SIAM J. NUMER. ANAL.
Vol. 45, No. 1, pp. 263–276

CONDITION ESTIMATES FOR PSEUDO-ARCLENGTH CONTINUATION*

K. I. DICKSON†, C. T. KELLEY†, I. C. F. IPSEN†, AND I. G. KEVREKIDIS‡

Abstract. We bound the condition number of the Jacobian in pseudo-arclength continuation problems, and we quantify the effect of this condition number on the linear system solution in a Newton-GMRES solve. Pseudo-arclength continuation solves parameter dependent nonlinear equations $G(u, \lambda) = 0$ by introducing a new parameter s, which approximates arclength, and viewing the vector $x = (u, \lambda)$ as a function of s. In this way simple fold singularities can be computed directly by solving a larger system $F(x, s) = 0$ by simple continuation in the new parameter s. It is known that the Jacobian F_x of F with respect to $x = (u, \lambda)$ is nonsingular if the path contains only regular points and simple fold singularities. We introduce a new characterization of simple folds in terms of the singular value decomposition, and we use it to derive a new bound for the norm of F_x^{-1}. We also show that the convergence rate of GMRES in a Newton step for $F(x, s) = 0$ is essentially the same as that of the original problem $G(u, \lambda) = 0$. In particular, we prove that the bounds on the degrees of the minimal polynomials of the Jacobians F_x and G_u differ by at most 2. We illustrate the effectiveness of our bounds with an example from radiative transfer theory.

Key words. pseudo-arclength continuation, singularity, GMRES, singular vectors, eigenvalues, rank-one update, turning point, simple fold, fold point, limit point

AMS subject classifications. 65H10, 65H17, 65H20, 65F10, 65F15

DOI. 10.1137/060654384

1. Introduction.

Numerical continuation is the process of solving systems of nonlinear equations $G(u, \lambda) = 0$ for various values of a real parameter λ. Here $u \in R^N$, λ is a real scalar, and $G : R^{N+1} \to R^N$. An obvious approach for implementing numerical continuation, called *parameter continuation* [11, 13, 19], traces out a solution path by repeatedly incrementing λ until the desired value of λ is reached. In each such iteration, the current solution u is used as an initial iterate for the next value of λ. Although parameter continuation is simple and intuitive, it fails at points (u, λ) where the Jacobian G_u is singular. In this paper we consider singularities which are simple folds.

The standard way to remedy the failure of parameter continuation at simple folds is to reparameterize the problem by introducing an approximate arclength parameter, s, so that both u and λ depend on s. This idea, known as *pseudo-arclength continuation* [11, 13, 19], introduces a new parameter s and treats the vector $x = (u, \lambda)$ as a function of s. We then solve a new system $F(x, s) = 0$ by parameter continuation in s. In order for this approach to succeed, the Jacobian F_x of F must be nonsingular. It is known that F_x is nonsingular at simple folds and points where G_u is nonsingular [13].

*Received by the editors March 15, 2006; accepted for publication (in revised form) August 29, 2006; published electronically January 22, 2007.
http://www.siam.org/journals/sinum/45-1/65438.html

†Center for Research in Scientific Computation and Department of Mathematics, North Carolina State University, Box 8205, Raleigh, NC 27695-8205 (kidickso@unity.ncsu.edu, Tim_Kelley@ncsu.edu, ipsen@math.ncsu.edu). The work of these authors has been partially supported by National Science Foundation grants DMS-0404537 and DMS-0209695, and Army Research Office grants DAAD19-01-1-0592, W911NF-04-1-0276, W911NF-05-1-0171, and W911NF-06-1-0096.

‡Department of Chemical Engineering and PACM, Princeton University, Princeton, NJ 08544 (yannis@princeton.edu). The work of this author was supported in part by AFOSR and an NSF/ITR grant.

Our first goal (section 3) is to quantify this nonsingularity. To this end we provide a new characterization of simple folds in terms of the singular value decomposition (SVD) of G_u. From the SVD, we derive a new bound for $\|F_x^{-1}\|_2$. This bound can be used to limit the arclength step in Newton's method. As a byproduct we obtain a refinement of Weyl's monotonicity theorem [23] for the smallest eigenvalue of a symmetric positive semidefinite matrix (section 3.1).

We also examine in section 4 how the conditioning of F_x affects the convergence of the inner GMRES [26] iteration in a Newton-GMRES solver [2, 3, 14, 15]. We show that the eigenvalue clustering of the Jacobian F_x in the reformulated problem is not much different from that of the Jacobian G_u in the original problem. This implies [4, 17] that the convergence speed of GMRES, when used as a linear solver for the Newton step, is not degraded when parameter continuation is replaced by pseudo-arclength continuation.

Finally, in section 5, we illustrate our findings with a numerical example from radiative transfer theory.

2. Background. We briefly review theory and algorithms for solving numerical continuation problems $G(u, \lambda) = 0$, where $\lambda \in R$, $u \in R^N$, and $G : R^{N+1} \to R^N$. We discuss parameter continuation in section 2.1 and pseudo-arclength continuation in section 2.2. We use the abbreviations

$$G_u \equiv \frac{\partial G}{\partial u}, \qquad G_\lambda \equiv \frac{\partial G}{\partial \lambda}.$$

2.1. Simple parameter continuation. Parameter continuation [11, 13, 19] is the simplest method for solving $G(u, \lambda) = 0$. The idea is to start at a point $\lambda = \lambda_{init}$ and solve $G(u, \lambda) = 0$ for $u(\lambda)$, say, by Newton's method. Use the solution $u(\lambda)$ as the initial iterate to solve the next problem $G(u, \lambda + d\lambda) = 0$. Algorithm **paramc** below is a simple implementation of parameter continuation from λ_{init} to $\lambda_{end} = \lambda_{init} + n\, d\lambda$ where n denotes the maximum number of continuation iterations.

