*Chin. Ann. of Math.*
**9B** (3) 1988

# DETECTING AND SWITCHING METHODS FOR SIMPLE BIFURCATION POINTS

Shi Miaogen (施妙根)*

**Abstract**

For a finite-dimensional equation $G(y, t)=0$, where $G: D\subset R^{n+1}\to R^n$, suppose that on its primary solution curve there is a simple bifurcation point $x^*=(y^*, t^*)$ from where a secondary solution curve is branching off. Then during the trace process of the primary curve by a continuation method, it is always necessary to locate $x^*$ and to find another point on the secondary curve for switching to trace it. This paper presents a proof of a practical criterion for detecting simple bifurcation points and constructs a simplified perturbation algorithm for switching branches. The convergence result of the algorithm is also given. Our experiments show that the new method is more effective than other perturbation methods, especially for large scale problems.

## § 1. Introduction

We consider an equation of the finite-dimensional form

$$G(y, t)=0 \tag{1.1}$$

with a vector $y\in R^n$ of state variables, a parameter $t\in R^1$ and a given map $G: D\subset R^n\times R^1\to R^n$ which is assumed to be sufficiently differentiable. Whenever feasible, we shall denote the vector $(y, t)\in R^n\times R^1$ simply by $x$. The derivatives of $G$ with repect to $x$, $y$ and $t$ are written as $G'(x)$, $G_y(x)$ and $G_t(x)$, respectively.

In many applications the interest centers on the critical points of the solution set of (1.1), that is, those solutions $x$ where $G_y(x)$ is singular. We restrict ourselves here to the special case of simple bifurcation points. Simple bifurcation points are solution points, say $x^*$, where

$$\text{rank } G'(x^*)=G_y(x^*)=n-1 \tag{1.2}$$

and a secondary solution is branching off nontangentially from the primary solution curve.

In 1971, M. G. Crandall and P. H. Rabinowitz proved a famous theorem about the solution structure in a neighborhood of simple bifurcation points (see [1]). For the finite dimensional real space considered here, this result has the following

Manuscript received April 3, 1986. Revised January 20, 1987.

* Department of Applied Mathematics, Tsinghua University, Beijing, China.

form.

**Theorem 1.** *Let $D$ be an open subset of $R^{n+1}$, $G$: $D\subset R^{n+1}\to R^n$ be a twice continuously differentiable map and $x(s)$: $J\subset R^1\to D$ with some parameter $s$ be a simple continuously differentiable arc in $D$ such that $G(x(s))=0$ for any $s\in J$, where $J$ is an open interval of $R^1$. For some $s^*\in$ int $J$ denote $x^*=x(s^*)$ and suppose that*

(a) $\dot{x}(s^*)\neq 0$,

(b) *rank* $G'(x^*)=n-1$,

(c) $N(G'(x^*))$ *can be spanned by* $\dot{x}(s^*)$ *and* $u^*$, *and* (1.3)

(d) $G''(x^*)(\dot{x}(s^*), u^*)\in R(G'(x^*))$.

*Then in a neighborhood of $x^*$ the totality of solutions of* (1.1) *forms two continuous curves intersecting only at $x^*$.*

Usually, we can use a standard continuation procedure, for example[9], to trace the primary solution curve and during the process to detect simple bifurcation points by some criteria. In Section 2 we will give a proof for a criterion which has been applied by some authors without theoretical analysis (see, e. g. [8, 9]).

Furthermore, if an approximation of a simple bifurcation point $x^*$ is located, then the problem is how to produce numerically a point on the secondary curve near the singularity from where a successful trace of that curve can be started. In recent years, a number of iteration processes for this purpose were proposed, e. g. [4, 7, 10, and 11]. We restrict ourselves here to the development of perturbation methods suggested in [4, 7]. Section 3 will give a survey of Keller–Rheinboldt method and in Section 4 we will give the representation and the convergence proof of our simplified perturbation method. Section 5 presents a brief description of the implementation of our algorithm. For numerical experiments we refer to reference [12].

## § 2. A Criterion for Detecting Simple Bifurcation Points

Theorem 1 confirms that $x^*$ is an isolated singular solution. Hence there is a number $\delta>0$ such that on the primary curve

$$\begin{matrix}\text{rank } G'(x(s))=n,\\ \dot{x}(s)\neq 0,\end{matrix}\qquad 0<|s-s^*|<\delta.$$

Regarding to $G'(x(s))\dot{x}(s)=0$, we have

$$d(s)\equiv\det\begin{pmatrix}G'(x(s))\\ (\dot{x}(s))^T\end{pmatrix}\begin{cases}=0, & s=s^*,\\ \neq 0, & 0<|s-s^*|<\delta.\end{cases}\tag{2.1}$$

Because $d(s)$ is a continuous function, if $d'(s^*)\neq 0$, then $d(s)$ must change its sign while the parameter $s$ passes through $s^*$. Actually, we have the following result.

