Journal of Complexity 25 (2009) 38–62

Contents lists available at ScienceDirect

Journal of Complexity

journal homepage: www.elsevier.com/locate/jco

General local convergence theory for a class of iterative processes and its applications to Newton's method

Petko D. Proinov

Faculty of Mathematics and Informatics, University of Plovdiv, Plovdiv 4000, Bulgaria

ARTICLE INFO

Article history:
Received 19 September 2007
Accepted 13 May 2008
Available online 7 June 2008

Keywords:
Iterative methods
Gauge functions
Fixed points
Local convergence
Metric spaces
Newton iteration
Banach spaces
Schröder iteration
Polynomial zeros
Zeros of analytic functions

ABSTRACT

General local convergence theorems with order of convergence $r \geq 1$ are provided for iterative processes of the type $x_{n+1} = Tx_n$, where $T\colon D \subset X \to X$ is an iteration function in a metric space X. The new local convergence theory is applied to Newton iteration for simple zeros of nonlinear operators in Banach spaces as well as to Schröder iteration for multiple zeros of polynomials and analytic functions. The theory is also applied to establish a general theorem for the uniqueness ball of nonlinear equations in Banach spaces. The new results extend and improve some results of [K. Dočev, Über Newtonsche Iterationen, C. R. Acad. Bulg. Sci. 36 (1962) 695–701; J.F. Traub, H. Woźniakowski, Convergence and complexity of Newton iteration for operator equations, J. Assoc. Comput. Mach. 26 (1979) 250–258; S. Smale, Newton's method estimates from data at one point, in: R.E. Ewing, K.E. Gross, C.F. Martin (Eds.), The Merging of Disciplines: New Direction in Pure, Applied, and Computational Mathematics, Springer, New York, 1986, pp. 185–196; P. Tilli, Convergence conditions of some methods for the simultaneous computation of polynomial zeros, Calcolo 35 (1998) 3–15; X.H. Wang, Convergence of Newton's method and uniqueness of the solution of equations in Banach space, IMA J. Numer. Anal. 20 (2000) 123–134; I.K. Argyros, J.M. Gutiérrez, A unified approach for enlarging the radius of convergence for Newton's method and applications, Nonlinear Funct. Anal. Appl. 10 (2005) 555–563; M. Giusti, G. Lecerf, B. Salvy, J.-C. Yakoubsohn, Location and approximation of clusters of zeros of analytic functions, Found. Comput. Math. 5 (3) (2005) 257–311], and others.

E-mail address: proinov@uni-plovdiv.bg.

0885-064X/$ – see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.jco.2008.05.006

Contents

1. Introduction

In this paper we establish some general local convergence theorems with order of convergence $r \geq 1$ for iterative processes of the type

$$x_{n+1} = Tx_n, \quad n = 0, 1, 2, \ldots, \tag{1.1}$$

where $T: D \subset X \to X$ is an iteration function in a metric space X satisfying the following condition

$$E(Tx) \leq \varphi(E(x)) \quad \text{for all } x \in D \text{ with } Tx \in D \text{ and } E(x) \in J, \tag{1.2}$$

where $E: D \to \mathbb{R}_+$, J is an interval on $\mathbb{R}_+$ containing 0, φ is a gauge function on J, i.e. $\varphi: J \to J$. We prove that if T satisfies (1.2) and some other assumptions, then the initial convergence conditions of Picard iteration (1.1) can be given in the form $E(x_0) \in J$. This is way we call a function $E: D \to \mathbb{R}_+$ satisfying (1.2) *function of initial conditions* of T. Using this notion we establish a new local convergence theory for iterative processes of the type (1.1). Applying this theory one can get local convergence theorems for many iterations. In this work we apply our theory to Newton iteration as well as to Schröder iteration (Newton iteration for multiple zeros) with respect to various functions of initial conditions.

The paper is structured as follows. The general local convergence theory for the iterative processes of the type (1.1) is presented in Sections 2 and 3. The main results here are formulated in Theorems 3.6 and 3.8. In Section 4 we apply Theorem 3.6 to Newton iteration for multiple zeros of a complex polynomial $f(z)$ with respect to the function of initial conditions E defined as follows $E(z) = |z - \xi|/d$, where ξ is a zero of f (simple or multiple) and d denotes the distance from ξ to the other zeros of f. The results in this section extend the corresponding results of Dočev [5]. In Section 5 we continue to study the local convergence of Newton iteration for multiple polynomial zeros but with respect to the function of initial conditions defined by $E(z) = |z - \xi|/\rho(z)$, where $\rho(z)$ denotes the distance from z to the nearest zero of f which is not equal to ξ. The results of this section improve and extend a result of Tilli [18]. In Section 6 we apply our theory to Newton iteration for multiple zeros of a complex function $f(z)$ which is analytic in a neighborhood of ξ, where ξ is a zero of f with multiplicity $m \in \mathbb{N}$. The function of initial conditions here is defined as follows $E(z) = \gamma(\xi)\|z - \xi\|$, where $\gamma(\xi) = \sup_{k>m} \left|\frac{m! f^{(k)}(\xi)}{k! f^{(m)}(\xi)}\right|^{1/(k-m)}$. This quantity $\gamma(\xi)$ has been introduced in the case $m = 1$ by Smale [17] and in the case $m \geq 1$ by Yakoubsohn [24]. The first result for the convergence ball of Newton iteration with respect to this function of initial conditions is due to Traub and Woźniakowski [19] for simple zeros of analytic functions (even in Banach spaces). Later, Smale [17] (γ-theorem) improves Traub and Woźniakowski's result. In 2005, Giusti, Lecerf, Salvy and Yakoubsohn [6, Proposition 3.4] generalize γ-theorem to cluster of zeros. In this section we improve Proposition 3.4 of [6] in the case of multiple zeros of analytic functions. In Section 7 we investigate the local convergence of Newton iteration $x_{n+1} = x_n - F'(x_n)^{-1}F(x_n)$ for a simple zero ξ of a nonlinear Fréchet differentiable operator F defined on a subset D of a Banach space X with values in a Banach space Y. Here we study Newton iteration with respect to the standard function of initial conditions $E(x) = \|x - \xi\|$. The main result in this section gives a unified theory for the convergence ball of Newton's method and extends the corresponding results of Traub and Woźniakowski [19], Smale [17], Wang [21], Wang and Li [23], Argyros and Gutiérrez [2] and others. Finally, in Section 8 the new theory is applied to establish a general theorem for the uniqueness ball of nonlinear equations in Banach spaces around a simple zero. The main result in this section extends a resent result of Wang [21]. Section 7 and Section 8 can also be considered as a survey for convergence ball of Newton's method in Banach spaces.

Throughout the paper $S_n(t)$ denotes the polynomial $S_n(t) = 1 + t + \cdots + t^{n-1}$, where $n \in \mathbb{N}$; if $n = 0$ we assume that $S_n(t) \equiv 0$.

2. Functions of initial conditions and initial points

The main purpose of this section is to introduce the notion *function of initial conditions of an iteration function* which play a central role in the convergence theorems given in the next section. Throughout the paper J denotes an interval on $\mathbb{R}_+$ containing 0, that is an interval of the form $[0, R]$, $[0, R)$ or $[0, \infty)$. We begin this section with the following definition of gauge function of order $r \geq 1$.

Definition 2.1 (*[11]*). A function $\varphi: J \to J$ is said to be a *gauge function of order* $r \geq 1$ on J if it satisfies the following conditions:

(i) $\varphi(\lambda t) \leq \lambda^r \varphi(t)$ for all $\lambda \in (0, 1)$ and $t \in J$;
(ii) $\varphi(t) \leq t$ for all $t \in J$.

A gauge function of order r on J is said to be *strict gauge function* if $\varphi(t) < t$ for all $t \in J \setminus \{0\}$

Let us consider some simple but useful properties of nonnegative functions on J satisfying the first condition of Definition 2.1:

- A nonnegative function φ on J satisfies (i) if and only if $\varphi(0) = 0$ and the function $\varphi(t)/t^r$ is nondecreasing on $J \setminus \{0\}$.
- If two nonnegative function φ and ψ on J satisfy (i), then the function $\varphi + \psi$ also satisfies (i).
- If a nonnegative function φ on J satisfies (i) and if ψ is nonnegative nondecreasing functions on J, then the function $\varphi\psi$ satisfies (i).

Note that if $\varphi \not\equiv 0$ is a gauge function of order $r > 1$ on J, then J is a bounded interval. In this case the second condition of Definition 2.1 can be dropped. Moreover, every gauge function φ of order $r > 1$ on an interval $J = [0, R)$ is strict gauge function on J. Let us give two examples for gauge functions of order $r \geq 1$ which play an important role in Sections 7 and 8.

Example 2.2. Let ω be a nonnegative nondecreasing function on J such that $\omega(t)/t^p$ is nondecreasing on $J \setminus \{0\}$ for some $p \geq 0$. Then the function φ defined by $\varphi(t) = \int_0^t \omega(u)\mathrm{d}u$ is a gauge function of order $p + 1$ on J provided that $\varphi(t) \leq t$ for all $t \in J$.

Example 2.3. Let ω be a nonnegative nondecreasing function on J such that for $\lambda \in (0, 1)$ and $t, u \in J$ with $t \geq u$ it satisfies $\omega(\lambda t) - \omega(\lambda u) \leq \lambda^p[\omega(t) - \omega(u)]$ for some $p \geq 0$. Then the function φ defined by $\varphi(t) = t\omega(t) - \int_0^t \omega(u)\mathrm{d}u$ is a gauge function of order $p + 1$ on J provided that $\varphi(t) \leq t$ for all $t \in J$.

In the following lemma we consider a simple but useful property of gauge functions of order r.

Lemma 2.4 (*[11]*). *For every gauge function φ of order r on J there exists a nondecreasing nonnegative function ϕ on J such that*

$$\varphi(t) = t\phi(t) \quad \text{for all } t \in J. \tag{2.1}$$

The function ϕ has the following properties:

$$0 \leq \phi(t) \leq 1 \quad \text{for all } t \in J; \tag{2.2}$$

$$\phi(\lambda t) \leq \lambda^{r-1}\phi(t) \quad \text{for all } \lambda \in (0, 1) \text{ and all } t \in J. \tag{2.3}$$

If φ is a strict gauge function, then (2.2) can be replaced by

$$0 \leq \phi(t) < 1 \quad \text{for all } t \in J. \tag{2.4}$$

Remark 2.5. One can always defined a function ϕ on J satisfying (2.1) by

$$\phi(t) = \begin{cases} \varphi(t)/t & \text{if } t \in J \setminus \{0\}, \\ 0 & \text{if } t = 0. \end{cases} \tag{2.5}$$

Note that in the case $r > 1$ this is the unique nondecreasing nonnegative function on J which satisfies (2.1). Throughout the paper we assume for convenience that $0^0 = 1$. Then Lemma 2.4 remains true if we replace $\lambda \in (0, 1)$ in (2.3) by $\lambda \in [0, 1]$.

Definition 2.6 (*Function of Initial Conditions*). Let $T: D \subset X \to X$ be a map on an arbitrary set X. A function $E: D \to \mathbb{R}_+$ is said to be *function of initial conditions* of T (with a gauge function φ on J) if there exists a function $\varphi: J \to J$ such that

$$E(Tx) \le \varphi(E(x)) \quad \text{for all } x \in D \text{ with } Tx \in D \text{ and } E(x) \in J. \tag{2.6}$$

Definition 2.7 (*Initial Points*). Let $T: D \subset X \to X$ be a map on an arbitrary set X and let $E: D \to \mathbb{R}_+$ be a function of initial conditions of T (with a gauge function on J). Then a point $x \in D$ is said to be *initial point* of T if $E(x) \in J$ and all of the iterates $T^n x$ $(n = 0, 1, \ldots)$ are well-defined and belong to D.

It is easy to prove that if x_0 is an initial point of T, then each iterate x_n of Picard iteration (1.1) is also an initial point of T. Note that Definition 2.7 was given in [11] in the case when $E(x) = d(x, Tx)$. The following lemma was proved in [11] in the case $E(x) = d(x, Tx)$. The proof of the general case is the same.

