A modification of Newton's method for analytic mappings having multiple zeros

Peter Kravanja and Ann Haegemans

Report TW 266, August 1997

Katholieke Universiteit Leuven
Department of Computer Science
Celestijnenlaan 200A – B-3001 Heverlee (Belgium)

A modification of Newton's method for analytic mappings having multiple zeros

Peter Kravanja and Ann Haegemans

Report TW 266, August 1997

Department of Computer Science, K.U.Leuven

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.

Keywords : systems of analytic equations, multiple roots, Newton's method, Van de Vel's method
AMS(MOS) Classification : 65H10

A MODIFICATION OF NEWTON'S METHOD FOR ANALYTIC MAPPINGS HAVING MULTIPLE ZEROS

PETER KRAVANJA AND ANN HAEGEMANS

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.

1. INTRODUCTION

Consider a smooth function $f : \mathbb{C} \to \mathbb{C}$ 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 [35, 36] devised an iteration in which not only an approximation for the zero is refined, but also an estimate of its multiplicity. King [18] analysed Van de Vel's method and proved that its order of convergence is 1.554. He improved the algorithm and obtained an iteration that has order of convergence 1.618. In this paper we will generalize these results to the multivariate case. We will consider systems of analytic equations, i.e., analytic mappings $f : \mathbb{C}^n \to \mathbb{C}^n$. 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. A lot of numerical examples illustrate our results.

At a multiple zero the Jacobian matrix of f is singular. The set of points $z \in \mathbb{C}^n$ 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 will assume throughout this paper that this does not happen. In other words, we will assume that the initial iterate is such that the iteration is well defined at every step.

The behaviour of Newton's method in case the Jacobian is singular at the zero has been analysed extensively in the literature [30, 31, 32, 3, 4, 8, 5, 10, 11, 2, 17, 7, 9]. 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.

In [5] 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

---

Date: August 1997.

1991 *Mathematics Subject Classification.* 65H10.

have been proposed in [2, 17]. These methods result in superlinear or quadratic convergence but again require rather restrictive hypotheses to be satisfied and need additional information (certain constants, projectors, ...) that is usually not available.

In [29, 27, 28] 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 [12, 13, 19, 20, 14], enlargement methods [37, 26, 34] and homotopy continuation methods [24, 22, 23].

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 $f = f(z) : \mathbb{C}^n \to \mathbb{C}^n$ be an analytic mapping, with $z = (z_1, \dots, z_n)$ and $f = (f_1, \dots, f_n)$. A point $z^\star \in \mathbb{C}^n$ 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 $\zeta \in \mathbb{C}^n$, $\|\zeta\|_2 < \epsilon$, the mapping*

$$z \mapsto f(z) - \zeta \tag{2}$$

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 (2) 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* if $\det f'(z^\star) = 0$.

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, $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$, $|\alpha| = \alpha_1 + \dots + \alpha_n$ 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

$$P = P(z) := (P_1(z), \dots, P_n(z)) \tag{3}$$

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 (3). 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*

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

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.

Proof. Define

$$\hat{f}_j(z) := \sum_{|\alpha| \geq k_j+1} c_{j,\alpha}(z - z^\star)^\alpha$$

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

$$f_j(z) = P_j(z) + \hat{f}_j(z) = \sum_{|\alpha| = k_j} c_{j,\alpha}(z - z^\star)^\alpha + \hat{f}_j(z) \tag{4}$$

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

$$f'(z^\star + e^{(p)}) e^{(p+1)} = f'(z^\star + e^{(p)}) e^{(p)} - \operatorname{diag}(k_1, \dots, k_n) f(z^\star + e^{(p)}). \tag{5}$$

The jth component of the vector appearing in the right hand side of (5) is given by

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

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 (4) we can write $\det f'$ as a sum of 2^n determinants, including $\det P'$, and it follows readily that

$$\det f'(z^\star + e^{(p)}) = \begin{cases} \mathcal{O}(\|e^{(p)}\|^{\sum_{j=1}^{n} k_j - n}) & \text{if } \det P' \not\equiv 0, \\ \mathcal{O}(\|e^{(p)}\|^{\sum_{j=1}^{n} k_j - n + 1}) & \text{if } \det P' \equiv 0. \end{cases} \tag{6}$$

