Computing 62, 129–145 (1999)

Computing

A Modification of Newton's Method for Analytic Mappings Having Multiple Zeros

P. Kravanja and **A. Haegemans,** Heverlee

Received August 17, 1998

Abstract

We propose a modification of Newton's method for computing multiple roots of systems of analytic equations. Under mild assumptions the iteration converges quadratically. It involves certain constants whose product is a lower bound for the multiplicity of the root. As these constants are usually not known in advance, we devise an iteration in which not only an approximation for the root is refined, but also approximations for these constants. Numerical examples illustrate the effectiveness of our approach.

AMS Subject Classification: 65H10.

Key Words: Newton's method, analytic mappings, multiple zeros, Van de Vel's iteration.

1. Introduction

Consider a smooth function $f : \mathbf{C} \to \mathbf{C}$ that has a zero of multiplicity μ at the point $z^\star$. If $\mu = 1$, then Newton's method converges quadratically to $z^\star$ if the initial iterate is sufficiently close to $z^\star$. If $\mu > 1$, then the convergence is only linear. In the latter case, if μ is known in advance, quadratic convergence can be regained by considering the iteration

$$z^{(p+1)} = z^{(p)} - \mu \frac{f(z^{(p)})}{f'(z^{(p)})}, \qquad p = 0, 1, 2, \ldots . \tag{1}$$

Van de Vel [40, 41] devised an iteration in which not only an approximation for the zero is refined, but also an estimate of its multiplicity. King [20] analysed Van de Vel's method and proved that its order of convergence is 1.554. He rearranged the order of the calculations and obtained an iteration that has order of convergence 1.618. This iteration proceeds as follows:

$$\begin{cases} \mu^{(p+1)} &= \dfrac{u(z^{(p)})}{u(z^{(p)}) - u(z^{(p+1)})}\mu^{(p)} \\ z^{(p+2)} &= z^{(p+1)} - \mu^{(p+1)} u(z^{(p+1)}) \end{cases} \tag{2}$$

for $p = 0, 1, 2, \ldots$, with initial $z^{(0)}$ and $\mu^{(0)}$, and after one preliminary quasi-Newton step $z^{(1)} = z^{(0)} - \mu^{(0)} u(z^{(0)})$, and where $u(z) := f(z)/f'(z)$.

In this paper we generalize these two results to the multivariate case. We consider systems of analytic equations, i.e., analytic mappings $f : \mathbf{C}^n \to \mathbf{C}^n$. The multidimensional version of (1) is formulated in Theorem 5. Instead of μ the iteration now involves a diagonal matrix containing certain constants $k_1, \ldots, k_n$ that are called orders. The product of these orders is a lower bound for μ (Theorem 4). In Section 4 we present an algorithm in which an approximation for the zero as well as approximations for the orders are refined iteratively. This iteration is our multidimensional generalization of (2). A lot of numerical examples illustrate our results.

At a multiple zero the Jacobian matrix of f is singular. The set of points $z \in \mathbf{C}^n$ such that $\det f'(z) = 0$ is a codimension one smooth manifold through the zero. As soon as an iterate lies on this manifold, the iteration breaks down. We assume throughout this paper that this does not happen. In other words, we assume that the initial iterate is such that the iteration is well defined at every step. This assumption enables us to focus entirely on the order of convergence.

The behaviour of Newton's method in case the Jacobian is singular at the zero has been analysed extensively in the literature [4–7, 9–13, 19, 34–36]. Many sufficient conditions for its convergence have been formulated. Under certain regularity and smoothness assumptions, the existence of special regions (cones, starlike regions) about the zero $z^\star$ has been proven. The Jacobian f' is regular in every point of these regions except in $z^\star$. If the initial iterate lies in such a region, then the Newton iterates will remain in this region and converge (linearly) to $z^\star$. We have not investigated the existence of such regions for the iterations presented in this paper. Indeed, the proofs for the classical cases are extremely complex and it is not a priori clear how to extend them to the iterations of this paper.

In [7] a modification of Newton's method is proposed that produces a sequence $\{z^{(p)}\}_{p\geq 0}$ such that the subsequence $\{z^{(2p)}\}_{p\geq 0}$ converges quadratically to the zero. However, this method works only in case the dimension of the null space of the Jacobian at the zero is equal to 1 or 2, and the projector onto this null space is known explicitely. Other modifications of Newton's method have been proposed in [4, 19]. These methods result in superlinear or quadratic convergence but again require rather restrictive hypotheses to be satisfied and need additional information (certain constants, projectors, etc.) that is usually not available.

In [30–32] a so-called deflation algorithm was proposed for computing multiple roots of systems of nonlinear algebraic equations. The system to be solved is replaced by another one having the same root but with a lower multiplicity. While the deflation algorithm proceeds, the multiplicity is systematically reduced until it is equal to one and classical methods can be applied. However, this algorithm requires symbolic calculation and works only for systems of algebraic equations.

Other approaches that have been proposed include so-called bordering methods [14–16, 21, 22], enlargement methods [29, 38, 42] and homotopy continuation methods [25–27].

All these methods require deciding whether the problem is singular: one should know in advance that the zero is multiple. This probably makes these methods unsuitable for general purpose use. As we will illustrate in Example 7, our method also works in case the zero is simple. Moreover, it requires no additional information, works under mild assumptions, and provides a lower bound for the multiplicity of the zero.

2. Preliminaries and Notation

Let $f = f(z) : \mathbf{C}^n \to \mathbf{C}^n$ be an analytic mapping, with $z = (z_1, \ldots, z_n)$ and $f = (f_1, \ldots, f_n)$. A point $z^\star \in \mathbf{C}^n$ is called a *zero* of f if $f(z^\star) = 0$. An isolated zero $z^\star$ of f is called *simple* if the Jacobian matrix of f at $z^\star$ is regular, $\det f'(z^\star) \neq 0$.

