THE QUASI-NEWTON METHOD IN PARALLEL CIRCULAR ITERATION*1)

WANG XING-HUA (王兴华) ZHENG SHI-MING (郑士明)
(Department of Mathematics, Hangzhou University, Hangzhou, China)

Newton iteration

$$x^{(k+1)}=x^{(k)}-\frac{f(x^{(k)})}{f'(x^{(k)})},\quad k=0,\ 1,\ \cdots \tag{1}$$

is a most basic iteration for solving the numerical equation $f(x)=0$. If $f(x)$ is a monic polynomial of degree $n(n>1)$ with complex coefficients:

$$f(x)=x^n+a_1x^{n-1}+\cdots+a_n=\prod_{i=1}^{n}(x-\xi_i) \tag{2}$$

and its zeros $\xi_1,\ \cdots,\ \xi_n$ are different each other, we have

$$f(x)\approx\prod_{i=1}^{n}(x-x_i^{(k)}),\quad f'(x_i^{(k)})\approx\prod_{\substack{j=1\\ j\neq i}}^{n}(x_i^{(k)}-x_j^{(k)}) \tag{3}$$

for some approximation $x_1^{(k)},\ \cdots,\ x_n^{(k)}$ of $\xi_1,\ \cdots,\ \xi_n$. Then the further approximation is

$$x_i^{(k+1)}=x_i^{(k)}-\frac{f(x_i^{(k)})}{\prod\limits_{\substack{j=1\\ j\neq i}}^{n}(x_i^{(k)}-x_j^{(k)})},\quad i=1,\ \cdots,\ n;\ k=0,\ 1,\ \cdots. \tag{4}$$

This is just the parallel iteration proposed by Durand[1] and Kerner[2]. We see that this is a Newton method, which aims at the concrete task for finding all zeros of polynomial and applied approximation (3).

Let $W_i^{(k)}=[x_i^{(k)};\ r_i^{(k)}]$ denote disks in complex plane $\mathscr{C}$ with center $x_i^{(k)}$ and radius $r_i^{(k)}$

$$\{x\in\mathscr{C}:\ |x-x_i^{(k)}|\leqslant r_i^{(k)}\}.$$

Then under the operation rule of circular arithmetic2)

$$[x_1;r_1]\pm[x_2;\ r_2]=[x_1\pm x_2;\ r_1+r_2],$$

$$[x_1;\ r_1]\cdot[x_2;\ r_2]=[x_1x_2;\ |x_1|r_2+|x_2|r_1+r_1r_2],$$

$$\frac{1}{[x_2;\ r_2]}=\frac{1}{|x_2|^2-r_2^2}[\bar{x}_2;\ r_2],$$

$$\frac{[x_1;\ r_1]}{[x_2;\ r_2]}=[x_1;\ r_1]\cdot\frac{1}{[x_2;\ r_2]},$$

$0\overline{\in}[x_2;\ r_2]$, $\bar{x}_2$ denotes conjugate complex of x_2, an analogy of iteration (4) is

$$W_i^{(k+1)}=x_i^{(k)}-f(x_i^{(k)})\prod_{\substack{j=1\\ j\neq i}}^{n}\frac{1}{x_i^{(k)}-W_j^{(k)}},\quad i=1,\ \cdots,\ n;\ k=0,\ 1,\ \cdots, \tag{5}$$

* Received September 10, 1983.
1) Projects supported by the Science Fund of the Chinese Academy of Sciences.
2) Any complex is regarded as a disk with radius 0.

which is called the quasi-Newton method in parallel circular iteration. In comparison with the parallel circular iterations proposed by Braess, Hadeler[3], Petković[4], Gargantini, Henrici[5] etc., the construction of iteration (5) is simpler.