By setting $g := (g_1, \ldots, g_n)$ we can write (5) as

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

Cramer's rule implies that

$$e_j^{(p+1)} = \frac{1}{\det f'(z^\star + e^{(p)})} \times \tag{7}$$

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

for $j = 1, \ldots, n$. The denominator in the right hand side of (7) was examined in (6). 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 have used Mathematica 2.2. All the calculations have been 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. The following table illustrates this. The initial iterate was $z^{(0)} := (0.2, 0.2)$.

p	$-\log_{10}\lvert z_1^{(p)}\rvert$	$-\log_{10}\lvert z_2^{(p)}\rvert$	$-\log_{10}\sqrt{\lvert z_1^{(p)}\rvert^2 + \lvert z_2^{(p)}\rvert^2}$
0	0.7	0.7	0.6
1	2.0	1.5	1.5
2	2.8	3.6	2.8
3	8.4	6.4	6.4
4	11.1	14.8	11.1
5	27.1	26.0	26.0
6	∞	∞	∞

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

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

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

p	$-\log_{10}\lvert z_1^{(p)}\rvert$	$-\log_{10}\lvert z_2^{(p)}\rvert$	$-\log_{10}\sqrt{\lvert z_1^{(p)}\rvert^2+\lvert z_2^{(p)}\rvert^2}$
0	0.7	0.7	0.6
1	1.5	1.4	1.3
2	3.1	3.0	2.9
3	6.3	6.3	6.2
4	12.9	12.8	12.6
5	25.8	25.7	25.6
6	∞	∞	∞

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. The following table illustrates this. The initial iterate was $z^{(0)} := (0.2, 0.2)$.

p	$-\log_{10}\lvert z_1^{(p)}\rvert$	$-\log_{10}\lvert z_2^{(p)}\rvert$	$-\log_{10}\sqrt{\lvert z_1^{(p)}\rvert^2+\lvert z_2^{(p)}\rvert^2}$
0	0.7	0.7	0.6
1	0.3	0.2	0.1
2	0.9	0.6	0.6
3	1.0	0.9	0.8
4	1.6	1.5	1.4
5	2.2	2.2	2.0
$\vdots$	$\vdots$	$\vdots$	$\vdots$
11	5.8	5.8	5.6
12	6.4	6.4	6.2
13	7.0	7.0	6.8
14	7.6	7.6	7.4
15	8.2	8.2	8.0
$\vdots$	$\vdots$	$\vdots$	$\vdots$

Remark. In the previous examples we did not consider 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 $d_1, \dots, d_n \in \mathbb{C}_0$. Consider the iteration

$$z^{(p+1)} = z^{(p)} - [f'(z^{(p)})]^{-1} \operatorname{diag}(d_1, \dots, d_n) f(z^{(p)}), \qquad p = 0, 1, 2, \dots, \tag{8}$$

or, equivalently,

$$f'(z^\star + e^{(p)})e^{(p+1)} = f'(z^\star + e^{(p)})e^{(p)} - \operatorname{diag}(d_1, \dots, d_n) f(z^\star + e^{(p)}), \tag{9}$$

where $e^{(p)} := z^{(p)} - z^\star$ for $p = 0, 1, 2, \dots$. Let $g_j(e^{(p)})$ be the jth component of the vector appearing in the right hand side of (9). Using the same reasoning as in the proof of Theorem 5, one can easily show that

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

It follows that $|g_j(e^{(p)})| = \mathcal{O}(\|e^{(p)}\|^{k_j})$ for $j = 1, \dots, n$ and thus the iteration (8) 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 $e \in \mathbb{C}^n$. Then*