The following material is taken from the well-known book by Aĭzenberg and Yuzhakov [1].

Proposition 1. *If the closure of a neighbourhood $U_{z^\star}$ of a zero $z^\star$ of f does not contain other zeros of f, then there exists an $\epsilon > 0$ such that for almost all $\zeta \in \mathbf{C}^n$, $\|\zeta\|_2 < \epsilon$, the mapping*

$$z \mapsto f(z) - \zeta \tag{3}$$

has only simple zeros in $U_{z^\star}$ and their number depends neither on ζ nor on the choice of the neighbourhood $U_{z^\star}$.

The number of zeros of the mapping (3) in $U_{z^\star}$ indicated in this proposition is called the *multiplicity* of the zero $z^\star$ of f and is denoted by $\mu_{z^\star}(f)$. In other words, the multiplicity of an isolated zero of an analytic mapping is given by the number of simple zeros into which this zero desintegrates under a sufficiently small perturbation of the mapping.

The next result follows from the local invertibility of an analytic mapping at points where its Jacobian matrix is regular.

Proposition 2. *The multiplicity of a simple zero is equal to* 1.

Proposition 3. *If $z^\star$ is an isolated zero of f and* $\det f'(z^\star) = 0$, *then its multiplicity $\mu_{z^\star}(f)$ is larger than* 1. This statement justifies calling an isolated zero $z^\star$ of f *multiple* in case $\det f'(z^\star) = 0$.

Now let $z^\star = (z_1^\star, \dots, z_n^\star)$ be an isolated zero of $f = (f_1, \dots, f_n)$ such that

$$f_j(z) = \sum_{|\alpha| \geq k_j} c_{j,\alpha}(z - z^\star)^\alpha$$

for $j = 1, \dots, n$ where α is a multi-index, $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbf{N}^n$, $|\alpha| = \alpha_1 + \dots + \alpha_n$ and $(z - z^\star)^\alpha = (z_1 - z_1^\star)^{\alpha_1} \cdots (z_n - z_n^\star)^{\alpha_n}$. We call k_j the *order* of $z^\star$ as a zero of f_j. Define

$$P_j(z) := \sum_{|\alpha| = k_j} c_{j,\alpha}(z - z^\star)^\alpha$$

for $j = 1, \dots, n$. The homogeneous polynomial mapping

$$P = P(z) := (P_1(z), \dots, P_n(z)) \tag{4}$$

is called the *homogeneous principal part* of f at $z^\star$. The following theorem by Tsikh and Yuzhakov relates the multiplicity $\mu_{z^\star}(f)$ to the orders $k_1, \dots, k_n$.

Theorem 4. *The multiplicity of an isolated zero $z^\star$ of f is equal to the product of the orders of $z^\star$ as a zero of $f_1, \dots, f_n$ if and only if $z^\star$ is an isolated zero of the mapping* (4). *Moreover, the inequality $\mu_{z^\star}(f) \geq k_1 \cdots k_n$ always holds.*

3. A Modification of Newton's Method

Theorem 5. *Suppose that $z^{(0)}$ is such that the iteration*

$$z^{(p+1)} = z^{(p)} - [f'(z^{(p)})]^{-1}\, diag\,(k_1, \dots, k_n)\, f(z^{(p)}), \qquad p = 0, 1, 2, \dots,$$

is well defined for every p. If $\det P'(z) \not\equiv 0$ *and if $z^{(0)}$ is sufficiently close to $z^\star$, then $z^{(p)}$ converges quadratically to $z^\star$. If* $\det P'(z) \equiv 0$, *then the convergence is only linear.*

As already mentioned in Section 1, the problem of analysing the possible structure of the set of initial iterates $z^{(0)}$ that lie in a neighbourhood of $z^\star$ and guarantee convergence of the iteration is of considerable difficulty and would require a paper in itself. In this paper we therefore restrict our attention entirely to the *order* of convergence.

Proof: Define

$$\hat{f}$$

for $j = 1, \dots, n$. As

$$f_j(z) = P_j(z) + \hat{f}_j(z) = \sum_{|\alpha|=k_j} c_{j,\alpha}(z - z^\star)^\alpha + \hat{f}_j(z) \tag{5}$$

for $j = 1, \dots, n$, it follows that

$$\frac{\partial f_j}{\partial z_k}(z) = \sum_{|\alpha|=k_j} \alpha_k c_{j,\alpha}(z_1 - z_1^\star)^{\alpha_1} \cdots (z_k - z_k^\star)^{\alpha_k - 1} \cdots (z_n - z_n^\star)^{\alpha_n} + \frac{\partial \hat{f}_j}{\partial z_k}(z)$$

for $j, k = 1, \dots, n$. Let $e^{(p)} = (e_{1,p}, \dots, e_{n,p}) := z^{(p)} - z^\star$. Then the iteration can be written as

$$f'(z^\star + e^{(p)})e^{(p+1)} = f'(z^\star + e^{(p)})e^{(p)} - \operatorname{diag}(k_1, \dots, k_n) f(z^\star + e^{(p)}). \tag{6}$$