---

paramc$(u, G, \lambda_{init}, \lambda_{end}, d\lambda)$
  Set $\lambda = \lambda_{init}$, $u_0 = u$
  while $\lambda \leq \lambda_{end}$ **do**
    Solve $G(u, \lambda) = 0$ with u_0 as the initial iterate to obtain $u(\lambda)$
    $u_0 = u(\lambda)$
    $\lambda = \lambda + d\lambda$
  end while

Corollary 2.1 is a consequence of the implicit function theorem [13, 22] and states that parameter continuation, as realized in Algorithm **paramc**, will succeed near a solution at which G_u is nonsingular. Parameter continuation may fail if the arc of solutions contains singular points, i.e., solutions at which G_u is singular.

COROLLARY 2.1. *Let G be Lipschitz continuously differentiable, $G(u_0, \lambda_0) = 0$, and $G_u(u_0, \lambda_0)$ be nonsingular. Then there is $\delta > 0$, which depends only on $\|G_u^{-1}(u_0, \lambda_0)\|$ and the Lipschitz constants of G_u and G_λ, such that if $|\lambda - \lambda_0| < \delta$ then Newton's method with initial iterate u_0 converges q-quadratically to the solution $u(\lambda)$ of $G(u, \lambda) = 0$, i.e.,*

$$\|u_{n+1} - u(\lambda)\| = O(\|u_n - u(\lambda)\|^2), \tag{2.1}$$

where, for $n \geq 0$,

$$u_{n+1} = u_n - G_u(u_n, \lambda)^{-1} G(u_n, \lambda).$$

Proof. Define the Lipschitz constant

$$\|G_u(u, \lambda) - G_u(v, \mu)\| \leq \gamma_G(\|u - v\| + |\lambda - \mu|).$$

Differentiating $G(u, \lambda) = 0$ with respect to λ gives

$$du/d\lambda = -G_u^{-1} G_\lambda.$$

The implicit function theorem implies that there is δ_1 such that if

$$|\lambda - \lambda_0| \leq \delta_1$$

then there is a solution arc $u(\lambda)$ defined for $|\lambda - \lambda_0| \leq \delta_1$. Since $G_u^{-1} G_\lambda$ is Lipschitz continuous, there is γ_u, which depends only on $\|G_u^{-1}(u_0, \lambda_0)\|$ and the Lipschitz constants of G_u and G_λ, such that

$$\|du/d\lambda\| = \|G_u^{-1} G_\lambda\| \leq \gamma_u.$$

A lower bound for the radius of the ball of attraction for the Newton iteration is [14]

$$\frac{1}{2\gamma_G \|G_u^{-1}(u_0, \lambda_0)\|},$$

so choosing

$$\delta = \min\left(\delta_1, \frac{1}{2\gamma_u \gamma_G \|G_u^{-1}(u_0, \lambda_0)\|}\right)$$

completes the proof. ∎

The implicit function theorem and Corollary 2.1 fail near most singular points. Our objective in this paper is to investigate the simplest class of singular points at which the implicit function theorem fails.

2.2. Pseudo-arclength continuation. Pseudo-arclength continuation [11, 13, 19] avoids the problems of Algorithm **paramc** at singular points by using an approximation of arclength parameterization. The curve in Figure 5.1, for instance, has a singularity with respect to the parameter λ. If we choose arclength s as the parameter λ, and $x = (u^T, \lambda)^T$ in place of u, we can compute the curve with simple parameter continuation. The curve in Figure 5.1 has a simple fold, which is the singularity of interest for this paper. Formally, a simple fold (or fold point, turning point, or limit point) is defined as follows [5, 13, 20, 24].

DEFINITION 2.2. *A solution* (u_0, λ_0) *of* $G(u, \lambda) = 0$ *is a simple fold if*

- $\dim(\mathrm{Ker}(G_u(u_0, \lambda_0))) = 1$ *and*
- $G_\lambda(u_0, \lambda_0) \notin \mathrm{Range}(G_u(u_0, \lambda_0))$.

To develop a pseudo-arclength continuation method, we assume that x depends smoothly on s. Then one can differentiate $G(u, \lambda) = 0$ with respect to s and obtain

$$\frac{dG(u(s), \lambda(s))}{ds} = G_u \dot{u} + G_\lambda \dot{\lambda} = 0. \tag{2.2}$$

Equivalently, one can differentiate $G(x) = 0$ and obtain $G_x\dot{x} = 0$. Here, $\dot{x}$ denotes the derivative with respect to s. Since s is arclength in the Euclidean norm,

$$\|\dot{x}\|^2 = \|\dot{u}\|^2 + |\dot{\lambda}|^2 = 1. \tag{2.3}$$

Having introduced a new parameter s, one adds an equation to $G(u, \lambda) = 0$ so that the number of equations equals the number of unknowns. To do this one introduces the extended system

$$F(x, s) = \begin{pmatrix} G(x) \\ \mathcal{N}(x, s) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{2.4}$$

The normalization equation $\mathcal{N} = 0$ is an approximation of (2.3) where

$$\mathcal{N}(x, s) = \dot{x}_0^T(x - x_0) - (s - s_0) = 0. \tag{2.5}$$

Equation (2.5) says that the new point on the path lies on a hyperplane orthogonal to the tangent vector through the current point x_0, and the intersection of that hyperplane with the tangent vector is a distance $ds = s - s_0$ from x_0.

While we prove our results using the normalization (2.5), the bounds also apply to other normalizations [7, 11, 13, 25, 27], which are asymptotically equivalent to (2.5).

Given a known point (x_0, s_0), the pseudo-arclength continuation method increments arclength by ds, and solves (2.4) with the normalization (2.5) by Newton's method with initial iterate x_0. Algorithm **psarc** is a simple implementation of pseudo-arclength continuation.