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

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

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

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

$$z^{(p+1)} := z^{(p)} - [f'(z^{(p)})]^{-1} D^{(p)} f(z^{(p)}) \tag{10}$$

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

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

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

$$z^{(p)} - z^\star = [P'(z^{(p)})]^{-1} KP(z^{(p)}) \tag{11}$$

and

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

By subtracting (11) and (12), and using (10) we obtain that

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

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

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

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

$$\left[F(z^{(p)}) - f'(z^{(p)})[f'(z^{(p+1)})]^{-1} F(z^{(p+1)}) \right] d^{(p+1)} = F(z^{(p)}) d^{(p)}. \tag{13}$$

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

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

for every $z \in \mathbb{C}^n$ such that $f'(z)$ is regular, (13) can be written as

$$\left[U(z^{(p)}) - U(z^{(p+1)}) \right] d^{(p+1)} = U(z^{(p)}) d^{(p)}. \tag{14}$$

This is a multidimensional version of the iteration formula discovered by Van de Vel [35, 36]. 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 (10) can be written as

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

From the foregoing considerations we extract the following iterative procedure:

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

for $p = 0, 1, 2, \ldots$, starting with initial estimates $z^{(0)}$ for the zero $z^\star$ and $d^{(0)}$ for the orders $[k_1 \quad \cdots \quad k_n]^T$. 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 $f'(z^{(p)})\Delta z^{(p)} = -\operatorname{diag}(d^{(p)})f(z^{(p)})$
   $z^{(p+1/2)} \leftarrow z^{(p)} + \Delta z^{(p)}$
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 $f'(z^{(p+1/2)})\Delta z^{(p+1/2)} = -\operatorname{diag}(d^{(p+1)})f(z^{(p+1/2)})$
   $z^{(p+1)} \leftarrow z^{(p+1/2)} + \Delta z^{(p+1/2)}$

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 [18].

Algorithm (one-point version).

input $z^{(0)}$, $d^{(0)}$
Solve $f'(z^{(0)})\Delta z^{(0)} = -\operatorname{diag}(d^{(0)})f(z^{(0)})$
$z^{(1)} \leftarrow z^{(0)} + \Delta z^{(0)}$
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 $f'(z^{(p+1)})\Delta z^{(p+1)} = -\operatorname{diag}(d^{(p+1)})f(z^{(p+1)})$
   $z^{(p+2)} \leftarrow z^{(p+1)} + \Delta z^{(p+1)}$

Example 4. Let us reconsider the mapping of Example 1. Table 1 compares the two-point version of our algorithm with the one-point version. The columns labelled "k_1" and "k_2" contain $-\log_{10}|d_1^{(p)} - k_1|/k_1$ and $-\log_{10}|d_2^{(p)} - k_2|/k_2$, respectively. The columns labelled "$z_1^\star$" and "$z_2^\star$"

step	two-point version						one-point version						step
p	k_1	k_2	$\|\|k\|\|$	$z_1^\star$	$z_2^\star$	$\|\|z^\star\|\|$	k_1	k_2	$\|\|k\|\|$	$z_1^\star$	$z_2^\star$	$\|\|z^\star\|\|$	p
0				0.7	0.7	0.5				0.7	0.7	0.5	0
1/2				0.9	2.0	0.9	0.7	0.7	0.7	0.9	2.0	0.9	1
1	0.7	0.7	0.7	1.6	3.4	1.6	2.6	0.9	1.3	1.6	3.4	1.6	2
3/2				2.3	4.0	2.3	4.0	1.6	2.0	4.2	4.2	4.0	3
2	4.0	1.7	2.0	6.4	5.6	5.6	4.6	4.2	4.4	8.0	5.8	5.8	4
5/2				10.2	7.3	7.3	1.7	8.0	1.8	9.8	10.0	9.7	5
3	4.4	6.3	4.5	12.0	13.6	11.9	10.8	9.8	10.4	11.5	18.0	11.5	6
7/2				16.3	20.0	16.3	23.8	11.5	11.9	22.3	27.8	22.3	7
4	17.4	12.0	12.3	33.7	31.9	31.9	∞	22.3	22.7	46.1	39.3	39.3	8
9/2				51.0	44.0	43.8	∞	∞	∞	72.2	61.6	61.6	9
5	28.5	33.7	28.6	79.5	77.5	77.5				∞	∞	∞	10
11/2				∞	∞	∞							

