Journal of Computational Mathematics, Vol.10, No.3, 1992, 263–272.

A NEW TYPE OF REDUCED DIMENSION PATH FOLLOWING METHODS*

Zhang Shu-gong
(*Northeast University of Technology, Shenyang, China*)

Feng Guo-chen
(*Jilin University, Changchun, China*)

Abstract

To solve $F(x) = 0$ numerically, we first prove that there exists a tube-like neighborhood around the curve in R^n defined by the Newton homotopy in which $F(x)$ possesses some good properties. Then in this neighborhood, we set up an algorithm which is numerically stable and convergent. Since we can ensure that the iterative points are not far from the homotopy curve while computing, we need not apply the predictor-corrector which is often used in path following methods.

§1. Introduction

Suppose $F : R^n \to R^n$ is a smooth mapping. Let us consider solving $F(x) = 0$ globally. Recently, homotopy methods are used, that is, homotopy $H(t,x) = 0$ implicitly defines a path or a curve which leads to the root x^* of $F(x) = 0$. By following this path, we can finally reach x^*. But on computing, to follow the curve closely, we must use the predictor-corrector, which, of course, may bring us some trouble.

Here we take the advantadge of the Newton homotopy, and set up a new path following algorighm which (i) is numerically stable and (ii) does not use the correction technique. With it we can also judge whether or not we are going along the path we are following. Thus, the algorithm might make up for the deficiencies in current path following methods.

In this paper, we will use the following notations:

$$\|x\|^2 = \sum_{i=1}^n x_i^2 = (x_1, x_2, ..., x_n)^T \in R^n, \quad \|A\| = \max_{\|x\|=1} \|Ax\|, A \in L(R^n, R^m),$$

$$DQ = (\partial Q_i/\partial x_j), \quad Q : R^n \to R^m, \quad B(x,\delta) = \{y; \|y-x\| < \delta\}, x \in R^n,$$

$$d(y,E) = \inf\{\|y-x\|; x \in E\}, y \in R^n, E \subset R^n.$$

If $F(x)$ is a mapping from R^n to R^n, we then denote its last $n-1$ exponents by $G(x)$:

$$G(x) = (F_2(x), F_3(x), ..., F_n(x))^T.$$

* Received April 18, 1989.

§2. Basic Curve and Its Properties

Let $x_0 \in R^n$ be a given point. Consider the Newton homotopy

$$H(t,x) = F(x) - (1-t)F(x_0), \qquad 0 \le t \le 1 \tag{2.1}$$

chow etal. [1] discussed the conditions under which the connected exponents of one-dimensional manifolds determined by $H(t,x) = 0$ will be a diffeomorphism. [2] then proved the following theorem:

Theorem 2.1. *If $F(x): R^n \to R^n$ and $\bar{M} > 0$ satisfy*

(1) *F is a C^2 and proper mapping;*

(2) *the Lebesque measure of $\{x; \det DF(x) = 0\}$ is zero;*

(3) $\det DF(x) \ge 0$, *as $\|F(x)\| \ge \bar{M}$.*

Then for almost all $x_0 \in E_+(M) = \{x; \|F(x)\| \ge \bar{M}\}$, the projection $x(t)$: $[0,1] \to R^n$, of curve $(t, x(t))$ defined by the Newton homotopy $H(t,x) = 0$, is a diffeomorphism and connects x_0 with set $\{x; F(x) = 0\}$. Especially, $\{x; F(x) = 0\}$ is nonempty.

With the above condition, by Sard's theorem, we obtain

$$\operatorname{rank} DH(t,x) = n, \quad \text{for} \quad (t,x) \in \{(t,x); H(t,x) = 0\}. \tag{2.2}$$

In the following, we will always assume that $\det DF(x) \ne 0$ for $x \in \{x; F(x) = 0\}$ and $F(x)$, and x_0, and $H(t,x)$ satisfy the hypotheses in Theorem 2.1 and (2,2). Then, we have a smooth curve $x(t): [0,1] \to R^n$, denoted by $C(x_0)$.

We further assume its arc length $L < +\infty$ and parametrize by arc length s, i.e., $x(s): 0 \le s \le L, x(0) = x_0, x(L) = x^*$, instead of t.