**Theorem 2.** *In addition of assumptions of Theorem* 1, *suppose that $x(s)$ is twice*

*differentiable, then we have*

$$d'(s^*) \neq 0.$$

*Proof* For ease of notation we write

$$G'(x(s)) = \begin{bmatrix} d_1(s) \\ \vdots \\ d_n(s) \end{bmatrix}, \quad (d_i(s))^T \in R^{n+1}$$

and

$$d_i(s^*) = d_i, \ d_i'(s^*) d_i', \ i = 1, \cdots, n,$$

where

$$(d_i')^T = H_i(x^*)\dot{x}(s^*), \ i = 1, \cdots, n, \tag{2.2}$$

and $H_i(x^*)$ is the $i$-th Hessian matrix of $G(x)$ at $x^*$. From (1.3b) we have

$$\det \begin{bmatrix} d_1 \\ \vdots \\ d_n \\ (\ddot{x}(s^*))^T \end{bmatrix} = 0,$$

so that

$$d'(s^*) = \sum_{i=1}^{n} \det \begin{bmatrix} d_1 \\ \vdots \\ d_i' \\ \vdots \\ (\dot{x}(s^*))^T \end{bmatrix}.$$

Without loss of generality, we can assume that $d_1, \cdots, d_{n-1}$ are linearly independent due to (1.3b), and so there are constants $v_i (i = 1, \cdots, n-1)$; $\alpha_{ij}$, $\beta_i$ and $\nu_i (i = 1, \cdots, n;\ j = 1, \cdots, n-1)$ such that

$$d_n = \sum_{i=1}^{n-1} v_i d_i, \tag{2.3}$$

$$d_i' = \sum_{j=1}^{n-1} \alpha_{ij} d_j + \beta_i (\dot{x}(s^*))^T + \nu_i (u^*)^T. \tag{2.4}$$

The latter representation comes from (1.3b, c), that imply $R^{n+1} = \text{span}\,(d_1^T, \cdots, d_{n-1}^T, \dot{x}(s^*), u^*)$. Now it is easy to know that

$$d'(s^*) = \Big(\nu_n - \sum_{i=1}^{n-1} \nu_i v_i\Big) \det \begin{bmatrix} d_1 \\ \vdots \\ d_{n-1} \\ (u^*)^T \\ (\dot{x}(s^*))^T \end{bmatrix},$$

where the determinant is not equal to zero, so that we only need to prove

$$p \equiv \nu_n - \sum_{i=1}^{n-1} \nu_i v_i \neq 0.$$

Let $\dot{x}(s^*)$ and $u^*$ be such that

$$\|\dot{x}(s^*)\|_2 = \|u^*\|_2 = 1$$

and denote $b = (u^*)^T \dot{x}(s^*)$. Because $u^* \neq \dot{x}(s^*)$, it is obvious that $|b| < 1$. Now from (2.4) we have

$$b\beta_i+\nu_i=d_i'u^*,$$
$$\beta_i+b\nu_i=d_i'\dot{x}(s^*),$$

and hence

$$(1-b^2)\nu_i=d_i'(u^*-b\dot{x}(s^*)),\ i=1,\ \cdots,\ n. \tag{2.5}$$

From (2.5), (2.2) and the definition of $G''(x^*)$ we know that

$$(1-b^2)(\nu_1,\ \cdots,\ \nu_n)^T=G''(x^*)(\dot{x}(s^*),\ u^*-b\dot{x}(s^*)).$$

Noting

$$G''(x^*)(\dot{x}(s^*),\ \dot{x}(s^*))+G'(x^*)\ddot{x}(s^*)=0,$$

we obtain

$$(1-b^2)(\nu_1,\ \cdots,\ \nu_n)^T=G''(x^*)(\dot{x}(s^*),\ u^*)+bG'(x^*)\ddot{x}(s^*).$$

For the nullspace of $(G'(x^*))^T$, we take

$$V=(-v_1,\ \cdots,\ -v_{n-1},\ 1)^T,$$

which satisfies $V^TG'(x^*)=0$ due to (2.3). Therefore we have

$$(1-b^2)p=V^T[G''(x^*)(\dot{x}(s^*),\ u)+bG'(x^*)\ddot{x}(s^*)]=V^T[G''(x^*)(\dot{x}(s^*),\ u^*)].$$

It follows from (1.3d) and $1-b^2\neq 0$ that $p\neq 0$, and the proof of the theorem is finished.

