Iterative Algorithms I

Ioannis K. Argyros
Á. Alberto Magreñán

Mathematics Research Developments

NOVA

MATHEMATICS RESEARCH DEVELOPMENTS

ITERATIVE ALGORITHMS I

MATHEMATICS RESEARCH DEVELOPMENTS

Additional books in this series can be found on Nova's website
under the Series tab.

Additional e-books in this series can be found on Nova's website
under the e-book tab.

MATHEMATICS RESEARCH DEVELOPMENTS

ITERATIVE ALGORITHMS I

IOANNIS K. ARGYROS
AND
Á. ALBERTO MAGREÑÁN

nova publishers
New York

Additional color graphics may be available in the e-book version of this book.

Library of Congress Cataloging-in-Publication Data

Names: Argyros, Ioannis K., editor. | Magreñán, Á. Alberto (Angel Alberto), editor.
Title: Iterative algorithms / editors, Ioannis K. Argyros and Á. Alberto Magreñán
(Cameron University, Department of Mathematical Sciences,
Lawton, OK, USA, and others).
Description: Hauppauge, New York: Nova Science Publishers, Inc., [2016]- |
Series: Mathematics research developments | Includes index.
Identifiers: LCCN 2016021559 (print) | LCCN 2016025038 (ebook) | ISBN 9781634854061 (hardcover: v. 1) | ISBN 9781634854221
Subjects: LCSH: Iterative methods (Mathematics) | Algorithms. | Numerical analysis.
Classification: LCC QA297.8 .I834 2016 (print) | LCC QA297.8 (ebook) | DDC 518/.26--dc23
LC record available at https://lccn.loc.gov/2016021559

Published by Nova Science Publishers, Inc. † New York

Contents

Dedicated to
My mother Anastasia

Dedicated to
My parents Alberto and Mercedes
My grandmother Ascensión
My beloved Lara

Preface

It is a well-known fact that iterative methods have been studied since problems where we cannot find a solution in a closed form. There exist methods with different behaviors when they are applied to different functions, methods with higher order of convergence, methods with great zones of convergence, methods which do not require the evaluation of any derivative, etc. and researchers are developing new iterative methods frequently.

Once these iterative methods appeared, several researchers have studied them in different terms: convergence conditions, real dynamics, complex dynamics, optimal order of convergence, etc. This phenomena motivated the authors to study the most used and classical ones as for example Newton's method or its derivative-free alternative the Secant method.

Related to the convergence of iterative methods, the most well known conditions are the Kantorovich ones, who developed a theory which has allow many researchers to continue and experiment with these conditions. Many authors in the recent years have studied modifications of theses conditions related, for example, to centered conditions, ω-conditions or even convergence in Hilbert spaces.

In this monograph, we present the complete recent work of the past decade of the authors on Convergence and Dynamics of iterative methods. It is the natural outgrowth of their related publications in these areas. Chapters are self-contained and can be read independently. Moreover, an extensive list of references is given in each chapter, in order to allow reader to use the previous ideas. For these reasons, we think that several advanced courses can be taught using this book.

The list of presented topic of our related studies follows.

Secant-type methods;
Efficient Steffensen-type algorithms for solving nonlinear equations;
On the semilocal convergence of Halley's method under a center-Lipschitz condition on the second Fréchet derivative;
An improved convergence analysis of Newton's method for twice Fréchet differentiable operators;
Expanding the applicability of Newton's method using Smale's α-theory;
Newton-type methods on Riemannian Manifolds under Kantorovich-type conditions;
Improved local convergence analysis of inexact Gauss-Newton like methods;
Expanding the Applicability of Lavrentiev Regularization Methods for Ill-posed Problems;
A semilocal convergence for a uniparametric family of efficient secant-like methods;
On the semilocal convergence of a two-step Newton-like projection method for ill-posed

equations;
New Approach to Relaxed Proximal Point Algorithms Based on $A-$maximal;
Newton-type Iterative Methods for Nonlinear Ill-posed Hammerstein-type Equations;
Enlarging the convergence domain of secant-like methods for equations;
Solving nonlinear equations system via an efficient genetic algorithm with symmetric and harmonious individuals;
On the Semilocal Convergence of Modified Newton-Tikhonov Regularization Method for Nonlinear Ill-posed Problems;
Local convergence analysis of proximal Gauss-Newton method for penalized nonlinear least squares problems;
On the convergence of a Damped Newton method with modified right-hand side vector;
Local convergence of inexact Newton-like method Under weak Lipschitz conditions;
Expanding the applicability of Secant method with applications;
Expanding the convergence domain for Chun-Stanica-Neta family of third order methods in Banach spaces;
Local convergence of modified Halley-like methods with less computation of inversion;
Local convergence for an improved Jarratt-type method in Banach space;
Enlarging the convergence domain of secant-like methods for equations.

The book's results are expected to find applications in many areas of applied mathematics, engineering, computer science and real problems. As such this monograph is suitable to researchers, graduate students and seminars in the above subjets, also to be in all science and engineering libraries.

The preparation of this book took place during 2015-2016 in Lawton, Oklahoma, USA and Logroño, La Rioja, Spain.

Ioannis K. Argyros
Á. Alberto Magreñán

April 2016

Chapter 1

Secant-Type Methods

1.1. Introduction

In this chapter we are concerned with the problem of approximating a locally unique solution $x^\star$ of the nonlinear equation

$$F(x) = 0, \tag{1.1.1}$$

where, F is a Fréchet-differentiable operator defined on a nonempty subset $\mathcal{D}$ of a Banach space $\mathcal{X}$ with values in a Banach space $\mathcal{Y}$. A lot of problems from Applied Sciences can be expressed in a form like (1.1.1) using mathematical modelling [3]. The solutions of these equations can be found in closed form only in special cases. That is why the most solution methods for these equations are iterative. The convergence analysis of iterative methods is usually divided into two categories: semilocal and local convergence analysis. In the semilocal convergence analysis one derives convergence criteria from the information around an initial point whereas in the local analysis one finds estimates of the radii of convergence balls from the information around a solution. If $\mathcal{X} = \mathcal{Y}$ and $Q(x) = F(x) + x$, then the solution x^* of equation (1.1.1) is very important in fixed point theory.

We study the convergence of the secant-type method

$$x_{n+1} = x_n - \mathcal{A}_n^{-1} F(x_n),\ \mathcal{A}_n = \delta F(x_n, y_n) \text{ for each } n = 1, 2, \cdots, \tag{1.1.2}$$

where x_{-1}, x_0 are initial points, $y_n = \theta_n x_n + (1 - \theta_n) x_{n-1}$, $\theta_n \in \mathbb{R}$. Here $\mathcal{A}_n \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$, $x, y \in \mathcal{D}$ is a consistent approximation of the Fréchet-derivative of F (see page 182 of [15] or the second estimate in condition $(\mathcal{D}_4)$ of Definition 3.1). $\mathcal{L}(\mathcal{X}, \mathcal{Y})$ stands for the space of bounded linear operators from $\mathcal{X}$ to $\mathcal{Y}$. Many iterative methods are special cases of (1.1.2). Indeed, if $\theta_n = 1$, then we obtain Newton's method

$$x_{n+1} = x_n - F'(x_n)^{-1} F(x_n) \text{ for each } n = 0, 1, 2 \ldots; \tag{1.1.3}$$

if $\theta_n = 0$, we obtain the secant method

$$x_{n+1} = x_n - \delta F(x_n, x_{n-1})^{-1} F(x_n) \text{ for each } n = 0, 1, 2 \ldots; \tag{1.1.4}$$

if $\theta_n = 2$, we obtain the Kurchatov method

$$x_{n+1} = x_n - \delta F(x_n, 2x_n - x_{n-1})^{-1} F(x_n) \text{ for each } n = 0, 1, 2 \ldots. \tag{1.1.5}$$

Other choices of θ_n are also possible [1, 2, 6, 8, 9, 12, 14, 15, 21, 22]. There is a plethora of sufficient convergence criteria for special cases of secant-type methods (1.1.3)-(1.1.5) under Lipschitz-type conditions (1.1.2) (see [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 15, 16, 17, 18, 19, 20, 21, 22] and the references there in) or even graphical tools to study them [13]. Therefore, it is important to study the convergence of the secant-type method in a unified way. It is interesting to notice that although we use very general majorizing sequences for $\{x_n\}$ our technique leads in the semilocal case to: weaker sufficient convergence criteria; more precise estimates on the distances $\|x_n - x_{n-1}\|$, $\|x_n - x^*\|$ and an at least as precise information on the location of the solution x^* in many interesting special cases such as Newton's method or the secant method (see Remark 3.3 and the Examples). Moreover, in the local case: a larger radius of convergence and more precise error estimates than in earlier studies such as [8, 10, 11, 12, 14, 15, 16, 17, 18, 19, 20, 21, 22] are obtained in this chapter (see Remark 4.2 and the Examples).

The chapter is organized as follows. In Section 1.2 we study the convergence of the majorizing sequences for $\{x_n\}$. Section 1.3 contains the semilocal and Section 1.4 the local convergence analysis for $\{x_n\}$. The numerical examples are given in the concluding Section 1.5. In particular, in the local case we present an example where the radius of convergence is larger than the one given by Rheinboldt [18] and Traub [19] for Newton's method. Moreover, in the semilocal case we provide an example involving a nonlinear integral equation of Chandrasekhar type [7] appearing in radiative transfer as well as an example involving a two point boundary value problem.

1.2. Majorizing Sequences for the Secant-Type Method

In this Section, we shall first study some scalar sequences which are related to the secant-type method.

Let there be parameters $c \geq 0$, $\nu \geq 0$, $\lambda \geq 0$, $\mu \geq 1$, $l_0 > 0$ and $l > 0$ with $l_0 \leq l$. Define the scalar sequence $\{\alpha_n\}$ by

$$\begin{cases} \alpha_{-1} = 0,\ \alpha_0 = c, \alpha_1 = c + \nu \\ \alpha_{n+2} = \alpha_{n+1} + \dfrac{l\,(\alpha_{n+1} - \alpha_n + \lambda(\alpha_n - \alpha_{n-1}))\,(\alpha_{n+1} - \alpha_n)}{1 - l_0\,[\mu(\alpha_{n+1} - c) + \lambda(\alpha_n - c) + c]} \quad \text{for each } n = 0, 1, 2, \cdots \end{cases} \tag{1.2.1}$$

Special cases of the sequence $\{\alpha_n\}$ have been used as majorizing sequences for secant-type method by several authors. For example: **Case 1** (secant method) $l_0 = l$, $\lambda = 1$ and $\mu = 1$ has been studied in [6, 8, 9, 12, 14, 15, 20, 21] and for $l_0 \leq l$ in [2, 4]. **Case 2** (Newton's method) $l_0 = l$, $\lambda = 0$, $c = 0$ and $\mu = 2$ has been studied in [1, 8, 10, 11, 12, 14, 15, 17, 18, 19, 21, 22] and for $l_0 \leq l$ in [2, 3, 4]. In the present chapter we shall study the convergence of sequence $\{\alpha_n\}$ by first simplifying it. Indeed, the purpose of the following transformations is to study the sequence (1.2.1) after using easier to study sequences defined by (1.2.3), (1.2.6) and (1.2.8). Let

$$L_0 = \frac{l_0}{1 + (\mu + \lambda - 1) l_0 c} \text{ and } L = \frac{l}{1 + (\mu + \lambda - 1) l_0 c}. \tag{1.2.2}$$

Using (1.2.1) and (1.2.2), sequence $\{\alpha_n\}$ can be written as

$$\begin{cases} \alpha_{-1}=0,\ \alpha_0=c, \alpha_1=c+\nu \\ \alpha_{n+2}=\alpha_{n+1}+\dfrac{L\left(\alpha_{n+1}-\alpha_n+\lambda(\alpha_n-\alpha_{n-1})\right)(\alpha_{n+1}-\alpha_n)}{1-L_0(\mu\alpha_{n+1}+\lambda\alpha_n)} \text{ for each } n=0,1,2,\cdots \end{cases} \tag{1.2.3}$$

Moreover, let

$$L=bL_0 \text{ for some } b\geq 1 \tag{1.2.4}$$

and

$$\beta_n=L_0\alpha_n. \tag{1.2.5}$$

Then, we can define sequence $\{\beta_n\}$ by

$$\begin{cases} \beta_{-1}=0,\ \beta_0=L_0c,\ \beta_1=L_0(c+\nu) \\ \beta_{n+2}=\beta_{n+1}+\dfrac{b\left(\beta_{n+1}-\beta_n+\lambda(\beta_n-\beta_{n-1})\right)(\beta_{n+1}-\beta_n)}{1-(\mu\beta_{n+1}+\lambda\beta_n)} \text{ for each } n=0,1,2,\cdots \end{cases} \tag{1.2.6}$$

Furthermore, let

$$\gamma_n=\frac{1}{\mu+\lambda}-\beta_n \text{ for each } n=0,1,2,\cdots. \tag{1.2.7}$$

Then, sequence $\{\gamma_n\}$ is defined by

$$\begin{cases} \gamma_{-1}=\frac{1}{\mu+\lambda}, \gamma_0=\frac{1}{\mu+\lambda}-L_0c,\ \gamma_1=\frac{1}{\mu+\lambda}-L_0(c+\nu) \\ \gamma_{n+2}=\gamma_{n+1}-\dfrac{b\left(\gamma_{n+1}-\gamma_n+\lambda(\gamma_n-\gamma_{n-1})\right)(\gamma_{n+1}-\gamma_n)}{\mu\gamma_{n+1}+\lambda\gamma_n} \text{ for each } n=0,1,2,\cdots \end{cases} \tag{1.2.8}$$

Finally, let

$$\delta_n=1-\frac{\gamma_n}{\gamma_{n-1}} \text{ for each } n=0,1,2,\cdots \tag{1.2.9}$$

Then, we define the sequence $\{\delta_n\}$ by

$$\begin{cases} \delta_0=1-\frac{\gamma_0}{\gamma_{-1}}, \delta_1=1-\frac{\gamma_1}{\gamma_0} \\ \delta_{n+2}=\dfrac{b\delta_{n+1}\left(\lambda\delta_n+(1-\delta_n)\delta_{n+1}\right)}{(1-\delta_n)(1-\delta_{n+1})\left(\mu(1-\delta_{n+1})+\lambda\right)} \text{ for each } n=0,1,2,\cdots \end{cases} \tag{1.2.10}$$

It is convenient for the study of the convergence of the sequence $\{\alpha_n\}$ to define polynomial p by

$$p(t)=\mu t^3-(\lambda+3\mu+b)t^2+(2\lambda+3\mu+b(\lambda+1))t-(\mu+\lambda). \tag{1.2.11}$$

We have that $p(0)=-(\mu+\lambda)<0$ and $p(1)=b\lambda>0$ for $\lambda>0$. It follows from the intermediate value theorem that p has roots in $(0,1)$. Denote the smallest root by δ. If

$\lambda = 0$, then $p(t) = (t-1)(\mu t^2 - (2\mu + b)t + \mu)$. Hence, we can choose the smallest root of p given by $\frac{2\mu+b-\sqrt{b^2+4\mu b}}{2\mu} \in (0,1)$ to be δ in this case. Note that in particular for Newton's method and secant method, respectively, we have that

$$p(t) = (t-1)(2t^2 - (b+4)t + 2)$$

and

$$p(t) = (t-2)(t^2 - (b+2)t + 1).$$

Hence, we obtain, respectively that

$$\delta = \frac{4}{b+4+\sqrt{b^2+8b}} \tag{1.2.12}$$

and

$$\delta = \frac{2}{b+2+\sqrt{b^2+4b}}. \tag{1.2.13}$$

Notice also that

$$p(t) \le 0 \text{ for each } t \in (-\infty, \delta]. \tag{1.2.14}$$

Next, we study the convergence of these sequences starting from $\{\delta_n\}$.

Lemma 1.2.1. *Let $\delta_1 > 0$, $\delta_2 > 0$ and $b \ge 1$ be given parameters. Suppose that*

$$0 < \delta_2 \le \delta_1 \le \delta, \tag{1.2.15}$$

where δ was defined in (1.2.11). *Let $\{\delta_n\}$ be the scalar sequence defined by* (1.2.10). *Then, the following assertions hold:*

($\mathcal{A}_1$) *If*

$$\delta_1 = \delta_2 \tag{1.2.16}$$

then,

$$\delta_n = \delta \text{ for each } n = 1, 2, 3, \cdots \tag{1.2.17}$$

($\mathcal{A}_2$) *If*

$$0 < \delta_2 < \delta_1 < \delta \tag{1.2.18}$$

then, sequence $\{\delta_n\}$ is decreasing and converges to 0.

Proof. It follows from (1.2.10) and $\delta_2 \le \delta_1$ that $\delta_3 > 0$. We shall show that

$$\delta_3 \le \delta_2. \tag{1.2.19}$$

In view of (1.2.10) for $n = 1$, it suffices to show that

$$p_1(\delta_2) = \mu(1-\delta_1)\delta_2^2 - (1-\delta_1)(2\mu+\lambda+b)\delta_2 - (\mu+(1+b)\lambda)\delta_1 + \mu + \lambda \ge 0. \tag{1.2.20}$$

The discriminant Δ of the quadratic polynomial p_1 is given by

$$\Delta = (1-\delta_1)\left[(1-\delta_1)(\lambda^2 + 2(2\mu+\lambda)b + b^2) + 4\mu\lambda b\delta_1\right] > 0. \tag{1.2.21}$$

Hence, p_1 has two distinct roots δ_s and δ_l with $\delta_s < \delta_l$. Polynomial p_1 is quadratic with respect to δ_2 and the leading coefficient $(\mu(1-\delta_1))$ is positive. Therefore, we have that

$$p_1(t) \geq 0 \text{ for each } t \in (-\infty, \delta_s] \cup [\delta_l, +\infty)$$

and

$$p_1(t) \leq 0 \text{ for each } t \in [\delta_s, \delta_l].$$

Then, (1.2.20) shall be true, if

$$\delta_2 \leq \delta_s. \tag{1.2.22}$$

By hypothesis (1.2.15) we have $\delta_1 \leq \delta_0$. Then by (1.2.14) we get that $p(\delta_1) \leq 0 \Rightarrow \delta_1 \leq \delta_s \Rightarrow$(1.2.22), since $\delta_2 \leq \delta_1$ by hypothesis (1.2.15). Hence, we showed (1.2.19). Therefore, relation

$$0 < \delta_{k+1} < \delta_k, \tag{1.2.23}$$

holds for $k = 2$. Then, we must show that

$$0 < \delta_{k+2} < \delta_{k+1}. \tag{1.2.24}$$

It follow from (1.2.10), $\delta_k < 1$ and $\delta_{k+1} < 1$ that $\delta_{k+2} > 0$. Then, in view of (1.2.10) the right hand side of (1.2.24) is true, if

$$\frac{b\delta_{k+1}\left[\lambda\delta_k + (1-\delta_k)\delta_{k+1}\right]}{(1-\delta_k)(1-\delta_{k+1})\left[\lambda + \mu(1-\delta_{k+1})\right]} \leq \delta_{k+1} \tag{1.2.25}$$

or

$$p(\delta_k) \leq 0, \tag{1.2.26}$$

which is true by (1.2.14) since $\delta_k \leq \delta_1 \leq \delta$. The induction for (1.2.23) is complete. If $\delta_1 = \delta_2 = \delta$, then it follows from (1.2.10) for $n = 1$ that $\delta_3 = \delta$ and $\delta_n = \delta$ for $n = 4, 5, \cdots$, which shows (1.2.17). If $\delta_2 < \delta_1$, the sequence $\{\delta_n\}$ is decreasing, bounded below by 0 and as such it converges to its unique largest lower bound denoted by γ. We then have from (1.2.10) that

$$\gamma = \frac{b\gamma[\lambda\gamma + (1-\gamma)\gamma]}{(1-\gamma)^2\left[\lambda + \mu(1-\gamma)\right]} \Rightarrow \gamma = \delta \text{ or } \gamma = 0. \tag{1.2.27}$$

But $\gamma \leq \delta_1 \leq \delta$. Hence, we conclude that $\gamma = 0$. ☒

Next, we present three results for the convergence of sequences $\{\alpha_n\}$, $\{\beta_n\}$ and $\{\gamma_n\}$ under conditions that are not all the same with the ones in Lemma 2.1 (see e.g. (1.2.28)).

Lemma 1.2.2. *Suppose that the hypothesis* (1.2.18) *is satified. Then, the sequence* $\{\gamma_n\}$ *is decreasingly convergent and sequences* $\{\alpha_n\}$ *and* $\{\beta_n\}$ *are increasingly convergent. Moreover, the following estimate holds:*

$$l_0 c < 1. \tag{1.2.28}$$

Proof. Using (1.2.2) and (1.2.9) we get that

$$\gamma_n = (1-\delta_n)\gamma_{n-1} = \cdots = (1-\delta_n)\cdots(1-\delta_1)\gamma_0 = (1-\delta_n)\cdots(1-\delta_1)\gamma_0 > 0.$$

In view of (1.2.18) we have in turn that

$$\begin{aligned}\delta_1 > 0 \quad &\Rightarrow \quad 1 - \frac{\gamma_1}{\gamma_0} > 0 \\ &\Rightarrow \quad \gamma_0 = \frac{1-(\mu+\lambda)L_0 c}{\mu+\lambda} > 0 \\ &\Rightarrow \quad \gamma_0 = \frac{1-l_0 c}{(\mu+\lambda)[1+(\mu+\lambda-1)l_0 c]} > 0 \\ &\Rightarrow \quad (2.28)\end{aligned}$$

and by the preceding equation we deduce that $\gamma_n > 0$ for each $n = 1,2,\ldots$ and

$$\gamma_n < \gamma_{n-1} \text{ for each } n = 1,2,\ldots,$$

since $\delta_n < 1$. Hence, sequence $\{\gamma_n\}$ converges to its unique largest lower bound denoted by γ^*. We also have that $\beta_n = \frac{1}{\mu+\lambda} - \gamma_n < \frac{1}{\mu+\lambda}$. Thus, the sequence $\{\beta_n\}$ is increasing, bounded from above by $\frac{1}{\mu+\lambda}$ and as such it converges to its unique least upper bound denoted by β^*. Then, in view of (1.2.5) sequence $\{\alpha_n\}$ is also increasing, bounded from above by $\frac{L_0^{-1}}{\mu+\lambda}$ and such it also converges to its unique least upper bound denoted by α^*. ⊠

Lemma 1.2.3. *Suppose that* (1.2.15) *and* (1.2.16) *are satisfied. Then, the following assertions hold for each* $n = 1,2,\cdots$

$$\delta_n = \delta$$

$$\gamma_n = (1-\delta)^n \gamma_0, \; \gamma^* = \lim_{n\to\infty} \gamma_n = 0,$$

$$\beta_n = \frac{1}{\mu+\lambda} - (1-\delta)^n \gamma_0, \; \beta^* = \lim_{n\to\infty} \beta_n = \frac{1}{\mu+\lambda}$$

and

$$\alpha_n = \frac{1}{L_0}\left[\frac{1}{\mu+\lambda} - (1-\delta)^n \gamma_0\right], \; \alpha^* = \lim_{n\to\infty} \alpha_n = \frac{1}{L_0(\mu+\lambda)}$$

Corollary 1.2.4. *Suppose that the hypotheses of Lemma 2.1 and Lemma 2.2 hold. Then, sequence* $\{\alpha_n\}$ *defined in* (1.2.1) *is nondecreasing and converges to*

$$\alpha^* = \beta^* \frac{1+(\mu+\lambda-1)l_0 c}{l_0}.$$

Next, we present lower and upper bounds on the limit point α^*.

Lemma 1.2.5. *Suppose that the condition* (1.2.18) *is satisfied. Then, the following assertion holds*

$$b_1^1 \leq \alpha^* \leq b_2^1, \tag{1.2.29}$$

where

$$b_1^1 = \frac{1+(\mu+\lambda-1)l_0 c}{l_0}\left[\frac{1}{\mu+\lambda} - \exp\left(-2\left(\frac{\delta_1}{2-\delta_1} + \frac{\delta_2}{2-\delta_2}\right)\right)\right],$$

$$b_2^1 = \frac{1+(\mu+\lambda-1)l_0c}{l_0}\left[\frac{1}{\mu+\lambda} - \exp(\delta^*)\right], \tag{1.2.30}$$

$$\delta^* = -\left[\frac{1}{1-\delta_1}\left(\delta_1 + \frac{\delta_2}{1-r}\right) + \ln\left(\frac{(\mu+\lambda)(1-(\mu+\lambda-1)l_0c)}{1-l_0c}\right)\right]$$

and

$$r = b\frac{\lambda\delta_1 + \delta_2(1-\delta_1)}{(1-\delta_1)(1-\delta_2)(\lambda+\mu(1-\delta_2))}.$$

Proof. Using (1.2.18) and (1.2.28) we have that $0 < \delta_3 < \delta_2 < \delta_1$. Let us assume that $0 < \delta_{k+1} < \delta_k < \cdots < \delta_1$. Then, it follows from the induction hypotheses and (1.2.34) that

$$\delta_{k+2} = \delta_{k+1} b \frac{\delta_k + \delta_{k+1}(1-\delta_k)}{(1-\delta_k)(1-\delta_{k+1})(2-\delta_{k+1})} < r\delta_{k+1} < r^2\delta_k \leq \cdots \leq r^{k-1}\delta_3 \leq r^k\delta_2.$$

We have that

$$\gamma^* = \lim_{n\to\infty}\gamma_n = \prod_{i=1}^{\infty}(1-\delta_n)\gamma_0.$$

This is equivalent to

$$\ln\left(\frac{1}{\gamma^*}\right) = \sum_{n=1}^{\infty}\ln\left(\frac{1}{1-\delta_n}\right) + \ln\left(\frac{(\mu+\lambda)(1+(\mu+\lambda-1)l_0c)}{1-l_0c}\right),$$

recalling that $\gamma_0 = (1-l_0c)/((\mu+\lambda)(1+(\mu+\lambda-1)l_0c))$. We shall use the following bounds for $\ln t$, $t > 1$:

$$2\left(\frac{t-1}{t+1}\right) \leq \ln t \leq \frac{t^2-1}{2t}.$$

First, we shall find an upper bound for $\ln(1/\gamma^*)$. We have that

$$\begin{aligned}
\ln(1/\gamma^*) &\leq \sum_{n=1}^{\infty}\frac{\delta_n(2-\delta_n)}{2(1-\delta_n)} + \ln\left(\frac{(\mu+\lambda)(1+(\mu+\lambda-1)l_0c)}{1-l_0c}\right)\\
&\leq \tfrac{1}{1-\delta_1}\sum_{n=1}^{\infty}\delta_n + \ln\left(\frac{(\mu+\lambda)(1+(\mu+\lambda-1)l_0c)}{1-l_0c}\right)\\
&\leq \tfrac{1}{1-\delta_1}(\delta_1+\delta_2+\delta_3+\cdots) + \ln\left(\tfrac{(\mu+\lambda)(1+(\mu+\lambda-1)l_0c)}{1-l_0c}\right)\\
&\leq \tfrac{1}{1-\delta_1}(\delta_1+\delta_2+r\delta_2+\cdots+r^n\delta_2+\cdots) + \ln\left(\tfrac{(\mu+\lambda)(1+(\mu+\lambda-1)l_0c)}{1-l_0c}\right)\\
&\leq \tfrac{1}{1-\delta_1}\left(\delta_1+\delta_2(r+r^2+\cdots+r^n+\cdots\right) + \ln\left(\tfrac{(\mu+\lambda)(1+(\mu+\lambda-1)l_0c)}{1-l_0c}\right)\\
&\leq \tfrac{1}{1-\delta_1}\left(\delta_1+\tfrac{\delta_2}{1-r}\right) + \ln\left(\tfrac{(\mu+\lambda)(1+(\mu+\lambda-1)l_0c)}{1-l_0c}\right) = -\delta^*.
\end{aligned}$$

As $\beta^* = 1/(\mu+\lambda) - \gamma^*$ and $\alpha^* = L_0^{-1}\beta^*$, we obtain the upper bound in (1.2.33). Moreover, in order to obtain the lower bound for $\ln(1/\gamma^*)$, we have that

$$\ln(1/\gamma^*) \geq 2\sum_{n=1}^{\infty}\frac{\delta_n}{2-\delta_n} > 2\left(\frac{\delta_1}{2-\delta_1} + \frac{\delta_2}{2-\delta_2}\right),$$

which implies the lower bound in (1.2.33). ☒

From now on we shall denote by (C^1) the hypothesis of Lemma 2.1 and Lemma 2.2.

Remark 1.2.6. **(a)** *Let us introduce the notation*

$$c^N = \alpha_{N-1} - \alpha_{N-2},\ \nu^N = \alpha_N - \alpha_{N-1}$$

for some integer $N \geq 1$. Notice that $c^1 = \alpha_0 - \alpha_{-1} = c$ and $\nu^1 = \alpha_1 - \alpha_0 = \nu$. The results in the preceding Lemmas can be weakened even further as follows. Consider the convergence criteria (C_*^N) for $N > 1$: (C^1) with c, ν replaced by c^N, ν^N, respectively

$$\alpha_{-1} < \alpha_0 < \alpha_1 < \cdots < \alpha_N < \alpha_{N+1},$$

$$l_0 \left[\mu(\alpha_{N+1} - c^N) + \lambda(\alpha_N - c^N) + c^N\right] < 1.$$

Then, the preceding results hold with $c, \nu, \delta_1, \delta_2, b_1^1, b_2^1$ replaced, respectively by $c^N, \nu^N, \delta_N, \delta_{N+1}, b_1^N, b_2^N$.

(b) *Notice that if*

$$l_0 \left[\mu(\alpha_{n+1} - c) + \lambda(\alpha_n - c) + c\right] < 1 \text{ holds for each } n = 0, 1, 2, \cdots, \tag{1.2.31}$$

then, it follows from (1.2.1) *that sequence $\{\alpha_n\}$ is increasing, bounded from above by $\frac{1+(\mu+\lambda-1)l_0 c}{l_0(\mu+\lambda)}$ and as such it converges to its unique least upper bound α^*. Criterion* (1.2.31) *is the weakest of all the preceding convergence criteria for sequence $\{\alpha_n\}$. Clearly all the preceding criteria imply* (1.2.31)*. Finally, define the criteria for $N \geq 1$*

$$(I^N) = \begin{cases} (C_*^N) \\ (1.2.31) \text{ if criteria } (C_*^N) \text{ fail.} \end{cases} \tag{1.2.32}$$

Lemma 1.2.7. *Suppose that the conditions* (1.2.18) *and* (1.2.28) *hold. Then, the following assertion holds*

$$b_1^1 \leq \alpha^* \leq b_2^1, \tag{1.2.33}$$

where

$$b_1^1 = \frac{1+(\mu+\lambda-1)l_0 c}{l_0}\left[\frac{1}{\mu+\lambda} - \exp\left(-2\left(\frac{\delta_1}{2-\delta_1} + \frac{\delta_2}{2-\delta_2}\right)\right)\right],$$

$$b_2^1 = \frac{1+(\mu+\lambda-1)l_0 c}{l_0}\left[\frac{1}{\mu+\lambda} - \exp(\delta^*)\right], \tag{1.2.34}$$

$$\delta^* = -\left[\frac{1}{1-\delta_1}\left(\delta_1 + \frac{\delta_2}{1-r}\right) + \ln\left(\frac{(\mu+\lambda)(1-(\mu+\lambda-1)l_0 c)}{1-l_0 c}\right)\right]$$

and

$$r = b\frac{\lambda\delta_1 + \delta_2(1-\delta_1)}{(1-\delta_1)(1-\delta_2)(\lambda+\mu(1-\delta_2))}.$$

Proof. Using (1.2.18) and (1.2.28) we have that $0 < \delta_3 < \delta_2 < \delta_1$. Let us assume that $0 < \delta_{k+1} < \delta_k < \cdots < \delta_1$. Then, it follows from the induction hypotheses and (1.2.34) that

$$\delta_{k+2} = \delta_{k+1} b \frac{\delta_k + \delta_{k+1}(1-\delta_k)}{(1-\delta_k)(1-\delta_{k+1})(2-\delta_{k+1})} < r\delta_{k+1} < r^2\delta_k \leq \cdots \leq r^{k-1}\delta_3 \leq r^k\delta_2.$$

We have that

$$\gamma^* = \lim_{n\to\infty}\gamma_n = \prod_{i=1}^{\infty}(1-\delta_n)\gamma_0.$$

This is equivalent to

$$\ln\left(\frac{1}{\gamma^*}\right) = \sum_{n=1}^{\infty}\ln\left(\frac{1}{1-\delta_n}\right) + \ln\left(\frac{(\mu+\lambda)(1+(\mu+\lambda-1)l_0c)}{1-l_0c}\right),$$

recalling that $\gamma_0 = (1-l_0c)/((\mu+\lambda)(1+(\mu+\lambda-1)l_0c))$. We shall use the following bounds for $\ln t$, $t>1$:

$$2\left(\frac{t-1}{t+1}\right) \le \ln t \le \frac{t^2-1}{2t}.$$

First, we shall find an upper bound for $\ln(1/\gamma^*)$. We have that

$$\begin{aligned}
\ln(1/\gamma^*) &\le \sum_{n=1}^{\infty}\frac{\delta_n(2-\delta_n)}{2(1-\delta_n)} + \ln\left(\frac{(\mu+\lambda)(1+(\mu+\lambda-1)l_0c)}{1-l_0c}\right)\\
&\le \tfrac{1}{1-\delta_1}\sum_{n=1}^{\infty}\delta_n + \ln\left(\frac{(\mu+\lambda)(1+(\mu+\lambda-1)l_0c)}{1-l_0c}\right)\\
&\le \tfrac{1}{1-\delta_1}(\delta_1+\delta_2+\delta_3+\cdots) + \ln\left(\tfrac{(\mu+\lambda)(1+(\mu+\lambda-1)l_0c)}{1-l_0c}\right)\\
&\le \tfrac{1}{1-\delta_1}(\delta_1+\delta_2+r\delta_2+\cdots+r^n\delta_2+\cdots) + \ln\left(\tfrac{(\mu+\lambda)(1+(\mu+\lambda-1)l_0c)}{1-l_0c}\right)\\
&\le \tfrac{1}{1-\delta_1}\left(\delta_1+\delta_2(r+r^2+\cdots+r^n+\cdots\right) + \ln\left(\tfrac{(\mu+\lambda)(1+(\mu+\lambda-1)l_0c)}{1-l_0c}\right)\\
&\le \tfrac{1}{1-\delta_1}\left(\delta_1+\tfrac{\delta_2}{1-r}\right) + \ln\left(\tfrac{(\mu+\lambda)(1+(\mu+\lambda-1)l_0c)}{1-l_0c}\right) = -\delta^*.
\end{aligned}$$

As $\beta^* = 1/(\mu+\lambda)-\gamma^*$ and $\alpha^* = L_0^{-1}\beta^*$, we obtain the upper bound in (1.2.33). Moreover, in order to obtain the lower bound for $\ln(1/\gamma^*)$, we have that

$$\ln(1/\gamma^*) \ge 2\sum_{n=1}^{\infty}\frac{\delta_n}{2-\delta_n} > 2\left(\frac{\delta_1}{2-\delta_1}+\frac{\delta_2}{2-\delta_2}\right),$$

which implies the lower bound in (1.2.33). ⊠

1.3. Semilocal Convergence of the Secant-Type Method

In this section, we first present the semilocal convergence of the secant-type method using $\{\alpha_n\}$ (defined in (1.2.1)) as a majorizing sequence. Let $U(x,R)$ stand for an open ball centered at $x\in X$ with radius $R>0$. Let $\overline{U}(x,R)$ denote its closure. We shall study the secant method for triplets $(\mathcal{F},x_{-1},x_0)$ belonging to the class $\mathcal{K}=\mathcal{K}(l_0,l,\nu,c,\lambda,\mu)$ defined as follows.

Definition 1.3.1. *Let l_0,l,ν,c,λ,μ be constants satisfying the hypotheses (I^N) for some fixed integer $N\ge 1$. A triplet $(\mathcal{F},x_{-1},x_0)$ belongs to the class $\mathcal{K}=\mathcal{K}(l_0,l,\nu,c,\lambda,\mu)$ if:*

$(\mathcal{D}_1)$ *$\mathcal{F}$ is a nonlinear operator defined on a convex subset D of a Banach space $\mathcal{X}$ with values in a Banach space $\mathcal{Y}$.*

($\mathcal{D}_2$) *x_{-1} and x_0 are two points belonging to the interior D^0 of D and satisfying the inequality*

$$\|x_0 - x_{-1}\| \leq c.$$

($\mathcal{D}_3$) *There exists a sequence $\{\theta_n\}$ of real numbers and λ, μ such that $|1-\theta_n| \leq \lambda$ and $1+|\theta_n| \leq \mu$ for each $n = 0,1,2,\cdots$.*

($\mathcal{D}_4$) *$\mathcal{F}$ is Fréchet-differentiable on D^0 and there exists an operator $\delta\mathcal{F} : \mathcal{D}^0 \times D^0 \to Ł(X,Y)$ such that $\mathcal{A}^{-1} = \delta\mathcal{F}(x_0,y_0)^{-1} \in Ł(Y,X)$ for all $x,y,z \in D$ then, the following hold*

$$\|\mathcal{A}^{-1}\mathcal{F}(x_0)\| \leq \nu,$$

$$\|\mathcal{A}^{-1}(\delta\mathcal{F}(x,y) - \mathcal{F}'(z))\| \leq l(\|x-z\| + \|y-z\|)$$

and

$$\|\mathcal{A}^{-1}(\delta\mathcal{F}(x,y) - \mathcal{F}'(x_0))\| \leq l_0(\|x-x_0\| + \|y-x_0\|),$$

where $y_0 = \theta_0 x_0 + (1-\theta_0)x_{-1}$.

($\mathcal{D}_5$)

$$\overline{U}(x_0, \alpha_0^*) \subseteq D_c = \{x \in D : \mathcal{F} \text{ is continuous at } x\} \subseteq D,$$

where $\alpha_0^* = (\mu + \lambda - 1)(\alpha^* - c)$ and α^* is given in Lemma 2.3.

Next, we present the semilocal convergence result for the secant method.

Theorem 1.3.2. *If $(\mathcal{F}, x_{-1}, x_0) \in \mathcal{K}(l_0, l, \nu, c, \lambda, \mu)$ then, the sequence $\{x_n\}$ ($n \geq -1$) generated by the secant-type method is well defined, remains in $\overline{U}(x_0, \alpha_0^*)$ for each $n = 0,1,2,\cdots$ and converges to a unique solution $x^* \in \overline{U}(x_0, \alpha^* - c)$ of* (1.1.1). *Moreover, the following assertions hold for each $n = 0,1,2,\cdots$*

$$\|x_n - x_{n-1}\| \leq \alpha_n - \alpha_{n-1} \tag{1.3.1}$$

and

$$\|x^* - x_n\| \leq \alpha^* - \alpha_n, \tag{1.3.2}$$

where sequence $\{\alpha_n\}$ ($n \geq 0$) is given in (1.2.1). *Furthermore, if there exists R such that*

$$\overline{U}(x_0, R) \subseteq D,\ R \geq \alpha^* - c \text{ and } l_0(\alpha^* - c + R) + \|\mathcal{A}^{-1}(\mathcal{F}^{-1}(x_0) - \mathcal{A})\| < 1, \tag{1.3.3}$$

then, the solution x^* is unique in $\overline{U}(x_0, R)$.

Proof. First, we show that $\mathcal{M} = \delta\mathcal{F}(x_{k+1}, y_{k+1})$ is invertible for $x_{k+1}, y_{k+1} \in \overline{U}(x_0, \alpha_0^*)$. By ($\mathcal{D}_2$),($\mathcal{D}_3$) and ($\mathcal{D}_4$), we have that

$$\|y_{k+1} - x_0\| \leq \|\theta_k(x_{k+1} - x_0) + (1-\theta_{k+1})(x_k - x_0)\|$$

$$\leq |\theta_{k+1}|\|x_{k+1} - x_0\| + |1-\theta_{k+1}|\|x_k - x_0\| \leq (\mu - 1)(\alpha^* - c) + \lambda(\alpha^* - c) = \alpha_0^*$$

and

$$\begin{aligned}\|I-\mathcal{A}^{-1}\mathcal{M}\| &= \|\mathcal{A}^{-1}(\mathcal{M}-\mathcal{A})\| \\ &\le \|\mathcal{A}^{-1}(\mathcal{M}-\mathcal{F}'(x_0))\|+\|\mathcal{A}^{-1}(\mathcal{F}'(x_0)-\mathcal{A})\| \\ &\le l_0(\|x_{k+1}-x_0\|+\|y_{k+1}-x_0\|+\|x_0-x_{-1}\|) \\ &\le l_0(\|x_{k+1}-x_0\|+|\theta_{k+1}|\|x_{k+1}-x_0\|+|1+\theta_{k+1}|\|x_{k+1}-x_0\|+c) \\ &\le l_0(\mu(\alpha_{k+1}-c)+\lambda(\alpha_{k+1}-c)+c)<1\end{aligned} \tag{1.3.4}$$

Using the Banach Lemma on invertible operators [9], [10], [15], [18], [20] and (1.3.4), we deduce that $\mathcal{M}$ is invertible and

$$\|\mathcal{M}^{-1}\mathcal{A}\| \le (1-l_0(\mu(\alpha_{k+1}-c)+\lambda(\alpha_{k+1}-c)+c))^{-1}. \tag{1.3.5}$$

By $(\mathcal{D}_4)$, we have

$$\|\mathcal{A}^{-1}(\mathcal{F}'(u)-\mathcal{F}'(v))\| \le 2l\|u-v\|,\ u,v\in D^0. \tag{1.3.6}$$

We can write the identity

$$\mathcal{F}'(x)-\mathcal{F}'(y)=\int_0^1 \mathcal{F}'(y+t(x-y))dt(x-y). \tag{1.3.7}$$

Then, for all $x,y,u,v\in D^0$, we obtain

$$\|\mathcal{A}^{-1}(\mathcal{F}(x)-\mathcal{F}(y)-\mathcal{F}'(u)(x-y))\| \le l(\|x-u\|+\|y-u\|)\|x-y\| \tag{1.3.8}$$

and

$$\|\mathcal{A}^{-1}(\mathcal{F}(x)-\mathcal{F}(y)-\delta\mathcal{F}(u,v)(x-y))\| \le l(\|x-v\|+\|y-v\|+\|u-v\|)\|x-y\|. \tag{1.3.9}$$

By a continuity argument (1.3.6)-(1.3.9) remain valid if x and/or y belong to D_c. Next, we show (1.3.1). If (1.3.1) holds for all $n\le k$ and if $\{x_n\}$ $(n\ge 0)$ is well defined for $n=0,1,2,\cdots,k$, then

$$\|x_n-x_0\| \le \alpha_n-\alpha_0<\alpha^*-\alpha_0,\ n\le k. \tag{1.3.10}$$

That is (1.1.2) is well defined for $n=k+1$. For $n=-1$ and $n=0$, (1.3.1) reduces to $\|x_{-1}-x_0\|\le c$ and $\|x_0-x_1\|\le\nu$. Suppose (1.3.1) holds for $n=-1,0,1,\cdots,k$ $(k\ge 0)$. By (1.3.5), (1.3.9), and

$$\mathcal{F}(x_{k+1})=\mathcal{F}(x_{k+1})-\mathcal{F}(x_k)-\mathcal{A}_k(x_{k+1}x_k) \tag{1.3.11}$$

we obtain in turn the following estimates

$$\begin{aligned}&\|\mathcal{A}^{-1}\mathcal{F}(x_{k+1})\|=\|\mathcal{A}^{-1}(\delta\mathcal{F}(x_{k+1},x_k)-\mathcal{A}_k)(x_{k+1}-x_k)\| \\ &\le \left(\|\mathcal{A}^{-1}(\delta\mathcal{F}(x_{k+1},x_k)-\mathcal{F}'(x_k))\|+\|\mathcal{A}^{-1}(\mathcal{F}'(x_k)-\mathcal{A}_k)\|\right)\|(x_{k+1}-x_k)\| \\ &\le l\,[\|(x_{k+1}-x_k)\|+\|(x_k-y_k)\|]\,\|(x_{k+1}-x_k)\| \\ &\le l(\alpha_{k+1}-\alpha_k+|1-\theta_k|(\alpha_k-\alpha_{k-1})(\alpha_{k+1}-\alpha_k))\end{aligned} \tag{1.3.12}$$

and

$$\begin{aligned}\|x_{k+2}-x_{k+1}\| &= \|\mathcal{A}_{k+1}^{-1}\mathcal{F}(x_{k+1})\|\\ &\le \|\mathcal{A}_{k+1}^{-1}\mathcal{A}\|\|\mathcal{A}^{-1}\mathcal{F}(x_{k+1})\|\\ &\le \frac{l(\alpha_{k+1}-\alpha_k+|1-\theta_k|(\alpha_k-\alpha_{k-1}))}{1-l_0[(1+|\theta_{k+1}|)(\alpha_{k+1}-c)+|1-\theta_{k+1}|(\alpha_k-c)+c]}(\alpha_{k+1}-\alpha_k)\\ &\le \alpha_{k+2}-\alpha_{k+1}.\end{aligned}$$

The induction for (1.3.1) is complete. It follows from (1.3.1) and Lemma 2.1 that $\{x_n\}$ $(n\ge -1)$ is a complete sequence in a Banach space $\mathcal{X}$ and as such it converges to some $x^*\in \overline{U}(x_0,\alpha^*-c)$ (since $\overline{U}(x_0,\alpha^*-c)$ is a closed set). By letting $k\to\infty$ in (1.3.12), we obtain $\mathcal{F}(x^*)=0$. Moreover, estimate (1.3.2) follows from (1.3.1) by using standard majoration techniques [8, 12, 14]. Finally, to show the uniqueness in $\overline{U}(x_0,R)$, let $y^*\in\overline{U}(x_0,R)$ be a solution (1.1.1). Set

$$\mathcal{T}=\int_0^1\mathcal{F}'(y^*+t(y^*-x^*))dt$$

Using $(\mathcal{D}_4)$ and (1.3.3) we get in turn that

$$\begin{aligned}\|\mathcal{A}^{-1}(\mathcal{A}-\mathcal{T})\| &= l_0(\|y^*-x_0\|+\|x^*-x_0\|)+\|\mathcal{A}^{-1}(\mathcal{F}'(x_0)-\mathcal{A})\|\\ &\le l_0[(\alpha^*-\alpha_0)+R]+\|\mathcal{A}^{-1}(\mathcal{F}'(x_0)-\mathcal{A})\|<1.\end{aligned} \tag{1.3.13}$$

If follows from (1.3.13) and the Banach lemma on invertible operators that $\mathcal{T}^{-1}$ exists. Using the identity:

$$\mathcal{F}(x^*)-\mathcal{F}'(y^*)=\mathcal{T}(x^*-y^*), \tag{1.3.14}$$

we deduce that $x^*=y^*$. ⊠

Remark 1.3.3. *If follows from the proof of Theorem 3.2 that sequences $\{r_n\}$, $\{s_n\}$ defined by*

$$\begin{cases} r_{-1}=0,\ r_0=c,\ r_1=c+\nu\\ r_2=r_1+\frac{l_0(r_1-r_0+|1-\theta_0|(r_0-r_{-1}))(r_1-r_0)}{1-l_0((1+|\theta_1|)(r_1-r_0))}\\ r_{n+2}=r_{n+1}+\frac{l(r_{n+1}-r_n+|1-\theta_n|(r_n-r_{n-1}))(r_{n+1}-r_n)}{1-l_0[(1+|\theta_{n+1}|)(r_{n+1}-r_0)+(|1-\theta_{n+1}|)(r_n-r_0)+c]}\end{cases} \tag{1.3.15}$$

and

$$\begin{cases} s_{-1}=0,\ s_0=c,\ s_1=c+\nu\\ s_2=s_1+\frac{l_0(s_1-s_0+\lambda(s_0-s_{-1}))(s_1-s_0)}{1-l_0(1+|\theta_1|)(s_1-s_0)}\\ s_{n+2}=s_{n+1}+\frac{l(s_{n+1}-s_n+\lambda(s_n-s_{n-1}))(s_{n+1}-s_n)}{1-l_0(\mu(s_{n+1}-s_0)+\lambda(s_n-s_0))+c}\end{cases} \tag{1.3.16}$$

respectively are more precise majorizing sequences for $\{x_n\}$. Clearly, these sequences also converge under the (I^N) hypotheses.

A simple inductive argument shows that if $l_0<l$ for each $n=2,3,\cdots$

$$r_n<s_n<\alpha_n \tag{1.3.17}$$

$$r_{n+1}-r_n<s_{n+1}-s_n<\alpha_{n+1}-\alpha_n \tag{1.3.18}$$

and

$$r^*=\lim_{n\to\infty}r_n\le s^*=\lim_{n\to\infty}s_n\le\alpha^*=\lim_{n\to\infty}\alpha_n. \tag{1.3.19}$$

In practice, one must choose $\{\theta_n\}$ so that the best error bounds are obtained (see also Section 4). Note also that sequences $\{r_n\}$ or $\{s_n\}$ may converge under even weaker hypotheses. The sufficient convergence criterion (1.2.15) determines the smallness of c and r. This criterion can be solved for c and r (see for example the h criteria or (1.3.29) that follow). Indeed, let us demonstrate the advantages in two popular cases:

Case 1. Newton's method. (i. e., if $c = 0, \lambda = 0, \mu = 1$). Then, it can easily be seen that $\{s_n\}$ (and consequently $\{r_n\}$) converges provided that (see also [3])

$$h_2 = l_2\nu \leq 1, \tag{1.3.20}$$

where

$$l_2 = \frac{1}{4}\left(4\kappa_0 + \sqrt{\kappa_0\kappa} + \sqrt{\kappa_0\kappa + 8\kappa_0^2}\right), \tag{1.3.21}$$

whereas sequence $\{x_n\}$ converges, if

$$h_1 = l_1\nu \leq 1 \tag{1.3.22}$$

where

$$l_1 = \frac{1}{4}\left(4\kappa_0 + \kappa + \sqrt{\kappa_0^2 + 8\kappa\kappa_0}\right), \tag{1.3.23}$$

In the case $\kappa_0 = \kappa$ (i. e. $b = 1$), we obtain the famous for its simplicity and clarity Kantorovich sufficient convergent criteria [2] given by

$$h = 2\kappa\nu \leq 1. \tag{1.3.24}$$

Notice however that

$$h \leq 1 \Rightarrow h_1 \leq 1 \Rightarrow h_2 \leq 1 \tag{1.3.25}$$

but not necessarily vice versa unless if $\kappa_0 = \kappa$. Moreover, we have that

$$\frac{h_1}{h} \to \frac{1}{4},\ \frac{h_1}{h} \to 0,\ \frac{h_2}{h_1} \to 0 \text{ as } \frac{\kappa_0}{\kappa} \to 0 \tag{1.3.26}$$

Case 2. Secant method. (i. e. for $\theta_n = 0$). Schmidt [20], Potra-Ptáck [15], Dennis [8], Ezquerro el at. [9], used the majorizing sequence $\{\alpha_n\}$ for $\theta_n \in [0,1]$ and $l_0 = l$. That is, they used the sequence $\{t_n\}$ given by

$$\begin{cases} t_{-1} = 0,\ t_0 = c,\ t_1 = c + \nu \\ t_{n+2} = t_{n+1} + \frac{l(t_{n+1} - t_{n-1})(t_{n+1} - t_n)}{1 - l(t_n - t_{n+1} + c)} \end{cases} \tag{1.3.27}$$

whereas our sequence $\{\alpha_n\}$ reduces to

$$\begin{cases} \alpha_{-1} = 0,\ \alpha_0 = c,\ \alpha_1 = c + \nu \\ \alpha_{n+2} = \alpha_{n+1} + \frac{l(\alpha_{n+1} - \alpha_{n-1})(\alpha_{n+1} - \alpha_n)}{1 - l_0(\alpha_{n+1} - \alpha_n + c)} \end{cases} \tag{1.3.28}$$

Then, in case $l_0 < l$ our sequence is more precise (see also (1.3.17)-(1.3.19)). Notice also that in the preceding references the sufficient convergence criterion associated to $\{t_n\}$ is given by

$$lc + 2\sqrt{l\nu} \leq 1 \tag{1.3.29}$$

Our sufficient convergence criteria can be also weaker in this case (see also the numerical examples). It is worth nothing that if $c = 0$ (1.3.29) reduces to (1.3.24) (since $\kappa = 2l$).

Similar observations can be made for other choices of parameters.

1.4. Local Convergence of the Secant-Type Method

In this section, we present the local convergence analysis of the secant-type method. Let $x^* \in \mathcal{X}$ be such that $\mathcal{F}(x^*) = 0$ and $\mathcal{F}'(x^*)^{-1} \in Ł(\mathcal{Y}, \mathcal{X})$. Using the identities

$$x_{n+1} - x^* = (\mathcal{A}_n^{-1}\mathcal{F}'(x^*))F'(x^*)^{-1}\left[(\delta\mathcal{F}(x_n, y_n) - F'(x_n)) + (F'(x_n) - \delta\mathcal{F}(x_n, x^*))\right](x_n - x^*),$$

$$y_n - x_n = (1 - \theta_n)(x_{n-1} - x_n),$$

and

$$y_n - x^* = \theta_n(x_n - x^*) + (1 - \theta_n)(x_{n-1} - x^*)$$

we easily arrive at:

Theorem 1.4.1. *Suppose that $(\mathcal{D}_1)$ and $(\mathcal{D}_3)$ hold. Moreover, suppose that there exist $x^* \in D, K_0 > 0, K > 0$ such that $\mathcal{F}(x^*) = 0$, $F'(x^*)^{-1} \in Ł(\mathcal{Y}, \mathcal{X})$,*

$$\|F'(x^*)^{-1}(\delta\mathcal{F}(x, y) - F'(x^*))\| \leq K_0(\|x - x^*\| + \|y - x^*\|)$$

$$\|F'(x^*)^{-1}(\delta\mathcal{F}(x, y) - F'(z))\| \leq K(\|x - z\| + \|y - z\|) \ \text{for each } x, y, z \in D,$$

and

$$\overline{U}(x^*, R_0^*) \subseteq D,$$

where

$$R^* = \frac{1}{(2\lambda + 1)K + (\lambda + \mu)K_0}$$

and

$$R_0^* = (\mu + \lambda - 1)R^*.$$

Then, sequence $\{x_n\}$ generated by the secant-type method is well defined, remains in $\overline{U}(x^*, R^*)$ for each $n = -1, 0, 1, 2, \cdots$ and converges to x^* provided that $x_{-1}, x_0 \in U(x^*, R^*)$. Moreover, the following estimates hold

$$\|x_{n+1} - x^*\| \leq \hat{e_n}\|x_n - x^*\| \leq \overline{e_n}\|x_n - x^*\| \leq e_n\|x_n - x^*\|,$$

where

$$\hat{e_n} = \frac{\overline{K}(\|x_n - x^*\| + |1 - \theta_n|\|x_{n-1} - x_n\|}{1 - K_0([(1 + |\theta_n|)\|x_n - x^*\| + |1 - \theta_n|\|x_{n-1} - x^*\|]}$$

$$\overline{e_n} = \frac{\overline{K}(\|x_n - x^*\| + \lambda\|x_{n-1} - x_n\|}{1 - K_0([(\mu\|x_n - x^*\| + \lambda\|x_{n-1} - x^*\|]}$$

$$e_n = \frac{\overline{K}(2\lambda+1)R^*}{1-K_0(\lambda+\mu)R^*}$$

and

$$\overline{K} = \begin{cases} \kappa_0, \ if\, n=0 \\ \kappa, \ if\, n>0 \end{cases}$$

Remark 1.4.2. *Comments similar to the one given in Remark 3.3 can also follow for this case. For example, notice again that in the case of Newton's method*

$$R_* = \frac{2}{2\kappa_0+\kappa},$$

whereas the convergence ball given independently by Rheinboldt [18] and Traub [19] is given by

$$R_*^1 = \frac{2}{3\kappa}.$$

Note that

$$R_*^1 \le R_*.$$

Strict inequality holds in the preceding inequality if $\kappa_0 < \kappa$. *Moreover, the error bounds are tighter, if* $\kappa_0 < \kappa$. *Finally, note that* $\frac{\kappa_0}{\kappa}$ *can be arbitrarily small and*

$$\frac{R_*}{R_*^1} \to 3 \ as \ \frac{\kappa_0}{\kappa} \to 0.$$

1.5. Numerical Examples

Related to the semilocal case we present the following examples.

Example 1.5.1. *Let* $\mathcal{X} = \mathcal{Y} = \mathbb{R}$ *and let consider the following function*

$$x^3 - 0.49 = 0, \tag{1.5.1}$$

and we are going to apply the secant method ($\lambda = 1$, $\mu = 1$, $\theta_n = 0$*) to find the solution of* (1.5.1). *We take the starting points* $x_{-1} = 1.14216\cdots$, $x_0 = 1$ *and we consider the domain* $\Omega = B(x_0, 2)$. *In this case, we obtain*

$$\nu = 0.147967\cdots, \tag{1.5.2}$$

$$\nu = 0.14216\cdots, \tag{1.5.3}$$

$$l = 2.61119\cdots, \tag{1.5.4}$$

$$l_0 = 1.74079\cdots. \tag{1.5.5}$$

Notice that hypothesis $lc + 2\sqrt{l\nu} \le 1$ *is not satisfied, but hypotheses of Theorem 3.2 are satisfied, so the convergence of secant method starting form* $x_0 \in B(x_0, 2)$ *converges to the solution of* (1.5.1).

Example 1.5.2. *Let $\mathcal{X} = \mathcal{Y} = \mathcal{C}[0,1]$, equipped with the max-norm. Consider the following nonlinear boundary value problem*

$$\begin{cases} u'' = -u^3 - \gamma\, u^2 \\ u(0) = 0, \quad u(1) = 1. \end{cases}$$

It is well known that this problem can be formulated as the integral equation

$$u(s) = s + \int_0^1 \mathcal{Q}(s,t)\, (u^3(t) + \gamma\, u^2(t))\, dt \tag{1.5.6}$$

where, $\mathcal{Q}$ is the Green's function:

$$\mathcal{Q}(s,t) = \begin{cases} t\,(1-s), & t \leq s \\ s\,(1-t), & s < t. \end{cases}$$

We observe that

$$\max_{0 \leq s \leq 1} \int_0^1 |\mathcal{Q}(s,t)|\, dt = \frac{1}{8}.$$

Then problem (1.5.6) is in the form (1.1.1), where, $F : \mathcal{D} \longrightarrow \mathcal{Y}$ is defined as

$$[F(x)]\,(s) = x(s) - s - \int_0^1 \mathcal{Q}(s,t)\, (x^3(t) + \gamma\, x^2(t))\, dt.$$

We define the divided difference by $\delta F(x,y) = \displaystyle\int_0^1 F'(y + t(x-y))dt$. Set $u_0(s) = s$ and $\mathcal{D} = U(u_0, R_0)$. It is easy to verify that $U(u_0,R_0) \subset U(0, R_0+1)$ since $\| u_0 \| = 1$. If $2\,\gamma < 5$, the operator F' satisfies conditions of Theorem 3.2, with

$$\theta_n = 0, \quad \nu = \frac{1+\gamma}{(1-l_0c)(5-2\,\gamma)}, \quad l = \frac{\gamma + 6\,R_0 + 3}{(1-l_0c)(5-2\,\gamma)}, \quad l_0 = \frac{2\,\gamma + 3\,R_0 + 6}{(1-l_0c)(5-2\,\gamma)}.$$

Since $\|\delta F(x_0,x_{-1})^{-1} F(x_0)\| \leq \|\delta F(x_0,x_{-1})^{-1} F'(x_0)\| \|F'(x_0) F(x_0)\| \leq \frac{1}{(1-l_0c)} \frac{1+\gamma}{5-2\gamma}$. Note that $l_0 < l$. Therefore, the hypothesis of Kantorovich may not be satisfied, but conditions of Theorem 3.2 may be satisfied.

Finally, for the local case we study the following one.

Example 1.5.3. *Let $X = Y = \mathbb{R}^3$, $D = U(0,1)$, $x^* = (0,0,0)$ and define function F on D by*

$$F(x,y,z) = (e^x - 1, y^2 + y, z). \tag{1.5.7}$$

We have that for $u = (x,y,z)$

$$F'(u) = \begin{pmatrix} e^x & 0 & 0 \\ 0 & 2y+1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{1.5.8}$$

Using the norm of the maximum of the rows and (1.5.7)–(1.5.8) we see that since $F'(x^*) = diag\{1,1,1\}$, we can define parameters for Newton's method by

$$K = e/2, \tag{1.5.9}$$

$$K_0 = 1, \tag{1.5.10}$$

$$R^* = \frac{2}{e+4}, \tag{1.5.11}$$

$$R_0^* = R^*, \tag{1.5.12}$$

since $\theta_n = 1$, $\mu = 2$, $\lambda = 0$. *Then the Newton's method starting form* $x_0 \in B(x^*, R^*)$ *converges to a solution of* (1.5.7). *Note that using only Lipschitz condition we obtain the Rheinboldt or Traub ball* $R^*_{TR} = \frac{2}{3e} < R^*$.

Example 1.5.4. *In this example we present an application of the previous analysis to the Chandrasekhar equation:*

$$x(s) = 1 + \frac{s}{4}x(s)\int_0^1 \frac{x(t)}{s+t}\,dt, \quad s \in [0,1], \tag{1.5.13}$$

which arises in the theory of radiative transfer [7]; $x(s)$ *is the unknown function which is sought in* $C[0,1]$. *The physical background of this equation is fairly elaborate. It was developed by Chandraseckhar [7] to solve the problem of determination of the angular distribution of the radiant flux emerging from a plane radiation field. This radiation field must be isotropic at a point, that is the distribution in independent of direction at that point. Explicit definitions of these terms may be found in the literature [7]. It is considered to be the prototype of the equation,*

$$x(s) = 1 + \lambda s x(s)\int_0^1 \frac{\varphi(s)}{s+t}x(t)\,dt, \quad s \in [0,1],$$

for more general laws of scattering, where $\varphi(s)$ *is an even polynomial in s with*

$$\int_0^1 \varphi(s)\,ds \le \frac{1}{2}.$$

Integral equations of the above form also arise in the other studies [7]. We determine where a solution is located, along with its region of uniqueness.

Note that solving (3.7) is equivalent to solve $F(x) = 0$, *where* $F : C[0,1] \to C[0,1]$ *and*

$$[F(x)](s) = x(s) - 1 - \frac{s}{4}x(s)\int_0^1 \frac{x(t)}{s+t}\,dt, \quad s \in [0,1]. \tag{1.5.14}$$

To obtain a numerical solution of (3.7), we first discretize the problem and approach the integral by a Gauss-Legendre numerical quadrature with eight nodes,

$$\int_0^1 f(t)\,dt \approx \sum_{j=1}^{8} w_j f(t_j),$$

where

$t_1 = 0.019855072, \quad t_2 = 0.101666761, \quad t_3 = 0.237233795, \quad t_4 = 0.408282679,$
$t_5 = 0.591717321, \quad t_6 = 0.762766205, \quad t_7 = 0.898333239, \quad t_8 = 0.980144928,$
$w_1 = 0.050614268, \quad w_2 = 0.111190517, \quad w_3 = 0.156853323, \quad w_4 = 0.181341892,$
$w_5 = 0.181341892, \quad w_6 = 0.156853323, \quad w_7 = 0.111190517, \quad w_8 = 0.050614268.$

If we denote $x_i = x(t_i)$, $i = 1,2,\ldots,8$, equation (3.7) is transformed into the following nonlinear system:

$$x_i = 1 + \frac{x_i}{4}\sum_{j=1}^{8} a_{ij}x_j, \quad i = 1,2,\ldots,8,$$

where, $a_{ij} = \dfrac{t_i w_j}{t_i + t_j}$.

Denote now $\overline{x} = (x_1, x_2, \ldots, x_8)^T$, $\overline{1} = (1,1,\ldots,1)^T$, $A = (a_{ij})$ *and write the last nonlinear system in the matrix form:*

$$\overline{x} = \overline{1} + \frac{1}{4}\overline{x} \odot (A\overline{x}), \tag{1.5.15}$$

where $\odot$ *represents the inner product. Set* $G(x) = x$. *If we choose* $\overline{x}_0 = (1,1,\ldots,1)^T$ *and* $\overline{x}_{-1} = (0,0,\ldots,0)^T$. *Assume sequence* $\{\overline{x}_n\}$ *is generated by secant-type mtehods with different choices of* θ_n. *Table 1 gives the comparison results for* $\|\overline{x}_{n+1} - \overline{x}_n\|$ *equipped with the max-norm for this example. The computational order of convergence (COC) is shown in Table 1.5.1 for various methods. Here (COC) is defined in [1],[4] by*

$$\rho \approx \ln\left(\frac{\|\overline{x}_{n+1} - \overline{x}^\star\|_\infty}{\|\overline{x}_n - \overline{x}^\star\|_\infty}\right) / \ln\left(\frac{\|\overline{x}_n - \overline{x}^\star\|_\infty}{\|\overline{x}_{n-1} - \overline{x}^\star\|_\infty}\right), \qquad n \in \mathbb{N},$$

The last line in Table 1.5.1 shows the (COC).

Table 1.5.1. The comparison results of $\|\overline{x}_{n+1} - \overline{x}_n\|$ for Example 3.3 using various methods

n	$\|x_{n+1} - x_n\|$ $\theta_n = 0$, **Newton**	$\|x_{n+1} - x_n\|$ $\theta_n = 1$, **secant**	$\|x_{n+1} - x_n\|$ $\theta_n = 2$, **Kurchatov**	$\|x_{n+1} - x_n\|$ $\theta_n = 1/2$, **midpoint**
1	9.49639×10^{-6}	4.70208×10^{-2}	4.33999×10^{-1}	1.42649×10^{-1}
2	8.18823×10^{-12}	7.77292×10^{-3}	3.28371×10^{-2}	1.51900×10^{-2}
3	5.15077×10^{-24}	5.14596×10^{-5}	2.33370×10^{-3}	1.66883×10^{-4}
4	1.79066×10^{-48}	3.89016×10^{-8}	9.32850×10^{-6}	1.34477×10^{-7}
5	1.95051×10^{-97}	1.77146×10^{-13}	2.214411×10^{-9}	1.03094×10^{-12}
6	2.12404×10^{-195}	5.35306×10^{-22}	1.801201×10^{-15}	5.63911×10^{-21}
ρ	2.00032	1.61815	1.61854	1.61817

References

[1] Amat, S., Busquier, S., Negra, M., Adaptive approximation of nonlinear operators, *Numer. Funct. Anal. Optim.*, 25 (2004), 397–405.

[2] Argyros, I.K., Computational theory of iterative methods. *Series: Studies in Computational Mathematics*, 15, Editors: C.K. Chui and L. Wuytack, 2007, Elsevier Publ. Co. New York, U.S.A.

[3] Argyros, I.K., Hilout, S., Weaker conditions for the convergence of Newton's method, *J. Complexity*, AMS, 28 (2012), 364–387.

[4] Argyros, I. K., Cho, Y. J., Hilout, S., *Numerical method for equations and its applications*. CRC Press/Taylor and Francis, New York, 2012.

[5] Argyros, I. K., George, S., Ball convergence for Steffensen-type fourth-order methods. *Int. J. Interac. Multim. Art. Intell.*, 3(4) (2015), 37–42.

[6] Cătinaş, E., The inexact, inexact perturbed, and quasi-Newton methods are equivalent models, *Math. Comp.*, 74(249) (2005), 291–301.

[7] Chandrasekhar, S., *Radiative transfer*, Dover Publ., New York, 1960.

[8] Dennis, J.E., Toward a unified convergence theory for Newton-like methods, in *Nonlinear Functional Analysis and Applications* (L. B. Rall, ed.) Academic Press, New York, (1971), 425–472.

[9] Ezquerro, J.A., Hernández, Rubio, M.J., Secant-like methods for solving nonlinear integral equations of the Hammerstein type, *J. Comput. Appl. Math.*, 115 (2000), 245–254.

[10] Ezquerro, J.A., Gutiérrez, J.M., Hernández, M.A., Romero, N., Rubio, M.J., The Newton method: from Newton to Kantorovich. (Spanish), *Gac. R. Soc. Mat. Esp.*, 13(1) (2010), 53–76.

[11] Gragg, W.B., Tapia, R.A., Optimal error bounds for the Newton-Kantorovich theorem, *SIAM J. Numer. Anal.*, 11 (1974), 10–13.

[12] Kantorovich, L.V., Akilov, G.P., *Functional Analysis*, Pergamon Press, Oxford, 1982.

[13] Magreñán, Á. A., A new tool to study real dynamics: The convergence plane, *App. Math. Comp.*, 248 (2014), 215–224.

[14] Ortega, L.M., Rheinboldt, W.C., *Iterative Solution of Nonlinear Equations in Several Variables*, Academic press, New York, 1970.

[15] Potra, F.A., Pták, V., Nondiscrete induction and iterative processes. *Research Notes in Mathematics*, 103. Pitman (Advanced Publishing Program), Boston, MA, 1984.

[16] Proinov, P.D., General local convergence theory for a class of iterative processes and its applications to Newton's method, *J. Complexity*, 25 (2009), 38–62.

[17] Proinov, P.D., New general convergence theory for iterative processes and its applications to Newton-Kantorovich type theorems, *J. Complexity*, 26 (2010), 3–42.

[18] Rheinboldt, W.C., An adaptative continuation process for solving systems of nonlinear equations, *Banach ctz. Publ.* 3 (1975), 129–142.

[19] Traub, J. F. , *Iterative method for solutions of equations*, Prentice-Hall, New Jersey, 1964.

[20] Schmidt, J. W., Untere Fehlerschranken fun Regula-Falsi Verhafren, *Period. Hungar.*, 9 (1978), 241–247.

[21] Yamamoto, T., A convergence theorem for Newton-like methods in Banach spaces, *Numer. Math.*, 51 (1987), 545–557.

[22] Zabrejko, P.P., Nguen, D.F., The majorant method in the theory of Newton-Kantorovich approximations and the Pták error estimates, *Numer. Funct. Anal. Optim.*, 9 (1987), 671–684.

Chapter 2

Efficient Steffensen-Type Algorithms for Solving Nonlinear Equations

2.1. Introduction

In this chapter we are concerned with the problem of approximating a locally unique solution $x^\star$ of an equation

$$F(x) = 0, \tag{2.1.1}$$

where F is an operator defined on a non–empty, open subset Ω of a Banach space $\mathcal{X}$ with values in a Banach space $\mathcal{Y}$.

Many problems in computational sciences can be brought in the form of equation (2.1.1). For example, the unknowns of engineering equations can be functions (difference, differential, and integral equations), vectors (systems of linear or nonlinear algebraic equations), or real or complex numbers (single algebraic equations with single unknowns). The solutions of these equations can rarely be found in closed form. That is why the most commonly used solution method are iterative. The practice of numerical analysis is usually connected to Newton-like methods [1,3,5,7–9, 10–16,18,19,21–27].

The study about convergence matter of iterative procedures is usually based on two types: semilocal and local convergence analysis. The semilocal convergence matter is, based on the information around an initial point, to give conditions ensuring the convergence of the iterative method; while the local one is, based on the information around a solution, to find estimates of the radii of convergence balls.

A classic iterative process for solving nonlinear equations is Chebyshev's method (see [5], [8], [17]):

$$\begin{cases} x_0 \in \Omega, \\ y_k = x_k - F'(x_k)^{-1}\, F(x_k), \\ x_{k+1} = y_k - \frac{1}{2}\, F'(x_k)^{-1} F''(x_k)(y_k - x_k)^2, \quad k \geq 0. \end{cases}$$

This one-point iterative process depends explicitly on the two first derivatives of F (namely, $x_{k+1} = \psi(x_k, F(x_k), F'(x_k), F''(x_k))$). Ezquerro and Hernández introduced in [15] some modifications of Chebyshev's method that avoid the computation of the second derivative of F and reduce the number of evaluations of the first derivative of F. Actually, these authors have obtained a modification of the Chebyshev iterative process which only need to evaluate the first derivative of F, (namely, $x_{k+1} = \overline{\Psi}(x_k, F'(x_k))$, but with third-order of convergence. In this chapter we recall this method as the Chebyshev–Newton–type method (CNTM) and it is written as follows:

$$\begin{cases} x_0 \in \Omega, \\ y_k = x_k - F'(x_k)^{-1}\, F(x_k), \\ z_k = x_k + a\,(y_k - x_k) \\ x_{k+1} = x_k - \dfrac{1}{a^2}\, F'(x_k)^{-1}\, ((a^2 + a - 1)\, F(x_k) + F(z_k)), \quad k \geq 0. \end{cases}$$

There is an interest in constructing families of iterative processes free of derivatives. To obtain a new family in [8] we considered an approximation of the first derivative of F from a divided difference of first order, that is, $F'(x_k) \approx [x_{k-1}, x_k, F]$, where, $[x, y; F]$ is a divided difference of order one for the operator F at the points x, $y \in \Omega$. Then, we introduce the Chebyshev–Secant–type method (CSTM)

$$\begin{cases} x_{-1},\, x_0 \in \Omega, \\ y_k = x_k - B_k^{-1}\, F(x_k), \quad B_k = [x_{k-1}, x_k; F], \\ z_k = x_k + a\,(y_k - x_k), \\ x_{k+1} = x_k - B_k^{-1}\,(b\,F(x_k) + c\,F(z_k)), \quad k \geq 0, \end{cases}$$

where a, b, c are non–negative parameters to be chosen so that sequence $\{x_k\}$ converges to $x^\star$. Note that (CSTM) is reduced to the secant method (SM) if $a = 0$, $b = c = 1/2$, and $y_k = x_{k+1}$. Moreover, if $x_{k-1} = x_k$, and F is differentiable on Ω, then, $F'(x_k) = [x_k, x_k; F]$, and (CSTM) reduces to Newton's method (NM).

We provided a semilocal convergence analysis for (CSTM) using recurrence sequences, and also illustrated its effectiveness through numerical examples. Dennis [14], Potra [23], Argyros [1]–[11], Ezquerro et al. [15] and others [16], [22], [25], have provided sufficient convergence conditions for the (SM) based on Lipschitz–type conditions on divided difference operator (see, also relevant works in [12]–[13], [17], [19], [20].

In this chapter, we continue the study of derivative free iterative processes. We introduce the Steffensen-type method (STTM):

$$\begin{cases} x_0 \in \Omega, \\ y_k = x_k - A_k^{-1}\, F(x_k), \quad A_k = [x_k, G(x_k); F], \\ z_k = x_k + a\,(y_k - x_k), \\ x_{k+1} = x_k - A_k^{-1}\,(b\,F(x_k) + c\,F(z_k)), \quad k \geq 0, \end{cases}$$

where, $G : X \to X$. Note that (STTM) reduces to (CNTM) if $G(x) = x$, $b = \frac{a^2+a-1}{a^2}$ and $c = \frac{1}{a^2}$ provided that F is Fréchet-differentiable on Ω.

In the special case $a = 0$, $b = c = \frac{1}{2}$, $x_{k+1} = y_k$ the quadratic convergence of (CNTM) is established in [15]. The semilocal convergence analysis of (CNTM) when A_k is replaced by the more general $G_k \in L(X, Y)$ is given by us in [8]. In the present chapter we provide a local convergence analysis for (STTM). Then, we give numerical examples to show that (STTM) is faster than (CSTM). In particular, three numerical examples are also provided. Firstly, we consider a scalar equation where the main study of the chapter is applied. Secondly, we give the radius of convergence of (STTM) for a nonlinear integral equation. Thirdly, we discretize a nonlinear integral equation and approximate a numerical solution by using (STTM).