According to the inclusion monotone of circular arithmetic, if $W_1^{(k)}, \cdots, W_n^{(k)}$ are isolate and contain $\xi_1, \cdots, \xi_n$ respectively, then $W_1^{(k+1)}, \cdots, W_n^{(k+1)}$ contain also $\xi_1, \cdots, \xi_n$ respectively. Therefore, the circular iteration is feasible if the isolation of $W_1^{(k+1)}, \cdots, W_n^{(k+1)}$ is guaranteed, and the rate of convergence may be described by the rate of

$$r^{(k)} = \max_{1 \le i \le n} r_i^{(k)} \tag{6}$$

tending to 0. We may denote the isolation of disks $W_1^{(k)}, \cdots, W_n^{(k)}$ by

$$\vartheta^{(k)} = r^{(k)}/\rho^{(k)}, \tag{7}$$

where

$$\rho^{(k)} = \min_{\substack{1 \le i,j \le n \\ j \ne i}} \min_{z \in W_j^{(k)}} |z - x_i^{(k)}|. \tag{8}$$

If $r^{(k)} \to 0$ $(k \to \infty)$, then

$$\rho^{(k)} \to \min_{\substack{1 \le i,j \le n \\ j \ne i}} |\xi_j - \xi_i| \ne 0, \quad k \to \infty$$

and

$$\vartheta^{(k)} \asymp r^{(k)}, \quad k \to \infty. \tag{9}$$

Hence, the rate tended to 0 of $\vartheta^{(k)}$ represents directly the rate of convergence of circular iteration. We have the following theorem on $\vartheta^{(k)}$ for circular iteration (5).

Theorem. *Suppose that the initial disks $W_1^{(0)}, \cdots, W_n^{(0)}$ include the roots $\xi_1, \cdots, \xi_n$ of equation (2) respectively, and*

$$\vartheta^{(0)} \le \frac{1}{3(n-1)}. \tag{10}$$

Then the sequences $\{W_i^{(k)}\}_{k=0}^{\infty}$ $(i=1, \cdots, n)$ produced by (5) satisfy

$$\vartheta^{(k+1)} \le 3(n-1)(\vartheta^{(k)})^2, \quad k=0, 1, \cdots \tag{11}$$

and contract to roots $\xi_1, \cdots, \xi_n$ of equation (2) respectively:

$$\bigcap_{k=0}^{\infty} W_i^{(k)} = \xi_i, \quad i=1, \cdots, n. \tag{12}$$

Proof. From (5) and (2), we have

$$W_i^{(k+1)} - x_i^{(k)} = (x_i^{(k)} - \xi_i) \prod_{\substack{j=1 \\ j \ne i}}^{n} \frac{x_i^{(k)} - \xi_j}{x_i^{(k)} - W_j^{(k)}}, \quad i=1, \cdots, n; \ k=0, 1, \cdots. \tag{13}$$

Let

$$x = \text{mid}[x; r], \quad r = \text{rad}[x; r]. \tag{14}$$

By circular arithmetic we know that

$$\text{mid} \frac{x_i^{(k)} - \xi_j}{x_i^{(k)} - W_j^{(k)}} = \frac{(x_i^{(k)} - \xi_j)(\bar{x}_i^{(k)} - \bar{x}_j^{(k)})}{|x_i^{(k)} - x_j^{(k)}|^2 - (r_j^{(k)})^2},$$

$$\text{rad} \frac{x_i^{(k)} - \xi_j}{x_i^{(k)} - W_j^{(k)}} = \frac{|x_i^{(k)} - \xi_j| r_j^{(k)}}{|x_i^{(k)} - x_j^{(k)}|^2 - (r_j^{(k)})^2}, \quad i, j=1, \cdots, n; \ j \ne i; \ k=0, 1, \cdots.$$

Clearly,

$$|x_i^{(k)} - \xi_j| \le |x_i^{(k)} - x_j^{(k)}| + r_j^{(k)},$$

$$\rho^{(k)} \le |x_i^{(k)} - x_j^{(k)}| - r_j^{(k)}, \quad i, j=1, \cdots, n; \ j \ne i; \ k=0, 1, \cdots.$$