The jth component of the vector appearing in the right hand side of (6) is given by

$$\begin{aligned}
g_j(e^{(p)}) &:= \sum_{k=1}^{n} \frac{\partial f_j}{\partial z_k}(z^\star + e^{(p)})e_{k,p} - k_j f_j(z^\star + e^{(p)}) \\
&= \sum_{k=1}^{n}\Big[\sum_{|\alpha|=k_j} \alpha_k c_{j,\alpha} e_{1,p}^{\alpha_1} \cdots e_{n,p}^{\alpha_n} + \frac{\partial \hat{f}_j}{\partial z_k}(z^\star + e^{(p)})e_{k,p}\Big] \\
&\qquad - k_j\Big[\sum_{|\alpha|=k_j} c_{j,\alpha} e_{1,p}^{\alpha_1} \cdots e_{n,p}^{\alpha_n} + \hat{f}_j(z^\star + e^{(p)})\Big] \\
&= \sum_{|\alpha|=k_j} \big(|\alpha| - k_j\big) c_{j,\alpha}\big[e^{(p)}\big]^\alpha + \sum_{k=1}^{n} \frac{\partial \hat{f}_j}{\partial z_k}(z^\star + e^{(p)})e_{k,p} - k_j \hat{f}_j(z^\star + e^{(p)}) \\
&= \sum_{k=1}^{n} \sum_{|\alpha|\ge k_j+1} \alpha_k c_{j,\alpha}\big[e^{(p)}\big]^\alpha - k_j \sum_{|\alpha|\ge k_j+1} c_{j,\alpha}\big[e^{(p)}\big]^\alpha \\
&= \sum_{|\alpha|\ge k_j+1} \big(|\alpha| - k_j\big) c_{j,\alpha}\big[e^{(p)}\big]^\alpha \\
&= \sum_{|\alpha|=k_j+1} c_{j,\alpha}\big[e^{(p)}\big]^\alpha + \sum_{|\alpha|>k_j+1} \big(|\alpha| - k_j\big) c_{j,\alpha}\big[e^{(p)}\big]^\alpha.
\end{aligned}$$

It follows that $|g_j(e^{(p)})| = \mathcal{O}(\, \|e^{(p)}\|^{k_j+1}\,)$ for $j = 1, \dots, n$.

One can easily verify that $\det P'(z)$ is a homogeneous polynomial in $z - z^\star$. Now there are two possibilities: either all its coefficients are equal to zero, $\det P' \equiv 0$, or $\det P'$ has degree $\sum_{j=1}^{n} k_j - n$. By using (5) we can write $\det f'$ as a sum of 2^n

determinants, including det P', and it follows readily that

$$\det f'(z^\star + e^{(p)}) = \begin{cases} \mathcal{O}(\|e^{(p)}\|^{\sum_{j=1}^n k_j - n}) & \text{if } \det P' \not\equiv 0, \\ \mathcal{O}(\|e^{(p)}\|^{\sum_{j=1}^n k_j - n + 1}) & \text{if } \det P' \equiv 0. \end{cases} \tag{7}$$

By setting $g := (g_1, \ldots, g_n)$ we can write (6) as

$$f'(z^\star + e^{(p)})e^{(p+1)} = g(e^{(p)}).$$

Cramer's rule implies that

$$e_j^{(p+1)} = \frac{1}{\det f'(z^\star + e^{(p)})} \times \tag{8}$$

$$\begin{vmatrix} \frac{\partial f_1}{\partial z_1}(z^\star + e^{(p)}) & \ldots & \frac{\partial f_1}{\partial z_{j-1}}(z^\star + e^{(p)}) & g_1(e^{(p)}) & \frac{\partial f_1}{\partial z_{j+1}}(z^\star + e^{(p)}) & \ldots & \frac{\partial f_1}{\partial z_n}(z^\star + e^{(p)}) \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \frac{\partial f_n}{\partial z_1}(z^\star + e^{(p)}) & \ldots & \frac{\partial f_n}{\partial z_{j-1}}(z^\star + e^{(p)}) & g_n(e^{(p)}) & \frac{\partial f_n}{\partial z_{j+1}}(z^\star + e^{(p)}) & \ldots & \frac{\partial f_n}{\partial z_n}(z^\star + e^{(p)}) \end{vmatrix}$$

for $j = 1, \ldots, n$. The denominator in the right hand side of (8) was examined in (7). One can easily verify that the numerator is $\mathcal{O}(\|e^{(p)}\|^\alpha)$ with $\alpha = (\sum_{j=1}^n k_j - n) - (k_j - 1) + (k_j + 1) = \sum_{j=1}^n k_j - n + 2$. This proves the theorem. □

In the following examples we used Mathematica 2.2. All the calculations were done in multiple precision arithmetic.

Example 1. The mapping $f = (f_1, f_2) = (z_1 \sin z_1 + z_2^3, z_2 + z_1 \sin z_2)$ has an isolated zero at $z^\star = (0, 0)$. The orders are $k_1 = 2$ and $k_2 = 1$. The homogeneous principal part of f at $z^\star$ is given by $P = P(z_1, z_2) = (z_1^2, z_2)$. It follows that $z^\star$ is an isolated zero of P and thus, according to Theorem 4, the multiplicity of $z^\star$ as a zero of f is equal to $k_1 \cdot k_2 = 2$. The Jacobian matrix of P is given by

$$P'(z_1, z_2) = \begin{bmatrix} 2z_1 & 0 \\ 0 & 1 \end{bmatrix}$$

and thus $\det P'(z_1, z_2) \not\equiv 0$. Therefore the iteration of Theorem 5 will converge quadratically, as illustrated in Table 1. The initial iterate was $z^{(0)} := (0.2, 0.2)$.

Example 2. The mapping $f = (f_1, f_2) = (z_1 z_2 + (\sin z_1)^2 + z_2^3, \sin z_1 \sin z_2)$ has an isolated zero at $z^\star = (0, 0)$. The orders are $k_1 = 2$ and $k_2 = 2$. The homogeneous principal part of f at $z^\star$ is given by $P = P(z_1, z_2) = (z_1^2 + z_1 z_2, z_1 z_2)$. As $P(z_1, z_2) = 0$ if and only if $z_1 = 0$ and z_2 arbitrary, it follows that $z^\star$ is not an isolated zero of P and thus, according to Theorem 4, the multiplicity of $z^\star$ as a zero of f is strictly larger than $k_1 \cdot k_2 = 4$. The Jacobian matrix of P is given by

$$P'(z_1, z_2) = \begin{bmatrix} 2z_1 + z_2 & z_1 \\ z_2 & z_1 \end{bmatrix}$$

Table 1

p	$-\log_{10}\sqrt{\lvert z_1^{(p)}\rvert^2+\lvert z_2^{(p)}\rvert^2}$
0	0.6
1	1.5
2	2.8
3	6.4
4	11.1
5	26.0
6	$\vdots$

and thus $\det P'(z_1, z_2) \not\equiv 0$. Therefore the iteration of Theorem 5 will converge quadratically (see Table 2). The initial iterate was $z^{(0)} := (0.2, 0.2)$.