2.2. Local Convergence of (STTM)

In this section we provide a local convergence analysis for (STTM). The radius of convergence is also found. The convergence order of (STTM) is at least quadratic if $(1-a)c = 1-b$ and at least cubic if $a=b=c=1$.

We need a result for zeros of functions related to our local convergence analysis of (STTM).

Lemma 2.2.1. *Suppose* $a \in [0,1]$, $b \in [0,1]$, $c \geq 0$ *are parameters satisfying* $(1-a)c = 1-b$, L_0, L *and* N *are positive constants with* $L_0 \leq L$. *Let* ψ *be a function defined on* $[0,+\infty)$ *by*

$$\begin{aligned}\psi(r) &= a^2c(N+2)^2L^3r^3 + 2[1-b+ac+c(N+2)]a(N+2)L^2r^2(1-\tfrac{N+1}{2}L_0r)\\ &\quad +4[|1-ac|(N+2)+a(1-b)]Lr(1-\tfrac{N+1}{2}L_0r)^2 - 8(1-\tfrac{N+1}{2}L_0r)^3.\end{aligned} \tag{2.2.1}$$

Then, ψ *has a least positive zero in* $(0,R_0]$ *with* R_0 *given by*

$$R_0 = \tfrac{2}{(N+2)L+(N+1)L_0}. \tag{2.2.2}$$

Proof. We shall consider two possibilities:
<u>*Case* $ac=0$.</u> If $a=0$, then function ψ further reduces to

$$\psi(r) = 4(1-\tfrac{N+1}{2}L_0r)^2[(N+2)Lr - 2(1-\tfrac{N+1}{2}L_0r)]$$

with minimal zero R_0 given by (2.2), since

$$R_0 < \frac{2}{(N+1)L_0}. \tag{2.2.3}$$

If $c=0$, then $b=1$ from the condition $(1-a)c = 1-b$ and function ψ becomes

$$\psi(r) = 4(1-\tfrac{N+1}{2}L_0r)^2[(N+2)Lr - 2(1-\tfrac{N+1}{2}L_0r)]$$

leading to the same value for minimal zero R_0.
<u>*Case* $ac>0$.</u> Using the definition of R_0 we get

$$1-\tfrac{N+1}{2}L_0R_0 = \tfrac{(N+2)L}{(N+2)L+(N+1)L_0}.$$

Then, we have $\psi(0) = -8 < 0$ and

$$\begin{aligned}\psi(R_0) &= 4[|1-ac|(N+2)+a(1-b)]\tfrac{2L}{(N+2)L+(N+1)L_0}\left(\tfrac{(N+2)L}{(N+2)L+(N+1)L_0}\right)^2\\ &\quad +2[1-b+ac+c(N+2)]\tfrac{4a(N+2)L^2}{[(N+2)L+(N+1)L_0]^2}\tfrac{(N+2)L}{(N+2)L+(N+1)L_0}\\ &\quad +\tfrac{8a^2c(N+2)^2L^3}{((N+2)L+(N+1)L_0)^3} - \tfrac{8(N+2)^3L^3}{((N+2)L+(N+1)L_0)^3}\\ &= \tfrac{8L^3(N+2)^2}{((N+2)L+(N+1)L_0)^3}[(|1-ac|+ac-1)(N+2)+2a(1-b+ac)].\end{aligned} \tag{2.2.4}$$

If $1-ac \geq 0$ the bracket in (2.2.4) becomes $2a(1-b+ac) = 2ac > 0$, whereas if $1-ac \leq 0$, we have

$$2[(ac-1)(N+2)+a(1-b+ac)] = 2[(ac-1)(N+2)+ac] > 0.$$

Hence, in either case $\psi(R_0) > 0$. It follows from the intermediate value theorem that there exists a zero of function ψ in $(0,R_0)$ and the minimal such zero must satisfy $0 < R < R_0$. That completes the proof of the lemma.

Remark 2.2.2. *We are especially interested in the case when $a = b = c = 1$. It follows from* (2.2.1) *that in this case we can write*

$$\begin{aligned} \psi(r) &= 8(1-\tfrac{N+1}{2}L_0 r)^3[(N+2)^2(\tfrac{Lr}{2(1-\frac{N+1}{2}L_0 r)})^3 + (N+3)(N+2)(\tfrac{Lr}{2(1-\frac{N+1}{2}L_0 r)})^2 \\ &\quad -1]. \end{aligned} \tag{2.2.5}$$

Define function ϕ on $[0,+\infty)$ by

$$\phi(r) = (N+2)^2 r^3 + (N+3)(N+2)r^2 - 1. \tag{2.2.6}$$

We have $\phi(0) = -1 < 0$ and $\phi(1) = (N+2)^2 + (N+3)(N+2) - 1 > 0$. Then, again by the intermediate value theorem there exists $R_1 \in (0,1)$ such that $\phi(R_1) = 0$. Moreover, we get

$$\phi'(r) = 3(N+2)r^2 + 2(N+3)(N+2)r > 0, \quad for \;\; r > 0.$$

That is ϕ is increasing on $[0,+\infty)$. Hence, ϕ crosses the $x-$axis only once. Therefore R_1 is the unique zero of ϕ in $(0,1)$. In this case, by setting

$$\frac{LR}{2(1-\frac{N+1}{2}L_o R)} = R_1 \tag{2.2.7}$$

and solving for R we obtain

$$R^\star := R = \frac{2R_1}{L+(N+1)L_0 R_1}. \tag{2.2.8}$$

We can show the main result of this section concerning the local convergence of (STTM).

Theorem 2.2.3. *Suppose:*
(a) $F : \Omega \subseteq X \to Y$ and there exists divided difference [x,y;F] satisfying

$$[x,y;F](x-y) = F(x) - F(y) \quad for\, all\; x,y \in \Omega; \tag{2.2.9}$$

(b) Point $x^\star$ is a solution of equation $F(x) = 0$, $F'(x^\star)^{-1} \in L(Y,X)$ and there exists a constant $L > 0$ such that

$$\|F'(x^\star)^{-1}([x,y;F]-[u,v;F])\| \leq \frac{L}{2}(\|x-u\|+\|y-v\|) \quad for\, all\; x,y,u,v \in \Omega; \tag{2.2.10}$$

(c) There exists a constant $L_0 > 0$ such that

$$\|F'(x^\star)^{-1}([x,y;F]-F'(x^\star))\| \leq \frac{L_0}{2}(\|x-x^\star\|+\|y-x^\star\|) \quad for\, all\; x,y \in \Omega; \tag{2.2.11}$$

(d) $G : \Omega \subseteq X \to X$ *is continuous and such that* $G(x^\star) = x^\star$.
(e) There exists $N \in (0,1]$ *such that*

$$\|G(x) - G(x^\star)\| \leq N\|x - x^\star\| \quad for\ all\ x \in \Omega; \tag{2.2.12}$$

(f) The relation $(1-a)c = 1-b$ *is true;*
(g)

$$U(x^\star, R) = \{x \in \Omega : \|x - x^\star\| < R\} \subseteq \Omega, \tag{2.2.13}$$

where R is the least positive zero of function ψ *given in* (2.2.1).
Then, sequence $\{x_n\}$ *generated by (STTM) is well defined, remains in* $U(x^\star, R)$ *for all* $n \geq 0$ *and converges to* $x^\star$ *provided that* $x_0 \in U(x^\star, R)$. *Moreover, the following error estimates are satisfied for* $e_n = x_n - x^\star$

$$\|e_{n+1}\| \leq \xi_n \|e_n\|^2 \leq \xi \|e_n\|^2, \tag{2.2.14}$$

where

$$h_n = \frac{L}{2(1 - \frac{L_0}{2}(N+1)\|e_n\|)}, \tag{2.2.15}$$

$$\begin{aligned} \xi_n &= [|1-ac|(N+2) + a(1-b)]h_n + [1-b+ac+c(N+2)]a(N+2)h_n^2\|e_n\| \\ &\quad + a^2c(N+2)^2h_n^3\|e_n\|^2, \end{aligned} \tag{2.2.16}$$

$$H = \frac{L}{2(1-\frac{L_0}{2}(N+1)R)}, \tag{2.2.17}$$

$$\begin{aligned} \xi &= [|1-ac|(N+2) + a(1-b)]H + [1-b+ac+c(N+2)]a(N+2)H^2R \\ &\quad + a^2c(N+2)^2H^3R^2. \end{aligned} \tag{2.2.18}$$

In particular, if

$$a = b = c = 1, \tag{2.2.19}$$

the optimal (STTM) is obtained which is cubically convergent. Furthermore, the error estimates (2.2.14) *are*

$$\|e_{n+1}\| \leq \lambda_n \|e_n\|^3 \leq \lambda \|e_n\|^3, \tag{2.2.20}$$

where

$$\lambda_n = (N+2)h_n^2[N+3+(N+2)h_n\|e_n\|] \tag{2.2.21}$$

and

$$\lambda = (N+2)H^2[N+3+(N+2)HR^\star], \tag{2.2.22}$$

where $R^\star$ *is given by* (2.2.8).

Proof. We shall show the assertions of the theorem using induction. Let $u_n = y_n - x^\star$ and $v_n = z_n - x^\star (n \geq 0)$. Using $x_0 \in U(x^\star, R)$, $G(x^\star) = x^\star$ and (2.2.12) we obtain

$$\|G(x_0) - x^\star\| = \|G(x_0) - G(x^\star)\| \leq N\|x_0 - x^\star\| \leq \|x_0 - x^\star\| < R, \tag{2.2.23}$$

which implies $G(x_0) \in U(x^\star, R)$. Then, we have by (2.2.11) and (2.2.3):

$$\begin{aligned} \|F'(x^\star)^{-1}(F'(x^\star) - [x_0, G(x_0); F])\| &\leq \tfrac{L_0}{2}(\|x^\star - x_0\| + \|x^\star - G(x_0)\|) \\ &= \tfrac{L_0}{2}(\|x^\star - x_0\| + \|G(x^\star) - G(x_0)\|) \\ &\leq \tfrac{L_0}{2}(\|x^\star - x_0\| + N\|x^\star - x_0\|) \\ &< \tfrac{L_0}{2}(N+1)R < \tfrac{L_0}{2}(N+1)R_0 < 1. \end{aligned} \tag{2.2.24}$$

It follows from (2.2.24) and the Banach lemma on invertible operators that $A_0^{-1} \in L(Y,X)$ and

$$\|A_0^{-1}F'(x^\star)\| \leq \frac{1}{1-\frac{L_0}{2}(N+1)\|e_0\|} < \frac{1}{1-\frac{L_0}{2}(N+1)R}. \tag{2.2.25}$$

Thus, y_0 is well defined. Using (2.2.9) we have

$$\begin{aligned} F(x_0) &= F(x_0)-F(x^\star) = -(F(x^\star)-F(x_0)) \\ &= -[x^\star,x_0;F](x^\star-x_0) = [x^\star,x_0;F]e_0 \\ &= [G(x^\star),x_0;F]e_0. \end{aligned} \tag{2.2.26}$$

So,

$$\begin{aligned} u_0 &= y_0-x^\star = x_0-x^\star-A_0^{-1}F(x_0) \\ &= A_0^{-1}F'(x^\star)F'(x^\star)^{-1}(A_0-[x^\star,x_0;F])e_0. \end{aligned} \tag{2.2.27}$$

By (2.2.10), (2.2.25) and (2.2.27) we get in turn

$$\begin{aligned} \|u_0\| &\leq \|A_0^{-1}F'(x^\star)\|\tfrac{L}{2}(\|x_0-x^\star\|+\|G(x_0)-x_0\|)\|e_0\| \\ &\leq \frac{1}{1-\frac{L_0}{2}(N+1)\|e_0\|}\frac{L}{2}(\|x_0-x^\star\|+\|G(x_0)-G(x^\star)\|+\|x^\star-x_0\|)\|e_0\| \\ &\leq \frac{1}{1-\frac{L_0}{2}(N+1)\|e_0\|}\frac{L}{2}(2\|x_0-x^\star\|+N\|x^\star-x_0\|)\|e_0\| \\ &\leq \frac{L(N+2)R_0}{2[1-\frac{L_0}{2}(N+1)R_0]}\|e_0\| = \|e_0\| < R, \end{aligned} \tag{2.2.28}$$

which implies $y_0 \in U(x^\star,R)$. Noting that

$$v_0 = z_0-x^\star = au_0+(1-a)e_0, \tag{2.2.29}$$

we get

$$\|v_0\| \leq a\|u_0\|+(1-a)\|e_0\| \leq \|e_0\| < R. \tag{2.2.30}$$

As in (2.2.26), we have

$$F(z_0) = [x^\star,z_0;F]v_0 = [G(x^\star),z_0;F]v_0. \tag{2.2.31}$$

Using (2.2.29) and (2.2.31), we get

$$\begin{aligned} e_1 &= e_0-A_0^{-1}(bF(x_0)+cF(z_0)) \\ &= A_0^{-1}\big([x_0,G(x_0);F]e_0-(b[x^\star,x_0;F]e_0+c[x^\star,z_0;F]v_0)\big) \\ &= A_0^{-1}\big([x_0,G(x_0);F]e_0-b[x^\star,x_0;F]e_0-c[x^\star,z_0;F](au_0+(1-a)e_0)\big) \\ &= A_0^{-1}\big([x_0,G(x_0);F]e_0-b[x^\star,x_0;F]e_0-(1-b)[x^\star,z_0;F]e_0-ac[x^\star,z_0;F]u_0\big) \\ &= A_0^{-1}\big(([x_0,G(x_0);F]-[x^\star,x_0;F])e_0+(1-b)([x^\star,x_0;F]-[x^\star,z_0;F])e_0\big) \\ &\quad +acA_0^{-1}([x^\star,x_0;F]-[x^\star,z_0;F]+[x_0,G(x_0);F]-[x^\star,x_0;F])u_0-acu_0. \end{aligned} \tag{2.2.32}$$

Define

$$\begin{aligned} D_0 &= A_0^{-1}([x_0,G(x_0);F]-[x^\star,x_0;F]), \\ E_0 &= A_0^{-1}([x^\star,x_0;F]-[x^\star,z_0;F]), \end{aligned} \tag{2.2.33}$$

then, we have from (2.2.27) that

$$u_0 = D_0e_0. \tag{2.2.34}$$

Moreover, we can rewrite (2.2.32) as

$$\begin{aligned} e_1 &= D_0e_0+(1-b)E_0e_0+acE_0u_0+acD_0u_0-acu_0 \\ &= (1-ac)D_0e_0+(1-b)E_0e_0+acE_0D_0e_0+acD_0^2e_0. \end{aligned} \tag{2.2.35}$$

We need to find upper bounds on the norms $\|D_0\|$ and $\|E_0\|$. Using (2.2.10) and (2.2.33) we get in turn

$$\begin{aligned} \|D_0\| &\leq \|A_0^{-1}F'(x^\star)\|\|F'(x^\star)^{-1}([x_0,G(x_0);F]-[x^\star,x_0;F])\| \\ &\leq \frac{1}{1-\frac{L_0}{2}(N+1)\|e_0\|}\frac{L}{2}(\|x_0-x^\star\|+\|G(x_0)-x_0\|) \\ &\leq \frac{\frac{L}{2}(N+2)\|x_0-x^\star\|}{1-\frac{L_0}{2}(N+1)\|x_0-x^\star\|}=\frac{L(N+2)\|e_0\|}{2(1-\frac{L_0}{2}(N+1)\|e_0\|)} \end{aligned} \tag{2.2.36}$$

and

$$\begin{aligned} \|E_0\| &\leq \|A_0^{-1}F'(x^\star)\|\|F'(x^\star)^{-1}([x^\star,x_0;F]-[x^\star,z_0;F])\| \\ &\leq \frac{\frac{L}{2}\|z_0-x_0\|}{1-\frac{L_0}{2}(N+1)\|x_0-x^\star\|}=\frac{\frac{aL}{2}\|y_0-x_0\|}{1-\frac{L_0}{2}(N+1)\|e_0\|}\leq\frac{\frac{aL}{2}(\|u_0\|+\|e_0\|)}{1-\frac{L_0}{2}(N+1)\|e_0\|} \\ &\leq \frac{\frac{aL}{2}(\|D_0\|+1)\|e_0\|}{1-\frac{L_0}{2}(N+1)\|e_0\|}\leq\frac{1}{1-\frac{L_0}{2}(N+1)\|e_0\|}\frac{aL}{2}\Big[\frac{\frac{L}{2}(N+2)\|e_0\|}{1-\frac{L_0}{2}(N+1)\|e_0\|}+1\Big]\|e_0\| \\ &\leq \frac{aL^2(N+2)\|e_0\|^2}{4(1-\frac{L_0}{2}(N+1)\|e_0\|)^2}+\frac{aL\|e_0\|}{2(1-\frac{L_0}{2}(N+1)\|e_0\|)}. \end{aligned} \tag{2.2.37}$$

Using (2.2.35)-(2.2.37) we get

$$\begin{aligned} \|e_1\| &\leq |1-ac|\frac{L(N+2)\|e_0\|^2}{2(1-\frac{L_0}{2}(N+1)\|e_0\|)}+\frac{(1-b)aL^2(N+2)\|e_0\|^3}{4(1-\frac{L_0}{2}(N+1)\|e_0\|)^2}+\frac{(1-b)aL\|e_0\|^2}{2(1-\frac{L_0}{2}(N+1)\|e_0\|)} \\ &\quad +ac\Big(\frac{aL^2(N+2)\|e_0\|^2}{4(1-\frac{L_0}{2}(N+1)\|e_0\|)^2}+\frac{aL\|e_0\|}{2(1-\frac{L_0}{2}(N+1)\|e_0\|)}\Big)\frac{L(N+2)\|e_0\|^2}{2(1-\frac{L_0}{2}(N+1)\|e_0\|)} \\ &\quad +ac\frac{L^2(N+2)^2\|e_0\|^3}{4(1-\frac{L_0}{2}(N+1)\|e_0\|)^2} \\ &\leq h_0|1-ac|(N+2)\|e_0\|^2+(1-b)a(N+2)h_0^2\|e_0\|^3+(1-b)a\|e_0\|^2h_0 \\ &\quad +ac(ah_0^2(N+2)\|e_0\|^2+a\|e_0\|h_0)\|e_0\|^2(N+2)h_0+ac(N+2)^2h_0^2\|e_0\|^3 \\ &= \xi_0\|e_0\|^2\leq\xi\|e_0\|^2\leq\{[|1-ac|(N+2)+a(1-b)]HR \\ &\quad +[1-b+ac+c(N+2)]a(N+2)H^2R^2+a^2c(N+2)^2H^3R^3\}\|e_0\| \\ &= \{[|1-ac|(N+2)+a(1-b)]\frac{LR}{2(1-L_0(\frac{N+1}{2})R)} \\ &\quad +[1-b+ac+c(N+2)]a(N+2)\frac{L^2R^2}{4(1-L_0(\frac{N+1}{2})R)^2} \\ &\quad +a^2c(N+2)^2\frac{L^3R^3}{8(1-L_0(\frac{N+1}{2})R)^3}\}\|e_0\| \\ &= \frac{\psi(R)+8(1-L_0(\frac{N+1}{2})R)^3}{8(1-L_0(\frac{N+1}{2})R)^3}\|e_0\|=\|e_0\|<R, \end{aligned} \tag{2.2.38}$$

which implies $x_1\in U(x^\star,R)$. Hence, assertion (2.2.14) is true for $n=0$.

Let us assume $\{x_n\}$ is well defined and $x_n\in U(x^\star,R)$ for all $0\leq n\leq k$ $(k\geq 1)$. Using an analogous way with x_0 replaced by x_k we deduce:
(i) $G(x_k)\in U(x^\star,R)$;
(ii) $A_k^{-1}\in L(Y,X)$ and

$$\|A_k^{-1}F'(x^\star)\|\leq\frac{1}{1-\frac{L_0}{2}(N+1)\|e_k\|}<\frac{1}{1-\frac{L_0}{2}(N+1)R}; \tag{2.2.39}$$

(iii) y_k is well defined, $y_k \in U(x^\star, R)$ and

$$\begin{array}{rcl} \|u_k\| & \leq & \|A_k^{-1}F'(x^\star)\|\frac{L}{2}(\|x_k - x^\star\| + \|G(x_k) - x_k\|)\|e_k\| \\ & \leq & \frac{1}{1-\frac{L_0}{2}(N+1)\|e_k\|}\frac{L}{2}(\|x_k - x^\star\| + \|G(x_k) - x^\star\| + \|x^\star - x_k\|)\|e_k\| \\ & \leq & \frac{L(N+2)\|e_k\|^2}{2(1-L_0(\frac{N+1}{2})\|e_k\|)} \leq \frac{L(N+2)R_0\|e_k\|}{2(1-L_0(\frac{N+1}{2})R_0)} = \|e_k\| < R; \end{array} \tag{2.2.40}$$

(iv) z_k is well defined, $z_k \in U(x^\star, R)$;
(v) x_{k+1} is well defined and

$$\begin{array}{rcl} \|e_{k+1}\| & \leq & h_k|1-ac|(N+2)\|e_k\|^2 + (1-b)a(N+2)h_k^2\|e_k\|^3 \\ & & +(1-b)ah_k\|e_k\|^2 + ac(ah_k^2(N+2)\|e_k\|^2 + a\|e_k\|h_k)(N+2)h_k\|e_k\|^2 \\ & & +ac(N+2)^2h_k^2\|e_k\|^3 \leq \xi_k\|e_k\|^2 \leq \xi\|e_k\|^2 \leq \|e_k\| < R. \end{array} \tag{2.2.41}$$

The induction is completed and by (2.2.41) $\lim_{k\to\infty} x_k = x^\star$. In the special case of $a = b = c = 1$, in view of $\xi_n = \lambda_n e_n$ and $R^* = R$, we can deduce that the error estimates (2.2.20) hold for any $n \geq 0$. That completes the proof of the theorem.

Remark 2.2.4. *(a) In view of* (2.2.10) *condition* (2.2.11) *always holds. Hence,* (2.2.11) *is not an additional to* (2.2.10) *hypothesis, since in practice the computation of L requires that of L_0.*
(b) It follows from (2.2.22) *that λ is directly proportional to $R^\star$ since L_0, L and N are constants. Clearly, the smaller $R^\star$ is the smaller the ratio of convergence in* (2.2.22) *will be.*
(c) Note that (2.2.10) *implies that F is a differentiable operator in Ω [5,17,19].*

2.3. Numerical Examples

In this section, we present numerical examples, where we verify the conditions of Theorem 2.2.3

Example 2.3.1. *Let $X = Y = \mathbb{R}$, $\Omega = (-1,1)$ and define F on Ω by*

$$F(x) = e^x - 1. \tag{2.3.1}$$

Then, $x^\star = 0$ is a solution of Eq. (2.1.1)*, and $F'(x^\star) = 1$. Note that for any $x, y, u, v \in \Omega$, we have*

$$\begin{array}{l} |F'(x^\star)^{-1}([x,y;F] - [u,v;F])| = |\int_0^1 (F'(tx+(1-t)y) - F'(tu+(1-t)v))dt| \\ = |\int_0^1\int_0^1 (F''\big(\theta(tx+(1-t)y) + (1-\theta)(tu+(1-t)v)\big)\big(tx+(1-t)y - (tu+(1-t)v)\big)d\theta dt| \\ = |\int_0^1\int_0^1 (e^{\theta(tx+(1-t)y)+(1-\theta)(tu+(1-t)v)}\big(tx+(1-t)y - (tu+(1-t)v)\big)d\theta dt| \\ \leq \int_0^1 e|t(x-u) + (1-t)(y-v)|dt \\ \leq \frac{e}{2}(|x-u| + |y-v|) \end{array} \tag{2.3.2}$$

and

$$\begin{array}{l} |F'(x^\star)^{-1}([x,y;F] - [x^\star,x^\star;F])| = |\int_0^1 F'(tx+(1-t)y)dt - F'(x^\star)| \\ = |\int_0^1 (e^{tx+(1-t)y} - 1)dt| \\ = |\int_0^1 (tx+(1-t)y)(1 + \frac{tx+(1-t)y}{2!} + \frac{(tx+(1-t)y)^2}{3!} + \cdots)dt| \\ \leq |\int_0^1 (tx+(1-t)y)(1 + \frac{1}{2!} + \frac{1}{3!} + \cdots)dt| \\ \leq \frac{e-1}{2}(|x-x^\star| + |y-x^\star|). \end{array} \tag{2.3.3}$$

That is to say, the Lipschitz condition (2.2.10) *and the center-Lipschitz condition* (2.2.11) *are true for* $L = e$ *and* $L_0 = e - 1$*, respectively.*

Choose $G(x) = x - hF(x)$*, where* $h \in (0, \frac{2}{e-1})$ *is a constant. Then,* $G : \Omega \subseteq X \to X$ *is continuous and such that* $G(x^\star) = x^\star$*. Moreover, for any* $x \in \Omega$*, we have*

$$\begin{aligned} |G(x) - G(x^\star)| &= |x - h(e^x - 1)| = |x - h(x + \tfrac{x^2}{2!} + \tfrac{x^3}{3!} + \cdots)| \\ &= |(1-h)x - h(\tfrac{x^2}{2!} + \tfrac{x^3}{3!} + \cdots)| \le (|1-h| + h|\tfrac{1}{2!} + \tfrac{1}{3!} + \cdots|)|x| \\ &= (|1-h| + h(e-2))|x - x^\star|, \end{aligned} \tag{2.3.4}$$

which means condition (2.2.12) *is true for* $N = |1-h| + h(e-2)$ *and* $N \in (0,1)$.

Table 2.3.1. The comparison results of R_0 and R for Example 2.3.1 using various choices of a, b, c and h

Case	h	R_0(using both (2.10) and (2.11))	R_0(using only (2.10))	R(using both (2.10) and (2.11))	R(using only (2.10))
$a = b = c = 1$	0.99	0.193161183	0.165629465	0.113779771	0.103632763
	1.00	0.193394634	0.165839812	0.113940329	0.103781113
	1.01	0.191979459	0.164565090	0.112967093	0.102882059
$a = b = 1, c = 0.5$	0.99	0.193161183	0.165629465	0.130274315	0.117141837
	1.00	0.193394634	0.165839812	0.130452367	0.117305161
	1.01	0.191979459	0.164565090	0.129373022	0.116315329
$a = b = 0.5, c = 1$	0.99	0.193161183	0.165629465	0.128109720	0.115388719
	1.00	0.193394634	0.165839812	0.128282409	0.115547600
	1.01	0.191979459	0.164565090	0.127235492	0.114584632

Table 2.3.3 gives the comparison results of R_0 *and* R *for Example 2.3.1 using various choices of* a, b, c *and* h*, which show that the convergence radius* R *is always enlarged by using both condition* (2.2.10) *and* (2.2.11) *than the one by using only condition* (2.2.10)*. The same result is true for* R_0*.*

Let us set $h = 1$ *and choose* $x_0 = 0.11$*. Suppose sequence* $\{x_n\}$ *is generated by (STTM). Table 2.3.3 gives the error estimates for Example 2.3.1 using various choices of* a, b *and* c*, which shows that all error estimates given by* (2.2.14) *(or* (2.2.14)*) are satisfied. Moreover, the error estimates of case* $a = b = c = 1$ *are smallest among all choices of* a, b, c*.*

Example 2.3.2. *Let* $X = Y = C[0,1]$*, the space of continuous functions defined on* $[0,1]$*, equipped with the max norm and* $\Omega = \overline{U}(0,1)$*. Define function* F *on* Ω*, given by*

$$F(x)(s) = x(s) - 5\int_0^1 stx^3(t)dt, \tag{2.3.5}$$

and the divided difference of F *is defined by*

$$[x, y; F] = \int_0^1 F'(tx + (1-t)y)dt. \tag{2.3.6}$$

Table 2.3.2. The comparison results of error estimates for Example 2.3.1 using various choices of a, b and c

Case	n	$\|e_{n+1}\|$	$\lambda_n\|e_n\|^3$(or $\xi_n\|e_n\|^2$)	$\lambda\|e_n\|^3$(or $\xi\|e_n\|^2$)
$a=b=c=1$	0	1.73843e-05	0.040042622	0.040806467
	1	1.32733e-15	9.80995e-14	1.61073e-13
$a=b=1, c=0.5$	0	0.000178345	0.039055027	0.053915139
	1	7.08930e-13	4.32947e-08	1.41726e-07
$a=b=0.5, c=1$	0	0.001347553	0.043408200	0.057192845
	1	2.26682e-07	3.11418e-06	8.58318e-06

Then, we have

$$[F'(x)y](s) = y(s) - 15\int_0^1 stx^2(t)y(t)dt, \ \ for \ all \ \ y \in \Omega. \tag{2.3.7}$$

We have $x^\star(s) = 0$ for all $s \in [0,1]$, $L_0 = 7.5$ and $L = 15$ [5].

Choose $G(x) = x$ and $a = b = c = 1$. Then, $N = 1$. Using Theorem 2.2.3 and Remark 2.2.4, we deduce that R_1 is the unique positive zero of function

$$\phi(r) = 9r^3 + 12r^2 - 1, \tag{2.3.8}$$

which leads to $R_1 = 0.263762616$. Moreover, the radius of convergence of (STTM) is given by

$$R^\star = R = \frac{2R_1}{L + (N+1)L_0R_1} = 0.027828287, \tag{2.3.9}$$

which is bigger than the corresponding radius

$$R' = \frac{2R_1}{L + (N+1)LR_1} = 0.023023089 \tag{2.3.10}$$

obtained by only using the Lipschitz condition (2.2.10).

In the last example we are not interested in checking if the hypotheses of Theorem 2.2.3 are satisfied or not, but comparing the numerical behavior of (STTM) with earlier methods.

Example 2.3.3. *In this example we present an application of the previous analysis to the significant Chandrasekhar integral equation [7]:*

$$x(s) = 1 + \frac{s}{4}x(s)\int_0^1 \frac{x(t)}{s+t}dt, \quad s \in [0,1]. \tag{2.3.11}$$

Integral equations of the form (2.3.11) *are very important and appear in the areas of neutron transport, radiative transfer and the Kinetic theory of gasses. We refer the interested reader to [1,11,17] where a detailed description of the physical phenomenon described by* (2.3.11) *can be found. We determine where a solution is located, along with its region of uniqueness. Later, the solution is approximated by an iterative method of (STTM).*

Note that solving ((2.3.11)) is equivalent to solve $F(x) = 0$, where $F : C[0,1] \to C[0,1]$ and

$$[F(x)](s) = x(s) - 1 - \frac{s}{4}x(s)\int_0^1 \frac{x(t)}{s+t}\,dt, \quad s \in [0,1]. \tag{2.3.12}$$

To obtain a numerical solution of (2.3.11)*, we first discretize the problem and we find it convenient by testing several number of nodes to approach the integral by a Gauss-Legendre numerical quadrature with eight nodes (see also [1], [11], [17])*

$$\int_0^1 f(t)\,dt \approx \sum_{j=1}^{8} w_j f(t_j),$$

where

$$\begin{array}{llll} t_1 = 0.019855072, & t_2 = 0.101666761, & t_3 = 0.237233795, & t_4 = 0.408282679, \\ t_5 = 0.591717321, & t_6 = 0.762766205, & t_7 = 0.898333239, & t_8 = 0.980144928, \\ w_1 = 0.050614268, & w_2 = 0.111190517, & w_3 = 0.156853323, & w_4 = 0.181341892, \\ w_5 = 0.181341892, & w_6 = 0.156853323, & w_7 = 0.111190517, & w_8 = 0.050614268. \end{array}$$

If we denote $x_i = x(t_i)$, $i = 1,2,\ldots,8$, equation (2.3.11) *is transformed into the following nonlinear system:*

$$x_i = 1 + \frac{x_i}{4}\sum_{j=1}^{8} a_{ij}x_j, \quad i = 1,2,\ldots,8,$$

where, $a_{ij} = \dfrac{t_i w_j}{t_i + t_j}$.

Table 2.3.3. The comparison results of $\|\overline{x}_{n+1} - \overline{x}_n\|$ for Example 2.3.3 using various methods

n	STTM ($a = b = c = 1$)	STTM ($a = 0.5, b = 0, c = 2$)	CSTM ($a = b = c = 1$)	CSTM ($a = 0.5, b = 0, c = 2$)
1	2.49e-01	2.45e-01	2.49e-01	2.45e-01
2	5.69e-04	4.85e-03	6.14e-04	4.87e-03
3	3.40e-12	1.33e-06	5.76e-07	6.18e-06
4	4.34e-37	8.02e-14	1.91e-15	3.28e-12
5	6.36e-112	2.46e-28	4.34e-30	1.33e-24
6	1.54e-336	2.04e-57	8.04e-62	1.40e-49

Denote now $\overline{x} = (x_1, x_2, \ldots, x_8)^T$, $\overline{1} = (1,1,\ldots,1)^T$, $A = (a_{ij})$ and write the last nonlinear system in the matrix form:

$$\overline{x} = \overline{1} + \frac{1}{4}\overline{x} \odot (A\overline{x}), \tag{2.3.13}$$

where $\odot$ represents the inner product. Set $G(x) = x$. If we choose $\overline{x}_0 = (1,1,\ldots,1)^T$ and $\overline{x}_{-1} = (.99,.99,\ldots,.99)^T$. Assume sequence $\{\overline{x}_n\}$ is generated by (STTM) (or (CSTM)) with different choices of parameters a, b and c. Table 2.3.3 gives the comparison results for $\|\overline{x}_{n+1} - \overline{x}_n\|$ equipped with the max-norm for this example, which show that (STTM) is faster *than (CSTM). Here, we perform the computations by Maple 11 in a computer equipped with Inter(R) Core(TM) i3-2310M CPU.*

In future results we shall use higher precision instead of a fixed number of digits in all computations. We shall also use an adaptive arithmetic in each step of the iterative method. Note that this higher precision is only necessary in the last step of the iterative process.

Table 2.3.3 shows the usefulness of (STTM) since it is faster than other relevant methods in the literature like (CSTM).

References

[1] Argyros, I.K., Polynomial operator equations in abstract spaces and applications, St.Lucie/CRC/Lewis Publ. *Mathematics series*, 1998, Boca Raton, Florida, U.S.A.

[2] Argyros, I.K., On the Newton–Kantorovich hypothesis for solving equations, *J. Comput. Appl. Math.*, 169 (2004), 315–332.

[3] Argyros, I.K., A unifying local–semilocal convergence analysis and applications for two–point Newton–like methods in Banach space, *J. Math. Anal. Appl.*, 298 (2004), 374–397.

[4] Argyros, I.K., New sufficient convergence conditions for the Secant method, *Chechoslovak Math. J.*, 55 (2005), 175–187.

[5] Argyros, I.K., *Convergence and applications of Newton–type iterations*, Springer–Verlag Publ., New–York, 2008.

[6] Argyros, I.K., Hilout, S., On the weakening of the convergence of Newton's method using recurrent functions, *J. Complexity*, 25 (2009), 530–543.

[7] Argyros, I.K., Hilout, S., On the convergence of two-step Newton-type methods of high efficiency order, *Applicationes Mathematicae*, 36(4) (2009), 465-499.

[8] Argyros, I.K., Ezquerro, J., Gutiérrez, J.M., Hernández, M., Hilout, S., On the semilocal convergence of efficient Chebyshev-Secant-type methods, *J. Comput. Appl. Math.*, 235 (2011), 3195–3206.

[9] Argyros, I.K., Cho, Y.J., Hilout, S., *Numerical Methods for Equations and Its Applications*, CRC Press/Taylor and Francis Group, New York, 2012.

[10] Argyros, I. K., George, S. Ball convergence for Steffensen-type fourth-order methods. *Int. J. Interac. Multim. Art. Intell.*, 3(4) (2015), 37–42.

[11] Argyros, I. K., González, D., Local convergence for an improved Jarratt-type method in Banach space. *Int. J. Interac. Multim. Art. Intell.*, 3(4) (2015), 20–25.

[12] Catinas, E., On some iterative methods for solving nonlinear equations, *Revue d' analyse numerique et de theorie de l'approximation*, 23(1) (1994), 47–53.

[13] Chandrasekhar, S., *Radiative transfer*, Dover Publ., New–York, 1960.

[14] Dennis, J.E., Toward a unified convergence theory for Newton–like methods, in *Nonlinear Funct. Anal. App.* (L.B. Rall, ed.), Academic Press, New York, (1971), 425–472.

[15] Ezquerro, J.A., Hernández, M.A., An optimization of Chebyshev's method, *J. Complexity*, 25 (2009), 343–361.

[16] Hernández, M.A., Rubio, M.J., Ezquerro, J.A., Solving a special case of conservative problems by Secant–like method, *Appl. Math. Comp.*, 169 (2005), 926–942.

[17] Hernández, M.A., Rubio, M.J., Ezquerro, J.A., Secant–like methods for solving nonlinear integral equations of the Hammerstein type, *J. Comp. Appl. Math.*, 115 (2000), 245–254.

[18] Grau, M., Noguera, M., A variant of Cauchy's method with accelerated fifth-order convergence. *Appl. Math. Lett.*, 17 (2004), 509–517.

[19] Kantorovich, L.V., Akilov, G.P., *Functional Analysis*, Pergamon Press, Oxford, 1982.

[20] Laasonen, P., Ein überquadratisch konvergenter iterativer algorithmus, *Ann. Acad. Sci. Fenn. Ser I*, 450 (1969), 1–10.

[21] Magreñán, Á. A. , Different anomalies in a Jarratt family of iterative root-finding methods, *App. Math. Comp.* 233 (2014), 29–38.

[22] Ortega, J.M., Rheinboldt, W.C., *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970.

[23] Potra, F.A., Sharp error bounds for a class of Newton–like methods, *Libertas Mathematica*, 5 (1985), 71–84.

[24] Petković, M.S., Petković, L.D., Construction of zero-finding methods by Weierstrass functions, *App. Math. Comp.*, 184 (2007), 351–359.

[25] Petković, M.S., Ilić, S., Dzunić, J., Derivative free two-point methods with and without memory for solving nonlinear equations, *App. Math. Comp.*, 217 (2010), 1887–1895.

[26] Schmidt, J.W., Untere Fehlerschranken fur Regula–Falsi Verhafren, *Period. Hungar.*, 9 (1978), 241–247.

[27] Yamamoto, T., A convergence theorem for Newton–like methods in Banach spaces, *Numer. Math.*, 51 (1987), 545–557.

[28] Wolfe, M.A., Extended iterative methods for the solution of operator equations, *Numer. Math.*, 31 (1978), 153–174.

Chapter 3

On the Semilocal Convergence of Halley's Method under a Center-Lipschitz Condition on the Second Fréchet Derivative

3.1. Introduction

Let X and Y be Banach spaces and D be a non-empty, open and convex subset of X. The aim of this chapter is to show using a numerical example that the convergence theorem of Ref. [15] is false under the stated hypotheses. Reference [15] was concerned with the semilocal convergence of Halley's method for solving a nonlinear operator equation

$$F(x) = 0, \tag{3.1.1}$$

where $F : D \subset X \to Y$ is continuously twice Fréchet differentiable.

Many problems from computational sciences and other disciplines can be brought in a form similar to equation (3.1.1) using mathematical modelling [1, 2, 3, 4, 7, 8, 10, 12]. The solutions of these equations can be rarely be found in closed form. That is why most solution methods for these equations are iterative. The study about convergence matter of iterative procedures is usually based on two types: semilocal and local convergence analysis. The semilocal convergence matter is, based on the information around an initial point, to give conditions ensuring the convergence of the iterative procedure; while the local one is, based on the information around a solution, to find estimates of the radii of convergence balls.

Halley's method with initial point $x_0 \in D$ is defined by [1, 5, 6, 9, 13, 14, 15]

$$x_{k+1} = x_k - [I - L_F(x_k)]^{-1} F'(x_k)^{-1} F(x_k), \quad k = 0, 1, 2, \ldots, \tag{3.1.2}$$

where, $L_F(x) = \frac{1}{2} F'(x)^{-1} F''(x) F'(x)^{-1} F(x)$. Let $U(x,R)$, $\overline{U}(x,R)$ stand, respectively, for the open and closed balls in X with center x and radius $R > 0$. Halley's method is cubically convergent and has been studied extensively (see [1-13] and the references therein). In particular, recurrence relations have been used by Parida [13], Parida and Gupta [14], Chun, Stănică and Neta [9] together with different continuity conditions on the second Fréchet derivative F'' of F such as F'' is Lipschitz or Hölder continuous to provider a semilocal convergence analysis for third order methods such as Halley's, Chebyshev's method, super-Halley's method and other high order methods. The sufficient conditions usually associated with the semilocal convergence of Halley's method are the (C) conditions [1, 5] given by

(C_1) $\|F'(x_0)^{-1}F(x_0)\| \leq \eta$,

(C_2) $\|F'(x_0)^{-1}F''(x)\| \leq \beta$,

(C_3) $\|F'(x_0)^{-1}(F''(x) - F''(y))\| \leq M\|x - y\|$,

(C_4) $h = \frac{3M^2}{(\beta^2+2M)^{\frac{3}{2}} - \beta(\beta^2+3M)}\eta \leq 1$,

(C_5) $\overline{U}(x_0, R_0) \subseteq D$, where R_0 is the small positive root of

$$p(t) = \frac{M}{6}t^3 + \frac{\beta}{2}t^2 - t + \eta.$$

Similar conditions but with different (C_4) and (C_5) have been given by us in [5, Theorem 2.3], where the corresponding to R_0 radius is given in closed form. There are many interesting examples in the literature (see [3, 4, 11, 15] and Example 3.5.2), where Lipschitz condition (C_3)(used in [9, 13, 14]) is violated but center-Lipschitz condition

$$\|F'(x_0)^{-1}[F''(x) - F''(x_0)]\| \leq L\|x - x_0\|, \quad for\ each\ x \in D \tag{3.1.3}$$

is satisfied. Note that

$$L \leq M$$

holds in general and $\frac{M}{L}$ can be arbitrarily large [4, 5]. A local convergence analysis for Halley's method under (1.3) and more general conditions has been given by us in [5, 6]. Relevant work but for Newton's method can be found in [3, 4, 11].

The following semilocal convergence theorem was established in [15].

Theorem 3.1.1. *Let $F : D \subset X \to Y$ be continuously twice Fréchet differentiable, D open and convex. Assume that there exists a starting point $x_0 \in D$ such that $F'(x_0)^{-1}$ exists, and the following conditions hold:*

(i) $\|F'(x_0)^{-1}F(x_0)\| \leq \eta$;

(ii) $\|F'(x_0)^{-1}F''(x_0)\| \leq \beta$;

(iii) (1.3) is true;

(iv) $\frac{1}{2}\beta\eta < \tau$, *where*

$$\tau = \frac{3s^\star + 1 - \sqrt{7s^\star + 1}}{9s^\star - 1} = 0.134065\ldots, \tag{3.1.4}$$

$s^\star = 0.800576\ldots$ *such that* $q(s^\star) = 1$, *and*

$$q(s) = \frac{(6s+2) - 2\sqrt{7s+1}}{(6s-2) + \sqrt{7s+1}}(1 + \frac{s}{1-s^2}); \tag{3.1.5}$$

(v) $\overline{U}(x_0, R) \subset D$, *where R is the positive solution of*

$$Lt^2 + \beta t - 1 = 0. \tag{3.1.6}$$

Then, the Halley sequence $\{x_k\}$ generated by (3.1.2) remains in the open ball $U(x_0,R)$, and converges to the unique solution $x^\star \in \overline{U}(x_0,R)$ of Eq. (3.1.1). Moreover, the following error estimate holds

$$\|x^\star - x_k\| \leq \frac{a}{c(1-\tau)\gamma}\sum_{i=k+1}^{\infty}\gamma^{2^i}, \tag{3.1.7}$$

where $a = \beta\eta$, $c = \frac{1}{R}$ and $\gamma = \frac{a(a+4)}{(2-3a)^2}$.

In the present chapter we expand the applicability of Halley's method using (3.1.3) instead of (C_3). The chapter is organized as follows: In Section 3.2 we present a counterexample to show that the result in [15] using (3.1.3) is false. The mistakes in the proof are pointed out in Section 3.3. Section 3.4 contains our semilocal convergence analysis of Halley's method using (3.1.3). The numerical examples are given in the concluding Section 3.5.

3.2. Motivational example

Example 3.2.1. *Let us define a scalar function $F(x) = 20x^3 - 54x^2 + 60x - 23$ on $D = (0,3)$ with initial point $x_0 = 1$. Then, we have that*

$$F'(x) = 12(5x^2 - 9x + 5), \quad F''(x) = 12(10x - 9). \tag{3.2.1}$$

So, $F(x_0) = 3$, $F'(x_0) = 12$, $F''(x_0) = 12$. We can choose $\eta = \frac{1}{4}$ and $\beta = 1$ in Theorem 3.1.1. Moreover, we have for any $x \in D$ that

$$|F'(x_0)^{-1}[F''(x) - F''(x_0)]| = 10|x - x_0|. \tag{3.2.2}$$

Hence, the center Lipschitz condition (3.1.3) is true for constant $L = 10$. We can also verify condition $\frac{1}{2}\beta\eta = \frac{1}{8} < \tau = 0.134065\ldots$ is true. By (3.1.6), we get

$$R = \frac{\sqrt{\beta^2 + 4L} - \beta}{2L} = \frac{\sqrt{41} - 1}{20} = 0.270156\ldots. \tag{3.2.3}$$

Then, condition $\overline{U}(x_0,R) = [x_0 - R, x_0 + R] \approx [0.729844, 1.270156] \subset D$ is also true. Hence, all conditions in Theorem 3.1.1 are satisfied. However, we can verify that the point x_1 generated by the Halley's method (3.1.2) doesn't remain in the open ball $U(x_0,R)$. In fact, we have that

$$|x_1 - x_0| = \frac{|F'(x_0)^{-1}F(x_0)|}{|1 - \frac{1}{2}F'(x_0)^{-1}F''(x_0)F'(x_0)^{-1}F(x_0)|} = \frac{2}{7} = 0.285714\ldots > R. \tag{3.2.4}$$

Clearly, the rest conclusions of Theorem 3.1.1 cannot be reached.

3.3. Mistakes in the Proof of Theorem 3.1.1

One crucial mistake exists in the proof of Theorem 3.1.1. To show this, let us introduce real constants a, b and c, and real sequences $\{a_k\}, \{b_k\}, \{c_k\}$ and $\{d_k\}$ given in Ref [15] as follows:

$$0 < a < 2\tau, \quad b > 0, \quad c > 0, \quad 2bc < 2 - a, \tag{3.3.1}$$

$$a_0 = 1, \quad b_0 = b, \quad c_0 = \tfrac{a}{2}, \quad d_0 = \tfrac{b}{1-\frac{a}{2}}, \tag{3.3.2}$$

and

$$\begin{cases} a_{k+1} = \frac{a_k}{1-ca_kd_k}, \\ b_{k+1} = c(1+\frac{c_k}{2})d_k^2, \\ c_{k+1} = \frac{c}{2}a_{k+1}^2 b_{k+1}, \\ d_{k+1} = \frac{a_{k+1}b_{k+1}}{1-c_{k+1}} \end{cases} \tag{3.3.3}$$

for all $k \geq 0$. In addition, author in [15] sets

$$r_k = d_0 + d_1 + \cdots + d_k = \sum_{i=0}^{k} d_i \tag{3.3.4}$$

and

$$r = \lim_{k\to\infty} r_k \tag{3.3.5}$$

provided the limit exists. Using induction, Ref. [15] shows that the following relation holds for $k \geq 0$ if $a_k \geq 1$:

$$a_{k+1} = \frac{1}{1-cr_k}. \tag{3.3.6}$$

Next Ref. [15] claims that from the initial relations, it follows by induction that, for $k = 0, 1, 2, \ldots$

$$ca_kd_k = \frac{ca_k^2 b_k}{1-c_k} = \frac{2c_k}{1-c_k}. \tag{3.3.7}$$

Here, we point out that the relation (3.3.7) is not always true. In fact, for $k = 0$, the first and second equality of (3.3.7) are obtained from

$$d_0 = \frac{a_0 b_0}{1-c_0} \tag{3.3.8}$$

and

$$c_0 = \frac{c}{2}a_0^2 b_0, \tag{3.3.9}$$

respectively. We can easily verify (3.3.8) is really true from (3.3.2). However, (3.3.9) is not true in general. Otherwise, using (3.3.2) and (3.3.9), we demand that

$$a = bc. \tag{3.3.10}$$

Clearly, this condition is introduced improperly and will be violated frequently. Since some lemmas of Ref. [15] are established on the basis of the above basic relation (3.3.7), they will be not always true. Therefore, the main theorem of Ref. [15] (Theorem 3.1.1 stated above) will be not always true, because it is based on these lemmas.

3.4. New Semilocal Convergence Theorem

In this section, we will give a new semilocal convergence theorem for Halley's method under the center Lipschitz condition. We first need some auxiliary lemmas.

Lemma 3.4.1. *Function $q(s)$ given by*

$$q(s) = \tfrac{2s}{1-s}\Big(1 + \frac{\frac{s(s+2)}{(1-3s)^2}}{1-(\frac{s(s+2)}{(1-3s)^2})^2}\Big) \tag{3.4.1}$$

increases monotonically on $[0, \frac{2-\sqrt{2}}{4})$ *and there exists a unique point* $\tau \approx 0.134065 \in (0, \frac{2-\sqrt{2}}{4})$ *such that* $q(\tau) = 1$.

Proof. It is easy to verify function $h(s)$ given by

$$h(s) = \tfrac{s(s+2)}{(1-3s)^2}. \tag{3.4.2}$$

increases monotonically on $[0, \frac{1}{3})$, and such that $h(\frac{2-\sqrt{2}}{4}) = 1$ and $h(s) < 1$ is true on $[0, \frac{2-\sqrt{2}}{4})$. By (4.1), $q(s)$ increases monotonically on $[0, \frac{2-\sqrt{2}}{4})$. Since $q(0) = 0$ and $q(s)$ tends to $+\infty$ as s tends to $\frac{2-\sqrt{2}}{4}$ from the left side. Therefore, there exists a unique point $\tau \in (0, \frac{2-\sqrt{2}}{4})$ such that $q(\tau) = 1$. We can use iterative methods such as the Secant method to obtain: $\tau \approx 0.134065$. The proof is complete. □

Lemma 3.4.2. *Let real constants a, b and c be defined by*

$$a > 0, \quad b > 0, \quad c > 0, \quad 2bc < 2 - a, \tag{3.4.3}$$

and real sequences $\{a_k\}, \{b_k\}, \{c_k\}$ and $\{d_k\}$ be defined by (3.3.2)-(3.3.3). Assume that

$$c_1 = \frac{c^2 b^2 (a+4)}{2(2-a-2bc)^2} < \tau, \tag{3.4.4}$$

where τ is a constant defined in Lemma 3.4.1. Then, sequence $\{c_k\}$ is a bounded strictly decreasing and $c_k \in (0, \tau)$ for all $k \geq 1$. Moreover, we have that

$$c_{k+1} = \frac{c_k^2}{(1-3c_k)^2}(c_k + 2), \quad k = 1, 2, \ldots. \tag{3.4.5}$$

Proof. By conditions (3.4.3) and (3.4.4), a_1, b_1, c_1 and d_1 are well defined. Since,

$$c a_1 d_1 = \frac{c a_1^2 b_1}{1-c_1} = \frac{2c_1}{1-c_1} < 1, \tag{3.4.6}$$

a_2, b_2 and c_2 are well defined. We have that

$$c_2 = \frac{c}{2} \frac{a_1^2}{(1 - \frac{2c_1}{1-c_1})^2} c(1 + \frac{c_1}{2}) \frac{a_1^2 b_1^2}{(1-c_1)^2} = \frac{c_1^2}{(1-3c_1)^2}(c_1 + 2). \tag{3.4.7}$$

Hence, (3.4.5) holds for $k=1$. By (3.4.7), $c_2 < c_1 < 1$ is true, since

$$(1-3c_1)^2 - c_1(c_1+2) = 8c_1^2 - 8c_1 + 1 > 0, \tag{3.4.8}$$

which is equivalent to

$$c_1 < \frac{2-\sqrt{2}}{4}. \tag{3.4.9}$$

Then, d_2 is well defined.

Suppose sequences $\{a_k\},\{b_k\},\{c_k\}$ and $\{d_k\}$ are well defined for $k=0,1,\ldots,n+1$, $c_{n+1} < c_n < \cdots < c_2 < c_1$, and (3.4.5) holds for $k=1,2,\ldots,n$, where $n \geq 1$ is a fixed integer. Since,

$$ca_{n+1}d_{n+1} = \frac{ca_{n+1}^2 b_{n+1}}{1-c_{n+1}} = \frac{2c_{n+1}}{1-c_{n+1}} < 1, \tag{3.4.10}$$

it follows that a_{n+2}, b_{n+2} and c_{n+2} are well defined. We have that

$$c_{n+2} = \frac{c}{2}\frac{a_{n+1}^2}{(1-\frac{2c_{n+1}}{1-c_{n+1}})^2}c(1+\frac{c_{n+1}}{2})\frac{a_{n+1}^2 b_{n+1}^2}{(1-c_{n+1})^2} = \frac{c_{n+1}^2}{(1-3c_{n+1})^2}(c_{n+1}+2), \tag{3.4.11}$$

thus (4.5) holds for $k=n+1$. So, $c_{n+2} < c_{n+1} < \cdots < c_2 < c_1 < 1$, and d_{n+2} is well defined. That completes the induction and the proof of the lemma. □

Lemma 3.4.3. *Under the assumptions of Lemma 3.4.2, if we set* $\bar{\gamma} = \frac{c_2}{c_1}$*, then for* $k \geq 0$

(i) $c_{k+1} \leq c_1 \bar{\gamma}^{2^k-1}$*;*
(ii) $d_{k+1} \leq \frac{2c_1}{ca_1(1-c_1)}\bar{\gamma}^{2^k-1}$*;*
(iii) $r - r_k \leq \sum_{i=k+1}^{\infty} \frac{2c_1}{ca_1(1-c_1)}\bar{\gamma}^{2^{i-1}-1}$*;*
where r is defined in (3.3.5).

Proof. Obviously, (i) is true for $k=0$. By Lemma 3.4.2, for any $k \geq 1$, we have that

$$c_{k+1} = \frac{c_k^2}{(1-3c_k)^2}(c_k+2) \leq \frac{c_k^2}{(1-3c_{k-1})^2}(c_{k-1}+2) \leq \cdots \leq \frac{c_k^2}{(1-3c_1)^2}(c_1+2) = \lambda c_k^2, \tag{3.4.12}$$

where

$$\lambda = \frac{c_1+2}{(1-3c_1)^2} = \frac{c_1^2(c_1+2)}{(1-3c_1)^2 c_1^2} = \frac{c_2}{c_1^2} = \frac{\bar{\gamma}}{c_1}. \tag{3.4.13}$$

Multiplying both side of (3.4.12) by λ yields

$$\lambda c_{k+1} \leq (\lambda c_k)^2 \leq (\lambda c_{k-1})^{2^2} \leq \cdots \leq (\lambda c_1)^{2^k} = \bar{\gamma}^{2^k}, \tag{3.4.14}$$

which shows (i). Consequently, for $k \geq 1$, we have

$$d_{k+1} = \frac{ca_{k+1}d_{k+1}}{ca_{k+1}} = \frac{ca_{k+1}^2 b_{k+1}}{ca_{k+1}(1-c_{k+1})} = \frac{2c_{k+1}}{ca_{k+1}(1-c_{k+1})} \leq \frac{2c_1\bar{\gamma}^{2^k-1}}{ca_1(1-c_1)}. \tag{3.4.15}$$

That is, (ii) is true for $k \geq 1$. For the case of $k=0$, we have that

$$d_1 = \frac{a_1 b_1}{1-c_1} = \frac{ca_1^2 b_1}{ca_1(1-c_1)} = \frac{2c_1}{ca_1(1-c_1)}, \tag{3.4.16}$$

which means (ii) is also true for $k=0$. Moreover, (iii) is true by using (ii) and the definitions of r_k and r. The proof is complete. □

Lemma 3.4.4. *Under the assumptions of Lemma 3.4.2, if we set $a=\beta\eta$, $b=\eta$, $c=\frac{1}{R}$, then $r=\lim_{k\to\infty} r_k < R$.*

Proof. By the definition of r_k, Lemmas 3.4.1-3.4.3, for any $k\geq 1$, we have that

$$\begin{aligned} r_k &\leq d_0+\sum_{i=1}^{k}\frac{2c_1}{ca_1(1-c_1)}\overline{\gamma}^{2^{i-1}-1} < d_0+\frac{2Rc_1}{a_1(1-c_1)}\left(1+\frac{\overline{\gamma}}{1-\overline{\gamma}^2}\right)\\ &= d_0+\frac{2(1-cd_0)Rc_1}{1-c_1}\left(1+\frac{\frac{c_1(c_1+2)}{(1-3c_1)^2}}{1-(\frac{c_1(c_1+2)}{(1-3c_1)^2})^2}\right)=d_0+q(c_1)(R-d_0)\\ &< d_0+q(\tau)(R-d_0)=R. \end{aligned} \tag{3.4.17}$$

Here, we used $R=\frac{1}{c}>\frac{b}{1-\frac{a}{2}}=d_0$ by (3.4.3). Hence, $r=\lim_{k\to\infty} r_k$ exists and $r<R$. The proof is complete. □

Lemma 3.4.5. *Set $a=\beta\eta$, $b=\eta$ and $c=\frac{1}{R}$. Let $\{a_k\},\{b_k\},\{c_k\},\{d_k\}$ be the sequences generated by (3.3.2)-(3.3.3). Suppose that conditions (3.4.3) and (3.4.4) are true. Then for any $k\geq 0$, we have*
(i) $F'(x_k)^{-1}$ exists and $\|F'(x_k)^{-1}F'(x_0)\|\leq a_k$;
(ii) $\|F'(x_0)^{-1}F(x_k)\|\leq b_k$;
(iii) $[I-L_F(x_k)]^{-1}$ exists and $\|L_F(x_k)\|\leq c_k$;
(iv) $\|x_{k+1}-x_k\|\leq d_k$;
(v) $\|x_{k+1}-x_0\|\leq r_k<R$.

Proof. The proof is similar to the one in [15], and we shall omit it. □

Now, we can state our main theorem.

Theorem 3.4.6. *Let $F: D\subset X\to Y$ be continuously twice Fréchet differentiable, D open and convex. Assume that there exists a starting point $x_0\in D$ such that $F'(x_0)^{-1}$ exists, and the following conditions hold:*
(i) $\|F'(x_0)^{-1}F(x_0)\|\leq\eta$;
(ii) $\|F'(x_0)^{-1}F''(x_0)\|\leq\beta$;
(iii) the center Lipschitz-condition (3.1.3) is true;
(iv) conditions (3.4.3) and (3.4.4) are true;
(v) $\overline{U}(x_0,R)\subset D$, where R is the positive solution of (3.1.6).
Then, the Halley sequence $\{x_k\}$ generated by (3.1.2) remains in the open ball $U(x_0,R)$, and converges to the unique solution $x^\star\in\overline{U}(x_0,R)$ of Eq. (3.1.1). Moreover, the following error estimate holds for any $k\geq 1$

$$\|x^\star-x_k\|\leq\sum_{i=k}^{\infty}\frac{2c_1}{ca_1(1-c_1)}\overline{\gamma}^{2^{i-1}-1}, \tag{3.4.18}$$

where $a=\beta\eta, b=\eta, c=\frac{1}{R}$ and $\overline{\gamma}=\frac{c_2}{c_1}=\frac{c_1(c_1+2)}{(1-3c_1)^2}$.

Proof. Using Lemma 3.4.5, we have for all $k\geq 0$, $x_k\in U(x_0,R)$. From Lemma 3.4.5 and Lemma 3.4.3, we get for any integer $k,m\geq 1$

$$\begin{aligned} \|x_{k+m}-x_k\| &\leq \sum_{i=k}^{k+m-1}\|x_{i+1}-x_i\|\leq\sum_{i=k}^{k+m-1}d_i\leq\sum_{i=k}^{k+m-1}\frac{2c_1}{ca_1(1-c_1)}\overline{\gamma}^{2^{i-1}-1}\\ &\leq \frac{2c_1}{ca_1(1-c_1)}\sum_{i=k}^{\infty}\overline{\gamma}^{2^{i-1}-1}\leq\frac{2c_1}{ca_1(1-c_1)\overline{\gamma}}\frac{\overline{\gamma}^{2^{k-1}}}{1-\overline{\gamma}^2}. \end{aligned} \tag{3.4.19}$$

That is, $\{x_k\}$ is a Cauchy sequence. So, there exists a point $x^\star \in \overline{U}(x_0,R)$ such that $\{x_k\}$ converges to $x^\star$ as $k \to \infty$. Using Lemma 3.4.3, clearly we have $d_k \to 0$ as $k \to \infty$. Using Lemma 3.4.5 and Lemma 3.4.2, for any $k \geq 0$, we have

$$\|F'(x_0)^{-1}F(x_{k+1})\| \leq b_{k+1} = c(1+\tfrac{c_k}{2})d_k^2 \leq c(1+\tfrac{c_1}{2})d_k^2 \to 0 \quad as\, k \to \infty. \tag{3.4.20}$$

The continuity of F gives

$$\|F'(x_0)^{-1}F(x^\star)\| = \lim_{k\to\infty}\|F'(x_0)^{-1}F(x_{k+1})\| = 0 \quad as\, k \to \infty, \tag{3.4.21}$$

that is $F(x^\star) = 0$. By let $m \to \infty$ in (3.4.19), (3.4.18) is obtained immediately.

Finally, we can show the uniqueness of $x^\star \in U(x_0,R)$ by using the same technique as in [2, 3, 4, 5, 15]. The proof is complete. □

Remark 3.4.7. (a) Let us compare our sufficient convergence condition (3.4.3) with condition (C_4). Condition (3.4.3) can be rewritten as

$$h_0 = \frac{2\beta+\sqrt{\beta^2+4L}}{2}\eta < 1 \tag{3.4.22}$$

if we use the choices of a,b,c given in Lemma 3.4.5 and R given by (3.2.3). Then, we have that

$$h_0 \leq h. \tag{3.4.23}$$

Estimate (3.4.23) shows that one of our convergence conditions is at least as weak as (C4). However a direct comparison between (3.4.4) and (C4) is not practical. A similar favorable comparison can be followed with all other sufficient convergence conditions of the form (C_4) already in the literature using M instead of L (see [4, 5, 10, 11, 13, 14, 15]) and the references therein).

(b) It is possible that (C_3) is satisfied (hence, (3.1.3) too) but not (C_4) (or (C_5)). In this case we test to see if our conditions are satisfied. If our conditions are satisfied although we predict only quadratic convergence of the Halley method (3.1.2) (see e.g. Lemma 3.4.3) after a certain iterate x_N, where N is a finite natural integer (C_4) and (C_5) will be satisfied for $x_0 = x_N$. Therefore, the usual error estimates for the cubical convergence of the Halley method (3.1.2) will hold. We refer the reader to [3, 4], where we show how to choose N in the case of Newton's method. The N for Halley's method (3.1.2) can be found in an analogous way.

3.5. Numerical Examples

In this section, we will give some examples to show the application of our Theorem 3.4.6.

Example 3.5.1. *Let us define a scalar function $F(x) = x^3 - 2.25x^2 + 3x - 1.585$ on $D = (0,3)$ with initial point $x_0 = 1$. Then, we have that*

$$F'(x) = 3x^2 - 4.5x + 3, \quad F''(x) = 6x - 4.5. \tag{3.5.1}$$

So, $F(x_0) = 0.165$, $F'(x_0) = 1.5$, $F''(x_0) = 1.5$. We can choose $\eta = 0.11$ and $\beta = 1$ in Theorem 3.4.6. Moreover, we have for any $x \in D$ that

$$|F'(x_0)^{-1}[F''(x) - F''(x_0)]| = 4|x - x_0|. \tag{3.5.2}$$

Hence, the weak Lipschitz condition (3.1.3) is true for constant $L = 4$. *By (3.1.6), we get*

$$R = \frac{\sqrt{\beta^2 + 4L} - \beta}{2L} = \frac{\sqrt{17} - 1}{8} = 0.390388\ldots. \tag{3.5.3}$$

Then, condition $\overline{U}(x_0, R) = [x_0 - R, x_0 + R] \approx [0.609612, 1.390388] \subset D$ *is true. We can also verify conditions* $2 - a - 2bc = 2 - \beta\eta - 2\eta/R \approx 1.326458 > 0$ *and* $c_1 = \frac{c^2b^2(a+4)}{2(2-a-2bc)^2} \approx 0.092729212 < \tau = 0.134065\ldots$ *is true. Hence, all conditions in Theorem 3.4.6 are satisfied.*

Example 3.5.2. *In this example we provide an application of our results to a special nonlinear Hammerstein integral equation of the second kind. Consider the integral equation*

$$u(s) = f(s) + \lambda \int_{a'}^{b'} k(s,t) u(t)^{2+\frac{1}{n}} dt, \quad \lambda \in \mathbb{R}, n \in \mathbb{N}, \tag{3.5.4}$$

where f *is a given continuous function satisfying* $f(s) > 0$ *for* $s \in [a', b']$ *and the kernel is continuous and positive in* $[a', b'] \times [a', b']$.

Let $X = Y = C[a', b']$ *and* $D = \{u \in C[a', b'] : u(s) \geq 0, s \in [a', b']\}$. *Define* $F : D \to Y$ *by*

$$F(u)(s) = u(s) - f(s) - \lambda \int_{a'}^{b'} k(s,t) u(t)^{2+\frac{1}{n}} dt, \quad s \in [a', b']. \tag{3.5.5}$$

We use the max-norm, The first and second derivatives of F *are given by*

$$F'(u)v(s) = v(s) - \lambda(2 + \frac{1}{n}) \int_{a'}^{b'} k(s,t) u(t)^{1+\frac{1}{n}} v(t) dt, \quad v \in D, s \in [a', b'], \tag{3.5.6}$$

and

$$F''(u)(vw)(s) = -\lambda(1 + \frac{1}{n})(2 + \frac{1}{n}) \int_{a'}^{b'} k(s,t) u(t)^{\frac{1}{n}} (vw)(t) dt, \quad v, w \in D, s \in [a', b'], \tag{3.5.7}$$

respectively.

Let $x_0(t) = f(t)$, $\alpha = \min_{s \in [a', b']} f(s)$, $\delta = \max_{s \in [a', b']} f(s)$ *and* $M = \max_{s \in [a', b']} \int_{a'}^{b'} |k(s,t)| dt$. *Then, for any* $v, w \in D$,

$$\begin{aligned}
&\|[F''(x) - F''(x_0)](vw)\| \leq |\lambda|(1 + \tfrac{1}{n})(2 + \tfrac{1}{n}) \max_{s \in [a', b']} \textstyle\int_{a'}^{b'} |k(s,t)| |x(t)^{\frac{1}{n}} - f(t)^{\frac{1}{n}}| dt \|vw\| \\
&= |\lambda|(1 + \tfrac{1}{n})(2 + \tfrac{1}{n}) \max_{s \in [a', b']} \textstyle\int_{a'}^{b'} |k(s,t)| \frac{|x(t) - f(t)|}{x(t)^{\frac{n-1}{n}} + x(t)^{\frac{n-2}{n}} f(t)^{\frac{1}{n}} + \cdots + f(t)^{\frac{n-1}{n}}} dt \|vw\| \\
&\leq |\lambda|(1 + \tfrac{1}{n})(2 + \tfrac{1}{n}) \max_{s \in [a', b']} \textstyle\int_{a'}^{b'} |k(s,t)| \frac{|x(t) - f(t)|}{f(t)^{\frac{n-1}{n}}} dt \|vw\| \\
&\leq \frac{|\lambda|(1+\frac{1}{n})(2+\frac{1}{n})}{\alpha^{\frac{n-1}{n}}} \max_{s \in [a', b']} \textstyle\int_{a'}^{b'} |k(s,t)| |x(t) - f(t)| dt \|vw\| \\
&\leq \frac{|\lambda|(1+\frac{1}{n})(2+\frac{1}{n})M}{\alpha^{\frac{n-1}{n}}} \|x - x_0\| \|vw\|,
\end{aligned} \tag{3.5.8}$$

which means

$$\|F''(x) - F''(x_0)\| \leq \frac{|\lambda|(1+\frac{1}{n})(2+\frac{1}{n})M}{\alpha^{\frac{n-1}{n}}} \|x - x_0\|. \tag{3.5.9}$$

Next, we give a bound for $\|F'(x_0)^{-1}\|$*. Using (3.5.6), we have that*

$$\|I - F'(x_0)\| \leq |\lambda|(2+\frac{1}{n})\delta^{1+\frac{1}{n}}M. \tag{3.5.10}$$

It follows from the Banach theorem that $F'(x_0)^{-1}$ *exists if* $|\lambda|(2+\frac{1}{n})\delta^{1+\frac{1}{n}}M < 1$*, and*

$$\|F'(x_0)^{-1}\| \leq \frac{1}{1-|\lambda|(2+\frac{1}{n})\delta^{1+\frac{1}{n}}M}. \tag{3.5.11}$$

On the other hand, we have from (3.5.5) and (3.5.7) that $\|F(x_0)\| \leq |\lambda|\delta^{2+\frac{1}{n}}M$ *and* $\|F''(x_0)\| \leq |\lambda|(1+\frac{1}{n})(2+\frac{1}{n})\delta^{\frac{1}{n}}M$*. Hence, if* $|\lambda|(2+\frac{1}{n})\delta^{1+\frac{1}{n}}M < 1$*, the weak Lipschitz condition (1.3) is true for*

$$L = \frac{|\lambda|(1+\frac{1}{n})(2+\frac{1}{n})M}{\alpha^{\frac{n-1}{n}}[1-|\lambda|(2+\frac{1}{n})\delta^{1+\frac{1}{n}}M]} \tag{3.5.12}$$

and constants η *and* β *in Theorem 3.4.6 can be given by*

$$\eta = \frac{|\lambda|\delta^{2+\frac{1}{n}}M}{1-|\lambda|(2+\frac{1}{n})\delta^{1+\frac{1}{n}}M}, \quad \beta = \frac{|\lambda|(1+\frac{1}{n})(2+\frac{1}{n})\delta^{\frac{1}{n}}M}{1-|\lambda|(2+\frac{1}{n})\delta^{1+\frac{1}{n}}M}. \tag{3.5.13}$$

Next we let $[a',b'] = [0,1]$*,* $n = 2$*,* $f(s) = 1$*,* $\lambda = 1.1$ *and* $k(s,t)$ *is the Green's kernel on* $[0,1]\times[0,1]$ *defined by*

$$G(s,t) = \begin{cases} t(1-s), & t \leq s; \\ s(1-t), & s \leq t. \end{cases} \tag{3.5.14}$$

Consider the following particular case of (3.5.4):

$$u(s) = f(s) + 1.1\int_0^1 G(s,t)u(t)^{\frac{5}{2}}dt, \quad s \in [0,1]. \tag{3.5.15}$$

Then, $\alpha = \delta = 1$ *and* $M = \frac{1}{8}$*. Moreover, we have that*

$$\eta = \frac{22}{105}, \quad \beta = \frac{11}{14}, \quad L = \frac{11}{14}. \tag{3.5.16}$$

Therefore $2-a-2bc \approx 1.264456 > 0$*,* $\tau - c_1 \approx 0.027938 > 0$ *and* $R \approx 0.733988$*. Hence,* $\overline{U}(x_0,R) \subset D$*. Thus, all conditions of Theorem 3.4.6 are satisfied. Consequently, sequence* $\{x_k\}$ *generated by Halley's method (3.1.2) with initial point* x_0 *converges to the unique solution* $x^\star$ *of Eq. (3.5.15) on* $\overline{U}(x_0,0.733988)$*. The Lipschitz condition* (C_3) *is not satisfied [3, 4, 11, 15]. Hence, we have expanded the applicability of Halley's method. Note also that verifying (3.1.3) is less expensive than verifying* (C_3)*.*

References

[1] Amat, S., Busquier, S, Third-order iterative methods under Kantorovich conditions, *J. Math. Anal. Appl.* 336 (2007), 243–261.

[2] Argyros, I.K., The convergence of Halley-Chebyshev type method under Newton-Kantorovich hypotheses, *Appl. Math. Lett.* 6 (1993), 71–74.

[3] Argyros, I.K., Approximating solutions of equations using Newton's method with a modified Newton's method iterate as a starting point, *Rev. Anal. Numer. Theor. Approx*, 36 (2007), 123–138.

[4] Argyros, I.K., Computational theory of iterative methods, *Series: Studies in Computational Mathematics* 15, Editors, C.K. Chui and L. Wuytack, Elservier Publ. Co. New York, USA, 2007.

[5] Argyros, I.K., Cho, Y.J., Hilout, S., On the semilocal convergence of the Halley method using recurrent functions, *J. Appl. Math. Computing.* 37 (2011), 221–246.

[6] Argyros, I.K., Ren, H.M., Ball convergence theorems for Halley's method in Banach spaces, *J. Appl. Math. Computing* 38 (2012), 453–465.

[7] Argyros, I. K., and S. George, Ball convergence for Steffensen-type fourth-order methods. *Int. J. Interac. Multim. Art. Intell.*, 3(4) (2015), 37–42.

[8] Argyros, I. K., and González, D., Local convergence for an improved Jarratt-type method in Banach space. *Int. J. Interac. Multim. Art. Intell.*, 3(4) (2015), 20–25.

[9] Chun, C., Stănică, P., Neta, B., Third-order family of methods in Banach spaces, *Comp. Math. with App.*, 61 (2011), 1665–1675.

[10] Deuflhard, P., *Newton Methods for Nonlinear Problems: Affine Invariance and Adaptive Algorithms*, Springer-Verlag, Berlin, Heidelberg, 2004.

[11] Gutiérrez, J.M., Hernández, M.A., Newton's method under weak Kantorovich conditions, *IMA J. Numer. Anal.* 20 (2000), 521–532.

[12] Magreñán, Á. A. , Different anomalies in a Jarratt family of iterative root-finding methods, *App. Math. Comp.* 233 (2014), 29–38.

[13] Parida, P.K., Study of third order methods for nonlinear equations in Banach spaces, PhD dissertation, 2007, IIT kharagpur, India.

[14] Parida, P.K., Gupta, P.K., Semilocal convergence of a family of third-order Chebyshev-type methods under a mild differentiability condition. *Int. J. Comput. Math.* 87 (2010), 3405–3419.

[15] Xu, X.B., Ling, Y.H., Semilocal convergence for Halley's method under weak Lipschitz condition, *App. Math. Comp.* 215 (2009), 3057–3067.

Chapter 4

An Improved Convergence Analysis of Newton's Method for Twice Fréchet Differentiable Operators

4.1. Introduction

In this chapter, we are concerned with the problem of approximating a locally unique solution $x^\star$ of equation

$$\mathcal{F}(x) = 0, \tag{4.1.1}$$

where, $\mathcal{F}$ is a twice Fréchet differentiable operator defined on a convex subset $\mathcal{D}$ of a Banach space $\mathcal{X}$ with values in a Banach space $\mathcal{Y}$. Numerous problems in science and engineering – such as optimization of chemical processes or multiphase, multicomponent flow – can be reduced to solving the above equation [7, 8, 9, 14, 15, 16]. Consequently, solving these equations is an important scientific field of research. For most problems, finding a closed form solution for the non-linear equation (4.1.1) is not possible. Therefore, iterative solution techniques are employed for solving these equations. The study about convergence analysis of iterative methods is usually divided into two categories: semilocal and local convergence analysis. The semilocal convergence analysis is based upon the information around an initial point to give criteria ensuring the convergence of the iterative procedure. While the local convergence analysis is based on the information around a solution to find estimates of the radii of convergence balls.