Thus,

$$\left|\operatorname{mid}\frac{x_i^{(k)}-\xi_j}{x_i^{(k)}-W_j^{(k)}}\right|\leqslant\frac{|\bar{x}_i^{(k)}-\bar{x}_j^{(k)}|}{|x_i^{(k)}-x_j^{(k)}|-r_j^{(k)}}=\frac{|x_i^{(k)}-x_j^{(k)}|}{|x_i^{(k)}-x_j^{(k)}|-r_j^{(k)}}$$

$$=1+\frac{r_j^{(k)}}{|x_i^{(k)}-x_j^{(k)}|-r_j^{(k)}}\leqslant 1+\frac{r_j^{(k)}}{\rho^{(k)}}\leqslant 1+\vartheta^{(k)},$$

$$\operatorname{rad}\frac{x_i^{(k)}-\xi_j}{x_i^{(k)}-W_j^{(k)}}\leqslant\frac{r_j^{(k)}}{|x_i^{(k)}-x_j^{(k)}|-r_j^{(k)}}\leqslant\frac{r_j^{(k)}}{\rho^{(k)}}\leqslant\vartheta^{(k)},\quad i,\ j=1,\ \cdots,\ n;\ j\neq i;\ k=0,\ 1,\ \cdots.$$

Hence, we obtain by (13)

$$|x_i^{(k+1)}-x_i^{(k)}|=|\operatorname{mid}(W_i^{(k+1)}-x_i^{(k)})|\leqslant r^{(k)}(1+\vartheta^{(k)})^{n-1},\tag{15}$$

$$r_i^{(k+1)}=\operatorname{rad}(W_i^{(k+1)}-x_i^{(k)})\leqslant r^{(k)}\{(1+2\vartheta^{(k)})^{n-1}-(1+\vartheta^{(k)})^{n-1}\},$$
$$i=1,\ \cdots,\ n;\ k=0,\ 1,\ \cdots.\tag{16}$$

Let

$$\delta^{(k)}=\max_{1\leqslant i\leqslant n}|x_i^{(k+1)}-x_i^{(k)}|,\tag{17}$$

$$\rho_{i,j}^{(k)}=\min_{z\in W_j^{(k)}}|z-x_i^{(k)}|.\tag{18}$$

We know from (15), (16)

$$\delta^{(k)}\leqslant r^{(k)}(1+\vartheta^{(k)})^{n-1},\tag{19}$$

$$r^{(k+1)}\leqslant r^{(k)}\{(1+2\vartheta^{(k)})^{n-1}-(1+\vartheta^{(k)})^{n-1}\},\quad k=0,\ 1,\ \cdots.\tag{20}$$

For $k=0,\ 1,\ \cdots,$ if $i,\ j\ (i\neq j)$ and $z\in W_j^{(k+1)}$ are choosed such that

$$\rho^{(k+1)}=\rho_{i,j}^{(k+1)}=|z-x_i^{(k+1)}|.$$

Then from

$$\rho_{i,j}^{(k)}\leqslant|\xi_j-x_i^{(k)}|\leqslant|\xi_j-x_j^{(k)}|+|x_j^{(k)}-z|+|z-x_i^{(k+1)}|+|x_i^{(k+1)}-x_i^{(k)}|$$
$$\leqslant 2r^{(k+1)}+\rho^{(k+1)}+\delta^{(k)}$$

we obtain

$$\rho^{(k+1)}\geqslant\rho^{(k)}-\delta^{(k)}-2r^{(k+1)}.\tag{21}$$

Hence, for $k=0,\ 1,\ \cdots,$ by (7), (18), (19), (17) we have

$$\vartheta^{(k+1)}=\frac{r^{(k+1)}}{\rho^{(k+1)}}\leqslant\frac{r^{(k)}\{(1+2\vartheta^{(k)})^{n-1}-(1+\vartheta^{(k)})^{n-1}\}}{\rho^{(k)}-\delta^{(k)}-2r^{(k+1)}}$$