For given $F(x_0)$, there exists an orthogonal matrix P, such that

$$PF(x_0) = (\|F(x_0)\|, 0, \cdots, 0)^T.$$

Obviously, solving $PF(x) = 0$ is equivalent to solving $F(x) = 0$, and $PF(x)$ and $F(x)$ have the same properties along $C(x_0)$. Later, we will consider solving $PF(x) = 0$ instead of $F(x) = 0$ and still denote $PF(x)$ by $F(x)$. Thus we have

$$F(x_0) = (\|F(x_0)\|, 0, \cdots, 0)^T. \tag{2.3}$$

From (2.2) and (2.3), we can easily get

Lemma 2.2. *If $F(x)$ and x_0 satisfy the conditions and (2.3), then*

$$\operatorname{rank} DG(x) = n-1, \quad \textit{for} \quad x \in \overline{C(x_0)}. \tag{2.4}$$

Hence, there is a constant $\gamma > 0$, such that

$$\|DG(x)^+\| \le \gamma, \quad x \in \overline{C(x_0)} \tag{2.5}$$

where $DG(x)^+ = DG(x)^T (DG(x) DG(x)^T)^{-1}$.

By the definitions of $H(t,x)$ and $G(x)$, we can get

Lemma 2.3. *If $F(x)$ and x_0 are as above, then*

(1) $F_1(x) \ge 0, G(x) \equiv 0$, *for* $x \in \overline{C(x_0)}$;

(2) *along* $C(x_0)$, *we have*

$$\begin{cases} \dot{x}(s) = \dfrac{dx(s)}{ds} = T(DG(x(s))), \\ x(0) = x_0, \end{cases}$$

where $T(DG(x))$ *is the solution of* $DG(x)T(DG(x)) = 0$ *satisfying*

$$\det \begin{pmatrix} T(DG(x))^T \\ DG(x) \end{pmatrix} < 0 \text{ and } \|T(DG(x))\| = 1.$$

Proof. We need only to show

$$\det \begin{pmatrix} \dot{x}(s)^T \\ DG(x(s)) \end{pmatrix} < 0.$$

The rest is trivial.

By $DG(x)\dot{x} = 0$ and rank $DG(x) = n - 1$, we have

$$\operatorname{sign} \det \begin{pmatrix} \dot{x}(s)^T \\ DG(x(s)) \end{pmatrix} = \operatorname{sign} \det \begin{pmatrix} \dot{x}(0)^T \\ DG(x(0)) \end{pmatrix}. \tag{2.6}$$

Let $\dot{t} = \dfrac{dt}{ds}$. From (2.3) and $DH(t,x)\begin{pmatrix} \dot{t} \\ \dot{x} \end{pmatrix} = 0$, it is clear that

$$\det \begin{pmatrix} \dot{t} & \dot{x}^T \\ \multicolumn{2}{c}{DH(t,x)} \end{pmatrix} \neq 0 \tag{2.7}$$

but

$$\begin{pmatrix} \dot{t} & \dot{x}^T \\ \multicolumn{2}{c}{DH(t,x)} \end{pmatrix} \begin{pmatrix} \dot{t} & 0^T \\ \dot{x} & I \end{pmatrix} = \begin{pmatrix} \dot{t}^2 + \|\dot{x}\|^2 & \dot{x}^T \\ 0 & DF(x) \end{pmatrix}. \tag{2.8}$$

Take determinants on both sides; then

$$\dot{t}(s) \det \begin{pmatrix} \dot{t} & \dot{x}^T \\ \multicolumn{2}{c}{DH(t,x)} \end{pmatrix} = (\dot{t}^2 + \|\dot{x}\|^2) \det DF(x). \tag{2.9}$$

Since $\|\dot{x}\| = 1$, $\dot{t}(s) = 0$ iff $\det DF(x) = 0$. On the other hand, $t \in [0,1]$, and $\det DF(x_0) > 0$. We have $\dot{t}(0) > 0$. From the first exponent of

$$DF(x_0)\dot{x}(0) = -\dot{t}(0)F(x_0)$$

we get

$$VF_1(x(0))\dot{x}(0) = -\dot{t}(0)\|F(x_0)\| < 0 \tag{2.10}$$

where $VF_1(x) = \left(\dfrac{\partial F_1}{\partial X_1}, \dfrac{\partial F_1}{\partial X_2}, \cdots, \dfrac{\partial F_1}{\partial X_n}\right)$. Consider matrix

$$DF(x) = \begin{pmatrix} VF_1(x) \\ DG(x) \end{pmatrix} = \begin{pmatrix} \dot{x}^T \\ DG(x) \end{pmatrix} \left(I - \dot{x}(\dot{x}^T - VF_1(x))\right). \tag{2.11}$$

We obtain

$$\det DF(x) = VF_1(x)\dot{x} \det \begin{pmatrix} \dot{x}^T \\ DG(x) \end{pmatrix}. \tag{2.12}$$

Hence, $\det DF(x_0) > 0$ and (2.10) yields the result.