Lemma 2.8. *Let $T: D \subset X \to X$ be a map on an arbitrary set X and let $E: D \to \mathbb{R}_+$ be a function of initial conditions of T with a gauge function φ of order r on an interval J. For every initial point $x_0 \in D$ of T and every $n \ge 0$ we have*

$$E(x_{n+1}) \le \varphi(E(x_n)) \quad \text{and} \quad E(x_n) \le E(x_0)\lambda^{S_n(r)},$$

where x_n denotes the n-th iterate of the iterative process (1.1), $\lambda = \phi(E(x_0))$ and ϕ is a nondecreasing nonnegative function on J satisfying (2.1).

3. General local convergence theorems for iterative processes

In this section we present three general convergence theorems for local convergence of iterative processes of the type

$$x_{n+1} = Tx_n, \quad n = 0, 1, 2, \ldots, \tag{3.1}$$

where $T: D \subset X \to X$ is an iteration function in a metric space (X, d). In these theorems we assume that the operator T has a function of initial conditions E with a gauge function φ on an interval J. Throughout the paper for a convergent sequence (x_n) in X we use the notion of Q-order of convergence as well as the notion of R-order of convergence (see, e.g. [9]). Let us recall these two notions.

Definition 3.1. A sequence (x_n) converges to ξ with Q-order (at least) $r \ge 1$ if there exists $c \ge 0$ such that $d(x_{n+1}, \xi) \le c(d(x_n, \xi))^r$ for sufficiently large n; in the case $r = 1$ we assume in addition that $c < 1$. In the cases $r = 1$ and $r = 2$ one say that (x_n) converges Q-linearly and Q-quadratically respectively.

Definition 3.2. A sequence (x_n) converges to ξ with R-order (at least) $r \ge 1$ if there exists a sequence of real numbers (α_n) converging to zero with Q-order (at least) r such that $d(x_n, \xi) \le \alpha_n$.

Now we shall establish some general convergence theorems for the iterative process (3.1) with order of convergence $r \ge 1$. We begin with the case when the gauge function $\varphi: J \to J$ satisfies only the following weak condition:

$$\varphi(t) \le t \quad \text{for all } t \in J. \tag{3.2}$$

Theorem 3.3. *Let $T: D \subset X \to X$ be an operator on a metric space (X, d). Suppose that $E: D \to \mathbb{R}_+$ is a function of initial conditions of T with a gauge function $\varphi: J \to J$ satisfying (3.2). Let ξ be a point in D such that $E(\xi) \in J$. Assume that*

$$d(Tx, \xi) \le \beta(E(x))d(x, \xi) \quad \text{for all } x \in D \text{ with } E(x) \in J, \tag{3.3}$$

where β is a nondecreasing function on J satisfying

$$0 \le \beta(t) < 1 \quad \text{for all } t \in J. \tag{3.4}$$

Then ξ is a unique fixed point of T in the set $U = \{x \in D : E(x) \in J\}$. Moreover, for each initial point x_0 of T Picard iteration (3.1) remains in U and converges Q-linearly to ξ with error estimates

$$d(x_{n+1}, \xi) \le hd(x_n, \xi) \quad \text{and} \quad d(x_n, \xi) \le h^n d(x_0, \xi), \tag{3.5}$$

where $h = \beta(E(x_0))$.

Proof. Applying (3.3) with $x = \xi$ we conclude that ξ is a fixed point of T. Suppose $\eta \in U$ is also a fixed point of T. Applying (3.3) with $x = \eta$ we obtain $d(\eta, \xi) \le \beta(E(\eta))d(\eta, \xi)$ which according to (3.4) means that $d(\eta, \xi) = 0$, i.e. $\eta = \xi$. From (3.3) and the fact that each iterate x_n is an initial point of T, we obtain

$$d(x_{n+1}, \xi) \le \beta(E(x_n))d(x_n, \xi). \tag{3.6}$$

By (2.6) and (3.2) we get $E(x_n) \le E(x_0)$ and so $\beta(E(x_n)) \le \beta(E(x_0))$. Now it follows from (3.6) that

$$d(x_{n+1}, \xi) \le hd(x_n, \xi), \tag{3.7}$$

which implies (3.5). To complete the proof it is sufficient to note that (3.4) implies $0 \le h < 1$. □

Remark 3.4. Let ξ is a point in D. It is easy to show that under the assumptions of Theorem 3.3 the condition $E(\xi) \in J$ is satisfied if and only if $E(\xi)$ is a fixed point of the gauge function φ. In particular, one can choose in Theorem 3.3 ξ to be a zero of the function E since 0 is a fixed point of φ.

Remark 3.5. It is easy to prove that Theorem 3.3 remains true but without the estimate (3.5) if one replace the assumption "β is a nondecreasing function on J" by "β is a right continuous function on J".

Theorem 3.6 (*Main Convergence Theorem*). *Let $T: D \subset X \to X$ be an operator on a metric space (X, d). Suppose that $E: D \to \mathbb{R}_+$ is a function of initial conditions of T with a gauge function φ of order $r \ge 1$ on J. Let ξ be a point in D satisfying $E(\xi) \in J$. Assume that (3.3) holds with a nondecreasing nonnegative function β on J such that*

$$t\beta(t) \text{ is strict gauge function of order } r \text{ on } J \tag{3.8}$$

and

$$\text{for } t \in J : \phi(t) = 0 \text{ implies } \beta(t) = 0, \tag{3.9}$$

where ϕ is a nondecreasing nonnegative function on J satisfying (2.1). Then ξ is a unique fixed point of T in the set $U = \{x \in D : E(x) \in J\}$. Moreover, for each initial point x_0 of T the following statements hold.

(i) *The iterative sequence (3.1) remains in the set U and converges to ξ. If $\phi(E(x_0)) < 1$, then the sequence (3.1) converges with R-order r.*
(ii) *For all $n \ge 0$ we have the estimate*

$$d(x_n, \xi) \le \theta^n \lambda^{S_n(r)} d(x_0, \xi), \tag{3.10}$$

where $\lambda = \phi(E(x_0))$, $\theta = \psi(E(x_0))$ and ψ is a nonnegative function on J such that $\beta(t) = \phi(t)\psi(t)$ for all $t \in J$.

(iii) *If there exists a function* $C: D \to \mathbb{R}_+$ *continuous at* ξ *and such that* $E(x) \le C(x)d(x, \xi)$ *for all* $x \in D$, *then the iterative sequence* (3.1) *converges to* ξ *with* Q*-order* r *as follows*

$$d(x_{n+1}, \xi) \le C_n(d(x_n, \xi))^r \quad \text{for all } n \ge 0, \tag{3.11}$$

where $C_n = C(x_n)^{r-1}\mu(E(x_n)) \to C(\xi)^{r-1}\mu(0)$ *as* $n \to \infty$, *and the real function* $\mu: J \to \mathbb{R}_+$ *is continuous at* 0 *and* $\beta(t) = t^{r-1}\mu(t)$ *for all* $t \in J$.

Proof. Let us note first that condition (3.8) guarantees the existence of a nonnegative function ψ on J such that $\beta = \phi\psi$. Indeed, choose an arbitrary nonnegative number b and define $\psi: J \to \mathbb{R}_+$ as follows:

$$\psi(t) = \begin{cases} \beta(t)/\phi(t) & \text{if } \phi(t) > 0, \\ b & \text{if } \phi(t) = 0. \end{cases}$$

Taking into account (3.9) we get that $\beta(t) = \phi(t)\psi(t)$ for all $t \in J$. It is easy to see that condition (3.8) guarantees the existence of the function μ as well. (i) The Q-linear convergence of the iterative sequence (3.1) follows immediately from Theorem 3.3. From Lemma 2.4 and the definitions of λ and θ we conclude that

$$0 \le \lambda \le 1, \quad \theta \ge 0 \quad \text{and} \quad 0 \le \theta\lambda < 1.$$

Therefore, if $\lambda = \phi(E(z_0)) < 1$, then (3.10) shows that Picard sequence (3.1) converges with R-order r. (ii) According to Lemma 2.4 and (3.8) the function β satisfies (3.4) and

$$\beta(ut) \le u^{r-1}\beta(t) \quad \text{for all } u \in [0, 1] \text{ and } t \in J. \tag{3.12}$$

Now we shall prove (3.10) by induction on n. If $n = 0$, then (3.10) reduces to an equality. Assuming (3.10) to hold for an integer $n \ge 0$, we shall prove it for $n + 1$. Using Lemma 2.8 and (3.12) we deduce

$$\beta(E(x_n)) \le \beta(E(x_0)\lambda^{S_n(r)}) \le \lambda^{(r-1)S_n(r)}\beta(E(x_0)) = \theta\lambda^{1+(r-1)S_n(r)} = \theta\lambda^{r^n}.$$

From (3.6), the last inequality and induction hypotheses, we get

$$d(x_{n+1}, \xi) \le \theta^{n+1}\lambda^{S_{n+1}(r)}d(x_0, \xi)$$

which proves (3.10). (iii) It follows from (3.3) that for every $x \in D$ such that $E(x) \in J$ we have

$$\begin{aligned} d(Tx, \xi) &\le \beta(E(x))d(x, \xi) = E(x)^{r-1}\mu(E(x))d(x, \xi) \\ &\le C(x)^{r-1}\mu(E(x))(d(x, \xi))^r. \end{aligned}$$

From this inequality we obtain (3.11) since x_n is an initial point of T. Now taking into account that C is continuous at ξ and μ is continuous at 0, we conclude that $C_n \to C(\xi)^{r-1}\mu(0)$ as $n \to \infty$ □

If the function of initial conditions is defined by $E(x) = d(x, \xi)$, where $\xi \in D$, we get the following simple but useful special case of Theorem 3.6.

Corollary 3.7. *Let* $T: D \subset X \to X$ *be an operator on a metric space* (X, d) *and let* $\xi \in D$. *Assume that*

$$d(Tx, \xi) \le \varphi(d(x, \xi)) \quad \text{for all } x \in D \text{ with } d(x, \xi) \in J, \tag{3.13}$$

where φ *is a strict gauge function of order* $r \ge 1$ *on* J. *Then* ξ *is a unique fixed point of* T *in the set* $U = \{x \in D : d(x, \xi) \in J\}$. *Moreover, if* $T: U \to U$, *then for each point* $x_0 \in U$ *the following statements hold.*

(i) *Picard sequence* (3.1) *remains in* U *and converges to* ξ *with* Q*-order* r.
(ii) *For all* $n \ge 0$ *we have the estimates*

$$d(x_n, \xi) \le \lambda^{S_n(r)}d(x_0, \xi), \tag{3.14}$$

where $\lambda = \phi(E(x_0))$ *and* ϕ *is a nondecreasing nonnegative function on* J *satisfying* (2.1)

(iii) *For all $n \geq 0$ we have the estimate*

$$d(x_{n+1}, \xi) \leq \varphi(d(x_n, \xi)). \tag{3.15}$$

The following general convergence theorem gives better error bounds than Theorem 3.6 but a smaller convergence domain.

Theorem 3.8. *Let $T: D \subset X \to X$ be an operator on a metric space (X, d). Suppose that $E: D \to \mathbb{R}_+$ is a function of initial conditions of T with a gauge function φ of order $r \geq 1$ on J. Let ξ be a point in D satisfying $E(\xi) \in J$. Assume that (3.3) hold with a nonnegative and nondecreasing function β on J satisfying (3.8) and (3.9) as well as*

$$\beta(\varphi(t)) \leq \beta(t)^r \quad \text{for all } t \in J. \tag{3.16}$$

Then ξ is a unique fixed point of T in the set $U = \{x \in D : E(x) \in J\}$. Moreover, for each initial point x_0 of T the following statements hold.

(i) *Picard sequence (3.1) remains in U and converges to ξ with R-order r. It converges with Q-order r provided that $E(x) \leq C(x)d(x, \xi)$ for all $x \in D$, where $C: D \to \mathbb{R}_+$ is continuous at ξ.*

(ii) *For all $n \geq 0$ we have the estimate*

$$d(x_n, \xi) \leq h^{S_n(r)} d(x_0, \xi), \tag{3.17}$$

where $h = \beta(E(x_0))$, ϕ is a nondecreasing nonnegative on J satisfying (2.1).