Table 2

p	$-\log_{10}\sqrt{\lvert z_1^{(p)}\rvert^2+\lvert z_2^{(p)}\rvert^2}$
0	0.6
1	1.3
2	2.9
3	6.2
4	12.6
5	25.6
6	$\vdots$

Example 3. The mapping $f = (f_1, f_2) = (z_1 + z_2 + z_1^2 + z_1 z_2 + 2z_2^3 + (\sin z_1)^3, 2(z_1 + z_2)^3 + z_1^4)$ has an isolated zero at $z^\star = (0, 0)$. The orders are $k_1 = 1$ and $k_2 = 3$. The homogeneous principal part of f at $z^\star$ is given by $P = P(z_1, z_2) = (z_1 + z_2, 2(z_1 + z_2)^3)$. As $P(z_1, z_2) = 0$ if and only if $z_2 = -z_1$, it follows that $z^\star$ is not an isolated zero of P and thus, according to Theorem 4, the multiplicity of $z^\star$ as a zero of f is strictly larger than $k_1 \cdot k_2 = 3$. The Jacobian matrix of P is given by

$$P'(z_1, z_2) = \begin{bmatrix} 1 & 1 \\ 6(z_1 + z_2)^2 & 6(z_1 + z_2)^2 \end{bmatrix}$$

and thus $\det P'(z_1, z_2) \equiv 0$. Therefore the iteration of Theorem 5 will converge linearly (Table 3). The initial iterate was $z^{(0)} := (0.2, 0.2)$.

Table 3

p	$-\log_{10}\sqrt{\lvert z_1^{(p)}\rvert^2 + \lvert z_2^{(p)}\rvert^2}$
0	0.6
1	0.1
2	0.6
3	0.8
4	1.4
5	2.0
$\vdots$	$\vdots$
11	5.6
12	6.2
13	6.8
14	7.4
15	8.0
$\vdots$	$\vdots$

Remark. In the previous examples we have not considered the case that det $P'(z) \equiv 0$ and $z^\star$ is an isolated zero of P. In fact, this situation cannot occur. Proposition 1 immediately implies that det $P'(z)$ cannot be identically equal to zero near $z^\star$ if $z^\star$ is an isolated zero of P. Thus det $P'(z) \equiv 0$ implies that $z^\star$ is not an isolated zero of P.

Let $d_1, \ldots, d_n \in \mathbf{C}_0$. Consider the iteration

$$z^{(p+1)} = z^{(p)} - [f'(z^{(p)})]^{-1} \operatorname{diag}(d_1, \ldots, d_n) f(z^{(p)}), \qquad p = 0, 1, 2, \ldots, \tag{9}$$

or, equivalently,

$$f'(z^\star + e^{(p)})e^{(p+1)} = f'(z^\star + e^{(p)})e^{(p)} - \operatorname{diag}(d_1, \ldots, d_n) f(z^\star + e^{(p)}), \tag{10}$$

where $e^{(p)} := z^{(p)} - z^\star$ for $p = 0, 1, 2, \ldots$. Let $g_j(e^{(p)})$ be the jth component of the vector appearing in the right hand side of (10). Using the same reasoning as in the proof of Theorem 5, one can easily show that

$$g_j(e^{(p)}) = \sum_{|\alpha| \geq k_j} \big(|\alpha| - d_j\big) c_{j,\alpha} \big[e^{(p)}\big]^\alpha.$$

It follows that $|g_j(e^{(p)})| = \mathcal{O}(\|e^{(p)}\|^{k_j})$ for $j = 1, \ldots, n$ and thus the iteration (9) converges only linearly to $z^\star$ (if det $P'(z) \not\equiv 0$ and if $z^{(0)}$ is sufficiently close to $z^\star$).

The special choice $d_1 = k_1, \dots, d_n = k_n$ gives quadratic convergence. But of course, the orders $k_1, \dots, k_n$ are usually not known in advance.

4. The Algorithm

The following proposition will help us to devise an iteration for the unknown orders $k_1, \dots, k_n$.

Proposition 6. *Let* $e \in \mathbf{C}^n$. *Then*

$$P'(z^\star + e)e = \mathit{diag}\,(k_1, \dots, k_n)P(z^\star + e).$$

Proof: Suppose $e =: (e_1, \dots, e_n)$. Then the jth component of $P'(z^\star + e)e$ is given by

$$\sum_{k=1}^{n} \frac{\partial P_j}{\partial z_k}(z^\star + e)e_k = \sum_{|\alpha|=k_j} |\alpha| c_{j,\alpha} e^\alpha = k_j P_j(z^\star + e)$$

for $j = 1, \dots, n$. This proves the proposition. □

Let $K := \operatorname{diag}(k_1, \dots, k_n)$ and suppose $z^{(p)}$ is our current approximation to $z^\star$. Then, by the previous proposition,

$$P'(z^{(p)})(z^{(p)} - z^\star) = K P(z^{(p)}).$$

The next iterate $z^{(p+1)}$ is defined as

$$z^{(p+1)} := z^{(p)} - [f'(z^{(p)})]^{-1} D^{(p)} f(z^{(p)}) \tag{11}$$

where $D^{(p)} := \operatorname{diag}(d_1^{(p)}, \dots, d_n^{(p)})$ contains our current approximations to the orders $k_1, \dots, k_n$. Then

