease criterion (2.6).

To illustrate the effect of damping on the corrector iteration, let us consider the two-variable system ($n = 2$ in (1.1))

$$G(u) = \mu - \lambda e^{\mu} \tag{2.12}$$

$$u = \begin{bmatrix} \mu \\ \lambda \end{bmatrix}$$

This example has a simple fold point at $\mu = 1$, $\lambda = 1/e$ as illustrated in figure 2.1.

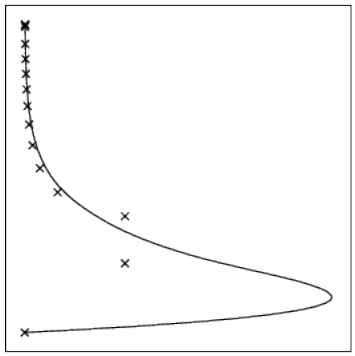

FIG. 2.1. *The solution curve for (2.12)*

At the point (μ, λ), where $\lambda = \mu e^{-\mu}$, the tangent vector is

$$\dot{u} = \begin{bmatrix} e^{\mu} \\ 1 - \mu \end{bmatrix} / \sqrt{e^{2\mu} + (1 - \mu)^2}$$

i	μ	λ	$\lambda - \mu e^{-\mu}$	$\|H\|$	t
0	$0.000E+00$	$0.000E+00$	$0.000E+00$	$0.100E+02$	
1	$0.104E+01$	$0.104E+01$	$-0.669E+00$	$0.916E+01$	0.104
2	$0.225E+01$	$0.121E+00$	$0.117E+00$	$0.783E+01$	0.135
3	$0.377E+01$	$0.121E+00$	$-0.339E-01$	$0.640E+01$	0.196
4	$0.456E+01$	$0.397E-01$	$0.813E-02$	$0.550E+01$	0.126
5	$0.533E+01$	$0.182E-01$	$0.758E-02$	$0.492E+01$	0.143
6	$0.608E+01$	$0.944E-02$	$0.449E-02$	$0.438E+01$	0.160
7	$0.675E+01$	$0.539E-02$	$0.249E-02$	$0.389E+01$	0.172
8	$0.735E+01$	$0.332E-02$	$0.139E-02$	$0.343E+01$	0.185
9	$0.789E+01$	$0.215E-02$	$0.790E-03$	$0.299E+01$	0.205
10	$0.840E+01$	$0.145E-02$	$0.447E-03$	$0.255E+01$	0.238
11	$0.888E+01$	$0.994E-03$	$0.244E-03$	$0.208E+01$	0.300
12	$0.937E+01$	$0.683E-03$	$0.119E-03$	$0.152E+01$	0.435
13	$0.993E+01$	$0.453E-03$	$0.320E-04$	$0.659E+00$	0.883
14	$0.100E+02$	$0.455E-03$	$-0.212E-07$	$0.304E-02$	0.960
15	$0.100E+02$	$0.454E-03$	$0.175E-08$	$0.381E-04$	1.000

TABLE 2.1
The damped Newton's iteration for (2.12)-(2.13)

We will use the normalization (1.4) with $k = 1$:

$$N - \sigma = (\mu - \bar{\mu}) \tag{2.13}$$

where $\bar{\mu}$ is a fixed target value (note we have dropped a constant $\sqrt{2}$ from the normalization equation). The initial guess is $u_0^t = (0, 0)$, $\dot{u}_0^t = (1, 1)/\sqrt{2}$. The target solution is $\mu = \bar{\mu}$, $\lambda = \bar{\mu}e^{-\bar{\mu}}$ In table 1, we have recorded the progress of damped Newton's method using exact line searches for the case $\bar{\mu} = 10$.