Proof. According to Theorem 3.6 we have to prove only (ii). First we will show by induction that

$$\beta(E(x_n)) \leq h^{r^n} \tag{3.18}$$

for all $n \geq 0$. It is trivial in the case $n = 0$. From Lemma 2.8, (3.16) and induction hypotheses we get

$$\beta(E(x_{n+1})) \leq \beta(\varphi(E(x_n))) \leq [\beta(E(x_n))]^r \leq \left(h^{r^n}\right)^r = h^{r^{n+1}}$$

which proves (3.18). Now we shall prove (3.17) again by induction. In the case $n = 0$ it is an equality. From (3.6) and (3.18) and induction hypotheses, we obtain

$$d(x_{n+1}, \xi) \leq h^{r^n} d(x_n, \xi) \leq h^{r^n + S_n(r)} d(x_0, \xi) = h^{r^{n+1}} d(x_0, \xi)$$

which completes the proof. □

Obviously, in the case $r = 1$ the assumption (3.16) holds always. Therefore, in this case both Theorems 3.6 and 3.8 coincide.

4. Newton's method for multiple polynomial zeros. I

Let $f: D \subset \mathbb{C} \to \mathbb{C}$ and ξ be a zero of f with multiplicity $m \in \mathbb{N}$. Let us define Newton's method for multiple zeros

$$z_{k+1} = Tz_k, \quad k = 0, 1, \ldots, \tag{4.1}$$

where $T: \mathbb{C} \to \mathbb{C}$ denotes Schröder's operator [16] defined as follows

$$Tz = \begin{cases} z - m\dfrac{f(z)}{f'(z)} & \text{if } f'(z) \neq 0, \\ z & \text{if } f'(z) = 0. \end{cases} \tag{4.2}$$

In this section we apply Theorem 3.6 to Newton's method for multiple polynomial zeros. Assume that f is a complex polynomial and ξ is a zero of f (simple or multiple). We study the convergence of Newton's method with respect to the following function of initial conditions:

$$E(z) = E(f, z) = \frac{|z - \xi|}{d}, \tag{4.3}$$

where d denotes the distance from ξ to the nearest zero of f which is not equal to ξ; if ξ is a unique zero of f we set $E(z) \equiv 0$.

In 1962, Dočev [5] established two convergence theorems for Newton iteration for a polynomial f with real and simple zeros. From the first Dočev's theorem one can get the the following result.

Theorem 4.1 (*Dočev [5]*). *Let f be a polynomial of degree $n \geq 2$ which has only real and simple zeros. Suppose ξ is a zero of f and $0 < \lambda < 1$. Assume $z_0 \in \mathbb{R}$ satisfies*

$$E(z_0) \leq \frac{\lambda}{(\lambda + 1)n - 1},$$

where the function $E\colon \mathbb{C} \to \mathbb{R}_+$ is defined by (4.3). Then Newton's method (4.1) converges to ξ with error estimate

$$|z_k - \xi| \leq \lambda^{2^k - 1} |z_0 - \xi| \quad \text{for all } k \geq 0. \tag{4.4}$$

The next result is the second Dočev's theorem.

Theorem 4.2 (*Dočev [5]*). *Let f be a polynomial of degree $n \geq 2$ which has only real and simple zeros and ξ be a zero of f which lie in an interval $[a - \varepsilon, a - \varepsilon]$ with $\varepsilon < d/(4n - 1)$, where d denotes the distance from ξ to the other zeros of f. Then for every $z_0 \in [a - \varepsilon, a - \varepsilon]$ Newton iteration (4.1) is convergent to ξ with error estimate*

$$|z_k - \xi| \leq 2\varepsilon \lambda^{2^k - 1} \quad \text{for all } k \geq 0,$$

where $\lambda = 2(n - m)\varepsilon/(md - (2n + 1)\varepsilon)$.

In this section we extend Theorems 4.1 and 4.2 for arbitrary complex polynomials.

Lemma 4.3. *Let f be a complex polynomial of degree $n \geq 2$ and let ξ be a zero of f with multiplicity m. Then for every $z \in \mathbb{C}$ satisfying*

$$E(z) < \frac{m}{n}, \tag{4.5}$$

where the function $E\colon \mathbb{C} \to \mathbb{R}_+$ is defined by (4.3), there exists a complex number σ such that

$$Tz - \xi = \frac{\sigma}{m + \sigma}(z - \xi) \quad \text{and} \quad |\sigma| \leq \frac{(n - m)E(z)}{1 - E(z)}, \tag{4.6}$$

where $T\colon \mathbb{C} \to \mathbb{C}$ is Schröder's operator defined by (4.2).

Proof. Let $z \in \mathbb{C}$ satisfy (4.5). If either $m = n$ or $z = \xi$, then $Tz = \xi$ and so the statement of the lemma holds with $\sigma = 0$. Suppose that $m < n$ and $z \neq \xi$. Let $\xi_1, \ldots, \xi_s$ be all the zeros of f with the multiplicities $m_1, \ldots, m_s$ respectively. Without loss of generality we can assume that $\xi = \xi_i$ and $m = m_i$ $(1 \leq i \leq s)$. By triangle inequality and (4.5) for every $j \neq i$ we get

$$|z - \xi_j| \geq |\xi - \xi_j| - |z - \xi| \geq d - |z - \xi| > 0. \tag{4.7}$$

It follows from (4.7) that z is not a zero of f. Now we shall prove that the complex number

$$\sigma = (z - \xi) \sum_{j=1, j \neq i}^{s} \frac{m_j}{z - \xi_j} \tag{4.8}$$

satisfies the statement of the lemma. From (4.7) we obtain

$$\frac{|z - \xi|}{|z - \xi_j|} \leq \frac{|z - \xi|}{d - |z - \xi|} = \frac{E(z)}{1 - E(z)}. \tag{4.9}$$

Combining (4.8) and (4.9) we obtain the inequality in (4.6). It remains to prove the equality in (4.6). From the well-known identity

$$\frac{f'(z)}{f(z)} = \sum_{j=1}^{s} \frac{m_j}{z - \xi_j}$$

we have

$$\frac{f'(z)}{f(z)} = \frac{m}{z-\xi} + \sum_{j=1, j\neq i}^{s} \frac{m_j}{z-\xi_j} = \frac{m+\sigma}{z-\xi}. \tag{4.10}$$

From the second part of (4.6) and (4.5)we conclude that

$$|m+\sigma| \geq m - |\sigma| \geq \frac{m - nE(z)}{1 - E(z)} > 0$$

which means that $m + \sigma \neq 0$. Then (4.10) implies that $f'(z) \neq 0$. Hence we have $Tz = z - f(z)/f'(z)$. Using (4.10) we obtain the first part of (4.6). □

From Lemma 4.3 we immediately get the following lemma.

Lemma 4.4. *Assume that the assumptions of Lemma 4.3 are satisfied. Then for every $z \in \mathbb{C}$ satisfying (4.5) we have*

$$E(Tz) \leq \varphi(E(z)) \quad \text{and} \quad |Tz - \xi| \leq \phi(E(z))|z - \xi|,$$

where the real functions φ and ϕ are defined on $[0, m/n)$ as follows $\varphi(t) = (n-m)t^2/(m-nt)$ and $\phi(t) = (n-m)t/(m-nt)$.

Theorem 4.5. *Let f be a complex polynomial of degree $n \geq 2$ and let ξ be a zero of f with multiplicity m. Suppose z_0 is a complex number satisfying*

$$E(z_0) < \frac{m}{2n-m},$$

where the function $E: \mathbb{C} \to \mathbb{R}_+$ is defined by (4.3). Then Newton's method (4.1) converges Q-quadratically to ξ with error estimate

$$|z_k - \xi| \leq \lambda^{2^k - 1}|z_0 - \xi| \quad \text{for all } k \geq 0.$$

where $\lambda = \phi(E(z_0))$ and $\phi(t) = (n-m)t/(m-nt)$,

Proof. It is easy to see that the function $\varphi(t) = t\phi(t)$ is a strict gauge function of the second order on the interval $J = [0, m/(2n-m))$. According to Lemma 4.4 the assumptions of Theorem 3.6 are fulfilled with $\beta(t) = \phi(t)$ and $\psi(t) \equiv 1$. □

Taking into account that the function ϕ is strictly increasing and continuous on the interval $J = [0, m/(2n-m)]$ and that $\phi(J) = [0, 1]$ we can reformulate Theorem 4.5 in the following equivalent form. Note that Theorem 4.6 is a extension of Theorem 4.1

Theorem 4.6. *Let f be a polynomial of degree $n \geq 2$, ξ a zero of f with multiplicity m and $0 < \lambda < 1$. Suppose $z_0 \in \mathbb{C}$ satisfies*

$$E(z_0) \leq \frac{m\lambda}{(\lambda+1)n - m},$$

where the function $E: \mathbb{C} \to \mathbb{R}_+$ is defined by (4.3). Then Newton's method (4.1) converges Q-quadratically to ξ with error estimate (4.4)

The following corollary of Theorem 4.6 improves and generalizes Theorem 4.2.

Corollary 4.7. *Let f be a complex polynomial of degree $n \geq 2$. Suppose ξ is a zero of f with multiplicity m which lie in a closed disc $\overline{U}(a, \varepsilon)$ with radius $\varepsilon < md/(4n - 2m)$, where d denotes the distance from ξ to the other zeros of f (if ξ is a unique zero of f we assume that $d = +\infty$). Then for every $z_0 \in \overline{U}(a, \varepsilon)$ Newton's method (4.1) converges Q-quadratically to ξ with error estimate*

$$|z_k - \xi| \leq 2\varepsilon\lambda^{2^k - 1} \quad \text{for all } k \geq 0,$$

where $\lambda = 2(n - m)\varepsilon/(md - 2n\varepsilon)$.

5. Newton's method for multiple polynomial zeros. II

In this section we again apply Theorem 3.6 as well as Theorem 3.8 to Newton's method (4.1) for multiple polynomial zeros but with respect to another function of initial conditions. Let f be a complex polynomial and ξ be a zero of f (simple or multiple). Define the following function of initial conditions

$$E(z) = E(f, z) = \frac{|z - \xi|}{\rho(z)}, \tag{5.1}$$

where $\rho(z)$ denotes the distance from z to the nearest zero of f which is not equal to ξ; if ξ is a unique zero of f we set $E(z) \equiv 0$. Suppose that f has at least two zeros. Then it is easy to prove that the function ρ satisfies $|\rho(x) - \rho(y)| \leq |x - y|$ for all $x, y \in \mathbb{C}$, i.e. ρ is Lipschitz continuous on $\mathbb{C}$.

In 1998, Tilli [18], improving a result of Renegar [14], proved that if f has only simple zeros and the starting point z_0 is such that $E(z_0) \leq 1/(3n - 3)$, then Newton iteration converges quadratically right from the first iteration (see Corollary 5.5). In this section we extend Tilli's result for arbitrary polynomials. The new results (Theorems 5.3 and 5.4) improve Tilli's one even in the case of polynomials with simple zeros.

Lemma 5.1. *Let f be a complex polynomial of degree $n \geq 2$ and let ξ be a zero of f with multiplicity m. Then for every $z \in \mathbb{C}$ satisfying*

$$E(z) < \frac{m}{n - m}, \tag{5.2}$$

where the function $E: \mathbb{C} \to \mathbb{R}_+$ is defined by (5.1), there exists a complex number σ such that

$$Tz - \xi = \frac{\sigma}{m + \sigma}(z - \xi) \quad \text{and} \quad |\sigma| \leq (n - m)E(z), \tag{5.3}$$

where $T: \mathbb{C} \to \mathbb{C}$ is Schröder's operator defined by (4.2).