## § 3. A Survey of Keller–Rheinboldt Method

The perturbation methods for switching branches are drived from the constructive existence theory (see, e. g. [1, 3, 5]). In [4], H. B. Keller presented the generic forms of these perturbation iterations. But unfortunately, these forms are too complicated to be used.

For the method III in [4], W. C. Rheinboldt [7] applied finite differences to avoid the evaluations of the second derivatives, and for the primary curve a parabolic approximation was introduced. In order to make the computation easier, he also changed the perturbation terms. We describe the process briefly as follows and refer for details to the cited articls.

During a continuation process it is always the case that in a neighborhood of a simple bifurcation point $x^*=(y^*,\ t^*)$ of (1.1), the primary curve can be locally parameterized by $t$, a suitably chosen coordinate, and so can be written as

$$\begin{gathered}p(t)\colon J\subset R^1\to R^n,\\ (p(t),\ t)\in D,\ G(p(t),\ t)=0,\ \forall t\in J,\\ x^*=(p(t^*),\ t^*),\ t^*\in \text{int } J,\end{gathered} \tag{3.1}$$

where $J$ is an open interval of $R^1$. Furthermore, we assume that $G$: $D\subset R^{n+1}\to R^n$ is of class $C^3$ and $p$: $J\subset R^1\to R^n$ is a simple path of class $C^2$. For ease of notation we write

$$\begin{gathered}q(t)=(p(t),\ t),\ x^*=q^*=q(t^*)\\ L=G_y(q^*),\end{gathered} \tag{3.2}$$

where $L$ has rank-deficiency one, that is,

$$\text{rank}(L)=n-1,\ Lu=L^Tv=0,\ u^Tu=v^Tv=1. \tag{3.3}$$

Denoting $r^*=\dot{q}(t^*)$ and $u^*=(u,\ 0)$, if

$$v^Tc^*\neq 0 \text{ with } c^*=G''(q^*)(r^*,\ u^*), \tag{3.4}$$

then Crandall-Rabinowitz Theorem 1 asserts that in a neighborhood of $q^*$ the totality of solutions of (1.1) consists of two continuous curves in $R^{n+1}$ intersecting only at $q^*$, one of which is the primary curve. Actually, here we have the parameter $s=t$ and hence $x(s)=q(t)$. $\dot{q}(t^*)=(\dot{p}(t^*),\ 1)\neq 0$ means (1.3a). $G'(q^*)\dot{q}(t^*)=0$ leads to $G_t(q^*)=-G_y(q^*)\dot{p}(t^*)$, which with (3.3) implies (1.3b). Obviously,

$$\dot{q}(t^*),\ u^*\in N(G'(q^*)),\ \text{rank}(\dot{q}(t^*),\ u^*)=2.$$

Hence $N(G'(q^*))=\text{span}(\dot{q}(t^*),\ u^*)$ which is (1.3c). At last, (3.3), (3.4) and $G_t(q^*)=-G_y(q^*)\dot{p}(t^*)$ imply (1.3d).

For given small $\varepsilon\neq 0$ and $\eta\neq 0$ we introduce the notations

$$\begin{aligned} &t_0=t^*+\eta,\ q^0=q(t_0),\\ &r_1=\dot{q}(t_0),\ r_2=\ddot{q}(t_0). \end{aligned} \tag{3.5}$$

Now the method proposed in [7] can be represented by the mappings

$$\begin{aligned} &\psi:\ R^1\to R^{n+1},\ \psi(s)=q^0+(\varepsilon s)r_1+\frac{1}{2}(\varepsilon s)^2r_2,\\ &\phi:\ R^{n+1}\to R^{n+1},\ \phi(x)=\psi(s)+\varepsilon u^*+\varepsilon^2\begin{pmatrix}w\\0\end{pmatrix},\\ &\forall x=\begin{pmatrix}w\\s\end{pmatrix}\in R^{n+1} \end{aligned} \tag{3.6}$$

and

$$\begin{aligned} &F:\ D_F\subset R^{n+1}\to R^{n+1},\ F(x)=VG(\phi(x)),\ \forall x\in D_F=\phi^{(-1)}(D)\\ &\text{with } V=\begin{pmatrix}I\\v^T\end{pmatrix}\in L(R^n,\ R^{n+1}), \end{aligned} \tag{3.7}$$