Let $\Re(G) = \{x; \mathrm{rank} DG(x) = n-1\}$. Assume that $DF(x)$ satisfies the Lipschitz condition

$$\|DF(x) - DF(y)\| \leq K\|x-y\|, \ x, y \in R^n \tag{2.13}$$

where K is a Lipschitz constant. It is easy to see that $DG(x)$ satisfies the same Lipschitz condition.

Proposition 2.3[3]. Assume that A is an $m \times n$ matrix, and $B = A + \delta A$.

1) If rank $A = \min(m,n)$, then B is of full rnak, as $\|\delta A\| \ \|A^+\| < 1$.

2) If rank $A =$ rank B, and $\|\delta A\|\|A^+\| < 1$, then

$$\|B^+\| \leq \frac{\|A^+\|}{1 - \|\delta A\| \ \|A\|} \tag{2.14}$$

where A^+ is the pseudo — inverse of A.

Corollary 2.4. If $DF(x)$ satisfies (2.13), then

$$\Re(G) \supset \left\{y; \|y-x\| \leq \frac{1}{K\|DG(x)^+\|}, \ x \in \overline{C(x_0)}\right\}.$$

It is easy to show

Lemma 2.5. *Suppose* $A_1, A_2 \in L(R^n, R^{n-1}), A_i u_i = 0, \|u_i\| = 1, \det\begin{pmatrix} u_i^T \\ A_i \end{pmatrix} < 0$, $i = 1, 2$. *If*

$$\min_i \|A_i^+\| \ \|A_1 - A_2\| < 1, \tag{2.15}$$

then $u_1^T u_2 > 0$.

Lemma 2.6[4]. *If* $A_i \in L(R^n, R^{n-1}), A_i u_i = 0, \|u_i\| = 1, \det\begin{pmatrix} u_i^t \\ A_i \end{pmatrix} \neq 0$, $i = 1, 2, u_1^T u_2 \neq -1$, *then*

$$\|u_1 - u_2\| \leq \sqrt{\frac{2}{1 + u_1^T u_2}} \min_i \|A_i^+\| \ \|A_1 - A_2\|. \tag{2.16}$$

In $\Re(G)$, consider

$$\text{(I)} \qquad \dot{y}(s) = v(y(s)), \quad y(0) = y_0 \in \Re(G),$$

where $v(y)$ satisfies $DG(y)v(y) = 0$, $\det\begin{pmatrix} v(y)^T \\ DG(y) \end{pmatrix} < 0$ and $\|v(y)\| = 1$. We have

Theorem 2.7. *Let* K, L, γ *be as above,* $\delta < (K\gamma)^{-1} e^{-\sqrt{2}KL\gamma}$. *Then, for any* $y_0 \in B(x_0, \delta)$, *Eq* (I) *has, at least for* $s \leq L$, *solution* $y(s) \subset \Re(G)$, *and the estimate*

$$\|y(s) - x(s)\| \leq \|y_0 - x_0\| e^{\sqrt{2}Ks\gamma}, \ 0 \leq s \leq L. \tag{2.17}$$

holds.

Proof. From Lemmas 2.5 and 2.6, $v(y)$ satisfies the local Lipschitz condition. We can easily show that, for sufficiently small s,

$$\begin{aligned} \|y(s) - x(s)\| &\leq \|y_0 - x_0\| + \int_0^s \|\dot{y}(\tau) - \dot{x}(\tau)\| d\tau \\ &\leq \|y_0 - x_0\| + \sqrt{2}K\gamma \int_0^s \|y(\tau) - x(\tau)\| d\tau. \end{aligned} \tag{2.18}$$

Therefore, the Gronwall lemma and the extension theorem of ODEs give the results.