---

psarc(u, F, s_{end}, ds)
  Set $s = 0$, $x_0 = (u_0^T, \lambda_0)^T$
  while $s \leq s_{end}$ **do**
    Approximate $\dot{x}$
    Solve $F(x, s) = 0$ with fixed s and x_0 as the initial iterate to obtain $x(s)$
    $x_0 = x(s)$
    $s = s + ds$
  end while

Since pseudo-arclength continuation is just simple parameter continuation applied to F with s as the parameter, Corollary 2.1 gives conditions for the convergence of Newton's method in pseudo-arclength continuation, and we restate the corollary in terms of F for completeness.

COROLLARY 2.3. *Let the assumptions of Corollary* 2.1 *hold for* F. *Then there is* $\delta > 0$, *which depends only on* $\|F_x^{-1}(x_0, s_0)\|$ *and the Lipschitz constants of* F_x *and* x, *such that if* $|s - s_0| < \delta$ *then Newton's method with initial iterate* x_0 *converges q-quadratically to the solution.*

One consequence of Corollary 2.3 is that a bound on $\|F_x^{-1}\|$ is an important factor in bounding the arclength step. In the next section we present the main result of this paper, a new bound on $\|F_x^{-1}\|$.

3. Nonsingularity of F_x. For a solution $x_0 = (u_0, \lambda_0)$ to $G(u, \lambda) = 0$, we present an upper bound on $\|F_x^{-1}(x_0, s_0)\|$ in the case that

- $G_u(u_0, \lambda_0)$ is nonsingular or
- (u_0, λ_0) is a simple fold of $G(u, \lambda) = 0$.

In order to derive the bound, we introduce a new characterization of simple folds, which is based on the SVD of G_u. We prove the bound in section 3.2. In section 3.1 we refine Weyl's monotonicity theorem for the smallest eigenvalue of a symmetric positive semidefinite matrix, which we need for the proof.

Let

$$G_u(u,\lambda) = U\Sigma V^T$$

be an SVD of $G_u(u,\lambda)$, where

$$\Sigma = \mathrm{diag}(\sigma_1,\sigma_2,\ldots,\sigma_N), \qquad \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_N, \qquad u_N \equiv Ue_N,$$

and e_N is the last column of the $N \times N$ identity matrix. The trailing column u_N of U is a left singular vector associated with the smallest singular value σ_N. When necessary, we will make the dependence on λ or u explicit, by writing, for example, $\sigma_N(u,\lambda)$ or $u_N(u,\lambda)$.

Since the singular values are continuous functions of the elements in $G_u(u,\lambda)$, they are also continuous in λ. If

$$\sigma_{N-1} \geq \bar{\sigma} > 0$$

for all (u,λ) then the nullity of $G_u(u,\lambda)$ is at most one. If in addition $\sigma_N = 0$ then u_N spans the left nullspace of $G_u(u,\lambda)$. From the direct sum

$$\mathrm{Ker}(G_u^T(u_0,\lambda_0)) \oplus \mathrm{Range}(G_u(u_0,\lambda_0)) = R^N$$

we see that $G_\lambda(u_0,\lambda_0)$ is not in the Range $(G_u(u_0,\lambda_0))$ if and only if $G_\lambda(u_0,\lambda_0)^T u_N \neq 0$. Hence we have a new, equivalent definition of a simple fold.

DEFINITION 3.1 (simple fold via SVD). *Let (u_0,λ_0) be a solution of $G(u,\lambda)=0$, and let $u_N(u_0,\lambda_0)$ be a left singular vector of $G_u(u_0,\lambda_0)$ associated with σ_N.*

Then (u_0,λ_0) is a simple fold if

- $\sigma_{N-1}(u_0,\lambda_0) > 0$,
- $\sigma_N = 0$, *and*
- $u_N(u_0,\lambda_0)^T G_\lambda(u_0,\lambda_0) \neq 0$.

We will use Definition 3.1 to motivate the assumptions in Theorem 3.2. Suppose (u_0,λ_0) is a regular point (G_u nonsingular) or a simple fold. Since G is Lipschitz continuously differentiable, we can, by requiring $u_N^T(\lambda_0)u_N(\lambda) > 0$, for example, define u_N as a continuous function of u and λ. Hence $G_\lambda(u,\lambda)^T u_N(u,\lambda)$ is a continuous function of (u,λ). So there is $\alpha > 0$ such that for all (u,λ) sufficiently near (u_0,λ_0),

$$\max\left(\sigma_N(u,\lambda)^2, |u_N(u,\lambda)^T G_\lambda(u,\lambda)|^2 \frac{\mathrm{gap}}{\mathrm{gap}+\xi^2}\right) \geq \alpha > 0, \tag{3.1}$$

where

$$\mathrm{gap} \equiv \sigma_{N-1}(u,\lambda)^2 - \sigma_N(u,\lambda)^2 \tag{3.2}$$

and

$$\xi \equiv |u_N(u,\lambda)^T G_\lambda(u,\lambda)| + \|(I - u_N(u,\lambda)u_N(u,\lambda)^T)G_\lambda(u,\lambda)\|. \tag{3.3}$$

Inequality (3.1) is a way to quantify the statement that all points on a solution arc are either regular points or simple folds by saying that either $\sigma_N > 0$ (regular point) or the conditions in Definition 3.1 hold.

The main result of this paper is the following theorem.

THEOREM 3.2. *Let $\bar{\Omega}$ be the closure of an open subset $\Omega \in R^{N+1}$, and let G be continuously differentiable in $\bar{\Omega}$. Let $x_0 = (u_0, \lambda_0)$ in $\bar{\Omega}$ be a solution to $G(u_0, \lambda_0) = 0$, and $\mathcal{N}(x_0, s_0) = 0$ with $\|\dot{x}_0\| = 1$. Let $\tau \geq 0$ be such that $\|G_u(u_0, \lambda_0)\dot{u}_0 + G_\lambda(u_0, \lambda_0)\dot{\lambda}_0\| \leq \tau$.*

Assume that for all (u, λ) in $\bar{\Omega}$ there exists $\alpha > 0$ such that

$$\sigma_{N-1}(u, \lambda) > 0 \text{ and } \max\left(\sigma_N(u, \lambda)^2, |u_N(u, \lambda)^T G_\lambda(u, \lambda)|^2 \frac{\text{gap}}{\text{gap} + \xi^2}\right) \geq \alpha,$$

where gap *and ξ are defined by* (3.2) *and* (3.3).