In figure 2.1, the locations of the Newton iterates are marked with an 'x' (the point (λ_1, μ_1) is off the scale of the figure to the right). From the picture, it is quite apparent that the damping strategy is forcing the iterates to follow the solution curve in some sense. To see why this is true, consider the level curves

$$\|H(u)\| = c$$

In this case

$$(\mu - \lambda e^{-\mu})^2 + (\mu - 10)^2 = c^2$$

or

$$\lambda = e^{-\mu}(\mu \pm \sqrt{c^2 - (\mu - 10)^2}) \tag{2.14}$$

The locus of the set of points such that

$$\|H(u)\| < c$$

is the interior of a "banana-shaped" region whose boundary is defined by (2.14) for $10 - c \leq \mu \leq 10 + c$. In practical terms, this means that if the current iterate has $\|H(u)\| = c$, the next iterate must lie within this region if it is to satisfy a sufficient decrease criterion; the exact location might depend on the particular criterion chosen.

In any event, the next iterate will be forced to lie close to the solution curve. Thus we see that in this example, the damped Newton procedure is obliged to produce iterates which implicitly follow the solution manifold in an approximate sense. It should be evident that this example contains the nucleus of the general case.

This example suggests that a continuation procedure which takes many small steps, explicitly following the curve, and using little or no damping in the corrector, and one taking fewer large steps but using damping with sufficient decrease, have much more in common than one might at first suspect. The damping in the latter case has roughly the effect of step-picking in the former.

We conclude this section with a sample of the type of convergence theory one could develop for a specific strategy. We seek to solve (1.7) by a damped Newton's method in which we compute a sequence of iterates u_j for $j = 0, 1, \ldots$. We assume $G(u_0) = 0$ and $N(u_0) = 0$. The sequence of iterates u_j is then defined by

$$H_u(u_j)\Delta_j = -H(u_j) \tag{2.15}$$

$$u_{j+1} = u_j + t_j\Delta_j \tag{2.16}$$

$$t_j = (1 + \kappa_j\|H(u_j)\|)^{-1} \tag{2.17}$$

This particular damping strategy is discussed in detail in [3]. The κ_j are nonnegative scalars described more fully below. We define

$$\mathcal{S} = \{u|\ \|H(u)\| \leq |\sigma|\}$$

We assume that $\mathcal{S}$ is closed, bounded, and that H_u is nonsingular on $\mathcal{S}$ with

$$\|H_u^{-1}\| \leq \gamma$$

for $u \in \mathcal{S}$.

We can infer the nonsingularity of H_u from our full rank assumption for G_u. In particular, if $H_u(u_0)x = 0$, then $G_u(u_0)x = 0$ and x must be a scalar multiple of $\dot{u}_0$. Now $N_u(u_0)x = 0$ implies that $x = 0$ for either (1.3) or (1.4) (assuming $\dot{u}_0^t e_k \neq 0$). Thus H_u must be nonsingular in some neighborhood of u_0; for $|\ \sigma\ |$ sufficiently small, $\mathcal{S}$ will be contained in this region.

Let $\mathcal{S}'$ be the closed convex ball

$$\mathcal{S}' = \{u|\ \|u\| \leq \sup_{v\in\mathcal{S}}\|v\| + \gamma|\sigma|\}$$

We assume that G is continuously differentiable and G_u satisfies (2.8) for $v, w \in \mathcal{S}'$.

THEOREM 2.2. *Let $\delta \in (0, 1)$, and let κ_j be chosen such that*

$$\kappa_0 \geq \kappa_j \geq \gamma^2 L(2(1-\delta))^{-1} - \|H(u_j)\|^{-1} \tag{2.18}$$

Then the sequence $\|H(u_j)\|$ is strictly decreasing, and the convergence $\|H(u_j)\|\to 0$ is asymptotically quadratic.

Proof. The theorem and proof are similar to those found in [3] (see Proposition 1). It is straightforward to see from Taylor's theorem and our assumptions that in general

$$\|H(u_{j+1})\| \leq (1 - t_j)\|H(u_j)\| + \gamma^2 t_j^2 L\,\|H(u_j)\|^2/2 \tag{2.19}$$