Assume that $y_0 \in B(x_0, \delta)$, and $y(s)$ is the solution curve of Eq(I). Let

$$H_{x_0} = \left\{z; (z - x_0)^T \dot{x}(0) = 0\right\}. \tag{2.19}$$

Then we have

Theorem 2.8. *Suppose δ is small enough. For all $y_0 \in H_{x_0} \cap \overline{B(x_0, \delta)}$, $0 \le s \le L$, all the integral curves of Eq(I) constitute a neighborhood, say $\mathcal{N}_\delta$, of $C(x_0)$, i.e., for any $0 < s < L$, $x(s)$ is an inner point of $\mathcal{N}_\delta$.* Furthermore, for each integral curve $y(s)$ of Eq(I), we have

$$\text{sign}\ \frac{dF_1(y(s))}{ds} = \text{-sign}\ \det DF(y(s)),\ s \le L, \tag{2.20}$$

$$G(y(s)) = G(y_0) \tag{2.21}$$

and

$$\det DF(x^*) > 0, \tag{2.22}$$

$$\delta^* = \inf_{x \in B} \|G(x)\| > 0 \tag{2.23}$$

where $B = \partial\mathcal{N}_\delta - \left\{B(x_0, \delta) \cup B(x^*, \delta e^{\sqrt{2}KL\gamma})\right\}$.

Proof. From Theorem 2.7, if δ is small enough, then for $y_0 \in H_{x_0} \cap \overline{B(x_0, \delta)}$, $0 \le s \le L$, all the integral curves $y(s)$ of Eq(I) form a close connected set $\mathcal{N}_\delta$, the continuity of $y(s) = y(s, y_0)$ with respect to s and y_0 ensures that $x(s) = y(s, x_0)$ is an inner point of $\mathcal{N}_\delta$. (2.21) can be easily got from

$$G(y(s)) = G(y_0) + \int_0^s DG(y(\tau))\dot{y}(\tau)d\tau = G(y_0). \tag{2.24}$$

To obtain (2.20), notice that $\frac{dF_1(y(s))}{ds} = VF_1(y(s))\dot{y}(s)$. Then (2.12) and det $\begin{pmatrix} \dot{y}(s)^T \\ DG(y(s)) \end{pmatrix} < 0$ give the result. $F_1(x_0) > 0$ and $F_1(x^*) = 0$ yield sign $\det DF(x^*) = -\text{sign}\frac{dF(x^*)}{ds} > 0$. (2.23) can be deduced from (2.24) and $\delta^* = \min\left\{\|G(x)\|;\ x \in H_{x_0} \cap \partial\overline{B(x_0, \delta)}\right\} > 0$.

§3. Algorithm

Consider

$$\text{(II)} \quad \begin{cases} \dot{z} = u(z),\ DF(z)u(z) = -\begin{pmatrix} \lambda(z)F_1(z) \\ |\lambda(z)|G(z) \end{pmatrix}, \\ z(0) = z_0 \in \mathcal{N}_\delta \cap \left\{z;\ \|G(z)\| \le \delta^*\right\} \end{cases}$$

where $\lambda(z) : R^n \longrightarrow R^{n-1}$ should satisfy

(1) $\text{sign}\lambda(z) = \text{sign}\ \det DF(z)$;

(2) $\lim_{z \to x^*} \lambda(z) = +\infty$, $x^* \in \{x;\ F(x) = 0\}$.

For example, $\lambda(z) = \det DF(z)/\|F(z)\|$ is an ideal one.

It is clear that if the solutions of Eq(II) exist, then they satisfy

$$G(z(s)) = e^{-\int_0^s |\lambda(z(\tau))| d\tau} G(z_0), \tag{3.1}$$

$$\text{sign } \frac{dF_1(z(s))}{ds} = \text{sign } \det DF(z(s)). \tag{3.2}$$

If $\det DF(z) = 0$, then the solutions of Eq(II) and Eq(I) are tangent to each other at z. If z_0 and $F(x)$ satisfy the condition in the Kantorovich theorem, then the Newton iteration which numerically solves $F(x) = 0$ converges to the root x^* of $F(x) = 0$.

Consider the discrete analogue of Eq(II). We have

Reduced Dimension Path Following Method. Given $x_0, \varepsilon > 0, d > 0, \eta > 0$, set $k = 0$.

Step 1. If $D_k = |\det DF(x_k)| \le d$, set $f = F_1(x_k)$, goto Step 2;

$$\bar{x} = x_k - \tau_k DF(x_k)^{-1} \begin{pmatrix} \sigma_k F_1(x_k) \\ G(x_k) \end{pmatrix},$$

where $\sigma_k = \sigma(x_k) = \text{sign } \det DF(x_k)$, and τ_k is to be determined. If $\|F(\bar{x})\| < \epsilon$, stop. $x_{k+1} = \bar{x}$, $k = k + 1$, go back to Step 1.