If $\tau < \alpha$, then for all $x = (u, \lambda)$ in $\bar{\Omega}$, the smallest singular value $\sigma_{\min}(F_x)$ of the Jacobian F_x of $F(x, s)$ is bounded from below with

$$\sigma_{\min}(F_x) \geq \sqrt{1 - \tau \max\left\{\frac{1}{\alpha}, 1\right\}}.$$

We postpone the proof of Theorem 3.2 until section 3.1 in order to derive an auxiliary result first.

3.1. Lower bound for the smallest eigenvalue. We derive a lower bound for the smallest eigenvalue of the rank-one update $A + yy^T$, where A is a real symmetric positive semidefinite matrix of order N, and y is a real $N \times 1$ vector.

Let $\beta_1 \geq \cdots \geq \beta_N \geq 0$ be the eigenvalues of A. Weyl's monotonicity theorem [23, Theorem (10.3.1)] implies bounds for the smallest eigenvalue of $A + yy^T$:

$$\beta_N \leq \lambda_{\min}(A + yy^T) \leq \beta_{N-1}.$$

Intuitively one would expect that $\lambda_{\min}(A+yy^T)$ is larger if y is close to an eigenvector of β_N. We confirm this by deriving lower bounds for $\lambda_{\min}(A + yy^T)$ that incorporate the contribution of y in the eigenspace of β_N.

THEOREM 3.3. *Let A be an $N \times N$ real symmetric positive semidefinite matrix, u_N an eigenvector of A associated with β_N, $\|u_N\| = 1$, and $y \neq 0$ a real $N \times 1$ vector. Set $y_N \equiv u_N^T y$. Then*

$$\lambda_{\min}(A + yy^T) \geq \max\left\{\beta_N,\ y_N^2\ \frac{\text{gap}}{\text{gap} + \xi^2}\right\}, \tag{3.4}$$

where gap $\equiv \beta_{N-1} - \beta_N$ *and* $\xi \equiv |y_N| + \sqrt{\|y\|^2 - y_N^2}$.

Proof. We first show that

$$\lambda_{\min}(A + yy^T) \geq \min\left\{\beta_N + y_N^2 \frac{\text{gap}}{\text{gap} + \xi^2},\ \beta_{N-1} \frac{y_N^2}{\xi^2}\right\} \tag{3.5}$$

is a lower bound for $\lambda_{\min}(A + yy^T) = \min_{\|x\|=1} x^T(A + yy^T)x$.

Let

$$A = U \begin{pmatrix} \beta_1 & & \\ & \ddots & \\ & & \beta_N \end{pmatrix} U^T$$

be an eigendecomposition of A, and x be any real vector with $\|x\| = 1$. Partition

$$U^T x = \begin{pmatrix} \bar{x} \\ x_N \end{pmatrix}, \qquad U^T y = \begin{pmatrix} \bar{y} \\ y_N \end{pmatrix}$$

so that $\xi = |y_N| + \|\bar{y}\|$. Then

$$x^T(A + yy^T)x \geq \beta_{N-1}\|\bar{x}\|^2 + \beta_N x_N^2 + (y^T x)^2.$$

If $\|\bar{x}\| \geq |y_N|/\xi$ then

$$x^T(A + yy^T)x \geq (\beta_{N-1} y_N^2)/\xi^2,$$

which proves the second part of the bound in (3.5).

If $\|\bar{x}\| < |y_N|/\xi$ then $|y_N| - \|\bar{x}\|\xi > 0$, and it makes sense to use $|x_N| \geq 1 - \|\bar{x}\|$ in

$$|y^T x| = |y_N x_N + \bar{y}^T \bar{x}| \geq |y_N x_N| - \|\bar{x}\|\|\bar{y}\| \geq |y_N| - \|\bar{x}\|\xi.$$

Hence

$$x^T(A + yy^T)x \geq \beta_{N-1}\|\bar{x}\|^2 + \beta_N x_N^2 + (y^T x)^2 \geq \beta_N + y_N^2 + (\text{gap} + \xi^2)\|\bar{x}\|^2 - 2\xi\|\bar{x}\||y_N|.$$

This is a function of $\|\bar{x}\|$ which has a minimum at $\|\bar{x}\| = |y_N|\xi/(\text{gap} + \xi^2)$. Hence

$$x^T(A + yy^T)x \geq \beta_N + y_N^2 \frac{\text{gap}}{\text{gap} + \xi^2},$$

which proves the first part of the bound in (3.5).

With the help of (3.5) we now show the desired bound (3.4). Weyl's theorem [23, Theorem (10.3.1)] implies $\lambda_{\min}(A + yy^T) \geq \beta_N$, which proves the first part of the bound in (3.4). For the second part of the bound in (3.4), we use the fact that the eigenvalues of A are nonnegative; hence $\beta_{N-1} \geq \text{gap}$ and

$$\frac{\beta_{N-1}}{\xi^2} \geq \frac{\text{gap}}{\text{gap} + \xi^2}.$$

Substituting this into (3.5) gives the second part of the bound in (3.4):

$$\begin{aligned} \min(A + yy^T) &\geq \min\left\{ \beta_N + y_N^2 \frac{\text{gap}}{\text{gap} + \xi^2},\, y_N^2 \frac{\beta_{N-1}}{\xi^2} \right\} \\ &\geq \min\left\{ \beta_N + y_N^2 \frac{\text{gap}}{\text{gap} + \xi^2},\, y_N^2 \frac{\text{gap}}{\text{gap} + \xi^2} \right\} = y_N^2 \frac{\text{gap}}{\text{gap} + \xi^2}. \qquad \square \end{aligned}$$

The quantity gap in Theorem 3.3 is the absolute gap between the smallest and next smallest eigenvalues, and ξ is an approximation for $\|y\|$ since $\|y\| \leq \xi \leq \sqrt{2}\|y\|$. The theorem shows that $\lambda_{\min}(A + yy^T)$ is likely to be larger if y has a substantial contribution in the eigenspace of β_N.