Proof. The proof is the same as the proof of Lemma 4.3. One must only replace the inequalities (4.7) and (4.9) by $|z - \xi_j| > 0$ and $|(z - \xi)/(z - \xi_j)| \leq E(z)$ respectively. □

Lemma 5.2. *Assume that the assumptions of Lemma 5.1 are satisfied. Then for every $z \in \mathbb{C}$ satisfying*

$$E(z) < \frac{m}{n} \tag{5.4}$$

we have the following two inequalities:

$$E(Tz) \leq \varphi(E(z)) \quad \text{and} \quad |Tz - \xi| \leq \beta(E(z))|z - \xi|, \tag{5.5}$$

where the real functions φ and β are defined by

$$\varphi(t) = \frac{(n - m)t^2}{m - nt} \quad \text{and} \quad \beta(t) = \frac{(n - m)t}{m - (n - m)t}.$$

Proof. Let z be a complex number satisfying (5.4). The second inequality in (5.5) follows immediately from Lemma 5.1. Let η be any zero of f which is not equal to ξ. By triangle inequality, the second inequality in (5.5) and the definition of $E(z)$ we obtain

$$\begin{aligned}|Tz-\eta| &\geq |z-\eta|-|z-\xi|-|Tz-\xi| \geq \rho(z)-[1+\beta(E(z))]|z-\xi| \\ &= [1-E(z)(1+\beta(E(z)))]\rho(z) = \psi(E(z))\rho(z),\end{aligned}$$

where $\psi(t)=(m-nt)/[m-(n-m)t]$. It follows from the last inequality and (5.4) that

$$\rho(Tz) \geq \psi(E(z))\rho(z) > 0. \tag{5.6}$$

Dividing by $\rho(Tz)$ both sides of the second inequality in (5.5) and applying (5.6) we get the first inequality in (5.5) with

$$\varphi(t) = t\beta(t)/\psi(t) = (n-m)t^2/(m-nt)$$

which completes the proof. □

Theorem 5.3. *Let f be a complex polynomial of degree $n \geq 2$ and let ξ be a zero of f with multiplicity m. Suppose z_0 is a complex number satisfying*

$$E(z_0) \leq \frac{m}{2n-m},$$

where the function $E\colon \mathbb{C} \to \mathbb{R}_+$ is defined by (5.1). Then Newton's method (4.1) converges Q-quadratically to ξ. Moreover, we have the estimate

$$|z_k-\xi| \leq \theta^k \lambda^{2^k-1}|z_0-\xi| \quad \text{for all } k \geq 0,$$

where $\lambda=\phi(E(z_0))$, $\theta=\psi(E(z_0))$, and the real functions ϕ and ψ are defined by

$$\phi(t)=\frac{(n-m)t}{m-nt} \quad \text{and} \quad \psi(t)=\frac{m-nt}{m-(n-m)t}.$$

Proof. Let $\varphi(t)=t\phi(t)$ and $\beta(t)=\phi(t)\psi(t)$. It is easy to see that both $\varphi(t)$ and $t\beta(t)$ are gauge functions of the second order on the interval $I=[0, m/(2n-m)]$ and $t\beta(t)$ is a strict gauge function. According to Lemma 5.2 the assumptions of Theorem 3.6 are satisfied. □

The next convergence theorem for Newton's method gives better error estimates than Theorem 5.3 but a smaller convergence domain. Note that the new theorem can be applied only to polynomial zeros with multiplicity $m < n/2$.

Theorem 5.4. *Let f be a complex polynomial of degree $n \geq 3$ and let ξ be a zero of f with multiplicity $m < n/2$. Suppose z_0 is a complex number such that*

$$E(z_0) \leq \frac{m(n-2m)}{2(n-m)^2},$$

where the function $E\colon \mathbb{C} \to \mathbb{R}_+$ is defined by (5.1). Then Newton's method (4.1) converges Q-quadratically to ξ. Moreover, for all $k \geq 0$ we have

$$|z_k-\xi| \leq \lambda^{2^k-1}|z_0-\xi|$$

where $\lambda=\beta(E(z_0))$ and $\beta(t)=(n-m)t/(m-(n-m)t)$.

Proof. Define the functions φ and β on the interval $I=[0, m/(2n-m)]$ as in Lemma 5.2. As we have mentioned in the proof of the previous theorem the assumptions of Theorem 3.6 are fulfilled with φ and β defined on I. Now to satisfy the assumptions of Theorem 3.8 we have to find a subinterval J of I such that

$$\beta(\varphi(t)) \leq \beta(t)^2 \quad \text{for all } t \in J.$$

It is easy to verify that the last inequality is equivalent to the following one

$$[m-(n-m)t]^2 \le m(m-nt)-(n-m)^2t^2$$

which holds for $t \in J = [0, m(n-2m)/2(n-m)^2]$ provided that $1 \le m < n/2$. Applying Theorem 3.8 we complete the proof. □

The following result of Tilli [18] is an immediate consequence of Theorem 5.4.

Corollary 5.5 (*Tilli [18]*). *Let f be a complex polynomial of degree $n \ge 4$ with only simple zeros. Assume ξ is a zero of f and $0 < \mu \le 1/3$. Suppose z_0 is a complex number such that*

$$E(z_0) \le \frac{\mu}{n-1},$$

where the function $E: \mathbb{C} \to \mathbb{R}_+$ is defined by (5.1). Then Newton's method (4.1) converges to ξ. Moreover, for all $k \ge 0$ we have

$$|z_k-\xi| \le \sigma^{2^k-1}|z_0-\xi|,$$

where $\sigma = \mu/(1-\mu)$.

6. Newton's method for multiple zeros of analytic functions

In this section we apply Theorem 3.6 to Newton's method (4.1) for multiple zeros of analytic functions. Let $f: D \subset \mathbb{C} \to \mathbb{C}$ and let $\xi \in D$ be a zero of f with multiplicity $m \in \mathbb{N}$. If f is analytic in a neighbourhood of ξ we define

$$\gamma(\xi) = \gamma(f,\xi) = \sup_{k>m} \left| \frac{m! f^{(k)}(\xi)}{k! f^{(m)}(\xi)} \right|^{1/(k-m)}.$$

This quantity has been introduced in the case $m = 1$ by Smale [17] and in the case $m \ge 1$ by Yakoubsohn [24]. In this section we study the local convergence of Newton's method (4.1) for analytic functions with respect to the following function of initial conditions

$$E(z) = E(f,z) = \gamma(\xi)|z-\xi|. \tag{6.1}$$

In 1979, Traub and Woźniakowski [19] proved that if ξ is a simple zero of f and f is analytic in the ball $\{z \in \mathbb{C} : E(z) < 1\} \subset D$, then there exists $r \in (0,1)$ such that for every z_0 in the ball $U = \{z \in \mathbb{C} : E(z) < r\}$ Newton iteration (4.1) converges quadratically to ξ. Traub and Woźniakowski obtained this result with $r = 0.182\ldots$, where r is the unique real solution of the equation $(1-t)^3 = 3t$. In 1986, Smale [17] (see also [3, Chapter 8, Proposition 1]) obtained the optimal value of $r = (5-\sqrt{17})/4 = 0.219\ldots$ This result of Smale is known in the literature as γ-theorem. A short proof of the γ-theorem is given in [7, Theorem 1.16]. Note that the results in [19,17] are obtained even for analytic functions in Banach spaces (see Corollary 7.9). In 2005, Giusti, Lecerf, Salvy and Yakoubsohn [6, Proposition 3.4] generalize the γ-theorem to cluster of zeros of analytic functions. In this section we present a generalization of the γ-theorem that improves the result of Giusti et al. [6] in the case of multiple zeros of analytic functions.

Lemma 6.1. *Let $f: D \subset \mathbb{C} \to \mathbb{C}$. Assume that $\xi \in D$ is a zero of f with multiplicity m and f is analytic in an open ball $U = \{z \in \mathbb{C} : E(z) < r\} \subset D$, where the function $E: \mathbb{C} \to \mathbb{R}_+$ is defined by (6.1) and*

$$0 < r \le R = \frac{3m+1-\sqrt{m^2+6m+1}}{4m}. \tag{6.2}$$

Then for every $z \in U$ there exists a complex number σ such that

$$|Tz-\xi| \le \frac{1}{|m+\sigma|} \frac{E(z)}{(1-E(z))^2}|z-\xi| \tag{6.3}$$

and

$$|\sigma| \le \frac{(m+1)E(z) - mE(z)^2}{(1-E(z))^2}, \tag{6.4}$$

where $T: \mathbb{C} \to \mathbb{C}$ *is Schröder's operator defined by* (4.2)*.*

Proof. Let $z \in U$. If $z = \xi$, then $Tz = \xi$ and so the statement of the lemma holds with $\sigma = 0$. Let us consider the case $z \ne \xi$. Using Taylor's expansion of $f'(z)$ around ξ we get

$$\begin{aligned} f'(z) &= \sum_{k=m}^{\infty} \frac{f^{(k)}(\xi)}{(k-1)!}(z-\xi)^{k-1} \\ &= \frac{f^{(m)}(\xi)}{(m-1)!}(z-\xi)^{m-1} + \sum_{k=m+1}^{\infty} \frac{f^{(k)}(\xi)}{(k-1)!}(z-\xi)^{k-1} \\ &= \frac{f^{(m)}(\xi)}{m!}(z-\xi)^{m-1}(m+\sigma), \end{aligned} \tag{6.5}$$

where

$$\sigma = \sum_{k=m+1}^{\infty} k \frac{m! f^{(k)}(\xi)}{k! f^{(m)}(\xi)}(z-\xi)^{k-m}.$$

By triangle inequality we obtain the estimate

$$\begin{aligned} |\sigma| &\le \sum_{k=m+1}^{\infty} k \left| \frac{m! f^{(k)}(\xi)}{k! f^{(m)}(\xi)} \right| |z-\xi|^{k-m} \le \sum_{k=m+1}^{\infty} kE(z)^{k-m} \\ &= \sum_{k=1}^{\infty} (m+k)E(z)^k = \frac{mE(z)}{1-E(z)} + \frac{E(z)}{(1-E(z))^2} \\ &= \frac{(m+1)E(z) - mE(z)^2}{(1-E(z))^2}, \end{aligned}$$

which proves (6.4). From the last inequality taking account (6.2) we deduce that

$$|m+\sigma| \ge m - |\sigma| \ge \frac{2mE(z)^2 - (3m+1)E(z) + m}{(1-E(z))^2} > 0 \tag{6.6}$$

which means that $m + \sigma \ne 0$. Then (6.5) implies $f'(z) \ne 0$ and

$$|f'(z)| = \frac{|f^{(m)}(\xi)|}{m!}|z-\xi|^{m-1}|m+\sigma|, \tag{6.7}$$

Now we shall prove (6.3). Using Taylor's expansion on $f(z)$ and $f'(z)$ around ξ we get

$$\begin{aligned} Tz - \xi &= f'(z)^{-1}[f'(z)(z-\xi) - mf(z)] \\ &= \frac{1}{m!}\frac{f^{(m)}(\xi)}{f'(z)} \sum_{k=m}^{\infty}(k-m)\frac{m! f^{(k)}(\xi)}{k! f^{(m)}(\xi)}(z-\xi)^k. \end{aligned}$$

From this we obtain

$$\begin{aligned} |Tz-\xi| &\le \frac{1}{m!}\left|\frac{f^{(m)}(\xi)}{f'(z)}\right| \sum_{k=m}^{\infty}(k-m)\left|\frac{m! f^{(k)}(\xi)}{k! f^{(m)}(\xi)}\right| |z-\xi|^k \\ &\le \frac{1}{m!}\left|\frac{f^{(m)}(\xi)}{f'(z)}\right| |z-\xi|^m \sum_{k=m}^{\infty}(k-m)E(z)^{k-m} \end{aligned}$$

$$= \frac{1}{m!}\left|\frac{f^{(m)}(\xi)}{f'(z)}\right| |z-\xi|^m \sum_{k=1}^{\infty} kE(z)^k$$
$$= \frac{1}{m!}\left|\frac{f^{(m)}(\xi)}{f'(z)}\right| |z-\xi|^m \frac{E(z)}{(1-E(z))^2}.$$

Combining this bound with (6.7) we obtain (6.3). □

The next lemma is an immediate consequence of the previous one.

Lemma 6.2. *Assume that the assumptions of Lemma 6.1 are satisfied. Then for every $z \in U$ we have*

$$E(Tz) \le \varphi(E(z)) \quad \text{and} \quad |Tz-\xi| \le \phi(E(z))|z-\xi|, \tag{6.8}$$

where φ and ϕ are real functions defined by $\varphi(t) = t^2/(2mt^2 - (3m+1)t + m)$ and $\phi(t) = t/(2mt^2 - (3m+1)t + m)$

The following theorem generalizes the above mentioned results of Traub–Woźniakowski [19] and Smale [17] and improves Proposition 3.4 of Giusti et al. [6] in the case of multiple zeros.

Theorem 6.3. *Let $f: D \subset \mathbb{C} \to \mathbb{C}$. Assume that $\xi \in D$ is a zero of f with multiplicity m and f is analytic in an open ball $U = \{z \in \mathbb{C} : E(z) < r\} \subset D$, where the function $E: \mathbb{C} \to \mathbb{R}_+$ is defined by (6.1) and*

$$0 < r \le R = \frac{3m+2-\sqrt{m^2+12m+4}}{4m}. \tag{6.9}$$

Then for every $z_0 \in U$ Newton iteration (4.1) converges Q-quadratically to ξ as follows

$$|z_{k+1}-\xi| \le C_k |z_k-\xi|^2 \quad \text{for all } k \ge 0,$$

where $C_k = \gamma(\xi)\mu(E(x_k))$, $C_k \to \gamma(\xi)/m$ as $k \to \infty$, and the real function μ is defined by $\mu(t) = 1/(2mt^2 - (3m+1)t + m)$. Moreover, for all $k \ge 0$ we have

$$|z_k - \xi| \le \lambda^{2^k-1}|z_0-\xi|,$$

where $\lambda = \phi(E(z_0))$ and $\phi(t) = t/(2mt^2-(3m+1)t+m)$.