The most popular iterative method for solving problem (4.1.1) is the Newton's method

$$x_{n+1} = x_n - \mathcal{F}'(x_n)^{-1}\mathcal{F}(x_n) \quad \text{for each} \quad n = 0, 1, 2, \ldots, \tag{4.1.2}$$

where $x_0 \in \mathcal{D}$ is an initial point. There exists extensive local as well as semilocal convergence analysis results under various Lipschitz type conditions for Newton's method (4.1.2) [1–17]. The following four conditions have been used to perform semilocal convergence analysis of Newton's method (4.1.2) [3, 5, 7, 8, 9, 13, 14]

C$_1$. there exists $x_0 \in \mathcal{D}$ such that $\mathcal{F}'(x_0)^{-1} \in \mathbf{L}(\mathcal{Y}, \mathcal{X})$,

C$_2$. $\left\|\mathcal{F}'(x_0)^{-1}\mathcal{F}(x_0)\right\| \leq \eta$,

C$_3$. $\left\|\mathcal{F}'(x_0)^{-1}\mathcal{F}''(x)\right\| \leq \mathcal{K}$ for each $x \in \mathcal{D}$,

C$_4$. $\left\|\mathcal{F}'(x_0)^{-1}(\mathcal{F}''(x)-\mathcal{F}''(y))\right\| \leq \mathcal{M}\,\|x-y\|$ for each $x, y \in \mathcal{D}$.

Let us also introduce the center-Lipschitz condition

C$_5$. $\left\|\mathcal{F}'(x_0)^{-1}(\mathcal{F}'(x)-\mathcal{F}'(x_0))\right\| \leq \mathcal{L}_0\,\|x-x_0\|$ for each $x \in \mathcal{D}$.

We shall refer to $(\mathbf{C}_1)$ – $(\mathbf{C}_5)$ as the $(\mathbf{C})$ conditions. The following conditions have also been employed [9, 10, 11, 12, 17, 14]

C$_6$. $\|\mathcal{F}'(x_0)^{-1}\mathcal{F}''(x_0)\| \leq \mathcal{K}_0$

C$_7$. $\left\|\mathcal{F}'(x_0)^{-1}(\mathcal{F}''(x)-\mathcal{F}''(x_0))\right\| \leq \mathcal{M}_0\,\|x-x_0\|$ for each $x \in \mathcal{D}$.

Here onwards, the conditions ($\mathbf{C}_1$), ($\mathbf{C}_2$), ($\mathbf{C}_5$), ($\mathbf{C}_6$), ($\mathbf{C}_7$) are referred as the (**H**) conditions.

For the semilocal convergence of Newton's method the conditions ($\mathbf{C}_1$), ($\mathbf{C}_2$), ($\mathbf{C}_3$) together with the following sufficient conditions are given [1, 2, 3, 4, 9, 10, 11, 12, 17, 14, 15, 16, 18]

$$\eta \leq \frac{4\mathcal{M}+\mathcal{K}^2-\mathcal{K}\sqrt{\mathcal{K}^2+2\mathcal{M}}}{3\mathcal{M}(\mathcal{K}+\sqrt{\mathcal{K}^2+2\mathcal{M}})}, \tag{4.1.3}$$

$$\overline{U}(x_0,R_1) \subseteq \mathcal{D} \tag{4.1.4}$$

where R_1 is the smallest positive root of

$$\mathcal{P}_1(t) = \frac{\mathcal{M}}{6}t^3 + \frac{\mathcal{K}}{2}t^2 - t + \eta. \tag{4.1.5}$$

Whereas the conditions ($\mathbf{C}_1$), ($\mathbf{C}_2$), ($\mathbf{C}_6$), ($\mathbf{C}_7$) together with

$$\eta \leq \frac{4\mathcal{M}_0+\mathcal{K}_0^2-\mathcal{K}_0\sqrt{\mathcal{K}_0^2+2\mathcal{M}_0}}{3\mathcal{M}_0(\mathcal{K}_0+\sqrt{\mathcal{K}_0^2+2\mathcal{M}_0})} \tag{4.1.6}$$

$$\overline{U}(x_0,R_2) \subseteq \mathcal{D} \tag{4.1.7}$$

where R_2 is the small positive root of

$$\mathcal{P}_2(t) = \frac{\mathcal{M}_0}{6}t^3 + \frac{\mathcal{K}_0}{2}t^2 - t + \eta. \tag{4.1.8}$$

have also been used for the semilocal convergence of Newton's method. Conditions (4.1.3) and (4.1.6) cannot be directly be compared with ours given in Sections 4.2 and 4.3, since we use $\mathcal{L}_0$ that does not appear in (4.1.3) and (4.1.6). However, comparisons can be made on concrete numerical examples. Let us consider $\mathcal{X} = \mathcal{Y} = \mathbb{R}$, $x_0 = 1$ and $\mathcal{D} = [\zeta, 2-\zeta]$ for $\zeta \in (0,1)$. Define function $\mathcal{F}$ on $\mathcal{D}$ by

$$\mathcal{F}(x) = x^5 - \zeta. \tag{4.1.9}$$

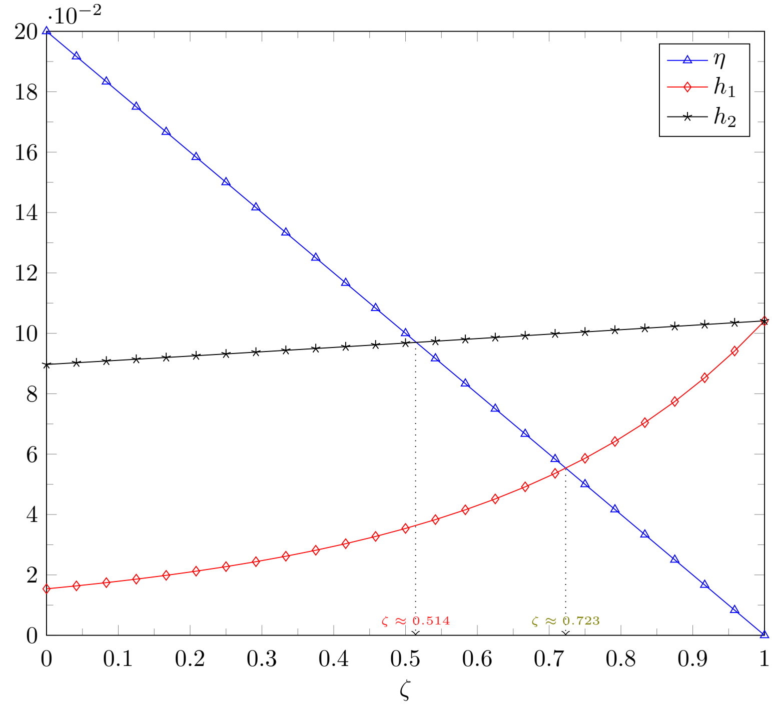


Figure 4.1.1. Convergence criteria (4.1.3) and (4.1.6) for the equation (4.1.9). Here, h_1 and h_2 stands respectively for the right hand side of the conditions (4.1.3) and (4.1.6).

Then, through some simple calculations, the conditions ($\mathbf{C}_2$), ($\mathbf{C}_3$), ($\mathbf{C}_4$), ($\mathbf{C}_5$), ($\mathbf{C}_6$) and ($\mathbf{C}_7$) yield

$$\left.\begin{array}{c} \eta = \dfrac{(1-\zeta)}{5}, \quad \mathcal{K} = 4(2-\zeta)^3, \quad \mathcal{M} = 12(2-\zeta)^2, \quad \mathcal{K}_0 = 4, \\ \mathcal{M}_0 = 4\zeta^2 - 20\zeta + 28, \quad \mathcal{L}_0 = 15 - 17\zeta + 7\zeta^2 - \zeta^3. \end{array}\right\}$$

Figure 4.1.1 plots the criteria (4.1.3) and (4.1.6) for the problem (4.1.9). In the Figure 4.1.1, h_1 stands for the right hand side of the condition (4.1.3) and h_2 stands for the the right hand side of the condition (4.1.6). In the Figure 4.1.1, we observe that for $\zeta < 0.723$ the criterion (4.1.3) does not hold while for $\zeta < 0.514$ the criterion (4.1.6) does not hold. However, one may see that the method (4.1.2) is convergent.

In this chapter, we expand the applicability of Newton's method (4.1.2) first under the (**C**) conditions and secondly under the (**H**) conditions. The local convergence of Newton's method (4.1.2) is also performed under similar conditions.

The chapter is organized as follows. In the Section 4.2 and Section 4.3, we study majorizing sequences for the Newton's iterate $\{x_n\}$. Section 4.4 contains the semilocal convergence of Newton's method. The local convergence is given in Section 4.5. Finally, numerical examples are given in Section 4.6.

4.2. Majorizing Sequences I

In this section, we present scalar sequences and prove that these sequences are majorizing for Newton's method (4.1.2). We need the following convergence results for majorizing sequences under the (**C**) conditions.

Lemma 4.2.1. *Let* $\mathcal{K}$*,* $\mathcal{L}_0$*,* $\mathcal{M} > 0$ *and* $\eta > 0$*. Define parameters* α, η_0 *and* η_1 *by*

$$\alpha = \frac{2\mathcal{K}}{\mathcal{K} + \sqrt{\mathcal{K}^2 + 8\mathcal{L}_0\mathcal{K}}}, \tag{4.2.1}$$

$$\eta_0 = \frac{2}{\frac{\mathcal{K}}{2} + (1+\alpha)\mathcal{L}_0 + \sqrt{\left(\frac{\mathcal{K}}{2} + (1+\alpha)\mathcal{L}_0\right)^2 + \frac{2\mathcal{M}\alpha}{3}}} \tag{4.2.2}$$

and

$$\eta_1 = \frac{2\alpha}{\frac{\mathcal{K}}{2} + \alpha\mathcal{L}_0 + \sqrt{\left(\frac{\mathcal{K}}{2} + \alpha\mathcal{L}_0\right)^2 + \frac{2\mathcal{M}\alpha}{3}}}. \tag{4.2.3}$$

Suppose that

$$\eta \leq \begin{cases} \eta_1 & \text{if} \quad \mathcal{L}_0\eta \leq \dfrac{1-\alpha^2}{2+2\alpha-\alpha^2} \\ \eta_0 & \text{if} \quad \dfrac{1-\alpha^2}{2+2\alpha-\alpha^2} \leq \mathcal{L}_0\eta. \end{cases} \tag{4.2.4}$$

Then, sequence $\{t_n\}$ *generated by*

$$t_0 = 0, \quad t_1 = \eta, \quad t_{n+2} = t_{n+1} + \frac{\mathcal{K} + \frac{\mathcal{M}}{3}(t_{n+1} - t_n)}{2(1 - \mathcal{L}_0 t_{n+1})}(t_{n+1} - t_n)^2 \tag{4.2.5}$$

is well defined, increasing, bounded from above

$$t^{\star\star} = \frac{\eta}{1-\alpha} \tag{4.2.6}$$

and converges to its unique least upper bound $t^\star$ *which satisfies* $t^\star \in [\eta, t^{\star\star}]$*. Moreover the following estimates hold*

$$t_{n+1} - t_n \leq \alpha^n\eta \tag{4.2.7}$$

and

$$t^\star - t_n \leq \frac{\alpha^n\eta}{1-\alpha}. \tag{4.2.8}$$

Proof. We use mathematical induction to prove (4.2.7). Set

$$\alpha_k = \frac{\mathcal{K} + \frac{\mathcal{M}}{3}(t_{k+1} - t_k)}{2(1 - \mathcal{L}_0 t_{k+1})}. \tag{4.2.9}$$

According to (4.2.5) and (4.2.9), we must prove that

$$\alpha_k \leq \alpha. \tag{4.2.10}$$

Estimate (4.2.10) holds for $k = 0$ by (4.2.4) and the choice of η_1 given in (4.2.3). Then, we also have

$$t_2 - t_1 \leq \alpha(t_1 - t_0)$$

and

$$t_2 \leq t_1 + \alpha(t_1 - t_0) = \eta + \alpha\eta = (1+\alpha)\eta = \frac{1-\alpha^2}{1-\alpha}\eta < \frac{\eta}{1-\alpha} = t^{\star\star}.$$

Let us assume that (4.2.9) holds for all $k \leq n$. Then, we also have by (4.2.5) that

$$t_{k+1} - t_k \leq \alpha^k \eta$$

and

$$t_{k+1} \leq \frac{1-\alpha^{k+1}}{1-\alpha}\eta < t^{\star\star}.$$

Then, we must prove that

$$\left(\frac{\mathcal{K}}{2} + \frac{\mathcal{M}}{6}\alpha^k\eta\right)\alpha^k\eta + \alpha\mathcal{L}_0\frac{1-\alpha^{k+1}}{1-\alpha}\eta - \alpha \leq 0. \tag{4.2.11}$$

Estimate (4.2.11) motivates us to define recurrent functions f_k on $[0,1)$ for each $k = 1,2,\ldots$ by

$$f_k(t) = \frac{1}{2}\left(\mathcal{K} + \frac{\mathcal{M}}{3}t^k\eta\right)t^{k-1}\eta + \mathcal{L}_0(1 + t + \cdots + t^k)\eta - 1. \tag{4.2.12}$$

We need a relationship between two consecutive functions f_k. Using (4.2.12) we get that

$$f_{k+1}(t) = f_k(t) + g_k(t), \tag{4.2.13}$$

where

$$\begin{aligned} g_k(t) &= \left[\frac{1}{2}\left(\mathcal{K} + \frac{\mathcal{M}}{3}t^{k+1}\eta\right)t - \frac{1}{2}\left(\mathcal{K} + \frac{\mathcal{M}}{3}t^k\eta\right) + \mathcal{L}_0 t^2\right]t^{k-1}\eta \\ &= \left[\frac{1}{2}(2\mathcal{L}_0 t^2 + \mathcal{K}t - \mathcal{K}) + \frac{\mathcal{M}}{6}t^k\eta(t^2-1)\right]t^{k-1}\eta. \end{aligned} \tag{4.2.14}$$

In particular, we get that

$$g_k(\alpha) \leq 0, \tag{4.2.15}$$

since $\alpha \in (0,1)$ and

$$2\mathcal{L}_0\alpha^2 + \mathcal{L}\alpha - \mathcal{K} = 0 \tag{4.2.16}$$

by the choice of α. Evidently (4.2.11) holds if

$$f_k(\alpha) \leq 0 \quad \text{for each} \quad k = 1, 2, \ldots. \tag{4.2.17}$$

But in view of (4.2.13), (4.2.14) and (4.2.15) we have that

$$f_k(\alpha) \leq f_{k-1}(\alpha) \leq \cdots \leq f_1(\alpha). \tag{4.2.18}$$

Hence, (4.2.17) holds if

$$f_1(\alpha) \leq 0 \tag{4.2.19}$$

which is true by the choice of η_0. The induction for (4.2.7) is complete. Hence, sequence $\{t_n\}$ is increasing, bounded from above by $t^{\star\star}$ and as such it converges to $t^\star$. Estimates (4.2.8) follows from (4.2.7) and by standard majorization techniques [7, 8, 14, 15, 16, 18]. □

Let us denote by γ_0 and γ_1, respectively, the minimal positive zeros of the following equations with respect to η

$$\left[\frac{\mathcal{K}}{2} + \frac{\mathcal{M}}{6}\alpha(t_2 - t_1)\right](t_2 - t_1) + \mathcal{L}_0(1+\alpha)(t_2 - t_1) + \mathcal{L}_0 t_1 - 1 = 0 \tag{4.2.20}$$

and

$$\left[\frac{\mathcal{K}}{2} + \frac{\mathcal{M}}{6}(t_2 - t_1)\right](t_2 - t_1) + \alpha\mathcal{L}_0 t_2 - \alpha = 0. \tag{4.2.21}$$

Let us set

$$\gamma = \min\{\gamma_0, \gamma_1, 1/\mathcal{L}_0\}. \tag{4.2.22}$$

Then, we can show the following result.

Lemma 4.2.2. *Suppose that*

$$\eta \begin{cases} \leq \gamma & \text{if} \quad \gamma \neq \dfrac{1}{\mathcal{L}_0} \\ < \gamma & \text{if} \quad \gamma = \dfrac{1}{\mathcal{L}_0} \end{cases} \tag{4.2.23}$$

Then, sequence $\{t_n\}$ *generated by* (4.2.5) *is well defined, increasing, bounded from above by*

$$t_1^{\star\star} = t_1 + \frac{t_2 - t_1}{1 - \alpha} \tag{4.2.24}$$

and converges to its unique least upper bound $t_1^\star \in [0, t_1^{\star\star}]$. *Moreover, the following estimates hold for each* $n = 1, 2, \ldots$

$$t_{n+2} - t_{n+1} \leq \alpha^n (t_2 - t_1). \tag{4.2.25}$$

Proof. As in Lemma 4.2.1 we shall prove (4.2.25) using mathematical induction. We have by the choice of γ_1 that

$$\alpha_1 = \frac{\mathcal{K}+\frac{\mathcal{M}}{3}(t_2-t_1)}{2(1-\mathcal{L}_0 t_2)}(t_2-t_1) \le \alpha. \tag{4.2.26}$$

Then, it follows from (4.2.26) and (4.2.20) that

$$\begin{gathered} 0 < t_3 - t_2 \le \alpha(t_2-t_1) \\ t_3 \le t_2 + \alpha(t_2-t_1) \\ t_3 \le t_2 + (1+\alpha)(t_2-t_1) - (t_2-t_1) \\ t_3 \le t_1 + \frac{1-\alpha^2}{1-\alpha}(t_2-t_1) < t^{\star\star}. \end{gathered}$$

Assume that

$$0 < \alpha_k \le \alpha \tag{4.2.27}$$

holds for all $n \le k$. Then, we get by (4.2.5) and (4.2.27) that

$$0 < t_{k+2} - t_{k+1} \le \alpha^k (t_2 - t_1) \tag{4.2.28}$$

and

$$t_{k+2} \le t_1 + \frac{1-\alpha^{k+1}}{1-\alpha}(t_2-t_1) < t_1^{\star\star}. \tag{4.2.29}$$

Estimate (4.2.27) is true, if k is replaced by $k+1$ provided that

$$\left[\frac{\mathcal{K}}{2} + \frac{\mathcal{M}}{6}(t_{k+2}-t_{k+1})\right](t_{k+2}-t_{k+1}) \le \alpha(1-\mathcal{L}_0 t_{k+2})$$

or

$$\left[\frac{\mathcal{K}}{2} + \frac{\mathcal{M}}{6}\alpha^k(t_2-t_1)\right]\alpha^k(t_2-t_1) + \alpha\mathcal{L}_0\left[t_1 + \frac{1-\alpha^{k+1}}{1-\alpha}(t_2-t_1)\right] - \alpha \le 0. \tag{4.2.30}$$

Estimate (4.2.30) motivates us to define recurrent functions f_k on $[0,1)$ by

$$\begin{aligned} f_k(t) = \left[\frac{\mathcal{K}}{2} + \frac{\mathcal{M}}{6}t^k(t_2-t_1)\right]t^k(t_2-t_1) + t\mathcal{L}_0(1+t+\cdots+t^k)(t_2-t_1) \\ - t(1-\mathcal{L}_0 t_1). \end{aligned} \tag{4.2.31}$$

We have that

$$f_{k+1}(t) = f_k(t) + \left[\frac{1}{2}(2\mathcal{L}_0 t^2 + \mathcal{K}t - \mathcal{K}) + \frac{\mathcal{M}}{6}t^k(t^2-1)(t_2-t_1)\right]t^k(t_2-t_1). \tag{4.2.32}$$

In particular, we have the choice of α that

$$f_{k+1}(\alpha) \le f_k(\alpha) \le \cdots \le f_1(\alpha) \le 0. \tag{4.2.33}$$

Evidently, estimate (4.2.30) holds if

$$\begin{aligned} f_k(\alpha) \le 0 \quad \text{or by (4.2.33) if} \\ f_1(\alpha) \le 0 \end{aligned} \tag{4.2.34}$$

which is true by the choice of η_0. The proof of the Lemma is complete. □

Lemma's 4.2.1 and 4.2.2 admit the following useful extensions. The proofs are omitted since they can simply be obtained by replacing $\eta = t_1 - t_0$ with $t_{N+1} - t_N$ where $N = 1, 2, \ldots$ for Lemma 4.2.3 and $N = 2, 3, \ldots$ for Lemma 4.2.4.

Lemma 4.2.3. *Suppose there exists $N = 1, 2, \ldots$ such that*

$$t_0 < t_1 < t_2 < \cdots < t_N < t_{N+1} < \frac{1}{\mathcal{L}_0}$$

and

$$t_{N+1} - t_N \le \begin{cases} \eta_1 & \text{if} \quad \mathcal{L}_0 \eta \le \dfrac{1-\alpha^2}{2+2\alpha-\alpha^2} \\ \eta_0 & \text{if} \quad \dfrac{1-\alpha^2}{2+2\alpha-\alpha^2} \le \mathcal{L}_0 \eta. \end{cases}$$

Then, the conclusions of Lemma 4.2.1 for sequence $\{t_n\}$ hold.

Lemma 4.2.4. *Suppose there exists $N = 2, 3, \ldots$ such that*

$$t_0 < t_1 < t_2 < \cdots < t_N < t_{N+1} < \frac{1}{\mathcal{L}_0}$$

and

$$\eta \begin{cases} \le \gamma & \text{if} \quad \gamma \ne \dfrac{1}{\mathcal{L}_0} \\ < \gamma & \text{if} \quad \gamma = \dfrac{1}{\mathcal{L}_0}, \end{cases}$$

where γ is defined by (4.2.22) *where $t_2 - t_1$, t_1, t_2 are replaced, respectively, by $t_{N+1} - t_N$, t_N, t_{N+1}. Then, the conclusions of Lemma 4.2.1 for sequence $\{t_n\}$ hold.*

Remark 4.2.5. *Another sequence related to Newton's method* (4.1.2) *is given by (see Theorem 4.4.1)*

$$\left.\begin{aligned} s_0 = 0, \quad s_1 = \eta, \quad s_2 = s_1 + \frac{\mathcal{K}_0 + \dfrac{\mathcal{M}_1}{3}(s_1 - s_0)}{2(1 - \mathcal{L}_0 s_1)}(s_1 - s_0)^2 \\ s_{n+2} = s_{n+1} + \frac{\mathcal{K} + \dfrac{\mathcal{M}}{3}(s_{n+1} - s_n)}{2(1 - \mathcal{L}_0 s_{n+1})}(s_{n+1} - s_n)^2 \end{aligned}\right\} \tag{4.2.35}$$

for each $n = 1, 2, \ldots$ *and some* $\mathcal{K}_0 \in (0, \mathcal{K}]$, $\mathcal{M}_1 \in (0, \mathcal{M}]$. *Then, a simple inductive argument shows that*

$$s_n \leq t_n \tag{4.2.36}$$

$$s_{n+1} - s_n \leq t_{n+1} - t_n \tag{4.2.37}$$

and

$$s^\star = \lim_{n\to\infty} s_n \leq t^\star. \tag{4.2.38}$$

Moreover, if $\mathcal{K}_0 < \mathcal{K}$ or $\mathcal{M}_1 < \mathcal{M}$ then (4.2.36) and (4.2.37) hold as strict inequalities. Clearly, sequence $\{s_n\}$ converges under the hypotheses of Lemma 4.2.1 or Lemma 4.2.2. However, $\{s_n\}$ can converge under weaker hypotheses than those of Lemma 4.2.2. Indeed, denote by γ_0^1 and γ_1^1, respectively, the minimal positive zeros of equations

$$\left[\frac{\mathcal{K}}{2} + \frac{\mathcal{M}}{6}\alpha(s_2 - s_1)\right](s_2 - s_1) + \mathcal{L}_0(1+\alpha)(s_2 - s_1) + \mathcal{L}_0 s_1 - 1 = 0 \tag{4.2.39}$$

and

$$\left[\frac{\mathcal{K}_0}{2} + \frac{\mathcal{M}_1}{6}(s_2 - s_1)\right](s_2 - s_1) + \alpha\mathcal{L}_0 s_2 - \alpha = 0. \tag{4.2.40}$$

Set

$$\gamma^1 = \min\{\gamma_0^1, \gamma_1^1, 1/\mathcal{L}_0\}. \tag{4.2.41}$$

Then, we have that

$$\gamma \leq \gamma^1. \tag{4.2.42}$$

Moreover, the conclusions of Lemma 4.2.2 hold for sequence $\{s_n\}$ if (4.2.42) replaces (4.2.23).

Note also that strict inequality can hold in (4.2.42) which implies that the sequence $\{s_n\}$ – which is tighter than $\{t_n\}$ – converges under weaker conditions.

4.3. Majorizing Sequences II

We show convergence of sequences that are majorizing for Newton's method (4.1.2) under the (**H**) conditions.

Lemma 4.3.1. *Let* $\mathcal{K}_0 > 0$, $\mathcal{L}_0 > 0$, $\mathcal{M}_0 > 0$ *and* $\eta > 0$ *with* $\mathcal{K}_0 \leq \mathcal{L}_0$. *Define parameters* a, θ_0 *and* η_1 *by*

$$a = \frac{2\mathcal{K}_0}{\mathcal{K}_0 + \sqrt{\mathcal{K}_0^2 + 8\mathcal{K}_0\mathcal{L}_0}}, \tag{4.3.1}$$

$$\theta_0 = \frac{2}{\frac{\mathcal{K}_0}{2} + (1+a)\mathcal{L}_0 + \sqrt{\left(\frac{\mathcal{K}_0}{2} + (1+a)\mathcal{L}_0\right)^2 + \frac{2\mathcal{M}_0(a+3)}{3}}} \tag{4.3.2}$$

and

$$\theta_1 = \frac{2a}{\frac{\mathcal{K}_0}{2} + a\mathcal{L}_0 + \sqrt{\left(\frac{\mathcal{K}_0}{2} + a\mathcal{L}_0\right)^2 + \frac{2\mathcal{M}_0 a}{3}}}. \tag{4.3.3}$$

Suppose that

$$\eta \leq \begin{cases} \theta_1 & \text{if} \quad \mathcal{L}_0\eta \leq \dfrac{1-a^2}{2+2a-a^2} \\ \theta_0 & \text{if} \quad \dfrac{1-a^2}{2+2a-a^2} \leq \mathcal{L}_0\eta \end{cases} \tag{4.3.4}$$

Then, sequence $\{v_n\}$ generated by

$$v_0 = 0, \quad v_1 = \eta, \quad v_{n+2} = v_{n+1} + \frac{\frac{\mathcal{M}_0}{6}(v_{n+1}-v_n) + \frac{\mathcal{M}_0}{2}v_n + \frac{\mathcal{K}_0}{2}}{1-\mathcal{L}_0 v_{n+1}}(v_{n+1}-v_n) \tag{4.3.5}$$

is well defined, increasing, bounded from above

$$v^{\star\star} = \frac{\eta}{1-a} \tag{4.3.6}$$

and converges to its unique least upper bound $v^\star$ which satisfies $v^\star \in [0, v^{\star\star}]$. Moreover the following estimates hold

$$v_{n+1} - v_n \leq a^n \eta \tag{4.3.7}$$

and

$$v^\star - v_n \leq \frac{a^n\eta}{1-a}. \tag{4.3.8}$$

Proof. As in Lemma 4.2.1 we use mathematical induction to prove that

$$\beta_k = \frac{\frac{\mathcal{K}_0}{2} + \frac{\mathcal{M}_0}{2}v_k + \frac{\mathcal{M}_0}{6}(v_{k+1}-v_k)}{1-\mathcal{L}_0 v_{k+1}}(v_{k+1}-v_k) \leq a. \tag{4.3.9}$$

Estimate (4.3.9) holds for $k=0$ by the choice of θ_1. Let us assume that (4.3.9) holds for all $k \leq n$. Then, we must prove that

$$\left(\frac{\mathcal{K}_0}{2} + \frac{\mathcal{M}_0}{2}\frac{1-a^k}{1-a}\eta + \frac{\mathcal{M}_0}{6}a^k\eta\right)a^k\eta + a\mathcal{L}_0\frac{1-a^{k+1}}{1-a}\eta - a \leq 0. \tag{4.3.10}$$

Define recurrent functions f_k on $[0,1)$ for each $k = 1, 2, \ldots$ by

$$f_k(t) = \left(\frac{\mathcal{K}_0}{2} + \frac{\mathcal{M}_0}{2}(1+t+\cdots++t^{k-1})\eta + \frac{\mathcal{M}_0}{6}a^k\eta\right)t^{k-1}\eta + \mathcal{L}_0(1+t+\cdots+t^k)\eta - 1. \tag{4.3.11}$$

Using (4.3.11), we get that

$$f_{k+1}(a) = f_k(a) + \Big[\frac{1}{2}(2\mathcal{L}_0 a^2 + \mathcal{K}_0 a - \mathcal{K}_0) + \frac{\mathcal{M}_0}{6}(a^{k+2} + 3a^{k+1} + 2a^k - 3)\eta\Big] a^{k-1}\eta \leq f_k(a), \quad (4.3.12)$$

since a given by (4.3.1) solves the equation $2\mathcal{L}_0 a^2 + \mathcal{K}_0 a - \mathcal{K}_0 = 0$ and $a^{k+2} + 3a^{k+1} + 2a^k - 3 \leq 0$ for each $k = 1,2,\ldots$, if $a \in [0,1/2]$. Evidently, it follows from (4.3.12) that (4.3.10) holds which is true by the choice of θ_0. □

Denote by δ_0 and δ_1, respectively, the minimal positive zeros of equations

$$\Big[\frac{\mathcal{K}_0}{2} + \frac{\mathcal{M}_0}{2}\Big(v_2 + \frac{\mathcal{M}_0}{6}a(v_2 - v_1)\Big)\Big](v_2 - v_1) + \mathcal{L}_0(v_1 + (1+a)(v_2 - v_1)) - 1 = 0 \quad (4.3.13)$$

and

$$\Big[\frac{\mathcal{M}_0}{6}(v_2 - v_1) + \frac{\mathcal{M}_0}{2}v_1 + \frac{\mathcal{K}_0}{2}\Big](v_2 - v_1) + a\mathcal{L}_0 v_2 - a = 0. \quad (4.3.14)$$

Set

$$\delta = \min\{\delta_0, \delta_1, 1/\mathcal{L}_0\}. \quad (4.3.15)$$

Then, we can show:

Lemma 4.3.2. *Suppose that*

$$\eta \begin{cases} \leq \delta & \text{if} \quad \delta \neq \dfrac{1}{\mathcal{L}_0} \\ < \delta & \text{if} \quad \delta = \dfrac{1}{\mathcal{L}_0} \end{cases} \quad (4.3.16)$$

Then, sequence $\{v_n\}$ *generated by equation* (4.3.5) *is well defined, increasing, bounded from above by*

$$v_1^{\star\star} = v_1 + \frac{v_2 - v_1}{1 - a} \quad (4.3.17)$$

and converges to its unique least-upper bound $v_1^\star$ *which satisfies* $v_1^\star \in [0, v_1^{\star\star}]$. *Moreover, the following estimates hold for each* $n = 1,2,3,\ldots$

$$v_{n+2} - v_{n+1} \leq a^n(v_2 - v_1). \quad (4.3.18)$$

Proof. We have that $\beta_1 \leq a$ by the choice of δ_1. This time we must have

$$\Big[\frac{\mathcal{K}_0}{2} + \frac{\mathcal{M}_0}{2}\Big(v_1 + \frac{1-a^k}{1-a}(v_2 - v_1)\Big) + \frac{\mathcal{M}_0}{6}a^k(v_2 - v_1)\Big]a^k(v_2 - v_1) + a\mathcal{L}_0\Big[v_1 + \frac{1-a^{k+1}}{1-a}(v_2 - v_1)\Big] - a \leq 0. \quad (4.3.19)$$

Define functions f_k on $[0,1)$ by

$$f_k(t) = \Big[\frac{\mathcal{K}_0}{2} + \frac{\mathcal{M}_0}{2}\Big(v_1 + \frac{1-t^k}{1-t}(v_2-v_1)\Big) + \frac{\mathcal{M}_0}{6}a^k(t_2-t_1)\Big]t^k(v_2-v_1) + t\mathcal{L}_0\Big[v_1 + \frac{1-t^{k+1}}{1-t}(v_2-v_1)\Big] - t. \quad (4.3.20)$$

We have that

$$f_{k+1}(a) = f_k(a) + \Big[\frac{1}{2}(2\mathcal{L}_0a^2 + \mathcal{K}_0a - \mathcal{K}_0) + \frac{\mathcal{M}_0}{6}(v_2-v_1)\Big(3(a-1)a + (a^{k+2}+3a^{k+1}+2a^k-3)\Big)\Big]a^k(v_2-v_1).$$

Thus

$$f_{k+1}(a) \le f_k(a) \le \cdots \le f_1(a). \quad (4.3.21)$$

But by the choice of η_0 we have that $f_1(a) \le 0$. □

Remark 4.3.3. *A sequence related to Newton's method* (4.1.2) *under the **(H)** conditions is defined by*

$$\left.\begin{aligned} u_0 = 0, \quad u_1 = \eta, \quad u_2 &= u_1 + \frac{\mathcal{K}_0 + \frac{\mathcal{M}_1}{3}(u_1-u_0)}{2(1-\mathcal{L}_0u_1)}(u_1-u_0)^2 \\ u_{n+2} &= u_{n+1} + \frac{\mathcal{K}_0 + \frac{\mathcal{M}_0}{3}(u_{n+1}-u_n)}{2(1-\mathcal{L}_0u_{n+1})}(u_{n+1}-u_n)^2 \end{aligned}\right\} \quad (4.3.22)$$

for each $n = 1,2,\ldots$ and $\mathcal{M}_1 \in (0,\mathcal{M}]$. Then, a simple inductive argument shows that for each $n = 2,3,\ldots$

$$u_n \le v_n \quad (4.3.23)$$

$$u_{n+1} - u_n \le v_{n+1} - v_n \quad (4.3.24)$$

and

$$u^\star = \lim_{n\to\infty} \le v^\star. \quad (4.3.25)$$

Moreover, if $\mathcal{K}_0 < \mathcal{K}$ or $\mathcal{M}_1 < \mathcal{M}_0$ then (4.3.23) *and* (4.3.24) *hold as strict inequalities. Sequence $\{u_n\}$ converges under the hypotheses of Lemma 4.3.1 or 4.3.2. However, $\{u_n\}$ can converge under weaker hypotheses than those of Lemma 4.3.2. Indeed, denote by δ_0^1 and δ_1^1, respectively, the minimal positive zeros of equations*

$$\Big[\frac{\mathcal{K}_0}{2} + \frac{\mathcal{M}_0}{2}\Big(u_2 + \frac{\mathcal{M}_0}{6}(u_2-u_1)\Big)\Big](u_2-u_1) + \mathcal{L}_0(u_1 + (1+a)(u_2-u_1)) - 1 = 0 \quad (4.3.26)$$

and

$$\Big[\frac{\mathcal{M}_0}{6}(u_2-u_1) + \frac{\mathcal{M}_0}{2}u_1 + \frac{\mathcal{K}_0}{2}\Big](u_2-u_1) + a\mathcal{L}_0u_2 - a = 0. \quad (4.3.27)$$

Set

$$\delta^1 = \min\{\delta_0^1, \delta_1^1, 1/\mathcal{L}_0\}. \tag{4.3.28}$$

Then, we have that

$$\delta \leq \delta^1.$$

Moreover, the conclusions of Lemma 4.3.2 hold for sequence $\{u_n\}$ *if* (4.3.28) *replaces* (4.3.16). *Note also that strict inequality may hold in* (4.3.28) *which implies that, tighter than* $\{v_n\}$*, sequence* $\{u_n\}$ *converges under weaker conditions. Finally note that sequence* $\{t_n\}$ *is tighter than* $\{v_n\}$ *although the sufficient convergence conditions for* $\{v_n\}$ *are weaker than those of* $\{t_n\}$.

Lemmas similar to Lemma 2.3 and Lemma 2.4 for sequence $\{v_n\}$ can follow in an analogous way.

4.4. Semilocal Convergence

We present the semilocal convergence of Newton's method (4.1.2) first under the (**C**) and then under the (**H**) conditions. Let $u(x,R)$ and $\overline{U}(x,R)$ stand, respectively, for the open and closed balls in $\mathcal{X}$ centered at $x \in \mathcal{X}$ and of radius $R > 0$.

Theorem 4.4.1. *Let* $\mathcal{F} : \mathcal{D} \subseteq \mathcal{X} \longrightarrow \mathcal{Y}$ *be twice Fréchet differentiable. Suppose that the (**C**) conditions, hypotheses of Lemma 4.2.1 hold and*

$$\overline{U}(x_0, t^\star) \subseteq \mathcal{D}. \tag{4.4.1}$$

Then, the sequence $\{x_n\}$ *defined by Newton's method* (4.1.2) *is well defined, remains in* $\overline{U}(x_0, t^\star)$ *for all* $n \geq 0$ *and converges to a unique solution* $x^\star \in \overline{U}(x_0, t^\star)$ *of equation* $\mathcal{F}(x) = 0$. *Moreover, the following estimates hold for all* $n \geq 0$

$$\|x_{n+2} - x_{n+1}\| \leq t_{n+2} - t_{n+1} \tag{4.4.2}$$

and

$$\|x_n - x^\star\| \leq t^\star - t_n, \tag{4.4.3}$$

where, sequence $\{t_n\}$ *(*$n \geq 0$*) is given by* (4.2.5). *Furthermore, if there exists* $R \geq t^\star$*, such that*

$$U(x_0, R) \subseteq \mathcal{D} \tag{4.4.4}$$

and

$$\mathcal{L}_0(t^\star + R) \leq 2. \tag{4.4.5}$$

The solution $x^\star$ *is unique in* $U(x_0, R)$.

Proof. Let us prove that

$$\|x_{k+1}-x_k\| \le t_{k+1}-t_k \tag{4.4.6}$$

and

$$\overline{U}(x_{k+1},t^\star-t_{k+1}) \subseteq \overline{U}(x_k,t^\star-t_k) \tag{4.4.7}$$

hold for all $k \ge 0$. For every $z \in \overline{U}(x_1,t^\star-t_1)$

$$\begin{aligned}\|z-x_0\| &\le \|z-x_1\|+\|x_1-x_0\|\\ &\le (t^\star-t_1)+(t_1-t_0)=t^\star-t_0,\end{aligned}$$

implies that $z \in \overline{U}(x_0,t^\star-t_0)$. Since, also

$$\|x_1-x_0\| = \left\|\mathcal{F}'(x_0)^{-1}\mathcal{F}(x_0)\right\| \le \eta = t_1-t_0.$$

Thus estimate (4.4.6) and (4.4.7) hold for $k=0$. Given they hold for $n=0,1,2,\ldots,k$, then we have

$$\|x_{k+1}-x_0\| \le \sum_{i=1}^{k+1}\|x_i-x_{i-1}\| \le \sum_{i=1}^{k+1}(t_i-t_{i-1}) = t_{k+1}-t_0 = t_{k+1} \tag{4.4.8}$$

and

$$\|x_k+\theta(x_{k+1}-x_k)-x_0\| \le t_k+\theta(t_{k+1}-t_k) \le t^\star, \tag{4.4.9}$$

for all $\theta \in [0,1]$. Using (4.1.2), we obtain the approximation

$$\begin{aligned}\mathcal{F}(x_{k+1}) &= \mathcal{F}(x_{k+1})-\mathcal{F}(x_k)-\mathcal{F}'(x_k)(x_{k+1}-x_k)\\ &= \int_0^1 [\mathcal{F}'(x_k+\theta(x_{k+1}-x_k))-\mathcal{F}'(x_k)](x_{k+1}-x_k)\mathrm{d}\theta \\ &= \int_0^1 \mathcal{F}''(x_k+\theta(x_{k+1}-x_k))(1-\theta)(x_{k+1}-x_k)^2\mathrm{d}\theta.\end{aligned} \tag{4.4.10}$$

Then, we get by $(\mathbf{C}_3)$, $(\mathbf{C}_4)$ and (4.4.1)

$$\begin{aligned}\left\|\mathcal{F}'(x_0)^{-1}\mathcal{F}(x_{k+1})\right\| &\le \int_0^1 \Big(\left\|\mathcal{F}'(x_0)^{-1}[\mathcal{F}''(x_k+\theta(x_{k+1}-x_k))-\mathcal{F}''(x^\star)]\right\| \\ &\qquad + \left\|\mathcal{F}'(x_0)^{-1}\mathcal{F}''(x^\star)\right\|\Big)\|x_{k+1}-x_k\|^2(1-\theta)\mathrm{d}\theta \\ &\le \Big[\mathcal{M}\Big(\int_0^1\|x_{k+1}-x_k\|(1-\theta)\mathrm{d}\theta\Big)+\frac{\mathcal{K}}{2}\Big]\|x_{k+1}-x_k\|^2 \\ &\le \frac{\mathcal{M}}{6}\|x_{k+1}-x_k\|^3+\frac{\mathcal{K}}{2}\|x_{k+1}-x_k\|^2 \\ &\le \Big[\overline{\mathcal{M}}\Big(\frac{1}{6}(t_{k+1}-t_k)\Big)+\frac{\overline{\mathcal{K}}}{2}\Big](t_{k+1}-t_k)^2\end{aligned} \tag{4.4.11}$$

where

$$\overline{\mathcal{K}} = \begin{cases} \mathcal{K}_0, & \mathcal{K} = 0, \\ \mathcal{K}, & \mathcal{K} > 0, \end{cases} \quad \text{and} \quad \overline{M} = \begin{cases} \mathcal{M}_0, & \mathcal{K} = 0, \\ \mathcal{M}, & \mathcal{K} > 0. \end{cases}$$

Using ($\mathbf{C}_5$), we obtain that

$$\begin{aligned} \left\| \mathcal{F}'(x_0)^{-1}(\mathcal{F}'(x_{k+1}) - \mathcal{F}'(x_0)) \right\| &\leq \mathcal{L}_0 \left\| x_{k+1} - x_0 \right\| \\ &\leq \mathcal{L}_0 t_{k+1} \leq \mathcal{L}_0 t^\star < 1. \end{aligned} \tag{4.4.12}$$

It follows from the Banach lemma on invertible operators [7, 8, 14, 15, 16] and (4.4.12) that $\mathcal{F}'(x_{k+1})^{-1}$ exists and

$$\begin{aligned} \left\| \mathcal{F}'(x_{k+1})^{-1} \mathcal{F}'(x_0) \right\| &\leq (1 - \mathcal{L}_0 \left\| x_{k+1} - x_0 \right\|)^{-1} \\ &\leq (1 - \mathcal{L}_0 t_{k+1})^{-1}. \end{aligned} \tag{4.4.13}$$

Therefore by (4.1.2), (4.4.11) and (4.4.13), we obtain in turn

$$\begin{aligned} \left\| x_{k+2} - x_{k+1} \right\| &\leq \left\| \mathcal{F}'(x_{k+1})^{-1} \mathcal{F}'(x_{k+1}) \right\| \\ &\leq \left\| \mathcal{F}'(x_{k+1})^{-1} \mathcal{F}'(x_0) \right\| \left\| \mathcal{F}'(x_0)^{-1} \mathcal{F}(x_{k+1}) \right\| \\ &\leq t_{k+2} - t_{k+1}. \end{aligned} \tag{4.4.14}$$

Thus for every $z \in \overline{U}(x_{k+2}, t^\star - t_{k+2})$, we have

$$\begin{aligned} \left\| z - x_{k+1} \right\| &\leq \left\| z - x_{k+2} \right\| + \left\| x_{k+2} - x_{k+1} \right\| \\ &\leq t^\star - t_{k+2} + t_{k+2} - t_{k+1} = t^\star - t_{k+1}. \end{aligned} \tag{4.4.15}$$

That is,

$$z \in \overline{U}(x_{k+1}, t^\star - t_{k+1}). \tag{4.4.16}$$

Estimates (4.4.13) and (4.4.16) imply that (4.4.6) and (4.4.7) hold for $n = k+1$. The proof of (4.4.6) and (4.4.7) is now complete by induction.

Lemma 4.2.1 implies that sequence $\{t_n\}$ is a Cauchy sequence. From (4.4.6) and (4.4.7), $\{x_n\}$ $(n \geq 0)$ becomes a Cauchy sequence too and as such it converges to some $x^\star \in \overline{U}(x_0, t^\star)$ (since $\overline{U}(x_0, t^\star)$ is a closed set). Estimate (4.4.3) follows from (4.4.2) by using standard majorization techniques [7, 8, 14, 15, 16, 18]. Moreover, by letting $k \to \infty$ in (4.4.11), we obtain $\mathcal{F}(x^\star) = 0$. Finally, to show uniqueness let $y^\star$ be a solution of equation $\mathcal{F}(x) = 0$ in $U(x_0, R)$. It follows from ($\mathbf{C}_5$) for $x = y^\star + \theta(x^\star - y^\star)$, $\theta \in [0,1]$, the estimate

$$\begin{aligned} &\left\| \mathcal{F}'(x_0)^{-1} \int_0^1 (\mathcal{F}'(y^\star + \theta(x^\star - y^\star)) - \mathcal{F}'(x_0)) \right\| \mathrm{d}\theta \\ &\leq \mathcal{L}_0 \int_0^1 \left\| y^\star + \theta(x^\star - y^\star) - x_0 \right\| \mathrm{d}\theta \\ &\leq \mathcal{L}_0 \int_0^1 (\theta \left\| x^\star - x_0 \right\| + (1-\theta) \left\| y^\star - x_0 \right\|) \mathrm{d}\theta \\ &\leq \frac{\mathcal{L}_0}{2}(t^\star + R) \leq 1, \quad \text{(by (4.6.2))} \end{aligned}$$

and the Banach lemma on invertible operators implies that the linear operator $T^{\star\star} = \int_0^1 \mathcal{F}'(y^\star + \theta(x^\star - y^\star))\mathrm{d}\theta$ is invertible. Using the identity $0 = \mathcal{F}(x^\star) - \mathcal{F}'(y^\star) = T^{\star\star}(x^\star - y^\star)$, we deduce that $x^\star = y^\star$.

Similarly, we show the uniqueness in $\overline{U}(x_0, t^\star)$ by setting $t^\star = R$. That completes the proof of Theorem 4.4.1. □

Remark 4.4.2. *The conclusions of Theorem 4.4.1 hold if $\{t_n\}$, $t^\star$ are replaced by $\{r_n\}$, $r^\star$, respectively.*

Using the approximation

$$\mathcal{F}'(x_0)^{-1}\mathcal{F}(x_{k+1}) = \int_0^1 \mathcal{F}'(x_0)^{-1}[\mathcal{F}''(x_k + \theta(x_{k+1} - x_k)) - \mathcal{F}''(x_0)](x_{k+1} - x_k)^2(1-\theta)\mathrm{d}\theta$$
$$+ \int_0^1 \mathcal{F}'(x_0)^{-1}\mathcal{F}''(x_0)(1-\theta)\mathrm{d}\theta\, \|x_{k+1} - x_k\|^2 \quad (4.4.17)$$

instead of (4.4.11) and ($\mathbf{C}_6$), ($\mathbf{C}_7$) instead of, respectively, ($\mathbf{C}_3$), ($\mathbf{C}_4$), we arrive at the following semilocal convergence result under the (**H**) conditions [8, Theorem 6.3.7 p. 210 for proof].

Theorem 4.4.3. *Let $\mathcal{F} : \mathcal{D} \subseteq \mathcal{X} \longrightarrow \mathcal{Y}$ be twice Fréchet differentiable. Furthermore suppose that the (**H**) conditions,*

$$\overline{U}(x_0, v^\star) \subseteq \mathcal{D}, \quad (4.4.18)$$

and hypotheses of Lemma 4.3.1 hold. Then, the sequence $\{x_n\}$ generated by Newton's method (4.1.2) *is well defined, remains in $\overline{U}(x_0, t^\star)$ for all $n \geq 0$ and converges to a unique solution $x^\star \in \overline{U}(x_0, t^\star)$ of equation $\mathcal{F}(x) = 0$. Moreover, the following estimates hold for all $n \geq 0$:*

$$\|x_{n+2} - x_{n+1}\| \leq v_{n+2} - v_{n+1} \quad (4.4.19)$$

and

$$\|x_n - x^\star\| \leq v^\star - v_n \quad (4.4.20)$$

where, sequence $\{v_n\}$ ($n \geq 0$) is given by (4.3.5). *Furthermore, if there exists $R \geq t^\star$, such that*

$$U(x_0, R) \subseteq \mathcal{D} \quad (4.4.21)$$

and

$$\mathcal{L}_0(t^\star + R) \leq 2. \quad (4.4.22)$$

The solution $x^\star$ is unique in $U(x_0, R)$.

4.5. Local Convergence

We study the local convergence of Newton's method under the (**A**) conditions

A$_1$. there exists $x^\star \in \mathcal{D}$ such that $\mathcal{F}(x^\star) = 0$ and $\mathcal{F}'(x^\star)^{-1} \in \mathbf{L}(\mathcal{Y}, \mathcal{X})$

A$_2$. $\left\|\mathcal{F}'(x^\star)^{-1}\mathcal{F}''(x^\star)\right\| \le b$

A$_3$. $\left\|\mathcal{F}'(x^\star)^{-1}[\mathcal{F}''(x) - \mathcal{F}''(x^\star)]\right\| \le c\,\|x - x^\star\|$ for each $x \in \mathcal{D}$

A$_4$. $\left\|\mathcal{F}'(x^\star)^{-1}[\mathcal{F}'(x) - \mathcal{F}'(x^\star)]\right\| \le d\,\|x - x^\star\|$ for each $x \in \mathcal{D}$.

Note also that in view of (**A**$_3$) and (**A**$_4$), respectively, there exist $c_0 \in (0, c]$ and $d_0 \in (0, d]$ such that for each $\theta \in [0, 1]$

A$'_3$. $\left\|\mathcal{F}'(x^\star)^{-1}\Big(\mathcal{F}''(x_0 + \theta(x^\star - x_0)) - \mathcal{F}''(x^\star)\Big)\right\| \le c_0(1-\theta)\,\|x_0 - x^\star\|$

A$'_4$. $\left\|\mathcal{F}'(x^\star)^{-1}\Big(\mathcal{F}'(x_0) - \mathcal{F}'(x^\star)\Big)\right\| \le d_0(1-\theta)\,\|x_0 - x^\star\|$.

Then, we can show:

Theorem 4.5.1. *Suppose that (**A**) conditions hold and*

$$\overline{U}(x^\star, r) \subseteq \mathcal{D}, \tag{4.5.1}$$

where

$$r = \frac{2}{\dfrac{b}{2} + d + \sqrt{\left(\dfrac{b}{2} + d\right)^2 + \dfrac{4c}{3}}}. \tag{4.5.2}$$

Then, sequence $\{x_n\}$ (starting from $x_0 \in U(x^\star, r)$) generated by Newton's method (4.1.2) *is well defined, remains in $U(x^\star, r)$ for all $n \ge 0$ and converges to $x^\star$. Moreover the following estimates hold*

$$\|x_{n+1} - x^\star\| \le e_n\,\|x_n - x^\star\|^2, \tag{4.5.3}$$

$$e_n = \frac{\dfrac{\overline{c}}{3}\,\|x_n - x^\star\| + \dfrac{b}{2}}{1 - \overline{d}\,\|x_n - x^\star\|} \quad \text{and} \quad q(t) = \frac{\dfrac{ct}{3} + \dfrac{b}{2}}{1 - dt}t \tag{4.5.4}$$

where

$$\overline{c} = \begin{cases} c_0 & \text{if} \quad n = 0, \\ c & \text{if} \quad n > 0, \end{cases} \qquad \overline{d} = \begin{cases} d_0 & \text{if} \quad n = 0, \\ d & \text{if} \quad n > 0. \end{cases}$$

Proof. The starting point $x_0 \in U(x^\star, r)$. Then, suppose that $x_k \in U(x^\star, r)$ for all $k \le n$. Using (**A**$_4$) and the definition of r we get that

$$\left\|\mathcal{F}'(x^\star)^{-1}(\mathcal{F}'(x_k) - \mathcal{F}'(x^\star))\right\| \le d\,\|x_k - x^\star\| < dr < 1. \tag{4.5.5}$$

It follows from (4.5.5) and the Banach lemma on invertible operators that $\mathcal{F}'(x_k)^{-1}$ exists and

$$\left\|\mathcal{F}'(x_k)^{-1}\mathcal{F}'(x^\star)\right\| \leq \frac{1}{1-d\,\|x_k-x^\star\|}. \tag{4.5.6}$$

Hence, x_{k+1} exists. Using (4.1.2), we obtain the approximation

$$x^\star - x_{k+1} = -\mathcal{F}'(x_k)^{-1}\mathcal{F}'(x^\star)\Big[\int_0^1 \mathcal{F}'(x^\star)^{-1}\Big(\mathcal{F}''(x_k+\theta(x^\star-x_k)) - \mathcal{F}''(x^\star)\Big) + \mathcal{F}''(x^\star)\Big] (x^\star - x_k)^2(1-\theta)\mathrm{d}\theta \tag{4.5.7}$$

In view of $(\mathbf{A}_2)$, $(\mathbf{A}_3)$, $(\mathbf{A}_4)$, (4.5.6), (4.5.7) and the choice of r we have in turn that

$$\begin{aligned}\|x_{k+1}-x^\star\| &\leq \frac{c\int_0^1(1-\theta)^2\,\|x_k-x^\star\|^3\,\mathrm{d}\theta + b\int_0^1(1-\theta)\mathrm{d}\theta\,\|x_k-x^\star\|^2}{1-d\,\|x_k-x^\star\|}\\ &\leq e_k\,\|x_k-x^\star\|^2 < q(r)\,\|x_k-x^\star\| = \|x_k-x^\star\|\end{aligned} \tag{4.5.8}$$

which implies that $x_{k+1}\in U(x^\star,r)$ and $\lim_{k\to\infty}x_k = x^\star$. □

Remark 4.5.2. *The local results can be used or projection methods such as Arnolds, the generalized minimum residual method (GMRES), the generalized conjugate method (GCR) for combined Newton/finite projection methods and in connection with the mesh independence principle to develop the cheapest and most efficient mesh refinement strategies [7, 8, 4, 15, 16]. These results can also be used to solve equations of the form* (4.1.1), *where $\mathcal{F}'$, $\mathcal{F}''$ satisfy differential equations of the form*

$$\mathcal{F}'(x) = \mathcal{P}(\mathcal{F}(x)) \quad \textit{and} \quad \mathcal{F}''(x) = \mathcal{Q}(\mathcal{F}(x)). \tag{4.5.9}$$

where, $\mathcal{P}$ and $\mathcal{Q}$ are known operators. Since, $\mathcal{F}'(x^\star) = \mathcal{P}(\mathcal{F}(x^\star)) = \mathcal{P}(0)$ and $\mathcal{F}''(x^\star) = \mathcal{Q}(\mathcal{F}(x^\star)) = \mathcal{Q}(0)$ we can apply our results without actually knowing the solution $x^\star$ of equation (4.1.1).

4.6. Numerical Examples

Example 1. *Let $\mathcal{X} = \mathcal{Y} = \mathbb{R}$ be equipped with the max-norm, $x_0 = \omega$, $\mathcal{D} = [-\exp(1), \exp(1)]$. Let us define $\mathcal{F}$ on $\mathcal{D}$ by*

$$\mathcal{F}(x) = x^3 - \exp(1). \tag{4.6.1}$$

Here, $\omega \in \mathcal{D}$. Through some algebraic manipulations, we obtain

$$\begin{cases} \eta = \dfrac{|\omega^3-\exp(1)|}{3\omega^2}, \quad \mathcal{K} = \dfrac{4\exp(1)}{\omega^2}, \quad \mathcal{L}_0 = \dfrac{2\exp(1)+\omega}{\omega^2}, \quad \mathcal{K}_0 = \dfrac{2}{\omega} \\ \mathcal{M} = \dfrac{2}{\omega^2}, \quad \mathcal{M}_0 = \dfrac{2}{\omega^2}. \end{cases}$$

For $\omega = 0.48\exp(1)$, the criteria (4.1.3) *and* (4.1.6) *yield*

$$0.09730789545 \leq 0.07755074734 \quad \textit{and} \quad 0.09730789545 \leq 0.2856823952$$

Table 4.6.1. Newton's method applied to (4.4.11)

n	x_n	$\|x_{n+2}-x_{n+1}\|$	$\|x_n - x^\star\|$
0	$1.30e+00$	$6.44e-03$	$9.08e-02$
1	$1.40e+00$	$2.98e-05$	$6.47e-03$
2	$1.40e+00$	$6.37e-10$	$2.98e-05$
3	$1.40e+00$	$2.91e-19$	$6.37e-10$
4	$1.40e+00$	$6.06e-38$	$2.91e-19$
5	$1.40e+00$	$2.63e-75$	$6.06e-38$
6	$1.40e+00$	$4.95e-150$	$2.63e-75$
7	$1.40e+00$	$1.76e-299$	$4.95e-150$
8	$1.40e+00$	$2.22e-598$	$1.76e-299$
9	$1.40e+00$	$3.52e-1196$	$2.22e-598$

respectively. Thus we observe that the criterion (4.1.3) *fails while the criterion* (4.1.6) *holds. From the hypothesis of Lemma 4.2.1, we get*

$$0.09730789545 \le \begin{cases} 0.2017739733 & \text{if} \quad 0.08268226632 \le 0.2499999999 \\ 0.2036729480 & \text{if} \quad 0.2499999999 \le 0.08268226632. \end{cases}$$

Thus the hypothesis of Lemma 4.2.1 hold. As a consequence, we can apply the Theorem 4.4.1. The table 4.6.1 reports convergence behavior of Newton's method (4.1.2) *applied to* (4.4.11) *with* $x_0 = 1$ *and* $\psi = 0.55$. *Numerical computations are performed to the decimal point accuracy of* 2005 *by employing the high-precision library* ARPREC. *The Table 4.6.2 reports behavior of series* $\{t_n\}$ (4.2.5). *Comparing Tables 4.6.1 and 4.6.2, we observe that*

Table 4.6.2. Sequences $\{t_n\}$ (4.2.5)

n	t_n	$t_{n+2}-t_{n+1}$	$t^\star - t_n$
0	$0.00e+00$	$4.95e-02$	$1.69e-01$
1	$9.73e-02$	$1.87e-02$	$7.16e-02$
2	$1.47e-01$	$3.26e-03$	$2.21e-02$
3	$1.66e-01$	$1.02e-04$	$3.36e-03$
4	$1.69e-01$	$1.01e-07$	$1.02e-04$
5	$1.69e-01$	$9.75e-14$	$1.01e-07$
6	$1.69e-01$	$9.16e-26$	$9.75e-14$
7	$1.69e-01$	$8.08e-50$	$9.16e-26$
8	$1.69e-01$	$6.30e-98$	$8.08e-50$
9	$1.69e-01$	$3.82e-194$	$6.30e-98$

the estimates of Theorem 4.4.1 hold.

Example 2. *In this example, we provide an application of our results to a special nonlinear*

Hammerstein integral equation of the second kind. Consider the integral equation

$$x(s) = 1 + \frac{4}{5}\int_0^1 G(s,t)x(t)^3\,dt, \quad s \in [0,1], \tag{4.6.2}$$

where, G is the Green kernel on $[0,1]\times[0,1]$ *defined by*

$$G(s,t) = \begin{cases} t(1-s), & t \le s; \\ s(1-t), & s \le t. \end{cases} \tag{4.6.3}$$

Let $\mathcal{X} = \mathcal{Y} = \mathcal{C}[0,1]$ *and* $\mathcal{D}$ *be a suitable open convex subset of* $\mathcal{X}_1 := \{x \in \mathcal{X} : x(s) > 0, s \in [0,1]\}$*, which will be given below. Define* $\mathcal{F} : \mathcal{D} \to \mathcal{Y}$ *by*

$$[\mathcal{F}(x)](s) = x(s) - 1 - \frac{4}{5}\int_0^1 G(s,t)x(t)^3\,dt, \quad s \in [0,1]. \tag{4.6.4}$$

The first and second derivatives of $\mathcal{F}$ *are given by*

$$[\mathcal{F}(x)'y](s) = y(s) - \frac{12}{5}\int_0^1 G(s,t)x(t)^2 y(t)\,dt, \quad s \in [0,1], \tag{4.6.5}$$

and

$$[\mathcal{F}(x)''yz](s) = \frac{24}{5}\int_0^1 G(s,t)x(t)y(t)z(t)\,dt, \quad s \in [0,1], \tag{4.6.6}$$

respectively. We use the max-norm. Let $x_0(s) = 1$ *for all* $s \in [0,1]$*. Then, for any* $y \in \mathcal{D}$*, we have*

$$[(I - \mathcal{F}'(x_0))(y)](s) = \frac{12}{5}\int_0^1 G(s,t)y(t)\,dt, \quad s \in [0,1], \tag{4.6.7}$$

which means

$$\|I - \mathcal{F}'(x_0)\| \le \frac{12}{5}\max_{s\in[0,1]}\int_0^1 G(s,t)\,dt = \frac{12}{5\times 8} = \frac{3}{10} < 1. \tag{4.6.8}$$

It follows from the Banach theorem that $\mathcal{F}'(x_0)^{-1}$ *exists and*

$$\|\mathcal{F}'(x_0)^{-1}\| \le \frac{1}{1-\dfrac{3}{10}} = \frac{10}{7}. \tag{4.6.9}$$

On the other hand, we have from (4.4.7) *that*

$$\|\mathcal{F}(x_0)\| = \frac{4}{5}\max_{s\in[0,1]}\int_0^1 G(s,t)\,dt = \frac{1}{10}.$$

Then, we get $\eta = 1/7$*. Note that* $\mathcal{F}''(x)$ *is not bounded in* $\mathcal{X}$ *or its subset* $\mathcal{X}_1$*. Take into account that a solution* $x^\star$ *of equation* (4.1.1) *with* $\mathcal{F}$ *given by* (4.4.6) *must satisfy*

$$\|x^\star\| - 1 - \frac{1}{10}\|x^\star\|^3 \le 0, \tag{4.6.10}$$

i.e., $\|x^\star\| \le \rho_1 = 1.153467305$ *and* $\|x^\star\| \ge \rho_2 = 2.423622140$, *where* ρ_1 *and* ρ_2 *are the positive roots of the real equation* $z - 1 - z^3/10 = 0$. *Consequently, if we look for a solution such that* $x^\star < \rho_1 \in \mathcal{X}_1$, *we can consider* $\mathcal{D} := \{x : x \in \mathcal{X}_1 \quad and \quad \|x\| < r\}$, *with* $r \in (\rho_1, \rho_2)$, *as a nonempty open convex subset of* $\mathcal{X}$. *For example, choose* $r = 1.7$. *Using* (4.3.7) *and* (4.3.8), *we have that for any* $x, y, z \in \mathcal{D}$

$$\begin{aligned}\left\|\left[(\mathcal{F}'(x) - \mathcal{F}'(x_0))y\right](s)\right\| &= \frac{12}{5}\left\|\int_0^1 G(s,t)(x(t)^2 - x_0(t)^2)y(t)\,dt\right\| \\ &\le \frac{12}{5}\int_0^1 G(s,t)\|x(t) - x_0(t)\|\,\|x(t) + x_0(t)\|y(t)\,dt \\ &\le \frac{12}{5}\int_0^1 G(s,t)\,(r+1)\|x(t) - x_0(t)\|y(t)\,dt, \quad s \in [0,1]\end{aligned} \tag{4.6.11}$$

and

$$\|(F''(x)yz)(s)\| = \frac{24}{5}\int_0^1 G(s,t)x(t)y(t)z(t)\,dt, \quad s \in [0,1]. \tag{4.6.12}$$

Then, we get

$$\|\mathcal{F}'(x) - \mathcal{F}'(x_0)\| \le \frac{12}{5}\frac{1}{8}(r+1)\|x - x_0\| = \frac{81}{100}\|x - x_0\|, \tag{4.6.13}$$

$$\|F''(x)\| \le \frac{24}{5} \times \frac{r}{8} = \frac{51}{50} \tag{4.6.14}$$

and

$$\left\|\left[\left[F''(x) - \mathcal{F}''(\overline{x})\right]yz\right](s)\right\| = \frac{24}{5}\left\|\int_0^1 G(s,t)\,(x(t) - \overline{x}(t)))y(t)z(t)\right\|dt \tag{4.6.15}$$

$$\le \frac{24}{5}\frac{1}{8}\|x - \overline{x}\| = \frac{3}{5}\|x - \overline{x}\|. \tag{4.6.16}$$

Now we can choose constants as follows:

$$\eta = \frac{1}{7}, \quad \mathcal{M} = \frac{6}{7}, \quad \mathcal{M}_0 = \frac{6}{7}, \quad \mathcal{K} = \frac{51}{35}, \quad \mathcal{L}_0 = \frac{49}{70} \quad and \quad \mathcal{K}_0 = \frac{11}{15}.$$

From (4.1.3) *and* (4.1.5), *we obtain*

$$0.1428571429 < 0.3070646192 \quad and \quad R_1 = 0.1627780248.$$

From (4.1.6) *and* (4.1.8), *we get*

$$0.1428571429 < 0.4988741112 \quad and \quad R_2 = 0.1518068730.$$

From the hypotheses (4.2.4) *and* (4.3.4) *we get*

$$\frac{1}{7} \le \begin{cases} 0.5047037049 & if \quad 0.1000000000 \le 0.2131833880 \\ 0.5228360736 & if \quad 0.2131833880 \le 0.1000000000 \end{cases}$$

Table 4.6.3. Comparison among the sequences (4.2.5)**,** (4.2.35)**,** (4.3.5) **and** (4.3.22)

n	t_n	s_n	v_n	u_n
0	$0.000000e+00$	$0.000000e+00$	$0.000000e+00$	$0.000000e+00$
1	$1.428571e-01$	$1.428571e-01$	$1.428571e-01$	$1.428571e-01$
2	$1.598408e-01$	$1.514801e-01$	$2.042976e-01$	$1.516343e-01$
3	$1.600782e-01$	$1.515408e-01$	$2.356037e-01$	$1.516661e-01$
4	$1.600783e-01$	$1.515408e-01$	$2.527997e-01$	$1.516661e-01$
5	$1.600783e-01$	$1.515408e-01$	$2.626215e-01$	$1.516661e-01$
6	$1.600783e-01$	$1.515408e-01$	$2.683548e-01$	$1.516661e-01$
7	$1.600783e-01$	$1.515408e-01$	$2.717435e-01$	$1.516661e-01$
8	$1.600783e-01$	$1.515408e-01$	$2.737612e-01$	$1.516661e-01$
9	$1.600783e-01$	$1.515408e-01$	$2.749678e-01$	$1.516661e-01$

and

$$\frac{1}{7} \leq \begin{cases} 0.6257238049 & \text{if} \quad 0.1000000000 \leq 0.2691240473 \\ 0.5832936968 & \text{if} \quad 0.2691240473 \leq 0.1000000000 \end{cases}$$

respectively. Thus hypotheses (4.2.4) *and* (4.3.4) *hold. Comparison – among sequences* (4.2.5)*,* (4.2.35)*,* (4.3.5) *and* (4.3.22) *is reported in Table 4.6.3. In the Table 4.6.3, we observe that the estimates* (4.2.36) *and* (4.3.23) *hold.*

Concerning the uniqueness balls. From equation (4.1.5)*, we get* $R_1 = 0.1627780248$ *and from equation* (4.1.8)*, we get* $R_2 = 0.1518068730$*. Whereas from Theorem 4.4.1, we get* $R \leq 1.257142857$*. Therefore, the new approach provides the largest uniqueness ball.*

Example 3. *Let us consider the case when* $\mathcal{X} = \mathcal{Y} = \mathbb{R}$*,* $\mathcal{D} = U(0,1)$ *and define* $\mathcal{F}$ *on* $\mathcal{D}$ *by*

$$\mathcal{F}(x) = e^x - 1. \tag{4.6.17}$$

Then, we can define $\mathcal{P}(x) = x+1$ *and* $\mathcal{Q}(x) = x+1$*. In order for us to compare our radius of convergence with earlier ones, let us introduce the Lipschitz condition*

$$\left\| \mathcal{F}'(x^\star)^{-1}(\mathcal{F}'(x) - \mathcal{F}'(y)) \right\| \leq \mathcal{L} \|x - y\| \quad \text{for each} \quad x, y \in \mathcal{D}. \tag{4.6.18}$$

The radius of convergence given by Traub-Wozniakowski [7, 8, 16] is

$$r_0 = \frac{2}{3\mathcal{L}} \tag{4.6.19}$$

The radius of convergence given by us in [5, 6, 7, 8]

$$r_1 = \frac{2}{2d + \mathcal{L}} \tag{4.6.20}$$

Using ($\boldsymbol{A}_2$)*,* ($\boldsymbol{A}_3$)*,* ($\boldsymbol{A}_4$) *and* (4.6.18)*, we get that* $b = 1$*,* $c = d = e - 1$ *and* $\mathcal{L} = e$*. Then, using* (4.5.2)*,* (4.6.19) *and* (4.6.20)*, we obtain*

$$r = 0.4078499356 > r_1 = 0.324947231 > r_0 = 0.245252961.$$

Example 4. *Let $X = Y = \mathbb{R}^3$, $D = U(0,1)$, $x^* = (0,0,0)$ and define function $\mathcal{F}$ on $\mathcal{D}$ by*

$$\mathcal{F}(x,y,z) = \left(e^x - 1, \frac{e-1}{2}y^2 + y, z\right)^T. \tag{4.6.21}$$

We have that for $u = (x,y,z)$

$$\mathcal{F}'(u) = \begin{pmatrix} e^x & 0 & 0 \\ 0 & (e-1)y+1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

$$\mathcal{F}''(u) = \left(\begin{array}{ccc|ccc|ccc} e^x & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & e-1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

and

$$\mathcal{F}'''(u) = \left(\begin{array}{c|c|c|c|c|c|c|c|ccc} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \begin{array}{ccc} e^x & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array} \end{array}\right).$$

Using the ($\mathbf{A}$) and ($\mathbf{A}'$) conditions – and $\mathcal{F}'(x^*) = diag\{1,1,1\}$ – we set

$$b = 1.0, \quad \overline{c} = c_0 = c = \overline{d} = d_0 = d = e-1, \quad \mathcal{L} = e, \quad \text{and} \quad \mathcal{L}_0 = e-1.$$

We obtain

$$r = 0.4078499356.$$

Thus, $r_0 < r$.

Table 4.6.4. Comparison among various iterative procedures

n	$\|x_{n+1} - x^\star\|$	$e_n \|x_n - x^\star\|^2$	$\lambda_n \|x_n - x^\star\|^2$	$\mu_n \|x_n - x^\star\|^2$
1	0.034624745433299	0.292667362771974	0.479494429606589	15.944478671072201
2	0.000669491177317	0.000677347930013	0.001732513344520	0.001798733838791
3	0.000000347374133	0.000000224639537	0.000000609893622	0.000000610302684
4	0.000000000000103	0.000000000000060	0.000000000000164	0.000000000000164
5	0.000000000000000	0.000000000000000	0.000000000000000	0.000000000000000

The following iterations have been used before

$$\|x_{n+1} - x^\star\| \leq p_n \|x_n - x^\star\|^2 \quad \text{[6, 7, 8, 12]},$$
$$\|x_{n+1} - x^\star\| \leq \lambda_n \|x_n - x^\star\|^2 \quad \text{[6, 7, 8]}$$
$$\|x_{n+1} - x^\star\| \leq \mu_n \|x_n - x^\star\|^2 \quad \text{[16]}$$

and

$$\|x_{n+1} - x^\star\| \leq \xi_n \|x_n - x^\star\|^2 \quad \text{[6, 7, 8, 16]}$$

Table 4.6.5. Comparison among various iterative procedures

n	$\xi_n \left\|x_n - x^\star\right\|^2$
1	0.240445748047369
2	0.000661013573819
3	0.000000224531576
4	0.000000000000060
5	0.000000000000000

where

$$p_n = \frac{\mathcal{L}/3\,\|x_n - x^\star\| + b/2}{1 - d\,\|x_n - x^\star\|}, \quad \lambda_n = \frac{\mathcal{L}/2}{1 - \mathcal{L}_0\,\|x_n - x^\star\|},$$

$$\mu_n = \frac{\mathcal{L}/2}{1 - \mathcal{L}\,\|x_n - x^\star\|} \quad \text{and} \quad \xi_n = \frac{\mathcal{L}/3\,\|x_n - x^\star\| + b/2}{1 - \Big(\mathcal{L}/2\,\|x_n - x^\star\| + b\Big)\,\|x_n - x^\star\|}.$$

To compare the above iterations with the iteration (4.5.3)*, we produce the comparison table 4.6.4 and 4.6.5, we apply Newton's method* (4.1.2) *to the equation* (4.6.21) *with* $x_0 = (0.21, 0.21, 0.21)^T$. *In the Table 4.6.4, we note that the estimate* (4.5.3) *– of Theorem 4.5.1 – hold.*

References

[1] Amat, S., C. Bermúdez, Busquier, S., Legaz, M. J., Plaza, S., On a Family of High-Order Iterative Methods under Kantorovich Conditions and Some Applications, *Abst. Appl. Anal.*, 2012, (2012).

[2] Amat, S., Bermúdez, C., Busquier, S., Plaza, S., On a third-order Newton-type method free of bilinear operators, *Numer. Linear Alg. with Appl.*, 17(4) (2010), 639-653.

[3] Amat, S., Busquier, S., Third-order iterative methods under Kantorovich conditions, *J. Math. Anal. App.*, 336(1) (2007), 243-261.

[4] Amat, S., Busquier, S., Gutiérrez, J. M., Third-order iterative methods with applications to Hammerstein equations: A unified approach, *J. Comput. App. Math.*, 235(9) (2011), 2936-2943.

[5] Argyros, I.K., A Newton-Kantorovich theorem for equations involving m-Fréchet differentiable operators and applications in radiative transfer, *J. Comp. App. Math.*, 131 (2001), 149-159.

[6] Argyros, I.K., On the Newton-Kantorovich hypothesis for solving equations, *Journal of Computation and Applied Mathematics*, 169(2) (2004), 315-332.

[7] Argyros, I.K., Cho, Y.J., Hilout, S., Numerical methods for equations and its applications, *CRC Press/Taylor and Francis Publications*, New York, 2012.

[8] Argyros, I.K., *Computational theory of iterative methods*, Elsevier, 2007.

[9] Ezquerro, J.A., González, D., Hernández, M. A., Majorizing sequences for Newton's method from initial value problems, *J. Comp. Appl. Math.*, 236 (2012), 2246-2258.

[10] Ezquerro, J.A., Hernández, M.A., Generalized differentiability conditions for Newton's method, *IMA J. Numer. Anal.*, 22(2) (2002), 187-205.

[11] Ezquerro, J.A., González, D., Hernández, M.A., On the local convergence of Newton's method under generalized conditions of Kantorovich, *App. Math. Let.*, 26 (2013), 566-570.

[12] Gutiérrez, J.M., A new semilocal convergence theorem for Newton's method, *J. Comp. App. Math.*, 79 (1997), 131 - 145.

[13] Gutiérrez, J.M., Magreñán, Á.A., Romero, N., On the semilocal convergence of Newton-Kantorovich method under center-Lipschitz conditions, *App. Math. Comp.*, 221 (2013), 79-88.

[14] Kantorovich,L.V., The majorant principle and Newton's method, *Doklady Akademii Nauk SSSR*, 76, 17-20, 1951. (In Russian).