$$P'(z^{(p+1)})(z^{(p+1)} - z^\star) = K P(z^{(p+1)}).$$

Suppose that the matrices $P'(z^{(p)})$ and $P'(z^{(p+1)})$ are regular. Then

$$z^{(p)} - z^\star = [P'(z^{(p)})]^{-1} K P(z^{(p)}) \tag{12}$$

and

$$z^{(p+1)} - z^\star = [P'(z^{(p+1)})]^{-1} K P(z^{(p+1)}). \tag{13}$$

By subtracting (12) and (13), and using (11) we obtain that

$$[P'(z^{(p)})]^{-1} K P(z^{(p)}) - [P'(z^{(p+1)})]^{-1} K P(z^{(p+1)}) = [f'(z^{(p)})]^{-1} D^{(p)} f(z^{(p)}).$$

If we replace P by f then this relation will be satisfied only approximatively. We use the resulting equation to define our next approximation $D^{(p+1)}$ to K:

$$D^{(p+1)}f(z^{(p)}) - f'(z^{(p)})[f'(z^{(p+1)})]^{-1}D^{(p+1)}f(z^{(p+1)}) = D^{(p)}f(z^{(p)}).$$

This equation is solved for the diagonal matrix $D^{(p+1)}$ in the following way. Let

$$F(z) := \mathrm{diag}\,(f_1(z), \ldots , f_n(z)), \qquad d^{(p)} := \begin{bmatrix} d_1^{(p)} \\ \vdots \\ d_n^{(p)} \end{bmatrix}$$

and define $d^{(p+1)}$ in a similar way. Obviously

$$D^{(p)}f(z^{(p)}) = F(z^{(p)})d^{(p)}, \qquad D^{(p+1)}f(z^{(p+1)}) = F(z^{(p+1)})d^{(p+1)}, \qquad \text{etc.}$$

Therefore

$$\left[\, F(z^{(p)}) - f'(z^{(p)})[f'(z^{(p+1)})]^{-1}F(z^{(p+1)}) \,\right]d^{(p+1)} = F(z^{(p)})d^{(p)}. \qquad (14)$$

This is the formula that we will use to calculate $d^{(p+1)}$ from $d^{(p)}$, $z^{(p)}$ and $z^{(p+1)}$. If we define the matrix-valued function

$$U(z) := [f'(z)]^{-1}\,\mathrm{diag}\,(f_1(z), \ldots , f_n(z))$$

for every $z \in \mathbf{C}^n$ such that $f'(z)$ is regular, (14) can be written as

$$\left[\, U(z^{(p)}) - U(z^{(p+1)}) \,\right]d^{(p+1)} = U(z^{(p)})d^{(p)}. \qquad (15)$$

This is a multidimensional version of the iteration formula that was discovered by Van de Vel [40, 41]. Note the diagonal matrix in the definition of $U(z)$. If $n = 1$ then $U(z) = f(z)/f'(z)$. In the multidimensional case it is tempting to consider the vector-valued function $[f'(z)]^{-1}f(z)$ but, as we have just found out, one should replace the vector $f(z)$ by its corresponding diagonal matrix, to obtain a matrix-valued function $U(z)$. Now the iteration (11) can be written as

$$z^{(p+1)} = z^{(p)} - U(z^{(p)})d^{(p)}.$$

From the foregoing considerations we extract the following iterative procedure:

$$\begin{aligned} d^{(p+1)} &= \left[\, U(z^{(p)}) - U\big(z^{(p)} - U(z^{(p)})d^{(p)}\big) \,\right]^{-1} U(z^{(p)})d^{(p)} \\ z^{(p+1)} &= \big(z^{(p)} - U(z^{(p)})d^{(p)}\big) - U\big(z^{(p)} - U(z^{(p)})d^{(p)}\big)d^{(p+1)} \end{aligned}$$

for $p = 0, 1, 2, \ldots$, starting with initial estimates $z^{(0)}$ for the zero $z^\star$ and $d^{(0)}$ for the orders $[k_1 \;\cdots\; k_n]^T$. This is our generalization of Van de Vel's method. It is a two-point method with memory. An equivalent formulation is the following:

$$\begin{aligned} z^{(p+1/2)} &= z^{(p)} - U(z^{(p)})d^{(p)} \\ d^{(p+1)} &= \left[U(z^{(p)}) - U(z^{(p+1/2)}) \right]^{-1} U(z^{(p)})d^{(p)} \\ z^{(p+1)} &= z^{(p+1/2)} - U(z^{(p+1/2)})d^{(p+1)} \end{aligned}$$

for $p = 0, 1, 2, \ldots$. This leads to the following algorithm.

Algorithm (two-point version)
input $z^{(0)}, d^{(0)}$
for $p = 0, 1, 2, \ldots$

1. Solve $f'(z^{(p)})\Delta z^{(p)} = -\,\mathrm{diag}\,(d^{(p)})f(z^{(p)})$

   $z^{(p+1/2)} \leftarrow z^{(p)} + \Delta z^{(p)}$

2. Solve $\left[F(z^{(p)}) - f'(z^{(p)})[f'(z^{(p+1/2)})]^{-1}F(z^{(p+1/2)}) \right] d^{(p+1)} = F(z^{(p)})d^{(p)}$

3. Solve $f'(z^{(p+1/2)})\Delta z^{(p+1/2)} = -\,\mathrm{diag}\,(d^{(p+1)})f(z^{(p+1/2)})$

   $z^{(p+1)} \leftarrow z^{(p+1/2)} + \Delta z^{(p+1/2)}$

This method can be improved by noting that after step 1 is completed the first time, there is no reason ever to return to it. Instead the estimate of the orders can be improved (step 2) before each and every further quasi-Newton step (step 3). Thus the improved iteration may be written as

$$\begin{aligned} d^{(p+1)} &= \left[U(z^{(p)}) - U(z^{(p+1)}) \right]^{-1} U(z^{(p)})d^{(p)} \\ z^{(p+2)} &= z^{(p+1)} - U(z^{(p+1)})d^{(p+1)} \end{aligned}$$

for $p = 0, 1, 2, \ldots$, with initial $z^{(0)}$ and $d^{(0)}$, and after one preliminary quasi-Newton step $z^{(1)} = z^{(0)} - U(z^{(0)})d^{(0)}$. This is a one-point method with memory. It corresponds to King's improvement of Van de Vel's method [20].