where, $I$ is the $n$-dimensional identity and $v\in R^n$ is defined as in (3.3). The process considered there has the form

$$\begin{aligned} &A(x^{k+1}-x^k)+F(x^k)=0,\\ &(u^*)^Tx^{k+1}=0,\ k=0,\ 1,\ \cdots,\ x^0=0, \end{aligned} \tag{3.8a}$$

with the matrix

$$A=\begin{bmatrix}\varepsilon^2L & \varepsilon c\\ 0 & \varepsilon v^Tc\end{bmatrix}\in L(R^{n+1}), \tag{3.8b}$$

where $c\doteq\varepsilon G''(q^*)(r^*,\ u^*)$ is a suitably chosen vector $c=c(\varepsilon,\ \eta)\in R^n$ such that $v^Tc\neq 0$, and the readily computable vector

$$c=G'(z^0)r_1 \text{ with } z^0=q^0+\varepsilon u^* \tag{3.9}$$

is used. Indeed,

$$c\doteq G'(q^0)r_1+\varepsilon G''(q^0)(r_1,\ u^*)\doteq\varepsilon G''(q^*)(r^*,\ u^*)$$

and the fact $v^Tc\neq 0$ will be proved in next section.

Because of rank $(A)=n$, (3.8)—(3.9) is a singular chord process in which every step involves numerical solution of

$$Lw=b,\ v^Tw=0,\ b\in R(L). \tag{3.10}$$

For solving (3.10) the following approach is employed. Starting from $w=0$ we iterate according to formulas

$$\begin{aligned} &1.\ d:=P_v(b-L_0w),\\ &2.\ \text{solve } L_0\tilde{w}=d,\\ &3.\ w:=w+P_u\tilde{w}, \end{aligned} \tag{3.11}$$

where $L_0\doteq L$ is the neighboring nonsingular matrix and $P_u=I-uu^T$, $P_v=I-vv^T$ are the orthogonal projections onto $N(L)^\perp$ and $R(L)$ respectively.

## § 4. Simplified Perturbation Method

Using a technique based on a modification [2] of Crandall–Rabinowitz proof of bifurcation, H. B. Keller suggested another perturbation approach, that is, method IV in [4]. In this case Rheinboldt's trick (3.6) with (3.5) can be applied directly. But instead of (3.7) we introduce the mapping

$$H:\ D_H\subset R^{n+1}\to R^{n+1},\ H(x)=\begin{pmatrix} G(\phi(x))\\ \varepsilon^2v^Tw\end{pmatrix}, \tag{4.1}$$

$$\forall x=\begin{pmatrix} w\\ s\end{pmatrix}\in D_H=\phi^{(-1)}(D).$$

The process considered here has the form

$$M(x^{k+1}-x^k)+H(x^k)=0,\ k=0,\ 1,\ \cdots,\ x^0=0 \tag{4.2a}$$

with the matrix

$$M=\begin{bmatrix} \varepsilon^2L & \varepsilon c\\ c^2v^T & 0\end{bmatrix}\in L(R^{n+1}), \tag{4.2b}$$

where $c=c(\varepsilon,\ \eta)\in R^n$ is chosen as in (3.9). Obviously we have $M\doteq H'(x^0)$.

To begin with theoretical analysis of our method we give the following lemma first.

**Lemma 3.** *Let $B\in L(R^{n+1})$ be of the form*

$$B=\begin{bmatrix} L & c\\ \tilde{v}^T & 0\end{bmatrix}\ \textit{with}\ \tilde{v},\ c\in R^n,\ L\in L(R^n).$$

*If* $\operatorname{rank}(L)=n-1$, $L^T\tilde{v}=0$ *and* $\tilde{v}^Tc\neq 0$, *then $B$ is nonsingular.*

*Proof* Because $\operatorname{rank}(L)=n-1$, there is a nonsingular matrix $P\in L(R^n)$ such that

$$L=P\begin{bmatrix} 0 & \\ & L_1\end{bmatrix}P^{-1}.$$

So the right and left eigenvectors of $L$ corresponding to the zero eigenvalue are $\tilde{u}=Pe_1$ and $\tilde{v}=P^{-T}e_1$, respectively. Now we assume that

$$Bx=0 \text{ for some } x=\begin{pmatrix} w \\ s \end{pmatrix}\in R^{n+1},$$

namely, $Lw+sc=0$ and $\tilde{v}^Tw=0$. Then the first equation means $\tilde{v}^TLw+s\tilde{v}^Tc=s\tilde{v}^Tc=0$. Hence $s=0$ and $w=\alpha\tilde{u}$ for a scalar $\alpha$. The second equation implies $\alpha\tilde{v}^T\tilde{u}=0$. Hence $\alpha=0$ due to $\tilde{v}^T\tilde{u}=1$. Therefore, the matrix $B$ is nonsingular.