Now we are in a position to complete the proof of Theorem 3.2.

3.2. Proof of Theorem 3.2. Define the residual

$$r \equiv G_u(u_0, \lambda_0)\dot{u}_0 + G_\lambda(u_0, \lambda_0)\dot{\lambda}_0.$$

Letting $G_u = G_u(u, \lambda)$, $G_\lambda = G_\lambda(u, \lambda)$, and $F_x = F_x(x, s)$, we have

$$F_x F_x^T = \begin{pmatrix} G_u & G_\lambda \\ \dot{u}_0^T & \dot{\lambda}_0 \end{pmatrix} \begin{pmatrix} G_u^T & \dot{u}_0 \\ G_\lambda^T & \dot{\lambda}_0 \end{pmatrix} = \begin{pmatrix} G_u G_u^T + G_\lambda G_\lambda^T & r \\ r^T & 1 \end{pmatrix}.$$

The eigenvalues of $F_x F_x^T$ are the squares of the singular values of F_x. Applying Theorem 3.3 to $G_u G_u^T + G_\lambda G_\lambda^T$ with $A = G_u G_u^T$, $y = G_\lambda$, $\beta_N = \sigma_N(u,\lambda)^2$, $\beta_{N-1} = \sigma_{N-1}^2(u,\lambda)$, and gap $= \sigma_{N-1}(u,\lambda)^2 - \sigma_N(u,\lambda)^2$ shows $\lambda_{\min}(G_u G_u^T + G_\lambda G_\lambda^T) \geq \alpha$. Hence we can write

$$\begin{pmatrix} G_u G_u^T + G_\lambda G_\lambda^T & 0 \\ 0 & 1 \end{pmatrix}^{-1} F_x F_x^T = I + E,$$

where $\|E\| \leq \tau \max\left\{\frac{1}{\alpha}, 1\right\}$. If $\tau < \min\{\alpha, 1\}$ then $\|E\| < 1$, $I + E$ is nonsingular, and

$$\frac{1}{\|(F_x F_x^T)^{-1}\|} \geq 1 - \tau \max\left\{\frac{1}{\alpha}, 1\right\}.$$

4. Newton-GMRES and eigenvalue clustering. This section discusses the performance of the inner GMRES iteration in the context of continuation with a Newton-GMRES nonlinear solver. Theorem 3.2 gives bounds on the smallest singular value of F_x in terms of the singular values of G_u. These lower bounds lead to bounds on the condition number of F_x. While the results in the previous section address conditioning, they do not directly translate into the performance of iterative methods [12, 14, 30], especially in the nonnormal case. However, we can go further to see that the eigenvalue clustering properties of the matrix F_x do not stray far from those of G_u.

Suppose the eigenvalues of G_u are nicely clustered (in the sense of [4, 17]). Even in the singular case, this would mean that the zero eigenvalue of G_u is an "outlier." We seek to show that adding the row and column does not significantly increase the number of outliers, and that we can then use the estimates in [4, 17].

The idea is that [16]

$$G_u = I + K(u) + E, \tag{4.1}$$

where K_u is a low-rank operator, say, of rank p, and E is small. We then want to write F_x in the same way, and then compare the number of outliers by comparing the ranks of the K-terms. The assumption that (4.1) holds is clearly valid if G_u is a compact perturbation of the identity; examples of this are nonlinear integral equations as well as the compact maps which are implicitly defined by the time-steppers as described in [1, 9, 10, 18, 28, 29].

Assume that E is small enough so that the eigenvalues of $I - K$ are "outliers" in the sense of [4]. Since the degree of the minimal polynomial of $I - K$ is at most $p + 1$, we have a bound for the sequence of residuals $\{r_l\}$ of the GMRES iteration of the form

$$\|r_{\hat{p}+k}\| \leq C\|E\|^k \|r_0\|, \tag{4.2}$$

where $\hat{p} \leq p + 1$ GMRES iterations are needed to remove the contribution of the outlying eigenvalues.

Theorem 4.1 states that the spectral properties of F_x are similar to those of G_u.

THEOREM 4.1. *Let the assumptions of Theorem* 3.2 *hold. Assume that* (4.1) *holds with* $\operatorname{rank}(K(u)) = p$. *Then there is* $\mathcal{K}(u)$ *having rank at most* $p + 2$ *such that*

$$\|F_x - I - \mathcal{K}(u)\| \leq \|E\|.$$

Proof. We write [16]

$$F_x = I_{(N+1)\times(N+1)} + \begin{pmatrix} K & G_\lambda \\ \dot{u}^T & \dot{\lambda} \end{pmatrix} + \begin{pmatrix} E & 0 \\ 0 & 0 \end{pmatrix}.$$

The range of

$$\mathcal{K} = \begin{pmatrix} K & G_\lambda \\ \dot{u}^T & \dot{\lambda} \end{pmatrix}$$

is

$$\begin{pmatrix} \text{Range}(K) \\ 0 \end{pmatrix} + \text{span}\left\{ \begin{pmatrix} G_\lambda \\ 0 \end{pmatrix} \right\} + \text{span}\left\{ \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\},$$

and hence the rank of $\mathcal{K}$ is at most $p+2$. ∎

So, while the eigenvalues may change, we have not increased the degree of the minimal polynomial of the main term (K versus $\mathcal{K}$) beyond $p+3$. Hence, the methods of [4] can be applied to obtain a bound like (4.2) with $\hat{p} \le p+3$.