[15] Ostrowski, A.M., *Solution of equations in Euclidean and Banach spaces*, Academic Press, New York, 3rd Ed. 1973.

[16] Traub, J. F., Iterative Methods for the Solution of Equations, *American Mathematical Soc.*, 1982.

[17] Werner, W., Some improvements of classical iterative methods for the solution of nonlinear equations, *Numerical Solution of Nonlinear Equations Lecture Notes in Mathematics*, 878/1981 (1981), 426-440.

[18] Yamamoto, T., On the method of tangent hyperbolas in Banach spaces, *Journal of Computational and Applied Mathematics*, 21(1), (1988), 75-86.

Chapter 5

Expanding the Applicability of Newton's Method Using Smale's α-Theory

5.1. Introduction

Let $\mathcal{X}$, $\mathcal{Y}$ be Banach spaces. Let $U(x,r)$ and $\overline{U}(x,r)$ stand, respectively, for the open and closed ball in $\mathcal{X}$ with center x and radius $r > 0$. Denote by $\mathcal{L}(\mathcal{X},\mathcal{Y})$ the space of bounded linear operators from $\mathcal{X}$ into $\mathcal{Y}$. In the present chapter we are concerned with the problem of approximating a locally unique solution $x^\star$ of equation

$$F(x) = 0, \tag{5.1.1}$$

where F is a Fréchet continuously differentiable operator defined on $\overline{U}(x_0,R)$ for some $R > 0$ with values in $\mathcal{Y}$.

A lot of problems from computational sciences and other disciplines can be brought in the form of equation (5.1.1) using Mathematical Modelling [5, 13]. The solution of these equations can rarely be found in closed form. That is why the solution methods for these equations are iterative. In particular, the practice of numerical analysis for finding such solutions is essentially connected to variants of Newton's method [5, 13, 21, 22]. The study about convergence matter of Newton methods is usually centered on two types: semilocal and local convergence analysis. The semilocal convergence matter is, based on the information around an initial point, to give criteria ensuring the convergence of Newton methods; while the local one is, based on the information around a solution, to find estimates of the radii of convergence balls. We find in the literature several studies on the weakness and/or extension of the hypothesis made on the underlying operators. There is a plethora on local as well as semil-local convergence results, we refer the reader to [1]–[34]. The most famous among the semilocal convergence of iterative methods is the celebrated Kantorovich theorem for solving nonlinear equations. This theorem provides a simple and transparent convergence criterion for operators with bounded second derivatives F'' or the Lipschitz continuous first derivatives [5, 13, 21, 22]. Another important theorem inaugurated by Smale at the International Conference of Mathematics (cf. [28]), where the concept of an approximate zero was proposed and the convergence criteria were provided to determine an approximate zero for analytic function, depending on the information at the initial point. Wang and Han [32, 31] generalized Smale's result by introducing the γ-condition (see (5.1.3)). For more details on Smale's theory, the reader can refer to the excellent Dedieu's book [15, Chapter 3.3].

Newton's method defined by

$$\begin{aligned} &x_0 \text{ is an initial point} \\ &x_{n+1} = x_n - F'(x_n)^{-1} F(x_n) \qquad \text{for each} \quad n = 0, 1, 2, \cdots \end{aligned} \tag{5.1.2}$$

is undoubtedly the most popular iterative process for generating a sequence $\{x_n\}$ approximating $x^\star$. Here, $F'(x)$ denotes the Fréchet-derivative of F at $x \in \overline{U}(x_0, R)$.

In the present chapter we expand the applicability of Newton's method under the γ-condition by introducing the notion of the center γ_0-condition (to be precised in Definition 5.3.1) for some $\gamma_0 \leq \gamma$. This way we obtain tighter upper bounds on the norms of $\| F'(x)^{-1} F'(x_0) \|$ for each $x \in \overline{U}(x_0, R)$ (see (5.2.4), (5.2.2) and (5.2.3)) leading to weaker sufficient convergence conditions and a tighter convergence analysis than in earlier studies such as [14, 19, 27, 28, 31, 32]. The approach of introducing center-Lipschitz condition has already been fruitful for expanding the applicability of Newton's method under the Kantorovich-type theory [3, 9, 11, 13].

Wang in his work [31] on approximate zeros of Smale (cf. [28, 29]) used the γ-Lipschitz condition at x_0

$$\begin{aligned} &\| F'(x_0)^{-1} F''(x) \| \leq \frac{2\gamma}{(1 - \gamma \| x - x_0 \|)^3} \\ &\text{for each} \quad x \in U(x_0, r),\ 0 < r \leq R, \end{aligned} \tag{5.1.3}$$

where $\gamma > 0$ and x_0 are such that $\gamma \| x - x_0 \| < 1$ and $F'(x_0)^{-1} \in \mathcal{L}(\mathcal{Y}, \mathcal{X})$ to show the following semilocal convergence result for Newton's method.

Theorem 5.1.1. *Let $F : \overline{U}(x_0, R) \subseteq \mathcal{X} \longrightarrow \mathcal{Y}$ be twice-Fréchet differentiable. Suppose there exists $x_0 \in U(x_0, R)$ such that $F'(x_0)^{-1} \in \mathcal{L}(\mathcal{Y}, \mathcal{X})$ and*

$$\| F'(x_0)^{-1} F(x_0) \| \leq \eta; \tag{5.1.4}$$

condition (5.1.3) holds and for $\alpha = \gamma\eta$

$$\alpha \leq 3 - 2\sqrt{2}; \tag{5.1.5}$$

$$t^\star \leq R, \tag{5.1.6}$$

where

$$t^\star = \frac{1 + \alpha - \sqrt{(1+\alpha)^2 - 8\alpha}}{4\gamma} \leq \left(1 - \frac{1}{\sqrt{2}}\right) \frac{1}{\gamma}. \tag{5.1.7}$$

Then, sequence $\{x_n\}$ generated by Newton's method is well defined, remains in $\overline{U}(x_0, t^\star)$ for each $n = 0, 1, \cdots$ and converges to a unique solution $x^\star \in \overline{U}(x_0, t^\star)$ of equation $F(x) = 0$. Moreover, the following error estimates hold

$$\| x_{n+1} - x_n \| \leq t_{n+1} - t_n \tag{5.1.8}$$

and

$$\| x_{n+1} - x^\star \| \leq t^\star - t_n, \tag{5.1.9}$$

where scalar sequence $\{t_n\}$ *is defined by*

$$\begin{aligned} & t_0 = 0, \quad t_1 = \eta, \\ & t_{n+1} = t_n + \frac{\gamma(t_n - t_{n-1})^2}{\left(2 - \frac{1}{(1-\gamma t_n)^2}\right)(1-\gamma t_n)(1-\gamma t_{n-1})^2} = t_n - \frac{\varphi(t_n)}{\varphi'(t_n)} \\ & for\ each\ n = 1, 2, \cdots, \end{aligned} \tag{5.1.10}$$

where

$$\varphi(t) = \eta - t + \frac{\gamma t^2}{1 - \gamma t}. \tag{5.1.11}$$

Notice that $t^\star$ *is the small zero of equation* $\varphi(t) = 0$*, which exists under the hypothesis (5.1.5).*

The chapter is organized as follows: sections 5.2. and 5.3. contain the semilocal and local convergence analysis of Newton's method. Applications and numerical examples are given in the concluding section 5.4.

5.2. Semilocal Convergence of Newton's Method

We need some auxiliary results. We shall use the Banach lemma on invertible operators [5, 13, 21, 22]

Lemma 5.2.1. *Let* A, B *be bounded linear operators, where* A *is invertible. Moreover,* $\| A^{-1} \| \| B \| < 1$*. Then,* $A + B$ *is invertible and*

$$\| (A+B)^{-1} \| \leq \frac{\| A^{-1} \|}{1 - \| A^{-1} \| \| B \|}. \tag{5.2.1}$$

We shall also use the following definition of Lipschitz and local Lipschitz conditions.

Definition 5.2.2. *(see [14, p. 634], [34, p. 673]) Let* $F : \overline{U}(x_0, R) \longrightarrow \mathcal{Y}$ *be Fréchet-differentiable on* $U(x_0, R)$*. We say that* F' *satisfies the Lipschitz condition at* x_0 *if there exists an increasing function* $\ell : [0, R] \longrightarrow [0, +\infty)$ *such that*

$$\begin{aligned} & \| F'(x_0)^{-1} (F'(x) - F'(y)) \| \leq \ell(r) \| x - y \| \\ & for\ each\ x, y \in \overline{U}(x_0, r),\ 0 < r \leq R. \end{aligned} \tag{5.2.2}$$

In view of (5.2.2), there exists an increasing function $\ell_0 : [0, R] \longrightarrow [0, +\infty)$ such that the center-Lipschitz condition

$$\begin{aligned} & \| F'(x_0)^{-1} (F'(x) - F'(x_0)) \| \leq \ell_0(r) \| x - x_0 \| \\ & for\ each\ x \in \overline{U}(x_0, r),\ 0 < r \leq R \end{aligned} \tag{5.2.3}$$

holds. Clearly,

$$\ell_0(r) \leq \ell(r) \text{ for each } r \in (0, R] \tag{5.2.4}$$

holds in general and $\ell(r)/\ell_0(r)$ can be arbitrarily large [3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13].

Lemma 5.2.3. *(see [14, p. 638]) Let* $F : \overline{U}(x_0,R) \longrightarrow \mathcal{Y}$ *be Fréchet-differentiable on* $U(x_0,R)$. *Suppose* $F'(x_0)^{-1} \in \mathcal{L}(\mathcal{Y},\mathcal{X})$ *and there exist* $\gamma_0 \geq 0$, $\gamma \geq 0$ *such that* $\gamma_0 R < 1$, $\gamma R < 1$. *Then,* F' *satisfies conditions (5.2.2) and (5.2.3), respectively, with*

$$\ell(r) := \frac{2\gamma}{(1-\gamma r)^3} \tag{5.2.5}$$

and

$$\ell_0(r) := \frac{\gamma_0(2-\gamma_0 r)}{(1-\gamma_0 r)^2}. \tag{5.2.6}$$

Notice that with preceding choices of functions ℓ and ℓ_0 and since condition (5.2.4) is satisfied, we can always choose γ_0, γ such that

$$\gamma_0 \leq \gamma. \tag{5.2.7}$$

From now on we assume that condition (5.2.7) is satisfied. We also need a result by Zabrejko and Nguen.

Lemma 5.2.4. *(see [34, p. 673]) Let* $F : \overline{U}(x_0,R) \longrightarrow \mathcal{Y}$ *be Fréchet-differentiable on* $U(x_0,R)$. *Suppose* $F'(x_0)^{-1} \in \mathcal{L}(\mathcal{Y},\mathcal{X})$ *and*

$$\begin{array}{l} \| F'(x_0)^{-1}(F'(x)-F'(y)) \| \leq \lambda(r) \, \| x-y \| \\ for\ each\ x,y \in \overline{U}(x_0,r),\ 0 < r \leq R \end{array}$$

for some increasing function $\lambda : [0,R] \longrightarrow [0,+\infty)$. *Then, the following assertion holds*

$$\begin{array}{l} \| F'(x_0)^{-1}(F'(x+p)-F'(x)) \| \leq \Lambda(r+\| p \|)-\Lambda(r) \\ for\ each\ x \in \overline{U}(x_0,r),\ 0 < r \leq R\ and\ \| p \| \leq R-r, \end{array}$$

where

$$\Lambda(r) = \int_0^r \lambda(t)\,dt.$$

In particular, if

$$\begin{array}{l} \| F'(x_0)^{-1}(F'(x)-F'(x_0)) \| \leq \lambda_0(r) \, \| x-x_0 \| \\ for\ each\ x \in \overline{U}(x_0,r),\ 0 < r \leq R \end{array}$$

for some increasing function $\lambda_0 : [0,R] \longrightarrow [0,+\infty)$. *Then, the following assertion holds*

$$\begin{array}{l} \| F'(x_0)^{-1}(F'(x_0+p)-F'(x_0)) \| \leq \Lambda_0(\| p \|) \\ for\ each\ 0 < r \leq R\ and\ \| p \| \leq R-r, \end{array}$$

where

$$\Lambda_0(r) = \int_0^r \lambda_0(t)\,dt.$$

Using the center-Lipschitz condition and Lemma 5.2.3, we can show the following result on invertible operators.

Lemma 5.2.5. *Let $F : \overline{U}(x_0,R) \longrightarrow \mathcal{Y}$ be Fréchet-differentiable on $U(x_0,R)$. Suppose $F'(x_0)^{-1} \in \mathcal{L}(\mathcal{Y},\mathcal{X})$ and $\gamma_0 R < 1$ for some $\gamma_0 > 0$ and $x_0 \in \mathcal{X}$; center-Lipschitz (5.2.3) holds on $U_0 = U(x_0,r_0)$, where $\ell_0(r)$ is given by (5.2.6) and $r_0 = (1-\frac{1}{\sqrt{2}})\frac{1}{\gamma_0}$. Then $F'(x)^{-1} \in \mathcal{L}(\mathcal{Y},\mathcal{X})$ on U_0 and satisfies*

$$\| F'(x)^{-1}F'(x_0) \| \leq \left(2-\frac{1}{(1-\gamma_0 r)^2}\right)^{-1}. \tag{5.2.8}$$

Proof. We have by (5.2.3), (5.2.6) and $x \in U_0$ that

$$\| F'(x_0)^{-1}(F'(x)-F'(x_0)) \| \leq \ell_0(r)\, r = \frac{1}{(1-\gamma_0 r)^2} - 1 < 1.$$

The result now follows from Lemma 5.2.1. The proof of Lemma 5.2.5 is complete. □

Using (5.1.3) a similar to Lemma 5.2.1, Banach lemma was given in [31] (see also [27, 28, 29]).

Lemma 5.2.6. *Let $F : \overline{U}(x_0,R) \longrightarrow \mathcal{Y}$ be twice Fréchet-differentiable on $U(x_0,R)$. Suppose $F'(x_0)^{-1} \in \mathcal{L}(\mathcal{Y},\mathcal{X})$ and $\gamma R < 1$ for some $\gamma > 0$ and $x_0 \in \mathcal{X}$; condition (5.1.3) holds on $V_0 = U(x_0,\overline{r_0})$, where $\overline{r_0} = (1-\frac{1}{\sqrt{2}})\frac{1}{\gamma}$. Then $F'(x)^{-1} \in \mathcal{L}(\mathcal{Y},\mathcal{X})$ on $\overline{V_0}$ and satisfies*

$$\| F'(x)^{-1}F'(x_0) \| \leq \left(2-\frac{1}{(1-\gamma r)^2}\right)^{-1}. \tag{5.2.9}$$

Remark 5.2.7. *It follows from (5.2.7)–(5.2.9) that (5.2.8) is more precise upper bound on the norm of $F'(x)^{-1}F'(x_0)$. This observation leads to a tighter majorizing sequence for $\{x_n\}$ (see Proposition 5.2.10).*

We can show the main following semilocal convergence theorem for Newton's method.

Theorem 5.2.8. *Suppose that*

(a) There exist $x_0 \in \mathcal{X}$ and $\eta > 0$ such that

$$F'(x_0)^{-1} \in \mathcal{L}(\mathcal{Y},\mathcal{X}) \quad \text{and} \quad \| F'(x_0)^{-1}F(x_0) \| \leq \eta;$$

(b) Operator F' satisfies Lipschitz and center-Lipschitz conditions (5.2.2) and (5.2.3) on $U(x_0,r_0)$ with $\ell(r)$ and $\ell(r)$ are given by (5.2.5) and (5.2.6), respectively;

(c) $U_0 \subseteq \overline{U}(x_0,R)$;

(d) Scalar sequence $\{s_n\}$ defined by

$$\begin{aligned} &s_0 = 0,\ s_1 = \eta, \\ &s_2 = s_1 + \frac{\gamma_0 (s_1-s_0)^2}{\left(2-\frac{1}{(1-\gamma_0 s_1)^2}\right)(1-\gamma s_1)} \\ &s_{n+2} = s_{n+1} + \frac{\gamma(s_{n+1}-s_n)^2}{\left(2-\frac{1}{(1-\gamma_0 s_{n+1})^2}\right)(1-\gamma s_{n+1})(1-\gamma s_n)^2} \\ &\mathit{for\ each} \quad n = 1,2,\cdots \end{aligned} \tag{5.2.10}$$

satisfies for each $n = 1, 2, \cdots$

$$s_n < b = \begin{cases} \dfrac{1}{\gamma} & if \quad \dfrac{\gamma_0}{\gamma} \leq 1 - \dfrac{1}{\sqrt{2}} \\ (1 - \dfrac{1}{\sqrt{2}})\dfrac{1}{\gamma_0} & if \quad \dfrac{\gamma_0}{\gamma} \geq 1 - \dfrac{1}{\sqrt{2}}. \end{cases} \tag{5.2.11}$$

Then, the following assertions hold

(i) Sequence $\{s_n\}$ *is increasingly convergent to its unique least upper bound* $s^\star$ *which satisfies* $s^\star \in [s_2, b]$*, where b is given in (5.2.11).*

(ii) Sequence $\{x_n\}$ *generated by Newton's method is well defined, remains in* $\overline{U}(x_0, s^\star)$ *for each* $n = 0, 1, \cdots$ *and converges to a unique solution* $x^\star \in \overline{U}(x_0, s^\star)$ *of equation* $F(x) = 0$*. Moreover, the following estimates hold*

$$\| x_{n+1} - x_n \| \leq s_{n+1} - s_n \tag{5.2.12}$$

and

$$\| x_n - x^\star \| \leq s^\star - s_n \quad \text{for each} \quad n = 0, 1, 2, \cdots. \tag{5.2.13}$$

Proof. (i) It follows from (5.2.8) and (5.2.9) that sequence $\{s_n\}$ is increasing and bounded above by $1/\gamma$. Hence, it converges to $s^\star \in [s_2, b]$.

(ii) We use Mathematical Induction to prove that

$$\| x_{k+1} - x_k \| \leq s_{k+1} - s_k \tag{5.2.14}$$

and

$$\overline{U}(x_{k+1}, s^\star - s_{k+1}) \subseteq \overline{U}(x_k, s^\star - s_k) \quad \text{for each} \quad k = 1, 2, \cdots. \tag{5.2.15}$$

Let $z \in \overline{U}(x_1, s^\star - s_1)$. Then, we obtain that

$$\| z - x_0 \| \leq \| z - x_1 \| + \| x_1 - x_0 \| \leq s^\star - s_1 + s_1 - s_0 = s^\star - s_0,$$

which implies $z \in \overline{U}(x_0, s^\star - s_0)$. Note also that

$$\| x_1 - x_0 \| = \| F'(x_0)^{-1} F(x_0) \| \leq \eta = s_1 - s_0.$$

Hence, estimates (5.2.14) and (5.2.15) hold for $k = 0$. Suppose these estimates hold for natural integers $n \leq k$. Then, we have that

$$\| x_{k+1} - x_0 \| \leq \sum_{i=1}^{k+1} \| x_i - x_{i-1} \| \leq \sum_{i=1}^{k+1} (s_i - s_{i-1}) = s_{k+1} - s_0 = s_{k+1}$$

and

$$\| x_k + \theta (x_{k+1} - x_k) - x_0 \| \leq s_k + \theta (s_{k+1} - s_k) \leq s^\star \quad \text{for all} \quad \theta \in (0, 1).$$

Using (5.2.2), Lemma 5.2.1 for $x = x_{k+1}$ and the induction hypotheses we get that

$$\begin{aligned} \| F'(x_0)^{-1}(F'(x_{k+1}) - F'(x_0)) \| &\leq \frac{1}{(1-\gamma_0 \| x_{k+1} - x_0 \|)^2} - 1 \\ &\leq \frac{1}{(1-\gamma_0 s_{k+1})^2} - 1 < 1. \end{aligned} \tag{5.2.16}$$

It follows from (5.2.16) and the Banach lemma 5.2.1 on invertible operators that $F'(x_{k+1})^{-1}$ exists and

$$\| F'(x_{k+1})^{-1} F'(x_0) \| \leq \left(2 - \frac{1}{(1-\gamma_0 s_{k+1})^2} \right)^{-1}. \tag{5.2.17}$$

Using (5.1.2), we obtain the approximation

$$\begin{aligned} F(x_{k+1}) &= F(x_{k+1}) - F(x_k) - F'(x_k)(x_{k+1} - x_k) \\ &= \int_0^1 (F'(x_k^\tau) - F'(x_k))\, d\tau\, (x_{k+1} - x_k), \end{aligned} \tag{5.2.18}$$

where $x_k^\tau = x_k + \tau(x_{k+1} - x_k)$ and $x_k^{\tau s} = x_k + \tau s (x_{k+1} - x_k)$ for each $0 \leq \tau, s \leq 1$. Then by (5.2.18) for $k = 0$, (5.2.3) and (5.2.6), we get that

$$\begin{aligned} &\| F'(x_0)^{-1} F(x_1) \| \\ &\leq \int_0^1 \| F'(x_0)^{-1}(F'(x_0 + \tau(x_1 - x_0)) - F'(x_0)) \| \, d\tau \, \| x_1 - x_0 \| \\ &\leq \int_0^1 \left(\frac{1}{(1-\gamma_0 \tau (\| x_1 - x_0 \|)^2} - 1 \right) d\tau \, \| x_1 - x_0 \| \\ &= \frac{\gamma_0 \| x_1 - x_0 \|^2}{1 - \gamma_0 \| x_1 - x_0 \|} \leq \frac{\gamma_0 (s_1 - s_0)^2}{1 - \gamma_0 s_1}. \end{aligned}$$

Moreover, it follows from Lemma 5.2.4, (5.2.2) and (5.2.18) in turn for $k = 1, 2, \cdots$ that

$$\begin{aligned} &\| F'(x_0)^{-1} F(x_{k+1}) \| \\ &\leq \int_0^1 \| F'(x_0)^{-1}(F'(x_k^\tau) - F'(x_k)) \| \, d\tau \, \| x_{k+1} - x_k \| \\ &\leq \int_0^1 \int_0^1 \frac{2\gamma\tau\, ds\, d\tau \, \| x_{k+1} - x_k \|^2}{(1 - \gamma \| x_k^{\tau s} - x_0 \|)^3} \\ &\leq \int_0^1 \int_0^1 \frac{2\gamma\tau\, ds\, d\tau \, \| x_{k+1} - x_k \|^2}{(1 - \gamma \| x_k - x_0 \| - \gamma\tau s \| x_{k+1} - x_k \|)^3} \\ &= \frac{\gamma \| x_{k+1} - x_k \|^2}{(1 - \gamma \| x_k - x_0 \| - \gamma \| x_{k+1} - x_k \|)(1 - \gamma \| x_k - x_0 \|)^2} \\ &\leq \frac{\gamma (s_{k+1} - s_k)^2}{(1 - \gamma s_{k+1})(1 - \gamma s_k)^2} \left(\frac{\| x_{k+1} - x_k \|}{s_{k+1} - s_k} \right)^2 \\ &\leq \frac{\gamma (s_{k+1} - s_k)^2}{(1 - \gamma s_{k+1})(1 - \gamma s_k)^2}. \end{aligned} \tag{5.2.19}$$

(see also [27, p. 33, estimate (3.19)]) Then, in view of (5.1.2), (5.2.10), (5.2.17) and the preceding two estimates we obtain that

$$\begin{aligned}\| x_2 - x_1 \| &\leq \| F'(x_1)^{-1} F(x_0) \| \| F'(x_0)^{-1} F(x_1) \| \\ &\leq \frac{1}{2 - \dfrac{1}{(1-\gamma_0 s_1)^2}} \frac{\gamma_0 (s_1 - s_0)^2}{1 - \gamma_0 s_1} = s_2 - s_1\end{aligned}$$

and for $k = 1, 2, \cdots$

$$\begin{aligned}&\| x_{k+2} - x_{k+1} \| = \| (F'(x_{k+1})^{-1} F'(x_0)) (F'(x_0)^{-1} F(x_{k+1})) \| \\ &\leq \| F'(x_{k+1})^{-1} F'(x_0) \| \| F'(x_0)^{-1} F(x_{k+1}) \| \\ &\leq \frac{1}{2 - \dfrac{1}{(1-\gamma_0 s_{k+1})^2}} \frac{\gamma (s_{k+1} - s_k)^2}{(1-\gamma s_{k+1})(1-\gamma s_k)^2} = s_{k+2} - s_{k+1}.\end{aligned} \tag{5.2.20}$$

Hence, we showed (5.2.14) holds for all $k \geq 0$. Furthermore, let $w \in \overline{U}(x_{k+2}, s^\star - s_{k+2})$. Then, we have that

$$\begin{aligned}\| w - x_{k+1} \| &\leq \| w - x_{k+2} \| + \| x_{k+2} - x_{k+1} \| \\ &\leq s^\star - s_{k+2} + s_{k+2} - s_{k+1} = s^\star - s_{k+1}.\end{aligned} \tag{5.2.21}$$

That is $w \in \overline{U}(x_{k+1}, s^\star - s_{k+1})$. The induction for (5.2.14) and (5.2.15) is now completed. Lemma 5.2.5 implies that $\{s_n\}$ is a complete sequence. It follows from (5.2.14) and (5.2.15) that $\{x_n\}$ is also a complete sequence in a Banach space $\mathcal{X}$ and as such it converges to some $x^\star \in \overline{U}(x_0, s^\star)$ (since $\overline{U}(x_0, s^\star)$ is a closed set). By letting $k \longrightarrow \infty$ in (5.2.19) we get $F(x^\star) = 0$. Estimate (5.2.13) is obtained from (5.2.12) by using standard majorization techniques (cf. [5, 13, 21, 28, 29]). Finally, to show the uniqueness part, let $y^\star \in U(x_0, s^\star)$ be a solution of equation (5.1.1). Using (5.2.3) for x replaced by $z^\star = x^\star + \tau (y^\star - x^\star)$ and $\mathcal{G} = \int_0^1 F'(z^\star)\, d\tau$ we get as in (5.2.9) that $\| F'(x_0)^{-1} (\mathcal{G} - F'(x_0)) \| < 1$. That is $\mathcal{G}^{-1} \in \mathcal{L}(\mathcal{Y}, \mathcal{X})$. Using the identity $0 = F(x^\star) - F(y^\star) = \mathcal{G}(x^\star - y^\star)$, we deduce $x^\star = y^\star$.

□

Remark 5.2.9. *(a) The convergence criteria in Theorem 5.2.8 are weaker than in Theorem 5.1.1. In particular, Theorem 5.1.1 requires that operator F is twice Fréchet-differentiable but our Theorem 5.2.8 requires only that F is Fréchet-differentiable. Notice also that if F is twice Fréchet-differentiable, then (5.2.2) implies (5.1.3). Moreover, in view of (5.1.7) and (5.2.9), we have that* (5.1.5) $\Longrightarrow$ (5.2.11) *but not necessarily vice versa. Therefore, Theorem 5.2.8 can apply in cases when Theorem 5.1.1 cannot.*

(b) Estimate (5.2.11) can be checked, since scalar sequence is based on the initial data γ_0, γ *and* η, *especially in the case when* $s_i = s_{i+n}$ *for some finite i. At this point, we would like to know if it is possible to find convergence criteria stronger than (5.2.11) but weaker than (5.1.5). To this extend we first compare our majorizing sequence* $\{s_n\}$ *to the old majorizing sequence* $\{t_n\}$.

Proposition 5.2.10. *Let $F : \overline{U}(x_0,R) \longrightarrow \mathcal{Y}$ be twice Fréchet-differentiable on $U(x_0,R)$. Suppose that hypotheses of Theorem 5.1.1 and the center-Lipschitz condition hold on $\overline{U}(x_0,r_0)$. Then, the following assertions hold*

(a) *Scalar sequences $\{t_n\}$ and $\{s_n\}$ are increasingly convergent to $t^\star$, $s^\star$, respectively.*

(b) *Sequence $\{x_n\}$ generated by Newton's method is well defined, remains in $\overline{U}(x_0,r_0)$ for each $n = 0,1,\cdots$ and converges to a unique solution $x^\star \in \overline{U}(x_0,r_0)$ of equation $F(x) = 0$. Moreover, the following estimates hold for each $n = 0,1,\cdots$*

$$s_n \le t_n, \tag{5.2.22}$$

$$s_{n+1} - s_n \le t_{n+1} - t_n, \tag{5.2.23}$$

$$s^\star \le t^\star, \tag{5.2.24}$$

$$\| x_{n+1} - x_n \| \le s_{n+1} - s_n$$

and

$$\| x_n - x^\star \| \le s^\star - s_n.$$

Proof. According to Theorems 5.1.1 and 5.2.8 we only need to show (5.2.22) and (5.2.23), since (5.2.24) follows from (5.2.22) by letting $n \to \infty$. It follows from the definition of sequences $\{t_n\}$ and $\{s_n\}$ (see (10) and (21)) that $t_0 = s_0$, $t_1 = s_1$, $s_2 \le t_2$ and $s_2 - s_1 \le t_2 - t_1$, since $\gamma_0 \le \gamma$,

$$\frac{1}{1-\gamma_0 s_0} \le \frac{1}{1-\gamma t_0}, \quad \frac{1}{1-\gamma s_1} = \frac{1}{1-\gamma t_1} \tag{5.2.25}$$

and

$$\frac{1}{2-\dfrac{1}{(1-\gamma_0 s_1)^2}} \le \frac{1}{2-\dfrac{1}{(1-\gamma_0 t_1)^2}}. \tag{5.2.26}$$

Hence, (5.2.22) and (5.2.23) hold for $n = 0,1,2$. Suppose that (5.2.22) and (5.2.23) hold for all $k \le n$. Then, we have that $s_{k+1} \le t_{k+1}$ and $s_{k+1} - s_k \le t_{k+1} - t_k$, since $\gamma_0 \le \gamma$,

$$\frac{1}{1-\gamma s_{k-1}} \le \frac{1}{1-\gamma t_{k-1}}, \quad \frac{1}{1-\gamma s_k} \le \frac{1}{1-\gamma t_k} \tag{5.2.27}$$

and

$$\frac{1}{2-\dfrac{1}{(1-\gamma_0 s_k)^2}} \le \frac{1}{2-\dfrac{1}{(1-\gamma_0 t_k)^2}}. \tag{5.2.28}$$

□

Remark 5.2.11. *In view of (5.2.22)–(5.2.24), our error analysis is tighter and the information on the location of the solution $x^\star$ is at least as precise as the old one. Notice also that estimates (5.2.22) and (5.2.23) hold as strict inequalities for $n > 1$ if $\gamma_0 < \gamma$ (see also the numerical examples) and these advantages hold under the same or less computational cost as before (see Remark 5.2.9).*

Next, we present our [11, Theorem 3.2]. This theorem shall be used to show that (5.1.5) can be weakened.

Theorem 5.2.12. *Let $F : \overline{U}(x_0,R) \subseteq \mathcal{X} \longrightarrow \mathcal{Y}$ be Fréchet-differentiable. Suppose there exist parameters $L \geq L_0 > 0$ and $\eta > 0$ such that for all $x,y \in \overline{U}(x_0,R)$*

$$F'(x_0)^{-1} \in \mathcal{L}(\mathcal{Y},\mathcal{X}), \quad \| F'(x_0)^{-1} F(x_0) \| \leq \eta,$$

$$\| F'(x_0)^{-1} (F'(x) - F'(x_0)) \| \leq L_0 \| x - x_0 \|, \tag{5.2.29}$$

$$\| F'(x_0)^{-1} (F'(x) - F'(y)) \| \leq L \| x - y \|, \tag{5.2.30}$$

$$\overline{s^\star} := \lim_{n \longrightarrow \infty} \overline{s_n} \leq R$$

and

$$h_1 = 2 L_1 \eta \leq 1, \tag{5.2.31}$$

where

$$\overline{s_0} = 0, \quad \overline{s_1} = \eta, \quad \overline{s_2} = \eta + \frac{L_0 \eta^2}{2(1 - L_0 \eta)},$$
$$\overline{s_{n+1}} = \overline{s_n} + \frac{L(\overline{s_n} - \overline{s_{n-1}})^2}{2(1 - L_0 \overline{s_n})} \quad \text{for each} \quad n = 2,3,\cdots$$

and $L_1 = \dfrac{1}{8}(4L_0 + \sqrt{L_0 L} + \sqrt{L_0 L + 8 L_0^2})$. Then, the following assertions hold

(a) *Sequence $\{\overline{s_n}\}$ is increasing convergent to its unique least upper bound $s^\star$, which satisfies*

$$\overline{s_2} \leq \overline{s^\star} \leq \overline{s^{\star\star}} = \delta \eta,$$

where

$$\delta = 1 + \frac{L_0 \eta}{2(1-\beta)(1 - L_0 \eta)}$$

and

$$\beta = \frac{2L}{L + \sqrt{L^2 + 8 L_0 L}}.$$

(b) *Sequence $\{x_n\}$ generated by Newton's method is well defined, remains in $\overline{U}(x_0,\overline{s^\star})$ for each $n = 0,1,\cdots$ and converges to a solution $x^\star \in \overline{U}(x_0,\overline{s^\star})$ of equation $F(x) = 0$. Moreover, the following estimates hold for each $n = 0,1,\cdots$*

$$\| x_{n+1} - x_n \| \leq \frac{L \| x_n - x_{n-1} \|}{2(1 - L_0 \| x_n - x_0 \|)} \leq \overline{s_{n+1}} - \overline{s_n}$$

and

$$\| x_n - x^\star \| \leq \overline{s^\star} - \overline{s_n}.$$

(c) *If there exists* $\varsigma > \overline{s^\star}$ *such that* $\varsigma < R$ *and* $L_0(\overline{s^\star}+\varsigma) \leq 2$, *then, the solution* $x^\star$ *of equation* $F(x)=0$ *is unique in* $U(x_0,\varsigma)$.

Remark 5.2.13. (a) *If* $L_0 = L$, *convergence criterion (5.2.31) reduces to the famous for its simplicity and clarity Kantorovich hypothesis [21] for solving equations*

$$h = 2L\eta \leq 1. \tag{5.2.32}$$

Notice that

$$L_0 \leq L \tag{5.2.33}$$

holds in general and L_0/L *can be arbitrarily small [11, 13]. We also have that*

$$h \leq 1 \Longrightarrow h_1 \leq 1 \tag{5.2.34}$$

and

$$\frac{h_1}{h_0} \longrightarrow 0 \quad \text{as} \quad \frac{L_0}{L} \longrightarrow 0.$$

Moreover, the Kantorovich majorizing sequence is given by

$$\overline{\overline{s}}_0 = 0, \quad \overline{\overline{s}}_{n+1} = \overline{\overline{s}}_n - \frac{p(\overline{\overline{s}}_n)}{p'(\overline{\overline{s}}_n)} = \overline{\overline{s}}_n - \frac{L(\overline{\overline{s}}_n - \overline{\overline{s}}_{n-1})^2}{p'(\overline{\overline{s}}_n)} \quad \text{for each} \quad n = 0,1,2,\cdots,$$

where $p(t) = (L/2)t^2 - t + \eta$. *If (5.2.32) is satisfied then (see [11])*

$$\overline{s}_n \leq \overline{\overline{s}}_n,$$

$$\overline{s}_{n+1} - \overline{s}_n \leq \overline{\overline{s}}_{n+1} - \overline{\overline{s}}_n$$

and

$$\overline{s^\star} \leq \overline{\overline{s^\star}} = \lim_{n\longrightarrow\infty} \overline{\overline{s}}_n = \frac{2\eta}{1+\sqrt{1-h}}.$$

(b) *Let us show that Wang's convergence criterion (5.1.5) can be weakened under the Kantorovich hypotheses. In particular, suppose that (5.2.30) and (5.2.32) are satisfied. Then, (5.2.31) is also satisfied. Moreover, if* F *is twice Fréchet-differentiable on* $U(x_0, 1/\gamma)$, *then Wang's condition (5.1.3) is certainly satisfied, if we choose* $\gamma = L/2$. *Then, condition (5.2.32) becomes*

$$\gamma\eta \leq \frac{1}{4}, \tag{5.2.35}$$

which improves (5.1.5). We must also show that $\overline{\overline{s^\star}} \leq 1/\gamma$. *But the preceding inequality reduces to showing that* $h-2 \leq 2\sqrt{1-h}$, *which is true by (5.2.32). Clearly, in view of (5.2.31), (5.2.33)–(5.2.35), criterion (5.1.5) (i.e., (5.2.35)) can be improved even further, for* $\gamma_0 = L_0/2$, *if* $L_0 \leq L$.

(c) *Suppose Wang's condition (5.1.3) is satisfied as well as criterion (5.1.5) on $\overline{U}(x_0,\overline{r}_0)$. Recall that $\overline{r}_0 = (1-\frac{1}{\sqrt{2}})\frac{1}{\gamma}$. Then, for $r \in [0,\overline{r}_0]$, we have that $1/(1-\gamma r) \leq \sqrt{2}$. Then, in view of (5.1.3) and (5.2.30), we can choose $L = 4\sqrt{2}\gamma$. Then, (5.1.5) becomes*

$$L\eta \leq 4(3\sqrt{2}-4) = .970562748.$$

However, we must also have that $t^\star \leq 1/L$, where $t^\star$ is given in (5.1.7). By direct algebraic manipulation, we see that the preceding inequality is satisfied, if

$$.078526267 = \frac{2\sqrt{2}(2\sqrt{2}-1)}{8-\sqrt{2}} \leq L\eta \leq \frac{4\sqrt{2}}{8\sqrt{2}-1} = .548479169.$$

Hence, the last two estimates on "$L\eta$" are satisfied, if the preceding inequality is satisfied, which is weaker than (5.2.32) for $L\eta \in (.5, .548479169]$. However, the preceding inequality may not be weaker than (5.2.31) (if $\gamma_0 = L_0/2$) for L_0 sufficiently smaller than L.

Next, we present the following specialization of Theorem 5.2.12 for Newton-Kantorovich method and analytic operators defined on $\overline{U}(x_0,R)$. In the following theorem and for $\varepsilon = \gamma_0/\gamma$, interval I is defined by

$$I = \begin{cases} (0,1) & if \quad \varepsilon \leq (\sqrt{2}-1)/\sqrt{2} \\ (0,\varepsilon^{-1}(\sqrt{2}-1)/\sqrt{2}) & if \quad \varepsilon > (\sqrt{2}-1)/\sqrt{2}. \end{cases}$$

Theorem 5.2.14. *Let $F : \overline{U}(x_0,R) \subseteq \mathcal{X} \to \mathcal{Y}$ be Fréchet-differentiable in $U(x_0,R)$. Define functions f, H and H_1 on interval I by*

$$f(r) = g(r)\,r\alpha - \frac{1}{2}, \; H_1(r) = \frac{\alpha\varepsilon(2-r)}{(1-\varepsilon r)^2}$$

and

$$H(r) = \left[1 + \frac{\dfrac{\alpha\varepsilon(2-r)}{(1-\varepsilon r)^2}}{2(1-\beta)\left(1-\dfrac{\alpha\varepsilon(2-r)}{(1-\varepsilon r)^2}\right)}\right]\alpha - r,$$

where

$$g(r) = \frac{1}{8}\left[\frac{4\varepsilon(2-r)}{(1-\varepsilon r)^2} + \sqrt{\frac{2\varepsilon(2-r)}{(1-\varepsilon r)^2(1-r)^3}} + \sqrt{\frac{2\varepsilon(2-r)}{(1-\varepsilon r)^2(1-r)^3} + \frac{8\varepsilon^2(2-r)^2}{(1-\varepsilon r)^4}}\right].$$

Suppose that there exist intervals I_f, I_H and I_{H_1} such that for some $\alpha \in I$

$$I_f \subset I, \qquad I_H \subset I, \qquad I_{H_1} \subset I,$$

$$f(r) \leq 0 \text{ for each } r \in I_f, \tag{5.2.36}$$

$$H(r) \leq 0 \text{ for each } r \in I_H, \tag{5.2.37}$$

$$H_1(r) \leq 1 \text{ for each } r \in I_{H_1} \tag{5.2.38}$$

and

$$I_0 = I_f \cap I_H \cap I_{H_1} \neq \emptyset.$$

Denote by $r^\star = r^\star(\alpha)$ *the largest element in* I_0. *Moreover, suppose there exists a point* $x_0 \in \overline{U}(x_0, R)$ *such that* $F'(x_0)^{-1} \in \mathcal{L}(\mathcal{Y}, \mathcal{X})$ *and*

$$\frac{r^\star}{\gamma} \leq R. \tag{5.2.39}$$

Then, the following assertions hold

(a) Scalar sequence $\{\overline{s_n}\}$ *is increasingly convergent to* $\overline{s^\star}$ *which satisfies*

$$\eta \leq \overline{s^\star} \leq \overline{s^{\star\star}} = \delta\eta,$$

where

$$\delta = 1 + \frac{L_0\eta}{2(1-\beta)(1-L_0\eta)},$$

$$\beta = \frac{2L}{L+\sqrt{L^2+8L_0L}} = \frac{2M}{M+\sqrt{M^2+8M_0M}},$$

$$M_0 = \frac{L_0}{\gamma}, \quad M = \frac{L}{\gamma}$$

$$L_0 = \frac{\gamma(2-\varepsilon r^\star)}{(1-\varepsilon r^\star)^2} \quad \text{and} \quad L = \frac{2\gamma}{(1-r^\star)^3}, \tag{5.2.40}$$

where sequence $\{\overline{s_n}\}$, $\overline{s^\star}$ *and* $\overline{s^{\star\star}}$ *are given in Theorem 5.2.12.*

(b) The conclusions (a) and (b) of Theorem 5.2.12 hold.

Proof. Notice that if follows from (5.2.38) that $H(r) + r \geq 0$ for each $r \in I_H$. We have by $\overline{s^\star} \leq \overline{s^{\star\star}}$ and (5.2.36) that $\gamma\overline{s^\star} \leq \gamma\overline{s^{\star\star}} \leq H(r^\star) \leq r^\star < 1$. Then, we showed in [5] (see also [13]) that (5.2.2) and (5.2.3) are satisfied for functions L_0 and L given by (5.2.40). Using these choices of L_0 and L we must show that (5.2.9) is satisfied. That is we must have

$$h_3 = g(r^\star)\alpha \leq \frac{\alpha}{2r^\star} \leq \frac{1}{2},$$

which is true by the choice of $r^\star$ in (5.2.36) and (5.2.39). Notice also that by (5.2.37) and (5.2.38), we have $\overline{s^\star} \leq r^\star/\gamma$. The rest follows from Theorem 5.2.8. □

Remark 5.2.15. *(a) It follows from the proof of Theorem 5.2.8 that function* f *can be replaced by* f^1 *defined by*

$$f^1(r) = g(r)\alpha - \frac{1}{2}. \tag{5.2.41}$$

In practice, we shall employ both functions to see which one will produce the largest possible upper bound $r^\star$ *for* α.

(b) It is worth noticing that

$$L_0(r) < L(r) \text{ for all } r \in (0,1).$$

(c) Notice that it follows from (5.2.37) *and* (5.2.38) *that* $\alpha \leq r^{\star}$.

In the case when F is Fréchet-differentiable on $\mathcal{X}$, we have the following result.

Proposition 5.2.16. *Let $F : \mathcal{X} \to \mathcal{Y}$ be analytic. Suppose that there exists a point $x_0 \in \mathcal{D}$ such that $F'(x_0)^{-1} \in \mathcal{L}(\mathcal{Y},\mathcal{X})$ and for each r in some interval I_{H_2} such that $\emptyset \neq I_{H_2} \subset I$, we have that*

$$H_2(r) = g(r) - \frac{1}{2r} \leq 0 \tag{5.2.42}$$

Denote by r_1 the largest element in I_{H_2}. Moreover, suppose

$$\alpha \leq r_1 = 0.179939475\ldots. \tag{5.2.43}$$

Then, the conclusions of Theorem 5.2.8 hold.

Proof. It follows by the choice of r_1 that

$$g(r_1) \leq \frac{1}{2r_1}. \tag{5.2.44}$$

Using (5.2.43) and (5.2.44) we get

$$h_3 = g(r_1)\alpha = \frac{1}{2r_1}\alpha \leq \frac{1}{2}.$$

Notice that condition (5.2.39) is satisfied automatically. □

The results obtained in this chapter can be connected to the following notion [14].

Definition 5.2.17. *A point x_0 is said to be an approximate zero of the first kind for F if $\{x_n\}$ is well defined for each $n = 0,1,\cdots$ and satisfies*

$$\|x_{n+1} - x_n\| \leq \Xi^{2^n-1}\|x_1 - x_0\| \quad for\ some \quad \Xi \in (0,1). \tag{5.2.45}$$

Notice that if we start from an approximate zero x_0 of the first kind then, the convergence of Newton-Kantorovich method to $x^{\star}$ is very fast.

In view of the estimate

$$\|x_{n+1} - x_n\| \leq \frac{L}{2(1 - L_0\,\overline{s_n})}\|x_n - x_{n-1}\|^2$$

we get that

$$\frac{L}{2(1-L_0\overline{s})} \leq \frac{L}{2\,(1-L_0\,\overline{s^{\star}})} \leq \frac{\gamma}{(1-r)^3}\,\frac{1}{1-\frac{2-r}{(1-r)^2}(H(r)+r)} \leq \frac{1}{\eta}$$

provided that

$$\frac{2-r}{(1-r)^2}(H(r)+r) < 1 \tag{5.2.46}$$

and

$$0 \leq \alpha\,\psi(r) \leq \Xi < 1, \tag{5.2.47}$$

where
$$\psi(r) = \frac{1}{(1-r)^3[1-\frac{2-r}{(1-r)^2}(H(r)+r)]}.$$

Conditions (5.2.46) and (5.2.47) must hold respectively in Theorem 5.2.8 and Proposition 5.2.10 for $r = r^\star, r_1$. Then, x_0 is an approximate zero of the first kind in all these results. If $\gamma_0 = \gamma$, then (5.2.42) holds for $r_1 = .179939475\cdots$. Using (5.2.45) we notice that (5.2.46) and (5.2.47) hold at r_1. It then follows that (5.2.45) is satisfied with factor Ξ/η, where Ξ is given by $\Xi = \alpha\psi(r_1)$.

Remark 5.2.18. *If $F : \mathcal{D} \subseteq \mathcal{X} \to \mathcal{Y}$ is an analytic operator and $x_0 \in \mathcal{D}$. Let γ be defined by (see [28, 29])*
$$\gamma = \sup_{j>1} \| \frac{1}{j!} F'(x_0)^{-1} F^{(j)}(x_0) \|^{\frac{1}{j-1}},$$

or $\gamma = \infty$ if $F'(x_0)$ is not invertible or the supremum in γ does not exist. Then, if $\mathcal{D} = \mathcal{X}$, the sufficient convergence condition of Newton-Kantorovich method is given by $\alpha \leq 0.130707$. Rheinboldt in [25] improved Smale's result by showing convergence of Newton's method when $\alpha \leq 0.15229240$. Here, we showed convergence for $\alpha \leq r_1 = .179939475$.

5.3. Local Convergence Analysis of Newton's Method

We shall use similar definitions to (5.1.3), (5.2.2) and (5.2.3) to study the local convergence of Newton's method.

Definition 5.3.1. *(see [30]) Let $F : \overline{U}(x^\star, R) \subseteq \mathcal{X} \to \mathcal{Y}$ be twice Fréchet-differentiable on $U(x^\star, R)$ and $F(x^\star) = 0$. Let $\gamma > 0$ and let $0 < r \leq 1/\gamma$ be such that $r \leq R$. The operator F'' is said to satisfy the γ-Lipschitz condition at $x^\star$ on $U(x^\star, r)$ if*

$$\| F'(x^\star)^{-1} F''(x) \| \leq \frac{2\gamma}{(1-\gamma \| x - x^\star \|)^3} \quad for\ each \quad x \in U(x^\star, r). \tag{5.3.1}$$

Definition 5.3.2. *Let $F : \overline{U}(x^\star, R) \longrightarrow \mathcal{Y}$ be Fréchet-differentiable on $U(x^\star, R)$ and $F(x^\star) = 0$. We say that F' satisfies the γ-Lipschitz condition at $x^\star$ if there exists an increasing function $\ell : [0, R] \longrightarrow [0, +\infty)$ such that*

$$\begin{aligned} &\| F'(x^\star)^{-1} (F'(x) - F'(y)) \| \leq \ell(r) \| x - y \| \\ &for\ each\ x, y \in \overline{U}(x^\star, r),\ 0 < r \leq R. \end{aligned} \tag{5.3.2}$$

Definition 5.3.3. *Let $F : \overline{U}(x^\star, R) \longrightarrow \mathcal{Y}$ be Fréchet-differentiable on $U(x^\star, R)$ and $F(x^\star) = 0$. We say that F' satisfies the $\gamma^\star$-center-Lipschitz condition at $x^\star$ if there exists an increasing function $\ell^\star : [0, R] \longrightarrow [0, +\infty)$ such that*

$$\begin{aligned} &\| F'(x^\star)^{-1} (F'(x) - F'(x^\star)) \| \leq \ell^\star(r) \| x - x^\star \| \\ &for\ each\ x \in \overline{U}(x^\star, r),\ 0 < r \leq R. \end{aligned} \tag{5.3.3}$$

Remark 5.3.4. *(a) Notice again that $\ell^\star(r) \leq \ell(r)$ and $\ell/\ell^\star$ can be arbitrarily large.*

(b) *In order for us to cover the local convergence analysis of Newton's method, let us define function* $f_\varepsilon : I_\varepsilon = [0, \frac{1}{\varepsilon}(1-\frac{1}{\sqrt{2}})] \longrightarrow \mathbb{R}$ *by*

$$f_\varepsilon(t) = (1-\varepsilon t)^2(2-t)t - (1-t)^2(2(1-\varepsilon t)^2-1). \tag{5.3.4}$$

Suppose that

$$\varepsilon > \frac{1}{2}(1-\frac{1}{\sqrt{2}}). \tag{5.3.5}$$

Then, we have that

$$f_\varepsilon(0) = -1 < 0 \quad \textit{and} \quad f_\varepsilon(\frac{1}{\varepsilon}(1-\frac{1}{\sqrt{2}})) = \frac{1}{\varepsilon\sqrt{2}}(1-\frac{1}{\sqrt{2}})(2-\frac{1}{\varepsilon}(1-\frac{1}{\sqrt{2}})) > 0.$$

Hence, it follows from the intermediate value theorem that function f_ε *has a zero in* $\overset{\circ}{I}_\varepsilon$. *Denote by* $\mu_\varepsilon^\star$ *the minimal such zero. Define function* $g_\varepsilon : I_\varepsilon \longrightarrow \mathbb{R}$ *by*

$$g_\varepsilon(t) = \frac{(1-\varepsilon t)^2(2-t)t}{(1-t)^2(2(1-\varepsilon t)^2-1)}. \tag{5.3.6}$$

Then, we have that

$$0 \le g_\varepsilon(t) < 1 \quad for\, each \quad t \in [0, \mu_\varepsilon^\star]. \tag{5.3.7}$$

Set

$$R_\varepsilon = \frac{\mu_\varepsilon^\star}{\gamma}. \tag{5.3.8}$$

It follows from the definition of f_ε, $\mu_\varepsilon^\star$ *and* g_ε *that*

$$R_1 = \frac{3-\sqrt{6}}{3\gamma} \le R_\varepsilon. \tag{5.3.9}$$

Moreover, strict inequality holds if $\varepsilon \neq 1$. *Let us assume that* F *satisfies the* $\gamma^\star$*-center-Lipschitz condition at* $x^\star$ *on* $U(x^\star, \frac{1}{\varepsilon\gamma}(1-\frac{1}{\sqrt{2}}))$ *with* $F(x^\star) = 0$ *and the* γ*-Lipschitz condition at* $x^\star$ *on* $U(x^\star, \frac{1}{\gamma})$. *Then, for* $x_0 \in U(x^\star, R_\varepsilon)$, *we have the identity*

$$\begin{aligned} x_{n+1} - x^\star &= F'(x_n)^{-1}(F(x^\star) - F(x_n) - F'(x_n)(x^\star - x_n)) \\ &= F'(x_n)^{-1}F'(x^\star)\int_0^1 F'(x^\star)^{-1}(F'(x_{n,\star}^\tau) - F'(x_n))\, d\tau\,(x^\star - x_n), \end{aligned} \tag{5.3.10}$$

where $x_{n,\star}^\tau = x_n + \tau(x^\star - x_n)$. *Set also*

$$x_{n,\star}^{\tau s} = x_n + \tau s(x^\star - x_n) \quad for\, each \quad 0 \le t \le 1, \quad 0 \le s \le 1.$$

As in Theorem 5.2.8 but using (5.3.2) and (5.3.3) for

$$\ell(r) = \frac{2\gamma}{(1-\gamma r)^3} \quad \text{and} \quad \ell^\star(r) = \frac{\gamma^\star(2-r)}{(1-\gamma^\star r)^2},$$

we get in turn as in (5.2.17) and (5.2.19), respectively, that

$$\| F'(x_n)^{-1}F'(x^\star) \| \leq \left(2 - \frac{1}{(1-\varepsilon\gamma \| x_n - x^\star \|)^2}\right)^{-1} \tag{5.3.11}$$

and

$$\begin{aligned} &\| \int_0^1 (F'(x_{n,\star}^\tau) - F'(x_n))\, d\tau\, (x^\star - x_n) \| \leq \\ &\int_0^1 \int_0^1 \frac{2\gamma \| x_{n,\star}^{\tau s} - x^\star \| \, ds\, d\tau}{(1-\gamma s \| x_{n,\star}^{\tau s} - x^\star \|)^3} \| x_{n,\star}^{\tau s} - x^\star \| \leq \\ &\left(\frac{1}{(1-\gamma \| x_n - x^\star \|)^2} - 1\right) \| x_n - x^\star \| . \end{aligned} \tag{5.3.12}$$

That is we have by (5.3.10)–(5.3.12) that

$$\begin{aligned} &\| x_{n+1} - x^\star \| \\ &\leq \frac{(1-\varepsilon\gamma \| x_n - x^\star \|)^2 (1-(1-\gamma \| x_n - x^\star \|)^2)}{(2(1-\varepsilon\gamma \| x_n - x^\star \|)^2 - 1)(1-\gamma \| x_n - x^\star \|)^2} \| x_n - x^\star \| \\ &< g_\varepsilon(\mu_\varepsilon^\star) \| x_n - x^\star \| = \| x_n - x^\star \| < R_\varepsilon. \end{aligned} \tag{5.3.13}$$

Estimate (5.3.13) shows that $x_{n+1} \in U(x^\star, R_\varepsilon)$ *and* $\lim_{n\to\infty} x_n = x^\star$.

Hence we arrived at the following result on the local convergence for Newton's method.

Theorem 5.3.5. *Let* $F : \overline{U}(x^\star, R) \subseteq \mathcal{X} \longrightarrow \mathcal{Y}$ *be Fréchet-differentiable on* $U(x^\star, R)$. *Suppose that*

(a) There exists $x^\star \in \overline{U}(x_0, R)$ *such that* $F(x^\star) = 0$ *and* $F'(x^\star)^{-1} \in \mathcal{L}(\mathcal{Y}, \mathcal{X})$.

(b) Operator F satisfies the center $\gamma^\star$*-center-Lipschitz condition at* $x^\star$ *on* $U(x^\star, \frac{1}{\varepsilon\gamma}(1 - \frac{1}{\sqrt{2}}))$ *for* ε *satisfying (5.3.5) and the* γ*-Lipschitz condition at* $x^\star$ *on* $U(x^\star, \frac{1}{\gamma})$.

Then, if $x_0 \in U(x^\star, R_\varepsilon)$, *sequence* $\{x_n\}$ *generated by Newton's method is well defined, remains in* $U(x^\star, R_\varepsilon)$ *for each* $n = 0, 1, \cdots$ *and converges to* $x^\star$. *Moreover, the following estimate holds*

$$\| x_{n+1} - x_n \| \leq \frac{\gamma(2-\gamma \| x_n - x^\star \|)(1-\varepsilon\gamma \| x_n - x^\star \|)^2}{(1-\gamma \| x_n - x^\star \|)^2 (2(1-\varepsilon \| x_n - x^\star \|)^2 - 1)} \| x_n - x^\star \|^2 . \tag{5.3.14}$$

Remark 5.3.6. *If* $\varepsilon = 1$ *(i.e.* $\gamma^\star = \gamma$*), our results reduces to the ones given by Wang [31] (see also [30, 32]). Otherwise, if*

$$\frac{1}{2}(1 - \frac{1}{\sqrt{2}}) \leq \frac{\gamma^\star}{\gamma} < 1, \tag{5.3.15}$$

then, according to (5.3.9) our convergence radius is larger. Moreover, our error bounds are tighter if $\gamma^\star < \gamma$.

Remark 5.3.7. *Let us define function* $f_{\varepsilon_1}\,:\, I_{\varepsilon_1}=[0,\frac{1}{\varepsilon_1}]\longrightarrow\mathbb{R}$ *for* $\varepsilon_1>0$ *by*

$$f_{\varepsilon_1}(t)=(2-t)\,t-(1-t)^2\,(1-\varepsilon_1\,t). \tag{5.3.16}$$

Suppose that

$$\varepsilon_1>\frac{1}{2}. \tag{5.3.17}$$

Then, we have

$$f_{\varepsilon_1}(0)=-1<0 \quad \text{and} \quad f_{\varepsilon_1}(\frac{1}{\varepsilon_1})>0.$$

Denote by $\mu^\star_{\varepsilon_1}$ *the minimal zero of* f_{ε_1} *on* $\overset{\circ}{I}_{\varepsilon_1}$*. Define function* $g_{\varepsilon_1}\,:\, I_{\varepsilon_1}\longrightarrow\mathbb{R}$ *by*

$$g_{\varepsilon_1}(t)=\frac{(2-t)\,t}{(1-t)^2\,(1-\varepsilon_1\,t)}. \tag{5.3.18}$$

Then, we have that

$$0\leq g_{\varepsilon_1}(t)<1 \quad \text{for each} \quad t\in[0,\mu^\star_{\varepsilon_1}].$$

Set

$$L_0=\varepsilon_1\,\gamma \quad \text{and} \quad R_{\varepsilon_1}=\frac{\mu^\star_{\varepsilon_1}}{\gamma}. \tag{5.3.19}$$

Hence, we arrived at the following result.

Theorem 5.3.8. *Suppose that*

(a) *There exists* $x^\star\in\overline{U}(x_0,R)$ *such that* $F(x^\star)=0$ *and* $F'(x^\star)^{-1}\in\mathcal{L}(\mathcal{Y},\mathcal{X})$.

(b) *Operator* F *satisfies the center* L_0*-Lipschitz condition at* $x^\star$ *on* $U(x^\star,\frac{1}{\varepsilon_1})$ *for* ε_1 *satisfying (5.3.17) and the* γ*-Lipschitz condition at* $x^\star$ *on* $U(x^\star,\frac{1}{\gamma})$.

Then, if $x_0\in U(x^\star,R_{\varepsilon_1})$*, sequence* $\{x_n\}$ *generated by Newton's method is well defined, remains in* $U(x^\star,R_{\varepsilon_1})$ *for each* $n=0,1,\cdots$ *and converges to* $x^\star$*. Moreover, the following estimate holds*

$$\|\,x_{n+1}-x_n\,\|\leq\frac{\gamma(2-\gamma\,\|\,x_n-x^\star\,\|)\,\;\|\,x_n-x^\star\,\|^2}{(1-\gamma\,\|\,x_n-x^\star\,\|)^2\,(1-\varepsilon_1\;\|\,x_n-x^\star\,\|)}. \tag{5.3.20}$$

5.4. Numerical Examples

In this section we provide numerical examples.

Example 5.4.1. (a) *Consider* $\gamma=1.8$, $\gamma_0=.44$ *and* $\eta=.1$*. Using (5.2.10) and (5.2.11), we get that*

$$\frac{\gamma_0}{\gamma}=.2444444444\leq 1-\frac{1}{\sqrt{2}}=.2928932190, \quad \frac{1}{\gamma}=.5555555556,$$

Table 5.4.1. Comparison Table

n	s_n	t_n	$s_{n+1}-s_n$	$t_{n+1}-t_n$
1	.1	.1	.0051130691	.0059006211
2	.1051130691	.1059006211	.0000169735	.0000230132
3	.1051300426	.1059236343	2e-10	4e-10
4	.1051300428	.1059236347	0	0
5	$\sim$	$\sim$	$\sim$	$\sim$

$$s_2 = .1059236776, \quad s_3 = .1060526606, \quad s_4 = .1060527234$$

and

$$s_n = s_4 = .1060527234 \quad \text{for each} \quad n = 5,6,7,\cdots.$$

That is $s_n < 1/\gamma$ for each $n = 1,2,\cdots$ and condition (5.2.11) holds. Hence, our Theorem 5.2.8 is applicable. We have that

$$\alpha = .18 > 3-2\sqrt{2} = .171572876.$$

Hence the older convergence criteria in [32] do not hold.

(b) *Consider now $\gamma = .5$, $\gamma_0 = .44$ and $\eta = .1$. Using (5.2.10) and (5.2.11), we get that*

$$\frac{\gamma_0}{\gamma} = .88 > 1-\frac{1}{\sqrt{2}} = .2928932190, \quad (1-\frac{1}{\sqrt{2}})\frac{1}{\gamma_0} = .665666406,$$

$$s_2 = .1051130691, \quad s_3 = .1051300426, \quad s_4 = .1051300428$$

and

$$s_n = s_4 = .1051300428 \quad \text{for each} \quad n = 5,6,\cdots.$$

That is $s_n < (1-(1/\sqrt{2}))/\gamma_0$ for each $n = 1,2,\cdots$ and condition (5.2.11) holds. Hence, our Theorem 5.2.8 is applicable. we also have that

$$\alpha = .05 \leq .171572876.$$

Hence the convergence criterion in [31] is also satisfied. We can now compare our results of Theorem 5.2.8 (see also sequence $\{s_n\}$ given by (5.2.10)) to ones given in [31, 32] (see also $\{t_n\}$ given by (5.1.10)). Table 5.4.1 shows that our error bounds using sequence $\{s_n\}$ are tighter than those given in [32].

Example 5.4.2. *Let function $h : \mathbb{R} \longrightarrow \mathbb{R}$ be defined by*

$$h(x) = \begin{cases} 0 & if \quad x \leq 0 \\ x & if \quad x \geq 0. \end{cases}$$

Define function F by

$$F(x) = \begin{cases} \varpi - x + \frac{1}{18}x^3 + \frac{x^2}{1-x} & if \quad x \leq \frac{1}{2} \\ \varpi - \frac{71}{144} + 2x^2 & if \quad x \geq \frac{1}{2}, \end{cases} \tag{5.4.1}$$

where $\varpi > 0$ *is a constant. Then, we have that*

$$F'(x) = \begin{cases} -2 + \dfrac{1}{(1-x)^2} + \dfrac{x^2}{6} & if \quad x \leq \dfrac{1}{2} \\ 4x & if \quad x \geq \dfrac{1}{2} \end{cases} \tag{5.4.2}$$

and

$$F''(x) = \begin{cases} \dfrac{2}{(1-x)^2} + \dfrac{x}{3} & if \quad x \leq \dfrac{1}{2} \\ 4 & if \quad x \geq \dfrac{1}{2}. \end{cases} \tag{5.4.3}$$

We shall first show that F' *satisfies the L-Lipschitz condition (5.2.2) on* $U(0,1)$*, where*

$$L(u) = \frac{2}{(1-u)^3} + \frac{1}{6} \quad for\ each\ u \in [0,1) \tag{5.4.4}$$

and the L_0*-center-Lipschitz condition (5.2.3) on* $U(0,1)$*, where*

$$L_0(u) = \frac{2}{(1-u)^3} + \frac{1}{12} \quad for\ each\ u \in [0,1). \tag{5.4.5}$$

It follows from (5.4.3) that

$$L(u) < L(v) \quad for\ each \quad 0 \leq u < v < 1 \tag{5.4.6}$$

and

$$0 < F''(u) < F''(|u|) < L(|u|) \quad for\ each \quad \frac{1}{2} \neq u < 1. \tag{5.4.7}$$

Let $x, y \in U(0,1)$ *with* $|y| + |x-y| < 1$*. Then, it follows from (5.4.6) and (5.4.7) that*

$$\begin{aligned} |F'(x) - F'(y)| &\leq |x-y| \int_0^1 F''(y + t\,(x-y))\,dt \\ &\leq |x-y| \int_0^1 L(|y| + t\,|x-y|)\,dt. \end{aligned} \tag{5.4.8}$$

Hence, F' *satisfies the L-Lipschitz condition (5.2.2) on* $U(0,1)$*. Similarily, using (5.4.2) and (5.4.5), we deduce that* F' *satisfies the* L_0*-center-Lipschitz condition (5.2.3) on* $U(0,1)$*. Notice that*

$$L_0(u) < L(u) \quad for\ each \quad u \in [0,1). \tag{5.4.9}$$

Table 5.4.2 show that our error bounds $s_{n+1} - s_n$ *are finer than* $t_{n+1} - t_n$*.*

Example 5.4.3. *Let* $\mathcal{X} = \mathcal{Y} = \mathbb{R}^2$*,* $x_0 = (1,0)$*,* $\mathcal{D} = \overline{U}(x_0, 1-\kappa)$ *for* $\kappa \in (0,1)$*. Let us define function F on* $\mathcal{D}$ *as follows*

$$F(x) = (\zeta_1^3 - \zeta_2 - \kappa, \zeta_1 + 3\zeta_2 - \sqrt[3]{\kappa}) \quad with \quad x = (\zeta_1, \zeta_2). \tag{5.4.10}$$

Using (5.4.10) *we see that the* γ*-Lipschitz condition is satisfied for* $\gamma = 2 - \kappa$*. We also have that* $\eta = (1-\kappa)/3$*.*

Table 5.4.2. Comparison Table

n	s_n	t_n	$s_{n+1}-s_n$	$t_{n+1}-t_n$
0	0	0	.05	.05
1	.05	.05	.00308148876	.00321390287
2	.05308148876	.05321390287	.00001307052	.00001479064
3	.05309455928	.05322869351		

Table 5.4.3. Comparison Table

n	s_n	t_n	$s_{n+1}-s_n$	$t_{n+1}-t_n$
0	$0.000000e+00$	$0.000000e+00$	$1.000000e-01$	$1.000000e-01$
1	$1.000000e-01$	$1.000000e-01$	$1.52215005e-02$	$2.201246e-02$
2	$1.152215005e-01$	$1.220125e-01$	$6.507434e-04$	$1.683820e-03$
3	$1.158722439e-01$	$1.236963e-01$	$1.2499e-06$	$1.069600e-05$
4	$1.158734938e-01$	$1.237070e-01$	0	$4.338887e-10$
5	$\sim$	$1.237070e-01$	$\sim$	$7.140132e-19$
6	$\sim$	$1.237070e-01$	$\sim$	$1.933579e-36$
7	$\sim$	$1.237070e-01$	$\sim$	$1.417992e-71$
8	$\sim$	$1.237070e-01$	$\sim$	$7.626002e-142$
9	$\sim$	$1.237070e-01$	$\sim$	$2.205685e-282$

Case I. *Let* $\kappa = .6255$. *Then we notice that (5.1.5) is not satisfied since* $\alpha = .1715834166 > 3-2\sqrt{2} = .171572875$. *Hence there is no guarantee that Newton's method starting from* x_0 *will converge to* $x^\star = (\sqrt[3]{\kappa}, 0) = (.85521599, 0)$ *(cf. [14, 19, 26, 27, 30, 31, 32]). However, our results can apply. Indeed using the definition of Lipschitz and center-Lipschitz conditions we have that* $L_0 = 3-\kappa$ *and* $L = 4\sqrt{2}(2-\kappa)$. *Hence,* (5.2.31) *is satisfied since* $h = L_1\eta = .3396683409 < .5$. *We conclude that Theorem 5.2.12 is applicable and iteration* $\{s_n\}$ *converges to* $x^\star$.

Case II. *Let* $\kappa = .7$. *It can be seen that the condition* (5.1.5) *holds since* $\alpha = .13 \leq 3-2\sqrt{2}$. *We also obtain that* $h = .2626128133 < .5$. *We get in turn that* $1/\gamma = 0.7692307$,

$$\left(1-\frac{1}{\sqrt{2}}\right)\frac{1}{\gamma_0} = 0.2899932 \quad \textit{and} \quad 1-\frac{1}{\sqrt{2}} = .29289321 < 0.776923.$$

Then condition (5.2.31) *also holds. Using Theorem 5.2.8, the* γ_0*-center-Lipschitz condition is satisfied if*

$$\left\|F'(x_0)^{-1}\left(F'(x)-F'(x_0)\right)\right\| < \frac{1}{(1-\gamma_0\|x-x_0\|)^2}-1,$$

which is certainly satisfied for say $\gamma_0 = 1.01$. *Note that* $\gamma_0 < 1.3 = \gamma$. *Table 5.4.3 compare the sequences* $\{s_n\}$, $\{t_n\}$ *and the error bounds* $t_{n+1}-t_n$, $s_{n+1}-s_n$. *We also observe that* $\{s_n\}$ *is finer majorizing sequence than* $\{t_n\}$.

Conclusion

A convergence analysis of Newton's method is provided for approximating a locally unique solution of nonlinear equation in a Banach space setting. Using Smale's α-theory and the center-Lipschitz condition, we presented a new convergence analysis with larger convergence domain and weaker sufficient convergence conditions. Moreover, these advantages are obtained under the same computational cost as in earlier studies such as [14, 19, 27, 30, 31, 32]. Numerical examples validating the theoretical results are also provided in this chapter.

References

[1] Amat, S., Busquier, S., Negra, M., Adaptive approximation of nonlinear operators. *Numer. Funct. Anal. Optim.* 25 (2004), 397–405.

[2] Argyros, I.K., A convergence analysis for Newton's method based on Lipschitz, center Lipschitz conditions and analytic operators. *PanAmer. Math. J.* 13 (2003), 35–42.

[3] Argyros, I.K., A unifying local-semilocal convergence analysis and applications for two-point Newton-like methods in Banach space. *J. Math. Anal. Appl.* 298 (2004), 374–397.

[4] Argyros, I.K., On the Newton-Kantorovich hypothesis for solving equations. *J. Comp. Appl. Math.* 169 (2004), 315–332.

[5] Argyros, I.K., Computational theory of iterative methods. Series: Studies in *Comp. Math.*, 15, Editors: C.K. Chui and L. Wuytack, Elsevier Publ. Co. New York, U.S.A, 2007.

[6] Argyros, I.K., On the semilocal convergence of a Newton-type method in Banach spaces under a gamma-type condition. *J. Concr. Appl. Math.* 6 (2008), 33–44.

[7] Argyros, I.K., A new semilocal convergence theorem for Newton's method under a gamma-type condition. *Atti Semin. Mat. Fis. Univ. Modena Reggio Emilia* 56 (2008/09), 31–40.

[8] Argyros, I.K., Semilocal convergence of Newton's method under a weak gamma condition. *Adv. Nonlinear Var. Inequal.* 13 (2010), 65–73.

[9] Argyros, I.K., A semilocal convergence analysis for directional Newton methods. *Math. Comput.* 80 (2011), 327–343.

[10] Argyros, I.K., Hilout, S., Extending the Newton-Kantorovich hypothesis for solving equations. *J. Comput. Appl. Math.* 234 (2010), 2993–3006.

[11] Argyros, I.K., Hilout, S., Weaker conditions for the convergence of Newton's method. *J. Complexity* 28 (2012), 364–387.

[12] Argyros, I.K., Hilout, S., convergence of Newton's method under weak majorant condition. *J. Comput. Appl. Math.* 236 (2012), 1892–1902.

[13] Argyros, I.K., Cho, Y.J., Hilout, S., *Numerical methods for equations and its applications*, CRC Press/Taylor and Francis Publ., New York, 2012.

[14] Cianciaruso, F. Convergence of Newton-Kantorovich approximations to an approximate zero. *Numer. Funct. Anal. Optim.* 28 (2007), 631–645.

[15] Dedieu, J.P., *Points fixes, zéros et la méthode de Newton*, 54, Springer, Berlin, 2006.

[16] Ezquerro, J.A., Gutiérrez, J.M., Hernández, M.A., Romero, N., Rubio, M.J., The Newton method: from Newton to Kantorovich. (Spanish) *Gac. R. Soc. Mat. Esp.* 13 (2010), 53–76.

[17] Ezquerro, J.A., Hernández, M.A., An improvement of the region of accessibility of Chebyshev's method from Newton's method. *Math. Comp.* 78 (2009), 1613–1627.

[18] Ezquerro, J.A., Hernández, M.A., Romero, N., Newton-type methods of high order and domains of semilocal and global convergence. *App. Math. Comp.* 214 (2009), 142–154.

[19] Guo, X., On semilocal convergence of inexact Newton methods. *J. Comput. Math.* 25 (2007), 231–242.

[20] Hernández, M.A., A modification of the classical Kantorovich conditions for Newton's method. *J. Comp. Appl. Math.* 137 (2001), 201–205.

[21] Kantorovich, L.V., Akilov, G.P., *Functional Analysis*, Pergamon Press, Oxford, 1982.

[22] Ortega, L.M., Rheinboldt, W.C., *Iterative Solution of Nonlinear Equations in Several Variables*, Academic press, New York, 1970.

[23] Proinov, P.D., General local convergence theory for a class of iterative processes and its applications to Newton's method. *J. Complexity* 25 (2009) 38–62.

[24] Proinov, P.D., New general convergence theory for iterative processes and its applications to Newton-Kantorovich type theorems. *J. Complexity* 26 (2010), 3–42.

[25] Rheinboldt, W.C., On a theorem of S. Smale about Newton's method for analytic mappings. *Appl. Math. Let.* 1 (1988), 3–42.

[26] Shen, W., Li, C., Kantorovich-type convergence criterion for inexact Newton methods. *Appl. Numer. Math.* 59 (2009) 1599–1611.

[27] Shen, W., Li, C., Smale's α-theory for inexact Newton methods under the γ-condition. *J. Math. Anal. Appl.* 369 (2010), 29–42.

[28] Smale, S., *Newton's method estimates from data at one point. The merging of disciplines: new directions in pure, applied, and computational mathematics* (Laramie, Wyo., 1985) (1986), 185–196, Springer, New York.

[29] Smale, S., Algorithms for solving equations. Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Berkeley, Calif., 1986) (1987), 172–195, *Amer. Math. Soc.*, Providence, RI.

[30] Wang, D.R., Zhao, F.G. The theory of Smale's point estimation and its applications, Linear/nonlinear iterative methods and verification of solution (Matsuyama, 1993). *J. Comput. Appl. Math.* 60 (1995), 253–269.

[31] Wang, X.H., Convergence of Newton's method and inverse function theorem in Banach space. *Math. Comp.* 68 (1999), 169–186.

[32] Wang, X.H., Han, D.F., On dominating sequence method in the point estimate and Smale theorem. *Sci. China Ser. A* 33 (1990), 135–144.

[33] Yakoubsohn, J.C., Finding zeros of analytic functions: α–theory for Secant type method. *J. Complexity* 15 (1999), 239–281.

[34] Zabrejko, P.P., Nguen, D.F., The majorant method in the theory of Newton-Kantorovich approximations and the Pták error estimates. *Numer. Funct. Anal. Optim.* 9 (1987), 671–684.

Chapter 6

Newton-Type Methods on Riemannian Manifolds under Kantorovich-Type Conditions

6.1. Introduction

Let us suppose that F is an operator defined on an open convex subset Ω of a Banach space E. Let us denote by $DF(x_n)$ the first Fréchet derivatives of F at x_n.