**Corollary 4.** *For sufficiently small $\varepsilon$ and $\eta$, the matrix $M$ of (4.2b) is nonsingular.*

*Proof* By Lemma 3 we only need to verify $v^Tc\neq 0$. Let $\rho_0>0$, $\sigma_0>0$ be such that $\bar{B}(q^*, \rho_0)\subset D$ and $J_0=[t^*-\sigma_0, t^*+\sigma_0]\subset J$, and suppose by the hypothesis of continuity that under the 2-norm

$$\begin{matrix} \|G'(z)\|\leqslant\alpha_1, \ \|G''(z)\|\leqslant\alpha_2, \\ \|G'''(z)\|\leqslant\alpha_3, \end{matrix} \quad \forall z\in\bar{B}(q^*, \rho_0) \tag{4.3}$$

and

$$\|\dot{q}(t)\|\leqslant\beta_1, \ \|\ddot{q}(t)\|\leqslant\beta_2, \ \forall t\in J_0. \tag{4.4}$$

It follows from (3.4), (3.9) and $G'(q^0)r_1=0$ that

$$\begin{aligned} \|c-\varepsilon c^*\| &\leqslant \|(G'(z^0)-G'(q^0)-G''(q^0)\varepsilon u^*)r_1\| \\ &\quad +|\varepsilon|\,\|(G''(q^0)-G''(q^*))(r_1, u^*)\|+|\varepsilon|\,\|G''(q^*)(r_1-r^*, u^*)\|. \end{aligned}$$

In what follows we restrict $\varepsilon$ and $\eta$ to

$$0<|\eta|\leqslant|\varepsilon|\leqslant\varepsilon_0\leqslant\sigma_0. \tag{4.5}$$

Then by application of the mean-value theorem, for sufficiently small $\varepsilon_0$ we have

$$\|c-\varepsilon c^*\|\leqslant\nu_1\varepsilon^2, \tag{4.6}$$

where $\nu_1$ depends only on the constants of (4.3) and (4.4). Now from

$$v^Tc=\varepsilon\left[v^Tc^*+\frac{1}{\varepsilon}v^T(c-\varepsilon c^*)\right]$$

we have

$$|v^Tc|\geqslant|\varepsilon|\left[|v^Tc^*|-\frac{1}{|\varepsilon|}|v^T(c-\varepsilon c^*)|\right]\geqslant|\varepsilon|\,[|v^Tc^*|-\nu_1|\varepsilon|].$$

Therefore, for sufficiently small $\varepsilon_0$ we have $v^Tc\neq 0$ due to (3.4).

We conclude from the corollary that the process (4.2) with (3.9) is a general chord iteration which can be performed without using iteration (3.11). The convergence of our simplified perturbation method is based on the following result about the chord method which is a simple corollary to Theorem 15.5.5 of [6].

**Theorem 5.** *Let $F$: $D_F\subset R^m\to R^n$ be of class $c^1$ on $D_F$ with*

$$\|F'(x)-F'(\bar{x})\|\leqslant\nu\|x-\bar{x}\|, \ \forall x, \bar{x}\in D_F, \tag{4.7}$$

*where $\nu>0$ is a constant, and suppose that $A\in L(R^n)$ is nonsingular. If the starting point $x^0\in D_F$ is chosen such that*

$$\begin{aligned} &\text{(a)} \ \|A^{-1}\|\,\|A-F'(x^0)\|\leqslant\delta<1, \\ &\text{(b)} \ \|A^{-1}\|\nu\|x^1-x^0\|\leqslant(1-\delta)^2/2, \\ &\text{(c)} \ \bar{B}(x^0, \rho)\subset D_F \text{ for } \rho=\frac{2}{1-\delta}\|x^1-x^0\|, \end{aligned} \tag{4.8}$$

*then the iterations*

$$A(x^{k+1}-x^k)+F(x^k)=0,\ k=0,\ 1,\ \cdots$$

*remain in the set* $S=\overline{B}(x^0,\ \rho)$ *and converge to the unique solution of* $F(x)=0$ *in* $S$.

Now we give the main conclusion.