Algebraic manipulation of (2.18), using (2.17) leads to

$$(1 - t_j) + \gamma^2 t_j^2 L \, \|H(u_j)\|/2 \leq (1 - t_j\delta)$$

Thus

$$\|H(u_{j+1})\| \leq (1 - t_0\delta)\|H(u_j)\|$$

From (2.19) and $(1 - t_j) = \kappa_j t_j \, \|H(u_j)\|$, we see that

$$\|H(u_{j+1})\| \leq (\kappa_0 + \gamma^2 L/2)\|H(u_j)\|^2$$

showing that the convergence is asympotically quadratic. □ Using Theorem 2.2, one can prove convergence of the iterates u_j with the same techniques as in [3]. The scalars κ_j can be found by familiar line search techniques. One could, for instance, start with $\kappa_j = 0$; if the sufficient decrease criterion (2.6) is satisfied, then that value of κ_j is accepted. Otherwise, κ_j is systematically increased and tested until the sufficient decrease criterion is satisfied. Finally, note that by taking small steps, for instance

$$| \, \sigma \, | \leq 2(1 - \delta)/(\gamma^2 L)$$

one can take $\kappa_j = 0$, and Newton's method will satisfy the sufficient decrease criterion without damping.

3. Numerical Illustrations. The performance of the continuation procedure will be illustrated here with several examples frequently cited in the literature. For completeness, each system will be explicitly given. The results were obtained with a special version of the continuation routine of [2], adapted to problems of the form (1.1)-(1.4). Philosophically, our approach to continuation is somewhat unusual in that we do not have a procedure for automatically computing a complete solution curve. Rather, the user is required to (interactively) provide a sequence of target points. In the case of problem (1.1)-(1.4), this conists of an index k and a target value $\bar{u}_k$. The program then attempts to reach the target point using one or more continuation steps. This is somewhat analagous to the case of ordinary differential equations, where one might provide a sequence of target times where the solution is required, and the ode solver may use one or more automatically computed internal time steps to reach each target value. While this approach demands more of the user than a completely automatic procedure, it allows him to exploit whatever *a priori* knowledge he may have of the solution, and it also provides a great degree of flexibility for interactively exploring solution curves by trial and error (see [2] [6]).

The step σ is initally chosen to be $\bar{\sigma}$, the value of σ which would allow the target to be reached in one step. This value is accepted as the step if the sufficient decrease criterion (2.11) is satisfied. If $\bar{\sigma}$ fails to satisfy (2.11), it is damped by standard line search techniques until a suitable step is found. In these illustrations, we took $\bar{t} = .9$ and $\delta = 10^{-4}$.

Rather than explicitly rescaling G and N, we used a weighted norm of the form

$$\|H(u)\|^2 = \|G(u)\|^2 + s^2|N(u) - \sigma|^2$$

where

$$s = w\|G(u_0 + \bar{t}\bar{\sigma}\dot{u}_0/\dot{N}_0)\|/|\bar{\sigma}|$$

k	6	6	6	6	6	7	7
$\bar{u}_k$	1	2	5	8	11	.5	1
iterations	8	4	4	5	4	4	2
evaluations	9	7	5	6	5	5	3

TABLE 3.1
Continuation results for the trigger circuit

We took the scaler $w = 10$, which gave more weight to the normalization equation. The corrector iteration was stopped when

$$\|\Delta_j\| \leq \epsilon \|u_j\|$$

or

$$\|H(u_j)\| \leq \epsilon \|H(u_0)\|$$

For these examples, we took $\epsilon = 10^{-4}$. The calculations were done on a Microvax II workstation using single precision arithmetic.

3.1. The Trigger Circuit [7].

$$\begin{aligned}
(u_1 - u_3)/10^4 + (u_1 - u_2)/39 + (u_1 + u_7)/51 &= 0 \\
(u_2 - u_6)/10 + (u_2 - u_1)/39 + I(u_2) &= 0 \\
(u_3 - u_1)/10^4 + (u_3 - u_4)/25.5 &= 0 \\
(u_4 - u_3)/25.5 + u_4/.62 + u_4 - u_5 &= 0 \\
(u_5 - u_6)/13 + u_5 - u_4 + I(u_5) &= 0 \\
(u_6 - u_2)/10 + (u_6 - u_5)/13 + (u_6 - U(u_3 - u_1))/.201 &= 0
\end{aligned}$$

where

$$\begin{aligned}
I(u) &= 5.6 \cdot 10^{-8}(\exp(25u) - 1) \\
U(u) &= 7.65 \arctan(1962u)
\end{aligned}$$