5. Example: Chandrasekhar *H*-equation. We now present an example of a solution path containing a simple fold. The equation of interest is called the Chandrasekhar H-equation [6, 14, 21] from radiative transfer theory:

$$H(\mu) = 1 - \left(\frac{c}{2} \int_0^1 H(\nu) \frac{d\nu\mu}{\mu+\nu} \right)^{-1}. \tag{5.1}$$

The goal is to compute $H(\mu)$ for $\mu \in [0,1]$ as a function of c. There is a simple fold at $c = 1$ [21], and the same analysis shows that this is also the case for any discretization of the equation which uses a quadrature rule that integrates constant functions exactly.

In this section we use a Newton-GMRES version of pseudo-arclength continuation [8], fixing the step in arclength to $ds = .02$, using a secant predictor [13], and beginning the continuation at $c = 0$, where $H = 1$ is the solution. The vector with components all equal to one is the solution of the discrete problem as well. We discretize the integral with the composite midpoint rule using 200 nodes. A consequence of this discretization is that all scalar products of discretized functions in the GMRES solves were scaled by $1/200$. Because we do this, all the singular value results will converge as the quadrature rule is refined.

Figure 5.1 is a plot of $\|H\|_1$ against c. A fixed value of ds, as we use here, causes problems as the L^1 norm of H increases. The reason for this is that the solution develops very large derivatives, and the predictor becomes very poor. We stopped the continuation at $c = .9$ on the upper branch for that reason.

For this example we can also compute the L^1 norm as a function of c analytically, and verify the results in Figure 5.1. We can rewrite (5.1) as

$$H(\mu) = 1 + \frac{c}{2} \int_0^1 H(\mu) H(\nu) \frac{d\nu\mu}{\mu+\nu}. \tag{5.2}$$

Integrating (5.2) with respect to μ yields

$$\|H\|_1 = 1 + \frac{c}{2} \int_0^1 \int_0^1 \frac{H(\mu)H(\mu)\mu \, d\mu \, d\nu}{\mu+\nu} = 1 + \frac{c}{4} \|H\|_1^2,$$

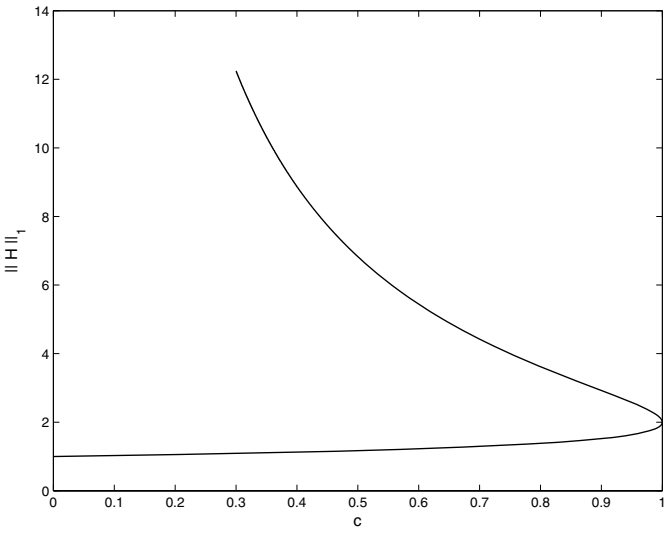

FIG. 5.1. $\|H\|_1$ *as a function of* c.

and so

$$\|H\|_1 = \frac{1 \pm \sqrt{1-c}}{c/2}. \tag{5.3}$$

As a demonstration of the result in section 3, we calculate the smallest singular value of the Jacobian matrix associated with the augmented system for the H-equation with each continuation iteration. We used the MATLAB `svds` command for this. In the language of section 3, we find $\sigma_{\min}(F_{(H,c)})$ for various c where $F_{(H,c)}$ denotes the Jacobian of

$$\begin{pmatrix} G(H,c) \\ \mathcal{N}(H,c,s) \end{pmatrix}.$$

Figure 5.2 shows that the smallest singular value of $F_{(H,c)}$ for each c stays away from zero keeping $F_{(H,c)}$ nonsingular, even at the simple fold ($c = 1$).

It is interesting to compare Figure 5.2 with a plot of the smallest singular value of G_H, which we can also compute on the path. In Figure 5.3, one can see the singularity at $c = 1$ and also see that G_H is becoming more and more poorly conditioned as the L^1 norm of H increases.

The consequences of the remarks in section 4 are that for a problem like the H-equation, which is a nonlinear compact fixed point problem, the number of GMRES iterations per Newton step should be bounded. One must take this expectation with a grain of salt because as one moves along the path, the norm of the solution increases, and so the number of outliers may increase slowly. The observations we present illustrate this.

In Figure 5.4 we plot the average number of GMRES iterations per Newton iteration as a function of c. The lower curve corresponds to the continuation from $c = 0$ to $c = 1$, and the upper from $c = 1$ to $c = .3$. The computation in this figure was done for pseudo-arclength continuation, which we compare with parameter continuation in Figure 5.5. As one moves further on the path, the predictor becomes less effective,

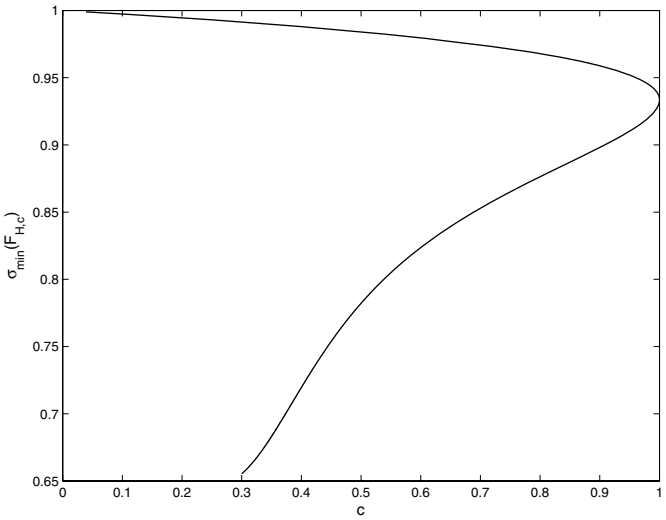

FIG. 5.2. $\sigma_{\min}(F_{(H,c)})$ *as a function of* c.