**Theorem 6.** *Let* $G\cdot D\subset R^{n+1}\to R^n$ *is of class* $C^3$ *and the primary solution of* (1.1) *can be written as* $q(t)=(p(t),\ t)$, *where* $p\colon J\subset R^1\to R^n$ *is a simple path of class* $C^2$. *Suppose that* (3.3) *and* (3.4) *hold. Then there exist constants* $\varepsilon_0>0$, $\rho>0$ *such that for any given* $\varepsilon$ *and* $\eta$ *with* $0<|\eta|\leqslant|\varepsilon|\leqslant\varepsilon_0$, *the iterations* $\{x^k\}$ *specified by* (4.2) *and* (3.9) *remain in the ball* $\overline{B}(0,\ \rho)\subset\phi^{(-1)}(D)$ *and converge to the unique solution* $\hat{x}=x(\varepsilon)$ *of* $H(x)=0$ *in that ball. Furthermore,* $\hat{z}=\phi(\hat{x})$ *is a solution of* (1.1) *and for small* $\varepsilon_0$ *the point* $\hat{z}$ *must be on the secondary solution curve.*

*Proof* At first, we confirm the condition (4.7). Regarding to

$$\phi'(x)=\left[\begin{array}{c|c} \varepsilon^2 I & \\ O^T & \varepsilon r_1+\varepsilon^2 s r_2 \end{array}\right],$$

we have from (3.5), (3.6) and (4.3)—(4.5)

$$\begin{aligned} &\|\phi'(x)\|\leqslant(\nu_2+\nu_3\|x\|)\,|\varepsilon|,\\ &\|\phi''(x)\|\leqslant\beta_2\varepsilon^2, \qquad\qquad \forall x\in R^{n+1}\\ &\|\phi(x)-q^*\|\leqslant(1+\beta_1+\nu_2\|x\|+\nu_3\|x\|^2/2)\,|\varepsilon|, \end{aligned} \tag{4.9}$$

where

$$\nu_2=\sigma_0+\beta_1,\ \nu_3=\sigma_0\beta_2.$$

Hence for any given ball $B_\rho=\overline{B}(0,\ \rho)$, $\rho>0$, we have

$$\phi(B_\rho)\subset\overline{B}(q^*,\ \rho_0),\ \forall\varepsilon_0\leqslant\varepsilon_1(\rho)\equiv\rho_0/(1+\beta_1+\nu_2\rho+\nu_3\rho^2/2), \tag{4.10}$$

where $\rho_0$ is a given constant such that $\overline{B}(q^*,\ \rho_0)\subset D$. Now let $x,\ \bar{x}\in B_\rho$ and thus $z,\ \bar{z}\in\overline{B}(q^*,\ \rho_0)$ for $z=\phi(x)$, $\bar{z}=\phi(\bar{x})$. Then for any $h\in R^{n+1}$ we have

$$(H'(x)-H'(\bar{x}))h=V^0[G'(z)\phi'(x)-G'(\bar{z})\phi'(\bar{x})]h,$$

where

$$V^0=\begin{pmatrix} I \\ O^T \end{pmatrix}\in L(R^n,\ R^{n+1}),$$

and by the same argument described in [7], we find that

$$\|H'(x)-H'(\bar{x})\|\leqslant\nu_4(\rho)\,|\varepsilon|^3\|x-\bar{x}\|,\ \forall x,\ \bar{x}\in B_\rho,$$

where $\nu_4(\rho)$ is a cubic polynomial in $\rho$ with coefficients that depend only on the constants of (4.3), (4.4) and (4.9). So (4.7) holds with $\nu=\nu_4(\rho)\,|\varepsilon|^3$.

In what follows we verify the conditions (4.8a, b, c). Let $M_0$ be the matrix $M$ with $c=\varepsilon c^*$, that is,

$$M_0=\left[\begin{array}{cc} \varepsilon^2 L & \varepsilon^2 c^* \\ \varepsilon^2 v^T & 0 \end{array}\right].$$

Clearly, we have

$$\|M_0^{-1}\|=\nu_5/\varepsilon^2 \text{ with } \nu_5=\left\|\begin{pmatrix} L & c^* \\ v^T & 0 \end{pmatrix}^{-1}\right\|,$$

and (4.6) implies $\|M-M_0\|\leqslant\nu_1|\varepsilon|^3$. Thus for small enough $\varepsilon_0$, it follows from

$$M^{-1}=M^{-1}(M_0-M)M_0^{-1}+M_0^{-1}$$

that

$$\|M^{-1}\|(1-\|M-M_0\|\,\|M_0^{-1}\|)\leqslant\|M_0^{-1}\|$$

and hence

$$\|M^{-1}\|\leqslant\frac{\nu_5/\varepsilon^2}{1-\nu_1\nu_5|\varepsilon|}\leqslant 2\nu_5/\varepsilon^2.$$