Given an integer m and an initial point $x_0 \in E$, we move from x_n to x_{n+1} through an intermediate sequence $\{y_n^i\}_{i=0}^m$, $y_n^0 = x_n$, which is a generalization of Newton ($m = 1$) and simplified Newton ($m = \infty$) methods

$$\begin{cases} y_n^1 = y_n^0 - \mathcal{D}F\left(y_n^0\right)^{-1} F\left(y_n^0\right) \\ y_n^2 = y_n^1 - \mathcal{D}F\left(y_n^0\right)^{-1} F\left(y_n^1\right) \\ \vdots \\ y_n^m = x_{n+1} = y_n^{m-1} - \mathcal{D}F\left(y_n^0\right)^{-1} F\left(y_n^{(m-1)}\right). \end{cases}$$

This family of methods was introduced by E. Shamanskii [43]. Under appropriate conditions, these iterative methods converge to a root x_* of the equation $F(x) = 0$. Moreover, if x_0 is sufficiently near x_* the method has order of convergence at least $m+1$. See [33, 38, 43, 46]. In particular, Notice that in [38] a modification of $\mathcal{D}F(x_n)$ at each substep. In [39, 40, 41], Parida and Gupta provided some recurrence relations to establish a convergence analysis for a third order Newton-type methods under Lipschitz or Hölder conditions on the second Fréchet derivative. A modification of the approach used in [39] and some applications are presented by Chun et al. in [19]. Recently, Argyros and Ren [17] expanded the applicability of Halley's method using a center-Lipschitz condition on the second Fréchet derivative instead of Lipschitz's condition.

On the other hand, in the last years, attention has been paid in studying Newton's method on manifolds, since there are many numerical problems posed on manifolds that arise naturally in many contexts. Some examples include eigenvalue problems, minimization problems with orthogonality constraints, optimization problems with equality constraints, invariant subspace computations. See for instance [1, 2, 3, 7, 15, 20, 21, 27, 29, 35, 36, 48, 49]. For these problems, one has to compute solutions of equations or to find zeros of a vector field on Riemannian manifolds.

The study about convergence matter of iterative methods is usually centered on two types: semilocal and local convergence analysis. The semilocal convergence matter is, based on the information around an initial point, to give criteria ensuring the convergence of iterative methods; while the local one is, based on the information around a solution, to find estimates of the radii of convergence balls. There is a plethora of studies on the weakness and/or extension of the hypothesis made on the underlying operators; see for example [4, 5, 6, 12, 14, 32, 34, 48, 49].

The semilocal convergence analysis of Newton's method is based on celebrated Kantorovich theorem. This theorem is a fundamental result in numerical analysis, e.g., for providing an iterative method for computing zeros of polynomials or of systems of nonlinear equations. Moreover, this theorem is a very usefull result in nonlinear functional analysis, e.g., for establishing that a nonlinear equation in an abstract space has a solution. Let us recall Kantorovich's theorem in a Banach space setting.

Theorem 6.1.1. *[32] Let E be a Banach space, $\Omega \subseteq E$ be an open convex set, $F : \Omega \longrightarrow \Omega$ be a continuous operator, such that, $F \in C^1$ and DF is Lipschitz on Ω*

$$\|DF(x) - DF(y)\| \leq l\, \|x-y\|\,,\ \textit{for all}\ x,y \in \Omega, l>0.$$

Suppose that for some $x_0 \in \Omega$, $DF(x_0)$ is invertible and that for some $a>0$ and $b \geq 0$:

$$\left\|DF(x_0)^{-1}\right\| \leq a,$$

$$\left\|DF(x_0)^{-1} F(x_0)\right\| \leq b,$$

$$h = a\,b\,l \leq \frac{1}{2} \tag{6.1.1}$$

and

$$B(x_0,t_*) \subseteq \Omega \ \textit{where}\ t_* = \frac{1}{a\,l}\left(1-\sqrt{1-2\,h}\right).$$

If

$$\begin{aligned} v_k &= -DF(x_k)^{-1} F(x_k), \\ x_{k+1} &= x_k + v_k. \end{aligned}$$

Then $\{x_k\}_{k\in\mathbb{N}} \subseteq B(x_0,t_*)$ and $x_k \longrightarrow p_*$, which is the unique zero of F in $B[x_0,t_*]$. Furthermore , if $h<\frac{1}{2}$ and $B(x_0,r) \subseteq \Omega$ with

$$t_* < r \leq t_{**} = \frac{1}{a\,l}\left(1+\sqrt{1-2\,h}\right),$$

then p_* is also the unique zero of F in $B(x_0,r)$. Also, the error bound is:

$$\|x_k - x_*\| \leq (2\,h)^{2^k} \frac{b}{h}; \quad k=1,2,\ldots$$

Although the concepts will be defined later on, to extend the method on Riemannian manifolds, preliminarily we will say that the derivative of F at x_n is replaced by the covariant derivative of X at p_n :

$$\begin{aligned} \nabla_{(.)}X(p_n): \quad & T_{p_n}M \longrightarrow T_{p_n}M \\ & v \longrightarrow \nabla_Y X, \end{aligned}$$

where Y is a vector field satisfying $Y(p) = v$. We adopt the notation $\mathcal{D}X(p)v = \nabla_y X(p)$; hence $\mathcal{D}X(p)$ is a linear mapping of T_pM into T_pM. So, in this new context

$$-F'(x_n)^{-1} F(x_n)$$

is written as

$$-\mathcal{D}X(p_n)^{-1} X(p_n)$$

or

$$-\left(\nabla_{X(p_n)}X\right)^{-1}(p_n).$$

Now we can write Kantorovich's theorem in the new context. A proof of this theorem can be found in [27]. We will say that a singularity of a vector field X, is a point $p \in M$ for which $X(p) = 0$.

Theorem 6.1.2. *[27] (Kantorovich's theorem on Riemannian manifold) Let M be a Riemannian manifold, $\Omega \subseteq M$ be an open convex set, $X \in \chi(M)$ and $\mathcal{D}X \in Lip_l(\Omega)$. Suppose that for some $p_0 \in \Omega$, $\mathcal{D}X(p_0)$ is invertible and that for some $a > 0$ and $b \geq 0$:*

$$\left\| \mathcal{D}X(p_0)^{-1} \right\| \leq a \qquad \left(\left\| \left(\nabla_{(.)}X(p_o)\right)^{-1} \right\| \leq a \right)$$

$$\left\| \mathcal{D}X(p_0)^{-1} X(p_0) \right\| \leq b \qquad \left(\left\| \left(\nabla_{X(p_o)}X(p_o)\right)^{-1} \right\| \leq b \right)$$

$$h = a\,b\,l \leq \frac{1}{2} \tag{6.1.2}$$

$$B(p_0,t_*) \subseteq \Omega \text{ where } t_* = \frac{1}{a\,l}\left(1 - \sqrt{1-2h}\right).$$

If

$$\begin{aligned} v_k &= -\mathcal{D}X(p_k)^{-1} X(p_k), \\ p_{k+1} &= \exp_{p_k}(v_k), \end{aligned}$$

then $\{p_k\}_{k\in\mathbb{N}} \subseteq B(p_0,t_*)$ and $p_k \longrightarrow p_*$ which is the unique singularity of X in $B[p_0,t_*]$, where $\exp_{p_k}$ is defined in (6.2.6). Furthermore , if $h < \frac{1}{2}$ and $B(p_0,r) \subseteq \Omega$ with

$$t_* < r \leq t_{**} = \frac{1}{a\,l}\left(1 + \sqrt{1-2h}\right),$$

then p_* is also the unique singularity of F in $B(p_0,r)$. The error bound is:

$$d(p_k,p_*) \leq \tfrac{b}{h}(2h)^{2^k}; \quad k = 1,2,\ldots \tag{6.1.3}$$

The Kantorovich hypothesis (6.1.2) in Theorem 6.1.2 is only a sufficient convergence criterion for Newton method as well as the modified Newton's method. There are numerical examples in the literature showing that Newton's method converges but the Kantorovich hypothesis is not satisfied (see [6, 9, 12, 14, 32] and the references therein). In the present chapter we show how to expand the convergence domain of Newton's method without additional hypotheses. We achieve this goal by introducing more precise majorizing sequences for Newton's method than in the earlier studies in the field. Notice that if DF is Lipschitz on Ω, there exists a constant $l_0 > 0$, such that DF is center-Lipschitz on Ω

$$\|DF(x) - DF(x_0)\| \leq l_0 \|x - x_0\|, \text{ for all } x \in \Omega, l_0 > 0.$$

Clearly,

$$l_0 \leq l \tag{6.1.4}$$

holds in general and l/l_0 can be arbitrarily large [6, 12, 14]. In particular, we show that in the case of the modified Newton's method, condition (3) of Theorem 6.1.2 can be replaced by

$$h_0 = a b l_0 \leq \frac{1}{2}, \tag{6.1.5}$$

whereas in the case of Newton's method, condition (3) of Theorem 6.1.2 can be replaced by

$$h_1 = \frac{ab}{8}(l + 4l_0 + \sqrt{l^2 + 8l_0 l}) \leq \frac{1}{2}, \tag{6.1.6}$$

or by

$$h_2 = \frac{ab}{8}(4l_0 + \sqrt{l l_0 + 8l_0^2} + \sqrt{l_0 l}) \leq \frac{1}{2}. \tag{6.1.7}$$

Notice that

$$h \leq \frac{1}{2} \Longrightarrow h_1 \leq \frac{1}{2} \Longrightarrow h_2 \leq \frac{1}{2} \Longrightarrow h_0 \leq \frac{1}{2} \tag{6.1.8}$$

but not necessarily vice versa unless if $l_0 = l$. Moreover, we have that

$$\frac{h_1}{h} \longrightarrow \frac{1}{4}, \quad \frac{h_2}{h_1} \longrightarrow 0, \quad \frac{h_2}{h} \longrightarrow 0 \quad \text{and} \quad \frac{h_0}{h} \longrightarrow 0 \quad \text{as} \quad \frac{l_0}{l} \longrightarrow 0. \tag{6.1.9}$$

The preceding estimates show by how many times (at most) the applicability of the modified Newton's method or Newton's method can be extended. Moreover, we show that under the new convergence conditions, the error estimates on the distances $d(p_n, p_{n-1})$, $d(p_n, p_*)$ can be tighter and the information on the location of the solution at least as precise as in Theorem 6.1.2.

The chapter in organized as follows: Section 6.2. contains some definitions and fundamental properties of Riemannian manifolds. The convergence of simplified Newton's method and the order of convergence using normal coordinates are given in Sections 6.3. and 6.4.. Family of high order Newton-type methods, precise majorizing sequences and the corresponding convergence results are provided in Sections 6.5. and 6.6..

6.2. Basic Definitions and Preliminary Results

In this section, we introduce some definitions and fundamental properties of Riemannian manifolds in order to make this chapter as selfcontained as possible. These definitions and properties can be found in [20, 21, 22, 28, 32, 35, 37, 38, 47]. The preceding references are recommended to the interested reader for further study.

Definition 6.2.1. *[28, 37, 38, 47] A differentiable manifold of dimension m is a set M and a family of injective mappings $x_\alpha : U_\alpha \subset \mathbb{R}^m \longrightarrow M$ of open sets U_α of $\mathbb{R}^m$ into M such that:*

(i) $\bigcup\limits_{\alpha} x_\alpha(U_\alpha) = M$.

(ii) for any pair α, β with $x_\alpha(U_\alpha) \cap x_\beta(U_\beta) = W \neq \emptyset$, the sets $x_\alpha^{-1}(W)$ and $x_\beta^{-1}(W)$ are open sets in $\mathbb{R}^m$ and the mappings $x_\beta^{-1} \circ x_\alpha$ are differentiable.

(iii) The family $\{(U_\alpha, x_\alpha)\}$ is maximal relative to the conditions (i) and (ii).

The pair (U_α, x_α) (or the mapping x_α) with $p \in x_\alpha(U_\alpha)$ is called a parametrization (or system of coordinates) of M at p; $x_\alpha(U_\alpha)$ is then called neighborhood at p and $\left(x_\alpha(U_\alpha), x_\alpha^{-1}\right)$ is called a coordinate chart. A family $\{(U_\alpha, x_\alpha)\}$ satisfying (i) and (ii) is called a differentiable structure on M.

Let M be a real manifold, $p \in M$ and denote by T_pM the *tangent space* at p to M. Let $x : U \subset \mathbb{R}^m \longrightarrow M$ be a system of coordinates around p whit $x(x_1, x_2, \cdots, x_m) = p$ and its associated basis

$$\left\{ \left.\frac{\partial}{\partial x_1}\right|_p, \left.\frac{\partial}{\partial x_2}\right|_p, \cdots, \left.\frac{\partial}{\partial x_m}\right|_p \right\}$$

in T_pM. The *tangent bundle* TM is defined as

$$TM = \{(p, v);\ p \in M \text{ and } v \in T_pM\} = \bigcup_{p \in M} T_pM$$

and provides a differentiable structure of dimension $2m$ [22]. Next, we define the concept of Riemannian metric:

Definition 6.2.2. *A Riemannian metric on a differentiable manifold M is a correspondence which associates to each point p of M an inner product $\langle ., . \rangle_p$ (that is, a symmetric, bilinear, positive-definite form) on the tangent space T_pM, which varies differentiabily in the following sense: $x : U \subset \mathbb{R}^m \longrightarrow M$ is a system of coordinates around p with $x(x_1, x_2, \ldots, x_m) = p$, then*

$$\begin{aligned} g_{ij}(x_1, x_2, \cdots, x_m) &:= \left\langle \left.\frac{\partial}{\partial x_i}\right|_p, \left.\frac{\partial}{\partial x_j}\right|_p \right\rangle_p \\ &= \left\langle dx^{-1}\left(\left.\frac{\partial}{\partial x_i}\right|_p \right), dx^{-1}\left(\left.\frac{\partial}{\partial x_j}\right|_p \right) \right\rangle, \end{aligned}$$

in which dx^{-1} is the tangent map of x^{-1} and is a differentiable operator on U for each $i, j = 1, 2, .., n$. The operators g_{ij} are called the local representatives of the Riemannian metric.

The inner product $\langle .,. \rangle_p$ induces in a natural way the norm $||.||_p$. The subscript p is usually deleted whenever there is not possibility of confusion.

If p and q are two elements of the manifold M and $c : [0,1] \longrightarrow M$ is a piecewise smooth curve connecting p and q, then the *arc length of* c is defined by

$$l(c) = \int_0^1 \left|\left|c'(t)\right|\right| dt \tag{6.2.1}$$
$$= \int_0^1 \left\langle \frac{dc}{dt}, \frac{dc}{dt} \right\rangle^{1/2} dt,$$

and the *Riemannian distance from p to q* by

$$d(p,q) = \inf_c l(c). \tag{6.2.2}$$

Definition 6.2.3. *Let $\chi(M)$ be the set of all vector fields of class C^∞ on M and $\mathcal{D}(M)$ the ring of real-valued operators of class C^∞ defined on M, that is:*

$$\chi(M) = C^\infty\left(M, T_{(.)}M\right),$$
$$\mathcal{D}(M) = C^\infty(M, \mathbb{R}).$$

An affine connection ∇ on M is a mapping

$$\begin{aligned} \nabla : \quad \chi(M) \times \chi(M) &\longrightarrow \chi(M) \\ (X,Y) &\longmapsto \nabla_X Y \end{aligned} \tag{6.2.3}$$

which satisfies the following properties:

$$\begin{array}{ll} i) & \nabla_{fX+gY} Z = f\nabla_X Z + g\nabla_Y Z. \\ ii) & \nabla_X (Y+Z) = \nabla_X Y + \nabla_X Z. \\ iii) & \nabla_X (fY) = f\nabla_X Y + X(f) Y, \end{array}$$

where $X, Y, Z \in \chi(M)$ and $f, g \in \mathcal{D}(M)$.

Definition 6.2.4. *Let X be a C^1 vector field on M, the covariant derivative of X determined by the Levi-Civita connection ∇ defines on each $p \in M$ a linear application of T_pM itself*

$$\begin{array}{llll} \mathcal{D}X(p) : & T_pM & \longrightarrow & T_pM \\ & v & \longmapsto & \mathcal{D}X(p)(v) = \nabla_Y X(p) \end{array} \tag{6.2.4}$$

where Y is a vector field satisfying $Y(p) = v$. The value $\mathcal{D}X(p)(v)$ depends only on the tangent vector $v = Y(p)$ since ∇ is linear in Y, thus we can write

$$\mathcal{D}X(p)(v) = \nabla_v X(p).$$

Let us consider a curve $c : [a,b] \longrightarrow M$ and a vector field X along c, i.e. $X(p) \in T_{c'(t)}M$ where $c(t) = p$ for all t. We say that a vector field X is parallel along c (with respect to ∇) if $\mathcal{D}X(p)(c'(t)) = 0$, the affine connection is *compatible* with the metric $\langle .,. \rangle$, when for any smooth curve c and any pair of parallel vector fields P and P' along c, we have that $\langle P, P'\rangle$ is constant or equivalently

$$\frac{d}{dt}\langle X,Y\rangle = \left\langle \nabla_{c'(t)}X, Y\right\rangle + \left\langle X, \nabla_{c'(t)}Y\right\rangle,$$

where X and Y are vector fields along the differentiable curve $c : I \longrightarrow M$ (see [22], [45]). We say that ∇ is *symmetric* if

$$\nabla_X Y - \nabla_Y X = [X,Y] \text{ for all } X,Y \in \chi(M).$$

The theorem of Levi-Civita (see [45]), establishes that there exists a unique symmetric affine connection ∇ on M compatible with the metric. This connection is called *connection of Levi-Civita.*

Definition 6.2.5. *[28, 37, 38, 47] A parameterized curve $\gamma : \mathfrak{I} \longrightarrow M$ is a geodesic at $t_0 \in \mathfrak{I}$ if $\nabla_{\gamma(t)}\gamma'(t) = 0$ in the point t_0. If γ is a geodesic at t, for all $t \in \mathfrak{I}$, we say that γ is a geodesic. If $[a,b] \subseteq \mathfrak{I}$, the restriction of γ to $[a,b]$ is called a geodesic segment joining $\gamma(a)$ to $\gamma(b)$.*

Some times, by abuse of the language, we refer to the image $\gamma(\mathfrak{I})$, of a geodesic γ, as a geodesic. A basic property of a geodesic is that, $\gamma'(t)$ is parallel along of $\gamma(t)$; this implies that $||\gamma'(t)||$ is constant.

Let $B(p,r)$ and $B[p,r]$ be respectively the *open geodesic and the closed geodesic ball* with center p and radius r, that is:

$$B(p,r) = \{q \in M : d(p,q) < r\}$$
$$B[p,r] = \{q \in M : d(p,q) \leq r\}.$$

We define an open set U of M to be *convex* if given $p, q \in U$ there exists a unique geodesic in U joining p to q, and such that the length of the geodesic is $d(p,q)$.

The Hopf and Rinow theorem (see [22]) gives necessary and sufficient conditions for M to be a complete metric space. In particular, if M is a complete metric space, then for any $q \in M$ there exists a geodesic γ, called *minimizing geodesic,* joining p to q with

$$l(\gamma) = d(p,q), \tag{6.2.5}$$

also if $v \in T_pM$, there exists a unique minimizing geodesic γ such that $\gamma(0) = p$ and $\gamma'(0) = v$. The point $\gamma(1)$ is called the image of v by the *exponential map at p*, that is, there exist a well defined map

$$\exp_p : \quad T_pM \longrightarrow M \tag{6.2.6}$$

such that

$$\exp_p(v) = \gamma(1),$$

and for any $t \in [0,1]$

$$\gamma(t) = \exp_p(tv).$$

It can be shown that $\exp_p$ defines a diffeomorphism of a neighborhood $\widehat{U}$ of the origin $0_p \in T_pM$ onto a neighborhood U of $p \in M$, called *normal neighborhood* of p, (see [22]).

Let $p \in M$ and U a normal neighborhood of p. Let us consider an orthonormal basis $\{e_i\}_{i=1}^m$ of T_pM. This basis gives the isomorphism $f : \mathbb{R}^m \longrightarrow T_pM$ defined by $f(u_1, \cdots, u_n) = \sum_{i=1}^m u_i e_i$. If $q = \exp_p\left(\sum_{i=1}^m u_i e_i\right)$, we say that $(u_1, \ldots, u_n)$ are *normal coordinates* of q in the normal neighborhood U of p and the coordinate chart is the composition

$$\varphi := \exp_p \circ f : \mathbb{R}^m \longrightarrow U.$$

One of the most important properties of the normal coordinates is that the geodesies passing through p are given by linear equations, (see [44]).

The exponential map has many important properties [22], [44]. When the exponential map is defined for each value of the parameter $t \in \mathbb{R}$ we will say that the Riemannian manifold M is geodesically complete (or simply complete). The Hopf and Rinow theorem (see [22]), also establishes that the property of the Riemannian manifold of being geodesically complete is equivalent to being complete as a metric space.

Definition 6.2.6. *[28, 37, 38, 47] Let c be a piecewise smooth curve. For any pair $a, b \in \mathbb{R}$, we define the parallel transport along c which is denoted by P_c as*

$$\begin{array}{rccc} P_{c,a,b}: & T_{c(a)}M & \longrightarrow & T_{c(b)}M \\ & v & \longmapsto & V(c(b)), \end{array} \tag{6.2.7}$$

where V is the unique vector field along c such that $\nabla_{c'(t)}V = 0$ and $V(c(a)) = v$.

It is easy to show that $P_{c,a,b}$ is linear and one-one, thus $P_{c,b,a}$ is an isomorphism between every two tangent spaces $T_{c(a)}M$ and $T_{c(b)}M$. Its inverse is the parallel translation along the reversed portion of c from $V(c(b))$ to $V(c(a))$, actually $P_{c,a,b}$ is a isometry between $T_{c(a)}M$ and $T_{c(b)}M$. Moreover, for a positive integer i and for all $(v_1, v_2, \ldots, v_i) \in \left(T_{c(a)}M\right)^i$, we define P_c^i as

$$P^i_{c,a,b}: \quad \left(T_{c(a)}M\right)^i \quad \longrightarrow \quad \left(T_{c(b)}M\right)^i,$$

where

$$P^i_{c,a,b}(v_1, v_2, \ldots, v_i) = \left(P_{c,a,b}(v_1), P_{c,a,b}(v_2), \ldots, P_{c,a,b}(v_i)\right).$$

The parallel transport has the important properties:

$$\begin{array}{l} P_{c,a,b} \circ P_{c,b,d} = P_{c,a,d}, \\ P^{-1}_{c,b,a} = P_{c,a,b}. \end{array} \tag{6.2.8}$$

Next, we generalize the concept of covariant derivative. We observe that

$$\begin{array}{rccc} \mathcal{D}X: & C^k(TM) & \longrightarrow & C^{k-1}(TM) \\ & (v,.) & \longmapsto & \mathcal{D}X(Y) = \nabla_Y X, \end{array} \tag{6.2.9}$$

where TM is the tangent bundle. Similar to the higher order Fréchet derivative, see [18]. We define the *higher order covariant derivatives*, see [45], as the multilinear map or j-tensor:

$$\mathcal{D}^j X: \quad \left(C^k(TM)\right)^j \quad \longrightarrow \quad C^{k-j}(TM)$$

given by

$$\mathcal{D}^j X(Y_1, Y_2, \ldots, Y_{j-1}, Y) = \nabla_Y \mathcal{D}^{j-1}(X(Y_1, Y_2, \ldots, Y_{j-1})) - \sum_{i=1}^{j-1} \mathcal{D}^{j-1} X(Y_1, Y_2, \ldots, \nabla_Y Y_i, \ldots, Y_{j-1},) \tag{6.2.10}$$

for each $Y_1, Y_2, \ldots, Y_{j-1} \in C^k(TM)$. In the case of $j = 2$ we have

$$\mathcal{D}^2 X : \quad C^k(TM) \times C^k(TM) \longrightarrow C^{k-2}(TM)$$

and

$$\begin{aligned} \mathcal{D}^2 X(Y_1, Y) &= \nabla_Y \mathcal{D}X(Y_1) - \mathcal{D}X(\nabla_Y Y_1) \\ &= \nabla_Y(\nabla_{Y_1} X) - \nabla_{\nabla_Y Y_1} X. \end{aligned} \tag{6.2.11}$$

The multilinearity refers to the structure of $C^k(M)$-module, such that, the value of

$$\mathcal{D}^j X(Y_1, Y_2, \ldots, Y_{j-1}, Y)$$

at $p \in M$ only depends on the j-tuple of tangent vectors

$$(v_1, v_2, \ldots, v_j) = (Y_1(p), Y_2(p), \ldots, Y_{j-1}(p), Y(p)) \in (T_p M)^j.$$

Therefore, for any $p \in M$, we can define the map

$$\mathcal{D}^j X(p) : \quad (T_p M)^j \longrightarrow T_p M$$

by

$$\mathcal{D}^j X(p)(v_1, v_2, \ldots, v_j) = \mathcal{D}^j X(Y_1, Y_2, \ldots, Y_{j-1}, Y)(p). \tag{6.2.12}$$

Definition 6.2.7. *[28, 37, 38, 47] Let M be a Riemannian manifold, $\Omega \subseteq M$ an open convex set and $X \in \chi(M)$. The covariant derivative $\mathcal{D}X = \nabla_{(.)}X$ is Lipschitz with constant $l > 0$, if for any geodesic γ and $a, b \in \mathbb{R}$ so that $\gamma[a,b] \subseteq \Omega$, it holds that:*

$$\left\| P_{\gamma,b,a} \mathcal{D}X(\gamma(b)) P_{\gamma,a,b} - \mathcal{D}X(\gamma(a)) \right\| \leq l \int_a^b \left\| \gamma'(t) \right\| dt. \tag{6.2.13}$$

We will write $\mathcal{D}X \in Lip_l(\Omega)$.

Note that $P_{\gamma,b,a}\mathcal{D}X(\gamma(b)) P_{\gamma,b,a}$ and $\mathcal{D}X(\gamma(a))$ are both operators defined in the same tangent plane $T_{\gamma(a)}M$. If M is an Euclidean space, the above definition coincides with the usual Lipschitz definition for the operator $DF : M \longrightarrow M$.

Proposition 6.2.8. *[28, 37, 38, 47] Let c be a curve in M and X be a C^1 vector field on M, then the covariant derivative of X in the direction of $c'(s)$ is*

$$\begin{aligned} \mathcal{D}X(c(s))\, c'(s) &= \nabla_{c'(s)} X_{c(s)} \\ &= \lim_{h \to 0} \frac{1}{h} \left(P_{c,s+h,s} X(c(s+h)) - X(c(s)) \right). \end{aligned} \tag{6.2.14}$$

Note that if $M = \mathbb{R}^n$ the previous proposition agrees with the definition of classic directional derivative in $\mathbb{R}^n$; (see [45]).

It is also possible to obtain a version of the fundamental theorem of calculus for manifolds:

Theorem 6.2.9. *[27] Let c be a geodesic in M and X be a C^1 vector field on M, then*

$$P_{c,t,0}X\left(c\left(t\right)\right) = X\left(c\left(0\right)\right) + \int_0^t P_{c,s,0}\left(\mathcal{D}X\left(c\left(s\right)\right)c'\left(s\right)\right)ds. \tag{6.2.15}$$

Theorem 6.2.10. *Let c be a geodesic in M and X be a C^2 vector field on M, then*

$$P_{c,t,0}\mathcal{D}X\left(c\left(t\right)\right)c'\left(t\right) = \mathcal{D}X\left(c\left(0\right)\right)c'\left(0\right) + \int_0^t P_{c,s,0}\left(\mathcal{D}^2X\left(c\left(s\right)\right)\left(c'\left(s\right),c'\left(s\right)\right)\right)ds. \tag{6.2.16}$$

Proof. Let us consider the vector field along of the geodesic $c\left(s\right)$

$$Y\left(c\left(s\right)\right) = \mathcal{D}X\left(c\left(s\right)\right)c'\left(s\right).$$

By the previous theorem

$$P_{c,t,0}Y\left(c\left(t\right)\right) = Y\left(c\left(0\right)\right) + \int_0^t P_{c,s,0}\left(\mathcal{D}Y\left(c\left(s\right)\right)c'\left(s\right)\right)ds$$

hence

$$P_{c,t,0}\mathcal{D}X\left(c\left(t\right)\right)c'\left(t\right) = \mathcal{D}X\left(c\left(0\right)\right)c'\left(0\right) + \int_0^t P_{c,s,0}\left(\mathcal{D}\left(\mathcal{D}X\left(c\left(s\right)\right)c'\left(s\right)\right)c'\left(s\right)\right)ds$$

by (6.2.11)

$$\begin{aligned}\mathcal{D}^2X\left(c\left(s\right)\right)\left(c'\left(s\right),c'\left(s\right)\right) &= \nabla_{c'(s)}\mathcal{D}\left(X\left(c\left(s\right)\right)\left(c'\left(s\right)\right)\right) - \mathcal{D}X\left(c\left(s\right)\right)\left(\nabla_{c'(s)}c'\left(s\right)\right)\\ &\quad \mathcal{D}\left(\mathcal{D}X\left(c\left(s\right)\right)c'\left(s\right)\right)c'\left(s\right) - \mathcal{D}X\left(c\left(s\right)\right)\left(\nabla_{c'(s)}c'\left(s\right)\right),\end{aligned}$$

since $c\left(s\right)$ is a geodesic, we have $\nabla_{c'(s)}c'\left(s\right) = 0$, hence

$$\mathcal{D}^2X\left(c\left(s\right)\right)\left(c'\left(s\right),c'\left(s\right)\right) = \mathcal{D}\left(\mathcal{D}X\left(c\left(s\right)\right)c'\left(s\right)\right)c'\left(s\right).$$

Therefore

$$P_{c,t,0}\mathcal{D}X\left(c\left(t\right)\right)c'\left(t\right) = \mathcal{D}X\left(c\left(0\right)\right)c'\left(0\right) + \int_0^t \mathcal{D}^2X\left(c\left(s\right)\right)\left(c'\left(s\right),c'\left(s\right)\right)ds.$$

□

In a similar way, using an induction strategy, we can prove that

$$P_{c,t,0}\mathcal{D}^nX\left(c\left(s\right)\right)P^n_{c,0,t} - \mathcal{D}^nX\left(c\left(0\right)\right) = \int_0^s P_{c,t,0}\left(\mathcal{D}^nX\left(c\left(t\right)\right)P^n_{c,0,t}\left(c'\left(0\right),\ldots,c'\left(0\right)\right)\right)dt \tag{6.2.17}$$

Theorem 6.2.11. *Let c be a geodesic in M, $[0,1] \subseteq Dom(c)$ and X be a C^2 vector field on M, then*

$$P_{c,1,0}X(c(1)) = X(c(0)) + \mathcal{D}X(c(0))c'(0) + \int_0^1 (1-t)P_{c,s,0}\mathcal{D}^2X(c(t))\left(c'(t),c'(t)\right)dt. \tag{6.2.18}$$

Proof. Consider the curve

$$f(s) = P_{c,s,0}X(c(s))$$

in $T_{c(0)}M$. We have that

$$f^{(n)}(s) = \quad P_{c,s,0}\mathcal{D}^{(n)}X(c(s))\underbrace{\left(c'(s),c'(s),\cdots,c'(s)\right)}_{n-times} \tag{6.2.19}$$

Then

$$f''(s) = P_{c,s,0}\mathcal{D}^2X(c(s))\left(c'(s),c'(s)\right),$$

and from Taylor's theorem

$$f(1) = f(0) + f'(0)(1-0) + \int_0^1 (1-t)f''(t)dt.$$

Therefore

$$P_{c,1,0}X(c(1)) = X(c(0)) + \mathcal{D}X(c(0))c'(0) + \int_0^1 (1-t)P_{c,t,0}\mathcal{D}^2X(c(t))\left(c'(t),c'(t)\right)dt.$$

□

Let us recall that if $A : T_pM \longrightarrow T_pM$, we can define $\|A\| = \sup\{\|Av\| : v \in T_pM,\ \|v\| = 1\}$.

The following is an important lemma, that allows to know when an operator is invertible and also allows to give on estimate for its inverse.

Lemma 6.2.12. *(Banach's Lemma [32]) Let A be an invertible bounded linear operator in a Banach space E and B be an bounded linear operator B in E, if*

$$\left\|A^{-1}B - I\right\| < 1$$

then B^{-1} exists and

$$\begin{aligned}\left\|B^{-1}\right\| &\le \frac{\left\|A^{-1}\right\|}{1-\|A^{-1}B-I\|}\\ &\le \frac{\left\|A^{-1}\right\|}{1-\|A^{-1}\|\,\|B-A\|}.\end{aligned}$$

Moreover,

$$\begin{aligned}\left\|B^{-1}A\right\| &\le \frac{1}{1-\|A^{-1}B-I\|}\\ &\le \frac{1}{1-\|A^{-1}\|\,\|B-A\|}.\end{aligned}$$

6.3. Simplified Newton's Method on Riemannian Manifolds ($m = \infty$)

Next we will prove the semilocal convergence of the simplified Newton's method in Riemannian manifolds (fixing $\mathcal{D}X(p_0)^{-1}$ in each iteration). Our main result is:

Theorem 6.3.1. *Let M be a Riemannian manifold, $\Omega \subseteq M$ be an open convex set, $X \in \chi(M)$, and $\mathcal{D}X \in Lip_l(\Omega)$. Suppose that for some $p_0 \in \Omega$, $\mathcal{D}X(p_0)$ is invertible and that for some $a > 0$ and $b \geq 0$:*

(1) $\left\| \mathcal{D}X(p_0)^{-1} \right\| \leq a$,
(2) $\left\| \mathcal{D}X(p_0)^{-1} X(p_0) \right\| \leq b$,
(3) $h = abl \leq \frac{1}{2}$,
(4) $B(p_0, t_*) \subseteq \Omega$ where $t_* = \frac{1}{al}\left(1 - \sqrt{1-2h}\right)$.

If

$$\begin{aligned} v_k &= -P_{\sigma_k,0,1}\mathcal{D}X(p_0)^{-1} P_{\sigma_k,1,0} X(p_k), \\ p_{k+1} &= \exp_{p_k}(v_k), \end{aligned} \tag{6.3.1}$$

where $\{\sigma_k : [0,1] \longrightarrow M\}_{k\in\mathbb{N}}$ is the minimizing geodesic family connecting p_0, p_k, then $\{p_k\}_{k\in\mathbb{N}} \subseteq B(p_0, t_*)$ and $p_k \longrightarrow p_*$ which is the only one singularity of X in $B[p_0, t_*]$. Furthermore, if $h < \frac{1}{2}$ and $B(p_0, r) \subseteq \Omega$ with

$$t_* < r \leq t_{**} = \frac{1}{al}\left(1 + \sqrt{1-2h}\right),$$

then p_* is also the only singularity of F in $B(p_0, r)$. The error bound is:

$$d(p_k, p_*) \leq \frac{b}{h}\left(1 - \sqrt{1-2h}\right)^{k+1}; \quad k = 1, 2, \ldots \tag{6.3.2}$$

First, we establish some results that are of primary relevance in this proof.

Lemma 6.3.2. *Let M be a Riemannian manifold, $\Omega \subseteq M$ an open convex set, $X \in \chi(M)$ and $\mathcal{D}X \in Lip_l(\Omega)$. Take $p \in B(p_0, r) \subseteq \Omega$, $v \in T_pM$, $\sigma : [0,1] \longrightarrow M$ be a minimizing geodesic connecting p_0, p and*

$$\gamma(t) = \exp_p(tv).$$

Then

$$P_{\gamma,t,0} X(\gamma(t)) = X(p) + P_{\sigma,0,1} t\mathcal{D}X(p_0) P_{\sigma,1,0} v + R(t)$$

with

$$\|R(t)\| \leq l\left(\frac{t}{2}\|v\| + d(p_0, p)\right) t\|v\|.$$

Proof. From Theorem 6.2.9, it follows that

$$P_{\gamma,t,0} X(\gamma(t)) - X(\gamma(0)) = \int_0^t P_{\gamma,s,0}\left(\mathcal{D}X(\gamma(s))\gamma'(s)\right) ds,$$

since γ is a minimizing geodesic, then $\gamma'(t)$ is parallel and $\gamma'(s) = P_{\gamma,0,s}\gamma'(0)$. Moreover $\gamma'(0) = v$ then

$$P_{\gamma,t,0}X(\gamma(t)) - X(p) = \int_0^t P_{\gamma,s,0}\left(\mathcal{D}X(\gamma(s))P_{\gamma,0,s}v\right)ds.$$

Thus

$$\begin{aligned}
&P_{\gamma,t,0}X(\gamma(t)) - X(p) - P_{\sigma,0,1}t\mathcal{D}X(p_0)P_{\sigma,1,0}v \\
&= \textstyle\int_0^t P_{\gamma,s,0}\left(\mathcal{D}X(\gamma(s))P_{\gamma,0,s}v\right)ds - P_{\sigma,0,1}\mathcal{D}X(p_0)P_{\sigma,1,0}v \\
&= \textstyle\int_0^t \left(P_{\gamma,s,0}\mathcal{D}X(\gamma(s))P_{\gamma,0,s}v - P_{\sigma,0,1}\mathcal{D}X(p_0)P_{\sigma,1,0}v\right)ds,
\end{aligned}$$

letting

$$R(t) = \int_0^t \left(P_{\gamma,s,o}\mathcal{D}X(\gamma(s))P_{\gamma,o,s}v - P_{\sigma,0,1}\mathcal{D}X(p_0)P_{\sigma,1,0}v\right)ds,$$

and since $\mathcal{D}X \in Lip_l(\Omega)$, we obtain

$$\begin{aligned}
\|R(t)\| &\le \int_0^t \left\|\left(P_{\gamma,s,o}\mathcal{D}X(\gamma(s))P_{\gamma,o,s} - \mathcal{D}X(p) + \mathcal{D}X(p) - P_{\sigma,0,1}\mathcal{D}X(p_0)P_{\sigma,1,0}\right)\right\| \|v\|\,ds \\
&\le \int_0^t \left(\left\|P_{\gamma,s,o}\mathcal{D}X(\gamma(s))P_{\gamma,o,s} - \mathcal{D}X(p)\right\| + \left\|\mathcal{D}X(p) - P_{\sigma,0,1}\mathcal{D}X(p_0)P_{\sigma,1,0}\right\|\right)\|v\|\,ds \\
&= \int_0^1 \left(\left\|P_{\gamma,s,o}\mathcal{D}X(\gamma(s))P_{\gamma,o,s} - \mathcal{D}X(\gamma(0))\right\| + \left\|\mathcal{D}X(\sigma(1)) - P_{\sigma,0,1}\mathcal{D}X(\sigma(0))P_{\sigma,1,0}\right\|\right)\|v\|\,ds \\
&\le l\int_0^t \left(\int_0^s \left\|\gamma'(\tau)\right\|d\tau + d(p_0,p)\right)\|v\|\,ds \\
&= l\int_0^t \left(\int_0^s \left\|\gamma'(0)\right\|d\tau + d(p_0,p)\right)\|v\|\,ds \\
&= l\int_0^t \left(s\left\|\gamma'(0)\right\| + d(p_0,p)\right)\|v\|\,ds \\
&= l\left(\frac{t^2}{2}\|v\| + td(p_0,p)\right)\|v\|.
\end{aligned}$$

Therefore,

$$\|R(t)\| \le l\left(\frac{t}{2}\|v\| + d(p_0,p)\right)t\|v\|.$$

□

Corollary 6.3.3. *Let M be a Riemannian manifold, $\Omega \subseteq M$ be an open convex set, $X \in \chi(M)$, $\mathcal{D}X \in Lip_l(\Omega)$. Take $p \in \Omega$, $v \in T_pM$ and let be*

$$\gamma(t) = \exp_p(tv).$$

If $\gamma[0,t) \subseteq \Omega$ and $P_{\sigma,0,1}\mathcal{D}X(p_0)P_{\sigma,1,0}v = -X(p)$, then

$$\left\|P_{\gamma,1,0}X(\gamma(1))\right\| \le l\left(\frac{1}{2}\|v\| + d(p_0,p)\right)\|v\|. \tag{6.3.3}$$

Now we can prove the simplified Kantorovich theorem on Riemannian manifolds. The proof of this theorem will be divided in two parts. First, we will prove that simplified Kantorovich method is well defined, i.e. $\{p_k\}_{k\in\mathbb{N}} \subseteq B(p_0,t_*)$; we will also prove the convergence of the method. In the second part, we will establish uniqueness.

- **CONVERGENCE**

We consider the auxiliary real function $f:\mathbb{R}\longrightarrow\mathbb{R}$, defined by

$$f(t)=\frac{l}{2}t^2-\frac{1}{a}t+\frac{b}{a}. \tag{6.3.4}$$

Its discriminant is

$$\triangle=\frac{1}{a^2}(1-2lba),$$

which is positive, because $abl\le\frac{1}{2}$. Thus f has a least one real root (unique when $h=\frac{1}{2}$). If t_* is the smallest root, a direct calculation show that $f'(t)<0$ for $0\le t<t_*$, so f is strictly decreasing in $[0,t_*]$. Therefore the (scalar) Newton's method can be applied to f, in other words:

If $t_0\in[0,t_*)$, for $k=0,1,2,\ldots$, we define

$$t_{k+1}=t_k-\frac{f(t_k)}{f'(0)}.$$

Then $\{t_k\}_{k\in\mathbb{N}}$ is well defined, strictly increasing and converges to t_*. Furthermore, if $h=abl<\frac{1}{2}$, then

$$t_*-t_k\le\frac{b}{h}\left(1-\sqrt{1-2h}\right)^{k+1},\;\; k=1,2,\ldots,\text{ see [32].} \tag{6.3.5}$$

Let us take as starting point $t_0=0$. We want to show that Newton's iteration are well defined for any $q\in B(p_0,t_*)\subseteq\Omega$.

We define

$$\mathbf{K}(t)=\left\{q\in\overline{B(p_0,t)}:\left\|P_{\sigma,0,1}\mathcal{D}X(p_0)^{-1}P_{\sigma,1,0}X(q)\right\|\le\frac{f(t)}{|f'(0)|}=af(t),\;\;0\le t<t_*\right\}, \tag{6.3.6}$$

where $\sigma:[0,1]\longrightarrow M$ is the minimizing geodesic connecting p_0 and q. Note that $\mathbf{K}(t)\neq\emptyset$ since $p_0\in\mathbf{K}(t)$.

Proposition 6.3.4. *Under the hypotheses of either the Kantorovich or the simplified Kantorovich method, if $q\in B(p_0,t_*)$, then $\mathcal{D}X(q)$ is nonsingular and*

$$\left\|\mathcal{D}X(q)^{-1}\right\|\le\frac{1}{|f'(\lambda)|}\text{ where }\lambda=d(p_0,q)<t_*.$$

Proof. Let $\lambda=d(p_0,q)$ and $\alpha:[0,1]\longrightarrow M$ be a geodesic with $\alpha(0)=p_0$, $\alpha(1)=q$ and $\|\alpha'(0)\|=\lambda$. Define $\phi:T_qM\longrightarrow T_qM$ by letting

$$\phi=P_{\alpha,1,0}\mathcal{D}X(p_0)P_{\alpha,0,1}. \tag{6.3.7}$$

Since $P_{\alpha,1,0}$ and $P_{\alpha,0,1}$ are linear, isometric and $\mathcal{D}X(p_0)$ is nonsingular, we have that ϕ is linear, nonsingular and

$$\|\phi^{-1}\|=\left\|\mathcal{D}X(p_0)^{-1}\right\|\le a=\frac{1}{|f'(0)|},$$

with $\alpha([0,1]) \subseteq B(p_0,t_*)$. Since $d(p_0,q) < t_*$, $\mathcal{D}X \in Lip_L(\Omega)$ and $\|\alpha'(0)\| = \lambda$. Therefore

$$\|\mathcal{D}X(q) - \phi\| \le l\lambda. \tag{6.3.8}$$

By (6.3.7) and (6.3.8), we have

$$\begin{aligned} \left\|\phi^{-1}\right\| \|\mathcal{D}X(q) - \phi\| &\le al\lambda \\ &\le alt_* \\ &= al\frac{1}{al}\left(1 - \sqrt{1-2abl}\right) \\ &\le 1. \end{aligned}$$

Using Banach's lemma, we conclude that $\mathcal{D}X(q)$ is nonsingular, and

$$\begin{aligned} \left\|\mathcal{D}X(q)^{-1}\right\| &\le \frac{\left\|\phi^{-1}\right\|}{1 - \left\|\phi^{-1}\right\| \|\mathcal{D}X(q) - \phi\|} \\ &\le \frac{a}{1 - al\lambda} \\ &\le \frac{1}{|f'(\lambda)|}. \end{aligned}$$

□

Therefore, for any $q \in B(p_0,t_*)$, we can apply the Kantorovich methods.

Lemma 6.3.5. *Let $q \in \mathbf{K}(t)$, define*

$$\begin{aligned} t_+ &= t - \tfrac{f(t)}{|f'(0)|} \\ q_+ &= \exp_q\left(-P_{\sigma,0,1}\mathcal{D}X(p_0)^{-1}P_{\sigma,1,0}X(q)\right). \end{aligned}$$

Then $t < t_+ < t_*$ and $q_+ \in \mathbf{K}(t_+)$.

Proof. Consider the geodesic $\gamma\colon [0,1] \longrightarrow M$ defined by

$$\gamma(\theta) = \exp_q\left(-\theta P_{\sigma,0,1}\mathcal{D}X(p_0)^{-1}P_{\sigma,1,0}X(q)\right),$$

we have

$$\begin{aligned} d(p_0,\gamma(\theta)) &\le d(p_0,q) + d(q,\gamma(\theta)) \\ &\le t + \left\|\theta P_{\sigma,0,1}\mathcal{D}X(p_0)^{-1}P_{\sigma,1,0}X(q)\right\| \\ &\le t + \theta\frac{f(t)}{|f'(0)|}. \end{aligned}$$

Since

$$\gamma(1) = \exp_q\left(-P_{\sigma,0,1}\mathcal{D}X(p_0)^{-1}P_{\sigma,1,0}X(q)\right) = q_+,$$

this implies that

$$d(p_0, q_+) = d(p_0, \gamma(1)) \leq t + \frac{f(t)}{|f'(0)|} = t_+,$$

therefore

$$q_+ \in B(p_0, t_+) \subset B(p_0, t_*).$$

Moreover, if $\sigma_+ [0,1] \longrightarrow M$ is the minimizing geodesic connecting p_0 and q_+, then

$$\left\| -P_{\sigma_+,0,1} \mathcal{D}X(p_0)^{-1} P_{\sigma_+,1,0} X(q_+) \right\| \leq \left\| \mathcal{D}X(p_0)^{-1} \right\| \|X(q_+)\|.$$

Furthermore, if $v = -P_{\sigma,0,1} \mathcal{D}X(p_0)^{-1} P_{\sigma,1,0} X(q)$, then

$$\begin{aligned} P_{\sigma,0,1} \mathcal{D}X(p_0) P_{\sigma,1,0} v &= P_{\sigma,0,1} \mathcal{D}X(p_0) P_{\sigma,1,0} \left(-P_{\sigma,0,1} \mathcal{D}X(p_0)^{-1} P_{\sigma,1,0} X(q) \right) \\ &= -P_{\sigma,0,1} \mathcal{D}X(p_0) \mathcal{D}X(p_0)^{-1} P_{\sigma,1,0} X(q) \\ &= -X(q). \end{aligned}$$

By Theorem 6.2.10,

$$\begin{aligned} \|X(q_+)\| &= \|X(\gamma(1))\| \\ &\leq l \left(\frac{1}{2} \|v\| + d(p_0, p) \right) \|v\| \\ &\leq l \left(\frac{1}{2} \left\| -P_{\sigma_+,0,1} \mathcal{D}X(p_0)^{-1} P_{\sigma_+,1,0} X(q) \right\| + t \right) \left\| -P_{\sigma_+,0,1} \mathcal{D}X(p_0)^{-1} P_{\sigma_+,1,0} X(q) \right\| \\ &\leq l \left(\frac{1}{2} \left(\frac{f(t)}{|f'(0)|} \right) + t \right) \left(\frac{f(t)}{|f'(0)|} \right). \end{aligned}$$

Thus, by (6.3.6), after some calculations

$$\begin{aligned} \left\| -P_{\sigma_+,0,1} \mathcal{D}X(p_0)^{-1} P_{\sigma_+,1,0} \right\| \|X(q_+)\| &\leq \left(\frac{1}{|f'(0)|} \right) l \left(\frac{1}{2} \left(\frac{f(t)}{|f'(0)|} \right) + t \right) \left(\frac{f(t)}{|f'(0)|} \right) \\ &= \frac{1}{8} l \left(2b - 2t + alt^2 \right) \left(2b + 2t + alt^2 \right) \\ &= \frac{f(t_+)}{|f'(0)|}, \end{aligned}$$

we thus conclude

$$\|X(q_+)\| \leq \frac{f(t_+)}{|f'(0)|},$$

and therefore

$$q_+ \in \mathbf{K}(t_+).$$

□

Now we are going to prove that starting from any point of $\mathbf{K}(t)$ the simplified Newton method converges.

Corollary 6.3.6. *Take* $0 \le t < t_*$ *and* $q \in \mathbf{K}(t)$, *and define*

$$\begin{aligned} &\tau_0 = t \\ &\tau_{k+1} = \tau_k - \tfrac{f(\tau_k)}{|f'(0)|} \quad \text{for each} \quad k = 0, 1, \cdots \end{aligned}$$

Then the sequence generated by Newton's method starting with the point $q_0 = q$ *is well defined for any k and*

$$q_k \in \mathbf{K}(\tau_k). \tag{6.3.9}$$

Moreover $\{q_k\}_{k\in\mathbb{N}}$ *converges to some* $q_* \in B(p_0, t_*)$, $X(q_*) = 0$ *and*

$$d(q_k, q_*) \le t_* - \tau_k \quad \text{for each} \quad k = 0, 1, \cdots.$$

Proof. It is clear that the sequence $\{\tau_k\}_{k\in\mathbb{N}}$ is the sequence generated by Newton's method for solving $f(t) = 0$. Therefore, $\{\tau_k\}_{k\in\mathbb{N}}$ is well defined, strictly increasing and it converges to the root t_* (see the definition of f). By hypothesis, $q_0 \in \mathbf{K}(\tau_0)$; suppose that the points $q_0, q_1, \ldots, q_k$ are well defined. Then, using Banach's Lemma, we conclude that q_{k+1} is well defined. Furthermore,

$$d(q_{k+1}, q_k) \le \left\| -P_{\sigma_k,0,1} \mathcal{D}X(p_0)^{-1} P_{\sigma_k,1,0} X(q_k) \right\|.$$

Since

$$q_{k+1} = \exp_{q_k}\left(-P_{\sigma_k,0,1} \mathcal{D}X(p_0)^{-1} P_{\sigma_k,1,0} X(q_k) \right)$$

and $\sigma_k : [0,1] \longrightarrow M$ is the minimizing geodesic connecting p_0, q_k, from Lemma 6.3.5 and using (6.3.9) we obtain

$$d(q_{k+1}, q_k) \le \frac{f(\tau_k)}{|f'(0)|} = \tau_{k+1} - \tau_k. \tag{6.3.10}$$

Hence, for $k \ge s$, $s \in \mathbb{N}$,

$$d(q_k, q_s) \le \tau_s - \tau_k. \tag{6.3.11}$$

It follows that $\{q_k\}_{k\in\mathbb{N}}$ is a Cauchy sequence. Since M is complete, it converges to the some $q_* \in M$. Moreover $q_k \in \mathbf{K}(\tau_k) \subseteq B[p_0, t_*]$, therefore $q_* \in B[p_0, t_*]$.

Next, we prove that $X(q_*) = 0$. We have

$$\begin{aligned} \|X(q_k)\| &= \left\| P_{\sigma_k,0,1} \mathcal{D}X(p_0) P_{\sigma_k,1,0} P_{\sigma_k,0,1} \mathcal{D}X(p_0)^{-1} P_{\sigma_k,1,0} X(q_k) \right\| \\ &\le \|\mathcal{D}X(p_0)\| \left\| \mathcal{D}X(p_0)^{-1} X(q_k) \right\| \\ &\le (\|\mathcal{D}X(p_0)\|) \frac{f(\tau_k)}{|f'(0)|} \\ &= (\|\mathcal{D}X(p_0)\|)(\tau_{k+1} - \tau_k). \end{aligned}$$

Passing to the limit in k, we conclude $X(q_*) = 0$. Finally, letting $s \longrightarrow \infty$ in (6.3.11), we get

$$d(q_*, q_k) \le t_* - \tau_k.$$

□

- Finally, from (6.3.5)

$$d(q_*, q_k) \leq \frac{b}{h}\left(1 - \sqrt{1-2h}\right)^{k+1}, \quad k = 1, 2, \ldots$$

By hypothesis, $p_0 \in \mathbf{K}(0)$, thus by the Lemma 6.3.5, the sequence $\{p_k\}_{k\in\mathbb{N}}$ generated by (6.3.1) is well defined, contained in $B(p_0, t_*)$ and converges to some p_*, which is a singular point of X in $B[p_0, t_*]$. Moreover, if $h < 1/2$, then

$$d(p_k, p_*) \leq \frac{b}{h}\left(1 - \sqrt{1-2h}\right)^{k+1}.$$

- **UNIQUENESS**

This proof will be made in an indirect way, by contradiction. But before we are going to establish some results.

Lemma 6.3.7. *Take* $0 \leq t < t_*$ *and* $q \in \mathbf{K}(t)$, *let*

$$\begin{aligned} A^{-1} &= -P_{\sigma,0,1}\mathcal{D}X(p_0)^{-1}P_{\sigma,1,0} \\ v &= A^{-1}X(q), \end{aligned}$$

where $\sigma : [0,1] \longrightarrow M$ *is a minimizing geodesic connecting* p_0, p_k. *Define for* $\theta \in \mathbb{R}$

$$\begin{aligned} \tau(\theta) &= t + \theta a f(t), \\ \gamma(\theta) &= \exp_q(\theta v). \end{aligned}$$

Then, for $\theta \in [0,1]$

$$t < \tau(\theta) < t_* \quad \textit{and} \quad \gamma(\theta) \in \mathbf{K}(\tau(\theta)).$$

Proof. Because γ is a minimizing geodesic, for all $\theta \in [0,1]$ we have

$$\begin{aligned} d(p_0, \gamma(\theta)) &\leq d(p_0, q) + d(q, \gamma(\theta)) \\ &\leq t + \theta\|v\| \\ &\leq t + \theta a f(t) \\ &= \tau(\theta). \end{aligned}$$

This implies that

$$t \leq \tau(\theta) \leq \tau(1) \leq t_* \text{ and } \gamma([0,\theta]) \subset B(p_0, t_*). \tag{6.3.12}$$

Using the Lemma 6.3.2, we obtain

$$X(\gamma(\theta)) = P_{\gamma,0,\theta}\left(X(p) + P_{\sigma,0,1}\theta\mathcal{D}X(p_0)P_{\sigma,1,0}v + R(\theta)\right),$$

with

$$R(\theta) = \int_0^\theta \left(P_{\gamma,s,0}\mathcal{D}X(\gamma(s))P_{\gamma,0,s}v - P_{\sigma,0,1}\mathcal{D}X(p_0)P_{\sigma,1,0}v\right)ds,$$

and

$$\|R(\theta)\| \leq L\left(\frac{\theta}{2}\|v\| + d(p_0, q)\right)\theta\|v\|.$$

This yields to

$$\begin{aligned}
\left\|A^{-1}X(\gamma(\theta))\right\| &= \left\|A^{-1}P_{\gamma,0,\theta}\left(X(q) - \int_0^\theta P_{\gamma,s,0}\mathcal{D}X(\gamma(s))P_{\gamma,0,s}v\right)ds\right\| \\
&= \left\|A^{-1}P_{\gamma,0,\theta}\left((1-\theta)X(q) - \int_0^\theta \left(P_{\gamma,s,0}\mathcal{D}X(\gamma(s))P_{\gamma,0,s} - \mathcal{D}X(q)\right)v\right)ds\right\| \\
&\leq \left\|A^{-1}P_{\gamma,0,\theta}(1-\theta)X(q)\right\| + \left\|A^{-1}P_{\gamma,0,\theta}\int_0^\theta \left(P_{\gamma,s,0}\mathcal{D}X(\gamma(s))P_{\gamma,0,s} - \mathcal{D}X(q)\right)vds\right\| \\
&\leq (1-\theta)af(t) + a\|R(\theta)\| \\
&\leq (1-\theta)af(t) + aL\left(\frac{\theta}{2}\|v\| + d(p_0,q)\right)\theta\|v\| \\
&\leq (1-\theta)af(t) + aL\left(\frac{\theta}{2}af(t) + t\right)\theta af(t) \\
&= \frac{1}{8}\left(2b - 2t + aLt^2\right)\left(-4\theta + a^2L^2\theta^2t^2 + 4aL\theta t + 2abL\theta^2 - 2aL\theta^2 t + 4\right) \\
&= af(\tau(\theta)).
\end{aligned}$$

Therefore

$$\gamma(\theta) \in \mathbf{K}(\tau(\theta)),$$

and the Lemma is proved. □

Lemma 6.3.8. *Let $0 \leq t < t_*$ and $q \in \mathbf{K}(t)$. Suppose that $q_* \in B[p_0, t_*]$ is a singularity of the vector field X and*

$$t + d(q, q_*) = t_*.$$

Then

$$d(p_0, q) = t.$$

Moreover, letting

$$\begin{aligned}
t_+ &= t + af(t), \\
q_+ &= \exp_q\left(A^{-1}X(q)\right),
\end{aligned}$$

then $t < t_+ < t_*$, $q_+ \in \mathbf{K}(t_+)$ and

$$t_+ + d(q_+, q_*) = t_*.$$

Proof. Consider the minimizing geodesic $\alpha : [0,1] \longrightarrow M$ joining q to q_*. Since $q \in \mathbf{K}(t)$, we have

$$\begin{aligned}
d(p_0, \alpha(\theta)) &\leq d(p_0, q) + d(q, \alpha(\theta)) \\
&\leq t + \theta d(q, q_*) \\
&\leq t + d(q, q_*) \\
&= t_*.
\end{aligned}$$

It follows that $\alpha([0,1]) \subset B(p_0, t_*)$. Taking $u = \alpha'(0)$, by Lemma 6.3.2 we have

$$P_{\alpha,1,0}X(\alpha(1)) = X(q) + P_{\sigma,0,1}\mathcal{D}X(p_0)P_{\sigma,1,0}u + R(1),$$

with
$$\|R(1)\| \leq L\left(\frac{1}{2}\|u\| + d(p_0,q)\right)\|u\|.$$

Therefore
$$\begin{aligned}
\|R(1)\| &\leq L\left(\frac{1}{2}d(q,q_*) + d(p_0,q)\right)d(q,q_*) && (6.3.13)\\
&= L\left(\frac{1}{2}(t_*-t) + d(p_0,q)\right)(t_*-t)\\
&\leq L\left(\frac{1}{2}(t_*-t) + t\right)(t_*-t) && (6.3.14)\\
&= L\frac{1}{2}(t_*+t)(t_*-t).
\end{aligned}$$

On the other hand, since $|f(t)|$ is strictly decreasing in $[0,t_*]$ and $0 \leq d(p_0,q) \leq t < t_*$,
$$\begin{aligned}
\|R(1)\| &= \|X(q) + Au\| && (6.3.15)\\
&\geq \frac{1}{\left\|\mathcal{D}X(p_0)^{-1}\right\|}\left\|A^{-1}X(q)+u\right\|\\
&\geq \left|f'(0)\right|\left\|A^{-1}X(q)+u\right\|\\
&\geq \left|f'(0)\right|\left(\|u\| - \left\|A^{-1}X(q)\right\|\right)\\
&\geq \left|f'(0)\right|(\|u\| - af(t))\\
&= -f'(0)(t_*-t) - f(t) > 0.
\end{aligned}$$

Because
$$\begin{array}{l}
f''(t) = L,\\
0 = f'(t_*) = f(t) + f'(t)(t_*-t) + \frac{1}{2}f''(t)(t_*-t)^2,
\end{array}$$

and
$$f'(t) = f'(0) + \int_0^t f''(t)\,dt,$$

therefore
$$0 = f(t) + \left(f'(0) + tL\right)(t_*-t) + \frac{1}{2}L(t_*-t)^2,$$

hence
$$\frac{1}{2}L(t_*+t)(t_*-t) = -f'(0)(t_*-t) - f(t).$$

Thus, the last term in (6.3.13) is equal to the last term in the inequality (6.3.15), we conclude that all these inequalities in (6.3.15) are equalities, in particular

$$\begin{array}{l}
\frac{1}{\left\|\mathcal{D}X(p_0)^{-1}\right\|} = |f'(0)| = a,\\
\|u\| - \left\|A^{-1}X(q)\right\| = \left\|A^{-1}X(q)+u\right\| > 0,\\
\left\|A^{-1}X(q)\right\| = af(t),\\
L\left(\frac{1}{2}(t_*-t) + d(p_0,q)\right)(t_*-t) = L\left(\frac{1}{2}(t_*-t)+t\right)(t_*-t).
\end{array} \qquad (6.3.16)$$

From the last equation in (6.3.16), we obtain

$$d(p_0, q) = t,$$

the second equation in (6.3.16) implies that u and $A^{-1}X(q)$ are linearly dependent vectors in T_qM, so that there exists $r \in \mathbb{R}$ such that

$$A^{-1}X(q) = -ru.$$

Thus, the second equation implies

$$1 - |r| = |1 - r|,$$

and because $r \neq 0$ and $r \neq 1$, we have $0 < r < 1$, thus

$$q_+ = \exp_q(ru) = \alpha(r).$$

Moreover, given that α is a minimizing geodesic joining q to q_*, we have that q, $\alpha(r)$ and q_* are in the same geodesic line, thus

$$d(q, \alpha(r)) + d(\alpha(r), q_*) = d(q, q_*),$$

therefore,

$$d(q, q_+) + d(q_+, q_*) = d(q, q_*).$$

Moreover,

$$d(q, q_+) = \|ru\| = \left\|A^{-1}X(q)\right\| = af(t) = t_+ - t,$$

hence

$$d(q_+, q_*) = d(q, q_*) - d(q, q_+) = (t_* - t) - (t_+ - t) = t_* - t_+,$$

that is

$$d(q_+, q_*) + t_+ = t_*.$$

□

Corollary 6.3.9. *Suppose that $q_* \in B[p_0, t_*]$ is a zero of the vector field X. If there exist $\tilde{t}$ and $\tilde{q}$ such that*

$$0 \leq \tilde{t} < t_*, \quad \tilde{q} \in \mathbf{K}(\tilde{t}) \quad \text{and} \quad \tilde{t} + d(\tilde{q}, q_*) = t_*,$$

then

$$d(p_0, q_*) = t_*.$$

Proof. Changing τ_0 by $\tilde{t}$ and q_0 by $\tilde{q}$ in Corollary 6.3.6, we obtain that

$$q_k \in \mathbf{K}(\tau_k), \text{ for all } k \in \mathbb{N},$$

$\{\tau_k\}_{k\in\mathbb{N}}$ converges to t_*, $\{q_k\}_{k\in\mathbb{N}}$ converges to some $\tilde{q}_* \in B(p_0, t_*)$, and $X(q_*) = 0$. Moreover, by Lemma 6.3.8 and applying induction, it is easy to show that for all k,

$$d(p_0, q_k) = \tau_k \quad \text{and} \quad d(q_k, q_*) = t_* - \tau_k.$$

Passing to the limit, we obtain

$$d(p_0, \tilde{q}_*) = t_* \quad \text{and} \quad d(\tilde{q}_*, q_*) = 0.$$

Therefore $\tilde{q}_* = q_*$ and

$$d(p_0, q_*) = t_*.$$

□

The two following Lemmas complete the proof of the uniqueness.

Lemma 6.3.10. *The limit p_* of the sequence $\{p_k\}_{k\in\mathbb{N}}$ is the unique singularity of X in $B[p_0, t_*]$.*

Proof. Let $q_* \in B[p_0, t_*]$ a singularity of the vector field X. Using induction, we will show that

$$d(p_k, q_*) + t_k \leq t_*.$$

We need to consider two cases:

Case 1. ($d(p_0, q_*) < t_*$). First we show by induction that for all $k \in \mathbb{N}$,

$$d(p_k, q_*) + t_k < t_*. \tag{6.3.17}$$

Indeed, for $k = 0$ (6.3.17) is immediately true, because $t_0 = 0$. Now, suppose the property is true for some k. Let us take the geodesic

$$\gamma_k(\theta) = \exp_{p_k}(-\theta v_k),$$

where v_k is defined in (6.3.1). From Lemma 6.3.7, for all $\theta \in [0, 1]$,

$$\gamma_k(\theta) \in \mathbf{K}(t_k + \theta(t_{k+1} - t_k)). \tag{6.3.18}$$

Define $\phi : [0, 1] \longrightarrow M$ by

$$\phi(\theta) = d(\gamma_k(\theta), q_*) + t_k + \theta(t_{k+1} - t_k). \tag{6.3.19}$$

We know that

$$\phi(0) = d(p_k, q_*) + t_k < t_*.$$

We next show, by contradiction, that $\phi(\theta) \neq t_*$ for all $\theta \in [0, 1]$.

Suppose that there exists a $\tilde{\theta} \in [0, 1]$ such that $\phi(\tilde{\theta}) = t_*$, and let $\tilde{q} = \gamma_k(\tilde{\theta})$ and $\tilde{t} = t_k + \tilde{\theta}(t_{k+1} - t_k)$. By (6.3.18) and (6.3.19),

$$\tilde{q} \in \mathbf{K}(\tilde{t}) \text{ and } d(\tilde{q}, q_*) + \tilde{t} = t_*.$$

Applying Corollary 6.3.9, we conclude that

$$d(p_0, q_*) = t_*,$$

which contradicts our assumption. Thus $\phi(\theta) \neq t_*$ for all $\theta \in [0,1]$, Since $\phi(0) < t_*$ and ϕ is continuous, we have that $\phi(\theta) < t_*$ for all $\theta \in [0,1]$. In particular, by (6.3.19),

$$d(\gamma_k(1), q_*) + t_{k+1} = \phi(1) < t_*.$$

Thus,

$$d(p_{k+1}, q_*) + t_{k+1} < t_*,$$

in this way (6.3.17) is true for all $k \in \mathbb{N}$.

Case 2. ($d(p_0, q_*) = t_*$). Using induction, let us prove that for all $k \in \mathbb{N}$,

$$d(p_k, q_*) + t_k = t_*. \tag{6.3.20}$$

Indeed, for $k = 0$, this is immediately true, because $t_0 = 0$. Now, suppose that (6.3.20) is true for some k. Since $p_k \in \mathbf{K}(t_k)$, by Lemma 6.3.8 we conclude that

$$d(p_{k+1}, q_*) + t_{k+1} = t_*.$$

Finally, by (6.3.17) and (6.3.20) we conclude that for all $k \in \mathbb{N}$,

$$d(p_k, q_*) + t_k \leq t_*,$$

and passing to the limit $k \longrightarrow \infty$, we obtain $d(p_*, q_*) = 0$, and therefore

$$p_* = q_*.$$

□

Lemma 6.3.11. *If* $h = abL < \frac{1}{2}$ *and* $B(p_0, r) \subseteq \Omega$, *with*

$$t_* < r \leq t_{**} = \frac{1}{aL}\left(1 + \sqrt{1 - 2h}\right),$$

then the limit p_* *of the sequence* $\{p_k\}_{k \in \mathbb{N}}$ *is the unique singularity of the vector field* X *in* $B(p_0, r)$.

Proof. Let $q_* \in B(p_0, r)$ be a singularity of the vector field X in $B(p_0, r)$. Let us consider the minimizing geodesic $\alpha : [0,1] \longrightarrow M$ joining p_0 to q_*. By Lemma 6.3.2,

$$P_{\alpha,1,0} X(\alpha(1)) = X(p_0) + P_{\sigma,0,1} \mathcal{D}X(p_0) P_{\sigma,1,0} u + R(1),$$

where

$$\|R(1)\| \leq L\left(\frac{1}{2}\|u\| + d(p_0, p_0)\right)\|u\| = \frac{L}{2} d(p_0, q_*)^2 \text{ and } \|u\| = d(p_0, q_*). \tag{6.3.21}$$

In a similar way to the inequality (6.3.15), is easy to prove that

$$\begin{aligned} \|R(1)\| &\geq \frac{1}{a}\left(\|u\| - \left\|\mathcal{D}X(p_0)^{-1} X(p_0)\right\|\right) \\ &\geq \frac{1}{a} d(p_0, q_*) - \frac{b}{a}. \end{aligned}$$

Therefore

$$\frac{L}{2}d(p_0,q_*)^2 \geq \frac{1}{a}d(p_0,q_*) - \frac{b}{a},$$

hence

$$f(d(p_0,q_*)) \geq 0,$$

since $d(p_0,q_*) \leq r \leq t_{**}$, then

$$d(p_0,q_*) \leq t_*.$$

Finally, from Lemma 6.3.10,

$$p_* = q_*.$$

□

6.4. Order of Convergence of Newton-Type Methods

The analysis of the order of convergence is performed in a local way, that is, in a neighborhood of the zero of the vector field. Then, we can define the order of convergence in Riemannian manifolds in following way:

Definition 6.4.1. *Let M be a manifold and let $\{p_k\}_{k\in\mathbb{N}}$ be a sequence on M converging to p_*. If there exists a system of coordinates (U,x) of M with $p_* \in U_\alpha$, constants $p > 0, c \geq 0$ and $K \geq 0$ such that, for all $k \geq K$, $\{p_k\}_{k=K}^{\infty} \subseteq U_\alpha$ the following inequality holds:*

$$\left\|x^{-1}(p_{k+1}) - x^{-1}(p_*)\right\| \leq c\left\|x^{-1}(p_k) - x^{-1}(p_*)\right\|^p, \tag{6.4.1}$$

then we said that $\{p_k\}_{k\in\mathbb{N}}$ converges to p_* with order at least p.

It can be shown that the definition above do not depend on the choice of the coordinates system and the multiplicative constant c depends on the chart, but for any chart, there exists such a constant, (see [1]).

Notice that in normal coordinates of 0_{p_k},

$$\left\|\exp_{p_k}^{-1}(p) - \exp_{p_k}^{-1}(q)\right\| = d(p,q),$$

thus, in normal coordinates, (6.4.1) is transformed into

$$d(p_{k+1},p_*) \leq c d(p_k,p_*)^p.$$

Lemma 6.4.2. *Let M be an Riemannian manifold, $\Omega \subseteq M$ be an open set, $X \in \chi x(M)$ and $\mathcal{D}X \in Lip_l(\Omega)$. Let us take $p \in \Omega$, $v \in T_pM$ and*

$$\gamma(t) = \exp_p(tv).$$

If $\gamma[0,t) \subseteq \Omega$, then

$$P_{\gamma,t,0}X(\gamma(t)) = X(p) + t\mathcal{D}X(p)v + R(t)$$

with

$$\|R(t)\| \leq \frac{l}{2}t^2\|v\|^2$$

Proof. From Theorem 6.2.9, it follows that

$$P_{\gamma,t,0}X(\gamma(t)) - X(\gamma(0)) = \int_0^t P_{\gamma,s,0}\left(\mathcal{D}X(\gamma(s))\gamma'(s)\right)ds.$$

Given that γ is a geodesic, we have that $\gamma'(t)$ is parallel and $\gamma'(s) = P_{\gamma,0,s}\gamma'(0)$. Moreover, since $\gamma'(0) = v$ then

$$P_{\gamma,t,0}X(\gamma(t)) - X(p) = \int_0^t P_{\gamma,s,0}\left(\mathcal{D}X(\gamma(s))P_{\gamma,0,s}v\right)ds.$$

Therefore

$$\begin{aligned}
&P_{\gamma,t,0}X(\gamma(t)) - X(p) - t\mathcal{D}X(p)v \\
&= \int_0^t P_{\gamma,s,0}\left(\mathcal{D}X(\gamma(s))P_{\gamma,0,s}v\right)ds - t\mathcal{D}X(p)v \\
&= \int_0^t \left(P_{\gamma,s,0}\left(\mathcal{D}X(\gamma(s))P_{\gamma,0,s}v\right) - \mathcal{D}X(p)v\right)ds,
\end{aligned}$$

let

$$R(t) = \int_0^t \left(P_{\gamma,s,0}\mathcal{D}X(\gamma(s))P_{\gamma,0,s}v - \mathcal{D}X(p)v\right)ds.$$

By hypothesis, $\mathcal{D}X \in Lip_L(\Omega)$, hence

$$\begin{aligned}
\|R(t)\| &\leq \int_0^t \left\|\left(P_{\gamma,s,0}\mathcal{D}X(\gamma(s))P_{\gamma,0,s}v - \mathcal{D}X(p)v\right)\right\|ds \\
&\leq \int_0^t \left\|\left(P_{\gamma,s,0}\mathcal{D}X(\gamma(s))P_{\gamma,0,s} - \mathcal{D}X(p)\right)\right\|\|v\|ds \\
&\leq \int_0^t \left(L\int_0^s \left\|\gamma'(\tau)\right\|d\tau\right)\|v\|ds.
\end{aligned}$$

Since γ is a geodesic, $\|\gamma'(\tau)\|$ is constant. Therefore,

$$\left\|\gamma'(\tau)\right\| = \left\|\gamma'(0)\right\| = \|v\|,$$

thus

$$\|R(t)\| \leq \int_0^t \left(L\int_0^s \|v\|d\tau\right)\|v\|ds = \int_0^t L\|v\|s\|v\|ds = \frac{L}{2}t^2\|v\|^2.$$

□

Lemma 6.4.3. *(Order of convergence)*

i) The convergence order of the Newton method in Riemannian manifold is two (quadratic convergence).

ii) The convergence order of the simplified Newton method in Riemannian manifold is one (linear convergence).

Proof. Let k be sufficiently large in such a way that p_*, p_k, $p_{k+1}, \ldots,$ belong to a normal neighborhood U of p_k. Let us consider the geodesic γ_k joining p_k to p_* defined by

$$\gamma_k(t) = \exp_{p_k}(tu_k),$$

where $u_k \in T_{p_k}M$ and $d(p_k, p_*) = \|u_k\|$.