$$\leqslant\frac{r^{(k)}\{(1+2\vartheta^{(k)})^{n-1}-(1+\vartheta^{(k)})^{n-1}\}}{\rho^{(k)}-r^{(k)}(1+\vartheta^{(k)})^{n-1}-2r^{(k)}\{(1+2\vartheta^{(k)})^{n-1}-(1+\vartheta^{(k)})^{n-1}\}}$$

$$=\frac{\vartheta^{(k)}\{(1+2\vartheta^{(k)})^{n-1}-(1+\vartheta^{(k)})^{n-1}\}}{1+\vartheta^{(k)}(1+\vartheta^{(k)})^{n-1}-2\vartheta^{(k)}(1+2\vartheta^{(k)})^{n-1}}$$

$$=3(n-1)(\vartheta^{(k)})^2$$

$$\times\left(1-\frac{1-\left\{\left(\frac{1}{3(n-1)\vartheta^{(k)}}+2\vartheta^{(k)}\right)(1+2\vartheta^{(k)})^{n-1}-\left(\frac{1}{3(n-1)\vartheta^{(k)}}+\vartheta^{(k)}\right)(1+\vartheta^{(k)})^{n-1}\right\}}{1-\{2\vartheta^{(k)}(1+2\vartheta^{(k)})^{n-1}-\vartheta^{(k)}(1+\vartheta^{(k)})^{n-1}\}}\right)$$

$$=3(n-1)(\vartheta^{(k)})^2\left(1-\frac{1-A(n-1,\ (n-1)\vartheta^{(k)})}{1-B(n-1,\ (n-1)\vartheta^{(k)})}\right),\tag{22}$$

where

$$A(t,\ \alpha)=\left(\frac{1}{3\alpha}+\frac{2\alpha}{t}\right)\left(1+\frac{2\alpha}{t}\right)^t-\left(\frac{1}{3\alpha}+\frac{\alpha}{t}\right)\left(1+\frac{\alpha}{t}\right)^t,$$

$$B(t,\ \alpha)=\frac{2\alpha}{t}\left(1+\frac{2\alpha}{t}\right)^t-\frac{\alpha}{t}\left(1+\frac{\alpha}{t}\right)^t.$$

It may be seen that $A(t,\ \alpha)$ is a decreasing function about $t\geqslant 1$ if $0<\alpha\leqslant\frac{1}{3}$ (see following lemma). Therefore

$$A(t,\alpha)\leqslant A(1,\alpha)=\frac{1}{3}+\alpha+3\alpha^2\leqslant 1,$$

$$B(t,\alpha)\leqslant A(t,\alpha)\leqslant 1, \quad \text{if } t\geqslant 1,\ 0<\alpha\leqslant\frac{1}{3}.$$

Thus, if

$$\vartheta^{(k)}\leqslant\frac{1}{3(n-1)} \tag{23}$$

holds for some $k\in\{0, 1, \cdots\}$, then

$$\vartheta^{(k+1)}\leqslant 3(n-1)(\vartheta^{(k)})^2 \tag{24}$$

from (22), and

$$\vartheta^{(k+1)}\leqslant\frac{1}{3(n-1)}.$$

Hence (24) and (23) hold for all $k=0, 1, \cdots$ by (10). The conclusion (11) of the theorem is proved.