Algorithm (one-point version)
input $z^{(0)}, d^{(0)}$
Solve $f'(z^{(0)})\Delta z^{(0)} = -\,\mathrm{diag}\,(d^{(0)})f(z^{(0)})$
$z^{(1)} \leftarrow z^{(0)} + \Delta z^{(0)}$
for $p = 0, 1, 2, \ldots$

1. Solve $\left[F(z^{(p)}) - f'(z^{(p)})[f'(z^{(p+1)})]^{-1}F(z^{(p+1)}) \right] d^{(p+1)} = F(z^{(p)})d^{(p)}$

2. Solve $f'(z^{(p+1)})\Delta z^{(p+1)} = -\,\mathrm{diag}\,(d^{(p+1)})f(z^{(p+1)})$

$z^{(p+2)} \leftarrow z^{(p+1)} + \Delta z^{(p+1)}$

Define

$$\zeta_p := \begin{cases} -\log_{10} \frac{\|z^{(p)} - z^\star\|}{\|z^\star\|} & \text{if } z^\star \neq 0 \\ -\log_{10} \|z^{(p)}\| & \text{if } z^\star = 0 \end{cases}$$

for $p = 0, 1, 2, \ldots$. Also, let

$$\vec{k} := \begin{bmatrix} k_1 \\ \vdots \\ k_n \end{bmatrix}$$

and define

$$\delta_p := -\log_{10} \frac{\|d^{(p)} - \vec{k}\|}{\|\vec{k}\|}$$

for $p = 0, 1, 2, \ldots$. These notations will be used in the tables that correspond to the following examples.

Example 4. Let us reconsider the mapping of Example 1. Table 4 compares the two-point version of our algorithm with the one-point version. The initial iterates were $z^{(0)} := (0.2, 0.2)$ and $d^{(0)} := (1, 1)$. The one-point version is superior. Intuitively this is not a surprise, of course.

Table 4. Example 4

step	two-point version		one-point version		step
p	δ_p	ζ_p	δ_p	ζ_p	p
0		0.5		0.5	0
1/2		0.9	0.7	0.9	1
1	0.7	1.6	1.3	1.6	2
3/2		2.3	2.0	4.0	3
2	2.0	5.6	4.4	5.8	4
5/2		7.3	1.8	9.7	5
3	4.5	11.9	10.4	11.5	6
7/2		16.3	11.9	22.3	7
4	12.3	31.9	22.7	39.3	8
9/2		43.8	$\vdots$	61.6	9
5	28.6	77.5		$\vdots$	10
11/2		$\vdots$			

Example 5. Next we reconsider the mapping of Example 2, but shifted to the point $z^\star = (1, 3)$. In other words, suppose $f = (f_1, f_2) = (uv + (\sin u)^2 + v^3, \sin u \sin v)$ where $u = z_1 - 1$ and $v = z_2 - 3$. Table 5 compares both versions of our algorithm. The initial iterates were $z^{(0)} := (1.1, 3.1)$ and $d^{(0)} := (1, 1)$. Again the one-point version is superior.

Table 5. Example 5

step	two-point	version	one-point	version	step
p	δ_p	ζ_p	δ_p	ζ_p	p
0		1.3		1.3	0
1/2		1.7	1.6	1.7	1
1	1.6	2.9	2.2	2.9	2
3/2		5.5	1.9	5.9	3
2	2.0	6.9	7.4	7.9	4
5/2		8.0	9.6	15.0	5
3	6.2	13.6	16.5	24.6	6
7/2		19.0		⋮	7
4	13.0	⋮			

Example 6. The mapping $f = (f_1, f_2, f_3) = (u^2 + u^2 \sin v + u^3 \sin w, v + uv + v^2 + u^2 \sin u, w^2 + u^3 + vw \sin w + v^4 + u^5)$ where $u = z_1 - 1$, $v = z_2 - 2$ and $w = z_3 - 5$ has an isolated zero at $z^\star = (1, 2, 5)$. The orders are $k_1 = 2$, $k_2 = 1$ and $k_3 = 2$. The homogeneous principal part of f at $z^\star$ is given by $P = P(z_1, z_2, z_3) = ((z_1 - 1)^2, z_2 - 2, (z_3 - 5)^2)$. It follows that $z^\star$ is an isolated zero of P and thus, according to Theorem 4, the multiplicity of $z^\star$ as a zero of f is equal to $k_1 \cdot k_2 \cdot k_3 = 4$. Table 6 gives the results obtained by the one-point version of our algorithm. The initial iterates were $z^{(0)} := (1.2, 2.2, 5.2)$ and $d^{(0)} := (1, 1, 1)$.