We know that if p, q be in one normal neighborhood U of p_k, then

$$\left\|\exp_{p_k}^{-1}(p) - \exp_{p_k}^{-1}(q)\right\| = d(p,q).$$

i) By Lemma 6.3.2,

$$P_{\gamma,t,o}X(p_*) = X(p_k) + \mathcal{D}X(p_k)u_k + R(1),$$

with

$$\|R(1)\| \leq \frac{L}{2}\|u_k\|^2 \text{ and } \|u_k\| = d(p_k, p_*).$$

Hence,

$$0 = \mathcal{D}X(p_k)^{-1}X(p_k) + u_k + \mathcal{D}X(p_k)^{-1}R(1).$$

Since

$$-\mathcal{D}X(p_k)^{-1}X(p_k) = \exp_{p_k}^{-1}(p_{k+1}) \quad \text{and} \quad u_k = \exp_{p_k}^{-1}(p_*),$$

we have

$$\exp_{p_k}^{-1}(p_{k+1}) - \exp_{p_k}^{-1}(p_*) = \mathcal{D}X(p_k)^{-1}R(1),$$

thus

$$d(p_{k+1}, p_*) \leq \left\|\mathcal{D}X(p_k)^{-1}\right\| \frac{L}{2}\|u_k\|^2.$$

Moreover, by Banach's Lemma,

$$\left\|\mathcal{D}X(p_k)^{-1}\right\| \leq \frac{a}{1 - ald(p_k, p_0)} \leq \frac{a}{1 - al\tau_k} \leq \frac{a}{1 - alt_*} = \frac{a}{\sqrt{1-2abl}}.$$

Therefore

$$d(p_{k+1}, p_*) \leq Cd(p_k, p_*)^2,$$

with

$$C = \frac{La}{2\sqrt{1-2abL}}.$$

ii) Let p_0 be sufficiently near to p_* in such a way that p_0 is in the normal neighborhood U of 0_{p_k}, By Lemma 6.3.2, if $\sigma_k : [0,1] \longrightarrow M$ is the minimizing geodesic connecting p_0, p_k, then

$$P_{\gamma,1,0}X(p_*) = X(p_k) + P_{\sigma_k,0,1}\mathcal{D}X(p_0)P_{\sigma_k,1,0}u_k + R(1),$$

with

$$\|R(1)\| \leq L\left(\frac{1}{2}\|u_k\| + d(p_0, p_k)\right)\|u_k\|.$$

Therefore

$$0 = P_{\sigma_k,0,1}\mathcal{D}X(p_0)^{-1}P_{\sigma_k,1,0}X(p_k) + u_k + P_{\sigma_k,0,1}\mathcal{D}X(p_0)^{-1}P_{\sigma_k,1,0}R(1).$$

Since

$$-P_{\sigma_k,0,1}\mathcal{D}X(p_0)^{-1}P_{\sigma_k,1,0}X(p_k) = \exp_{p_k}^{-1}(p_{k+1}) \quad \text{and} \quad u_k = \exp_{p_k}^{-1}(p_*),$$

we have

$$\exp_{p_k}^{-1}(p_{k+1}) - \exp_{p_k}^{-1}(p_*) = P_{\sigma_k,0,1}\mathcal{D}X(p_0)^{-1} P_{\sigma_k,1,0}R(1).$$

We thus conclude that

$$\begin{aligned}
d(p_{k+1},p_*) &= \left\|\exp_{p_k}^{-1}(p_{k+1}) - \exp_{p_k}^{-1}(p_*)\right\| \\
&= \left\|P_{\sigma_k,0,1}\mathcal{D}X(p_0)^{-1} P_{\sigma_k,1,0}R(1)\right\| \\
&\leq \left\|\mathcal{D}X(p_0)^{-1}\right\| \|R(1)\| \\
&\leq al\left(\frac{1}{2}\|u_k\| + d(p_0,p_k)\right)\|u_k\| \\
&= al\left(\frac{1}{2}d(p_k,p_*) + d(p_0,p_k)\right)d(p_k,p_*) \\
&= al\left(\frac{1}{2}\frac{d(p_k,p_*)}{d(p_0,p_k)} + 1\right)d(p_0,p_k)\,d(p_k,p_*).
\end{aligned}$$

If k is sufficiently large, then $d(p_k,p_*) \leq d(p_0,p_k)$, and therefore

$$\left(\frac{1}{2}\frac{d(p_k,p_*)}{d(p_0,p_k)} + 1\right) \leq \frac{3}{2},$$

and then, for p_0 sufficiently close to p_*,

$$d(p_{k+1},p_*) \leq K_0 d(p_0,p_k)\,d(p_k,p_*),$$

with $K_0 \leq \frac{3aL}{2}$. □

Remark 6.4.4. *Note that if instead of putting in the Kantorovich method the point p_0, we fix p_j sufficiently close to p_*, we will obtain a new convergent method. Indeed the calculations made in the previous lemma become in*

$$d(p_{k+1},p_*) \leq K_j d(p_j,p_*)\,d(p_k,p_*),$$

with $K_j \leq \frac{3aL}{2}$. Thus,

$$d(p_{k+1},p_*) \leq K d(p_j,p_*)\,d(p_k,p_*), \tag{6.4.2}$$

with $K \leq \frac{3aL}{2}$.

6.5. One Family of High Order Newton-Type Methods

We recall the Shamanskii family of iterative methods. Given an integer m and an initial point x_0 in a Banach Space, we move from x_n to x_{n+1} through an intermediate sequence $\{y_n^i\}_{i=0}^m$, $y_n^0 = x_n$, which is a generalization of Newton ($m = 1$) and simplified Newton ($m = \infty$) methods

$$\begin{cases}
y_n^1 = y_n^0 - \mathcal{D}F\left(y_n^0\right)^{-1} F\left(y_n^0\right) \\
y_n^2 = y_n^1 - \mathcal{D}F\left(y_n^0\right)^{-1} F\left(y_n^1\right) \\
\vdots \\
y_n^m = x_{n+1} = y_n^{m-1} - \mathcal{D}F\left(y_n^0\right)^{-1} F\left(y_n^{(m-1)}\right),
\end{cases}$$

For a problem on Riemannian manifolds, let us consider the family

$$\begin{cases} q_n^1 = \exp_{p_n}\left(-\mathcal{D}X(p_n)^{-1}X(p_n)\right) \\ q_n^2 = \exp_{q_n^1}\left(-P_{\sigma_1,0,1}\mathcal{D}X(p_n)^{-1}P_{\sigma_1,1,0}X\left(q_n^1\right)\right) \\ \vdots \\ q_n^m = p_{n+1} = \exp_{q_n^{m-1}}\left(-P_{\sigma_{m-1},0,1}\mathcal{D}X(p_n)^{-1}P_{\sigma_{m-1},1,0}X\left(q_n^{(m-1)}\right)\right), \end{cases} \tag{6.5.1}$$

where $\sigma_k : [0,1] \longrightarrow M$ be the minimizing geodesic joining the points p_n and q_n^k; $k = 1,2,\ldots,(m-1)$, thus:

$$\sigma_k(0) = p_n \quad \text{and} \quad \sigma_k(1) = q_n^k.$$

Theorem 6.5.1. *Under the hypotheses of Kantorovich's theorem, the method described in (6.5.1) converges with order of convergence* $m+1$.

Proof. Let us observe that

$$d(p_{n+1},p_n) \le d\left(p_{n+1},q_n^{(m-1)}\right) + d\left(q_n^{(m-1)},q_n^{(m-2)}\right) + \cdots + d\left(q_n^2,q_n^1\right) + d\left(q_n^1,p_n\right).$$

Now, if we define $p_{n+1} = q_n^m$, $p_n = q_n^1$, looking at each step as a different method according to (6.5.1) , then by Kantorovich theorem in the first step and by the simplified Kantorovich theorem for the following steps, each one of the sequences $\{q_n^m\}_{m\in\mathbb{N}}$ for fixed n, is convergent to the same point $p_* \in M$. Therefore, $\{p_n\}_{p\in\mathbb{N}}$ is convergent to p_*. Moreover, for Lemma 6.4.3 i) and (6.4.2),

$$\begin{aligned} d(p_{n+1},p_*) &\le Kd(p_n,p_*)\,d\left(q_n^{(m-1)},p_*\right) \le Kd(p_n,p_*)\,Kd(p_n,p_*)\,d\left(q_n^{(m-2)},p_*\right) \\ &\le \cdots \le K^{m-1}d(p_n,p_*)^{m-1}d\left(q_n^1,p_*\right) \le K^{m-1}d(p_n,p_*)^{m-1}Cd(p_n,p_*)^2. \end{aligned}$$

Therefore,

$$d(p_{n+1},p_*) \le CK^{m-1}d(p_n,p_*)^{m+1}.$$

□

6.6. Expanding the Applicability of Newton Methods

We have used Lipschitz condition (6.2.13) and the famous Kantorovich sufficient convergence criterion (6.1.1) in connection to majorizing function f for the semilocal convergence of both simplified Newton and Newton methods. According to the proof of Lemma 6.3.2, corresponding majorizing sequences for these methods are given by (see [27])

$$\begin{aligned} &t_0 = 0,\ t_1 = b \\ &t_{k+1} = t_k + \frac{f(t_k)}{|f'(0)|} = t_k + \frac{a\,l}{2}(t_k - t_{k-1})^2 \quad \text{for each} \quad k = 1,2,\cdots \end{aligned} \tag{6.6.1}$$

for the simplified Newton method and

$$\begin{aligned} &u_0 = 0,\ u_1 = b \\ &u_{k+1} = u_k + \frac{f(u_k)}{|f'(u_k)|} = u_k + \frac{a\,l\,(u_k - u_{k-1})^2}{2\,(1 - a\,l\,u_k)} \quad \text{for each} \quad k = 1,2,\cdots \end{aligned} \tag{6.6.2}$$

for the Newton method. Kantorovich criterion (6.1.1) may be not satisfied on a particular problem but Newton methods may still converges to p_* [27]. Next, we shall show that condition (6.1.1) can be weakened by introducing the center Lipschitz condition and relying on tighter majorizing sequences instead of majorizing function f.

Definition 6.6.1. *Let E be a Banach space, $\Omega \subseteq E$ be an open convex set, $F : \Omega \longrightarrow \Omega$ be a continuous operator, x_0 be a point in Ω, such that $F \in C^1$ and DF is center-Lipschitz in Ω at x_0*

$$\| DF(x) - DF(x_0) \| \leq l_0 \, \| x - x_0 \| \quad \text{for each} \quad x \in \Omega \quad \text{and some} \quad l_0 > 0.$$

As in the case of Definition 6.2.7 we will write $DF \in Lip_{l_0}(\Omega)$ at $x_0 \in \Omega$.

Note that

$$l_0 \leq l \tag{6.6.3}$$

holds in general and l/l_0 can be arbitrarily large [**?**], [14].

We present the semilocal convergence of the simplified Newton method using only the center-Lipschitz condition.

Theorem 6.6.2. *Let M be a Riemannian manifold, $\Omega \subseteq M$ be an open convex set, $X \in \chi(M)$. Suppose that for some $p_0 \in \Omega$, $DX \in Lip_{l_0}(\Omega)$ at p_0, $DX(p_0)$ is invertible and that for some $a > 0$ and $b \geq 0$, the following hold*

$$\left\| DX\left(p_0\right)^{-1} \right\| \leq a,$$

$$\left\| DX\left(p_0\right)^{-1} X\left(x_0\right) \right\| \leq b,$$

$$h_0 = a\, b\, l_0 \leq \frac{1}{2} \tag{6.6.4}$$

and

$$B\left(p_0, t_*^0\right) \subseteq \Omega \quad \text{where} \quad t_*^0 = \frac{1}{a\, l_0}\left(1 - \sqrt{1 - 2\, h_0}\right).$$

Then, sequence $\{p_k\}$ generated by (6.3.1) is such that $\{p_k\} \subseteq B(p_0, t_*^0)$ and $p_k \to p_*$, which the only singularity of X in $B(p_0, t_*^0)$. Moreover, if $h_0 < 1/2$ and $B(p_0, r) \subseteq \Omega$ with

$$t_*^0 < r \leq t_{**}^0 = \frac{1}{a\, l_0}\left(1 + \sqrt{1 - 2\, h_0}\right)$$

and p_* is also the only singularity of F in $B(p_0, r)$. Furthemore, the following error bounds are satisfied fr each $k = 1, 2, \cdots$

$$d(p_k, p_{k-1}) \leq t_k^0 - t_{k-1}^0,$$

$$d(p_k, p_*) \leq t_*^0 - t_k^0$$

and

$$d(p_k, p_*) \leq \frac{b}{h_0}(1 - \sqrt{1 - 2\, h_0})^{k+1},$$

where sequence $\{t_k^0\}$ is defined by

$$t_0^0 = 0,\ t_1^0 = b$$
$$t_{k+1}^0 = t_k^0 + \frac{a\,l_0}{2}\,(t_k^0 - t_{k-1}^0)^2 \quad \text{for each} \quad k = 1, 2, \cdots.$$

Proof. Simply notice that l_0, h_0, $\{t_k^0\}$, t_*^0, t_{**}^0 can replace l, h, $\{t_k\}$, t_*, t_{**}, respectively, in the proof of Theorem 6.3.1. □

Remark 6.6.3. *Under Kantorovich criterion (6.1.1) a simple inductive argument shows that*

$$t_k^0 \le t_k \quad \text{and} \quad t_{k+1}^0 - t_k^0 \le t_{k+1} - t_k \quad \text{for each} \quad k = 0, 1, \cdots.$$

Moreover, we have that

$$t_*^0 \le t_*, \quad t_{**} \le t_{**}^0, \quad h \le \frac{1}{2} \Longrightarrow h_0 \le \frac{1}{2}$$

and

$$\frac{h_0}{h} \longrightarrow 0 \quad \text{as} \quad \frac{l_0}{l} \longrightarrow 0.$$

Furthemore, strict inequality holds in these estimates (for $k > 1$) if $l_0 < l$.

The convergence order of simplified method is only linear, whereas the convergence order of Newton method is quadratic if $h < 1/2$. If criterion $h \le 1/2$ is not satisfied but weaker $h_0 \le 1/2$ is satisfied, we can start with the simplified method until a certain iterate x_N (N a finite natural integer) at which criterion $h \le 1/2$ is satisfied. Such an integer N exists. Since the simplified Newton method converges [8], [12], [14]. This approach was not possible before since $h \le 1/2$ was used at the convergence criterion for both methods.

Remark 6.6.4. *Under the hypotheses of Theorem 6.1.2, we see in the proof of this Theorem, sequences $\{r_k\}$, $\{s_k\}$ defined by*

$$r_0 = 0,\ r_1 = b,\ r_2 = r_1 + \frac{a\,l_0\,(r_1 - r_0)^2}{2\,(1 - a\,l_0\,r_1)^2}$$
$$r_{k+1} = r_k + \frac{a\,l\,(r_k - r_{k-1})^2}{2\,(1 - a\,l_0\,r_k)} \quad \text{for each} \quad k = 2, 3, \cdots \tag{6.6.5}$$

$$s_0 = 0,\ s_1 = b$$
$$s_{k+1} = s_k + \frac{a\,l\,(s_k - s_{k-1})^2}{2\,(1 - a\,l_0\,s_k)} \quad \text{for each} \quad k = 1, 2, \cdots \tag{6.6.6}$$

are also majorizing sequences for $\{p_k\}$ such that

$$r_k \le s_k \le u_k, \quad d(p_k, p_{k-1}) \le r_k - r_{k-1} \le s_k - s_{k-1} \le u_k - u_{k-1}$$

and

$$r_* = \lim_{k \to \infty} r_k \le s_* = \lim_{k \to \infty} s_k \le t_* = \lim_{k \to \infty} u_k.$$

Simply notice that for the computation of the upper bound on the norms $\| DX(q)^{-1} \|$ *(see (6.3.8)), we can have using the center-Lipschitz condition*

$$\| DX(q)^{-1} \| \leq \frac{\| \phi^{-1} \|}{1 - \| \phi^{-1} \| \| DX(q) - \phi \|} \leq \frac{a}{1 - a l_0 \lambda}$$

instead of the less tight (if $l_0 < l$*) and more expensive to compute estimate*

$$\| DX(q)^{-1} \| \leq \frac{a}{1 - a l \lambda}$$

obtained in the proofs of Theorems 6.1.2 and 6.3.1 using the Lipschitz condition. Hence, the results of Theorem 6.1.2 involving sequence $\{u_k\}$ *can be rewritten using tighter sequences* $\{r_k\}$ *or* $\{s_k\}$*. Note that the introduction of the center-Lipschitz condition is not an additional hypothesis to Lipschitz condition since in practice, the computation of* l *requires the computation of* l_0*. So far we showed that under Kantorovich criterion (6.1.1) the estimates of the distances* $d(p_k, p_{k-1})$, $d(p_k, p_*)$ *are improved (if* $l_0 < l$*) using tighter sequences* $\{r_k\}$, $\{s_k\}$ *for the computation on the upper bounds of these distances. Moreover, the information on the location of the solution is at least as precise.*

Next, we shall show that Kantorovich criterion (6.1.1) can be weakened if one directly (and not through majorizing function f) studies the convergence of sequences $\{r_k\}$ and $\{s_k\}$. First, we present the results for sequence $\{s_k\}$.

Lemma 6.6.5. *[13] Assume there exist constants* $l_0 \geq 0$, $l \geq 0$, $a > 0$ *and* $b \geq 0$ *with* $l_0 \leq l$ *such that*

$$h_1 = \bar{l} b \begin{cases} \leq 1/2 & \text{if} \quad l_0 \neq 0 \\ < 1/2 & \text{if} \quad l_0 = 0 \end{cases} \tag{6.6.7}$$

where $\bar{l} = \frac{a}{8}\left(l + 4 l_0 + \sqrt{l^2 + 8 l_0 l}\right)$*. Then, sequence* $\{s_n\}$ *given by (6.6.6) is nondecreasing, bounded from above by* $s^{\star\star}$ *and converges to its unique least upper bound* $s^\star \in [0, s^{\star\star}]$*, where*

$$s^{\star\star} = \frac{2b}{2 - \delta} \quad \text{and} \quad \delta = \frac{4l}{l + \sqrt{l^2 + 8 l_0 l}} < 1 \quad \text{for} \quad l_0 \neq 0. \tag{6.6.8}$$

Moreover the following estimates hold

$$a l_0 s^\star \leq 1, \tag{6.6.9}$$

$$0 \leq s_{n+1} - s_n \leq \frac{\delta}{2}(s_n - s_{n-1}) \leq \cdots \leq \left(\frac{\delta}{2}\right)^n b \quad \text{for each} \quad n = 1, 2, \cdots, \tag{6.6.10}$$

$$s_{n+1} - s_n \leq \left(\frac{\delta}{2}\right)^n (2h_1)^{2^n - 1} b \quad \text{for each} \quad n = 0, 1, \cdots \tag{6.6.11}$$

and

$$0 \leq s^\star - s_n \leq \left(\frac{\delta}{2}\right)^n \frac{(2h_1)^{2^n - 1} b}{1 - (2h_1)^{2^n}}, \quad (2h_1 < 1) \quad \text{for each} \quad n = 0, 1, \cdots. \tag{6.6.12}$$

Lemma 6.6.6. *[16] Suppose that hypotheses of Lemma 6.6.5 hold. Assume that*

$$h_2 = l_2 b \le \frac{1}{2}, \tag{6.6.13}$$

where $l_2 = \frac{a}{8}(4l_0 + (l\,l_0 + 8l_0^2)^{1/2} + (l_0 l)^{1/2})$. *Then, scalar sequence* $\{r_n\}$ *given by (6.6.5) is well defined, increasing, bounded from above by*

$$r^{\star\star} = b + \frac{a\,l_0 b^2}{2(1-(\delta/2))(1-a\,l_0 b)} \tag{6.6.14}$$

and converges to its unique least upper bound $r^\star$ *which satisfies* $0 \le r^\star \le r^{\star\star}$. *Moreover, the following estimates hold*

$$0 < r_{n+2} - r_{n+1} \le (\delta/2)^n \frac{a\,l_0 b^2}{2(1-a\,l_0 b)} \quad \text{for each} \quad n = 1, 2, \cdots. \tag{6.6.15}$$

Lemma 6.6.7. *[16] Suppose that hypotheses of Lemma 6.6.5 hold and there exists a minimum integer* $N > 1$ *such that iterates* r_i $(i = 0, 1, \cdots, N-1)$ *given by (6.6.5) are well defined,*

$$r_i < r_{i+1} < \frac{1}{a\,l_0} \quad \text{for each} \quad i = 0, 1, \cdots, N-2 \tag{6.6.16}$$

and

$$r_N \le \frac{1}{a\,l_0}(1-(1-a\,l_0\,r_{N-1})\frac{\delta}{2}). \tag{6.6.17}$$

Then, the following assertions hold

$$a\,l_0\,r_N < 1, \tag{6.6.18}$$

$$r_{N+1} \le \frac{1}{a\,l_0}(1-(1-a\,l_0\,r_N)\frac{\delta}{2}), \tag{6.6.19}$$

$$\delta_{N-1} \le \frac{\delta}{2} \le 1 - \frac{a\,l_0(r_{N+1}-r_N)}{1-a\,l_0\,r_N}, \tag{6.6.20}$$

sequence $\{r_n\}$ *given by (6.6.5) is well defined, increasing, bounded from above by*

$$r^{\star\star} = r_{N-1} + \frac{2}{2-\delta}(r_N - r_{N-1})$$

and converges to its unique least upper bound $r^\star$ *which satisfies* $0 \le r^\star \le r^{\star\star}$, *where* δ *is given in Lemma 6.6.5 and*

$$\delta_n = \frac{a\,l(r_{n+2}-r_{n+1})}{2(1-a\,l_0\,r_{n+2})}.$$

Moreover, the following estimates hold

$$0 < r_{N+n} - r_{N+n-1} \le \left(\frac{\delta}{2}\right)^{n-1}(r_{N+1}-r_N) \quad \text{for each} \quad n = 1, 2, \cdots.$$

Remark 6.6.8. *If $N = 2$ we must have*

$$r_2 = b + \frac{a\,l_0\,b}{2\,(1 - a\,l_0\,b)} \le \frac{a\,l\,b + \delta}{a\,l + \delta\,a\,l_0},$$

which is (6.6.13). When $N > 2$ we do not have closed form inequalities (solved for n) anymore given by

$$c_0\,\eta \le c_1,$$

where c_0 and c_1 may depend on l_0 and l, see e.g. (6.6.7) or (6.6.13). However, the corresponding inequalities can also be checked out, since only computations involving b, l_0, and l are carried out (see also [16]). Clearly, the sufficient convergence conditions of the form (6.6.17) become weaker as N increases.

Remark 6.6.9. *In [14], [16], tighter upper bounds on the limit points of majorizing sequence $\{r_n\}$, $\{s_n\}$, $\{u_k\}$ than [6, 8, 32] are given. Indeed, we have that*

$$r^\star = \lim_{n\to\infty} r_n \le r_3 = \left(1 + \frac{a\,l_0\,b}{(2-\delta)\,(1 - a\,l_0\,b)}\right) b.$$

Note that

$$r_3 \begin{cases} \le r_2 & \text{if} \quad l_0 \le l \\ < r_2 & \text{if} \quad l_0 < l \end{cases} \quad \text{and} \quad r_3 \begin{cases} \le r_1 & \text{if} \quad l_0 \le l \\ < r_1 & \text{if} \quad l_0 < l \end{cases}$$

where

$$r_2 = \frac{2\,b}{2-\delta} \quad \text{and} \quad r_1 = 2\,b.$$

Moreover, r_2 can be smaller than $s^\star$ for sufficiently small l_0. We have also that

$$h \le \frac{1}{2} \Longrightarrow h_1 \le \frac{1}{2} \Longrightarrow h_2 \le \frac{1}{2},$$

but not necessarily vice versa unless if $l_0 = l$. Moreover, we have that

$$\frac{h_1}{h} \longrightarrow \frac{1}{4}, \quad \frac{h_2}{h} \longrightarrow 0 \quad \text{and} \quad \frac{h_2}{h_1} \longrightarrow 0 \quad \text{as} \quad \frac{l_0}{l} \longrightarrow 0.$$

Example 6.6.10. *We consider a simple example to test the "h" conditions in one dimension. Let $X = \mathbb{R}$, $x_0 = 1$, $\Omega = [d, 2-d]$, $d \in [0, .5)$. Define function F on Ω by*

$$F(x) = x^3 - d. \tag{6.6.21}$$

We get that

$$b = \frac{1}{3}\,(1-d) \quad \text{and} \quad l = 2\,(2-d).$$

Kantorovich condition (6.1.1) is given by

$$h = \frac{2}{3}\,(1-d)\,(2-d) > .5 \quad \text{for all} \quad d \in (0, .5).$$

Hence, there is no guarantee that Newton's method starting at $x_0 = 1$ *converges to* $x^\star$. *However, one can easily see that if for example* $d = .49$, *Newton's method converges to* $x^\star = \sqrt[3]{.49}$. *In view of (6.6.21), we deduce the center-Lipschitz condition*

$$l_0 = 3 - d < l = 2(2-d) \quad \text{for all} \quad d \in (0, .5). \tag{6.6.22}$$

We consider the "h" conditions of Remark 6.6.9. Then, we obtain that

$$h_1 = \frac{1}{12}(8 - 3d + (5d^2 - 24d + 28)^{1/2})(1-d) \le .5 \quad \text{for all} \quad d \in [.450339002, .5)$$

and

$$h_2 = \frac{1}{24}(1-d)(12 - 4d + (84 - 58d + 10d^2)^{1/2} + (12 - 10d + 2d^2)^{1/2}) \le .5$$
$$\text{for all} \quad d \in [.4271907643, .5).$$

In Fig. 6.6.1, we compare the "h" conditions for $d \in (0, .999)$.

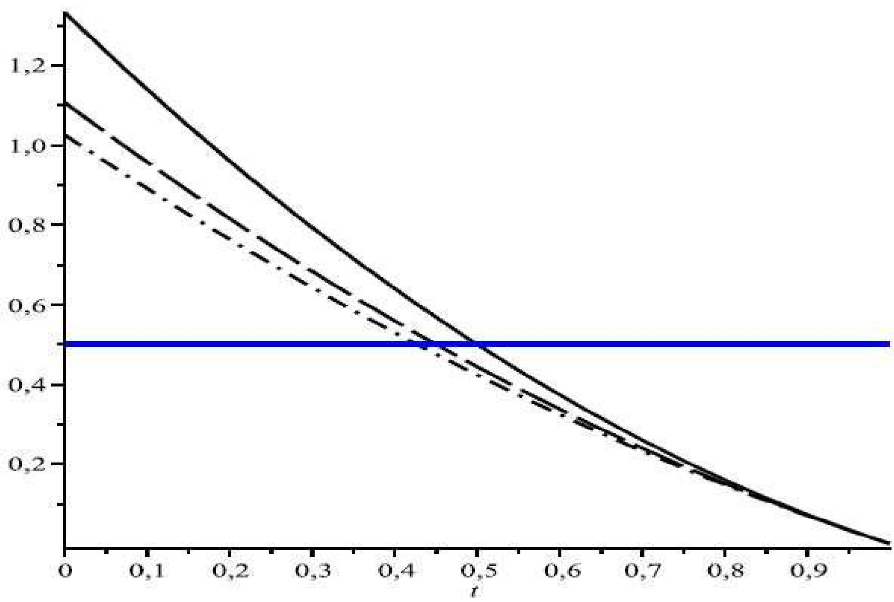


Figure 6.6.1. Functions h, h_1, h_2 (from top to bottom) with respect to d in interval $(0, .999)$, respectively. The horizontal blue line is of equation $y = .5$.

References

[1] Absil, P.A. Mahony, R., Sepulchre, R., *Optimization Algorithms on Matrix Manifolds*, Princeton University Press, Princeton NJ, 2008.

[2] Adler, R.L., Dedieu, J.P., Margulies, J.Y., Martens, M., Shub,M., Newton's method on Riemannian manifolds and a geometric model for the human spine, *IMA J. Numer. Anal.*, 22 (2002), 359–390.

[3] Alvarez, F., Bolte, J., Munier, J., A unifying local convergence result for Newton's method in Riemannian manifolds, *Foundations Comput. Math.*, 8 (2008), 197–226.

[4] Amat, S., Busquier, S., Third-order iterative methods under Kantorovich conditions, *J. Math. Anal. Appl.*, 336 (2007), 243–261.

[5] Amat, S., Busquier, S., Gutiérrez, J. M., Third-order iterative methods with applications to Hammerstein equations: A unified approach, *J. App. Math. Comp.*, 235 (2011), 2936–2943.

[6] Argyros, I.K., A unifying local-semilocal convergence analysis and applications for two-point Newton-like methods in Banach space, *J. Math. Anal. Appl.*, 298 (2004), 374–397.

[7] Argyros, I.K., An improved unifying convergence analysis of Newton's method in Riemannian manifolds, *J. App. Math. Comp.*, 25 (2007), 345–351.

[8] Argyros, I.K., Approximating solutions of equations using Newton's method with a modified Newton's method iterate as a starting point, *Rev. Anal. Numér. Théor. Approx.*, 36 (2007), 123–138.

[9] Argyros, I.K., Chebysheff-Halley like methods in Banach spaces, *Korean J. Comp. Appl. Math.*, 4 (1997), 83–107.

[10] Argyros, I.K., Improved error bounds for a Chebysheff-Halley-type method, *Acta Math. Hungar.*, 84 (1999), 209–219.

[11] Argyros, I.K., Improving the order and rates of convergence for the super-Halley method in Banach spaces, *Korean J. Comput. Appl. Math.*, 5 (1998), 465–474.

[12] Argyros, I.K., *Convergence and applications of Newton–type iterations*, Springer–Verlag Publ., New–York, 2008.

[13] Argyros, I.K., A semilocal convergence analysis for directional Newton methods, *Math. Comp.*, 80 (2011), 327–343.

[14] Argyros, I.K., Cho, Y.J., Hilout, S., *Numerical methods for equations and its applications*, CRC Press/Taylor and Francis Publ., New York, 2012.

[15] Argyros, I.K., Hilout, S., Newton's method for approximating zeros of vector fields on Riemannian manifolds, *J. App. Math. Comp.*, 29 (2009), 417–427.

[16] Argyros, I.K., Hilout, S., Weaker conditions for the convergence of Newton's method, *J. Complexity*, 28 (2012), 364–387.

[17] Argyros, I.K., Ren, H., On the semilocal convergence of Halley's method under a center-Lipschitz condition on the second Fréchet derivative, *App. Math. Comp.*, 218 (2012), 11488–11495.

[18] Averbuh, V.I., Smoljanov, O.G., Differentiation theory in linear topological spaces, *Uspehi Mat. Nauk* 6, 201-260 Russian Math. Surveys 6 (1967), 201–258.

[19] Chun, C., Stanica, P., Neta, B., Third-order family of methods in Banach spaces, *Comput. Math. Appl.*, 61 (2011), 1665–1675.

[20] Dedieu, J.P., Priouret, P., Malajovich, G., Newton's method on Riemannian manifolds: convariant alpha theory, *IMA J. Numer. Anal.*, 23 (2003), 395–419.

[21] Dedieu, J.P., Nowicki, D., Symplectic methods for the approximation of the exponential map and the Newton iteration on Riemannian submanifolds, *J. Complexity*, 21 (2005), 487–501.

[22] Do Carmo, M., *Riemannian geometry*, Birkhäuser, Boston, 1992.

[23] Ezquerro, J.A., Hernández, M.A., New Kantorovich-type conditions for Halley's method, *Appl. Numer. Anal. Comput. Math.*, 2 (2005), 70–77.

[24] Ezquerro, J.A., Gutiérrez, J. M., Hernández, M.A., Salanova, M.A., Chebyshev-like methods and quadratic equations, *Rev. Anal. Numér. Théor. Approx.*, 28 (2000), 23–35.

[25] Ezquerro, J.A., A modification of the Chebyshev method, *IMA J. Numer. Anal.*, 17 (1997) 511–525.

[26] Ezquerro, J.A., Hernández, M.A., A super-Halley type approximation in Banach spaces, *Approx. Theory Appl.*, 17 (2001), 14–25.

[27] Ferreira, O., Svaiter, B., Kantorovich's Theorem on Newton's Method in Riemannian Manifolds, *J. Complexity*, 18 (2002), 304–329.

[28] Gabay, D., Minimizing a differentiable function over a differential manifold, *J. Optim. Theory Appl.*, 37 (1982), 177–219.

[29] Groisser, D., Newton's method, zeros of vector fields, and the Riemannian center of mass, *Adv. Appl. Math.*, 33 (2004), 95–135.

[30] Hernández, M.A., Romero, N., General study of iterative processes of R-order at least three under weak convergence conditions, *J. Optim. Theory Appl.*, 133 (2007), 163–177.

[31] Hernández, M.A., Romero, N., On a characterization of some Newton-like methods of R-order at least three, *J. Comput. Appl. Math.*, 183 (2005), 53–66.

[32] Kantorovich, L.V., Akilov, G.P., *Functional Analysis in Normed Spaces*, Pergamon, Oxford, 1964.

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

[34] Kress, R., *Numerical Analysis*, Springer-Verlag, New York, 1998.

[35] Li, C., Wang, J., Newton's method on Riemannian manifolds: Smale's point estimate theory under the γ-condition, *IMA J. Numer*. Anal., 26 (2006), 228–251.

[36] Li, C., Wang, J., Newton's method for sections on Riemannian manifolds: Generalized covariant α-theory, *J. Complexity*, 25 (2009), 128–151.

[37] Manton, J.H., Optimization algorithms exploiting unitary constraints, *IEEE Trans. Signal Process.*, 50 (2002), 635–650.

[38] Neta, B., A new iterative method for the solution of systems of nonlinear equations. Approximation theory and applications (Proc. Workshop, Technion Israel Inst. Tech., Haifa, 1980), Academic Press, New York-London, (1981), 249–263.

[39] Parida, P.K., Gupta, D.K., *Recurrence relations for a Newton-like method in Banach spaces*, 206 (2007), 873–887.

[40] Parida, P.K., Gupta, D.K., Recurrence relations for semilocal convergence of a Newton-like method in Banach spaces, *J. Math. Anal. Appl.* 345 (2008), 350–361.

[41] Parida, P.K., Gupta, D.K., Semilocal convergence of a family of third-order methods in Banach spaces under Hölder continuous second derivative, *Nonlinear Anal.* 69 (2008), 4163–4173.

[42] Romero, N., PhD Thesis. Familias paramétricas de procesos iterativos de alto orden de convergencia. http://dialnet.unirioja.es/

[43] Shamanskii, V.E., A modification of Newton's method, *Ukrain. Mat. Zh.*, 19 (1967), 133–138.

[44] Spivak, M., *A comprehensive introduction to differential geometry*, Vol. I, third ed., Publish or Perish Inc., Houston, Texas., 2005.

[45] Spivak, M., *A comprehensive introduction to differential geometry*, Vol. II, third ed., Publish or Perish Inc., Houston, Texas, 2005.

[46] Traub, J.F., *Iterative methods for the solution of equations*, Prentice Hall, Englewood Cliffs, N. J., 1964.

[47] Udriste, C., Convex functions and optimization methods on Riemannian manifolds, *Mathematics and its Applications*, 297, Kluwer Academic Publishers Group, Dordrecht, 1994.

[48] Wang, J.H., Convergence of Newton's method for sections on Riemannian manifolds, *J. Optim. Theory Appl.*, 148 (2011), 125–145.

[49] Zhang, L.H., Riemannian Newton method for the multivariate eigenvalue problem, *SIAM J. Matrix Anal. Appl.*, 31 (2010), 2972–2996.

Chapter 7

Improved Local Convergence Analysis of Inexact Gauss-Newton Like Methods

7.1. Introduction

Let $\mathcal{X}$ and Y be Banach spaces. Let $\mathcal{D} \subseteq \mathcal{X}$ be open set and $F : \mathcal{D} \longrightarrow \mathcal{Y}$ be continuously differentiable. In this chapter we are concerned with the problem of approximating a locally unique solution $x^\star$ of nonlinear least squares problem

$$\min_{x \in \mathcal{D}} \| F(x) \|^2 . \tag{7.1.1}$$

A solution $x^\star \in \mathcal{D}$ of (7.1.1) is also called a least squares solution of the equation $F(x) = 0$.

Many problems from computational sciences and other disciplines can be brought in a form similar to equation (7.1.1) using mathematical modelling [8], [11]. For example in data fitting, we have $\mathcal{X} = \mathbb{R}^i$, $\mathcal{Y} = \mathbb{R}^j$, i is the number of parameters and j is the number of observations [23].

The solution of (7.1.1) can rarely be found in closed form. That is why the solution methods for these equations are usually iterative. In particular, the practice of numerical analysis for finding such solutions is essentially connected to Newton-type methods [8]. The study about convergence matter of iterative procedures is usually centered on two types: semilocal and local convergence analysis. The semilocal convergence matter is, based on the information around an initial point, to give criteria ensuring the convergence of iterative procedures; while the local one is, based on the information around a solution, to find estimates of the radii of convergence balls. A plethora of sufficient conditions for the local as well as the semilocal convergence of Newton-type methods as well as an error analysis for such methods can be found in [1]–[47].

In the present chapter we use the inexact Gauss-Newton like method

$$x_{n+1} = x_n + s_n, \quad \mathcal{B}(x_n)\, s_n = -F'(x_n)^\star F(x_n) + r_n \quad \text{for each} \quad n = 0, 1, \cdots, \tag{7.1.2}$$

where $x_0 \in \mathcal{D}$ is an initial point to generate a sequence $\{x_n\}$ approximating $x^\star$. Here, $A^\star$ denotes the adjoint of the operator A, $\mathcal{B}(x) \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ the space of bounded linear operators from $\mathcal{X}$ into $\mathcal{Y}$, is an approximation of the derivative $F'(x)^\star F'(x)$ ($x \in \mathcal{D}$); r_n is the residual tolerance and the preconditioning invertible matrix $\mathcal{P}$ for the linear systems defining the step s_n satisfy

$$\| \mathcal{P}_n r_n \| \leq \theta_n \| \mathcal{P}_n F'(x_n)^\star F(x_n) \| \quad \text{for each} \quad n = 0, 1, \cdots \tag{7.1.3}$$

If $\theta_n = 0$ for each $n = 0, 1, \cdots$, the inexact Gauss-Newton method reduces to Gauss-Newton method. If $x^\star$ is a solution of (7.1.1), $F(x^\star) = 0$ and $F'(x^\star)$ is invertible, then the theories of Gauss-Newton methods merge into those of Newton method. A survey of convergence results under various Lipschitz-type conditions for Gauss-Newton-type methods can be found in [8] (see also [5]–[15], [17]–[40]). The convergence of these methods requires among other hypotheses that F' satisfies a Lipschitz condition or F'' is bounded in $\mathcal{D}$. Several authors have relaxed these hypotheses [9]–[15]. In particular, Ferreira et al. [24]–[29] have used the majorant condition in the local as well as semilocal convergence of Newton-type method. Argyros and Hilout [12]–[16] have also used the majorant condition to provide a tighter convergence analysis and weaker convergence criteria for Newton-type method. The local convergence of inexact Gauss-Newton method was examined by Ferreira et al. [28] using the majorant condition. It was shown that this condition is better that Wang's condition [36], [47] in some sence. A certain relationship between the majorant function and operator F was established that unifies two previously unrelated results pertaining to inexact Gauss-Newton methods, which are the result for analytical functions and the one for operators with Lipschitz derivative.

In the present chapter, we are motivated by the elegant work in [28] and optimization considerations. Using more precise majorant condition and functions, we provide a new local convergence analysis for inexact Gauss-Newton-like methods under the same computational cost and the following advantages: larger radius of convergence; tighter error estimates on the distances $\| x_n - x^\star \|$ for each $n = 0, 1, \cdots$ and a clearer relationship between the majorant function and the associated least squares problems (7.1.1). These advantages are obtained because we use a center-type majorant condition (see (7.3.1)) for the computation of inverses involved which is more precise that the majorant condition used in [28]. Moreover, these advantages are obtained under the same computational cost, since as we will see in section 7.3. and section 7.4., the computation of the majorant function requires the computation of the center-majorant function. Furthemore, these advantages are very important in computational mathematics, since we have a wider choice of initial guesses x_0 and fewer computations to obtain a desired error tolerance on the distances $\| x_n - x^\star \|$ for each $n = 0, 1, \cdots$.

The chapter is organized as follows. In order to make the chapter as self contained as possible, we provide the necessary background in section 7.2.. Section 7.3. contains the local convergence analysis of inexact Gauss-Newton-like methods. Some proofs are abbreviated to avoid repetitions with the corresponding in [28]. Special cases and applications are given in the concluding section 7.4..

7.2. Background

Let $U(x,r)$ and $\overline{U}(x,r)$ stand, respectively, for the open and closed ball in $\mathcal{X}$ with center $x \in \mathcal{D}$ and radius $r > 0$. Let $A : \mathcal{X} \longrightarrow \mathcal{Y}$ be continuous linear and injective with closed image, the Moore-Penrose inverse [8], [11], [34] $A^+ : \mathcal{Y} \longrightarrow \mathcal{X}$ is defined by $A^+ = (A^\star A)^{-1}A^\star$. I denotes the identity operator on $\mathcal{X}$ (or $\mathcal{Y}$).

Lemma 7.2.1. *[8, 11, 35] (Banach's Lemma) Let $A : \mathcal{X} \longrightarrow \mathcal{X}$ be a continuous linear operator. If $\| A - I \| < 1$ then $A^{-1} \in \mathcal{L}(\mathcal{X},\mathcal{X})$ and $\| A^{-1} \| \leq 1/(1 - \| A - I \|)$.*

Lemma 7.2.2. *[8, ?] Let $A, E : \mathcal{X} \longrightarrow \mathcal{Y}$ be two continuous linear operators with closed images. Suppose $B = A + E$, A is injective and $\| EA^+ \| < 1$. Then, B is injective.*

Lemma 7.2.3. *[8, 11, 35] Let $A, E : \mathcal{X} \longrightarrow \mathcal{Y}$ be two continuous linear operators with closed images. Suppose $B = A + E$ and $\| A^+ \| \| E \| < 1$. Then, the following estimates hold*

$$\| B^+ \| \leq \frac{\| A^+ \|}{1 - \| A^+ \| \| E \|} \quad \textit{and} \quad \| B^+ - A^+ \| \leq \frac{\sqrt{2} \| A^+ \|^2 \| E \|}{1 - \| A^+ \| \| E \|}.$$

Proposition 7.2.4. *[34] Let $R > 0$. Suppose $g : [0,R) \longrightarrow \mathbb{R}$ is convex. Then, the following holds*

$$D^+ g(0) = \lim_{u \to 0^+} \frac{g(u) - g(0)}{u} = \inf_{u > 0} \frac{g(u) - g(0)}{u}.$$

Proposition 7.2.5. *[34] Let $R > 0$ and $\theta \in [0,1]$. Suppose $g : [0,R) \longrightarrow \mathbb{R}$ is convex. Then, $h : (0,R) \longrightarrow \mathbb{R}$ defined by $h(t) = (g(t) - g(\theta t))/t$ is increasing.*

7.3. Local Convergence Analysis

We examine the local convergence of inexact Gauss-Newton-like method. In order for us to show the main Theorem 7.3.8, we need some auxiliary results. The proofs of some of the results are omitted, since these proofs can be found in [28] by simply replacing function f by f_0. Assume that $x \in \mathcal{D} \longrightarrow F(x)^\star F(x)$ has $x^\star$ as stationarily point. Let $R > 0$, $c = \| F(x^\star) \|$, $\beta = \| F'(x^\star)^+ \|$ and

$$\kappa = \sup\{t \in [0,R) : U(x^\star, t) \subseteq \mathcal{D}\}.$$

Suppose that $F'(x^\star)^\star F(x^\star) = 0$, $F'(x^\star)$ is injective and there exist functions $f_0, f : [0,R) \longrightarrow (-\infty, +\infty)$ continuously differentiable, such that the following assumptions hold

($\mathcal{H}_0$)

$$\| F'(x) - F'(x^\star) \| \leq f_0'(\| x - x^\star \|) - f_0'(0), \tag{7.3.1}$$

$$\| F'(x) - F'(x^\star + \tau(x - x^\star)) \| \leq f'(\| x - x^\star \|) - f'(\tau \| x - x^\star \|), \tag{7.3.2}$$

for all $x \in U(x^\star, \kappa)$ and $\tau \in [0,1]$;

($\mathcal{H}_1$) $f_0(0) = f(0) = 0$ and $f_0'(0) = f'(0) = -1$;

($\mathcal{H}_2$) f_0', f' are strictly increasing,

$$f_0(t) \leq f(t) \quad \text{and} \quad f_0'(t) \leq f'(t) \quad \text{for each} \quad t \in [0,R);$$

($\mathcal{H}_3$)

$$\alpha_0 = \sqrt{2} c \beta^2 D^+ f_0'(0) < 1.$$

Let

$$0 \le \vartheta < 1, \quad 0 \le \omega_2 < \omega_1 \quad \text{such that} \quad \omega_1 (\alpha_0 + \alpha_0 \vartheta + \vartheta) + \omega_2 < 1, \tag{7.3.3}$$

where α_0 is defined in $(\mathcal{H}_3)$. Define parameters ν_0, ρ_0 and r^0 by

$$\nu_0 := \sup\{t \in [0,R) : \beta (f_0'(t)+1) < 1\} \tag{7.3.4}$$

$$\rho_0 := \sup\{t \in [0,\nu_0) : (1+\vartheta)\,\omega_1\,\beta \frac{t f'(t) - f(t) + \sqrt{2} c \beta (f_0'(t)+1)}{t(1-\beta(f_0'(t)+1))} + \omega_1 \vartheta + \omega_2 < 1\} \tag{7.3.5}$$

and

$$r^0 := \min\{\kappa, \rho_0\}. \tag{7.3.6}$$

We provide the following auxiliary lemmas.

Lemma 7.3.1. *Suppose that $(\mathcal{H}_0)$–$(\mathcal{H}_3)$ hold. Then, the constant ν_0 defined by (7.3.4) is positive and $\beta(f_0'(t)+1) < 1$ for each $t \in (0,\nu_0)$.*

Lemma 7.3.2. *Suppose that $(\mathcal{H}_0)$–$(\mathcal{H}_3)$ hold. Then, the following real functions h_i ($i = 1,2,3$) defined on $(0,R)$ by*

$$h_1(t) = \frac{1}{1-\beta(f_0'(0)+1)}, \quad h_2(t) = \frac{t f'(t) - f(t)}{t^2} \quad \text{and} \quad h_3(t) = \frac{f_0'(t)+1}{t}$$

are increasing. Note also that $h_2 h_1$ and $h_3 h_1$ are increasing on $(0,R)$.

Lemma 7.3.3. *Suppose that $(\mathcal{H}_0)$–$(\mathcal{H}_3)$ hold. Then, the constant ρ_0 defined by (7.3.5) is positive and the following holds for each $t \in (0,\rho_0)$:*

$$(1+\vartheta)\,\omega_1\,\beta \frac{t f'(t) - f(t) + \sqrt{2} c \beta (f_0'(t)+1)}{t(1-\beta(f_0'(t)+1))} + \omega_1 \vartheta + \omega_2 < 1,$$

where ϑ, ω_1 and ω_2 are defined in (7.3.3).

Proof. Using $(\mathcal{H}_1)$, we have that

$$\begin{aligned}\lim_{t\to 0} \frac{t f'(t) - f(t)}{t(1-\beta(f_0'(t)+1))} &= \lim_{t\to 0} \frac{f'(t) - (f(t)-f(0))/t}{1-\beta(f'(t)+1)} \frac{1-\beta(f'(t)+1)}{1-\beta(f_0'(t)+1)} \\ &= \frac{1-\beta(f'(0)+1)}{1-\beta(f_0'(0)+1)} \lim_{t\to 0} \frac{f'(t) - (f(t)-f(0))/t}{1-\beta(f'(t)+1)} = 0.\end{aligned}$$

By the convexity of f' and f_0' and Proposition 7.2.4, we get that

$$\lim_{t\to 0} \frac{f'(t)+1}{t(1-\beta(f_0'(t)+1))} = \lim_{t\to 0} \frac{(f_0'(t) - f_0'(0))/t}{1-\beta(f_0'(t)+1)} = D^+ f_0'(t).$$

We deduce that

$$\begin{aligned}&\lim_{t\to 0} (1+\vartheta)\,\omega_1\,\beta \frac{t f'(t) - f(t) + \sqrt{2} c \beta (f_0'(t)+1)}{t(1-\beta(f_0'(t)+1))} + \omega_1 \vartheta + \omega_2 \\ &= (1+\vartheta)\,\omega_1\,\alpha_0 + \omega_1 \vartheta + \omega_2.\end{aligned}$$

By (7.3.3), we have that $\omega_1(\alpha_0+\alpha_0\vartheta+\vartheta)+\omega_2<1$. Then, there exists δ_0 such that

$$(1+\vartheta)\,\omega_1\,\beta\,\frac{t\,f'(t)-f(t)+\sqrt{2}\,c\,\beta\,(f_0'(t)+1)}{t\,(1-\beta\,(f_0'(t)+1))}+\omega_1\,\vartheta+\omega_2<1 \quad \text{for each} \quad t\in(0,\delta_0).$$

The definition of ρ_0 gives that $\delta_0\le\rho_0$. The proof of Lemma 7.3.3 is complete. □

Lemma 7.3.4. *Suppose that* ($\mathcal{H}_0$)–($\mathcal{H}_3$) *hold. Then, for each* $x\in\mathcal{D}$ *such that* $x\in U(x^\star,\min\{\nu_0,\kappa\})$, $F'(x)^\star F'(x)$ *is invertible and the following estimates hold*

$$\| F'(x)^+ \|\le\frac{\beta}{1-\beta\,(f_0'(\| x-x^\star \|)+1)}$$

and

$$\| F'(x)^+-F'(x^\star)^+ \|\le\frac{\sqrt{2}\,\beta^2\,(f_0'(\| x-x^\star \|)+1)}{1-\beta\,(f_0'(\| x-x^\star \|)+1)}.$$

In particular, $F'(x)^\star F'(x)$ *is invertible in* $U(x^\star,r^0)$.

Proof. Since $x\in\mathcal{D}$ such that $x\in U(x^\star,\min\{\nu_0,\kappa\})$, then $\| x-x^\star \|\le\nu_0$. By Lemma 7.3.1, (7.3.1) and the definition of β, we have that

$$\| F'(x^\star)^+ \|\,\| F'(x)-F'(x^\star) \|\le\beta\,(f_0'(\| x-x^\star \|)-f_0'(0))<1.$$

Consider operators $A=F'(x^\star)$, $B=F'(x)$ and $E=B-A$. Hence, we have that $\| E\,A^+ \|\le\| E \|\,\| A^+ \|<1$. Then, we deduce the desired result by Lemmas 7.2.2 and Lemma 7.2.3. That completes the proof of Lemma 7.3.4. ⊠

Newton's iteration at a point is a zero of the linearization of F at such a point. Hence, we shall study the linearization error at a point in $\mathcal{D}$:

$$E_F(x,y):=F(y)-(F(x)+F'(x)\,(x-y)) \quad \text{for each} \quad x,y\in\mathcal{D}. \tag{7.3.7}$$

We shall bound this error by the error in linearization of the majorant function f:

$$e_f(t,u):=f(u)-(f(t)+f'(t)\,(u-t)) \quad \text{for each} \quad t,u\in[0,R). \tag{7.3.8}$$

Define also the Gauss-Newton step to the operator F by

$$S_F(x)=-F'(x)^+F(x) \quad \text{for each} \quad x\in\mathcal{D}. \tag{7.3.9}$$

Lemma 7.3.5. *Suppose that* ($\mathcal{H}_0$)–($\mathcal{H}_3$) *hold. If* $\| x^\star-x \|<\kappa$, *then the following assertion holds*

$$\| E_F(x,x^\star) \|\le e_f(\| x-x^\star \|,0).$$

Lemma 7.3.6. *Suppose that* ($\mathcal{H}_0$)–($\mathcal{H}_3$) *hold. Then, for each* $x\in\mathcal{D}$ *such that* $x\in U(x^\star,\min\{\nu_0,\kappa\})$, *the following estimate holds*

$$\| S_F(x) \|\le\frac{\beta\,e_f(\| x-x^\star \|,0)+\sqrt{2}\,c\,\beta^2\,(f_0'(\| x-x^\star \|)+1)}{1-\beta\,(f_0'(\| x-x^\star \|)+1)}+\| x-x^\star \|.$$

Proof. Let $x \in \mathcal{D}$ such that $x \in U(x^\star, \min\{\nu_0, \kappa\})$. Using (7.3.7) and (7.3.9), we have that

$$\begin{aligned}
&\| S_F(x) \| \\
&= \| F'(x)^+ (F(x^\star) - (F(x) - F'(x)(x^\star - x))) - (F'(x)^+ - F'(x^\star)^+) F(x^\star) + (x^\star - x) \| \\
&\leq \| F'(x)^+ \| \| E_F(x, x^\star) \| + \| F'(x)^+ - F'(x^\star)^+ \| \| F(x^\star) \| + \| x^\star - x \| .
\end{aligned}$$

Then, we deduce the desired result by Lemmas 7.3.4 and Lemma 7.3.5. That completes the proof of Lemma 7.3.6. ⊠

Lemma 7.3.7. *Let parameters ϑ, ω_1 and ω_2 defined by (7.3.3). Let ν_0, ρ_0 and r^0 as defined in (7.3.4), (7.3.5) and (7.3.6), respectively. Suppose that $(\mathcal{H}_0)$–$(\mathcal{H}_3)$ hold. For each $x \in U(x^\star, r^0) \setminus \{x^\star\}$, define*

$$x_+ = x + s, \quad \mathcal{B}(x)\, s = -F'(x)^\star F(x) + r, \tag{7.3.10}$$

where $\mathcal{B}(x)$ is an invertible approximation of $F'(x)^\star F(x)$ satisfying

$$\| \mathcal{B}(x)^{-1} F'(x)^\star F'(x) \| \leq \omega_1, \quad \| \mathcal{B}(x)^{-1} F'(x)^\star F'(x) - I \| \leq \omega_2. \tag{7.3.11}$$

Suppose also that the forcing term θ and the residuals r (as defined in (7.1.3)) satisfy

$$\| \mathcal{P} r \| \leq \theta \, \| \mathcal{P} F'(x)^\star F(x) \| \quad \text{and} \quad \theta \,\mathrm{cond}\,(\mathcal{P} F'(x)^\star F'(x)) \leq \vartheta. \tag{7.3.12}$$

Then, x_+ is well defined and the following estimate holds

$$\begin{aligned}
&\| x_+ - x^\star \| \leq (1+\vartheta)\, \omega_1 \, \beta \frac{f'(\| x^\star - x \|)\, \| x^\star - x \| - f(\| x^\star - x \|)}{\| x^\star - x \|^2 \, (1 - \beta (f_0'(\| x^\star - x \|) + 1))} \, \| x^\star - x \|^2 + \\
&\left(\frac{(1+\vartheta)\, \omega_1 \sqrt{2} c \beta^2 \, (f_0'(\| x - x^\star \|) + 1)}{\| x^\star - x \| \, (1 - \beta (f_0'(\| x - x^\star \|) + 1))} + \omega_1 \vartheta + \omega_2 \right) \| x^\star - x \| .
\end{aligned} \tag{7.3.13}$$

In particular, $\| x_+ - x^\star \| < \| x^\star - x \|$.

Proof. By Lemma 7.3.4 and since $x \in U(x^\star, r^0)$, we have that $F'(x)^\star F'(x)$ is invertible. In view of (7.1.2) and (7.3.10), we obtain the identity

$$\begin{aligned}
x_+ - x^\star &= x - x^\star - \mathcal{B}(x)^{-1} F'(x)^\star (F(x) - F(x^\star)) + \mathcal{B}(x)^{-1} r + \\
&\quad \mathcal{B}(x)^{-1} F'(x)^\star F'(x) (F'(x^\star)^+ F(x^\star) - F'(x)^+ F(x^\star)) \\
&= \mathcal{B}(x)^{-1} F'(x)^\star F'(x) F'(x)^+ (F(x^\star) - (F(x) + F'(x)(x^\star - x))) + \\
&\quad \mathcal{B}(x)^{-1} r + \mathcal{B}(x)^{-1} (F'(x)^\star F'(x) - \mathcal{B}(x))(x - x^\star) + \\
&\quad \mathcal{B}(x)^{-1} F'(x)^\star F'(x) (F'(x^\star)^+ F(x^\star) - F'(x)^+ F(x^\star)).
\end{aligned} \tag{7.3.14}$$

Using (7.3.7), (7.3.9), (7.3.11), (7.3.12) and (7.3.14), we get that

$$\begin{aligned}
\| x_+ - x^\star \| &\leq \omega_1 \, \| F'(x)^+ \| \| E_F(x, x^\star) \| + \| \mathcal{B}(x)^{-1} r \| + \omega_2 \, \| x^\star - x \| + \\
&\quad \omega_1 \, \| F'(x)^+ - F'(x^\star)^+ \| \| F(x^\star) \| \\
&\leq \omega_1 \, \| F'(x)^+ \| \| E_F(x, x^\star) \| + \omega_1 \vartheta \, \| S_F(x) \| + \\
&\quad \omega_2 \, \| x^\star - x \| + \omega_1 \, \| F'(x)^+ - F'(x^\star)^+ \| \| F(x^\star) \| .
\end{aligned} \tag{7.3.15}$$

Using (7.3.8), (7.3.15) and Lemmas 7.3.4–7.3.6, we deduce that

$$\begin{aligned}
&\| x_+ - x^\star \| \leq (1+\vartheta)\,\beta\,\omega_1 \frac{e_f(\| x-x^\star \|,0)+\sqrt{2}c\,\beta\,(f_0'(\| x-x^\star \|)+1)}{1-\beta\,(f_0'(\| x-x^\star \|)+1)}+\\
&\omega_1\,\vartheta\,\| x-x^\star \| + \omega_2\,\| x^\star - x \|\\
&\leq (1+\vartheta)\,\beta\,\omega_1 \frac{f'(\| x-x^\star \|)\,\| x-x^\star \| - f(\| x-x^\star \|)+\sqrt{2}c\,\beta\,(f_0'(\| x-x^\star \|)+1)}{1-\beta\,(f_0'(\| x-x^\star \|)+1)}\\
&+\omega_1\,\vartheta\,\| x-x^\star \| + \omega_2\,\| x^\star - x \|\,.
\end{aligned} \tag{7.3.16}$$

Hence, (7.3.13) holds. Note that if we factorize by $\| x^\star - x \|$ in the right term in (7.3.13), then, we deduce that $\| x_+ - x^\star \| < \| x^\star - x \|$. The proof of Lemma 7.3.7 is complete. ⊠

Next, we provide the main local convergence result for inexact Gauss-Newton-like method.

Theorem 7.3.8. *Let $F : \mathcal{D} \subseteq \mathcal{X} \longrightarrow \mathcal{Y}$ be a continuously differentiable operator. Let parameters ϑ, ω_1 and ω_2 defined by (7.3.3). Let ν_0, ρ_0 and r^0 as defined in (7.3.4), (7.3.5) and (7.3.6), respectively. Suppose that ($\mathcal{H}_0$)–($\mathcal{H}_3$) hold. Then, sequence $\{x_n\}$ generated by (7.1.2), starting at $x_0 \in U(x^\star, r^0) \setminus \{x^\star\}$ for the the forcing term θ_n, the residual r_n and the invertible preconditioning matrix $\mathcal{P}_n$ satisfying the following estimates for each $n = 0, 1, \cdots$:*

$$\| \mathcal{P}_n r_n \| \leq \theta_n\,\| \mathcal{P}_n F'(x_n)^\star F(x_n) \|, \quad 0 \leq \theta_n \operatorname{cond}(\mathcal{P}_n F'(x_n)^\star F'(x_n)) \leq \vartheta,$$

$$\| \mathcal{B}(x_n)^{-1} F'(x_n)^\star F'(x_n) \| \leq \omega_1 \quad \text{and} \quad \| \mathcal{B}(x_n)^{-1} F'(x_n)^\star F'(x_n) - I \| \leq \omega_2$$

is well defined, remains in $U(x^\star, r^0)$ for all $n \geq 0$ and converges to $x^\star$. Moreover, the following estimate holds for each $n = 0, 1, \cdots$

$$\| x_{n+1} - x^\star \| \leq \Xi_n\,\| x_n - x^\star \|, \tag{7.3.17}$$

where

$$\begin{aligned}
\Xi_n \;=\;& (1+\vartheta)\,\omega_1\,\beta \frac{f'(\| x^\star - x_0 \|)\,\| x^\star - x_0 \| - f(\| x^\star - x_0 \|)}{\| x^\star - x_0 \|^2\,(1-\beta\,(f_0'(\| x^\star - x_0 \|)+1))}\,\| x^\star - x_n \| +\\
& \frac{(1+\vartheta)\,\omega_1\,\sqrt{2}c\,\beta^2\,(f_0'(\| x_0 - x^\star \|)+1)}{\| x^\star - x_0 \|\,(1-\beta\,(f_0'(\| x_0 - x^\star \|)+1))} + \omega_1\,\vartheta + \omega_2.
\end{aligned}$$

Proof. By induction argument, Lemmas 7.3.4 and 7.3.7, $\{x_n\}$ starting at $x_0 \in U(x^\star, r^0) \setminus \{x^\star\}$ is well defined in $U(x^\star, r^0)$. By letting $x_+ = x_{n+1}$, $x = x_n$, $r = r_n$, $\mathcal{P} = \mathcal{P}_n$, $\theta = \theta_n$ and $\mathcal{P} = \mathcal{P}_n$ in (7.3.10)–(7.3.12), we get that

$$\begin{aligned}
&\| x_{n+1} - x^\star \| \leq\\
&(1+\vartheta)\,\omega_1\,\beta \frac{f'(\| x^\star - x_n \|)\,\| x^\star - x_n \| - f(\| x^\star - x_n \|)}{\| x^\star - x_n \|^2\,(1-\beta\,(f_0'(\| x^\star - x_n \|)+1))}\,\| x^\star - x_n \|^2 +\\
&\left(\frac{(1+\vartheta)\,\omega_1\,\sqrt{2}c\,\beta^2\,(f_0'(\| x_n - x^\star \|)+1)}{\| x^\star - x_n \|\,(1-\beta\,(f_0'(\| x_n - x^\star \|)+1))} + \omega_1\,\vartheta + \omega_2 \right)\,\| x^\star - x_n \|\,.
\end{aligned}$$

We also have by Lemma 7.3.7 that $\| x_n - x^\star \| \leq \| x_0 - x^\star \|$ for each $n = 1, 2, \cdots$. Hence, (7.3.17) holds. Proposition 7.3.3 imply that $x_{n+1} \in U(x^\star, r^0)$ and $\lim_{n \longrightarrow \infty} x_n = x^\star$. The proof of Theorem 7.3.8 is complete. ⊠

Remark 7.3.9. *If $f(t) = f_0(t)$ for each $t \in [0,R)$, then, Theorem 7.3.8 reduces to [28, Theorem 7]. In particular, we have in this case that $\nu = \nu_0$, $\rho = \rho_0$, $\delta = \delta_0$, $\alpha = \alpha_0$, $r = r^0$ and $D^+ f(0) = D^+ f_0(0)$, where ν, ρ, δ, α, r and $D^+ f(0)$ are defined, respectively, as ν_0, ρ_0, δ_0, α_0, r^0 and $D^+ f_0(0)$ by setting $f_0(t) = f(t)$. Otherwise, i.e., if*

$$f_0(t) < f(t) \quad and \quad f_0'(t) < f'(t) \quad for\,each \quad t \in [0,R), \tag{7.3.18}$$

then, we have that

$$\nu \geq \nu_0, \quad \rho \leq \rho_0, \quad \delta \leq \delta_0, \quad \alpha \geq \alpha_0, \quad r \leq r^0 \quad and \quad D^+ f(0) \geq D^+ f_0(0). \tag{7.3.19}$$

Note that these advantages are obtained under the same computational cost, since in practice, the computation of function f requires that of f_0. Note also that the local results in [18], [19], [24]–[27] are also extended, since these are special cases of Theorem 7.3.8. In particular, if $\vartheta = 0$ (i.e., if $\theta_n = r_n = 0$ for each $n = 0, 1, \cdots$) in Theorem 7.3.8, we improve the convergence of Gauss-Newton like method under majorant condition, which for $\omega_1 = 1$ and $\omega_2 = 0$ has been obtained in [26, Theorem 7]. These results extend those the ones obtained by Chen and Li in [18], [19] given only for the the case $c = 0$. Moreover, if $c = 0$ and $F'(x^\star)$ is invertible, we extend the convergence of inexact Newton-like methods under majorant condition, which was obtained in [24, Theorem 4]. Furthemore, if $c = \vartheta = \omega_2 = 0$, $\omega_1 = 1$ and $F'(x^\star)$ is invertible in Theorem 7.3.8, we extend the convergence of Newton's method under majorant condition obtained in [24, Theorem 2.1].

In the next section, we shall show how to choose functions f_0 and f so that (7.3.18) is satisfied.

7.4. Special Case and Numerical Examples

We present two special cases of Theorem 7.3.8. The first one is based on the center-Lipschitz and Lipschitz conditions [8], [11]. The second one is based on Wang's condition [47], which generalized Smale's alpha theory for analytic functions [44].

Remark 7.4.1. *Let us define functions $f, f_0 : [0,\kappa] \longrightarrow \mathbb{R}$ by*

$$f_0(t) = \frac{L_0 t^2}{2} - t \quad and \quad f(t) = \frac{L t^2}{2} - t,$$

where L_0 and L are the center-Lipschitz and Lipschitz constants, respectively. We have that $f_0(0) = f(0) = 0$ and $f_0'(0) = f'(0) = -1$. Set also $R = 1/L$. Then, it can easily be seen that Theorem 7.3.8 specializes to the following proposition.

Proposition 7.4.2. *Let $F : \mathcal{D} \subset \mathcal{X} \longrightarrow \mathcal{Y}$ be continuously differentiable operator. Let $x^\star \in \mathcal{D}$, such that $F'(x^\star)^\star F(x^\star) = 0$, $F'(x^\star)$ is injective. Let $c = \parallel F(x^\star) \parallel$, $\beta = \parallel F'(x^\star)^+ \parallel$ and $\kappa = \sup\{t \geq 0 : U(x^\star, t) \subseteq \mathcal{D}\}$. Suppose that there exist $L_0 > 0$, $L > 0$ such that*

$$\parallel F'(x) - F'(x^\star) \parallel \leq L_0 \parallel x - x^\star \parallel \quad and \quad \parallel F'(x) - F'(y) \parallel \leq L \parallel x - y \parallel,$$

for each $x, y \in U(x^\star, \kappa)$. *Suppose that* $\alpha_0 = \sqrt{2}\, c\, \beta^2 L < 1$. *Let parameters* ϑ, ω_1 *and* ω_2 *defined by (7.3.3). Let*

$$r^0 = \min\{\kappa, \frac{2\,(1-\omega_1\,\vartheta-\omega_2) - 2\,\sqrt{2}\,c\,L_0\,\beta^2\,\omega_1\,(1+\vartheta)}{\beta\,(L\,(1+\vartheta)\,\omega_1 + 2\,L_0\,(1-\omega_1\,\vartheta-\omega_2))}\}.$$

Then, sequence $\{x_n\}$ *generated by (7.1.2) with* $r_n = 0$ *and* $\mathcal{B}(x_n) = F'(x_n)^\star F'(x_n)$, *starting at* $x_0 \in U(x^\star, r^0) \setminus \{x^\star\}$ *for the the forcing term* θ_n *and the invertible preconditioning matrix* $\mathcal{P}_n$ *satisfying the following estimates for each* $n = 0, 1, \cdots$:

$$\| \mathcal{P}_n\, r_n \| \leq \theta_n \, \| \mathcal{P}_n\, F'(x_n)^\star\, F(x_n) \|, \quad 0 \leq \theta_n \,\mathrm{cond}\,(\mathcal{P}_n F'(x_n)^\star\, F'(x_n)) \leq \vartheta,$$

$$\| \mathcal{B}(x_n)^{-1}\, F'(x_n)^\star\, F'(x_n) \| \leq \omega_1 \quad \text{and} \quad \| \mathcal{B}(x_n)^{-1}\, F'(x_n)^\star\, F'(x_n) - I \| \leq \omega_2$$

is well defined, remains in $U(x^\star, r^0)$ *for all* $n \geq 0$ *and converges to* $x^\star$. *Moreover, the following estimate holds for each* $n = 0, 1, \cdots$

$$\| x_{n+1} - x^\star \| \leq \Delta_n \, \| x_n - x^\star \|, \tag{7.4.1}$$

where

$$\Delta_n = \frac{(1+\vartheta)\,\omega_1\,\beta\,L}{2\,(1-\beta\,L_0\, \| x_0 - x^\star \|)}\, \| x^\star - x_n \| + \frac{(1+\vartheta)\,\omega_1\,\sqrt{2}\,c\,\beta^2\,L_0}{1-\beta\,L_0\, \| x_0 - x^\star \|} + \omega_1\,\vartheta + \omega_2.$$

Remark 7.4.3. *(a) If* $L_0 = L$, *Proposition 7.4.2 reduces to [28, Theorem 16]. Moreover, if* $\vartheta = 0$, *Proposition 7.4.2 reduces to [28, Corollary 17]. Furthemore, if* $c = 0$, *then, Proposition 7.4.2 reduces to [19, Corollary 6.1].*

(b) If $F(x^\star) = 0$, $F'(x^\star)^+ = F'(x^\star)^{-1}$ *and* $L_0 < L$, *then Theorem 7.3.8 improves the corresponding results on inexact Newton-like methods [30], [33], [37], [38]. In particular for Newton's method. Set* $c = \vartheta = \omega_1 = \omega_2 = 0$. *Then, we get that*

$$r^0 = \min\{\kappa, \frac{2}{\beta\,(2\,L_0 + L)}\}.$$

This radius is at least as large as the one provided by Traub [46], which is given by

$$r_0^0 = \min\{\kappa, \frac{2}{3\,\beta\,L}\}.$$

Let us provide a numerical example for this case.

Example 7.4.4. *Let* $\mathcal{X} = \mathcal{Y} = \mathcal{C}[0,1]$, *the space of continuous functions defined on* $[0,1]$ *be equipped with the max norm and* $\mathcal{D} = \overline{U}(0,1)$. *Define function* F *on* $\mathcal{D}$ *by*

$$F(h)(x) = h(x) - 5 \int_0^1 x\,\theta\,h(\theta)^3\,d\theta. \tag{7.4.2}$$

Then, we have that

$$F'(h[u])(x) = u(x) - 15 \int_0^1 x\,\theta\,h(\theta)^2\,u(\theta)\,d\theta \quad \textit{for each} \quad u \in \mathcal{D}.$$

Using Proposition 7.4.2, we see that the hypotheses hold for $x^\star(x) = 0$, where $x \in [0,1]$, $\beta = 1$, $L = 15$ and $L_0 = 7.5$. Then, we get that

$$r_0^0 = \min\{\kappa, \frac{2}{45}\} \leq \min\{\kappa, \frac{1}{15}\} = r_0.$$

Clearly, if $\min\{\kappa, \frac{2}{45}\} = \frac{2}{45}$, then we deduce in particular that $r_0^0 < r_0$.

Remark 7.4.5. *Let $\gamma \geq \gamma_0$. Let us define functions $f, f_0 : [0,\kappa] \longrightarrow \mathbb{R}$ by*

$$f_0(t) = \frac{t}{1-\gamma_0 t} - 2t \quad \text{and} \quad f(t) = \frac{t}{1-\gamma t} - 2t.$$

Then, we have that $f_0(0) = f(0) = 0$, $f_0'(0) = f'(0) = -1$. Set also $R = 1/\gamma$. Note also that

$$f_0'(t) = \frac{1}{(1-\gamma_0 t)^2} - 2, \quad f'(t) = \frac{1}{(1-\gamma t)^2} - 2,$$

$$f_0''(t) = \frac{2\gamma_0}{(1-\gamma_0 t)^3} \quad \text{and} \quad f''(t) = \frac{2\gamma}{(1-\gamma t)^3}.$$

We introduce the definition of the center γ_0-condition.

Definition 7.4.6. *Let $\gamma_0 > 0$ and let $0 < \mu \leq 1/\gamma_0$ be such that $U(x^\star,\mu) \subseteq \mathcal{D}$. The operator F is said to satisfy the center γ_0-condition at $x^\star$ on $U(x^\star,\mu)$ if*

$$\| F'(x) - F'(x^\star) \| \leq \frac{1}{(1-\gamma_0 \| x - x^\star \|)^2} - 1 \quad for\, each \quad x \in U(x^\star,\mu).$$

We also need the definition of γ-condition due to Wang [47].

Definition 7.4.7. *Let $\gamma > 0$ and let $0 < \mu \leq 1/\gamma$ be such that $U(x^\star,\mu) \subseteq \mathcal{D}$. The operator F is said to satisfy the γ-condition at $x^\star$ on $U(x^\star,\mu)$ if*

$$\| F''(x) \| \leq \frac{2\gamma}{(1-\gamma \| x - x^\star \|)^3} \quad for\, each \quad x \in U(x^\star,\mu).$$

Remark 7.4.8. *(a) Note that $\gamma_0 \leq \gamma$ holds in general and γ/γ_0 can be arbitrarily large [7]–[16].*

(b) If F is an analytic function, Smale [44] used the following choice

$$\gamma = \sup_{n \in \mathbb{N}^\star} \| \frac{F^{(n)}(x^\star)}{n!} \|^{\frac{1}{n}} < +\infty.$$

Using the above definitions and choices of functions (see Remark 7.4.5, Definitions 7.4.6 and 7.4.7), the corresponding specialization of Theorem 7.3.8 along the lines of Proposition 7.4.2 can be obtained. However, we leave this part to the interested reader. Note that clearly if $\gamma_0 = \gamma$, this result reduces to [28, Theorem 18], which in turn reduces to [19, Example 1] if $c = 0$. Otherwise (i.e., if $\gamma_0 < \gamma$), our result is an improvement.

Next, we provide an example, where $\gamma_0 < \gamma$ in the case when $F(x^\star) = 0$, $F'(x)^+ = F'(x)^{-1}$ and $c = \vartheta = \omega_1 = \omega_2 = 0$.

References

[1] Amat, S., Bermúdez, C., Busquier, S., Legaz, M.J., Plaza, S., On a family of high-order iterative methods under Kantorovich conditions and some applications, *Abstr. Appl. Anal.* 2012, Art. ID 782170, 14 pp.

[2] Amat, S., Bermúdez, C., Busquier, S., Plaza, S., On a third-order Newton-type method free of bilinear operators, *Numer. Linear Algebra Appl.*, 17 (2010), 639–653.

[3] Amat, S., Busquier, S., Gutiérrez, J.M., Geometric constructions of iterative functions to solve nonlinear equations, *J. Comput. Appl. Math.*, 157 (2003), 197–205.

[4] Amat, S., Busquier, S., Gutiérrez, J.M., Third-order iterative methods with applications to Hammerstein equations: a unified approach, *J. Comput. Appl. Math.*, 235 (2011), 2936–2943.

[5] Argyros, I.K., Forcing sequences and inexact Newton iterates in Banach space, *Appl. Math. Lett.*, 13 (2000), 69–75.

[6] Argyros, I.K., A unifying local–semilocal convergence analysis and applications for two–point Newton–like methods in Banach space, *J. Math. Anal. Appl.*, 298 (2004), 374–397.

[7] Argyros, I.K., On the semilocal convergence of the Gauss–Newton method, *Adv. Nonlinear Var. Inequal.*, 8 (2005), 93–99.

[8] Argyros, I.K., *Convergence and applications of Newton–type iterations*, Springer–Verlag Publ., New York, 2008.

[9] Argyros, I.K., On the semilocal convergence of inexact Newton methods in Banach spaces, *J. Comput. Appl. Math.*, 228 (2009), 434–443.

[10] Argyros, I.K., Cho, Y.J., Hilout, S., On the local convergence analysis of inexact Gauss–Newton–like methods, *Panamer. Math. J.*, 21 (2011), 11–18.

[11] Argyros, I.K., Cho, Y.J., Hilout, S., *Numerical methods for equations and its applications*, Science Publishers, New Hampshire, USA, 2012.

[12] Argyros, I.K., Hilout, S., On the local convergence of the Gauss-Newton method, *Punjab Univ. J. Math.*, 41 (2009), 23–33.

[13] Argyros, I.K., Hilout, S., Improved generalized differentiability conditions for Newton–like methods, *J. Complexity*, 26 (2010), 316–333.

[14] Argyros, I.K., Hilout, S., Extending the applicability of the Gauss-Newton method under average Lipschitz-type conditions, *Numer. Algorithms*, 58 (2011), 23–52.

[15] Argyros, I.K., Hilout, S., Improved local convergence of Newton's method under weak majorant condition, *J. Comput. Appl. Math.*, 236 (2012), 1892–1902.