TABLE 1. Example 4.

step	two-point version						one-point version						step
p	k_1	k_2	$\|\|k\|\|$	$z_1^\star$	$z_2^\star$	$\|\|z^\star\|\|$	k_1	k_2	$\|\|k\|\|$	$z_1^\star$	$z_2^\star$	$\|\|z^\star\|\|$	p
0				1.0	1.5	1.3				1.0	1.5	1.3	0
1/2				1.3	1.8	1.7	1.5	2.6	1.6	1.3	1.8	1.7	1
1	1.5	2.6	1.6	2.6	3.1	2.9	2.0	3.1	2.2	2.6	3.1	2.9	2
3/2				5.6	5.5	5.5	1.8	5.6	1.9	5.5	6.5	5.9	3
2	1.8	4.2	2.0	7.0	6.8	6.9	7.3	9.1	7.4	7.4	8.4	7.9	4
5/2				8.1	8.0	8.0	9.5	15.6	9.6	14.6	15.6	15.0	5
3	6.1	6.8	6.2	13.8	13.6	13.6	16.3	25.8	16.5	24.2	25.2	24.6	6
7/2				19.2	19.0	19.0				∞	∞	∞	7
4	12.8	17.7	13.0	∞	∞	∞							

TABLE 2. Example 5.

contain $-\log_{10}|z_1^{(p)}|$ and $-\log_{10}|z_2^{(p)}|$, respectively. (Remember that $z^\star = (0,0)$.) The columns labelled "$||k||$" and "$||z^\star||$" contain

$$-\log_{10}\sqrt{\frac{|d_1^{(p)}-k_1|^2+|d_2^{(p)}-k_2|^2}{k_1^2+k_2^2}} \quad \text{and} \quad -\log_{10}\sqrt{|z_1^{(p)}|^2+|z_2^{(p)}|^2}$$

respectively. 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.

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 2 compares both versions of our algorithm. The columns labelled "$z_1^\star$", "$z_2^\star$" and "$||z^\star||$" are now related to the (componentwise or normwise) *relative* errors. The initial iterates were $z^{(0)} := (1.1, 3.1)$ and $d^{(0)} := (1,1)$. Again the one-point version is superior.

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 $k_1 \cdot k_2 \cdot k_3 = 4$. We have used the one-point version of our algorithm. Table 3 contains minus the logarithm with base 10 of the (componentwise or normwise) relative errors. The initial iterates were $z^{(0)} := (1.2, 2.2, 5.2)$ and $d^{(0)} := (1,1,1)$.

p	k_1	k_2	k_3	$\|\|k\|\|$	$z_1^\star$	$z_2^\star$	$z_3^\star$	$\|\|z^\star\|\|$
0					0.7	1.0	1.4	1.2
1	0.6	0.4	0.5	0.5	0.9	1.7	1.6	1.5
2	1.6	0.8	1.3	1.2	1.7	2.5	2.1	2.1
3	2.7	1.7	2.2	2.1	3.2	3.3	3.4	3.4
4	3.3	2.8	3.2	3.1	5.8	5.0	5.6	5.3
5	4.9	4.6	4.9	4.8	9.1	7.7	8.8	8.1
6	7.7	7.4	7.7	7.6	14.0	12.3	13.7	12.7
7	12.3	12.0	12.3	12.2	21.7	19.7	21.4	20.1
8	19.7	19.3	19.7	19.6	34.0	32.0	33.7	32.1
9	⋮	⋮	⋮	⋮	54.0	51.0	53.3	51.4
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

TABLE 3. Example 6.