Step 2. $\Delta x_k = -\varphi(\|DG(x_k)^+ G(x_k)\|) DG(x_k)^+ G(x_k) + u_k$ where

$$\varphi(t) = \begin{cases} 1 & \text{as } t \le 1, \\ 1/t & \text{as } t > 1; \end{cases}$$

$$u_k;\ DG(x_k)u_k = 0, \det \begin{pmatrix} uk^T \\ DG(x_k) \end{pmatrix} < 0,\ \|u_k\| = 1;$$

$$\bar{x} = x_k + \tau_k \Delta x_k, \text{ and } \tau_k \text{ is to be determined.}$$

If $\|F(\bar{x})\| < \varepsilon$, stop. $x_{k+1} = \bar{x}$, $k = k + 1$. If $|\det DF(\bar{x})| > d$, $\sigma(\bar{x})F_1(\bar{x}) < \sigma(\bar{x})(f - \sigma(\bar{x})\eta)$, $\sigma(\bar{x}) = \det DF(\bar{x})$ go to Step 1; otherwise go to Step 2.

§4. The Determination of Parameters and Convergence

In this section, we discuss the determination of parameters appearing in §2. First, choose δ such that Theorem 1.8 holds. Thus, there exist s_i such that $|DF(x(s_i))| = 0$ for $i = 1, 2, \cdots, I$. By the smoothness of $F(x)$ and the hypotheses in Theorem 1.1, if δ and d are small enough, then there exist $\rho_i > 0$, such that

$$C_i(d) = \{x \in \mathcal{N}_\delta;\ |\det DF(x)| \le d\} \subset B_i = B(x(s_i), \rho_i), \quad i = 1, 2, \cdots, I.$$

$$B_{I+1} = B_{I+1}(x^*, \rho_{I+1}) \subset \{x; |\det DF(x)| > d\}, \quad B_i \cap B_j = \phi, \quad i \ne j,$$

$$i, j = 1, 2, \cdots, I + 1.$$

Therefore, $\mathcal{N}_\delta$ can be written as

$$\mathcal{N}_\delta = \mathcal{N}_\delta^0 \cup C_1(d) \cup \mathcal{N}_\delta^1 \cup \cdots \cup C_I(d) \mathcal{N}_\delta^I \tag{4.1}$$

In this case, if $y \in \mathcal{N}_\delta^i$, then $|\det DF(y)| > d$ and the minimum and the maximum of $F_1(x)$ on $\mathcal{N}_\delta^i$ must appear on $\bar{\mathcal{N}}_\delta^i \cap C_i(d)$ or $\bar{\mathcal{N}}_\delta^i \cap C_{i-1}(d)$.

Now we can choose δ and d so small that the above and the following facts hold:

$$d < \det DF(x_0), \tag{4.2}$$

$$\delta < \min\left\{\frac{1}{2}K^{-1}\gamma^{-1}e^{-\sqrt{2}K\gamma L},\ \frac{1}{2}\frac{e^{-\sqrt{2}K\gamma L}}{\max_{0\le s\le L}\|\ddot{x}(s)\|}\right\}, \tag{4.3}$$

$$\eta = \max\{\sup_{c_i(d)} F_1(x) - \inf_{c_i(d)} F_1(x)\} < \min_i \inf_{c_i(d)} F_1(x) \tag{4.4}$$

where $\ddot{x}(s) = \dfrac{d^2x(s)}{ds^2}$. If we set

$$M = \sup\left\{\|DF(x)^{-1}\|;\ x \in \left(\bigcup_{i=1}^{I} \mathcal{N}_\delta^i\right) \cup B_{I+1}\right\}, \tag{4.5}$$

$$\mu = \min\left\{1/4KM^2,\ \frac{1}{2}\min_i \inf_{C_i(d)} F_1(x),\ \frac{1}{2}\inf_{x\in\partial B_{I+1}}\|F(x)\|\right\}$$

and

$$\begin{cases} \tau_k^{(1)} = |F_1(x_k)|/KM^2\|F(x_k)\|^2, \\ \tau_k^{(2)} = [\|G(x_k)\| + (\|G(x_k)\|^2 + 2KM^2\|F(x_k)\|^2)(\mu - \|G(x_k)\|)]^{\frac{1}{2}}/KM^2\|F(x_k)\|^2, \\ \tau_k^{(3)} = 2F_1(x_k)/[\sqrt{F_1(x_k)^2 + 2F_1(x_k)KM^2\|F(x_k)\|^2} + F_1(x_k)] \end{cases} \tag{4.6}$$