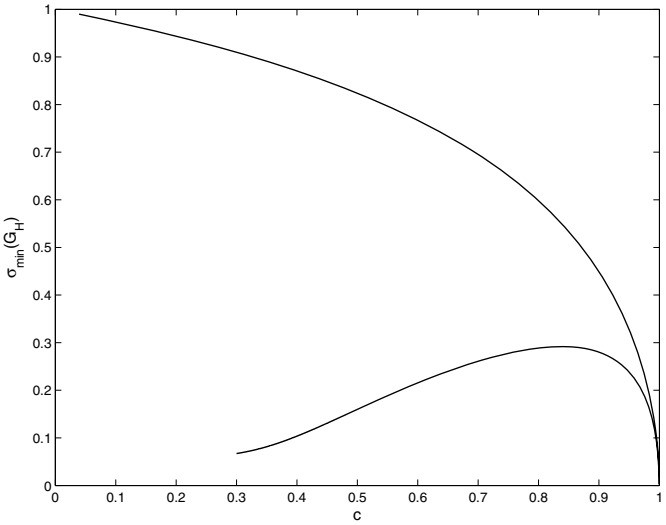

FIG. 5.3. $\sigma_{\min}(G_H)$ *as a function of* c.

and the number of Newton iterations increases. The predictor is also different for the first two points on the path, because we do not have the data we need to build the secant predictor before we have computed two points. The initial point for $c = 0$ is the vector with 1 in each component, which is the solution, so the plots begin with the first nonzero value of c.

Figure 5.5 is the result of a simple parameter continuation for each of the upper and lower branches. The lower curve is for values of $c \in [.3, .9]$, where the problem is quite easy. The linear solver takes fewer GMRES iterations per Newton iteration on this branch, and we observe that the difference in linear iterations from the lower branch in Figure 5.4 is at most 1, consistent with the theory. On the upper branch,

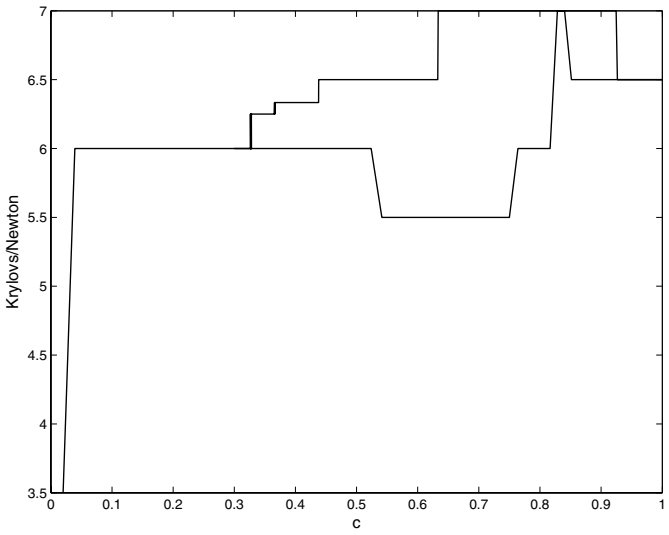

FIG. 5.4. *Krylov's per Newton: Pseudo-arclength continuation.*

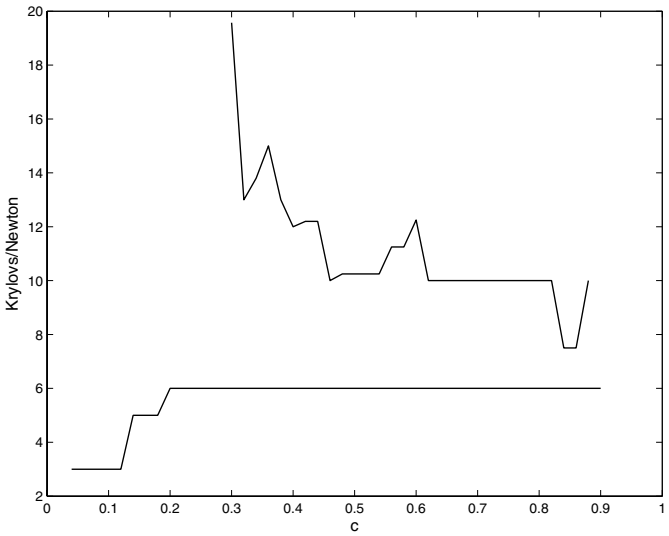

FIG. 5.5. *Krylov's per Newton: Parameter continuation in c.*

where $c \in [.3, .9]$, the performance of parameter continuation is significantly worse than that of pseudo-arclength continuation, and the linear solver performs significantly less well in the parameter continuation solver. This is consistent with the singular value results in Figure 5.3.

6. Conclusion. For simple fold singularities, we have given new bounds on the conditioning of the extended system of nonlinear equations that arise in pseudo-arclength continuation. The two bounds are a lower estimate on the smallest singular value of the Jacobian of the extended system, and an upper bound on the number of eigenvalues that lie outside a cluster of eigenvalues for the Jacobian of the origi-

nal system. The latter of these two bounds implies an upper bound on the number of GMRES iterations needed to achieve a certain termination criterion (4.2). We illustrate the bounds with a numerical experiment.

Acknowledgment. The authors thank Alastair Spence for many helpful discussions.