Moreover, from (23) we see that

$$\begin{aligned}(1+2\vartheta^{(k)})^{n-1}-(1+\vartheta^{(k)})^{n-1}&=\sum_{\nu=1}^{n-1}\binom{n-1}{\nu}(2^\nu-1)\vartheta^{(k)\nu}\\ &\leqslant\sum_{\nu=1}^{n-1}\frac{(n-1)(n-2)\cdots(n-\nu)}{(n-1)^\nu}\cdot\frac{2^\nu-1}{\nu!3^\nu}\\ &<\sum_{\nu=1}^{\infty}\frac{2^\nu-1}{\nu!3^\nu}=e^{2/3}-e^{1/3}<0.56.\end{aligned}$$

Hence, by (20) we have

$$r^{(k+1)}<0.56r^{(k)}, \quad k=0, 1, \cdots. \tag{25}$$

This means that $r^{(k)}\to 0$ $(k\to\infty)$. Because $\xi_i\in W_i^{(k)}$ $(i=1, \cdots, n)$ for all $k=0, 1, \cdots$, $\{W_i^{(k)}\}_{k=0}^{\infty}$ contract to ξ_i $(i=1, \cdots, n)$, respectively. The conclusion (12) of the theorem is proved.

In the proof of the theorem we have used the following lemma and now give a proof.

Lemma. *Suppose that* $0<\alpha\leqslant\frac{1}{3}$, $t\geqslant 1$, *then*

$$A(t,\alpha)=\left(\frac{1}{3\alpha}+\frac{2\alpha}{t}\right)\left(1+\frac{2\alpha}{t}\right)^t-\left(\frac{1}{3\alpha}+\frac{\alpha}{t}\right)\left(1+\frac{\alpha}{t}\right)^t$$

is a decreasing function of t.

Proof. Let

$$f_y(t)=\left(\frac{1}{3\alpha}+y\frac{\alpha}{t}\right)\left(1+y\frac{\alpha}{t}\right)^t, \quad 0\leqslant y\leqslant 3,\ t\geqslant 1$$

for fixed α, $0<\alpha\leqslant\frac{1}{3}$. Clearly,

$$A(t,\alpha)=f_2(t)-f_1(t).$$

Hence, it is enough to prove

$$f_2'(t)<f_1'(t), \quad t\geqslant 1.$$

Let

$$g(y)=f_y'(t)$$

for fixed $t\geqslant 1$. It is enough to prove

$$g'(y)\leqslant 0,\quad 0\leqslant y\leqslant 3.$$

Now

$$g(y)=\left(1+\frac{\alpha}{t}y\right)^t h(y),$$

where

$$h(y)=-\frac{\alpha}{t^2}y+\left(\frac{1}{3\alpha}+\frac{\alpha}{t}y\right)\left[\ln\left(1+\frac{\alpha}{t}y\right)-\frac{\alpha y}{t+\alpha y}\right].$$

From

$$\frac{1}{\alpha}h'(y)=-\frac{1}{t^2}+\frac{1}{t}\ln\left(1+\frac{\alpha}{t}y\right)+\left(\frac{1}{3}-\alpha\right)\frac{y}{(t+\alpha y)^2}$$
$$\leqslant\frac{1}{t^2}\left\{\alpha y+\left(\frac{1}{3}-\alpha\right)y-1\right\}=\frac{1}{t^2}\left(\frac{y}{3}-1\right)\leqslant 0,\quad 0\leqslant y\leqslant 3,$$

obtain

$$h(y)\leqslant h(0)=0,\quad 0\leqslant y\leqslant 3.$$

Hence

$$g'(y)=h(y)\frac{d}{dy}\left(1+\frac{\alpha}{t}y\right)^{\alpha}+\left(1+\frac{\alpha}{t}y\right)^t h'(y)\leqslant 0,\quad 0\leqslant y\leqslant 3.$$

The lemma is proved.

References

[1] E. Durand, Solutions Numérigues des Équations Algébrigues, Tome I: Équations du Type $F(x)=0$; Racines dun Polynôme, Masson, Paris, 1960.
[2] I. O. Kerner, *Numer. Math.*, **8** (1966), 290—294.
[3] D. Braess, K. P. Hadeler, *Numer. Math.*, **21** (1973), 161—165.
[4] M. S. Petković, *J. Comp. Appl. Math.*, **8** (1982), 51—56.
[5] I. Gargantini, P. Henrici, *Numer. Math.*, **18** (1972), 305—332.