then we have

Lemma 4.1. *If*

$$\tau_k = \begin{cases} \min\{\tau_k^{(1)},\ \tau_k^{(2)},\ \tau_k^{(3)},\ 1\}, & as\quad F_1(x_k) > \mu, \\ 1, & as\quad F_1(x_k) \le \mu \end{cases} \tag{4.7}$$

then in Step 1, there exists a constant $\tau^* > 0$, independent of k, such that $\tau_k \ge \tau^*$; moreover, if $\det DF(x_k) > d$, $k = N, \cdots, N+p$, then $\|G(x_k)\| \le \mu$ and

$$|F_1(x_{N+p})| \le \left(1 - \frac{1}{2}\tau^*\right)^p |F_1(x_N)|, \tag{4.8}$$

and if $\det DF(x_k) < -d$, $k = N_1, \cdots, N_1+q$, then $\|G(x_k)\| \le \mu$ and

$$|F_1(x_{N_1+q})| \ge \left(1 + \frac{1}{2}\tau^*\right)^q |F_1(x_{N_1})|. \tag{4.9}$$

Furthermore, either $F_1(x_k) \ge 0$ or the Newton iteration converges with initial x_k.

Proof. Consider the Taylor expansion of $F(x_{k+1})$ at x_k, and let

$$\sigma_k|F_1(x_{k+1})| \le \sigma_k\left(1 - \frac{1}{2}\sigma_k\tau_k\right)|F_1(x_k)|, \tag{4.10}$$

$$F_1(x_{k+1}) \ge 0 \text{ and } \|G(x_{k+1})\| \le \mu. \tag{4.11}$$

Then we can get $\tau_k^{(i)}$, $i = 1, 2, 3$, and hence (4.8), (4.9).

For $x_k \in \mathcal{N}_\delta$, there exists an F^* such that $\|F(x_k)\| \le F^*$. If $F_1(x_k) > \mu$, it is easy to show that there is a $\tau^* > 0$ independent of k, such that $\tau_k^{(i)} \ge \tau^*$, $i = 1, 2, 3$. If $0 \le F_1(x_k) \le \mu$, then $\tau_k = 1$, and $x_k \in \mathcal{N}_\delta^I \cup B_{I+1}$. Since $\|F(x_k)\|^2 = F_1(x_k)^2 +$

$\|G(x_k)\|^2 \le 2\mu$, i.e. $\|F(x_k)\| \le \inf_{x\in\partial B_{I+1}} \|F(x)\|$, by the Kantorovich theorem, the Newton iteration converges.

Lemma 4.2. *In Step 2, if we denote $v = DG(x_k)^+G(x_k)$, and choose*

$$\tau_k^{(4)} = \begin{cases} \min\left\{1,\ \left[\frac{\mu-\|G(x_k)\|}{K} + \frac{\|G(x_k)\|}{4K^2\|V\|}\right]^{\frac{1}{2}} + \frac{\|G(x_k)\|}{2K\|V\|}\right\}, & G(x_k)\neq 0, \\ \min\left\{1,\ 2\sqrt{\mu/K}\right\}, & \text{as } G(x_k)=0, \end{cases} \tag{4.12}$$

$$\tau_k = \max\{t;\ \ t\in[0,\ \tau_k^{(4)}],\ F_1(x_k+t\Delta x_k)\ge 0\}, \tag{4.13}$$

then $\|G(x_{k+1})\| \le \mu$, and either $\tau_k > \tau^* > 0$, τ^* independent of k, or the Newton iteration converges with initial x_{k+1}.

Proof. Let $z(t) = x_k + t\Delta x_k$, $t\in(0,\ \tau_k^{(4)})$. By the Taylor expansion of $G(z(t))$ at x_k, we can get $\|G(z(t))\|\le\mu$, and especially $\|G(x_{k+1})\|\le\mu$. Now we can inductively assume that $x_k\in\mathcal{N}_\delta$, $F_1(x_k)\ge 0$. If $\tau_k = \tau_k^{(4)}$, it is clear that $\tau_k>\tau^*>0$, τ^* independent of k. If $\tau_k<\tau_k^{(4)}$, then $F_1(x)$ must vanish at x_{k+1}. Hence $x_{k+1}\in\mathcal{N}_\delta^I\cup B_{I+1}$ and $\|F(x_{k+1})\|\le\mu$. Therefore, the Newton iteration converges with initial x_{k+1}.