p	k_1	k_2	k_3	$\|\|k\|\|$	$z_1^\star$	$z_2^\star$	$z_3^\star$	$\|\|z^\star\|\|$
0					0.7	1.0	1.4	1.2
1	0.3	0.2	0.3	0.2	1.2	1.5	2.0	1.7
2	1.1	1.1	1.2	1.1	1.7	1.9	2.4	2.2
3	1.2	1.2	1.5	1.3	2.5	2.7	3.4	3.0
4	2.1	2.0	2.2	2.1	3.7	3.8	4.8	4.2
5	3.4	3.2	3.5	3.3	5.7	5.8	7.0	6.2
6	5.5	5.3	5.6	5.4	9.1	9.1	10.5	9.5
7	8.9	8.6	9.0	8.8	14.6	14.4	16.2	14.8
8	14.4	14.0	14.6	14.1	23.5	23.0	25.2	23.4
9	23.3	22.6	23.5	22.8	38.0	36.9	39.8	37.3
10	∞	∞	∞	∞	61.2	59.5	63.3	69.9
11					∞	∞	∞	∞

TABLE 4. Example 7.

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 4 shows the performance of the one-point version of our algorithm. The initial iterates were $z^{(0)} := (1.2, 2.2, 5.2)$ and $d^{(0)} := (1, 1, 1)$.

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 [18] analysed Van de Vel's method [35], improved it, 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. This is the subject of ongoing research. ◇

Acknowledgements

The authors thank Prof. A. Griewank for pointing out a flaw in the proof of a previous version of Theorem 5. They also 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. I. A. Aĭzenberg and A. P. Yuzhakov, *Integral representations and residues in multidimensional complex analysis*, Translations of Mathematical Monographs, vol. 58, American Mathematical Society, Providence, Rhode Island, 1983.
2. D. W. Decker, H. B. Keller, and C. T. Kelley, *Convergence rates for Newton's method at singular points*, SIAM J. Numer. Anal. **20** (1983), no. 2, 296–314.
3. D. W. Decker and C. T. Kelley, *Newton's method at singular points: I*, SIAM J. Numer. Anal. **17** (1980), no. 1, 66–70.
4. ———, *Newton's method at singular points: II*, SIAM J. Numer. Anal. **17** (1980), no. 3, 465–471.
5. ———, *Convergence acceleration for Newton's method at singular points*, SIAM J. Numer. Anal. **19** (1981), no. 1, 219–229.
6. ———, *Broyden's method for a class of problems having singular Jacobian at the root*, SIAM J. Numer. Anal. **22** (1985), no. 3, 566–574.
7. ———, *Expanded convergence domains for Newton's method at nearly singular roots*, SIAM J. Sci. Stat. Comput. **6** (1985), no. 4, 951–966.
8. A. Griewank, *Starlike domains of convergence for Newton's method at singularities*, Numer. Math. **35** (1980), 95–111.
9. ———, *On solving nonlinear equations with simple singularities or nearly singular solutions*, SIAM Review **27** (1985), no. 4, 537–563.
10. A. Griewank and M. R. Osborne, *Newton's method for singular problems when the dimension of the null space is* > 1, SIAM J. Numer. Anal. **18** (1981), no. 1, 145–149.
11. ———, *Analysis of Newton's method at irregular singularities*, SIAM J. Numer. Anal. **20** (1983), no. 4, 747–773.
12. A. Hoy, *A relation between Newton and Gauss-Newton steps for singular nonlinear equations*, Computing **40** (1988), 19–27.
13. ———, *An efficiently implementable Gauss-Newton-like method for solving singular nonlinear equations*, Computing **41** (1989), 107–122.
14. A. Hoy and H. Schwetlick, *Some superlinearly convergent methods for solving singular nonlinear equations*, Computational Solution of Nonlinear Systems of Equations (Providence, Rhode Island) (E. L. Allgower and K. Georg, eds.), Lectures in Applied Mathematics, vol. 26, 1988 SIAM-AMS Summer Seminar on Computational Solution of Nonlinear Systems of Equations, July 18-29, 1988, Fort Collins, Colorado, American Mathematical Society, 1990, pp. 285–300.
15. H. B. Keller, *Geometrically isolated nonisolated solutions and their approximation*, SIAM J. Numer. Anal. **18** (1981), no. 5, 822–838.
16. C. T. Kelley, *A Shamanskii-like acceleration scheme for nonlinear equations at singular roots*, Math. Comput. **47** (1986), no. 176, 609–623.
17. C. T. Kelley and R. Suresh, *A new acceleration method for Newton's method at singular points*, SIAM J. Numer. Anal. **20** (1983), no. 5, 1001–1009.
18. R. F. King, *Improving the Van de Vel root-finding method*, Computing **30** (1983), 373–378.
19. P. Kunkel, *Efficient computation of singular points*, IMA J. Numer. Anal. **9** (1989), 421–433.
20. Z. Mei, *A special extended system and a Newton-like method for simple singular nonlinear equations*, Computing **45** (1990), 157–167.
21. R. Menzel and G. Pönisch, *A quadratically convergent method for computing simple singular roots and its application to determining simple bifurcation points*, Computing **32** (1984), 127–138.
22. A. P. Morgan, A. J. Sommese, and C. W. Wampler, *Computing singular solutions to nonlinear analytic systems*, Numer. Math. **58** (1991), 669–684.
23. ———, *Computing singular solutions to polynomial systems*, Adv. Appl. Math. **13** (1992), no. 3, 305–327.
24. ———, *A power series method for computing singular solutions to nonlinear analytic systems*, Numer. Math. **63** (1992), 391–409.
25. M. Z. Nashed and X. Chen, *Convergence of Newton-like methods for singular operator equations using outer inverses*, Numer. Math. **66** (1993), 235–257.
26. B. Neta and H. D. Victory, *A higher order method for determining nonisolated solutions of a system of nonlinear equations*, Computing **32** (1984), 163–166.
27. T. Ojika, *Modified deflation algorithm for the solution of singular problems. I. A system of nonlinear algebraic equations*, J. Math. Anal. Appl. **123** (1987), 199–221.