REFERENCES

[1] A. Armaou and I. G. Kevrekidis, *Equation-free optimal switching policies for bistable reacting systems using coarse time-steppers*, Internat. J. Robust Nonlinear Control, 15 (2005), pp. 713–726; also available online from http://arxiv.org/abs/math.OC/0410467, 2004.
[2] P. N. Brown and Y. Saad, *Hybrid Krylov methods for nonlinear systems of equations*, SIAM J. Sci. Statist. Comput., 11 (1990), pp. 450–481.
[3] P. N. Brown and Y. Saad, *Convergence theory of nonlinear Newton–Krylov algorithms*, SIAM J. Optim., 4 (1994), pp. 297–330.
[4] S. L. Campbell, I. C. F. Ipsen, C. T. Kelley, and C. D. Meyer, *GMRES and the minimal polynomial*, BIT, 36 (1996), pp. 664–675.
[5] T. F. C. Chan and H. B. Keller, *Arc-length continuation and multi-grid techniques for nonlinear elliptic eigenvalue problems*, SIAM J. Sci. Statist. Comput., 3 (1982), pp. 173–194.
[6] S. Chandrasekhar, *Radiative Transfer*, Dover, New York, 1960.
[7] E. J. Doedel and J. P. Kernévez, *AUTO: Software for Continuation and Bifurcation Problems in Ordinary Differential Equations*, Tech. rep., California Institute of Technology, Pasadena, CA, 1986.
[8] W. R. Ferng and C. T. Kelley, *Mesh independence of matrix-free methods for path following*, SIAM J. Sci. Comput., 21 (2000), pp. 1835–1850.
[9] C. W. Gear and I. G. Kevrekidis, *Telescopic projective methods for parabolic differential equations*, J. Comput. Phys., 187 (2003), pp. 95–109.
[10] C. W. Gear, I. G. Kevrekidis, and K. Theodoropoulos, *Coarse integration/bifurcation analysis via microscopic simulators: Micro-Galerkin methods*, Comp. Chem. Engrg., 26 (2002), pp. 941–963.
[11] W. J. F. Govaerts, *Numerical Methods for Bifurcations of Dynamic Equilibria*, SIAM, Philadelphia, 2000.
[12] I. Ipsen and C. Meyer, *The idea behind Krylov methods*, Amer. Math. Monthly, 105 (1998), pp. 889–99.
[13] H. B. Keller, *Lectures on Numerical Methods in Bifurcation Theory*, Tata Institute of Fundamental Research Lectures on Mathematics and Physics 79, Springer-Verlag, Berlin, 1987.
[14] C. T. Kelley, *Iterative Methods for Linear and Nonlinear Equations*, Frontiers Appl. Math. 16, SIAM, Philadelphia, 1995.
[15] C. T. Kelley, *Solving Nonlinear Equations with Newton's Method*, Fundamentals of Algorithms 1, SIAM, Philadelphia, 2003.
[16] C. T. Kelley, I. G. Kevrekidis, and L. Qiao, *Newton-Krylov Solvers for Time-Steppers*, Tech. rep. CRSC-TR04-10, Center for Research in Scientific Computation, North Carolina State University, Raleigh, NC, 2004.
[17] C. T. Kelley and Z. Q. Xue, *GMRES and integral operators*, SIAM J. Sci. Comput., 17 (1996), pp. 217–226.
[18] I. G. Kevrekidis, C. W. Gear, J. M. Hyman, P. G. Kevrekidis, O. Runborg, and C. Theodoropoulos, *Equation-free coarse-grained multiscale computation: Enabling microscopic simulators to perform system-level analysis*, Commun. Math. Sci., 1 (2003), pp. 715–762.
[19] Y. A. Kuznetsov, *Elements of Applied Bifurcation Theory*, Springer-Verlag, New York, 1998.
[20] R. Menzel and H. Schwetlick, *Zur Lösung parameterabhängiger nichtlinearer Gleichungen mit singulären Jacobi-Matrizen*, Numer. Math., 30 (1978), pp. 65–79.
[21] T. W. Mullikin, *Some probability distributions for neutron transport in a half space*, J. Appl. Probab., 5 (1968), pp. 357–374.
[22] J. M. Ortega and W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970.
[23] B. N. Parlett, *The Symmetric Eigenvalue Problem*, Prentice–Hall, Englewood Cliffs, NJ, 1980.

[24] W. C. Rheinboldt, *Solution fields of nonlinear equations and continuation methods*, SIAM J. Numer. Anal., 17 (1980), pp. 221–237.
[25] W. C. Rheinboldt, *Numerical Analysis of Parametrized Nonlinear Equations*, John Wiley and Sons, New York, 1986.
[26] Y. Saad and M. H. Schultz, *GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Statist. Comput., 7 (1986), pp. 856–869.
[27] A. G. Salinger, N. M. Bou-Rabee, R. P. Pawlowski, E. D. Wilkes, E. A. Burroughs, R. B. Lehoucq, and L. A. Romero, *LOCA* 1.0 *Library of Continuation Algorithms: Theory and Implementation Manual*, Tech. rep. SAND2002-0396, Sandia National Laboratory, Albuquergue, NM, 2002.
[28] C. I. Siettos, C. C. Pantelides, and I. G. Kevrekidis, *Enabling dynamic process simulators to perform alternative tasks: A time-stepper based toolkit for computer-aided analysis*, Ind. Eng. Chem. Res., 42 (2003), pp. 6795–6801.
[29] K. Theodoropoulos, Y.-H. Qian, and I. G. Kevrekidis, *Coarse stability and bifurcation analysis using timesteppers: A reaction diffusion example*, Proc. Natl. Acad. Sci., 97 (2000), pp. 9840–9843.
[30] L. N. Trefethen and D. Bau, *Numerical Linear Algebra*, SIAM, Philadelphia, 1997.