Proof. Obviously, the function $\varphi(t) = t\phi(t)$ is strictly increasing on $[0, R]$ and R is a fixed point of φ. Therefore, φ is a strict gauge functions of the second order on the interval $J = [0, R)$. According to Lemma 6.2 the assumptions of Theorem 3.6 are satisfied with $\beta(t) = \phi(t)$. □

Corollary 6.4 (*Giusti–Lecerf–Salvy–Yakoubsohn [6]*). *Let $f: D \subset \mathbb{C} \to \mathbb{C}$ be analytic. Assume that $\xi \in D$ is a zero of f with multiplicity m such that $U = \{z \in \mathbb{C} : E(z) < r\} \subset D$, where the function $E: \mathbb{C} \to \mathbb{R}_+$ is defined by (6.1) and*

$$0 < r \le \widetilde{R} = \frac{4m+1-\sqrt{8m^2+8m+1}}{4m}. \tag{6.10}$$

Then for every $z_0 \in U$ Newton iteration (4.1) converges to ξ with error estimate

$$|z_{k+1}-\xi| \le \gamma(\xi)\widetilde{\mu}(r)|z_k-\xi|^2,$$

where the real function $\widetilde{\mu}$ is defined by $\widetilde{\mu}(t) = 1/(2mt^2 - 4mt + m)$.

Proof. The proof follows immediately from Theorem 6.3 taking into account that $\widetilde{R} \le R$ and $\mu(E(x_k)) \le \mu(E(x_0)) \le \mu(r) \le \widetilde{\mu}(r)$. □

Note that in the case $m = 1$ both Theorem 6.3 and Corollary 6.4 coincide with the γ-theorem of Smale [17]. If $m \ge 2$, we have $\widetilde{R} < 1 - \sqrt{2}/2 \le R$ and $\widetilde{\mu}(r) > \mu(r)$. Hence, Theorem 6.3 gives larger convergence ball and better error bounds than Corollary 6.4.

Remark 6.5. From (6.9) it is easy to see that $R < m/(2m+1) < 1/2$. According to Theorem 6.3 Newton iteration is convergent in $U(\xi, R/\gamma(\xi))$ provided that f is analytic in this ball. Hence ξ is the unique zero of f in the ball $U(\xi, R/\gamma(\xi))$. This shows that two distinct zeros ξ and ξ' of f are at distance at least $R/\gamma(\xi)$. In fact Dedieu [4] (case $m = 1$) and Yakoubsohn [24,25] (case $m \geq 1$) have proved that this last result holds with $1/(2\gamma(\xi))$ instead of $R/\gamma(\xi)$.

Taking into account that the function ϕ defined in Theorem 6.3 is strictly increasing and continuous on the interval $J = [0, R]$ and that $\phi(J) = [0, 1]$ we can formulate the following theorem.

Theorem 6.6. *Let $f : D \subset \mathbb{C} \to \mathbb{C}$ and $0 < \lambda < 1$. Assume $\xi \in D$ is a zero of f with multiplicity m and f is analytic in a ball $U = \{z \in \mathbb{C} : E(z) < r\} \subset D$, where the function $E : \mathbb{C} \to \mathbb{R}_+$ is defined by (6.1) and*

$$0 < r \leq R_\lambda = \frac{3m + 1 + 1/\lambda - \sqrt{(3m + 1 + 1/\lambda)^2 - 8m^2}}{4m}. \tag{6.11}$$

Then for every $z_0 \in U$ Newton iteration (4.1) converges Q-quadratically to ξ with error estimate

$$|z_k - \xi| \leq \lambda^{2^k - 1} |z_0 - \xi|.$$

Setting $\lambda = 1/2$ in Theorem 6.6 we obtain the following generalization of the γ-theorem of Smale [17] (see also [3, Chapter 8, Theorem 1]). In the case $m = 1$ we get Smale's result.

Corollary 6.7. *Let $f : D \subset \mathbb{C} \to \mathbb{C}$. Assume $\xi \in D$ is a zero of f with multiplicity m and f is analytic in a ball $U = \{z \in \mathbb{C} : E(z) < r\} \subset D$, where the function $E : \mathbb{C} \to \mathbb{R}_+$ is defined by (6.1) and*

$$0 < r \leq R = \frac{3m + 3 - \sqrt{m^2 + 18m + 9}}{4m}. \tag{6.12}$$

Then every $z_0 \in U$ is an approximate zero of f with associated zero ξ, i.e. for all $k \geq 0$ Newton iteration (4.1) satisfies

$$|z_k - \xi| \leq \left(\frac{1}{2}\right)^{2^k - 1} |z_0 - \xi|.$$

7. Convergence ball of Newton's method in Banach spaces

In this section we apply Corollary 3.7 to Newton's method

$$x_{n+1} = x_n - F'(x_n)^{-1} F(x_n), \quad n = 0, 1, \ldots, \tag{7.1}$$

where $F : D \subset X \to Y$ is a Fréchet differentiable operator defined on a subset D of a Banach space X with values in a Banach space Y. Suppose that F has a simple zero ξ, i.e. $F(\xi) = 0$ and $F'(\xi)^{-1}$ exists and is bounded. Recall that an open ball $U(\xi, r) \subset D$ with center ξ and radius r is called a *convergence ball* of Newton's method, if Newton iteration (7.1) starting from each point $x_0 \in U(\xi, r)$ is well-defined and converges to ξ. An interesting problem is to establish the optimal radius of the convergence ball of Newton iteration. The classical works in this area are due to Rall [13], Rheinboldt [15], Traub and Woźniakowski [19], Ypma [26,27] and Smale [17]. In this section we study Newton iteration with respect to the standard function of initial conditions

$$E(x) = \|x - \xi\|.$$

The main result in this section gives a unified theory for convergence ball of Newton's method and extends the corresponding results of Traub and Woźniakowski [19], Smale [17], Wang [21], Wang and Li [23], Argyros and Gutiérrez [2] and others.

In 1979, Traub and Woźniakowski [19] proved that if ξ is a simple zero of F and F' satisfies an affine invariant Lipschitz condition

$$\|F'(\xi)^{-1}(F'(x) - F'(y))\| \leq L(\|x - y\|) \tag{7.2}$$

for all $x, y \in U(\xi, r)$, then Newton iteration starting from every point x_0 in the open ball $U(\xi, r)$ with radius $r = 2/(3L)$ is well-defined, remains in $U(\xi, r)$ and converges quadratically to ξ. Moreover, they proved also that $r = 2/(3L)$ is the optimal radius of the convergence ball of Newton's method under the condition (7.2). It is well-known that Traub–Woźniakowski's result remains true even if the condition (7.2) is satisfied only for the points y lying in the segment joining the points ξ and x. In what follows for two points $u, v \in X$ we denote by $[u, v]$ the segment joining these points, i.e.

$$[u, v] = \{x \in X : x = tu + (1 - t)v, 0 \leq t \leq 1\}. \tag{7.3}$$

In 2000, Wang [21] generalized Traub and Woźniakowski's result replacing (7.2) by

$$\|F'(\xi)^{-1}(F'(x) - F'(\xi + t(x - \xi)))\| \leq \int_{t\|x-\xi\|}^{\|x-\xi\|} L(u)\mathrm{d}u \tag{7.4}$$

for all $x \in U(\xi, r)$ and $0 \leq t \leq 1$, where L is a nondecreasing function on $[0, r]$. In 2005, Argyros and Gutiérrez [2] proposed the following more general condition

$$\|F'(\xi)^{-1}(F'(x) - F'(\xi + t(x - \xi)))\| \leq f(t, \|x - \xi\|) \tag{7.5}$$

for all $x \in U(\xi, r)$ and $0 \leq t \leq 1$, where f is a real function of two variables.

In this section we prove a general convergence theorem for Newton's method under the following two conditions:

$$\|F'(\xi)^{-1}(F'(x) - F'(y))\| \leq \Omega(\|x - y\|, \|x - \xi\|, \|y - \xi\|) \tag{7.6}$$

and

$$\|F'(\xi)^{-1}(F'(x) - F'(\xi))\| \leq \omega(\|x - \xi\|) \tag{7.7}$$

for all $x \in D$ with $\|x - \xi\| < r$ and all $y \in [\xi, x]$, where Ω and ω are nonnegative functions defined on $[0, r)^3$ and $[0, r)$ respectively. Without loss of generality we assume that $\omega(0) = 0$. It is easy to see that conditions (7.5) and (7.6) are equivalent.

Lemma 7.1. *Let $F: D \subset X \to Y$ be a continuously Fréchet differentiable operator on an open convex set D. Suppose $\xi \in D$ is a simple zero of F and the conditions (7.6) and (7.7) are satisfied. Assume that the real function*

$$\varphi(t) = \frac{\int_0^1 \Omega(t, u, t - u)\mathrm{d}u}{1 - \omega(t)} \tag{7.8}$$

is well-defined on the interval $[0, r)$. Then

$$\|Tx - \xi\| \leq \varphi(\|x - \xi\|) \quad \text{for all } x \in D \text{ with } \|\xi - x\| < r, \tag{7.9}$$

where T is Newton's operator, i.e. $Tx = x - F'(x)^{-1}F(x)$.

Proof. Let $x \in D$ be such that $\|x - \xi\| < r$. Note that since φ is well-defined on $[0, r)$, then $\omega(t) < 1$ for all $t \in [0, r)$. From this and (7.7) we get

$$\|F'(\xi)^{-1}F'(x) - I\| = \|F'(\xi)^{-1}[F'(x) - F'(\xi)]\| \leq \omega(\|x - \xi\|) < 1$$

since $0 \leq \|x - \xi\| < r$. It follows from Banach's theorem on invertible operator (see Kantorovich and Akilov [10, Theorem V.4.3]) that $F'(x)^{-1}$ exists and

$$\|F'(x)^{-1}F'(\xi)\| \leq \frac{1}{1 - \omega(\|x - \xi\|)}. \tag{7.10}$$

Further, we have

$$\begin{aligned} Tx - \xi &= -F'(x)^{-1}(F(x) - F(\xi) - F'(x)(x-\xi)) \\ &= F'(x)^{-1}F'(\xi)\int_0^1 F'(\xi)^{-1}(F'(x) - F'(x_t)(x-\xi))\mathrm{d}t, \end{aligned}$$

where $x_t = \xi + t(x - \xi)$. From this and (7.6) we obtain

$$\begin{aligned} \|Tx - \xi\| &\le \|F'(x)^{-1}F'(\xi)\| \int_0^1 \|F'(\xi)^{-1}(F'(x) - F'(x_t))\| \|x - \xi\| \mathrm{d}t \\ &\le \|F'(x)^{-1}F'(\xi)\| \int_0^1 \Omega(\|x-\xi\|, \|x_t - \xi\|, \|x - x_t\|)\mathrm{d}t \\ &= \|F'(x)^{-1}F'(\xi)\| \int_0^1 \Omega(\|x-\xi\|, t\|x-\xi\|, (1-t)\|x-\xi\|)\mathrm{d}t \\ &= \|F'(x)^{-1}F'(\xi)\| \int_0^{\|x-\xi\|} \Omega(\|x-\xi\|, u, \|x-\xi\| - u)\mathrm{d}t. \end{aligned}$$

Now combining this with (7.10) we get (7.9) which completes the proof. □

Theorem 7.2 (*Convergence Ball Theorem*). *Let $F: D \subset X \to Y$ be a continuously Fréchet differentiable operator on an open ball $U(\xi, r) \subset D$, where ξ is a simple zero of F. Suppose that F' satisfies conditions (7.6) and (7.7) and the real function φ defined by (7.8) is a strict gauge function of order $p+1$ for some $p \ge 0$ on the interval $[0, r)$. Then for each $x_0 \in U(\xi, r)$ the following statements hold true:*

(i) *Newton iteration (7.1) is well-defined, remains in $U(\xi, r)$ and converges to ξ with Q-order $p+1$.*

(ii) *For all $n \ge 0$ we have the following estimate*

$$\|x_{n+1} - \xi\| \le \varphi(\|x_n - \xi\|). \tag{7.11}$$

(iii) *For all $n \ge 0$ we have the following estimate*

$$\|x_n - \xi\| \le \lambda^{S_n(p+1)} \|x_0 - \xi\|, \tag{7.12}$$

where $\lambda = \phi(\|x_0 - \xi\|)$ and ϕ is a nondecreasing nonnegative function on J satisfying (2.1).