By the definition of vector $c$ we have

$$M-H'(0)=\varepsilon^2\begin{pmatrix}L-G_y(z^0) & 0\\ O^T & 0\end{pmatrix}$$

and hence by the mean-value theorem

$$\|M-H'(0)\|=\varepsilon^2\|L-G_y(z^0)\|\leqslant\alpha_2(1+\beta_1)|\varepsilon|^3.$$

After a further restriction of $\varepsilon_0$, we obtain

$$\|M^{-1}\|\,\|M-H'(0)\|\leqslant 2\alpha_2\nu_5(1+\beta_1)|\varepsilon|\leqslant\delta<1$$

which is (4.8a). For simplicity, suppose that $\delta<1/2$. Because of

$$x^0=0 \text{ and } x^1=-M^{-1}H(0),$$

we have

$$\|x^1-x^0\|\leqslant\|M^{-1}\|\,\|H(0)\|=\|M^{-1}\|\,\|V^0G(z^0)\|\leqslant\frac{2\nu_5}{\varepsilon^2}\|G(z^0)\|.$$

Noting $G(q^*)=0$ and $G'(q^*)u^*=0$, we may write

$$G(z^0)=G(z^0)-G(q^*)-G'(q^*)(\varepsilon u^*).$$

Then the mean-value theorem leads to

$$\|G(z^0)\|\leqslant\varepsilon^2\nu_6,$$

where $\nu_6$ depends only on the constants of (4.3) and (4.4). Now denoting $\nu_7=2\nu_5\nu_6$, we have

$$\|x^1-x^0\|\leqslant\nu_7.$$

The condition (4.8c) requires $\rho=4\nu_7$ and $B_\rho=\overline{B}(0,\rho)\subset D_H$. By (4.10) and $B(q^*,\rho_0)\subset D$ this is an allowable requirement provided only that $\varepsilon_0\leqslant\varepsilon_1(\rho)$, and with it we find, after a possible further decrease of $\varepsilon_0$, that

$$\frac{1}{(1-\delta)^2}\|M^{-1}\|\nu\|x^1-x^0\|\leqslant 8\nu_5\nu_7\nu_4(4\nu_7)|\varepsilon|\leqslant 1/2,$$

which is the condition (4.8b). Therefore, according to Theorem 5 we finish the first part proof of our result.

At last, suppose that $\hat{x}=\begin{pmatrix}\hat{w}\\ \hat{s}\end{pmatrix}$ is the limit point of process (4.2). Then $\hat{z}=\phi(\hat{x})$ is a solution of (1.1). Regarding to

$$\hat{t}=e_{n+1}^T\hat{z}=t_0+\varepsilon\hat{s}$$

and

$$\hat{z}-q(t_0+\varepsilon\hat{s})=\varepsilon\Big[u^*+\varepsilon\begin{pmatrix}\hat{w}\\ 0\end{pmatrix}+\frac{1}{2}\varepsilon\hat{s}^2(r_2-\ddot{q}(\xi))\Big],$$

where $0<\xi<\hat{s}$, we have

$$\|\hat{z}-q(t_0+\varepsilon\hat{s})\| \geqslant |\varepsilon|\,[1-(1+\rho\beta_2)\rho\varepsilon_0],$$

which implies that for small $\varepsilon_0$ the point $\hat{z}$ must be on the secondary curve. This completes the proof of our theroem.

## § 5. Implementation Remarks

We now give some remarks on the computational aspects of our methods.

· Determine an interval $\Delta=[t_1, t_2]$ such that $t_1<t^*<t_2$, where $q(t^*)$ is a simple bifurcation point. During the continuation process of tracing the primary curve $q(t)$, this can be done simply by monitoring the sign of the determinant $d(t)$ defined in (2.1). Because the process involves the computations of the derivative $G'(q(t))$ and the tangent vector $\dot{q}(t)$, $d(t)$ is an easy byproduct of the trace calculations (see [9]).

· Reduce the steplength $\Delta t=t_2-t_1$ by, for instance, the bisection technique to produce an approximation of $q(t^*)$, say, $q^0=(p(t_0), t_0)$. Numerical approximation of $r_1=\dot{q}(t_0)$ is a direct result of the process, and the difference quotient $(\dot{q}(t_2)-\dot{q}(t_1))/\Delta t$ based on the last interval may be expected to provide a satisfactory estimate of $r_2=\ddot{q}(t_0)$.