[16] Argyros, I.K., Hilout, S., Weaker conditions for the convergence of Newton's method, *J. Complexity*, 28 (2012), 364–387.

[17] Chen, J., The convergence analysis of inexact Gauss–Newton methods for nonlinear problems, *Comput. Optim. Appl.*, 40 (2008), 97–118.

[18] Chen, J., Li, W., Convergence of Gauss–Newton's method and uniqueness of the solution, *App. Math. Comp.*, 170 (2005), 686–705.

[19] Chen, J., Li, W., Local convergence results of Gauss-Newton's like method in weak conditions, *J. Math. Anal. Appl.*, 324 (2006), 1381–1394.

[20] Chen, J., Li, W., Convergence behaviour of inexact Newton methods under weak Lipschitz condition, *J. Comput. Appl. Math.*, 191 (2006), 143–164.

[21] Dedieu, J.P., Shub, M., Newton's method for overdetermined systems of equations, *Math. Comp.*, 69 (2000), 1099–1115.

[22] Dembo, R.S., Eisenstat, S.C., Steihaug, T., Inexact Newton methods, *SIAM J. Numer. Anal.*, 19 (1982), 400–408.

[23] Dennis, J.E., Schnabel, R.B., Numerical methods for unconstrained optimization and nonlinear equations (Corrected reprint of the 1983 original), *Classics in Appl. Math.*, 16, SIAM, Philadelphia, PA, 1996.

[24] Ferreira, O.P., Local convergence of Newton's method in Banach space from the viewpoint of the majorant principle, *IMA J. Numer. Anal.*, 29 (2009), 746–759.

[25] Ferreira, O.P., Local convergence of Newton's method under majorant condition, *J. Comput. Appl. Math.*, 235 (2011), 1515–1522.

[26] Ferreira, O.P., Gonçalves, M.L.N, Local convergence analysis of inexact inexact-Newton like methods under majorant condition, *Comput. Optim. Appl.*, 48 (2011), 1–21.

[27] Ferreira, O.P., Gonçalves, M.L.N, Oliveira, P.R., Local convergence analysis of Gauss–Newton like methods under majorant condition, *J. Complexity*, 27 (2011), 111–125.

[28] Ferreira, O.P., Gonçalves, M.L.N, Oliveira, P.R., Local convergence analysis of inexact Gauss–Newton like methods under majorant condition, *J. Comput. Appl. Math.*, 236 (2012), 2487–2498.

[29] Ferreira, O.P., Svaiter, B.F., Kantorovich's majorants principle for Newton's method, *Comput. Optim. Appl.*, 42 (2009), 213–229.

[30] Guo, X., On semilocal convergence of inexact Newton methods, *J. Comput. Math.*, 25 (2007), 231–242.

[31] Gutiérrez, J.M., Hernández, M.A., Newton's method under weak Kantorovich conditions, *IMA J. Numer. Anal.*, 20 (2000), 521–532.

[32] Häubler, W.M., A Kantorovich–type convergence analysis for the Gauss–Newton–method, *Numer. Math.*, 48 (1986), 119–125.

[33] Huang, Z.A., *Convergence of inexact Newton method*, J. Zhejiang Univ. Sci. Ed., 30 (2003), 393–396.

[34] Hiriart-Urruty, J.B, Lemaréchal, C., Convex analysis and minimization algorithms (two volumes). I. Fundamentals. II. *Advanced theory and bundle methods*, 305 and 306, Springer–Verlag, Berlin, 1993.

[35] Kantorovich, L.V., Akilov, G.P., Functional Analysis, Pergamon Press, Oxford, 1982.

[36] Li, C., Hu, N., Wang, J., Convergence bahavior of Gauss–Newton's method and extensions to the Smale point estimate theory, *J. Complexity*, 26 (2010), 268–295.

[37] Li, C., Shen, W.P., Local convergence of inexact methods under the Hölder condition, *J. Comput. Appl. Math.*, 222 (2008), 544–560.

[38] Li, C., Zhang, W–H., Jin, X–Q., Convergence and uniqueness properties of Gauss-Newton's method, *Comput. Math. Appl.*, 47 (2004), 1057–1067.

[39] Martínez, J.M., Qi, L.Q., Inexact Newton methods for solving nonsmooth equations. Linear/nonlinear iterative methods and verification of solution (Matsuyama, 1993), *J. Comput. Appl. Math.*, 60 (1995), 127–145.

[40] Morini, B., Convergence behaviour of inexact Newton methods, *Math. Comp.*, 68 (1999), 1605–1613.

[41] Potra, F.A., Sharp error bounds for a class of Newton–like methods, *Libertas Mathematica*, 5 (1985), 71–84.

[42] Proinov, P.D., General local convergence theory for a class of iterative processes and its applications to Newton's method, *J. Complexity*, 25 (2009), 38–62.

[43] Proinov, P.D., New general convergence theory for iterative processes and its applications to Newton–Kantorovich type theorems, *J. Complexity*, 26 (2010), 3–42.

[44] Smale, S., Newton's method estimates from data at one point. The merging of disciplines: new directions in pure, applied, and computational mathematics (Laramie, Wyo., 1985), 185-196, Springer, New York, 1986.

[45] Stewart, G.W., On the continuity of the generalized inverse, *SIAM J. Appl. Math.*, 17 (1969), 33–45.

[46] Traub, J.F., Iterative Methods for the Solution of Equations, Englewood Cliffs, New Jersey: Prentice Hall, 1964.

[47] Wang, X.H., Convergence of Newton's method and uniqueness of the solution of equations in Banach spaces, *IMA J. Numer. Anal.*, 20 (2000), 123–134.

Chapter 8

Expending the Applicability of Lavrentiev Regularization Methods for Ill-Posed Problems

8.1. Introduction

In this chapter, we are interested in obtaining a stable approximate solution for a nonlinear ill-posed operator equation of the form

$$F(x) = y, \tag{8.1.1}$$

where $F : D(F) \subset X \to X$ is a monotone operator and X is a Hilbert space. We denote the inner product and the corresponding norm on a Hilbert space by $\langle \cdot, \cdot \rangle$ and $\| \cdot \|$, respectively. Let $U(x,r)$ stand for the open ball in X with center $x \in X$ and radius $r > 0$. Note that F is a monotone operator if it satisfies the relation

$$\langle F(x_1) - F(x_2), x_1 - x_2 \rangle \geq 0 \tag{8.1.2}$$

for all $x_1, x_2 \in D(F)$.

We assume, throughout this chapter, that $y^\delta \in Y$ is the available noisy data with

$$\|y - y^\delta\| \leq \delta \tag{8.1.3}$$

and (8.1.1) has a solution $\hat{x}$. Since (8.1.1) is ill-posed, its solution need not depend continuously on the data, i.e., small perturbation in the data can cause large deviations in the solution. So the regularization methods are used ([9, 10, 11, 13, 14, 17, 19, 20]). Since F is monotone, the Lavrentiev regularization is used to obtain a stable approximate solution of (8.1.1). In the Lavrentiev regularization, the approximate solution is obtained as a solution of the equation

$$F(x) + \alpha(x - x_0) = y^\delta, \tag{8.1.4}$$

where $\alpha > 0$ is the regularization parameter and x_0 is an initial guess for the solution $\hat{x}$.

In [8], Bakushinskii and Seminova proposed an iterative method

$$x_{k+1}^\delta = x_k^\delta - (\alpha_k I + A_{k,\delta})^{-1}[(F(x_k^\delta) - y^\delta) + \alpha_k(x_k^\delta - x_0)], x_0^\delta = x_0, \tag{8.1.5}$$

where $A_{k,\delta} := F'(x_k^\delta)$ and (α_k) is a sequence of positive real numbers satisfying $\alpha_k \to 0$ as $k \to \infty$. It is important to stop the iteration at an appropriate step, say $k = k_\delta$, and show that x_k is well defined for $0 < k \le k_\delta$ and $x_{k_\delta}^\delta \to \hat{x}$ as $\delta \to 0$ (see [15]).

In [6]-[8], Bakushinskii and Seminova chose the stopping index k_δ by requiring it to satisfy

$$\|F(x_{k_\delta}^\delta) - y^\delta\|^2 \le \tau\delta < \|F(x_k^\delta) - y^\delta\|$$

for $k = 0, 1, \cdots$ and $k_\delta - 1, \tau > 1$. In fact, they showed that $x_{k_\delta}^\delta \to \hat{x}$ as $\delta \to 0$ under the following assumptions:

(1) There exists $L > 0$ such that $\|F'(x) - F'(y)\| \le L\|x - y\|$ for all $x, y \in D(F)$;

(2) There exists $p > 0$ such that

$$\frac{\alpha_k - \alpha_{k+1}}{\alpha_k \alpha_{k+1}} \le p \tag{8.1.6}$$

for all $k \in \mathbb{N}$;

(3) $\sqrt{(2 + L\sigma)\|x_0 - \hat{x}\|td} \le \sigma - 2\|x_0 - \hat{x}\|t \le d\alpha_0$, where

$$\sigma := (\sqrt{\tau} - 1)^2, \quad t := p\alpha_0 + 1, \quad d = 2(t\|x_0 - \hat{x}\| + p\sigma).$$

However, no the error estimate was given in [8] (see [15]).

In [15], Mahale and Nair, motivated by the work of Qi-Nian Jin [12] for an iteratively regularized Gauss-Newton method, considered an alternate stopping criterion which not only ensures the convergence, but also derives an order optimal error estimate under a general source condition on $\hat{x} - x_0$. Moreover, the condition that they imposed on $\{\alpha_k\}$ is weaker than (8.1.6).

In the present chapter, we are motivated by [15]. In particular, we expand the applicability of the method (8.1.5) by weakening one of the major hypotheses in [15] (see below Assumption 8.2.1 (ii) in the next section).

In Section 8.2, we consider some basic assumptions required throughout the chapter. Section 8.3 deals with the stopping rule and a result that establishes the existence of the stopping index. In Section 8.4, we prove results for the iterations based on the exact data and, in Section 8.5, the error analysis for the noisy data case is proved. The main order optimal result using the a posteriori stopping rule is provided in Section 8.6.

8.2. Basic Assumptions and Some Preliminary Results

We use the following assumptions to prove the results in this chapter.

Assumption 8.2.1. (1) There exists $r > 0$ such that $U(\hat{x}, r) \subseteq D(F)$ and $F : U(\hat{x}, r) \to X$ is Fréchet differentiable.

(2) There exists $K_0 > 0$ such that, for all $u_\theta = u + \theta(\hat{x} - u) \in U(\hat{x}, r)$, $\theta \in [0, 1]$ and $v \in X$, there exists an element, say $\phi(\hat{x}, u_\theta, v) \in X$, satisfying

$$[F'(\hat{x}) - F'(u_\theta)]v = F'(u_\theta)\phi(\hat{x}, u_\theta, v), \quad \|\phi(\hat{x}, u_\theta, v)\| \le K_0\|v\|\|\hat{x} - u_\theta\|$$

for all $u_\theta \in U(\hat{x}, r)$ and $v \in X$.

(3) $\|F'(u) + \alpha I)^{-1}F'(u_\theta)\| \le 1$ for all $u_\theta \in U(\hat{x}, r)$.

(4) $\|(F'(u) + \alpha I)^{-1}\| \le \frac{1}{\alpha}$ for all $u_\theta \in U(\hat{x}, r)$.

The condition (2) in Assumption 8.2.1 weakens the popular hypotheses given in [15], [16] and [18].

Assumption 8.2.2. There exists a constant $K > 0$ such that, for all $x, y \in U(\hat{x}, r)$ and $v \in X$, there exists an element denoted by $P(x, u, v) \in X$ satisfying

$$[F'(x) - F'(u)]v = F'(u)P(x,u,v), \quad \|P(x,u,v)\| \leq K\|v\|\|x-u\|.$$

Clearly, Assumption 8.2.2 implies Assumption 8.2.1 (2) with $K_0 = K$, but not necessarily vice versa. Note that $K_0 \leq K$ holds in general and $\frac{K_0}{K}$ can be arbitrarily large [1]-[5]. Indeed, there are many classes of operators satisfying Assumption 8.2.1 (2), but not Assumption 8.2.2 (see the numerical examples at the end of this chapter). Moreover, if K_0 is sufficiently smaller than K which can happen since $\frac{K}{K_0}$ can be arbitrarily large, then the results obtained in this chapter provide a tighter error analysis than the one in [15].

Finally, note that the computation of constant K is more expensive than the computation of K_0.

We need the auxiliary results based on Assumption 8.2.1.

Proposition 8.2.3. *For any $u \in U(\hat{x}, r)$ and $\alpha > 0$,*

$$\|(F'(u) + \alpha I)^{-1}[F(\hat{x}) - F(u) - F'(u)(\hat{x} - u)\| \leq \frac{3K_0}{2}\|\hat{x} - u\|^2.$$

Proof. Using the fundamental theorem of integration, for any $u \in U(\hat{x}, r)$, we get

$$F(\hat{x}) - F(u) = \int_0^1 F'(u + t(\hat{x} - u))(\hat{x} - u)dt.$$

Hence, by Assumption 8.2.2,

$$\begin{aligned}
& F(\hat{x}) - F(u) - F'(u)(\hat{x} - u) \\
= & \int_0^1 [F'(u + t(\hat{x} - u)) - F'(\hat{x}) + F'(\hat{x}) - F'(u)](\hat{x} - u)dt \\
= & \int_0^1 F'(\hat{x})[\phi(u + t(\hat{x} - u), \hat{x}, \hat{x} - u) - \phi(u, \hat{x}, \hat{x} - u)]dt.
\end{aligned}$$

Then, by (2), (3) in Assumptions 8.2.1 and the inequality $\|(F'(u) + \alpha I)^{-1}F'(u_\theta)\| \leq 1$, we obtain in turn

$$\begin{aligned}
& \|(F'(u) + \alpha I)^{-1}[F(\hat{x}) - F(u) - F'(u)(\hat{x} - u)\| \\
\leq & \int_0^1 \|\phi(u + t(\hat{x} - u), \hat{x}, \hat{x} - u) + \phi(u, \hat{x}, \hat{x} - u)\|dt. \\
\leq & \int_0^1 K_0\|\hat{x} - u\|^2 t dt + K_0\|\hat{x} - u\|^2 \\
\leq & \frac{3K_0}{2}\|\hat{x} - u\|^2.
\end{aligned}$$

This completes the proof.

Proposition 8.2.4. *For any $u \in U(\hat{x}, r)$ and $\alpha > 0$,*

$$\alpha \|(F'(\hat{x}) + \alpha I)^{-1} - (F'(u) + \alpha I)^{-1}\| \leq K_0 \|\hat{x} - u\|. \tag{8.2.1}$$

Proof. Let $T_{\hat{x},u} = \alpha(F'(\hat{x}) + \alpha I)^{-1} - (F'(u) + \alpha I)^{-1})$ for all $v \in X$. Then we have, by Assumption 8.2.2,

$$\begin{aligned}
\|T_{\hat{x},u} v\| &= \|\alpha(F'(\hat{x}) + \alpha I)^{-1}(F'(u) - F'(\hat{x})(F'(u) + \alpha I)^{-1})v\| \\
&= \|(F'(\hat{x}) + \alpha I)^{-1} F'(\hat{x}) \phi(u, \hat{x}, \alpha(F'(u) + \alpha I)^{-1} v)\| \\
&\leq K_0 \|\hat{x} - u\| \|v\|
\end{aligned}$$

for all $v \in X$. This completes the proof.

Assumption 8.2.5. There exists a continuous and strictly monotonically increasing function $\varphi : (0, a] \to (0, \infty)$ with $a \geq \|F'(\hat{x})\|$ satisfying

(1) $\lim_{\lambda \to 0} \varphi(\lambda) = 0$;

(2) $\sup_{\lambda \geq 0} \frac{\alpha \varphi(\lambda)}{\lambda + \alpha} \leq \varphi(\alpha)$ for all $\alpha \in (0, a]$;

(3) there exists $v \in X$ with $\|v\| \leq 1$ such that

$$\hat{x} - x_0 = \varphi(F'(\hat{x}))v. \tag{8.2.2}$$

Next, we assume a condition on the sequence $\{\alpha_k\}$ considered in (8.1.5).

Assumption 8.2.6. ([15], Assumption 2.6) The sequence $\{\alpha_k\}$ of positive real numbers is such that

$$1 \leq \frac{\alpha_k}{\alpha_{k+1}} \leq \mu, \quad \lim_{k \to 0} \alpha_k = 0 \tag{8.2.3}$$

for a constant $\mu > 1$.

Note that the condition (8.2.3) on $\{\alpha_k\}$ is weaker than (8.1.6) considered by Bakunshinskii and Smirnova [8] (see [15]). In fact, if (8.1.6) is satisfied, then it also satisfies (8.2.3) with $\mu = p\alpha_0 + 1$, but the converse need not be true (see [15]). Further, note that, for these choices of $\{\alpha_k\}$, α_k / α_{k+1} is bounded whereas $(\alpha_k - \alpha_{k+1})/\alpha_k \alpha_{k+1} \to \infty$ as $k \to \infty$. (2) in Assumption 8.2.1 is used in the literature for regularization of many nonlinear ill-posed problems (see [12], [13], [19]-[21]).

8.3. Stopping Rule

Let $c_0 > 4$ and choose k_δ to be the first non-negative integer such that x_k^δ in (8.1.5) is defined for each $k \in \{0, 1, 2, \cdots, k_\delta\}$ and

$$\|\alpha_{k_\delta}(A_{k_\delta}^\delta + \alpha_{k_\delta} I)^{-1}(F(x_{k_\delta}^\delta) - y^\delta)\| \leq c_0 \delta. \tag{8.3.1}$$

In the following, we establish the existence of such a k_δ. First, we consider the positive integer $N \in \mathbb{N}$ satisfying

$$\alpha_N \leq \frac{(c-1)\delta}{\|x_0 - \hat{x}\|} < \alpha_k \tag{8.3.2}$$

for all $k \in \{0, 1, \cdots, N-1\}$, where $c > 1$ and $\alpha_0 > (c-1)\delta/\|x_0 - \hat{x}\|$.

The following technical lemma from [15] is used to prove some of the results of this chapter.

Lemma 8.3.1. ([15], Lemma 3.1) *Let $a > 0$ and $b \geq 0$ be such that $4ab \leq 1$ and $\theta := (1 - \sqrt{1-4ab})/2a$. Let $\theta_1, \cdots, \theta_n$ be non-negative real numbers such that $\theta_{k+1} \leq a\theta_k^2 + b$ and $\theta_0 \leq \theta$. Then $\theta_k \leq \theta$ for all $k = 1, 2, \cdots, n$.*

The rest of the results in this chapter can be proved along the same lines of the proof in [15]. In order for us to make the chapter is a self contained as possible we present the proof of one of them and for the proof of the rest we refer the reader to [15].

Theorem 8.3.2. ([15], Theorem 3.2) *Let (8.1.2), (8.1.3), (8.2.3) and Assumption 8.2.1 be satisfied. Let N be as in (8.3.2) for some $c > 1$ and $6cK_0\|x_0 - \hat{x}\|/(c-1) \leq 1$. Then x_k^δ is defined iteratively for each $k \in \{0, 1, \cdots, N\}$ and*

$$\|x_k^\delta - \hat{x}\| \leq \frac{2c\|x_0 - \hat{x}\|}{c-1} \tag{8.3.3}$$

for all $k \in \{0, 1, \cdots, N\}$. In particular, if $r > 2c\|x_0 - \hat{x}\|/(c-1)$, then $x_k^\delta \in B_r(\hat{x})$ for $k \in \{0, 1, \cdots, N\}$. Moreover,

$$\|\alpha_N(A_N^\delta + \alpha_N I)^{-1}(F(x_N^\delta) - y^\delta)\| \leq c_0\delta \tag{8.3.4}$$

for $c_0 := \frac{7}{3}c + 1$.

Proof. We show (8.3.3) by induction. It is obvious that (8.3.3) holds for $k = 0$. Now, assume that (8.3.3) holds for some $k \in \{0, 1, \cdots, N\}$. Then it follows from (8.1.5) that

$$\begin{aligned}
& x_{k+1}^\delta - \hat{x} \\
= \;& x_k^\delta - \hat{x} - (A_k^\delta + \alpha_k I)^{-1}[F(x_k^\delta) - y^\delta + \alpha_k(x_k^\delta - x_0)] \\
= \;& (A_k^\delta + \alpha_k I)^{-1}((A_k^\delta + \alpha_k I)(x_k^\delta - \hat{x}) - [F(x_k^\delta) - y^\delta + \alpha_k(x_k^\delta - x_0)]) \\
= \;& (A_k^\delta + \alpha_k I)^{-1}[A_k^\delta(x_k^\delta - \hat{x}) + y^\delta - F(x_k^\delta) + \alpha_k(x_0 - \hat{x})] \\
= \;& \alpha_k(A_k^\delta + \alpha_k I)^{-1}(x_0 - \hat{x}) + (A_k^\delta + \alpha_k I)^{-1}(y^\delta - y) \\
& + (A_k^\delta + \alpha_k I)^{-1}[F(\hat{x}) - F(x_k^\delta) + A_k^\delta(x_k^\delta - \hat{x})]
\end{aligned} \tag{8.3.5}$$

Using (8.1.3), the estimates $\|(A_k^\delta + \alpha_k I)^{-1}\| \leq 1/\alpha_k$, $\|(A_k^\delta + \alpha_k I)^{-1}A_k^\delta\| \leq 1$ and Proposition 8.2.3, we have

$$\|\alpha_k(A_k^\delta + \alpha_k I)^{-1}(x_0 - \hat{x}) + (A_k^\delta + \alpha_k I)^{-1}(y^\delta - y)\| \leq \|x_0 - \hat{x}\| + \frac{\delta}{\alpha_k}$$

and

$$\|(A_k^\delta + \alpha_k I)^{-1}[F(\hat{x}) - F(x_k^\delta) + A_k^\delta(x_k^\delta - \hat{x})]\| \leq \frac{3K_0}{2}\|x_k^\delta - \hat{x}\|^2.$$

Thus we have

$$\|x_{k+1}^\delta - \hat{x}\| \leq \|x_0 - \hat{x}\| + \frac{\delta}{\alpha_k} + \frac{3K_0}{2}\|x_k^\delta - \hat{x}\|^2.$$

But, by (8.3.2), $\frac{\delta}{\alpha_k} \leq \|x_0 - \hat{x}\|/(c-1)$ and so

$$\|x_{k+1}^\delta - \hat{x}\| \leq \frac{c\|x_0 - \hat{x}\|}{c-1} + \frac{3K_0}{2}\|x_k^\delta - \hat{x}\|^2,$$

which leads to the recurrence relation

$$\theta_{k+1} \leq a\theta_k^2 + b,$$

where

$$\theta_k = \|x_k^\delta - \hat{x}\|, \quad a = \frac{3K_0}{2}, \quad b = \frac{c\|x_0 - \hat{x}\|}{c-1}.$$

From the hypothesis of the theorem, we have $4ab = 6cK_0 \frac{\|x_0 - \hat{x}\|}{c-1} < 1$. It is obvious that

$$\theta_0 \leq \|x_0 - \hat{x}\| \leq \theta := \frac{1 - \sqrt{1-4ab}}{2a} = \frac{2b}{1+\sqrt{1-4ab}} \leq 2b = \frac{2c\|x_0 - \hat{x}\|}{c-1}.$$

Hence, by Lemma 8.3.1, we get

$$\|x_k^\delta - \hat{x}\| \leq \theta \leq \frac{2c\|x_0 - \hat{x}\|}{c-1} \tag{8.3.6}$$

for all $k \in \{0, 1, \cdots, N\}$. In particular, if $r > 2c\|x_0 - \hat{x}\|/(c-1)$, then we have $x_k^\delta \in B_r(\hat{x})$ for all $k \in \{0, 1, \cdots, N\}$.

Next, let $\gamma = \|\alpha_N(A_N^\delta + \alpha_N I)^{-1}(F(x_N^\delta) - y^\delta)\|$. Then, using the estimates

$$\|\alpha_N(A_N^\delta + \alpha_N I)^{-1}\| \leq 1, \quad \|\alpha_N(A_N^\delta + \alpha_N I)^{-1} A_N^\delta\| \leq \alpha_k$$

and Proposition 8.2.3, we have

$$\begin{aligned}
& \gamma \\
\leq \;& \delta + \|\alpha_N(A_N^\delta + \alpha_N I)^{-1}(F(x_N^\delta) - y + A_N^\delta(x_N^\delta - \hat{x}) - A_N^\delta(x_N^\delta - \hat{x}))\| \\
= \;& \delta + \|\alpha_N(A_N^\delta + \alpha_N I)^{-1}[F(x_N^\delta) - F(\hat{x}) - A_N^\delta(x_N^\delta - \hat{x}) + A_N^\delta(x_N^\delta - \hat{x})]\| \\
\leq \;& \delta + \alpha_N[3K_0\frac{\|x_N^\delta - \hat{x}\|^2}{2} + \|x_N^\delta - \hat{x}\|] \\
\leq \;& \delta + \alpha_N\|x_N^\delta - \hat{x}\|[1 + 3K_0\frac{\|x_N^\delta - \hat{x}\|}{2}] \\
\leq \;& \delta + \frac{2\alpha_N c\|x_0^\delta - \hat{x}\|}{c-1}[1 + \frac{3K_0 c\|x_0^\delta - \hat{x}\|}{c-1}] \leq \delta + 2c\delta[1 + \frac{1}{6}] \leq (\frac{7c}{3} + 1)\delta.
\end{aligned} \tag{8.3.7}$$

Therefore, we have $\|\alpha_N(A_N^\delta + \alpha_N I)^{-1}(F(x_N^\delta) - y^\delta)\| \leq c_0\delta$, where $c_0 := \frac{7}{3}c + 1$. This completes the proof.

8.4. Error Bound for the Case of Noise-Free Data

Let

$$x_{k+1} = x_k - (A_k + \alpha_k I)^{-1}[F(x_k) - y + \alpha_k(x_k - x_0)] \tag{8.4.1}$$

for all $k \geq 0$.

We show that each x_k is well defined and belongs to $U(\hat{x}, r)$ for $r > 2\|x_0 - \hat{x}\|$. For this, we make use of the following lemma.

Lemma 8.4.1. ([15], Lemma 4.1) *Let Assumption 8.2.1 hold. Suppose that, for all $k \in \{0,1,\cdots,n\}$, x_k in (8.4.1) is well defined and $\rho_k := \|\alpha_k(A_k + \alpha_k I)^{-1}(x_0 - \hat{x})\|$ for some $n \in \mathbb{N}$. Then we have*

$$\rho_k - \frac{3K_0\|x_k - \hat{x}\|^2}{2} \leq \|x_{k+1} - \hat{x}\| \leq \rho_k + \frac{3K_0\|x_k - \hat{x}\|^2}{2} \tag{8.4.2}$$

for all $k \in \{0,1,\cdots,n\}$.

Theorem 8.4.2. ([15], Theorem 4.2) *Let Assumption 8.2.1 hold. If $6K_0\|x_0 - \hat{x}\| \leq 1$ and $r > 2\|x_0 - \hat{x}\|$, then, for all $k \in \mathbb{N}$, the iterates x_k in (8.4.1) are well defined and*

$$\|x_k - \hat{x}\| \leq \frac{2\|x_0 - \hat{x}\|}{1 + \sqrt{1 - 6K_0\|x_0 - \hat{x}\|}} \leq 2\|x_0 - \hat{x}\| \tag{8.4.3}$$

for all $k \in \mathbb{N}$.

Lemma 8.4.3. ([15], Lemma 4.3) *Let Assumptions 8.2.1 and 8.2.6 hold and let $r > 2\|x_0 - \hat{x}\|$. Assume that $\|A\| \leq \eta\alpha_0$ and $4\mu(1+\eta^{-1})K_0\|x_0 - \hat{x}\| \leq 1$ for some η with $0 < \eta < 1$. Then, for all $k \in \mathbb{N}$, we have*

$$\frac{1}{(1+\eta)\mu}\|x_k - \hat{x}\| \leq \|\alpha_k(A_k + \alpha_k I)^{-1}(x_0 - \hat{x})\| \leq \frac{1}{1-\eta}\|x_k - \hat{x}\| \tag{8.4.4}$$

and

$$\frac{1-\eta}{(1+\eta)\mu}\|x_k - \hat{x}\| \leq \|(x_{k+1} - \hat{x})\| \leq \left(\frac{1}{1-\eta} + \frac{\eta}{(1+\eta)\mu}\right)\|x_k - \hat{x}\|. \tag{8.4.5}$$

The following corollary follows from Lemma 8.4.3 by taking $\eta = 1/3$. We show that this particular case of Lemma 8.4.3 is better suited for our later results.

Corollary 8.4.4. ([15], Corollary 4.4) *Let Assumptions 8.2.1 and 8.2.6 hold and let $r > 2\|x_0 - \hat{x}\|$. Assume that $\|A\| \leq \alpha_0/3$ and $16\mu K_0\|x_0 - \hat{x}\| \leq 1$. Then, for all $k \in \mathbb{N}$, we have*

$$\frac{3}{4\mu}\|x_k - \hat{x}\| \leq \|\alpha_k(A + \alpha_k I)^{-1}(x_0 - \hat{x})\| \leq \frac{3}{2}\|x_k - \hat{x}\| \tag{8.4.6}$$

and

$$\frac{\|x_k - \hat{x}\|}{2\mu} \leq \|(x_{k+1} - \hat{x})\| \leq 2\|x_k - \hat{x}\|.$$

Theorem 8.4.5. ([15], Theorem 4.5) *Let the Assumptions of Lemma 8.4.3 hold. If x_0 is chosen such that $x_0 - \hat{x} \in N(F'(\hat{x}))^{\perp}$, then $\lim_{k\to\infty} x_k = \hat{x}$.*

Lemma 8.4.6. ([15], Lemma 4.6) *Let the assumptions of Lemma 8.4.3 hold for η satisfying*

$$(1-\sqrt{1-\frac{\eta}{(1+\eta)\mu}})[1+(2\mu-1)\eta+2\mu]+2\eta<\frac{4}{3}. \tag{8.4.7}$$

Then, for all $k,l \in \mathbb{N}\mathrm{U}\{0\}$ with $k \geq l$, we have

$$\|x_l-\hat{x}\| \leq c_\eta\left[\|x_k-\hat{x}\|+\frac{\|\alpha_l(A+\alpha_l I)^{-1}(F(x_l)-y)\|}{\alpha_k}\right],$$

where

$$c_\eta := (1-b_\eta)^{-1}\max\left\{\mu, 1+\frac{(3\varepsilon+1)\eta}{4(1-\eta)}\right\},$$

$$b_\eta := \frac{(3\varepsilon+1)\eta}{(1-\eta)}+\frac{3\varepsilon a}{4}, \quad \varepsilon := \frac{1-\sqrt{1-a}}{a}, \quad a := \frac{\eta}{(1+\eta})\mu.$$

Remark 8.4.7. ([15], Remark 4.7) It can be seen that (8.4.7) is satisfied if $\eta \leq 1/3+1/24$.

Now, if we take $\eta = 1/3$, that is, $K_0\|x_0-\hat{x}\|\mu \leq 1/16$ in Lemma 8.4.6, then it takes the following form.

Lemma 8.4.8. ([15], Lemma 4.8) *Let the assumptions of Lemma 8.4.3 hold with $\eta = 1/3$. Then, for all $k \geq l \geq 0$, we have*

$$\|x_l-\hat{x}\| \leq c_{1/3}\left[\|x_k-\hat{x}\|+\frac{\|\alpha_l(A+\alpha_l I)^{-1}(F(x_l)-y)\|}{\alpha_k}\right],$$

where

$$c_{1/3} = \left[1-\frac{8\mu+(8\mu+1)3\varepsilon}{16\mu}\right]^{-1}\max\left\{\mu, 1+\frac{3\varepsilon+1}{8}\right\},$$

$$\varepsilon := \frac{\sqrt{4\mu}}{\sqrt{4\mu}+\sqrt{4\mu-1}}.$$

8.5. Error Analysis with Noisy Data

The first result in this section gives an error estimate for $\|x_k^\delta - x_k\|$ under the Assumption 8.2.5, where $k = 0,1,2,\cdots,N$.

Lemma 8.5.1. ([15], Lemma 5.1) *Let Assumption 8.2.1 hold and let $K_0\|x_0-\hat{x}\| \leq 1/m$, where $m > (7+\sqrt{73})/2$, and N be the integer satisfying* (8.3.2) *with*

$$c > \frac{m^2-4m-6}{m^2-7m-6}.$$

Then, for all $k \in \{0, 1, \cdots, N\}$, *we have*

$$\|x_k^\delta - x_k\| \le \frac{\delta}{(1-\kappa)\alpha_k}, \tag{8.5.1}$$

where

$$\kappa := \frac{1}{m}\Big(4 + \frac{3c}{c-1} + \frac{6}{m}\Big).$$

If we take $m = 8$ in Lemma 8.5.1, then we get the following corollary as a particular case of Lemma 8.5.1. We make use of it in the following error analysis.

Corollary 8.5.2. ([15], Corollary 5.2) *Let Assumption 8.2.1 hold and let* $16K_0\|x_0 - \hat{x}\| \le 1$. *Let* N *be the integer defined by* (8.3.2) *with* $c > 13$. *Then, for all* $k \in \{0, 1, \cdots, N\}$, *we have*

$$\|x_k^\delta - x_k\| \le \frac{\delta}{(1-\kappa)\alpha_k},$$

where

$$\kappa := \frac{31c - 19}{32(c-1)}.$$

Lemma 8.5.3. ([15], Lemma 5.3) *Let the assumptions of Lemma 8.5.1 hold. Then we have*

$$\|\alpha_k(A + \alpha_{k_\delta} I)^{-1}(F(x_{k_\delta}) - y)\| \le c_1\delta.$$

Moreover, if $k_\delta > 0$, *then, for all* $0 \le k < k_\delta$, *we have*

$$\|\alpha_k(A + \alpha_k I)^{-1}(F(x_k) - y)\| \ge c_2\delta,$$

where

$$c_1 = \Big(1 + \frac{2cK_0\|x_0 - \hat{x}\|}{c-1}\Big)\Big(c_0 + \frac{2-\kappa}{1-\kappa} + \frac{3K_0\mu\|x_0 - \hat{x}\|}{2(1-\kappa)^2(c-1)}\Big),$$

$$c_2 = \frac{c_0 - ((2-\kappa)(1-\kappa)) - (3K_0\|x_0 - \hat{x}\|/2(1-\kappa)^2(c-1))}{1 + 2(cK_0\|x_0 - \hat{x}\|/(c-1))}$$

with $c_0 = \frac{7}{3}c + 1$ *and* κ *as in Lemma 8.5.1.*

Theorem 8.5.4. ([15], Theorem 5.4) *Let Assumptions 8.2.1 and 8.2.6 hold. If* $16k\mu\|x_0 - \hat{x}\| \le 1$ *and the integer* k_δ *is chosen according to stopping rule* (8.3.1) *with* $c_0 > \frac{94}{3}$, *then we have*

$$\|x_{k_\delta}^\delta - \hat{x}\| \le \xi \inf\Big\{\|x_k - \hat{x}\| + \frac{\delta}{\alpha_k} : k \ge 0\Big\}, \tag{8.5.2}$$

where $\xi = \max\{2\mu\rho, \frac{c_{1/3}c_1 + 1}{1-\kappa}, c\}$, $\rho := 1 + \frac{\mu(1 + 3K_0\|x_0 - \hat{x}\|)}{c_2(1-\kappa)}$ *with* $c_{1/3}$ *and* κ *as in Lemma 8.4.8 and Corollary 8.5.2, respectively, and* c_1, c_2 *as in Lemma 8.5.3.*

8.6. Order Optimal Result with an a Posterior Stopping Rule

In this section, we show the convergence $x^{\delta}_{k_\delta} \to \hat{x}$ as $\delta \to 0$ and also give an optimal error estimate for $\|x^{\delta}_{k_\delta} - \hat{x}\|$.

Theorem 8.6.1. ([15], Theorem 6.1) *Let the assumptions of Theorem 8.5.4 hold and let k_δ be the integer chosen by (8.3.1). If x_0 is chosen such that $x_0 - \hat{x} \in N(F'(\hat{x}))^{\perp}$, then we have $\lim_{\delta\to 0} x^{\delta}_{k_\delta} = \hat{x}$. Moreover, if Assumption 8.2.5 is satisfied, then we have*

$$\|x^{\delta}_{k_\delta} - \hat{x}\| \le \xi' \mu \psi^{-1}\Big(\delta\Big),$$

where $\xi' := 8\mu\xi/3$ with ξ as in Theorem 8.5.4 and $\psi : (0, \varphi(a)] \to (0, a\varphi(a)]$ is defined as $\psi(\lambda) := \lambda\varphi^{-1}(\lambda)$, $\lambda \in (0, \varphi(a)]$.

Proof. From (8.4.6) and (8.5.2), we get

$$\|x^{\delta}_{k_\delta} - \hat{x}\| \le \xi'' \inf\{\|\alpha_k(A+\alpha_k I)^{-1}(x_0-\hat{x})\| + \frac{\delta}{\alpha_k} : k = 0,1,\cdots\} \tag{8.6.1}$$

where $\xi'' = \frac{4\mu}{3}\max\{2\mu\Big(1+\frac{\mu(1+3k_0\|x_0-\hat{x}\|)}{c_2(1-\kappa)}\Big), \frac{c_{1/3}c_1+1}{1-\kappa}, c\}$. Now, we choose an integer m_δ such that $m_\delta = \max\{k : \alpha_k \ge \sqrt{\delta}\}$. Then, we have

$$\|x^{\delta}_{k_\delta} - \hat{x}\| \le \xi'' \inf\{\|\alpha_{m_\delta}(A+\alpha_{m_\delta} I)^{-1}(x_0-\hat{x})\| + \frac{\delta}{\alpha_{m_\delta}} : k = 0,1,\cdots\} \tag{8.6.2}$$

Note that $\frac{\delta}{\alpha_{m_\delta}} \le \sqrt{\delta}$, so $\frac{\delta}{\alpha_{m_\delta}} \to 0$ as $\delta \to 0$. Therefore by (8.6.2) to show that $x^{\delta}_{k_\delta} \to \hat{x}$ as $\delta \to 0$, it is enough to prove that $\|\alpha_{m_\delta}(A+\alpha_{m_\delta}I)^{-1}(x_0-\hat{x})\| \to 0$ as $\delta \to 0$. Observe that, for $w \in R(F'(\hat{x}))$, i.e., $w = F'(\hat{x})u$ for some $u \in D(F)$ we have $\|\alpha_{m_\delta}(A+\alpha_{m_\delta}I)^{-1}w\| \le \alpha_{m_\delta}\|u\| \to 0$ as $\delta \to 0$. Now since $R(F'(\hat{x}))$ is a dense subset of $N(F'(\hat{x}))^{\perp}$ it follows that $\|\alpha_{m_\delta}(A+\alpha_{m_\delta}I)^{-1}(x_0-\hat{x})\| \to 0$ as $\delta \to 0$. Using Assumption 8.2.5, we get that

$$\|\alpha_k(A+\alpha_k I)^{-1}(x_0-\hat{x})\| \le \varphi(\alpha_k). \tag{8.6.3}$$

So by (8.6.2) and (8.6.3) we obtain that

$$\|x^{\delta}_{k_\delta} - \hat{x}\| \le \xi'' \inf\{\varphi(\alpha_k) + \frac{\delta}{\alpha_k} : k = 0,1,\cdots\}. \tag{8.6.4}$$

Choose $\hat{k_\delta}$ such that

$$\varphi(\alpha_{\hat{k_\delta}})\alpha_{\hat{k_\delta}} \le \delta < \varphi(\alpha_k)\alpha_k \text{ for } k = 0,1,\cdots k_\delta - 1. \tag{8.6.5}$$

This also implies that

$$\psi(\varphi(\alpha_{\hat{k_\delta}})) \le \delta < \psi(\varphi(\alpha_k)) \text{ for } k = 0,1,\cdots k_\delta - 1. \tag{8.6.6}$$

From (8.6.4), $\|x^{\delta}_{k_\delta} - \hat{x}\| \le \xi''\{\varphi(\alpha_{\hat{k_\delta}}) + \frac{\delta}{\alpha_{\hat{k_\delta}}}\}$. Now using (8.6.5) and (8.6.6) we get $\|x^{\delta}_{k_\delta} - \hat{x}\| \le 2\xi''\frac{\delta}{\alpha_{\hat{k_\delta}}} \le 2\xi''\mu\frac{\delta}{\alpha_{\hat{k_\delta}-1}} \le 2\xi''\mu\psi^{-1}(\delta)$. This completes the proof.

8.7. Numerical Examples

We provide two numerical examples, where $K_0 < K$.

Example 8.7.1. Let $X = \mathbb{R}$, $D(F) = \overline{U(0,1)}$, $\hat{x} = 0$ and define a function F on $D(F)$ by

$$F(x) = e^x - 1. \tag{8.7.1}$$

Then, using (8.7.1) and Assumptions 8.2.1 (2) and 8.2.2, we get

$$K_0 = e - 1 < K = e.$$

Example 8.7.2. Let $X = C([0,1])$ (: the space of continuous functions defined on $[0,1]$ equipped with the max norm) and $D(F) = \overline{U(0,1)}$. Define an operator F on $D(F)$ by

$$F(h)(x) = h(x) - 5\int_0^1 x\theta h(\theta)^3 d\theta. \tag{8.7.2}$$

Then the Fréchet-derivative is given by

$$F'(h[u])(x) = u(x) - 15\int_0^1 x\theta h(\theta)^2 u(\theta) d\theta \tag{8.7.3}$$

for all $u \in D(F)$. Using (8.7.2), (8.7.3), Assumptions 8.2.1 (2), 8.2.2 for $\hat{x} = 0$, we get $K_0 = 7.5 < K = 15$.

Next, we provide an example where $\frac{K}{K_0}$ can be arbitrarily large.

Example 8.7.3. Let $X = D(F) = \mathbb{R}$, $\hat{x} = 0$ and define a function F on $D(F)$ by

$$F(x) = d_0 x - d_1 \sin 1 + d_1 \sin e^{d_2 x}, \tag{8.7.4}$$

where d_0, d_1 and d_2 are the given parameters. Note that $F(\hat{x}) = F(0) = 0$. Then it can easily be seen that, for d_2 sufficiently large and d_1 sufficiently small, $\frac{K}{K_0}$ can be arbitrarily large.

We now present two examples where Assumption 8.2.2 is not satisfied, but Assumption 8.2.1 (2) is satisfied.

Example 8.7.4. Let $X = D(F) = \mathbb{R}$, $\hat{x} = 0$ and define a function F on D by

$$F(x) = \frac{x^{1+\frac{1}{i}}}{1+\frac{1}{i}} + c_1 x - c_1 - \frac{i}{i+1}, \tag{8.7.5}$$

where c_1 is a real parameter and $i > 2$ is an integer. Then $F'(x) = x^{1/i} + c_1$ is not Lipschitz on D. Hence Assumption 8.2.2 is not satisfied. However, the central Lipschitz condition in Assumption 8.2.2 (2) holds for $K_0 = 1$. We also have that $F(\hat{x}) = 0$. Indeed, we have

$$\begin{aligned} \|F'(x) - F'(\hat{x})\| &= |x^{1/i} - \hat{x}^{1/i}| \\ &= \frac{|x-\hat{x}|}{\hat{x}^{\frac{i-1}{i}} + \cdots + x^{\frac{i-1}{i}}} \end{aligned}$$

and so

$$\|F'(x) - F'(\hat{x})\| \leq K_0 |x - \hat{x}|.$$

Example 8.7.5. We consider the integral equation

$$u(s) = f(s) + \lambda \int_a^b G(s,t)u(t)^{1+1/n}dt \tag{8.7.6}$$

for all $n \in \mathbb{N}$, where f is a given continuous function satisfying $f(s) > 0$ for all $s \in [a,b]$, λ is a real number and the kernel G is continuous and positive in $[a,b] \times [a,b]$.

For example, when $G(s,t)$ is the Green kernel, the corresponding integral equation is equivalent to the boundary value problem

$$\begin{aligned} &u'' = \lambda u^{1+1/n}, \\ &u(a) = f(a),\ u(b) = f(b). \end{aligned}$$

These type of the problems have been considered in [1]–[5]. The equation of the form (8.7.6) generalize the equation of the form

$$u(s) = \int_a^b G(s,t)u(t)^n dt, \tag{8.7.7}$$

which was studied in [1]-[5]. Instead of (8.7.6), we can try to solve the equation $F(u) = 0$, where

$$F : \Omega \subseteq C[a,b] \to C[a,b], \quad \Omega = \{u \in C[a,b] : u(s) \geq 0, s \in [a,b]\}$$

and

$$F(u)(s) = u(s) - f(s) - \lambda \int_a^b G(s,t)u(t)^{1+1/n}dt.$$

The norm we consider is the max-norm. The derivative F' is given by

$$F'(u)v(s) = v(s) - \lambda(1 + \frac{1}{n}) \int_a^b G(s,t)u(t)^{1/n}v(t)dt$$

for all $v \in \Omega$. First of all, we notice that F' does not satisfy the Lipschitz-type condition in Ω. Let us consider, for instance, $[a,b] = [0,1]$, $G(s,t) = 1$ and $y(t) = 0$. Then we have $F'(y)v(s) = v(s)$ and

$$\|F'(x) - F'(y)\| = |\lambda|(1 + \frac{1}{n}) \int_a^b x(t)^{1/n}dt.$$

If F' were the Lipschitz function, then we have

$$\|F'(x) - F'(y)\| \leq L_1\|x - y\|$$

or, equivalently, the inequality

$$\int_0^1 x(t)^{1/n}dt \leq L_2 \max_{x \in [0,1]} x(s) \tag{8.7.8}$$

would hold for all $x \in \Omega$ and for a constant L_2. But this is not true. Consider, for example, the function

$$x_j(t) = \frac{t}{j}$$

for all $j \geq 1$ and $t \in [0,1]$. If these are substituted into (8.7.7), then we have

$$\frac{1}{j^{1/n}(1+1/n)} \leq \frac{L_2}{j} \quad \Longleftrightarrow \quad j^{1-1/n} \leq L_2(1+1/n)$$

for all $j \geq 1$. This inequality is not true when $j \to \infty$. Therefore, Assumption 8.2.2 is not satisfied in this case. However, Assumption 8.2.1 (2) holds. To show this, suppose that $\hat{x}(t) = f(t)$ and $\gamma = \min_{s\in[a,b]} f(s)$. Then, for all $v \in \Omega$, we have

$$\begin{aligned} \|[F'(x) - F'(\hat{x})]v\| &= |\lambda|\left(1+\frac{1}{n}\right)\max_{s\in[a,b]}\left|\int_a^b G(s,t)(x(t)^{1/n} - f(t)^{1/n})v(t)dt\right| \\ &\leq |\lambda|\left(1+\frac{1}{n}\right)\max_{s\in[a,b]} G_n(s,t), \end{aligned}$$

where $G_n(s,t) = \frac{G(s,t)|x(t)-f(t)|}{x(t)^{(n-1)/n}+x(t)^{(n-2)/n}f(t)^{1/n}+\cdots+f(t)^{(n-1)/n}}\|v\|$. Hence it follows that

$$\begin{aligned} \|[F'(x) - F'(\hat{x})]v\| &= \frac{|\lambda|(1+1/n)}{\gamma^{(n-1)/n}}\max_{s\in[a,b]}\int_a^b G(s,t)dt\|x-\hat{x}\| \\ &\leq K_0\|x-\hat{x}\|, \end{aligned}$$

where $K_0 = \frac{|\lambda|(1+1/n)}{\gamma^{(n-1)/n}}N$ and $N = \max_{s\in[a,b]}\int_a^b G(s,t)dt$. Then Assumption 8.2.1 (2) holds for sufficiently small λ.

In the next remarks, we compare our results with the corresponding ones in [15].

Remark 8.7.6. Note that the results in [15] were shown using Assumption 8.2.2 whereas we used weaker Assumption 8.2.1 (2) in this chapter. Next, our result, Proposition 8.2.3, was shown with $3K_0$ replacing K. Therefore, if $3K_0 < K$ (see Example 8.7.3), then our result is tighter. Proposition 8.2.4 was shown with K_0 replacing K. Then, if $K_0 < K$, then our result is tighter. Theorem 8.3.2 was shown with $6K_0$ replacing $2K$. Hence, if $3K_0 < K$, our result is tighter. Similar favorable to us observations are made for Lemma 8.4.1, Theorem 8.4.2 and the rest of the results in [15].

Remark 8.7.7. The results obtained here can also be realized for the operators F satisfying an autonomous differential equation of the form

$$F'(x) = P(F(x)),$$

where $P : X \to X$ is a known continuous operator. Since $F'(\hat{x}) = P(F(\hat{x})) = P(0)$, we can compute K_0 in Assumption 8.2.1 (2) without actually knowing $\hat{x}$. Returning back to Example 8.7.1, we see that we can set $P(x) = x+1$.

References

[1] Argyros, I.K., *Convergence and Application of Newton-type Iterations*, Springer, 2008.

[2] Argyros, I.K., Approximating solutions of equations using Newton's method with a modified Newton's method iterate as a starting point, *Rev. Anal. Numer. Theor. Approx.* 36 (2007), 123–138.

[3] Argyros, I.K., A semilocal convergence for directional Newton methods, *Math. Comput. (AMS)* 80 (2011), 327–343.

[4] Argyros, I.K., and Hilout, S., Weaker conditions for the convergence of Newton's method, *J. Complexity* 28 (2012), 364–387.

[5] Argyros, I.K., Cho, Y.J., Hilout, S., *Numerical methods for equations and its applications*, CRC Press, Taylor and Francis, New York, 2012.

[6] Bakushinskii, A., Seminova, A., On application of generalized discrepancy principle to iterative methods for nonlinear ill-posed problems, *Numer. Funct. Anal. Optim.* 26(2005), 35–48.

[7] Bakushinskii, A., Seminova, A., A posteriori stopping rule for regularized fixed point iterations, *Nonlinear Anal.* 64(2006), 1255–1261.

[8] Bakushinskii, A., Seminova, A., Iterative regularization and generalized discrepancy principle for monotone operato equations, *Numer. Funct. Anal. Optim.* 28(2007), 13–25.

[9] Binder, A., Engl, H.W., Groetsch, C.W., Neubauer, A., Scherzer, O., Weakly closed nonlinear operators and parameter identification in parabolic equations by Tikhonov regularization, *Appl. Anal.* 55(1994), 215–235.

[10] Engl, H.W., Hanke, M., Neubauer, A., *Regularization of Inverse Problems*, Dordrecht, Kluwer, 1993.

[11] Engl, H.W., Kunisch, K., Neubauer, A., Convergence rates for Tikhonov regularization of nonlinear ill-posed problems, *Inverse Problems* 5(1989), 523–540.

[12] Jin, Q., On the iteratively regularized Gauss-Newton method for solving nonlinear ill-posed problems, *Math. Comp.* 69(2000), 1603–1623.

[13] Jin, Q., Hou, Z.Y., On the choice of the regularization parameter for ordinary and iterated Tikhonov regularization of nonlinear ill-posed problems, *Inverse Problems* 13(1997), 815–827.

[14] Jin, Q., Hou, Z.Y., On an a posteriori parameter choice strategy for Tikhonov regularization of nonlinear ill-posed problems, *Numer. Math.* 83(1990), 139–159.

[15] Mahale, P. Nair, M.T., Iterated Lavrentiev regularization for nonlinear ill-posed problems, *ANZIAM* 51(2009), 191–217.

[16] Mahale, P. Nair, M.T., General source conditions for nonlinear ill-posed problems, *Numer. Funct. Anal. Optim.* 28(2007), 111–126.

[17] Scherzer, O., Engl, H.W., Kunisch, K., Optimal a posteriori parameter choice for Tikhonov regularization for solving nonlinear ill-posed problems, *SIAM J. Numer. Anal.* 30(1993), 1796–1838.

[18] Semenova, E.V., Lavrentiev regularization and balancing principle for solving ill-posed problems with monotone operators, *Comput. Methods Appl. Math.* 4(2010), 444–454.

[19] Tautenhahn, U., Lavrentiev regularization of nonlinear ill-posed problems, *Vietnam J. Math.* 32(2004), 29–41.

[20] Tautenhahn, U., On the method of Lavrentiev regularization for nonlinear ill-posed problems, *Inverse Problems* 18(2002), 191–207.

[21] Tautenhahn, U., Jin, Q., Tikhonov regularization and a posteriori rule for solving nonlinear ill-posed problems, *Inverse Problems* 19(2003), 1–21.

Chapter 9

A Semilocal Convergence for a Uniparametric Family of Efficient Secant-Like Methods

9.1. Introduction

Let $U(x,r)$ and $\overline{U}(x,r)$ stand, respectively, for the open and closed ball in $\mathcal{X}$ with center $x \in \mathcal{X}$ and radius $r > 0$. Denote by $\mathcal{L}(\mathcal{X},\mathcal{Y})$ the space of bounded linear operators from $\mathcal{X}$ into $\mathcal{Y}$.

In this chapter we are concerned with the problem of approximating a locally unique solution x^* of nonlinear equation

$$F(x) = 0, \tag{9.1.1}$$

where F is a Fráchet-differentiable operator defined on a non-empty convex subset $\mathcal{D}$ of a Banach space $\mathcal{X}$ with values in a Banach space $\mathcal{Y}$.

Many problems from computational sciences, physics and other disciplines can be taken in the form of equation (9.1.1) using Mathematical Modelling [5, 6, 8, 9, 12, 22, 25]. The solution of these equations can rarely be found in closed form. That is why the solution methods for these equations are iterative. In particular, the practice of numerical analysis for finding such solutions is essentially connected to variants of Newton's method [5, 6, 9, 12, 19, 21, 22, 24, 25]. The study about the convergence of iterative procedures is usually focussed on two types: semilocal and local convergence analysis. The semilocal convergence is, based on the information around an initial point, to give criteria ensuring the convergence of iterative procedure; while the local one is, based on the information around a solution, to find estimates of the radii of convergence balls. There are a lot of studies on the weakness and/or extension of the hypothesis made on the underlying operators; see for example [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 17, 15, 16, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27] and the references therein.

Ezquerro and Rubio used in [17] the uniparametric family of secant-like methods defined by

$$\begin{cases} x_{-1}, x_0 \quad \text{given in } \mathcal{D}, \\ y_n = \mu x_n + (1-\mu) x_{n-1}, \qquad \mu \in [0,1], \\ x_{n+1} = x_n - B_n^{-1} F(x_n), \quad B_n = [y_n, x_n; F], \quad \text{for each} \quad n = 0,1,\ldots \end{cases} \tag{9.1.2}$$

and the method of recurrent relations to generate a sequence $\{x_n\}$ approximating x^*. Here, $[z,w;F]$ for each $z,w \in \mathcal{D}$ is a divided difference of order one, which is a bounded linear operator such that [4, 5, 7, 9, 19, 22, 25]

$$[z,w;F] : \mathcal{D} \to \mathcal{Y} \quad \text{and} \quad [z,w;F](z-w) = F(z) - F(w). \tag{9.1.3}$$

Secant-like method (9.1.2) can be considered as a combination of the secant and Newton's method. Indeed, if $\mu = 0$ we obtain the secant method and if $\mu = 1$ we get Newton's method provided that F' is Frchet-differentiable on $\mathcal{D}$, since, then $x_n = y_n$ and $[y_n, x_n; F] = F'(x_n)$.

It was shown in [15, 16] that the R-order of convergence is at least $\dfrac{1+\sqrt{5}}{2}$ for $\lambda \in [0,1)$, the same as that of the secant method. Later in [12] another uniparametric family of secant-like methods defined by

$$\begin{cases} x_{-1}, x_0 \quad \text{given in } \mathcal{D}, \\ y_n = \lambda x_n + (1-\lambda) x_{n-1}, \quad \lambda \geq 1, \\ x_{n+1} = x_n - A_n^{-1} F(x_n), \quad A_n = [y_n, x_{n-1}; F] \quad \text{for each} \quad n = 0,1,\ldots \end{cases} \tag{9.1.4}$$

was studied. It was shown that there exists $\lambda_0 \geq 2$ that the R-order of convergence is at least $\dfrac{1+\sqrt{5}}{2}$ if $\lambda \in [1, \lambda_0]$ and $\lambda \neq 2$ and if $\lambda = 2$ the R-order of convergence is quadratic. Note that if $\lambda = 1$ we obtain the secant method, whereas if $\lambda = 2$ we obtain the Kurchatov method [9, 12, 19, 25].

We present a semilocal convergence analysis for secant like method (9.1.2) using our idea of recurrent functions instead of recurrent relations and tighter majorizing sequences. This way our analysis provided the following advantages (A) over the work in [12] under the same computational cost:

(A_1) Weaker sufficient convergence conditions,

(A_2) Tighter estimates on the distances $\|x_{n+1} - x_n\|$ and $\|x_n - x^*\|$ for each $n = 0,1,\ldots$,

(A_3) At least as precise information on the location of the solution and

(A_4) The results are presented in affine invariant form, whereas the ones in [12] are given in a non-affine invariant forms. The advantages of affine versus non-affine results have been explained in [4, 5, 7, 9, 19, 22, 25]

Our hypotheses for the semilocal convergence of secant-like method (9.1.4) are:

(C_1) There exists a divided difference of order one $[z,w;F] \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ satisfying (9.1.3),

(C_2) There exist $x_0 \in \mathcal{D}$, $\eta \geq 0$ such that $A_0^{-1} \in \mathcal{L}(\mathcal{Y},\mathcal{X})$ and $\|A_0^{-1}F(x_0)\| \leq \eta$,

(C_3) There exist $x_{-1}, x_0 \in \mathcal{D}$ and $c \geq 0$ such that

$$\|x_0 - x_{-1}\| \leq c,$$

(C_4) There exists $K > 0$ such that

$$\|A_0^{-1}([x,y;F]-[v,w;F])\| \leq K(\|x-v\|+\|y-w\|) \quad \text{for each} \quad x,y,v,w \in \mathcal{D}.$$

We shall denote by (C) conditions (C_1)–(C_4). In view of (C_4) there exist $H_0, H_1, H > 0$ such that

(C_5) $\|A_0^{-1}([x_1,x_0;F]-A_0)\| \leq H_0(\|x_1-y_0\|+\|x_0-x_{-1}\|)$,

(C_6) $\|A_0^{-1}(A_1-A_0)\| \leq H_1(\|y_1-y_0\|+\|x_0-x_{-1}\|)$ and

(C_7) $\|A_0^{-1}([x,y;F]-A_0)\| \leq H(\|x-y_0\|+\|y-x_{-1}\|)$ for each $x,y \in \mathcal{D}$.

Clearly

$$H_0 \leq H_1 \leq H \leq K \tag{9.1.5}$$

hold in general and $\frac{K}{H}, \frac{H}{H_1}$ can be arbitrarily large [5, 6, 9]. Note that (C_5), (C_6), (C_7) are not additional to (C_4) hypotheses. In practise the computation of K requires the computation of H_0, H_1 and H. It also follows from (C_4) that F is differentiable [5, 6, 19, 21].

The chapter is organized as follows. In Section 9.2. we show that under the same hypotheses as in [18] and using recurrent relations, we obtain an at least as precise information on the location of the solution. Section 9.3. contains the semilocal convergence analysis using weaker hypotheses and recurrent functions. We also show the advantages (A). The results are also extended to cover the case of equations with nondifferentiable operators. Numerical examples are presented in the concluding Section 9.4..

9.2. Semilocal Convergence Using Recurrent Relations

As in [12] let us define sequences $\{a_n\}$ and $\{b_n\}$ for each $n = 0, 1, \ldots$ by

$$a_{-1} = \frac{\eta}{c+\eta}, \quad b_{-1} = \frac{Kc^2}{c+\eta},$$

$$a_n = f(a_{n-1})g(a_{n-1})b_{n-1}, \quad b_n = f(a_{n-1})^2 a_{n-1} b_{n-1}$$

and functions f, g on $[0,1)$ by

$$f(t) = \frac{1}{1-t} \quad \text{and} \quad g(t) = (2-\lambda) + \lambda f(t)t.$$

Next, we present the main result in this section in affine invariant form.

Theorem 9.2.1. *Under the (C) hypotheses further suppose that*

$$\overline{U}(x_0,R) \subseteq \mathcal{D}$$

and for $\lambda \in [1,\lambda_0]$

$$a_{-1} < \frac{3-\sqrt{5}}{2}, \quad b_{-1} < \frac{a_{-1}(1-a_{-1})^2}{2(1-a_{-1})-\lambda(1-2a_{-1})}$$

where

$$R = \frac{1-a_0}{1-2a_0}\lambda\eta$$

and

$$\lambda_0 \in \left[2, \frac{2c}{c-\eta}\right).$$

Then, sequence $\{x_n\}$ *generated by secant-like method (9.1.4) is well defined, remains in* $\overline{U}(x_0,R)$ *for each* $n = 0,1,2,\ldots$ *and converges to a solution* $x^* \in \overline{U}(x_0,R)$ *of equation* $F(x) = 0$. *Moreover, the following estimates hold*

$$\|x_{n+1}-x_n\| \le f(a_{n-1})a_{n-1}\|x_n - x_{n-1}\|$$

and

$$\|x_n - x^*\| \le \frac{(f(a_0)a_0)^n}{1-f(a_0)a_0}\|x_1 - x_0\|.$$

Furthermore, the solution x^* *is unique in* $\mathcal{D}_0 = U(x_0,\sigma_0)\cap D$, *where* $\sigma_0 = \dfrac{1}{H} - \lambda c - R$, *provided that*

$$R < \frac{1}{2}\left(\frac{1}{H} - \lambda c\right) = R_0.$$

Proof. The proof with the exception of the uniqueness part is given in Theorem 3 [12] if we use $A_0^{-1}F$ instead of F and set $b = 1$, where $\|A_0^{-1}\| \le b$.

To prove the uniqueness of the solution, let us assume $y^* \in \mathcal{D}_0$ is a solution of $F(x) = 0$. Let $L = [y^*,x^*;F]$. Then, using (C_7) and the definition of σ_0 we get in turn that

$$\|A_0^{-1}(L-A_0)\| \le H(\|y^*-y_0\| + \|x^*-x_{-1}\|) < H(\sigma_0+\lambda c+R) = 1. \tag{9.2.1}$$

It follows from (9.2.1) and the Banach lemma on invertible operators [4, 5, 6, 9, 19, 22, 25] that $L^{-1} \in \mathcal{L}(\mathcal{Y},\mathcal{X})$. Using the identity $0 = F(y^*) - F(x^*) = L(y^*-x^*)$ we deduce that $x^* = y^*$. That completes the proof of the Theorem. ■

Remark 9.2.2. *If* $K = H$, *Theorem 9.2.1 reduces to Theorem* 3 *in [12]. Otherwise, i. e. if* $H < K$, *then our Theorem 9.2.1 constitutes an improvement over Theorem* 3, *since*

$$\sigma < \sigma_0 \tag{9.2.2}$$

and

$$R_0 < R_1, \tag{9.2.3}$$

where

$$\sigma = \frac{1}{K} - \lambda c - R$$

and

$$R_0 = \frac{1}{2}\left(\frac{1}{K} - \lambda c\right)$$

where given in [12] (for $b = 1$). Hence, (9.2.2) and (9.2.3) justify our claim for this section made in the Introduction of this chapter.

9.3. Semilocal Convergence Using Recurrent Functions

We present the semilocal convergence of secant-like methods. First, we need some auxiliary results on majorizing sequences for secant-like method.

Lemma 9.3.1. *Let $c \geq 0$, $\eta > 0$, $H > 0$, $K > 0$ and $\lambda \geq 1$. Set $t_{-1} = 0$, $t_0 = c$ and $t_1 = c + \eta$. Define scalar sequences $\{q_n\}$, $\{t_n\}$, $\{\alpha_n\}$ for each $n = 0, 1, \ldots$ by*

$$q_n = H\lambda(t_{n+1} + t_n - c), \tag{9.3.1}$$

$$t_{n+2} = t_{n+1} + \frac{K(t_{n+1} - t_n + \lambda(t_n - t_{n-1}))}{1 - q_n}(t_{n+1} - t_n),$$

$$\alpha_n = \frac{K(t_{n+1} - t_n + \lambda(t_n - t_{n-1}))}{1 - q_n}, \tag{9.3.2}$$

functions f_n on $[0, 1)$ for each $n = 1, 2, \ldots$ by

$$f_n(t) = K(t^n + \lambda t^{n-1})\eta + H\lambda((1 + t + \cdots + t^{n+1})\eta + (1 + t + \cdots + t^n)\eta + c) - 1 \tag{9.3.3}$$

and polynomial p on $[0, 1)$ by

$$p(t) = H\lambda t^3 + (H\lambda + K)t^2 + K(\lambda - 1)t - \lambda K. \tag{9.3.4}$$

Denote by α the only root of polynomial p in $(0, 1)$. Suppose that

$$0 \leq \alpha_0 \leq \alpha \leq \frac{1 - H\lambda(c + 2\eta)}{1 - H\lambda c}. \tag{9.3.5}$$

Then, sequence $\{t_n\}$ is non-decreasing, bounded from above by t^{**} defined by

$$t^{**} = \frac{\eta}{1 - \alpha} + c \tag{9.3.6}$$

and converges to its unique least upper bound t^* which satisfies

$$c + \eta \leq t^* \leq t^{**}. \tag{9.3.7}$$

Moreover, the following estimates are satisfied for each $n = 0, 1, 2, \ldots$

$$0 \leq t_{n+1} - t_n \leq \alpha^n \eta \tag{9.3.8}$$

and

$$t^* - t_n \leq \frac{\alpha^n \eta}{1 - \alpha}. \tag{9.3.9}$$

Proof. We shall first show that polynomial p has roots in $(0,1)$. Indeed, we have $p(0) = -\lambda K < 0$ and $p(1) = 2H\lambda > 0$. Using the intermediate value theorem we deduce that there exists at least one root of p in $(0,1)$. Moreover $p'(t) > 0$. Hence p crosses the positive axis only once. Denote by α the only root of p in $(0,1)$. It follows from (9.3.1) and (9.3.2) that estimate (9.3.8) is certainly satisfied if

$$0 \leq \alpha_n \leq \alpha. \tag{9.3.10}$$

Estimate (9.3.10) is true by (9.3.5) for $n = 0$. Then, we have by (9.3.1) that

$$\begin{aligned} t_2 - t_1 &\leq \alpha(t_1 - t_0) \Rightarrow t_2 \leq t_1 + \alpha(t_1 - t_0) \Rightarrow t_2 \leq \eta + t_0 + \alpha\eta \\ &= c + (1+\alpha)\eta = c + \frac{1-\alpha^2}{1-\alpha}\eta < t^{**}. \end{aligned}$$

Suppose that

$$t_{k+1} - t_k \leq \alpha^k \eta \quad \text{and} \quad t_{k+1} \leq c + \frac{1-\alpha^{k+1}}{1-\alpha}\eta \quad \text{for each} \quad k \leq n. \tag{9.3.11}$$

Estimate (9.3.10) shall be true for $k+1$ replacing n if

$$0 \leq \alpha_{k+1} \leq \alpha \tag{9.3.12}$$

or

$$f_k(\alpha) \leq 0, \tag{9.3.13}$$

where f_k is defined by (9.3.3). We need a relationship between two consecutive recurrent functions f_k for each $k = 1,2\ldots$ Using (9.3.3) and (9.3.4) we deduce that

$$f_{k+1}(\alpha) = f_k(\alpha) + p(\alpha)\alpha^{k-1}\eta = f_k(\alpha), \tag{9.3.14}$$

since $p(\alpha) = 0$. Define function f_∞ on $(0,1)$ by

$$f_\infty(t) = \lim_{k\to+\infty} f_k(t). \tag{9.3.15}$$

Then, we get from (9.3.3) and (9.3.15) that

$$f_\infty(\alpha) = H\lambda\left(\frac{2\eta}{1-\alpha} + c\right) - 1. \tag{9.3.16}$$

Hence, by (9.3.14)–(9.3.16), (9.3.13) is satisfied if

$$f_\infty(\alpha) \leq 0 \tag{9.3.17}$$

which is true by (9.3.5). The induction for (9.3.8) is complete. That is sequence $\{t_n\}$ is non-decreasing, bounded from above by t^{**} given by (9.3.6) and as such it converges to some t^* which satisfies (9.3.7). Estimate (9.3.9) follows from (9.3.8) by using standard majorization techniques [4, 5, 6, 9, 19, 22, 25]. The proof of Lemma 9.3.1 is complete. ■

Lemma 9.3.2. *Let $c \geq 0$, $\eta > 0$, $H_0 > 0$, $H_1 > 0$, $H > 0$, $K > 0$ and $\lambda \geq 1$. Set $s_{-1} = 0$, $s_0 = c$, $s_1 = c + \eta$. Define scalar sequences $\{s_n\}$, $\{b_n\}$ for each $n = 1, 2, \ldots$ by*

$$\begin{cases} s_2 = s_1 + \dfrac{H_0(s_1 - s_0 + \lambda(s_0 - s_{-1}))}{1 - H_1\lambda(s_1 + s_0 - c)}(s_1 - s_0), \\ s_{n+2} = s_{n+1} + \dfrac{K(s_{n+1} - s_n + \lambda(s_n - s_{n-1}))}{1 - H\lambda(s_{n+1} + s_n - c)}(s_{n+1} - s_n), \end{cases} \tag{9.3.18}$$

$$\begin{cases} b_1 = \dfrac{H_0(s_1 - s_0 + \lambda(s_0 - s_{-1}))}{1 - H_1\lambda(s_1 + s_0 - c)}, \\ b_n = \dfrac{K(s_{n+1} - s_n + \lambda(s_n - s_{n-1}))}{1 - H\lambda(s_{n+1} + s_n - c)}, \end{cases} \tag{9.3.19}$$

and functions g_n on $[0, 1)$ by

$$\begin{aligned} g_n(t) = {} & K(t + \lambda)t^{n-1}(s_2 - s_1) + H\lambda t \left(2s_1 + \frac{1 - t^{n+1}}{1 - t}(s_2 - s_1) + \frac{1 - t^n}{1 - t}(s_2 - s_1) \right) \\ & -(1 + H\lambda c)t. \end{aligned} \tag{9.3.20}$$

Suppose that

$$0 \leq b_1 \leq \alpha \leq \frac{1 - H\lambda(2s_2 - c)}{1 - H\lambda(2s_1 - c)}, \tag{9.3.21}$$

where α is defined in Lemma 9.3.1. Then, sequence $\{s_n\}$ is non-decreasing, bounded from above by s^{**} defined by

$$s^{**} = c + \eta + \frac{s_2 - s_1}{1 - \alpha} \tag{9.3.22}$$

and converges to its unique least upper bound s^* which satisfies

$$c + \eta \leq s^* \leq s^{**}. \tag{9.3.23}$$

Moreover, the following estimates are satisfied for each $n = 1, 2, \ldots$

$$0 \leq s_{n+2} - s_{n+1} \leq \alpha^n (s_2 - s_1). \tag{9.3.24}$$

Proof. We shall show using induction that

$$0 \leq b_n \leq \alpha. \tag{9.3.25}$$

Estimate (9.3.25) is true for $n = 0$ by (9.3.21). Then, we have by (9.3.18) that

$$\begin{aligned} 0 & \leq s_3 - s_2 \leq \alpha(s_2 - s_1) \Rightarrow s_3 \leq s_2 + \alpha(s_2 - s_1) \Rightarrow \\ & \Rightarrow s_3 \leq s_2 + (1 + \alpha)(s_2 - s_1) - (s_2 - s_1) \Rightarrow \\ & \Rightarrow s_3 \leq s_1 + \frac{1 - \alpha^2}{1 - \alpha}(s_2 - s_1) \leq s^{**}. \end{aligned} \tag{9.3.26}$$

Suppose (9.3.25) holds for each $n \leq k$. Then using (9.3.18) we get that

$$0 \leq s_{k+2} - s_{k+1} \leq \alpha^k (s_2 - s_1) \tag{9.3.27}$$

and

$$s_{k+2} \leq s_1 + \frac{1+\alpha^{k+1}}{1-\alpha}(s_2 - s_1). \tag{9.3.28}$$

Estimate (9.3.25) shall be satisfied if

$$g_k(\alpha) \leq 0. \tag{9.3.29}$$

Using (9.3.20) we get the following relationship between two consecutive recurrent functions g_k:

$$g_{k+1}(\alpha) = g_k(\alpha) + p(\alpha)\alpha^{k-1}(s_2 - s_1) = g_k(\alpha). \tag{9.3.30}$$

Define function g_∞ on $[0,1)$ by

$$g_\infty(t) = \lim_{k\to+\infty} g_k(t). \tag{9.3.31}$$

Then, we get from (9.3.20) that

$$g_\infty(\alpha) = 2\alpha H\lambda \left[s_1 + \frac{s_2 - s_1}{1-\alpha} \right] - \alpha(1 + H\lambda c). \tag{9.3.32}$$

Then, (9.3.29) is satisfied if

$$g_\infty(\alpha) \leq 0, \tag{9.3.33}$$

which is true by the choice of α and the right hand side inequality in hypothesis (9.3.21). The induction for (9.3.25) (i. e. (9.3.24)) is complete. The rest of the proof as identical to Lemma 9.3.1 is omitted. The proof is complete. ■

Remark 9.3.3. *(a) Let us consider an interesting choice for λ. Let $\lambda = 1$ (secant method). Then, using (9.3.4) and (9.3.5) we have that*

$$\alpha = \frac{2K}{K + \sqrt{K^2 + 4HK}} \tag{9.3.34}$$

and

$$\frac{K(c+\eta)}{1 - H(c+\eta)} \leq \alpha \leq \frac{1 - H(c+2\eta)}{1 - Hc}. \tag{9.3.35}$$

The corresponding condition for the secant method is given by [6, 9, 18, 21]:

$$Kc + 2\sqrt{K\eta} \leq 1. \tag{9.3.36}$$

Condition (9.3.35) can be weaker than (9.3.36) (see also the numerical examples at the end of the chapter). Moreover, the majorizing sequence $\{u_n\}$ for the secant method related to (9.3.36) is given by

$$\begin{cases} u_{-1} = 0, \quad u_0 = c, \quad u_1 = c + \eta, \\ u_{n+2} = u_{n+1} + \dfrac{K(u_{n+1} - u_{n-1})}{1 - K(u_{n+1} + u_n - c)}(u_{n+1} - u_n). \end{cases} \tag{9.3.37}$$

A simple inductive argument shows that if $H < K$, then for each $n = 2,3,\ldots$

$$t_n < u_n, \quad t_{n+1} - t_n < u_{n+1} - u_n \quad \text{and} \quad t^* \le u^* = \lim_{n\to+\infty} u_n. \tag{9.3.38}$$

(b) *The majorizing sequence $\{v_n\}$ used in [12] is essentially given by*

$$\begin{cases} v_{-1} &= 0, \quad v_0 = c, \quad v_1 = c + \eta, \\ v_{n+2} &= v_{n+1} + \dfrac{K(v_{n+1} - v_n + \lambda(v_n - v_{n-1}))}{1 - K\lambda(v_{n+1} + v_n - c)}(v_{n+1} - v_n). \end{cases} \tag{9.3.39}$$

Then, again we have

$$t_n < v_n, \quad t_{n+1} - t_n < v_{n+1} - v_n \quad \text{and} \quad t^* \le v^* = \lim_{n\to+\infty} v_n. \tag{9.3.40}$$

Moreover, our sufficient convergence conditions can be weaker than [12].