(iv) *If r is a fixed point of φ, then r is the optimal radius of the convergence ball of Newton's method under the conditions (7.6) and (7.7) for some Ω and ω.*

Proof. The statements (i)–(iii) follows immediately from Corollary 3.7 and Lemma 7.1. The statement (iv) will be proved in Theorem 7.3. □

In the following theorems and corollaries in this section we consider some interesting special cases of Theorem 7.2.

Theorem 7.3. *Let $F: D \subset X \to Y$ be a continuously Fréchet differentiable operator on an open ball $U(\xi, r) \subset D$, where ξ is a simple zero of F. Suppose*

$$\|F'(\xi)^{-1}(F'(x) - F'(y))\| \le \omega(\|x - \xi\|) - \omega(\|y - \xi\|) \tag{7.13}$$

for all $x \in U(\xi, r)$ and all $y \in [\xi, x]$, where ω is a real function defined on $[0, r]$ with $\omega(0) = 0$. Assume that

$$\varphi(t) = \frac{t\omega(t) - \int_0^t \omega(u)\mathrm{d}u}{1 - \omega(t)}. \tag{7.14}$$

is a strict gauge function of order $p+1$ for some $p \ge 0$ on $[0, r)$. Then starting from every $x_0 \in U(\xi, r)$ Newton iteration (7.1) is well-defined, remains in $U(\xi, r)$, converges with Q-order $p+1$ to ξ and satisfies the estimates (7.11) and (7.12). Moreover, if r is a fixed point of φ and ω is continuous, then r is the optimal radius of the convergence ball of Newton's method.

Proof. The first part of the theorem follows immediately from Theorem 7.2. Assume that r is a fixed point of φ. We shall prove the exactness of r even in the case $X = Y = \mathbb{R}$. Let ξ be an arbitrary point in $\mathbb{R}$ and r be a positive number. Define a real function F on $D = \overline{U}(\xi, r)$ by

$$F(x) = x - \xi - \operatorname{sign}(x - \xi) \int_0^{|x-\xi|} \omega(t)\mathrm{d}t, \tag{7.15}$$

where ω is continuous on $[0, r]$ and such that $\omega(0) = 0$. It is easy to prove that F is continuously differentiable on $U(\xi, r)$ with $F'(x) = 1 - \omega(|x - \xi|)$. Note that for all $x \in U(\xi, r)$ and $y \in [\xi, x]$ we have

$$|F'(\xi)^{-1}(F'(x) - F'(y))| = \omega(|x - \xi|) - \omega(|y - \xi|)$$

which shows that (7.13) holds with usual norm in $\mathbb{R}$. It is easy to show that $T(\xi + r) = \xi - \varphi(r) = \xi - r$ and $T(\xi - r) = \xi + \varphi(r) = \xi + r$, where $Tx = x - F'(x)^{-1}F(x)$. Therefore, if $x_0 = \xi + r$, then $x_n = \xi + (-1)^n r$ and so Newton iteration (7.1) starting from $x_0 = \xi + r$ is not convergent. □

Remark 7.4. Let us give a sufficient condition for φ defined by (7.14) to be a gauge function of order $p+1$. It follows from Example 2.3that if ω is a nonnegative nondecreasing function on $[0, r)$ such that for all $\lambda \in (0, 1)$ and all $t, u \in [0, r)$ with $t \geq u$ it satisfies $\omega(\lambda t) - \omega(\lambda u) \leq \lambda^p[\omega(t) - \omega(u)]$ for some $p \geq 0$, then the function φ defined by (7.14) is a strict gauge function of order $p + 1$ on J provided that $\varphi(t) < t$ for all $t \in (0, r)$.

Note that in the case $\omega(t) = Lt$ condition (7.13) coincides with (7.2) and we obtain Traub and Woźniakowski's result [19]. Setting in Theorem 7.15 $\omega(t) = \int_0^t L(u)\mathrm{d}u$, where L is nondecreasing on $[0, r]$, we immediately get some results of Wang [21, Theorems 3.1 and 5.1]. Theorem 7.15 is also an improvement of a recent result of Wang and Li [23, Theorem 1.1].

Corollary 7.5. *Let $F: D \subset X \to Y$ be a continuously Fréchet differentiable operator on an open ball $U(\xi, r) \subset D$, where ξ is a simple zero of F. Suppose that (7.13) holds with*

$$\omega(t) = Lt^p \quad (p > 0, L > 0) \quad \text{and} \quad 0 < r \leq R = \left(\frac{p+1}{2p+1}\frac{1}{L}\right)^{1/p}. \tag{7.16}$$

Define the real functions

$$\varphi(t) = \frac{p}{p+1}\frac{Lt^{p+1}}{1 - Lt^p} \quad \text{and} \quad \phi(t) = \frac{p}{p+1}\frac{Lt^p}{1 - Lt^p}. \tag{7.17}$$

Then starting from every $x_0 \in U(\xi, r)$ Newton iteration (7.1) is well-defined, remains in $U(\xi, r)$ and converges with Q-order $p + 1$ to ξ and with the error estimates (7.11) and (7.12). Moreover, $r = R$ is optimal radius of the convergence ball of Newton's method under the condition (7.13) with ω defined by (7.16).

In the case when $p = 1$ we again get the above mentioned result of Traub and Woźniakowski [19] as well as the results of Rheinboldt [15], Wang [20] and Ypma [26]. If $0 < p \leq 1$, then Corollary 7.5 leads to the results of Wang and Li [23, Corollary 3.1] and Huang [8, Theorem 2].

The following example shows that for every positive number p there exist nonlinear operators $F: D \subset X \to Y$ (at least in the case $X = Y = \mathbb{R}$) such that Newton iteration (7.1) converges to a solution of the equation $F(x) = 0$ with the exact order of convergence $p + 1$.

Example 7.6. Let $X = Y = \mathbb{R}$ and let ξ be a point in $\mathbb{R}$. Further, let L and p be two positive numbers. Define the function $F: \mathbb{R} \to \mathbb{R}$ as follows

$$F(x) = (x - \xi)\left(1 - \frac{L|x - \xi|^p}{p+1}\right).$$

The function F is continuously differentiable on $\mathbb{R}$ and satisfies

$$|F'(\xi)^{-1}(F'(x) - F'(y))| = L(|x-\xi|^p - |y-\xi|^p)$$

for all $x \in \mathbb{R}$ and $y \in [\xi, x]$. Define the positive number R and the real functions φ and ϕ as in Corollary 7.5. By Corollary 7.5 it follows that Newton iteration (7.1) starting from every $x_0 \in U(\xi, R)$ is well defined, remains in $U(\xi, R)$ and converges with Q-order $p+1$ to ξ and with error estimates (7.11) and (7.12). Moreover, it is easy to show that for all $x \in U(\xi, (1/L)^{1/p})$ we have $|Tx - \xi| = \varphi(|x - \xi|)$, where again $Tx = x - F'(x)^{-1}F(x)$. Therefore, Newton iteration (7.1) starting from $x_0 \in U(\xi, R)$ satisfies

$$\lim_{n\to\infty} \frac{|x_{n+1} - \xi|}{|x_n - \xi|^{p+1}} = \frac{Lp}{p+1} \quad \text{if } x_n \neq \xi \text{ for sufficiently large } n$$

and so it converges to ξ with exact order $p+1$ and with asymptotic constant $c = Lp/(p+1)$. Note also that R is the optimal radius of the convergence ball since if $x_0 = \xi \pm R$, then Newton iteration (7.1) is not convergent.

Corollary 7.7. *Let $F: D \subset X \to Y$ be a continuously Fréchet differentiable operator on an open ball $U(\xi, r) \subset D$, where ξ is a simple zero of F. Suppose that (7.13) holds with*

$$\omega(t) = \frac{c}{(1-\gamma t)^2} \quad (c > 0, \gamma > 0) \tag{7.18}$$

and

$$0 < r \le R = \frac{3c + 2 - \sqrt{c(9c+8)}}{2(c+1)\gamma}. \tag{7.19}$$

Define the real functions

$$\varphi(t) = \frac{c\gamma t^2}{(c+1)(1-\gamma t)^2 - c} \quad \text{and} \quad \phi(t) = \frac{c\gamma t}{(c+1)(1-\gamma t)^2 - c}. \tag{7.20}$$

Then starting from every $x_0 \in U(\xi, r)$ Newton iteration (7.1) is well-defined, remains in $U(\xi, r)$ and converges Q-quadratically to ξ with the error estimates (7.11) and

$$\|x_n - \xi\| \le \lambda^{2^n - 1}\|x_0 - \xi\| \quad \text{for all } n \ge 0, \tag{7.21}$$

where $\lambda = \phi(\|x_0 - \xi\|)$. Moreover, $r = R$ is optimal radius of the convergence ball of Newton's method under the condition (7.13) with ω defined by (7.18).

Corollary 7.7 without the estimate (7.11) is due to Wang and Han [22] ($c = 1$) and Wang [21, Example 1] ($c > 0$). Note that the function ϕ defined in (7.20) is strictly increasing and continuous on the interval $J = [0, R]$ and it satisfies $\phi(J) = [0, 1]$. Therefore, if we ignore the estimate (7.11), then Corollary 7.7 is equivalent to the following result of Wang [21, p. 132].

Corollary 7.8. *Let $F: D \subset X \to Y$ be a continuously Fréchet differentiable operator on an open ball $U(\xi, r) \subset D$, where ξ is a simple zero of F. Let $0 < \lambda < 1$. Suppose that (7.13) holds with ω defined by (7.18) and*

$$0 < r \le R_\lambda = \frac{2c + 2 + c/\lambda - \sqrt{(2c + 2 + c/\lambda)^2 - 4(c+1)}}{2(c+1)\gamma}. \tag{7.22}$$

Then for every $x_0 \in U(\xi, r)$ Newton iteration (7.1) converges Q-quadratically to ξ with error estimate

$$\|x_n - \xi\| \le \lambda^{2^n - 1}\|x_0 - \xi\| \quad \text{for all } n \ge 0. \tag{7.23}$$

In general Newton iteration converges only locally. In 1979, Traub and Woźniakowski [19] established conditions under which Newton iteration enjoys a “type of global convergence”. The next corollary is a variant of Traub and Woźniakowski's result.

Corollary 7.9. *Let $F: D \subset X \to Y$ be analytic in an open ball $U(\xi, r) \subset D$, where $\xi \in D$ is a simple zero of F. Assume that*

$$\|F'(\xi)^{-1}F^{(k)}(\xi)\| \le ck!\gamma^{k-1} \quad \text{for all } k \ge 2 (c > 0, \gamma > 0) \tag{7.24}$$

and r satisfies (7.19). Define the real functions φ and ϕ by (7.20). Then starting from each $x_0 \in U(\xi, r)$ Newton iteration (7.1) is well-defined, remains in $U(\xi, r)$ and converges Q-quadratically to ξ with the error estimates (7.11) and (7.21). Moreover, $r = R$ is optimal radius of the convergence ball of Newton's method under the condition (7.24).

Proof. The statement follows from Corollary 7.7 since (7.24) implies (7.13) with ω defined by (7.18). □

Traub and Woźniakowski [19] obtained Corollary 7.9 in the case $c = 1$ with $r = 0.182\ldots$, where r is the unique real solution of the equation $(1-t)^3 = 3t$. In this case the optimal $r = (5-\sqrt{17})/4\gamma$ was obtained by Smale [17]. The case $c > 0$ was considered by Wang [21, p. 132].

Corollary 7.10. *Let $F: D \subset X \to Y$ be analytic in an open ball $U(\xi, r) \subset D$, where $\xi \in D$ is a simple zero of F. Let $0 < \lambda < 1$. Assume that (7.24) holds and r satisfies (7.22). Then for every $x_0 \in U(\xi, r)$ Newton iteration (7.1) converges Q-quadratically to ξ with error estimate (7.23).*

Setting $\lambda = 1/2$ in Corollary 7.10 we obtain the following generalization of the well-known γ-theorem of Smale [17] (see also [3, Chapter 8, Theorem 1]). In the case $c = 1$ we get Smale's result. The case $c > 0$ is due to Wang [21, p. 132].