Example 7. In our last example we consider a mapping that has a simple zero. The mapping $f = (f_1, f_2, f_3) = (u + u^2 + vw + \sin u \sin w + v^3, v + uv + v^2 + vw + (\sin u)^3 + vw^2, w + uw + w^2 + u^2 \sin v + w^3)$ where $u = z_1 - 1$, $v = z_2 - 2$ and $w = z_3 - 5$ has an isolated zero at $z^\star = (1, 2, 5)$. The orders are $k_1 = 1$, $k_2 = 1$ and $k_3 = 1$. The homogeneous principal part of f at $z^\star$ is given by $P = P(z_1, z_2, z_3) = (z_1 - 1, z_2 - 2, z_3 - 5)$. It follows that $z^\star$ is an isolated zero of P and thus, according to Theorem 4, the multiplicity of $z^\star$ as a zero of f is equal to $k_1 \cdot k_2 \cdot k_3 = 1$. Therefore $\det f'(z^\star) \neq 0$. The iteration of Theorem 5 reduces to the classical Newton's method. Table 7 shows the performance of the one-point version of our algorithm. The initial iterates were $z^{(0)} := (1.2, 2.2, 5.2)$ and $d^{(0)} := (1, 1, 1)$.

Table 6. Example 6

p	δ_p	ζ_p
0		1.2
1	0.5	1.5
2	1.2	2.1
3	2.1	3.4
4	3.1	5.3
5	4.8	8.1
6	7.6	12.7
7	12.2	20.1
8	19.6	32.1
9	⋮	51.4
⋮	⋮	⋮

Table 7. Example 7

p	δ_p	ζ_p
0		1.2
1	0.2	1.7
2	1.1	2.2
3	1.3	3.0
4	2.1	4.2
5	3.3	6.2
6	5.4	9.5
7	8.8	14.8
8	14.1	23.4
9	22.8	37.3
10	⋮	69.9
11		⋮

Remark. As soon as the iterates $d_1^{(p)}, \ldots, d_n^{(p)}$ are sufficiently close to integers, one can determine the correct values of $k_1, \ldots, k_n$ and use the iteration of Theorem 5, of course.

Remark. King [20] analysed Van de Vel's method 40, rearranged the order of the calculations, and gave a convergence proof for both iterations. He proved that Van de Vel's method has order of convergence 1.554, and that his modification (one-point version) has order of convergence 1.618. The previous examples indicate that our multidimensional generalizations of these methods have the same corresponding order of convergence. We have been unable to generalize King's proof to the multidimensional setting. The main problem in the multidimensional case is that matrices occur instead of scalars and therefore commutativity gets lost. This is the subject of ongoing research.

4. Conclusions

In this paper we presented two iterations for computing multiple zeros of analytic mappings and studied their order of convergence.

The iteration of Theorem 5 is a multidimensional generalization of iteration (1). It can be used if the orders, which are positive integers related to the multiplicity of the zero, are known in advance. Under certain assumptions this iteration converges quadratically.

Our generalization of Van de Vel's iteration is based on Eq. (15). The algorithm requires no prior knowledge about the multiplicity of the zero, which may even be simple, and proceeds by iterating on the zero as well as on the orders. Our numerical experiments indicate that the one-point version has order of convergence 1.618.

Several challenging open problems remain:

- We gave only an informal derivation of our multidimensional generalization (15) of Van de Vel's iteration, and the question remains whether it is possible to generalize King's proof.

- Local convergence regions: what is the structure of the set of initial iterates $z^{(0)}$ that lie in a neighbourhood of the zero $z^\star$ and that guarantee convergence?

- In the one-dimensional case the (discrete) dynamics of Newton's method are studied in terms of Julia and Fatou sets, Siegel disks, etc. (See, for example, [3] or [2].) The beautiful fractal-like images that illustrate these results are well-known. A global convergence analysis for the iterations presented in this paper, in particular with respect to the role played by the singular manifold $\{ z \in \mathbf{C}^n : \det f'(z) = 0 \}$, is a very interesting (but very difficult!) challenge.

- Newton's method has been studied in terms of its continuous counterpart: a system of autonomous differential equations whose Euler discretization yields (the damped) Newton's method. (See, for example, [39], [33] or [23].) Is it possible to formulate a continuous-time version of the iterations presented in this paper, in particular for our generalization of Van de Vel's method, and how do the regions of attraction of these dynamical systems relate to the convergence regions of the discrete methods?

We hope to address these problems in future publications.

Acknowledgements

The authors thank Dr. J. M. Gutiérrez, Dr. ir. P. Verlinden and Prof. H. Van de Vel for stimulating discussions on the topic of this paper. Peter Kravanja was supported by a grant from the Flemish Institute for the Promotion of Scientific and Technological Research in the Industry (IWT).

References

[1] Aĭzenberg, I. A., Yuzhakov, A. P.: Integral representations and residues in multidimensional complex analysis. Providence: AMS, 1983.
[2] Bergweiler, W.: Iteration of meromorphic functions. Bull. Am. Math. Soc. *29*, 151–188 (1993).
[3] Blanchard, P.: Complex analytic dynamics on the Riemann sphere. Bull. Amer. Math. Soc. *11*, 85–141 (1984).
[4] Decker, D. W., Keller, H. B., Kelley, C. T.: Convergence rates for Newton's method at singular points. SIAM J. Numer. Anal. *20*, 296–314 (1983).

[5] Decker, D. W., Kelley, C. T.: Newton's method at singular points: I. SIAM J. Numer. Anal. *17*, 66–70 (1980).

[6] Decker, D. W., Kelley, C. T.: Newton's method at singular points: II. SIAM J. Numer. Anal. *17*, 465–471 (1980).

[7] Decker, D. W., Kelley, C. T.: Convergence acceleration for Newton's method at singular points. SIAM J. Numer. Anal. *19*, 219–229 (1981).

[8] Decker, D. W., Kelley, C. T.: Broyden's method for a class of problems having singular Jacobian at the root. SIAM J. Numer. Anal. *22*, 566–574 (1985).

[9] Decker, D. W., Kelley, C. T.: Expanded convergence domains for Newton's method at nearly singular roots. SIAM J. Sci. Stat. Comput. *6*, 951–966 (1985).