The origin was the starting point for the continuation procedure. In table 2, the numbers of iterations, totaling 31, and function evaluations, totaling 40, are listed, while figure 3.1 shows the continuation points on the solution curve for (u_6, u_7). The number of function evaluations is larger than the number of corrector iterations by at least one, since $\|H(u)\|$ must be evaluated in order to check that the sufficient decrease criterion (2.11) was satisfied. In these examples, we did not encounter any cases where satisfying (2.11) required a line search or more than one function evaluation. Thus larger differences indicate that more than one function evaluation was needed in some corrector iterations to find a damping parameter yielding a sufficient decrease of $\|H(u)\|$.

While larger steps would have been possible, the above sequence permitted the cubic spline interpolation to provide a relatively accurate graph of the solution curve.

3.2. A Chemical Reaction Model [5].

$$\begin{aligned}
u_5(1 - u_3)E(u_1) - u_3 &= 0 \\
22u_5(1 - u_3)E(u_1) - 30u_1 &= 0 \\
u_3 - u_4 + u_5(1 - u_4)E(u_2) &= 0 \\
10u_1 - 30u_2 + 22u_5(1 - u_4)E(u_2) &= 0
\end{aligned}$$

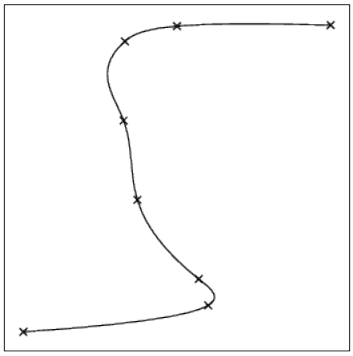

FIG. 3.1. (u_7, u_6) *graph of the solution for the trigger circuit*

k	2	2	2	5	2	2	2	2	5	4	4	5	5
$\bar{u}_k$	.1	.2	.65	.04	.6	.5	.3	.2	.03	.7	.9	.04	.06
iterations	4	4	4	7	4	3	5	2	3	3	2	8	3
evaluations	5	7	5	8	5	4	8	3	4	4	3	9	4

TABLE 3.2
Continuation results for the chemical reaction model

where

$$E(u) = \exp(10u/(1 + .01u))$$

The starting point was again the origin. Table 3 and figure 3.2 show the numerical and graphical results. A total of 52 corrector iterations and 69 function evaluations were necessary.

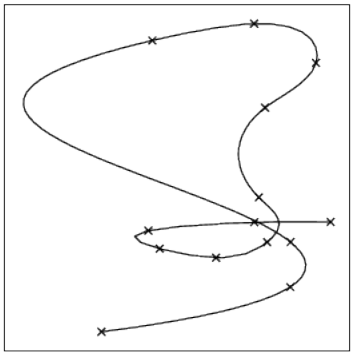

FIG. 3.2. (u_5, u_2) *graph of the solution for the chemical reaction model*

In this example, chosing steps too large could easily lead to a different part of the curve. Whether this is considered an advantage or disadvantage depends to some extent on the goal of the calculation. If the goal is to reach a particular target point as quickly as possible, then skipping intermediate portions of the curve is generally efficient. If the goal is to determine a more detailed structure of the curve, however,

some care must be taken in the selection of target points, to keep the resulting step-sizes sufficiently small. We remark that in both examples, no attempt was made to optimize the selection of target points. Our goals in the calculations were to provide the graphics program with enough data to make reasonably accurate drawings, but at the same time illustrate the robustness of a procedure using damped Newton's method with a sufficient decrease criterion and taking relatively large steps. There are undoubtedly many possible sequences of target values which could achieve these goals, some perhaps more efficiently than the ones we chose.