Corollary 7.11. *Let $F: D \subset X \to Y$ be analytic in an open ball $U(\xi, r) \subset D$, where $\xi \in D$ is a simple zero of F. Assume that (7.24) holds and*

$$0 < r \le R = \frac{2c + 1 - \sqrt{c(4c+3)}}{(c+1)\gamma}. \tag{7.25}$$

Then every $x_0 \in U(\xi, r)$ is an approximate zero of F with associated zero ξ, i.e. Newton iteration (7.1) satisfies

$$\|x_n - \xi\| \le \left(\frac{1}{2}\right)^{2^n - 1} \|x_0 - \xi\| \quad \text{for all } n \ge 0.$$

The following theorem is another natural generalization of Traub and Woźniakowski's result [19] mentioned in the beginning of the section.

Theorem 7.12. *Let $F: D \subset X \to Y$ be a continuously Fréchet differentiable operator on an open ball $U(\xi, r) \subset D$, where ξ is a simple zero of F. Suppose*

$$\|F'(\xi)^{-1}(F'(x) - F'(y))\| \le \omega(\|x - y\|) \tag{7.26}$$

for all $x \in U(\xi, r)$ and all $y \in [\xi, x]$, where ω is a real function defined on $[0, r]$ with $\omega(0) = 0$. Assume that

$$\varphi(t) = \frac{\int_0^t \omega(u)\mathrm{d}u}{1 - \omega(t)}. \tag{7.27}$$

is a strict gauge function of order $p+1$ for some $0 \le p \le 1$ on $[0, r)$. Then starting from every $x_0 \in U(\xi, r)$ Newton iteration (7.1) is well-defined, remains in $U(\xi, r)$ and converges with Q-order $p+1$ to ξ and with the error estimates (7.11) and (7.12).

Corollary 7.13. *Let $F: D \subset X \to Y$ be a continuously Fréchet differentiable operator on an open ball $U(\xi, r) \subset D$, where ξ is a simple zero of F. Suppose that (7.26) holds with*

$$\omega(t) = Lt^p \quad (0 \le p \le 1, L > 0) \quad \text{and} \quad 0 < r \le R = \left(\frac{p+1}{p+2}\frac{1}{L}\right)^{1/p}. \tag{7.28}$$

In the case $p = 0$ we assume that $L < 1/2$ and $R = \infty$. Define the functions

$$\varphi(t) = \frac{1}{p+1}\frac{Lt^{p+1}}{1 - Lt^p} \quad \text{and} \quad \phi(t) = \frac{1}{p+1}\frac{Lt^p}{1 - Lt^p}. \tag{7.29}$$

Then starting from every $x_0 \in U(\xi, r)$ Newton iteration (7.1) is well-defined, remains in $U(\xi, r)$ and converges with Q-order $p + 1$ to ξ with the error estimates (7.11) and (7.12).

From Corollary 7.13 in the case $p = 1$ we again get the classical results of Rheinboldt [15], Traub and Woźniakowski [19], Wang [21], Ypma [26]. In the case $0 \le p \le 1$ Corollary 7.13 is obtained by Ypma [27, Theorem 3.1], Huang [8, Theorem 1] and in a slightly different form by Argyros [1, Theorem 4].

Using the following proposition we can show that if the assumptions of Theorem 7.12 or Corollary 7.13 are satisfied with $p > 1$, then $F(x) = A(x - \xi)$, where A is a linear operator, i.e. in this case the equation $F(x) = 0$ is linear.

Proposition 7.14. *Let $F: D \subset X \to Y$ be an operator on a convex domain D and $\xi \in D$. Suppose*

$$\|F(x) - F(y)\| \le \omega(\|x - y\|) \quad \text{for all } x \in D \text{ and } y \in [\xi, x],$$

where $\omega: J \to \mathbb{R}$ is a nondecreasing function such that $\lim_{t\to 0} \omega(t)/t = 0$. Then the operator F is constant on D.

Proof. Define the function $\Omega: J \to \mathbb{R}$ as follows

$$\Omega(t) = \sup\{\|F(x) - F(y)\| : x \in D, y \in [\xi, x], \|x - y\| \le t\}.$$

Obviously, $\Omega(t) \le \omega(t)$ for all $t \in J$. This implies $\lim_{t\to 0} \Omega(t)/t = 0$. It is easy to show that Ω is nondecreasing and $\Omega(t_1 + t_2) \le \Omega(t_1) + \Omega(t_2)$ for all $t_1, t_2 \in J$ with $t_1 + t_2 \in J$. Using the last inequality one can show that $\Omega(ut) \le (u + 1)\Omega(t)$ for all $u, t \in J$ with $(u + 1)t \in J$. Now we shall prove that Ω is identical to zero. Without loss of generality we can assume that $J = [0, R)$, where $0 < R \le +\infty$. Suppose Ω is not identical to zero. Then there exists $\delta \in J$ such that $\Omega(\delta) > 0$. For every $t \in (0, \delta)$ with $\delta + t \in J$ we have

$$\Omega(\delta) = \Omega((\delta/t)t) \le (\delta/t + 1)\Omega(t) \le 2(\delta/t)\Omega(t)$$

which can be written in the form $\Omega(t)/t \ge (1/2)\Omega(\delta)/\delta > 0$. The last inequality contradicts to the fact that $\lim_{t\to 0} \Omega(t)/t = 0$. Consequently, $\Omega(t) \equiv 0$ which implies that $F(x) = F(\xi)$ for every $x \in D$. □

Now assume that an operator F satisfies assumptions of Corollary 7.13 (or Theorem 7.12) with $p > 1$. Then it follows from Proposition 7.14 that the operator $F'(\xi)^{-1}F'$ is constant in $U(\xi, r)$ which implies that F' is constant in this ball as well. Hence, for all $x \in U(\xi, r)$ we have $F(x) = A(x - \xi)$, where A is a linear operator.

The next theorem is s an improvement of a recent result of Wang and Li [23, Theorem 1.2].

Theorem 7.15. *Let $F: D \subset X \to Y$ be a continuously Fréchet differentiable operator on an open ball $U(\xi, r) \subset D$, where ξ is a simple zero of F. Suppose*

$$\|F'(\xi)^{-1}(F'(x) - F'(\xi))\| \le \omega(\|x - \xi\|) \tag{7.30}$$

for all $x \in U(\xi, r)$, where ω is a real function defined on $[0, r]$ with $\omega(0) = 0$. Assume that

$$\varphi(t) = \frac{t\omega(t) + \int_0^t \omega(u)\mathrm{d}u}{1 - \omega(t)}. \tag{7.31}$$

is a strict gauge function of order $p + 1$ for some $p \geq 0$ on $[0, r)$. Then starting from every $x_0 \in U(\xi, r)$ Newton iteration (7.1) is well-defined, remains in $U(\xi, r)$ and converges with Q-order $p+1$ to ξ with the error estimates (7.11) and (7.12).

Proof. Note that (7.30) implies that

$$\|F'(\xi)^{-1}(F'(x) - F'(y))\| \leq \omega(\|x - \xi\|) + \omega(\|y - \xi\|) \tag{7.32}$$

for all $x, y \in U(\xi, r)$. Now the proof follows from Theorem 7.2. □

Remark 7.16. Let φ be defined by (7.27) or (7.31). It follows from Example 2.2that if ω is nonnegative nondecreasing on $[0, r)$ and $\omega(t)/t^p$ is nondecreasing on $(0, r)$ for some $p \geq 0$, then φ is a strict gauge function of order $p + 1$ on $[0, r)$ provided that $\varphi(t) < t$ for all $t \in (0, r)$.

Corollary 7.17. *Let $F\colon D \subset X \to Y$ be a continuously Fréchet differentiable operator on an open ball $U(\xi, r) \subset D$, where ξ is a simple zero of F. Suppose that (7.30) holds with*

$$\omega(t) = Lt^p \quad (p \geq 0, L > 0) \quad \textit{and} \quad 0 < r \leq R = \left(\frac{p+1}{2p+3}\frac{1}{L}\right)^{1/p}. \tag{7.33}$$

In the case $p = 0$ we assume that $L < 1/3$ and $R = \infty$. Define the functions

$$\varphi(t) = \frac{p+2}{p+1}\frac{Lt^{p+1}}{1 - Lt^p} \quad \textit{and} \quad \phi(t) = \frac{p+2}{p+1}\frac{Lt^p}{1 - Lt^p}. \tag{7.34}$$

Then starting from every $x_0 \in U(\xi, r)$ Newton iteration (7.1) is well-defined, remains in $U(\xi, r)$ and converges with Q-order $p + 1$ to ξ with the error estimates (7.11) and (7.12).

From Corollary 7.17 in the case $p = 1$ we obtain $R = 2/(5L)$ which improves a recent result of Wang and Li [23, Corollary 3.2]. For example they have proved $R = 1/(3L)$. Note also that Example 7.6 shows that for every $p > 0$ there exist nonlinear functions $F\colon \mathbb{R} \to \mathbb{R}$ satisfying the assumptions of Corollary 7.17. The following corollary is an improvement of another result of Wang and Li [23, Corollary 3.3].

Corollary 7.18. *Let $F\colon D \subset X \to Y$ be a continuously Fréchet differentiable operator on an open ball $U(\xi, r) \subset D$, where ξ is a simple zero of F. Suppose that (7.30) holds with holds with ω defined by (7.18) and*

$$0 < r \leq R = \frac{5c + 2 - \sqrt{c(25c + 8)}}{2(3c + 1)\gamma}. \tag{7.35}$$

Define the real unctions

$$\varphi(t) = \frac{c\gamma t^2(3 - 2\gamma t)}{(c + 1)(1 - \gamma t)^2 - c} \quad \textit{and} \quad \phi(t) = \frac{c\gamma t(3 - 2\gamma t)}{(c + 1)(1 - \gamma t)^2 - c}. \tag{7.36}$$

Then starting from every $x_0 \in U(\xi, r)$ Newton iteration (7.1) is well-defined, remains in $U(\xi, r)$ and converges Q-quadratically to ξ with the error estimates (7.11) and (7.21).

In the convergence ball of Newton's method, the solution ξ of the equation $F(x) = 0$ is certainly unique. But it is well-known that the uniqueness ball of this equation may be larger. In Section 8 we study the problem of the uniqueness ball of the equation $F(x) = 0$.

8. Uniqueness ball of equations in Banach spaces

In this section we apply Corollary 3.7 to establish a general theorem for the uniqueness ball of nonlinear equations in Banach Spaces. Let X and Y be two Banach spaces. Suppose $F: D \subset X \to Y$ and $\xi \in D$ is a simple zero of the equation

$$F(x) = 0. \tag{8.1}$$

Then an open ball $U(\xi, r)$ with center ξ and radius r is called a *uniqueness ball* of equation (8.1) if ξ is a unique solution of (8.1) in $D \cap U(\xi, r)$.

Lemma 8.1. *Let $F: D \subset X \to Y$ be a continuous operator on a convex set D and continuously Fréchet differentiable on* int *D. Suppose $\xi \in \operatorname{int} D$ is a simple zero of F and*

$$\|F'(\xi)^{-1}(F'(x) - F'(\xi))\| \le \omega(\|x - \xi\|) \quad \textit{for all } x \in D \textit{ with } \|x - \xi\| < r, \tag{8.2}$$

where $r > 0$ and ω is such that the function

$$\varphi(t) = \int_0^t \omega(u)\mathrm{d}u \tag{8.3}$$

is well-defined on $[0, r)$ and right continuous at zero. Then

$$\|Tx - \xi\| \le \varphi(\|x - \xi\|) \quad \textit{for all } x \in D \textit{ with } \|x - \xi\| < r, \tag{8.4}$$

where $T: D \subset X \to X$ is defined by $Tx = x - F'(\xi)^{-1}F(x)$.