28. ______, *Modified deflation algorithm for the solution of singular problems. II. Nonlinear multipoint boundary value problems*, J. Math. Anal. Appl. **123** (1987), 222–237.
29. T. Ojika, S. Watanabe, and T. Mitsui, *Deflation algorithm for the multiple roots of a system of nonlinear equations*, J. Math. Anal. Appl. **96** (1983), 463–479.
30. L. B. Rall, *Convergence of the Newton process to multiple solutions*, Numer. Math. **9** (1966), 23–37.
31. G. W. Reddien, *On Newton's method for singular problems*, SIAM J. Numer. Anal. **15** (1978), no. 5, 993–996.
32. ______, *Newton's method and high order singularities*, Computers Math. Applic. **5** (1979), 79–86.
33. J. F. Traub, *Iterative methods for the solution of equations*, Prentice-Hall, 1964.
34. T. Tsuchiya, *Enlargement procedure for resolution of singularities at simple singular solutions of nonlinear equations*, Numer. Math. **52** (1988), 401–411.
35. H. Van de Vel, *A method for computing a root of a single nonlinear equation, including its multiplicity*, Computing **14** (1975), 167–171.
36. M. Vander Straeten and H. Van de Vel, *Multiple root-finding methods*, J. Comput. Appl. Math. **40** (1992), no. 1, 105–114.
37. H. Weber and W. Werner, *On the accurate determination of nonisolated solutions of nonlinear equations*, Computing **26** (1981), 315–326.
38. T. J. Ypma, *Finding a multiple zero by transformations and Newton-like methods*, SIAM Review **25** (1983), no. 3, 365–378.

Katholieke Universiteit Leuven, Department of Computer Science, Celestijnenlaan 200 A, B-3001 Heverlee, Belgium

E-mail address: Peter.Kravanja@na-net.ornl.gov

Katholieke Universiteit Leuven, Department of Computer Science, Celestijnenlaan 200 A, B-3001 Heverlee, Belgium

E-mail address: Ann.Haegemans@cs.kuleuven.ac.be