· The matrix $L=G_y(q^*)$ can be approximated by $L_0=G_y(q^0)$, and the eigenvectors $u$ and $v$ of $L$ can be numerically computed by applying the inverse power method to matrix $L_0$. For the test of the rank-deficiency assumption (3.3), we can restart the inverse power method with a vector orthogonal to the one originally found. Note that for any odd rank-deficiency of the matrix $L$, the determinant $d(t)$ may change its sign.

· For switching computation we give the following schematic algorithm by formulas (3.5), (3.6), (3.9), (4.1) and (4.2), where, $w$ and $s$ correspond to $s^2w^k$ and $\varepsilon s^k$, respectively; $q^0=(p^0, t_0)$, $r_1=(\dot{p}^0, 1)$ and $r_2=(\ddot{p}^0, 0)$.

1. Input$(t_0, p^0, \dot{p}^0, \ddot{p}^0, L_0, u, v, \varepsilon)$.
2. Initialize:

   $w:=0$, $s:=0$.

   $y:=p^0+\varepsilon u$, $t:=t_0$.
3. $c:=G_y(y, t)\dot{p}^0+G_t(y, t)$.
4. $\alpha:=|v^Tc|$, test whether $\alpha$ is too small.
5. Loop

(1) $h:=\begin{pmatrix} G(y, t) \\ v^Tw \end{pmatrix}$,

(2) Solve $\begin{pmatrix} L_0 & c \\ v^T & 0 \end{pmatrix}\begin{pmatrix} \Delta w \\ \Delta s \end{pmatrix}=-h$ for $\begin{pmatrix} \Delta w \\ \Delta s \end{pmatrix}$.

(3) $\begin{pmatrix} w \\ s \end{pmatrix} := \begin{pmatrix} w \\ s \end{pmatrix} + \begin{pmatrix} \Delta w \\ \Delta s \end{pmatrix}$.

(4) $y := p^0 + s\dot{p}^0 + \frac{1}{2}s^2\ddot{p}^0 + u + w$,

$t := t_0 + s$.

6. Output$(y, t)$.

In order to have some numerical experiments on switching methods, we tested a series of problems. The results were presented in [12], where a set of seven problems with a total of 31 simple bifurcation points was used. All of the switching methods selected there were divided into two classes:

(1) Methods based on calculating a rough branching direction, mainly, suggested by R. Seydel.

(2) Methods using standard perturbation arguments.

The numerical results show that for relatively lower dimension problems a Seydel method may be better than others, but for large scale problems it seems that our simplified perturbation method is the best one.

## References


[1] Crandall, M. G. and Rabinowitz, P. H., Bifurcation from simple eigenvalues, *J. Functional Analysis*, **8** (1971), 321—340.

[2] Keller, H. B., Bifurcation theory and nonlinear eigenvalue problem, Unpublished lecture notes, Caltech (1976).

[3] Keller, H. B., Nonlinear bifurcation, *J. Differential Equations*, **7** (1970), 417—434.

[4] Keller, H. B., Numerical solution of bifurcation and nonlinear eigenvalue problems, Applications of Bifurcation Teeory, P. H. Rabinowitz ed., Academic Press, New York, 1977, pp. 359—384.

[5] Keller, H. B. and Langford. W. F., Iterations, perturbations and multiplicities for nonlinear bifurcation problems, *Arch. Rational Mech. Anal.*, **48** (1972), 83—108.

[6] Ortega, J. M. and Rheinboldt, W. C., *Iterative solution of nonlinear equations in several variables* Academic Press, New York, 1970.

[7] Rheinboldt, W. C., Numerical methods for a class of finite dimensional bifurcation problems, *SIAM. J. Num. Anal.*, 15 (1978), 1—11.

[8] Rheinboldt, W. C., Numerical analysis of continuation methods for nonlinear structure problems, *Computers and Structures*, **13** (1981), 103—114.

[9] Rheinboldt, W. C. and Burkardt, J. V., A locally parameterized continuation process, *ACM Transaction on mathematical software*, **9** (1983), 215—241.

[10] Seydel, R., Numerical computation of branch points in nonlinear equations, *Num. Math.*, **33** (1979), 339—352.

[11] Seydel, R., Branch switching in bifurcation problems for ordinary differential equations, *Num. Math.*, 41 (1983), 93—116.

[12] Shi, M., A comparison of some methods for switching branches at simple bifurcation points, Tsinghua Univ., Dept. of Applied Math., Sci. Rep., No. 85012 (1985).