Proof. For each $x \in \operatorname{int} D$ such that $\|x - \xi\| < r$ we have

$$\begin{aligned} Tx - \xi &= -F'(\xi)^{-1}(F(x) - F(\xi) - F'(\xi)(x - \xi)) \\ &= -\int_0^1 F'(\xi)^{-1}(F'(x_t) - F'(\xi))(x - \xi)\mathrm{d}t, \end{aligned}$$

where $x_t = \xi + t(x - \xi)$. From this and (7.7) we obtain

$$\begin{aligned} \|Tx - \xi\| &\le \int_0^1 \|F'(\xi)^{-1}(F'(x_t) - F'(\xi))\| \|x - \xi\|\mathrm{d}t \\ &\le \int_0^1 \omega(t\|x - \xi\|)\|x - \xi\|\mathrm{d}t = \int_0^{\|x-\xi\|} \omega(u)\mathrm{d}u = \varphi(\|x - \xi\|) \end{aligned}$$

which proves (8.4) in the case when x is an interior point of D. Now let x be an arbitrary point of D with $\|x-\xi\| < r$. Take a real sequence (t_n) in the interval $(0, 1)$ that converges to 0 and set $x_n = \xi + t_n(x-\xi)$. Then the sequence (x_n) lies in $\operatorname{int} D$ and $\|x_n - \xi\| < R$. It follows from the first part of the proof that $\|Tx_n - \xi\| \le \varphi(\|x_n - \xi\|)$. Now by continuity of T and right continuity of φ we get (8.4). □

Theorem 8.2 (*Uniqueness Ball Theorem*). *Let $F: D \subset X \to Y$ be a continuous operator on a convex set D and continuously Fréchet differentiable on* int *D. Suppose $\xi \in \operatorname{int} D$ is a simple zero of F and the condition (8.2) holds. Assume that the real function φ defined by (8.3) is a strict gauge function of the first order on $[0, r)$ and right continuous at 0. Then the Eq. (8.1) has a unique solution in the set $D \cap U(\xi, r)$. Moreover, if r is a fixed point of φ and ω is continuous on $[0, r]$, then r is the optimal radius of the uniqueness ball of the equation $F(x) = 0$ under the condition (8.2).*

Proof. Consider again the modified Newton operator $T: D \subset X \to X$ defined by $Tx = x - F'(\xi)^{-1}F(x)$. Obviously, the set of fixed points of T coincides with the set of zeros of F. By Lemma 8.1 and the first part of Corollary 3.7 we conclude that ξ is a unique fixed point of T in $D \cap U(\xi, r)$. We shall prove the exactness of the radius of the uniqueness ball even in the case $X = Y = \mathbb{R}$. Let ξ be an arbitrary point in $\mathbb{R}$ and r be a positive number. Further, let ω be continuous on $[0, r]$ and $\omega(0) = 0$. Without loss of generality we can assume that ω is defined and continuous on $[0, \infty)$. Define a real function

$F:\mathbb{R}\to\mathbb{R}$ again by (7.15). Note that F is continuously differentiable on $\mathbb{R}$ and ξ is a simple zero of F. Moreover, $|F'(\xi)^{-1}(F'(x)-F'(\xi))|=\omega(|x-\xi|)$ for all $x\in\mathbb{R}$. Hence, condition (8.2) holds with usual norm in $\mathbb{R}$. It is easy to show that $F(\xi+r)=r-\varphi(r)$ and $F(\xi-r)=-r+\varphi(r)$. Now suppose that r is a fixed point of φ. Then $F(\xi\pm r)=0$, i.e. F has three zeros in $\overline{U}(\xi,r)$ which means that $U(\xi,r)$ is the optimal ball of uniqueness of the equation $F(x)=0$. □

Remark 8.3. From Example 2.2 it follows that if ω is nonnegative nondecreasing on $[0,r)$, then φ defined by (8.3) is a strict gauge function of the first order on $[0,r)$ provided that $\varphi(t)<t$ for all $t\in(0,r)$. In this case φ is right continuous at 0 as well.

From Theorem 8.2 one can obtain some results of Wang [21, Theorems 4.1 and 5.2].

Corollary 8.4. *Let $F:D\subset X\to Y$ be a continuous operator on a convex set D and continuously Fréchet differentiable on* int$D$*. Suppose $\xi\in$ int D is a simple zero of F and the condition (8.2) holds with*

$$\omega(t)=Lt^p \quad (p\ge 0, L>0) \quad \textit{and} \quad 0<r\le R=\left(\frac{p+1}{L}\right)^{1/p}. \tag{8.5}$$

In the case $p=0$ we assume that $L<1$ and $R=\infty$. Then the equation $F(x)=0$ has a unique solution in the set $D\cap U(\xi,r)$. Moreover, $r=R$ is the optimal radius of the uniqueness ball under the condition (8.2) with ω defined by (8.5).

Corollary 8.4 is obtained by Wang [21, Corollary 6.2] ($p=1$) and Huang [8, Theorem 3] ($0<p\le 1$). Note that Wang [21] and Huang [8] have formulated their results in a slightly weaker form assuming that $U(\xi,r)\subset D$. Note also that Example 7.6 shows that for every $p>0$ there exist nonlinear functions $F:\mathbb{R}\to\mathbb{R}$ satisfying the assumptions of Corollary 8.4.

Corollary 8.5. *Let $F:D\subset X\to Y$ be a continuous operator on a convex set D and continuously Fréchet differentiable on* int$D$*. Suppose $\xi\in$ int D is a simple zero of F and the condition (8.2) holds with*

$$\omega(t)=\frac{c}{(1-\gamma t)^2}-c \quad (c>0, \gamma>0) \quad \textit{and} \quad 0<r\le R=\frac{1}{(c+1)\gamma}. \tag{8.6}$$

Then the equation $F(x)=0$ has a unique solution in the set $D\cap U(\xi,r)$ and $r=R$ is the optimal radius of the uniqueness ball under the condition (8.2) with ω defined by (8.6).

Corollary 8.5 is due to Wang [21, p. 130]. When F is analytic and satisfies (7.24) with $c=1$ then this result is due to Dedieu [4].

9. Conclusion

In this paper we present a general method for proving convergence results for iterative processes of the type $x_{n+1}=Tx_n$, where $T:D\subset X\to X$ is an iteration function in a metric space (X,d). To establish an unified local convergence theory we introduce the notion of a *function of initial conditions* of T. A real-valued function $E:D\to\mathbb{R}_+$ is said to be a function of initial conditions of T (with a gauge function φ on an interval J) if there exists a function $\varphi:J\to J$ such that

$$E(Tx)\le\varphi(E(x)) \quad \text{for all } x\in D \text{ with } Tx\in D \text{ and } E(x)\in J. \tag{9.1}$$

Further, we assume that T has a fixed point $\xi\in X$ and that T satisfies a condition of the type

$$d(Tx,\xi)\le\beta(E(x))d(x,\xi) \quad \text{for all } x\in D \text{ with } E(x)\in J, \tag{9.2}$$

where β is a nondecreasing function on J satisfying $0\le\beta(t)<1$ for all $t\in J$. The main theorems of the paper (Theorems 3.6 and 3.8) give a method for obtaining local convergence theorems with order of convergence $r\ge 1$ for given iteration function T with respect to given function of initial conditions

E. We apply our theory to Newton's operator $Tx = x - F'(x)^{-1}F(x)$ with respect to four different functions of initial conditions of the type

$$E(x) = C(x)d(x, \xi), \tag{9.3}$$

where $C: D \to \mathbb{R}_+$. Let us give an easy algorithm for obtaining a local convergence result for given iteration function T with respect to given function of initial conditions E of the type (9.3).

Step 1. Consider $d(Tx, \xi)$ and derive an inequality of the type (9.2) with a nondecreasing function β on an interval $J = J_1$.
Step 2. Derive an inequality of the type $C(Tx) \le \gamma(E(x))C(x)$ for all $x \in D$ such that $E(x) \in J$ and $Tx \in D$, where γ is a real function defined on an interval $J = J_2$.
Step 3. Using Step 1 and Step 2 prove an inequality of the type (9.1) with the function φ defined by $\varphi(t) = t\beta(t)\gamma(t)$ on the interval $J = J_1 \cap J_2$.
Step 4. Find an interval $J \subset J_1 \cap J_2$ such that φ is a gauge function of order $r \ge 1$ on J and apply Theorem 3.6 or Theorem 3.8

Finally, let us note that the idea of this paper can be developed (see [12]) to establish general semilocal convergence theorems in which we do not assume the existence of fixed points or zeros.

References

[1] I.K. Argyros, On the convergence and application of Newton's method under weak Hölder continuity assumptions, Int. J. Comput. Math. 80 (2003) 767–780.
[2] I.K. Argyros, J.M. Gutiérrez, A unified approach for enlarging the radius of convergence for Newton's method and applications, Nonlinear Funct. Anal. Appl. 10 (2005) 555–563.
[3] L. Blum, F. Cucker, M. Shub, S. Smale, Complexity and Real Computation, Springer, New York, 1998.
[4] J.-P. Dedieu, Condition number analysis for sparse polynomial systems, in: F. Cucker, M. Shub (Eds.), Foundations of Computational Mathematics, Springer, Berlin, 1997, pp. 75–101.
[5] K. Dočev, Über Newtonsche Iterationen, C. R. Acad. Bulg. Sci. 36 (1962) 695–701.
[6] M. Giusti, G. Lecerf, B. Salvy, J.-C. Yakoubsohn, Location and approximation of clusters of zeros of analytic functions, Found. Comput. Math. 5 (3) (2005) 257–311.
[7] M. Giusti, G. Lecerf, B. Salvy, J.-C. Yakoubsohn, On location and approximation of clusters of zeros: Case of embedding dimension one, Found. Comput. Math. 7 (1) (2007) 1–58.
[8] Z.D. Huang, The convergence ball of Newton's method and uniqueness ball of equations under Hölder-type continuous derivatives, Comput. Math. Appl. 47 (2004) 247–251.
[9] L.O. Jay, A note on Q-order of convergence, BIT 41 (2001) 422–429.
[10] L.V. Kantorovich, G.P. Akilov, Functional Analysis in Normed Spaces, Pergamon, Oxford, 1982.
[11] P.D. Proinov, A generalization of the Banach contraction principle with high order of convergence of successive approximations, Nonlinear Anal. 67 (2007) 2361–2369.
[12] P.D. Proinov, New general convergence theory for iterative processes and its applications to Newton–Kantorovich theorem (in preparation).
[13] L.B. Rall, A note on the convergence of Newton's method, SIAM J. Numer. Anal. 11 (1974) 34–36.
[14] J. Renegar, On the worst-case arithmetic complexity of approximating zeros of polynomials, J. Complexity 3 (1987) 90–113.
[15] W.C. Rheinboldt, An adaptive continuation process for solving systems of nonlinear equations, in: Banach Center Publications, vol. 3, Polish Academy of Science, Warsaw, 1977, pp. 129–142.
[16] E. Schröder, Über unendlich viele Algorithmen zur Autlösung der Glleichungen, Math. Anal. 2 (1870) 317–365.
[17] S. Smale, Newton's method estimates from data at one point, in: R.E. Ewing, K.E. Gross, C.F. Martin (Eds.), The Merging of Disciplines: New Direction in Pure, Applied, and Computational Mathematics, Springer, New York, 1986, pp. 185–196.
[18] P. Tilli, Convergence conditions of some methods for the simultaneous computation of polynomial zeros, Calcolo 35 (1998) 3–15.
[19] J.F. Traub, H. Woźniakowski, Convergence and complexity of Newton iteration for operator equations, J. Assoc. Comput. Mach. 26 (1979) 250–258.
[20] X.H. Wang, The convergence of Newton's method, Kexue Tongbao 25 (1980) 36–37 (in Chinese).
[21] X.H. Wang, Convergence of Newton's method and uniqueness of the solution of equations in Banach space, IMA J. Numer. Anal. 20 (2000) 123–134.
[22] X.H. Wang, D.F. Han, Criterion α and Newton's method under weak conditions, Math. Numer. Sin. 19 (1997) 103–112 (in Chinese) English translation in: Chinese J. Numer. Math. Appl. 19 (1997) 96–105.
[23] X.H. Wang, C. Li, Convergence of Newton's method and uniqueness of the solution of equations in Banach spaces. II, Acta Math. Sinica (Engl. Ser.) 19 (2003) 405–412.
[24] J.-C. Yakoubsohn, Finding a cluster of zeros of univariate polynomials, J. Complexity 16 (2000) 603–638.
[25] J.-C. Yakoubsohn, Numerical elimination, Newton method and multiple roots, in: F. Chyzak (Ed.), Algorithms Seminar 2001–2002, INRIA, 2003, pp. 49–54. see http://algo.inria.fr/seminars/.
[26] T.J. Ypma, Affine invariant convergence results for Newton's method, BIT 22 (1982) 108–118.
[27] T.J. Ypma, Local convergence of inexact Newton methods, SIAM J. Numer. Anal. 21 (1984) 583–589.