Lemma 4.3. *In Step 2, if $x(s_k)$ is a point on $C(x_0)$ such that $d(x_k, C(x_0)) = \|x_k - x(s_k)\|$, then*

$$\langle\Delta x_k,\ \dot{x}(s_k)\rangle \le \frac{5}{12}\pi < \frac{\pi}{2}. \tag{4.14}$$

Moreover, there is a $\bar\tau>0$, independent of k, such that

$$s\ge s_k + c^*\tau, \qquad as\ \ 0\le\tau\le\bar\tau, \tag{4.15}$$

where s and τ satisfy $d(x_k+\tau\Delta x_k,\ C(x_0)) = \|x_k+\tau\Delta x_k - x(s)\|$, $\langle\Delta x_k,\ \dot{x}(s_k)\rangle$ is the angle between Δx_k and $\dot{x}(s_k)$, and c^* is a constant.

Proof. For $x_k\in\mathcal{N}_\delta$, there is a solution $y(s)$ of Eq(I) and $\bar{s}_k\le L$ such that $y(\bar{s}_k) = x_k$. But

$$\begin{aligned} d\ (x_k,\ C(x_0)) &= \|x_k - x(s_k)\| = \|y(\bar{s}_k) - x(s_k)\| \\ &\le \|y(s_k) - x(s_k)\| \le \frac{1}{2}K^{-1}\gamma^{-1}. \end{aligned} \tag{4.16}$$

By Lemmas 2.5 and 2.6, $\dot{y}(\bar{s}_k)^T\dot{x}(s_k)\ge\frac{\sqrt{3}}{2}$, hence

$$\langle\dot{y}(\bar{s}_k),\dot{x}(s_k)\rangle = \cos^{-1}(\dot{y}(\bar{s}_k)^T\dot{x}(s_k)) < \frac{\pi}{6}. \tag{4.17}$$

On the other hand, since

$$\Delta x_k = -\varphi(\|DG(x_k)^+G(x_k)\|)DG(x_k)^+G(x_k) + \dot{y}(\bar{S}_k) \tag{4.18}$$

(in this case, $u_k = \dot{y}(s_k)$, see §2), we have

$$\Delta x_k^T\dot{y}(\bar{s}_k) = 1 \quad\text{ and }\quad \|\Delta x_k\|\le\sqrt{2}. \tag{4.19}$$

Therefore

$$\langle\Delta x_k,\dot{y}(\bar{s}_k)\rangle = \cos^{-1}\left(\frac{\Delta x_k^T\dot{y}(s_k)}{\|\Delta x_k\|}\right)\ge\cos^{-1}\left(\frac{\sqrt{2}}{2}\right) = \frac{\pi}{4}. \tag{4.20}$$

Thus we get

$$\langle \Delta x_k, \dot{x}(s_k)\rangle \le \langle \Delta x_k, \dot{y}(\bar{s}_k)\rangle + \langle \dot{y}(\bar{s}_k), \dot{x}(s_k)\rangle \le \frac{\pi}{4} + \frac{\pi}{6} < \frac{\pi}{2}. \tag{4.21}$$

To obtain (4.15), let $\bar{x}(\tau) = x_k + \tau\Delta x_k$. Since $\|\bar{x}(\tau) - x(s)\| = d(\bar{x}(\tau),\ C(x_0))$, we have $(\bar{x}(\tau) - x(s))^T\dot{x}(s) = 0$. Thus

$$(\Delta x_k d\tau - \dot{x}(s)ds)^T\dot{x}(s) + (x_k + \tau\Delta x_k - x(s))^T\ddot{x}(s)ds = 0. \tag{4.22}$$

As τ is small enough, since $|(x_k - x(s_k))^T\ddot{x}(s_k)| \le \frac{1}{2}$, we get

$$\frac{ds}{d\tau} = \frac{\Delta x_k^T\dot{x}(s)}{1 - (\bar{x}(\tau) - x(s))^T\ddot{x}(s)}$$

and especially,

$$\left.\frac{ds}{d\tau}\right|_{\tau=0} = \frac{\Delta x_k^T\dot{x}(s_k)}{1 - (x_k - x(s_k))^T\ddot{x}(s_k)} \ge \frac{2}{3}\cdot\frac{\sqrt{3}}{2} = \frac{\sqrt{3}}{3} = C^*.$$

Thus, there exists a $\bar{\tau} > 0$, such that (4.15) holds. $\bar{\tau}$ and c^* are clearly independent of k.