3.3. Fold point on a curve of fold points. Finally, we demonstrate the computation of fold points. For instance, the Bratu problem

$$\begin{aligned} 2u_1 - u_2 - \lambda E(u_1, \epsilon) &= 0 \\ -u_1 + 2u_2 - u_3 - \lambda E(u_2, \epsilon) &= 0 \\ -u_2 + 2u_3 - \lambda E(u_3, \epsilon) &= 0 \end{aligned}$$

with

$$E(u, \epsilon) = \exp(u/(1 + \epsilon u))$$

possesses two quadratic fold points with respect to λ for $\epsilon < \epsilon^*$. For $\epsilon \to \epsilon^*$, the fold points coalesce into a cubic fold point. The curve of the quadratic fold points is given by

$$\begin{aligned} 2u_1 - u_2 - u_4 E(u_1, u_8) &= 0 \\ -u_1 + 2u_2 - u_3 - u_4 E(u_2, u_8) &= 0 \\ -u_2 + 2u_3 - u_4 E(u_3, u_8) &= 0 \\ (2 - u_4 F(u_1, u_8))u_5 - u_6 &= 0 \\ -u_5 + (2 - u_4 F(u_2, u_8))u_6 - u_7 &= 0 \\ -u_6 + (2 - u_4 F(u_3, u_8))u_7 &= 0 \\ u_5^2 + u_6^2 + u_7^2 - 1 &= 0 \end{aligned}$$

with $u_4 = \lambda$, $u_8 = \epsilon$, $(u_5, u_6, u_7)^t$ a null vector of the Jacobian for the Bratu problem, and

$$F(u, \epsilon) = \partial E(u, \epsilon)/\partial u$$

We started from an initial fold point for the Bratu problem for $\epsilon = 0$, i.e.,

$$u_0^t = (.825, 1.16, .825, .212, .482, .731, .482, 0.0) \tag{3.1}$$

We then continued using $k = 2$ to $u_2 = \bar{u}_2 = 5$. On this step, which required 6 corrector iterations and 13 function evaluations, the determinant changed sign (from $-.629$ to $.707$). A secant/bisection iteration was then applied to the equation

$$\dot{\epsilon}(\sigma) = \dot{u}_8(\sigma) = 0$$

to find the step σ corresponding to the fold point, strating from u_0 as in (3.1). This required 10 secant/bisection iterations, with each iteration requiring 1 or 2 corrector iterations and 2 or 3 function evaluations. The relatively large number of secant/bisection iterations reflects the large starting interval; for example, if we had

started from the point where $u_2 = 4$, rather than $u_2 = 1.16$ as in (3.1), the number of secant/bisection iterations required would have been 5. On the other hand, the large interval we used reflects the relatively large step sizes that damped Newton method allows. At the fold point we have

$$u_0^t = (3.39, 4.79, 3.39, .315, .5, .707, .5, .248)$$

REFERENCES

[1] E. L. Allgower and K. Georg, *Predictor-corrector and simplicial methods for approximating fixed points and zero points of nonlinear mappings*, in Mathematical Programming: The State of the Art, Springer-Verlag, New York, 1983, pp. 15–56.

[2] R. E. Bank, *PLTMG Users' Guide, Edition 5.0*, Tech. Rep., Department of Mathematics, University of California, San Diego, 1988.

[3] R. E. Bank and D. J. Rose, *Global approximate Newton methods*, Numer. Math., 37 (1981), pp. 279–295.

[4] H. B. Keller, *Numerical solution of bifurcation and nonlinear eigenvalue problems*, in Applications of Bifurcation Theory, Academic Press, New York, 1977.

[5] M. Kubiček, *Algorithm 502. Dependence of solution of nonlinear systems on a parameter*, ACM Trans. Math. Software, 2 (1976), pp. 98–107.

[6] H. D. Mittelmann, *A pseudo-arclength continuation method for nonlinear eigenvalue problems*, SIAM J. Numer. Anal., 23 (1986), pp. 1007–1016.

[7] G. Pönisch and H. Schwetlick, *Ein lokal überlinear konvergentes verfahren zur bestimmung von rückkehrpunkten implizit definierter raumkurven*, Numer. Math., 38 (1982), pp. 455–466.

[8] W. C. Rheinboldt and J. V. Burkhardt, *A locally parameterized continuation process*, ACM Trans. Math. Software, 9 (1983), pp. 215–235.