[10] Griewank, A.: Starlike domains of convergence for Newton's method at singularities. Numer. Math. *35*, 95–111 (1980).

[11] Griewank, A.: On solving nonlinear equations with simple singularities or nearly singular solutions. SIAM Rev. *27*, 537–563 (1985).

[12] Griewank, A., Osborne, M. R.: Newton's method for singular problems when the dimension of the null space is > 1. SIAM J. Numer. Anal. *18*, 145–149 (1981).

[13] Griewank, A., Osborne, M. R.: Analysis of Newton's method at irregular singularities. SIAM J. Numer. Anal. *20*, 747–773 (1983).

[14] Hoy, A.: A relation between Newton and Gauss-Newton steps for singular nonlinear equations. Computing *40*, 19–27 (1988).

[15] Hoy, A.: An efficiently implementable Gauss-Newton-like method for solving singular nonlinear equations. Computing *41*, 107–122 (1989).

[16] Hoy, A., Schwetlick, H.: Some superlinearly convergent methods for solving singular nonlinear equations. In: Computational solution of nonlinear systems of equations (Allgower, E. L., Georg, K., eds.), pp. 285–300. Rhode Island: SIAM, 1990.

[17] Keller, H. B.: Geometrically isolated nonisolated solutions and their approximation. SIAM J. Numer. Anal. *18*, 822–838 (1981).

[18] Kelley, C. T.: A Shamanskii-like acceleration scheme for nonlinear equations at singular roots. Math. Comput. *47*, 609–623 (1986).

[19] Kelley, C. T., Suresh, R.: A new acceleration method for Newton's method at singular points. SIAM J. Numer. Anal. *20*, 1001–1009 (1983).

[20] King, R. F.: Improving the Van de Vel root-finding method. Computing *30*, 373–378 (1983).

[21] Kunkel, P.: Efficient computation of singular points. IMA J. Numer. Anal. *9*, 421–433 (1989).

[22] Mei, Z.: A special extended system and a Newton-like method for simple singular nonlinear equations. Computing *45*, 157–167 (1990).

[23] Meier, H.-G.: Diskrete und kontinuierliche Newton-Systeme im Komplexen. PhD thesis, Rheinisch-Westfälische Technische Hochschule Aachen, 1991.

[24] Menzel, R., Pönisch, G.: A quadratically convergent method for computing simple singular roots and its application to determining simple bifurcation points. Computing *32*, 127–138 (1984).

[25] Morgan, A. P., Sommese, A. J., Wampler, C. W.: Computing singular solutions to nonlinear analytic systems. Numer. Math. *58*, 669–684 (1991).

[26] Morgan, A. P., Sommese, A. J., Wampler, C. W.: Computing singular solutions to polynomial systems. Adv. Appl. Math. *13*, 305–327 (1992).

[27] Morgan, A. P., Sommese, A. J., Wampler, C. W.: A power series method for computing singular solutions to nonlinear analytic systems. Numer. Math. *63*, 391–409 (1992).

[28] Nashed, M. Z., Chen, X.: Convergence of Newton-like methods for singular operator equations using outer inverses. Numer. Math. *66*, 235–257 (1993).

[29] Neta, B., Victory, H. D.: A higher order method for determining nonisolated solutions of a system of nonlinear equations. Computing *32*, 163–166 (1984).

[30] Ojika, T.: Modified deflation algorithm for the solution of singular problems. I. A system of nonlinear algebraic equations. J. Math. Anal. Appl. *123*, 199–221 (1987).

[31] Ojika, T.: Modified deflation algorithm for the solution of singular problems. II. Nonlinear multi-point boundary value problems. J. Math. Anal. Appl. *123*, 222–237 (1987).

[32] Ojika, T., Watanabe, S., Mitsui, T.: Deflation algorithm for the multiple roots of a system of nonlinear equations. J. Math. Anal. Appl. *96*, 463–479 (1983).

[33] Peitgen, H.-O., Prüfer, M., Schmitt, K.: Global aspects of the continuous and discrete Newton method: A case study. Acta Appl. Math. *13*, 123–202 (1988).

[34] Rall, L. B.: Convergence of the Newton process to multiple solutions. Numer. Math. *9*, 23–37 (1966).

[35] Reddien, G. W.: On Newton's method for singular problems. SIAM J. Numer. Anal. *15*, 993–996 (1978).

[36] Reddien, G. W.: Newton's method and high order singularities. Comput. Math. Appl. *5*, 79–86 (1979).
[37] Traub, J. F.: Iterative methods for the solution of equations. Englewood Cliffs: Prentice-Hall, 1964.
[38] Tsuchiya, T.: Enlargement procedure for resolution of singularities at simple singular solutions of nonlinear equations. Numer. Math. *52*, 401–411 (1988).
[39] Haeseler, F. v., Peitgen, H.-O.: Newton's method and complex dynamical systems. Acta Appl. Mat. *13*, 3–58 (1988).
[40] Vel, H. Van de: A method for computing a root of a single nonlinear equation, including its multiplicity. Computing *14*, 167–171 (1975).
[41] Straeten, M. Vander, Vel, H. Van de: Multiple root-finding methods. J. Comput. Appl. Math. *40*, 105–114 (1992).
[42] Weber, H., Werner, W.: On the accurate determination of nonisolated solutions of nonlinear equations. Computing *26*, 315–326 (1981).
[43] Ypma, T. J.: Finding a multiple zero by transformations and Newton-like methods. SIAM Rev. *25*, 365–378 (1983).

P. Kravanja and A. Haegemans
Department of Computer Science
Katholieke Universiteit Leuven
Celestijnenlaan 200 A
B-3001 Heverlee, Belgium
e-mail: Peter.Kravanja@na-net.ornl.gov
Ann.Haegemans@cs.kuleuven.ac.be