Theorem 4.4. *If the choices of all the parameters are as above, then the Reduced Dimension Path Following Methods converge.*

Proof. From the choice of d, the algorithm enters $\mathcal{N}_\delta^0$ by Step 1, since $\inf_{x\in\mathcal{N}_\delta^0} F_1(x) \ge \mu > 0$. By Lemma 4.1, the algorithm cannot stay in $\mathcal{N}_\delta^0$, but $\|x_{k+1} - x_k\| < \frac{1}{KM^2}$, $\|DF(x_k)^{-1}\| \le M$, and so there must be some $x_{k+1} \in C_1(d)$.

By Lemmas 4.2 and 4.3 and the choice of η, the algorithm neither stays in $C_1(d)$ nor goes back to $\mathcal{N}_\delta^0$; hence it must go into $\mathcal{N}_\delta - \{\mathcal{N}_\delta^0 \cup C_1(d)\}$.

Inductively, it must finally enter $\mathcal{N}_\delta^I \cup B_{I+1}$ and therefore, converge to x^*.

§5. Example

Consider

$$F(x_1,x_2,x_3) = \begin{pmatrix} x_1^4 - 2x_1^3 - 5x_1^2 + 12x_1 - 2 \\ x_2^2 - x_1 \\ x_1^2 + x_3 \end{pmatrix}.$$

Then

$$DG(x) = \begin{pmatrix} -1 & 2x_2 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

For any $x \in R^3$, we always have rank $DG(x) = 2$. Let $x_0 = (2.5, \sqrt{2.5}, -2.5)^T$. We get the following table. The increment Δx_k with asterisk is computed by Step 2. This example shows that although computation by Step 2 is more complicated than by Step 1, only a few points need to be computed by Step 2.

k	0	1	2	3
x_k	2.5 1.58113883 −2.5	2.119791667 1.460906398 −2.119791667	2.051862743 1.437348331 −2.051862743	1.366038104 1.193864513 −1.366038104
$F(x_k)$	4.5625 0 0	2.110970323 0.014455836 0	2.019672431 0.014107481 0	3.446118637 0.059274371 0
Δx_k	−0.380208333 −0.120232431 0.380208333	−1.086862791 −0.37692907 1.086862791	−0.685824639* −0.243483818 0.685824639	−1.295384871 −0.567342117 1.295384871
det DF	37.947332	5.674920656	2.228038779	−6.352087077
τ_k	1	2^{-4}	1	1/4

k	4	5	6	
x_k	1.042191886 1.052028984 −1.042191886	0.38014499 0.698871838 −0.38014499	0.181258712 0.425744891 −0.181258712	
$F(x_k)$	3.991252609 0.064573097 0	1.750202347 0.108276855 0	-3.55E-08 2.1E-10 0	
Δx_k	−0.662046896* −0.333157146 0.662046896	−0.075664844 −0.131599021 0.075664844	3.5449996E-09 3.9166643E-09 -3.5449996E-09	
det DF	−1.767360251	10.55468135	8.526908398	
τ_k	1	1	1	

$$x_{10} = \begin{pmatrix} 0.181258715 \\ 0.425744894 \\ -0.181258715 \end{pmatrix}, \quad F(x_{10}) = \begin{pmatrix} -5.4\text{E-09} \\ -2.4\text{E-10} \\ 0 \end{pmatrix}.$$

Take x_{10} as x^*; then we have $\|F(x^*)\| < 10^{-8}$.

References

[1] S.N. Chow, et al, Finding zeros of maps: Homotopy methods that are constructive with probability one, *Math. comput.*, **32** (1978), 887–899.

[2] M.W. Hirsch and S. Smale, On algorithms for solving $F(x) = 0$, *Comm. Pure and Applied Math.*, **32** (1979), 281–312.

[3] He Xu-chu, On singularity, ill-conditionness and related problems, *Numer. Math.*, **6** : 1 (1984), 74–85.

[4] W.C. Rheinboldt, Solution fields of nonlinear equations and continuation methods, *SIAM J. Numer. Anal.*, **17** (1980), 221–237.