(c) *Clearly, iteration $\{s_n\}$ is tighter than $\{t_n\}$ and we have as in (9.3.40) than for $H_0 < K$ or $H_1 < H$*

$$s_n < t_n, \quad s_{n+1} - s_n < t_{n+1} - t_n \quad \text{and} \quad s^* = \lim_{n\to+\infty} s_n < t^*. \tag{9.3.41}$$

Next, we present obvious and useful extensions of Lemma 9.3.1 and Lemma 9.3.2, respectively.

Lemma 9.3.4. *Let $N = 0,1,2,\ldots$ be fixed. Suppose that*

$$t_1 \le t_2 \le \cdots \le t_N \le t_{N+1}, \tag{9.3.42}$$

$$\frac{1}{H\lambda} > t_{N+1} - t_N + \lambda(t_N - t_{N-1}) \tag{9.3.43}$$

and

$$0 \le \alpha_N \le \alpha \le \frac{1 - H\lambda(t_N - t_{N-1} + 2(t_{N+1} - t_N))}{1 - H\lambda(t_N - t_{N-1})} \tag{9.3.44}$$

Then, sequence $\{t_n\}$ generated by (9.3.2) is nondecreasing, bounded from above by t^{**} and converges to t^* which satisfies $t^* \in [t_{N+1}, t^*]$. Moreover, the following estimates are satisfied for each $n = 0,1,\ldots$

$$0 \le t_{N+n+1} - t_{N+n} \le \alpha^n (t_{N+1} - t_N) \tag{9.3.45}$$

and

$$t^* - t_{N+n} \le \frac{\alpha^n}{1-\alpha}(t_{N+1} - t_N). \tag{9.3.46}$$

Lemma 9.3.5. *Let $N = 1,2,\ldots$ be fixed. Suppose that*

$$s_1 \le s_2 \le \cdots \le s_N \le s_{N+1}, \tag{9.3.47}$$

$$\frac{1}{H\lambda} > s_{N+1} - s_N + \lambda(s_N - s_{N-1}) \tag{9.3.48}$$

and

$$0 \le b_N \le \alpha \le \frac{1 - H\lambda(2s_{N+1} - s_{N-1})}{1 - H\lambda(2s_N - s_{N-1})}. \tag{9.3.49}$$

Then, sequence $\{s_n\}$ generated by (9.3.18) is nondecreasing, bounded from above by s^{**} and converges to s^* which satisfies $s^* \in [s_{N+1}, s^*]$. Moreover, the following estimates are satisfied for each $n = 0, 1, \ldots$

$$0 \le s_{N+n+1} - s_{N+n} \le \alpha^n (s_{N+1} - s_N) \tag{9.3.50}$$

and

$$s^* - s_{N+n} \le \frac{\alpha^n}{1-\alpha}(s_{N+1} - s_N). \tag{9.3.51}$$

Next, we present the following semilocal convergence result for secant-like method under the (C) conditions.

Theorem 9.3.6. *Suppose that the (C), Lemma 9.3.1 (or Lemma 9.3.4) conditions and*

$$U = \overline{U}(x_0, (2\lambda - 1)t^*) \subseteq \mathcal{D} \tag{9.3.52}$$

hold. Then, sequence $\{x_n\}$ generated by secant-like method is well defined, remains in U for each $n = -1, 0, 1, 2, \ldots$ and converges to a solution $x^* \in \overline{U}(x_0, t^* - c)$ of equation $F(x) = 0$. Moreover, the following estimates are satisfied for each $n = 0, 1, \ldots$

$$\|x_{n+1} - x_n\| \le t_{n+1} - t_n \tag{9.3.53}$$

and

$$\|x_n - x^*\| \le t^* - t_n. \tag{9.3.54}$$

Furthermore, if there exists $T \ge t^* - c$ such that

$$\overline{U}(x_0, r) \subseteq \mathcal{D} \tag{9.3.55}$$

and

$$H(T + t^* + (\lambda - 1)c) < 1, \tag{9.3.56}$$

then, the solution x^* is unique in $\overline{U}(x_0, T)$.

Proof. We use mathematical induction to prove that

$$\|x_{k+1} - x_k\| \le t_{k+1} - t_k \tag{9.3.57}$$

and

$$\overline{U}(x_{k+1}, t^* - t_{k+1}) \subseteq \overline{U}(x_k, t^* - t_k) \tag{9.3.58}$$

for each $k = -1, 0, 1, \ldots$. Let $z \in \overline{U}(x_0, t^* - t_0)$. Then we obtain that

$$\begin{aligned} \|z - x_{-1}\| &\le \|z - x_0\| + \|x_0 - x_{-1}\| \le t^* - t_0 + c = t^* \\ &= t^* - t_{-1}, \end{aligned}$$

which implies $z \in \overline{U}(x_{-1}, t^* - t_{-1})$. Let also $w \in \overline{U}(x_0, t^* - t_1)$. We get that

$$\begin{aligned} \|w - x_0\| &\leq \|w - x_1\| + \|x_1 - x_0\| \leq t^* - t_1 + t_1 - t_0 \\ &= t^* - t_0, \end{aligned}$$

hence, $w \in \overline{U}(x_0, t^*, t_0)$. Note that $\|x_{-1} - x_0\| \leq c = t_0 - t_{-1}$ and $\|x_1 - x_0\| = \|A_0^{-1} F(x_0)\| \leq \eta = t_1 - t_0 < t^*$. That is $x_1 \in \overline{U}(x_0, t^*) \subseteq \mathcal{D}$. Hence, estimates (9.3.57) and (9.3.58) hold for $k = -1$ and $k = 0$. Suppose that (9.3.57) and (9.3.58) hold for all $n \leq k$. Then, we obtain that

$$\begin{aligned} \|x_{k+1} - x_0\| &\leq \sum_{i=1}^{k+1} \|x_i - x_{i-1}\| \leq \sum_{i=1}^{k+1} (t_i - t_{i-1}) \\ &= t_{k+1} - t_0 = t^* - c \leq t^* \end{aligned}$$

and

$$\begin{aligned} \|y_k - x_0\| &= \|\lambda x_k + (1-\lambda)x_{k-1} - x_0\| = \|\lambda(x_k - x_0) + (1-\lambda)(x_{k-1} - x_0)\| \\ &\leq \lambda\|x_k - x_0\| + (\lambda - 1)\|x_{k-1} - x_0\| \\ &\leq \lambda t^* + (\lambda - 1)t^* = (2\lambda - 1)t^*. \end{aligned}$$

Hence, x_{k+1}, $y_k \in \overline{U}(x_0, t^*)$.

Using (C_7), Lemma 9.3.1 and the introduction hypotheses, we get that

$$\begin{aligned} \|A_0^{-1}(A_{k+1} - A_0)\| &\leq H(\|y_{k+1} - y_0\| + \|x_k - x_{-1}\|) \\ &\leq H(\lambda\|x_{k+1} - x_0\| + |1 - \lambda|\|x_k - x_{-1}\| + \|x_k - x_{-1}\|) \\ &\leq H\lambda(\|x_{k+1} - x_0\| + \|x_k - x_0\| + \|x_0 - x_{-1}\|) \qquad (9.3.59) \\ &\leq H\lambda(t_{k+1} - t_0 + t_k - t_0 + c) \\ &= H\lambda(t_{k+1} + t_k - c) < 1. \end{aligned}$$

It follows from (9.3.59) and the Banach lemma on invertible operators [4, 5, 6, 9, 19, 22, 25] that A_{k+1}^{-1} exists and

$$\|A_{k+1}^{-1} A_0\| \leq (1 - H\lambda(t_{k+1} + t_k - c))^{-1}. \qquad (9.3.60)$$

In view of (9.1.4), we obtain the identity

$$\begin{aligned} F(x_{k+1}) &= F(x_{k+1}) - F(x_k) - [y_k, x_{k-1}; F](x_{k+1} - x_k) \\ &= ([x_{k+1}, x_k; F] - [y_k, x_{k-1}; F])(x_{k+1} - x_k). \qquad (9.3.61) \end{aligned}$$

Using (9.1.4), (9.3.16) and the induction hypotheses we get in turn that

$$\begin{aligned} \|A_0^{-1} F(x_{k+1})\| &\leq K(\|x_{k+1} - y_k\| + \|x_k - x_{k-1}\|)\|x_{k+1} - x_k\| \\ &\leq K(\|x_{k+1} - x_k\| + \lambda\|x_k - x_{k-1}\|)\|x_{k+1} - x_k\| \qquad (9.3.62) \\ &\leq K(t_{k+1} - t_k + \lambda(t_k - t_{k+1}))(t_{k+1} - t_k). \end{aligned}$$

It now follows from (9.1.4), (9.3.1), (9.3.61) and (9.3.62) that

$$\begin{aligned} \|x_{k+2}-x_{k+1}\| &\leq \|A_{k+1}^{-1}A_0\|\|A_0^{-1}F(x_{k+1})\| \\ &\leq \frac{K(t_{k+1}-t_k+\lambda(t_k-t_{k-1}))(t_{k+1}-t_k)}{1-H\lambda(t_{k+1}+t_k-c)} \\ &= t_{k+2}-t_{k+1}, \end{aligned} \tag{9.3.63}$$

which completes the induction for (9.3.57). Moreover, let $v \in \overline{U}(x_{k+2}, t^*-t_{k+2})$. Then, we get that

$$\begin{aligned} \|v-x_{k+1}\| &\leq \|v-x_{k+2}\|+\|x_{k+2}-x_{k+1}\| \\ &\leq t^*-t_{k+2}+t_{k+2}-t_{k+1} = t^*-t_{k+1}, \end{aligned}$$

which implies $v \in \overline{U}(x_{k+1}, t^*-t_{k+1})$. The induction for (9.3.58) is complete.

Lemma 9.3.1 implies that $\{t_k\}$ is a complete sequence. It follows from (9.3.57) and (9.3.58) that $\{x_k\}$ is a complete sequence in a Banach space $\mathcal{X}$ and as such it converges to some $x^* \in \overline{U}(x_0, t^*-c)$ (since $\overline{U}(x_0, t^*-c)$ is a closed set). By letting $k \to +\infty$ in (9.3.62) we obtain $F(x^*) = 0$. Furthermore, estimate (9.3.54) follows from (9.3.53) by using standard majorization techniques [5, 6, 8, 9, 19, 22, 25]. To show the uniqueness part, let $y^* \in \overline{U}(x_0, T)$ be such that $F(y^*) = 0$. We have that

$$\begin{aligned} \|A_0^{-1}([y^*,x^*;F]-A_0)\| &\leq H(\|y^*-y_0\|+\|x^*-x_{-1}\|) \\ &\leq H(\|y^*-x_0\|+(\lambda-1)\|x_0-x_{-1}\| \\ &\quad +\|x^*-x_0\|+\|x_0-x_{-1}\|) \\ &\leq (R_0+t^*+(\lambda-1)c) < 1. \end{aligned} \tag{9.3.64}$$

It follows from (9.3.64) and the Banach lemma on invertible operators that $[y^*,x^*;F]^{-1}$ exists. Then, using the identity $0 = F(y^*)-F(x^*) = [y^*,x^*;F](y^*-x^*)$, we deduce that $x^* = y^*$. The proof of Theorem 9.3.6 is complete. ■

Remark 9.3.7. (a) *The limit point t^* can be replaced in Theorem 9.3.6 by t^{**} given in closed form by (9.3.6).*

(b) *It follows from the proof of Theorem 9.3.6 that $\{s_n\}$ is also a majorizing sequence for $\{x_n\}$. Hence, Lemma 9.3.2 (or Lemma 9.3.5), $\{s_n\}$, s^* can replace Lemma 9.3.1 (or Lemma 9.3.4) $\{t_n\}$, t^* in Theorem 9.3.6.*

Hence we arrive at:

Theorem 9.3.8. *Suppose that the (C) conditions, Lemma 9.3.2 (or Lemma 9.3.5) and*

$$U = \overline{U}(x_0, (2\lambda-1)s^*) \subseteq \mathcal{D}$$

hold. Then sequence $\{x_n\}$ generated by secant-like method is well defined, remains in U for each $n = -1, 0, 1, 2, \ldots$ and converges to a solution $x^* \in U(x_0, s^*-c)$ of equation $F(x) = 0$. Moreover, the following estimates are satisfied for each $n = 0, 1, \ldots$

$$\|x_{n+1}-x_n\| \leq s_{n+1}-s_n$$

and

$$\|x_n - x^*\| \leq s^* - s_n.$$

Furthermore, if there exists $T \geq s^* - c$ such that

$$\overline{U}(x_0, r) \subseteq \mathcal{D}$$

and

$$H(T + s^* + (\lambda - 1)c) < 1,$$

then, the solution x^* is unique in $\overline{U}(x_0, T)$.

Let us consider the equation

$$F(x) + G(x) = 0, \tag{9.3.65}$$

where F is a before and $G : \mathcal{D} \to \mathcal{Y}$ is continuous. The corresponding secant-like method is given by

$$x_{n+1} = x_n - A_n^{-1}(F(x_n) + G(x_n)) \quad \text{for each} \quad n = 0, 1, 2 \ldots, \tag{9.3.66}$$

where x_0 is an initial guess.

Suppose that

(C_8)

$$\|A_0^{-1}(G(x) - G(y))\| \leq M\|x - y\| \quad \text{for each} \quad x, y \in \mathcal{D}, \tag{9.3.67}$$

and

(C_9)

$$\|A_0^{-1}(G(x_1) - G(x_0))\| \leq M_0\|x_1 - x_0\|. \tag{9.3.68}$$

Clearly,

$$M_0 \leq M \tag{9.3.69}$$

holds and $\dfrac{M}{M_0}$ can be arbitrarily large [4, 5, 6, 8, 9].

We shall denote by (C^*) the conditions (C) and (C_8), (C_9). Then, we can present the corresponding result along the same lines as in Lemma 9.3.1, Lemma 9.3.2, Lemma 9.3.4, Lemma 9.3.5, Theorem 9.3.6 and Theorem 9.3.8. However, we shall only present the results corresponding to Lemma 9.3.2 and Theorem 9.3.8, respectively. The rest combination of results can be given in an analogous way.

Lemma 9.3.9. *Let $c \geq 0$, $\eta > 0$, $H_0 > 0$, $H_1 > 0$, $H > 0$, $M_0 > 0$, $M > 0$, $K > 0$ and $\lambda \geq 1$. Set $\gamma_{-1} = 0$, $\gamma_0 = c$, $\gamma_1 = c + \eta$. Define scalar sequences $\{\gamma_n\}$, $\{\delta_n\}$ by*

$$\begin{cases} \gamma_2 &= \gamma_1 + \dfrac{H_0(\gamma_1 - \gamma_0 + \lambda(\gamma_0 - \gamma_{-1})) + M_0}{1 - H_1\lambda(\gamma_1 + \gamma_0 - c)}(\gamma_1 - \gamma_0), \\ \gamma_{n+2} &= \gamma_{n+1} + \dfrac{K(\gamma_{n+1} - \gamma_n + \lambda(\gamma_n - \gamma_{n-1})) + M}{1 - H\lambda(\gamma_{n+1} + \gamma_n - c)}(\gamma_{n+1} - \gamma_n), \end{cases}$$

$$\begin{cases} \delta_1 & = & \dfrac{H_0(\gamma_1-\gamma_0+\lambda(\gamma_0-\gamma_{-1}))+M_0}{1-H_1\lambda(\gamma_1+\gamma_0-c)}, \\ \delta_n & = & \dfrac{K(\gamma_{n+1}-\gamma_n+\lambda(\gamma_n-\gamma_{n-1}))+M}{1-H\lambda(\gamma_{n+1}+\gamma_n-c)}, \end{cases}$$

and functions h_n on $[0,1)$ by

$$\begin{aligned} h_n(t) &= K(t+\lambda)t^{n-1}(\gamma_2-\gamma_1)+M \\ &\quad +H\lambda t\left[2\gamma_1+\frac{1-t^{n+1}}{1-t}(\gamma_2-\gamma_1)+\frac{1-t^n}{1-t}(\gamma_2-\gamma_1)\right] \\ &\quad -(1+H\lambda c)t. \end{aligned}$$

Suppose that function φ given by

$$\varphi(t)=2H\lambda\left(\gamma_1+\frac{\gamma_2-\gamma_1}{1-t}\right)t-(1+H\lambda c)t+M$$

has a minimal zero a in $[0,1)$ and

$$0\leq\delta_1\leq\alpha\leq a,$$

where α was defined in Lemma 9.3.1. Then, sequence $\{\gamma_n\}$ is non-decreasing, bounded from above by γ^{**} defined by

$$\gamma^{**}=c+\eta+\frac{\gamma_2-\gamma_1}{1-\alpha}$$

and converges to its unique least upper bound γ^* which satisfies

$$c+\eta\leq\gamma^*\leq\gamma^{**}.$$

Moreover, the following estimates are satisfied for each $n=1,2,\ldots$

$$0\leq\gamma_{n+2}-\gamma_{n+1}\leq\alpha^n(\gamma_2-\gamma_1).$$

Proof. Simply use $\{\gamma_n\}$, $\{\delta_n\}$, $\{h_n\}$, φ, a instead of $\{s_n\}$, $\{b_n\}$, $\{g_n\}$, p, α in the proof of Lemma 9.3.2 ■

Theorem 9.3.10. *Suppose that the (C^*), Lemma 9.3.9 conditions,*

$$U\subseteq\mathcal{D}$$

hold, where U was defined in Theorem 9.3.6 and $\|A_0^{-1}(F(x_0)+G(x_0))\|\leq\eta$. Then, sequence $\{x_n\}$ generated by the secant-like method (9.3.66) in well defined, remains in U for each $n=-1,0,1,2,\ldots$ and converges to a solution $x^*\in\overline{U}(x_0,\gamma^*-c)$ of equation $F(x)+G(x)=0$. Moreover, the following estimates are satisfied for each $n=0,1,\ldots$

$$\|x_{n+1}-x_n\|\leq\gamma_{n+1}-\gamma_n$$

and

$$\|x_n-x^*\|\leq\gamma^*-\gamma_n.$$

Furthermore, if there exists $\overline{\gamma} \geq \gamma^* - c$ *such that*

$$\overline{U}(x_0, \overline{\gamma}) \subseteq \mathcal{D}$$

and

$$0 < \frac{K((\lambda - 1)c + \overline{\gamma}) + M}{1 - H\lambda(2\overline{\gamma} - c)} \leq \eta, \quad \textit{for some} \quad \mu \in (0,1)$$

then, the solution x^* *is unique in* $\overline{U}(x_0, \overline{\gamma})$*.*

Proof. The proof until the uniqueness part follows as in Theorem 9.3.6 but using the identity

$$F(x_{k+1}) + G(x_{k+1}) = ([x_{k+1}, x_k; F] - A_k)(x_{k+1} - x_k) + (G(x_{k+1}) - G(x_k))$$

instead of (9.3.61). Finally, for the uniqueness part, let $y^* \in \overline{U}(x_0, \overline{\gamma})$ be such that $F(y^*) + G(y^*) = 0$. Then, we get from (9.3.66) the identity

$$\begin{aligned} x_{n+1} - y^* &= x_n - A_n^{-1}(F(x_n) + (x_n)) - y^* \\ &= -A_n^{-1}(F(x_n) - F(x^*) - A_n(x_n - y^*) + (G(x_n) - G(y^*))) \\ &= -A_n^{-1}(([x_n, y^*; F] - [y_n, x_{n-1}; F])(x_n - y^*) + (G(x_n) - G(y^*)). \end{aligned}$$

This identity leads to

$$\begin{aligned} \|x_{n+1} - y^*\| &\leq \frac{K(\|x_n - y_n\| + \|x_{n-1} - y^*\|) + M}{1 - H\lambda(\gamma_{n+1} + \gamma_n - c)} \|x_n - y^*\| \\ &\leq \frac{K((\lambda - 1)\|x_n - x_{n-1}\| + \|x_{n-1} - y^*\|) + M}{1 - H\lambda(2\overline{\gamma} - c)} \|x_n - y^*\| \\ &\leq \frac{K((\lambda - 1)c + \overline{\gamma}) + M}{1 - H\lambda(2\overline{\gamma} - c)} \|x_n - y^*\| \leq \mu \|x_n + y^*\| \\ &\leq \mu^{n+1} \|x_0 - y^*\| \leq \mu^{n+1} \overline{\gamma}. \end{aligned}$$

Hence, we deduce $\lim_{n \to +\infty} x_n = y^*$. But we know that $\lim_{n \to +\infty} x_n = x^*$. That is we conclude $x^* = y^*$, That completes the proof of the Theorem. ■

9.4. Numerical Examples

Example 9.4.1. *Let* $\mathcal{X} = \mathcal{Y} = \mathcal{C}[0,1]$*, equipped with the max-norm. Consider the following nonlinear boundary value problem*

$$\begin{cases} u'' = -u^3 - \gamma\, u^2 \\ u(0) = 0, \quad u(1) = 1. \end{cases}$$

It is well known that this problem can be formulated as the integral equation

$$u(s) = s + \int_0^1 Q(s,t)\, (u^3(t) + \gamma\, u^2(t))\, dt \tag{9.4.1}$$

where, Q is the Green function:

$$Q(s,t) = \begin{cases} t\,(1-s), & t \le s \\ s\,(1-t), & s < t. \end{cases}$$

We observe that

$$\max_{0\le s\le 1} \int_0^1 |Q(s,t)|\,dt = \frac{1}{8}.$$

Then problem (9.4.1) is in the form (9.1.1), where, $F : \mathcal{D} \longrightarrow \mathcal{Y}$ is defined as

$$[F(x)]\,(s) = x(s) - s - \int_0^1 Q(s,t)\,(x^3(t) + \gamma\, x^2(t))\,dt.$$

The Fréchet derivative of the operator F is given by

$$[F'(x)y]\,(s) = y(s) - 3\int_0^1 Q(s,t)x^2(t)y(t)dt - 2\gamma\int_0^1 Q(s,t)x(t)y(t)dt.$$

Then, we have that

$$[(I - F'(x_0))(y)](s) = 3\int_0^1 Q(s,t)x_0^2(t)y(t)dt + 2\gamma\int_0^1 Q(s,t)x_0(t)y(t)dt.$$

Hence, if $2\gamma < 5$, then

$$\|I - F'(x_0)\| \le 2(\gamma - 2) < 1.$$

It follows that $F'(x_0)^{-1}$ exists and

$$\|F'(x_0)^{-1}\| \le \frac{1}{5-2\gamma}.$$

We also have that $\|F(x_0)\| \le 1+\gamma$. Define the divided difference defined by

$$\delta F(x,y) = \int_0^1 F'(y + t(x-y))dt.$$

Choosing $x_{-1}(s)$ such that $|x_{-1} - x_0\| \le c$ and $l_0 c < 1$. Then we have for $\lambda = 1$

$$\|\delta F(x_{-1},x_0)^{-1}F(x_0)\| \le \|\delta F(x_{-1},x_0)^{-1}F'(x_0)\|\|F'(x_0)F(x_0)\|$$

and

$$\|\delta F(x_{-1},x_0)^{-1}F'(x_0)\| \le \frac{1}{(1-l_0c)},$$

where l_0 is such that

$$\|F'(x_0)^{-1}(F'(x_0) - A_0)\| \le l_0 c,$$

Set $u_0(s) = s$ and $\mathcal{D} = U(u_0,R_0)$. It is easy to verify that $U(u_0,R_0) \subset U(0,R_0+1)$ since $\| u_0 \| = 1$. If $2\,\gamma < 5$, and $l_0c < 1$ the operator F' satisfies conditions of Theorem 9.2.6, with

$$\eta = \frac{1+\gamma}{(1-l_0c)(5-2\,\gamma)}, \quad K = \frac{\gamma + 6\,R_0 + 3}{8(5-2\,\gamma)(1-l_0c)}, \quad H = \frac{2\,\gamma + 3\,R_0 + 6}{16(5-2\,\gamma)(1-l_0c)}.$$

Choosing $R_0 = 1$, $\gamma = 0.5$ *and* $c = 1$ *we obtain that*

$$l_0 = 0.1938137822\ldots,$$

$$\eta = 0.465153\ldots,$$

$$K = 0.368246\ldots$$

and

$$H = 0.193814\ldots.$$

Moreover we obtain that $a_{-1} = 0.317477$ *and* $b_{-1} = 0.251336$*, but conditions of Theorem 2.1 are not satisfied since*

$$b_{-1} = 0.251336 > 0.147893 = \frac{a_{-1}(1-a_{-1})^2}{2(1-a_{-1}) - \lambda(1-2a_{-1})}.$$

Notice also that the popular condition (9.3.36) *is also not satisfied, since* $Kc + 2\sqrt{K\eta} = 1.19599 > 1$*. Hence, there is no guarantee under the old conditions that the secant-type method converges to* x^**. However, conditions of Lemma 9.3.1 are satisfied, since*

$$0 < \alpha = 0.724067 \leq 0.776347 = \frac{1 - H\lambda(c+2\eta)}{1 - H\lambda c}$$

The convergence of the secant-type method is also ensured by Theorem 123.6.

Example 9.4.2. *Let* $\mathcal{X} = \mathcal{Y} = \mathbb{R}$ *and let consider the real function*

$$F(x) = x^3 - 2$$

and we are going to apply secant-type method with $\lambda = 2.5$*. We take the starting points* $x_0 = 1$, $x_{-1} = 0.25$ *and we consider the domain* $\Omega = B(x_0, 3/4)$*. In this case, we obtain*

$$c = 0.75,$$

$$\eta = 0.120301\ldots,$$

$$K = 0.442105\ldots,$$

$$H = 0.180451\ldots,$$

Notice that the conditions of Theorem 9.2.1 and Lemma 9.3.1 are satisfied, but since $H < K$ *Remark 9.2.2 ensures that our uniqueness ball is larger. It is clear as* $R_1 = 1.83333\ldots > 0.193452\ldots = R_0$.

References

[1] Amat, S., Bermúdez, C., Busquier, S., Gretay, J. O., Convergence by nondiscrete mathematical induction of a two step secant's method, *Rocky Mountain Journal of Mathematics*, 37(2) (2007), 359–369.

[2] Amat, S., Busquier, S., Gutiérrez, J. M., On the local convergence of secant-type methods, *International Journal of Computer Mathematics*, 81(8) (2004), 1153–1161.

[3] Amat, S., Busquier, S., Negra, M., Adaptive approximation of nonlinear operators, *Numerical Functional Analysis and Optimization*, 25(5) (2004), 397–405.

[4] Argyros, I. K., Polynomial operator equations in abstract spaces and applications, St.Lucie/CRC/Lewis Publ. Mathematics series, Boca Raton, Florida, U.S.A 1998.

[5] Argyros, I. K., Convergence and applications of Newton-type iterations, *Springer-Verlag Publications*, New York, 2008.

[6] Argyros, I. K., Cho, Y. J., Hilout, S., Numerical methods for equations and its applications, CRC Press/Taylor & Francis, New York, 2012.

[7] Argyros, I. K., Hilout, S., Convergence conditions for secant-type methods, *Czechoslovak Mathematical Journal*, 60 (2010), 11–18.

[8] Argyros, I. K., Hilout, S., Weaker conditions for the convergence of Newton's method, *Journal of Complexity*, 28(3) (2012), 346–387.

[9] Argyros, I. K., Hilout, S., Tabatabai, M.A., Mathematical modelling with applications in biosciences and engineering, Nova Publishers, New York, 2011.

[10] Bosarge, W. E., Falb, P. L. A multipoint method of third order, *Journal of Optimization Theory and Applications*, 4 (1969), 156–166.

[11] Dennis, J. E., Toward a unified convergence theory for Newton-like methods, Functional Analysis and Applications (L.B. Rall, ed.), Academic Press, New York, 1971.

[12] Ezquerro, J. A., Grau-Sánchez, M., Hernández, M. A., Noguera, M., Semilocal convergence of secant-like methods for differentiable and nondifferentiable operators equations, *J. Math. Anal. App.*, 398(1) (2013), 100-112.

[13] Ezquerro, J. A., Gutiérrez, J. M., Hernández, M.A., Romero, N., Rubio, M. J., The Newton method: from Newton to Kantorovich. (Spanish), *Gac. Real Soc. Mat. Esp.*, 13 (2010), 53–76.

[14] Ezquerro, J. A., Hernández, M.A., Romero, N., Velasco, A. I., *App. Math. Comp.*, 219(8) (2012), 3677-3692.

[15] Ezquerro, J. A., Hernández, M.A., Rubio, M.J., Secant-like methods for solving nonlinear integral equations of the Hammerstein type, *Proceedings of the 8th International Congress on Computational and Applied Mathematics, ICCAM-98 (Leuven), J. Comp. Appl. Math.*, 115 (2000), 245–254.

[16] Ezquerro, J. A., Hernández, M.A., Rubio, M.J., Solving a special case of conservative problems by secant-like methods, *App. Math. Comp.*, 169 (2005), 926–942.

[17] Ezquerro, J. A., Rubio, M.J.,, A uniparametric family of iterative processes for solving nondifferentiable equations, *J. Math. Anal. App.*, 275 (2002), 821–834.

[18] Kantorovich, L. V., Akilov, G. P., *Functional analysis*, Pergamon Press, Oxford, 1982.

[19] Laasonen, P.,, Ein überquadratisch konvergenter iterativer algorithmu, *Annales Academiae Scientiarum Fennicae Mathematica. Ser. I*, 450 (1969), 1–10.

[20] Magreñán, Á. A. A new tool to study real dynamics: The convergence plane *App. Math. Comp.*, 248 (2014), 215–224.

[21] Ortega, L. M., Rheinboldt, W. C., Iterative solution of nonlinear equations in several variables, Academic Press, New York, 1970.

[22] Potra, F. A., Pták, V.,, Nondiscrete induction and iterative processes. Research Notes in Mathematics, 103, Pitman (Advanced Publishing Program), Boston, Massachusetts, 1984.

[23] Proinov, P. D., General local convergence theory for a class of iterative processes and its applications to Newton's method, *Journal of Complexity*, 25 (2009), 38–62.

[24] Schmidt, J. W., Untere fehlerschranken fur regula-falsi verhafren, *Periodica Mathematica Hungarica*, 9 (1978), 241–247.

[25] Traub, J. F., Iterative Methods for the Solution of Equations, Prentice-Hall, Englewood Cliffs, New Jersay, 1964.

[26] Yamamoto, T., A convergence theorem for Newton-like methods in Banach spaces, *Numerische Mathematik*, 51, 545–557, 1987.

[27] Wolfe, M. A., Extended iterative methods for the solution of operator equations, *Numerische Mathematik*, 31 (1978), 153–174.

Chapter 10

On the Semilocal Convergence of a Two-Step Newton-Like Projection Method for Ill-Posed Equations

10.1. Introduction

Let X be a real Hilbert space with inner product $\langle .,. \rangle$ and norm $\|.\|$. Let $U(x,R)$ and $\overline{U(x,R)}$, stand respectively, for the open and closed ball in X with center x and radius $R > 0$. Let also $L(X)$ be the space of all bounded linear operators from X into itself.

In this chapter we are concerned with the problem of approximately solving the ill-posed equation

$$F(x) = y, \tag{10.1.1}$$

where $F : D(F) \subseteq X \to X$ is a nonlinear operator satisfying $\langle F(v) - F(w), v - w \rangle \geq 0, \quad \forall v, w \in D(F)$, and $y \in X$.

It is assumed that (10.1.1) has a solution, namely $\hat{x}$ and F possesses a locally uniformly bounded Fréchet derivative $F'(x)$ for all $x \in D(F)$ (cf. [18]) i.e.,

$$\|F'(x)\| \leq C_F, \quad x \in D(F)$$

for some constant C_F.

In application, usually only noisy data y^δ are available, such that

$$\|y - y^\delta\| \leq \delta.$$

Then the problem of recovery of $\hat{x}$ from noisy equation $F(x) = y^\delta$ is ill-posed, in the sense that a small perturbation in the data can cause large deviation in the solution. For solving (10.1.1) with monotone operators (see [12, 17, 18, 19]) one usually use the Lavrentiev regularization method. In this method the regularized approximation x_α^δ is obtained by solving the operator equation

$$F(x) + \alpha(x - x_0) = y^\delta. \tag{10.1.2}$$

It is known (cf. [19], Theorem 1.1) that the equation (10.1.2) has a unique solution x_α^δ for $\alpha > 0$, provided F is Fréchet differentiable and monotone in the ball $B_r(\hat{x}) \subset D(F)$ with radius $r = \|\hat{x} - x_0\| + \delta/\alpha$. However the regularized equation (10.1.2) remains nonlinear and one may have difficulties in solving them numerically.

In [6], George and Elmahdy considered an iterative regularization method which converges linearly to x_α^δ and its finite dimensional realization in [7]. Later in [8] George and Elmahdy considered an iterative regularization method which converges quadratically to x_α^δ and its finite dimensional realization in [9].

Recall that a sequence (x_n) in X with $\lim x_n = x^*$ is said to be convergent of order $p > 1$, if there exist positive reals β, γ, such that for all $n \in N$ $\|x_n - x^*\| \le \beta e^{-\gamma p^n}$.If the sequence (x_n) has the property that $\|x_n - x^*\| \le \beta q^n$, $0 < q < 1$ then (x_n) is said to be linearly convergent. For an extensive discussion of convergence rate (see [13]).

Note that the method considered in [6], [7], [8] and [9] are proved using a suitably constructed majorizing sequence which heavily depends on the initial guess and hence not suitable for practical consideration.

Recently, George and Pareth [10] introduced a two-step Newton-like projection method(TSNLPM) of convergence order four to solve (10.1.2). (TSNLPM) was realized as follows:

Let $\{P_h\}_{h>0}$ be a family of orthogonal projections on X. Our aim in this section is to obtain an approximation for x_α^δ, in the finite dimensional space $R(P_h)$, the range of P_h. For the results that follow, we impose the following conditions.

Let

$$\varepsilon_h(x) := \|F'(x)(I - P_h)\|, \quad \forall x \in D(F)$$

and $\{b_h : h > 0\}$ is such that $lim_{h\to 0} \frac{\|(I-P_h)x_0\|}{b_h} = 0$ and $lim_{h\to 0} b_h = 0$. We assume that $\varepsilon_h(x) \to 0$, $\forall x \in D(F)$ as $h \to 0$. The above assumption is satisfied if, $P_h \to I$ pointwise and if $F'(x)$ is a compact operator. Further we assume that $\varepsilon_h(x) \le \varepsilon_0$, $\forall x \in D(F)$, $b_h \le b_0$ and $\delta \in (0, \delta_0]$.

10.1.1. Projection Method

We consider the following sequence defined iteratively by

$$y_{n,\alpha}^{h,\delta} = x_{n,\alpha}^{h,\delta} - R_\alpha^{-1}(x_{n,\alpha}^{h,\delta}) P_h [F(x_{n,\alpha}^{h,\delta}) - f^\delta + \alpha(x_{n,\alpha}^{h,\delta} - x_0)] \tag{10.1.3}$$

and

$$x_{n+1,\alpha}^{h,\delta} = y_{n,\alpha}^{h,\delta} - R_\alpha^{-1}(y_{n,\alpha}^{h,\delta}) P_h [F(y_{n,\alpha}^{h,\delta}) - f^\delta + \alpha(y_{n,\alpha}^{h,\delta} - x_0)] \tag{10.1.4}$$

where $R_\alpha(x) := P_h F'(x) P_h + \alpha P_h$ and $x_{0,\alpha}^{h,\delta} := P_h x_0$, for obtaining an approximation for x_α^δ in the finite dimensional subspace $R(P_h)$ of X. Note that the iteration (10.1.3) and (10.1.4) are the finite dimensional realization of the iteration (10.1.3) and (10.1.4) in [16]. In [10], the parameter $\alpha = \alpha_i$ was chosen from some finite set

$$D_N = \{\alpha_i : 0 < \alpha_0 < \alpha_1 < \alpha_2 < \cdots < \alpha_N\}$$

using the adaptive method considered by Perverzev and Schock in [17].
The convergence analysis in [10] was carried out using the following assumptions.

Assumption 10.1.1. *(cf. [18], Assumption 3) There exists a constant $k_0 \geq 0$ such that for every $x,u \in D(F)$ and $v \in X$ there exists an element $\Phi(x,u,v) \in X$ such that $[F'(x) - F'(u)]v = F'(u)\Phi(x,u,v), \|\Phi(x,u,v)\| \leq k_0\|v\|\|x-u\|$.*

Assumption 10.1.2. *There exists a continuous, strictly monotonically increasing function $\varphi : (0,a] \to (0,\infty)$ with $a \geq \|F'(\hat{x})\|$ satisfying;*

(i) $lim_{\lambda\to 0}\varphi(\lambda) = 0,$

(ii) $sup_{\lambda\geq 0}\frac{\alpha\varphi(\lambda)}{\lambda+\alpha} \leq c_\varphi\varphi(\alpha) \qquad \forall\lambda \in (0,a]$ *and*

(iii) there exists $v \in X$ with $\|v\| \leq 1$ (cf. [15]) such that

$$x_0 - \hat{x} = \varphi(F'(\hat{x}))v.$$

In the present paper we extend the applicability of (TSNLPM) by weakening Assumption 10.1.1 which is very difficult to verify (or does not hold) in general (see numerical examples at the last section of the paper). In particular, we replace Assumption 10.1.1 by the weaker and easier to verify:

Assumption 10.1.3. *Let $x_0 \in X$ be fixed. There exists a constant $K_0 \geq 0$ such that for each $x,u \in D(F)$ and $v \in X$ there exists an element $\Phi(x,u,v) \in X$ depending on x_0 such that $[F'(x) - F'(u)]v = F'(u)\Phi(x,u,v), \|\Phi(x,u,v)\| \leq K_0\|v\|(\|x-P_hx_0\| + \|u-P_hx_0\|)$.*

Note that Assumption 10.1.1$\Rightarrow$ Assumption 10.1.3 but not necessarily vice versa. At the end of the chapter we have provided examples, where Assumption 10.1.3 is satisfied but not Assumption 10.1.1.

We also replace Assumption 10.1.2 by

Assumption 10.1.4. *There exists a continuous, strictly monotonically increasing function $\varphi : (0,a] \to (0,\infty)$ with $a \geq \|F'(x_0)\|$ satisfying;*

(i) $lim_{\lambda\to 0}\varphi(\lambda) = 0,$

(ii) $sup_{\lambda\geq 0}\frac{\alpha\varphi(\lambda)}{\lambda+\alpha} \leq \varphi(\alpha) \qquad \forall\lambda \in (0,a]$ *and*

(iii) there exists $v \in X$ with $\|v\| \leq 1$ (cf. [15]) such that

$$x_0 - \hat{x} = \varphi(F'(x_0))v.$$

Remark 10.1.5. *The hypotheses of Assumption 10.1.1 may not hold or may be very expensive or imposible to verify in general. In particular, as it is the case for well-posed nonlinear equations the computation of the Lipschitz constant k_0 even if this constant exists is very difficult. Moreover, there are classes of operators for which Assumption 10.1.1 is not satisfied but the (TSNLPM) converges.*

In this paper, we expand the applicability of (TSNLPM) under less computational cost. Let us explain how we achieve this goal.

(1) Assumption 10.1.3 is weaker than Assumption 10.1.1. Notice that there are classes of operators that satisfy Assumption 10.1.3 but do not satisfy Assumption 10.1.1;

(2) *The computational cost of constant K_0 is less than that of constant k_0, even when $K_0 = k_0$;*

(3) *The sufficient convergence criteria are weaker;*

(4) *The computable error bounds on the distances involved (including K_0) are less costly;*

(5) *The convergence domain of (TSNLPM) with Assumption 10.1.3 can be larger, since $\frac{K_0}{k_0}$ can be arbitrarily small (see Example 10.5.4);*

(6) *The information on the location of the solution is more precise;*

and

(7) *Note that the Assumption 10.1.2 involves the Fréchet derivative at the exact solution $\hat{x}$ which is unknown in practice. But Assumption 10.1.4 depends on the Fréchet derivative of F at x_0.*

These advantages are also very important in computational mathematics since they provide under less computational cost.

The paper is organization as follows: In Section 10.2 we present the convergence analysis of (TSNLPM). Section 10.3 contains the error analysis and parameter choice strategy. The algorithm for implementing(TSNLPM) is given in Section 10.4. Finally, numerical examples are presented in the concluding Section 10.5.

10.2. Semilocal Convergence

In order for us to present the semilocal convergence of (TSNLPM) it is convenient to introduce some parameters:

Let

$$e_{n,\alpha}^{h,\delta} \quad := \quad \|y_{n,\alpha}^{h,\delta} - x_{n,\alpha}^{h,\delta}\|, \qquad \forall n \geq 0. \tag{10.2.1}$$

Suppose that

$$0 < K_0 < \frac{1}{4(1+\frac{\varepsilon_0}{\alpha_0})} \tag{10.2.2}$$

and

$$\frac{4\delta_0}{\alpha_0}(1+\frac{\varepsilon_0}{\alpha_0}) < 1. \tag{10.2.3}$$

Define polynomial P on $(0,\infty)$ by

$$P(t) = (1+\frac{\varepsilon_0}{\alpha_0})\frac{K_0}{2}t^2 + (1+\frac{\varepsilon_0}{\alpha_0})t + \frac{\delta_0}{\alpha_0} - \frac{1}{4(1+\frac{\varepsilon_0}{\alpha_0})}. \tag{10.2.4}$$

It follows from (10.2.3) that P has a unique positive root given in closed form by the quadratic formula. Denote this root by p_0.

Let

$$b_0 < p_0, \|\hat{x} - x_0\| \leq \rho, \tag{10.2.5}$$

where

$$\rho < p_0 - b_0. \tag{10.2.6}$$

$$\gamma_\rho := (1 + \frac{\varepsilon_0}{\alpha_0}) \left[\frac{k_0}{2}(\rho + b_0)^2 + (\rho + b_0) \right] + \frac{\delta_0}{\alpha_0}, \tag{10.2.7}$$

$$r := \frac{4\gamma_\rho}{1 + \sqrt{1 + 32\gamma_\rho(1 + \frac{\varepsilon_0}{\alpha_0})}} \tag{10.2.8}$$

and

$$b := 4(1 + \frac{\varepsilon_0}{\alpha_0}) K_0 r. \tag{10.2.9}$$

Then we have by (10.2.2)-(10.2.9) that

$$0 < \gamma_\rho < \frac{1}{4}. \tag{10.2.10}$$

$$0 < r < 1 \tag{10.2.11}$$

and

$$0 < b < 1. \tag{10.2.12}$$

Indeed, we have by (10.2.4) and (10.2.12) that $\gamma_\rho - \frac{1}{4} \leq P(p_0) = 0 \Rightarrow 0 < \gamma_\rho < \frac{1}{4} \Rightarrow$ (10.2.10). Estimate (10.2.11) follows from (10.2.8) and (10.2.10). Moreover, estimate (10.2.12) follows from (10.2.2) and (10.2.11). We also have that

$$\gamma_\rho < r. \tag{10.2.13}$$

In view of (10.2.7) and (10.2.8), estimate (10.2.13) reduces to showing that $4\gamma_\rho(1 + \frac{\varepsilon_0}{\alpha_0}) < 1$ which is true by the choice of p_0 and (10.2.4). Finally it follows from (10.2.13) that

$$0 < \gamma_\rho < 1. \tag{10.2.14}$$

Lemma 10.2.1. *([10], Lemma1)Let $x \in D(F)$. Then*

$$\|R_\alpha^{-1}(x) P_h F'(x)\| \leq (1 + \frac{\varepsilon_0}{\alpha_0}).$$

Lemma 10.2.2. *([10], Lemma 2) Let $e_0 = e_{0,\alpha}^{h,\delta}$ and γ_ρ be as in (10.2.7). Then* $e_0 \leq \gamma_\rho$.

Lemma 10.2.3. *Suppose that (10.2.2), (10.2.3) and $\delta \in (0, \delta_0]$ hold and let Assumption 10.1.3 be satisfied. Then the following estimates hold for (TSNLPM):*

(a)

$$\|x_{n,\alpha}^{h,\delta} - y_{n-1,\alpha}^{h,\delta}\| \leq \frac{K_0}{2}(1+\frac{\varepsilon_0}{\alpha_0})[3\|x_{n-1,\alpha}^{h,\delta} - x_{0,\alpha}^{h,\delta}\| \tag{10.2.15}$$

$$+5\|y_{n-1,\alpha}^{h,\delta} - x_{0,\alpha}^{h,\delta}\|]e_{n-1,\alpha}^{h,\delta} \tag{10.2.16}$$

and

(b)

$$\|x_{n,\alpha}^{h,\delta} - x_{n-1,\alpha}^{h,\delta}\| \leq \{1+\frac{K_0}{2}(1+\frac{\varepsilon_0}{\alpha_0})[3\|x_{n-1,\alpha}^{h,\delta} - x_{0,\alpha}^{h,\delta}\| \tag{10.2.17}$$

$$+5\|y_{n-1,\alpha}^{h,\delta} - x_{0,\alpha}^{h,\delta}\|]\}e_{n-1,\alpha}^{h,\delta}. \tag{10.2.18}$$

Proof. Observe that,

$$\begin{aligned}
& x_{n,\alpha}^{h,\delta} - y_{n-1,\alpha}^{h,\delta} \\
= & \; y_{n-1,\alpha}^{h,\delta} - x_{n-1,\alpha}^{h,\delta} - R_\alpha^{-1}(y_{n-1,\alpha}^{h,\delta})P_h \\
& [F(y_{n-1,\alpha}^{h,\delta}) - f^\delta + \alpha(y_{n-1,\alpha}^{h,\delta} - x_0)] + R_\alpha^{-1}(x_{n-1,\alpha}^{h,\delta}) \\
& P_h[F(x_{n-1,\alpha}^{h,\delta}) - f^\delta + \alpha(x_{n-1,\alpha}^{h,\delta} - x_0)] \\
= & \; y_{n-1,\alpha}^{h,\delta} - x_{n-1,\alpha}^{h,\delta} - R_\alpha^{-1}(y_{n-1,\alpha}^{h,\delta})P_h \\
& [F(y_{n-1,\alpha}^{h,\delta}) - F(x_{n-1,\alpha}^{h,\delta}) + \alpha(y_{n-1,\alpha}^{h,\delta} - x_{n-1,\alpha}^{h,\delta})] \\
& +[R_\alpha^{-1}(x_{n-1,\alpha}^{h,\delta}) - R_\alpha^{-1}(y_{n-1,\alpha}^{h,\delta})]P_h[F(x_{n-1,\alpha}^{h,\delta}) - f^\delta \\
& +\alpha(x_{n-1,\alpha}^{h,\delta} - x_0)] \\
= & \; R_\alpha^{-1}(y_{n-1,\alpha}^{h,\delta})P_h[F'(y_{n-1,\alpha}^{h,\delta})(y_{n-1,\alpha}^{h,\delta} - x_{n-1,\alpha}^{h,\delta}) \\
& -(F(y_{n-1,\alpha}^{h,\delta}) - F(x_{n-1,\alpha}^{h,\delta}))] + R_\alpha^{-1}(y_{n-1,\alpha}^{h,\delta})P_h \\
& (F'(y_{n-1,\alpha}^{h,\delta}) - F'(x_{n-1,\alpha}^{h,\delta}))(x_{n-1,\alpha}^{h,\delta} - y_{n-1,\alpha}^{h,\delta}) \\
:= & \; \Gamma_1 + \Gamma_2
\end{aligned} \tag{10.2.19}$$

where

$$\begin{aligned}
\Gamma_1 := & \; R_\alpha^{-1}(y_{n-1,\alpha}^{h,\delta})P_h[F'(y_{n-1,\alpha}^{h,\delta})(y_{n-1,\alpha}^{h,\delta} - x_{n-1,\alpha}^{h,\delta}) \\
& -(F(y_{n-1,\alpha}^{h,\delta}) - F(x_{n-1,\alpha}^{h,\delta}))]
\end{aligned}$$

and

$$\begin{aligned}
\Gamma_2 := & \; R_\alpha^{-1}(y_{n-1,\alpha}^{h,\delta})P_h[F'(y_{n-1,\alpha}^{h,\delta}) - F'(x_{n-1,\alpha}^{h,\delta})] \\
& (x_{n-1,\alpha}^{h,\delta} - y_{n-1,\alpha}^{h,\delta}).
\end{aligned}$$

Note that,

$$\begin{aligned}
\|\Gamma_1\| &= \|R_\alpha^{-1}(y_{n-1,\alpha}^{h,\delta})P_h\int_0^1[F'(y_{n-1,\alpha}^{h,\delta})-F'(x_{n-1,\alpha}^{h,\delta}\\
&\quad +t(y_{n-1,\alpha}^{h,\delta}-x_{n-1,\alpha}^{h,\delta}))](y_{n-1,\alpha}^{h,\delta}-x_{n-1,\alpha}^{h,\delta})dt\|\\
&= \|R_\alpha^{-1}(y_{n-1,\alpha}^{h,\delta})P_hF'(y_{n-1,\alpha}^{h,\delta})\int_0^1[\phi(x_{n-1,\alpha}^{h,\delta}+\\
&\quad t(y_{n-1,\alpha}^{h,\delta}-x_{n-1,\alpha}^{h,\delta}),y_{n-1,\alpha}^{h,\delta},x_{n-1,\alpha}^{h,\delta}-y_{n-1,\alpha}^{h,\delta})]dt\|\\
&\le K_0(1+\frac{\varepsilon_0}{\alpha_0})[\int_0^1\|x_{n-1,\alpha}^{h,\delta}-x_{0,\alpha}^{h,\delta} \qquad (10.2.20)\\
&\quad -t(y_{n-1,\alpha}^{h,\delta}-x_{n-1,\alpha}^{h,\delta})\|dt+\|y_{n-1,\alpha}^{h,\delta}-x_{0,\alpha}^{h,\delta}\|]\\
&\quad \times\|y_{n-1,\alpha}^{h,\delta}-x_{n-1,\alpha}^{h,\delta}\|\\
&\le K_0(1+\frac{\varepsilon_0}{\alpha_0})[\int_0^1(1-t)\|x_{n-1,\alpha}^{h,\delta}-x_{0,\alpha}^{h,\delta}\|\\
&\quad +t\|y_{n-1,\alpha}^{h,\delta})-x_{0,\alpha}^{h,\delta}\|+\|y_{n-1,\alpha}^{h,\delta})-x_{0,\alpha}^{h,\delta}\|]dt\\
&\quad \times\|y_{n-1,\alpha}^{h,\delta}-x_{n-1,\alpha}^{h,\delta}\|\\
&\le \frac{K_0}{2}(1+\frac{\varepsilon_0}{\alpha_0})[\|x_{n-1,\alpha}^{h,\delta}-x_{0,\alpha}^{h,\delta}\|\\
&\quad +3\|y_{n-1,\alpha}^{h,\delta})-x_{0,\alpha}^{h,\delta}\|]e_{n-1,\alpha}^{h,\delta}
\end{aligned}$$

the last step follows from the Assumption 10.1.3 and Lemma 10.2.1. Similarly,

$$\|\Gamma_2\| \le K_0(1+\frac{\varepsilon_0}{\alpha_0})[\|y_{n-1,\alpha}^{h,\delta}-x_{0,\alpha}^{h,\delta}\|+\|x_{0,\alpha}^{h,\delta}-x_{n-1,\alpha}^{h,\delta}\|]e_{n-1,\alpha}^{h,\delta} \qquad (10.2.21)$$

So, (a) follows from (10.2.19), (10.2.20) and (10.2.21). And (b) follows from (a) and the triangle inequality;

$$\|x_{n,\alpha}^{h,\delta}-x_{n-1,\alpha}^{h,\delta}\|\le\|x_{n,\alpha}^{h,\delta}-y_{n-1,\alpha}^{h,\delta}\|+\|y_{n-1,\alpha}^{h,\delta}-x_{n-1,\alpha}^{h,\delta}\|.$$

Theorem 10.2.4. *Under the hypotheses of Lemma 10.2.3 the following estimates hold for (TSNLPM):*

$$\begin{aligned}
e_{n,\alpha}^{h,\delta} &\le \frac{K_0}{2}(1+\frac{\varepsilon_0}{\alpha_0})[5\|x_{n,\alpha}^{h,\delta}-x_{0,\alpha}^{h,\delta}\|\\
&\quad +3\|y_{n-1,\alpha}^{h,\delta})-x_{0,\alpha}^{h,\delta}\|]\|y_{n-1,\alpha}^{h,\delta}-x_{n,\alpha}^{h,\delta}\|\\
&\le b^2e_{n-1,\alpha}^{h,\delta}\le b^{2n}e_{0,\alpha}^{h,\delta}\le b^{2n}\gamma_\rho.
\end{aligned}$$

Proof. We have,

$$\begin{aligned}
y_{n,\alpha}^{h,\delta}-x_{n,\alpha}^{h,\delta} &= x_{n,\alpha}^{h,\delta}-y_{n-1,\alpha}^{h,\delta}-R_\alpha^{-1}(x_{n,\alpha}^{h,\delta})P_h \\
&\quad [F(x_{n,\alpha}^{h,\delta})-f^\delta+\alpha(x_{n,\alpha}^{h,\delta}-x_0)]+R_\alpha^{-1}(y_{n-1,\alpha}^{h,\delta}) \\
&\quad P_h[F(y_{n-1,\alpha}^{h,\delta})-f^\delta+\alpha(y_{n-1,\alpha}^{h,\delta}-x_0)] \\
&= x_{n,\alpha}^{h,\delta}-y_{n-1,\alpha}^{h,\delta}-R_\alpha^{-1}(x_{n,\alpha}^{h,\delta})P_h \\
&\quad [F(x_{n,\alpha}^{h,\delta})-F(y_{n-1,\alpha}^{h,\delta})+\alpha(x_{n,\alpha}^{h,\delta}-y_{n-1,\alpha}^{h,\delta})] \\
&\quad +[R_\alpha^{-1}(y_{n-1,\alpha}^{h,\delta})-R_\alpha^{-1}(x_{n,\alpha}^{h,\delta})]P_h[F(y_{n-1,\alpha}^{h,\delta})-f^\delta \\
&\quad +\alpha(y_{n-1,\alpha}^{h,\delta}-x_0)] \\
&= R_\alpha^{-1}(x_{n,\alpha}^{h,\delta})P_h[F'(x_{n,\alpha}^{h,\delta})(x_{n,\alpha}^{h,\delta}-y_{n-1,\alpha}^{h,\delta}) \\
&\quad -(F(x_{n,\alpha}^{h,\delta})-F(y_{n-1,\alpha}^{h,\delta}))]+R_\alpha^{-1}(x_{n,\alpha}^{h,\delta})P_h \\
&\quad [F'(x_{n,\alpha}^{h,\delta})-F'(y_{n-1,\alpha}^{h,\delta})]\times(y_{n-1,\alpha}^{h,\delta}-x_{n,\alpha}^{h,\delta}) \\
&:= \Gamma_3+\Gamma_4 \qquad (10.2.22)
\end{aligned}$$

where $\Gamma_3 = R_\alpha^{-1}(x_{n,\alpha}^{h,\delta})P_h[F'(x_{n,\alpha}^{h,\delta})(x_{n,\alpha}^{h,\delta} - y_{n-1,\alpha}^{h,\delta}) - (F(x_{n,\alpha}^{h,\delta}) - F(y_{n-1,\alpha}^{h,\delta}))]$ and $\Gamma_4 = R_\alpha^{-1}(x_{n,\alpha}^{h,\delta})P_h[F'(x_{n,\alpha}^{h,\delta})-F'(y_{n-1,\alpha}^{h,\delta})](y_{n-1,\alpha}^{h,\delta}-x_{n,\alpha}^{h,\delta})$. Analogous to the proof of (10.2.20) and (10.2.21) one can prove that

$$\begin{aligned}
\|\Gamma_3\| &\leq \frac{K_0}{2}(1+\frac{\varepsilon_0}{\alpha_0})[3\|x_{n,\alpha}^{h,\delta}-x_{0,\alpha}^{h,\delta}\| \\
&\quad +\|y_{n-1,\alpha}^{h,\delta}-x_{0,\alpha}^{h,\delta}\|]\|x_{n,\alpha}^{h,\delta}-y_{n-1,\alpha}^{h,\delta}\| \qquad (10.2.23)
\end{aligned}$$

and

$$\|\Gamma_4\| \leq K_0(1+\frac{\varepsilon_0}{\alpha_0})[\|x_{n,\alpha}^{h,\delta}-x_{0,\alpha}^{h,\delta}\|+\|y_{n-1,\alpha}^{h,\delta}-x_{0,\alpha}^{h,\delta}\|]\|x_{n,\alpha}^{h,\delta}-y_{n-1,\alpha}^{h,\delta}\|$$

Now

$$\begin{aligned}
e_{n,\alpha}^{h,\delta} &\leq \frac{K_0}{2}(1+\frac{\varepsilon_0}{\alpha_0})[5\|x_{n,\alpha}^{h,\delta}-x_{0,\alpha}^{h,\delta}\| \\
&\quad +3\|y_{n-1,\alpha}^{h,\delta}-x_{0,\alpha}^{h,\delta}\|]\|x_{n,\alpha}^{h,\delta}-y_{n-1,\alpha}^{h,\delta}\| \qquad (10.2.24) \\
&\leq \frac{K_0}{2}(1+\frac{\varepsilon_0}{\alpha_0})(8r)\frac{K_0}{2}(1+\frac{\varepsilon_0}{\alpha_0})(8r)\|x_{n-1,\alpha}^{h,\delta}-y_{n-1,\alpha}^{h,\delta}\| \\
&\leq b^2\|x_{n-1,\alpha}^{h,\delta}-y_{n-1,\alpha}^{h,\delta}\| \\
&\leq b^{2n}e_{0,\alpha}^{h,\delta}\leq b^{2n}\gamma_\rho.
\end{aligned}$$

This completes the proof of the theorem.

Theorem 10.2.5. *Suppose that the hypotheses of Theorem 10.2.4 hold. Then, sequences* $\{x_{n,\alpha}^{h,\delta}\}$, $\{y_{n,\alpha}^{h,\delta}\}$ *generated by (TSNLPM) are well defined and remain in* $U(P_hx_0,r)$ *for all* $n\geq 0$.

Proof. Note that by (b) of Lemma 10.2.3 we have,

$$\begin{aligned}\|x_{1,\alpha}^{h,\delta}-P_hx_0\| &= \|x_{1,\alpha}^{h,\delta}-x_{0,\alpha}^{h,\delta}\|\\ &\le [1+(1+\frac{\varepsilon_0}{\alpha_0})\frac{K_0}{2}(8r)]\gamma_\rho \\ &\le (1+b)\gamma_\rho\\ &\le \frac{1-b^2}{1-b}\gamma_\rho < r,\end{aligned} \tag{10.2.25}$$

i.e., $x_{1,\alpha}^{h,\delta} \in B_r(P_hx_0)$. Again note that from (10.2.25) and Theorem 10.2.4 we get,

$$\begin{aligned}\|y_{1,\alpha}^{h,\delta}-P_hx_0\| &\le \|y_{1,\alpha}^{h,\delta}-x_{1,\alpha}^{h,\delta}\|+\|x_{1,\alpha}^{h,\delta}-P_hx_0\|\\ &\le [1+(1+\frac{\varepsilon_0}{\alpha_0})4K_0r+((1+\frac{\varepsilon_0}{\alpha_0})4K_0r)^2]\gamma_\rho\\ &\le (1+b+b^2)\gamma_\rho\\ &\le \frac{1-b^2}{1-b}\gamma_\rho < r,\end{aligned}$$

i.e., $y_{1,\alpha}^{h,\delta} \in B_r(P_hx_0)$. Further by (10.2.25) and (b) of Lemma 10.2.3 we have,

$$\begin{aligned}\|x_{2,\alpha}^{h,\delta}-P_hx_0\| &\le \|x_{2,\alpha}^{h,\delta}-x_{1,\alpha}^{h,\delta}\|+\|x_{1,\alpha}^{h,\delta}-P_hx_0\|\\ &\le (1+b)\gamma_\rho+(1+b)\gamma_\rho\\ &= 2(1+b)\gamma_\rho < r\end{aligned}$$

and

$$\begin{aligned}\|y_{2,\alpha}^{h,\delta}-P_hx_0\| &\le \|y_{2,\alpha}^{h,\delta}-x_{2,\alpha}^{h,\delta}\|+\|x_{2,\alpha}^{h,\delta}-P_hx_0\|\\ &\le b^4\gamma_\rho+2(1+b)\gamma_\rho\\ &\le b^2\gamma_\rho+2(1+b)\gamma_\rho\\ &\le [\frac{1-b^3}{1-b}+\frac{1-b^2}{1-b}]\gamma_\rho\\ \text{(since } b<1) & \\ &< \frac{2\gamma_\rho}{1-b} < r\end{aligned}$$

by the choice of r, i.e., $x_{2,\alpha}^{h,\delta}$, $y_{2,\alpha}^{h,\delta} \in B_r(P_hx_0)$. Continuing this way one can prove that $x_{n,\alpha}^{h,\delta}$, $y_{n,\alpha}^{h,\delta} \in B_r(P_hx_0), \forall n \ge 0$. This completes the proof.

Theorem 10.2.6. *Suppose that the hypotheses of Theorem 10.2.5 hold. Then the following assertions hold*

(a) $\{x_{n,\alpha}^{h,\delta}\}$ *is a complete sequence in* $U(P_hx_0,r)$ *and converges to* $x_\alpha^{h,\delta} \in \overline{U(P_hx_0,r)}$.

(b) $P_h[F(x_\alpha^{h,\delta})+\alpha(x_\alpha^{h,\delta}-x_0)] = P_hy^\delta$.

(c)

$$\|x_{n,\alpha}^{h,\delta} - x_{\alpha}^{h,\delta}\| \leq \frac{(1+b)b^{2n}\gamma_{\rho}}{1-b^2}$$

where γ_{ρ} *and* b *are defined by (10.2.7) and (10.2.9), respectively.*

Proof. We have that

$$\begin{aligned}
\|x_{n+i+1,\alpha}^{h,\delta} - x_{n+i,\alpha}^{h,\delta}\| & \\
&\leq (1+b)b^0\|x_{n+i,\alpha}^{h,\delta} - y_{n+i,\alpha}^{h,\delta}\| \\
&\leq (1+b)b\|x_{n+i,\alpha}^{h,\delta} - y_{n+i-1,\alpha}^{h,\delta}\| \\
&\leq (1+b)b^2\|x_{n+i-1,\alpha}^{h,\delta} - y_{n+i,\alpha}^{h,\delta}\| \\
&\leq (1+b)b^{2(n+i)}e_{0,\alpha}^{h,\delta} \\
&\leq (1+b)b^{2(n+i)}\gamma_{\rho}.
\end{aligned}$$

So,

$$\begin{aligned}
\|x_{n+m,\alpha}^{h,\delta} - x_{n,\alpha}^{h,\delta}\| &\leq \sum_{i=0}^{m-1}\|x_{n+i+1,\alpha}^{h,\delta} - x_{n+i,\alpha}^{h,\delta}\| \\
&\leq (1+b)b^{2n}\sum_{i=0}^{m-1}b^{2i} \\
&= (1+b)b^{2n}\frac{1-b^{2m}}{1-b^2}\gamma_{\rho} \to \frac{(1+b)b^{2n}}{1-b^2}\gamma_{\rho},
\end{aligned}$$

as $m \to \infty$. Thus $x_{n,\alpha}^{h,\delta}$ is a Cauchy sequence in $U(P_hx_0, r)$ and hence it converges, say to $x_{\alpha}^{h,\delta} \in \overline{U(P_hx_0, r)}$.
Observe that,

$$\begin{aligned}
\|P_h[F(x_{n,\alpha}^{h,\delta}) - f^{\delta} + \alpha(x_{n,\alpha}^{h,\delta} - x_0)]\| & \\
&= \|R_{\alpha}(x_0)(x_{n,\alpha}^{h,\delta} - y_{n,\alpha}^{h,\delta})\| \\
&\leq \|R_{\alpha}(x_0)\|\|x_{n,\alpha}^{h,\delta} - y_{n,\alpha}^{h,\delta}\| \\
&= \|(P_hF'(x_{n,\alpha}^{h,\delta})P_h + \alpha P_h)\|e_{n,\alpha}^{h,\delta} \\
&\leq (C_F + \alpha)e_{n,\alpha}^{h,\delta}. \qquad (10.2.26)
\end{aligned}$$

Now by letting $n \to \infty$ in (10.2.26) we obtain

$$P_h[F(x_{\alpha}^{h,\delta}) + \alpha(x_{\alpha}^{h,\delta} - x_0)] = P_hy^{\delta}. \qquad (10.2.27)$$

This completes the proof.

Remark 10.2.7. *(a) The convergence order of (TSNLPM) is four [10], under Assumption 10.1.1. In Theorem 10.2.6 the error bounds are too pessimistic. That is why in* *practice we shall use the computational order of convergence (COC) (see eg. [5]) defined by*

$$\rho \approx \ln\left(\frac{\|x_{n+1}-x_\alpha^\delta\|}{\|x_n-x_\alpha^\delta\|}\right) / \ln\left(\frac{\|x_n-x_\alpha^\delta\|}{\|x_{n-1}-x_\alpha^\delta\|}\right).$$

The (COC) ρ will then be close to 4 which is the order of convergence of (TSNLPM).

(b) *Note that from the proof of the Theorem 10.2.5 a larger r can be obtained from solving the equation*

$$[b^4 t + 2(1+bt)]\gamma_\rho - rt = 0.$$

Note that this equation has a minimal root $r^* > r$. Then, r^* can replace r in Theorem 10.2.5. However, we decided to use r which is given in closed form. Using, Mathematica or Maple we found r^* in closed form. But it has a complicated and long form. That is why we decided not to include r in this paper.

10.3. Error Bounds under Source Conditions

The objective of this section is to obtain an error estimate for $\|x_{n,\alpha}^{h,\delta} - \hat{x}\|$ under a source condition on $x_0 - \hat{x}$.

Proposition 10.3.1. *Let $F : D(F) \subseteq X \to X$ be a monotone operator in X. Let $x_\alpha^{h,\delta}$ be the solution of (10.2.27) and $x_\alpha^h := x_\alpha^{h,0}$. Then*

$$\|x_\alpha^{h,\delta} - x_\alpha^h\| \le \frac{\delta}{\alpha}.$$

Proof. The result follows from the monotonicity of F and the relation;

$$P_h[F(x_\alpha^{h,\delta}) - F(x_\alpha^h) + \alpha(x_\alpha^{h,\delta} - x_\alpha^h)] \quad = \quad P_h(y^\delta - y).$$

Theorem 10.3.2. *Let $\rho < \frac{2}{K_0(1+\frac{\varepsilon_0}{\alpha_0})}$ and $\hat{x} \in D(F)$ be a solution of (10.1.1). And let Assumption 10.1.3, Assumption 10.1.4 and the assumptions in Proposition 10.3.1 be satisfied. Then*

$$\|x_\alpha^h - \hat{x}\| \le \tilde{C}(\varphi(\alpha) + \frac{\varepsilon_h}{\alpha})$$

where $\tilde{C} := \frac{\max\{1+(1+\frac{\varepsilon_0}{\alpha_0})K_0(2b_0+\rho), \rho+\|\hat{x}\|\}}{1-(1+\frac{\varepsilon_0}{\alpha_0})\frac{K_0}{2}\rho}$.

Proof. Let $M := \int_0^1 F'(\hat{x} + t(x_\alpha^h - \hat{x}))dt$. Then from the relation

$$P_h[F(x_\alpha^h) - F(\hat{x}) + \alpha(x_\alpha^h - x_0)] = 0$$

we have,

$$(P_h M P_h + \alpha P_h)(x_\alpha^h - \hat{x}) = P_h \alpha(x_0 - \hat{x}) + P_h M(I - P_h)\hat{x}.$$

Hence,

$$\begin{aligned} x_\alpha^h-\hat{x} &= [(P_hMP_h+\alpha P_h)^{-1}P_h-(F'(x_0)+\alpha I)^{-1}]\alpha(x_0-\hat{x}) \\ &\quad +(F'(x_0)+\alpha I)^{-1}\alpha(x_0-\hat{x}) \\ &\quad +(P_hMP_h+\alpha P_h)^{-1}P_hM(I-P_h)\hat{x} \\ &= (P_hMP_h+\alpha P_h)^{-1}P_h[F'(x_0)-M+M(I-P_h)] \\ &\quad (F'(x_0)+\alpha I)^{-1}\alpha(x_0-\hat{x}) \\ &\quad +(F'(x_0)+\alpha I)^{-1}\alpha(x_0-\hat{x}) \\ &\quad +(P_hMP_h+\alpha P_h)^{-1}P_hM(I-P_h)\hat{x} \\ &:= \zeta_1+\zeta_2 \end{aligned} \tag{10.3.1}$$

where $\zeta_1=(P_hMP_h+\alpha P_h)^{-1}P_h[F'(x_0)-M+M(I-P_h)](F'(x_0)+\alpha I)^{-1}\alpha(x_0-\hat{x})and\zeta_2=(F'(x_0)+\alpha I)^{-1}\alpha(x_0-\hat{x})+(P_hMP_h+\alpha P_h)^{-1}P_hM(I-P_h)\hat{x}$. Observe that,

$$\begin{aligned} \|\zeta_1\| &\leq \|(P_hMP_h+\alpha P_h)^{-1}P_h\int_0^1[F'(x_0)-F'(\hat{x} \\ &\quad +t(x_\alpha^h-\hat{x}))]dt(F'(x_0)+\alpha I)^{-1}\alpha(x_0-\hat{x})\| \\ &\quad +\|(P_hMP_h+\alpha P_h)^{-1}P_hM(I-P_h) \\ &\quad (F'(x_0)+\alpha I)^{-1}\alpha(x_0-\hat{x})\| \\ &\leq \|(P_hMP_h+\alpha P_h)^{-1}P_h \\ &\quad \int_0^1[F'(\hat{x}+t(x_\alpha^h-\hat{x}))(P_h+I-P_h) \\ &\quad \phi(x_0,\hat{x}+t(x_\alpha^h-\hat{x}),(F'(x_0)+\alpha I)^{-1}\alpha(x_0-\hat{x}))]dt\|+\frac{\varepsilon_h}{\alpha}\rho \end{aligned}$$

where, here and below $\varepsilon_h:=\varepsilon_h(\hat{x}+t(x_\alpha^h-\hat{x}))$. So

$$\begin{aligned} \|\zeta_1\| &\leq (1+\frac{\varepsilon_h}{\alpha})K_0\int_0^1[\|x_0-P_hx_0\|+\|\hat{x}+t(x_\alpha^h-\hat{x})-P_hx_0\|] \\ &\quad \|F'(x_0)+\alpha I)^{-1}\alpha(x_0-\hat{x}))\|+\frac{\varepsilon_h}{\alpha}\rho \\ &\leq (1+\frac{\varepsilon_h}{\alpha})K_0[(b_0+\|\hat{x}-x_0+x_0-P_hx_0\|)\varphi(\alpha)+\frac{1}{2}\|x_\alpha^h-\hat{x}\|\rho]+\frac{\varepsilon_h}{\alpha}\rho \\ &\leq (1+\frac{\varepsilon_h}{\alpha})K_0[(2b_0+\rho)\varphi(\alpha)+\frac{1}{2}\|x_\alpha^h-\hat{x}\|\rho]+\frac{\varepsilon_h}{\alpha}\rho \end{aligned} \tag{10.3.2}$$

and

$$\|\zeta_2\|\leq\varphi(\alpha)+\frac{\varepsilon_h}{\alpha}\|\hat{x}\|. \tag{10.3.3}$$

The result now follows from (10.3.1), (10.3.2) and (10.3.3).

Theorem 10.3.3. *Let $x_{n,\alpha}^{h,\delta}$ be as in (10.1.4). And the assumptions in Theorem 10.2.6 and Theorem 10.3.2 hold. Then*

$$\|x_{n,\alpha}^{h,\delta}-\hat{x}\|\leq\frac{1+b}{1-b^2}\gamma_\rho b^{2n}+max\{1,\tilde{C}\}(\varphi(\alpha)+\frac{\delta+\varepsilon_h}{\alpha}).$$

Proof. Observe that,

$$\|x_{n,\alpha}^{h,\delta}-\hat{x}\| \le \|x_{n,\alpha}^{h,\delta}-x_{\alpha}^{h,\delta}\|+\|x_{\alpha}^{h,\delta}-x_{\alpha}^{h}\|+\|x_{\alpha}^{h}-\hat{x}\|$$

so, by Proposition 10.3.1, Theorem 10.2.6 and Theorem 10.3.2 we obtain,

$$\begin{aligned}\|x_{n,\alpha}^{h,\delta}-\hat{x}\| &\le \frac{1+b}{1-b^2}\gamma_\rho b^{2n}+\frac{\delta}{\alpha}+\tilde{C}(\varphi(\alpha)+\frac{\varepsilon_h}{\alpha})\\ &\le \frac{1+b}{1-b^2}\gamma_\rho b^{2n}+max\{1,\tilde{C}\}(\varphi(\alpha)+\frac{\delta+\varepsilon_h}{\alpha}).\end{aligned}$$

Let

$$n_\delta := min\left\{n : b^{2n} \le \frac{\delta+\varepsilon_h}{\alpha}\right\} \tag{10.3.4}$$

and

$$C_0 = \frac{1+b}{1-b^2}\gamma_\rho + max\{1,\tilde{C}\}. \tag{10.3.5}$$

Theorem 10.3.4. *Let n_δ and C_0 be as in (10.3.4) and (10.3.5) respectively. And let $x_{n_\delta,\alpha}^{h,\delta}$ be as in (10.1.4) and the assumptions in Theorem 10.3.3 be satisfied. Then*

$$\|x_{n_\delta,\alpha}^{h,\delta}-\hat{x}\| \le C_0(\varphi(\alpha)+\frac{\delta+\varepsilon_h}{\alpha}). \tag{10.3.6}$$

10.3.1. A Priori Choice of the Parameter

Let $\psi(\lambda) := \lambda\varphi^{-1}(\lambda), 0 < \lambda \le a$. Then the choice

$$\alpha_\delta = \varphi^{-1}(\psi^{-1}(\delta+\varepsilon_h)),$$

gives the optimal order error estimate (see [10]) for $\varphi(\alpha)+\frac{\delta+\varepsilon_h}{\alpha}$. So the relation (10.3.6) leads to the following.

Theorem 10.3.5. *Let $\psi(\lambda) := \lambda\varphi^{-1}(\lambda)$ for $0 < \lambda \le a$, and the assumptions in Theorem 10.3.4 hold. For $\delta > 0$, let $\alpha_\delta = \varphi^{-1}(\psi^{-1}(\delta+\varepsilon_h))$ and let n_δ be as in (10.3.4). Then*

$$\|x_{n_\delta,\alpha}^{h,\delta}-\hat{x}\| = O(\psi^{-1}(\delta+\varepsilon_h)).$$

10.3.2. An Adaptive Choice of the Parameter

As in [10], the parameter α is chosen according to the balancing principle studied in [14], [17], i.e., the parameter α is selected from some finite set

$$D_N(\alpha) := \{\alpha_i = \mu^i\alpha_0, i = 0,1,\cdots,N\}$$

where $\mu > 1$, $\alpha_0 > 0$ and let

$$n_i := min\left\{n : b^{2n} \le \frac{\delta+\varepsilon_h}{\alpha_i}\right\}.$$

Then for $i = 0, 1, \cdots, N$, we have

$$\|x^{h,\delta}_{n_i,\alpha_i} - x^{h,\delta}_{\alpha_i}\| \leq C\frac{\delta+\varepsilon_h}{\alpha_i}, \quad \forall i = 0, 1, \cdots N.$$

Let $x_i := x^{h,\delta}_{n_i,\alpha_i}$. In this paper we select $\alpha = \alpha_i$ from $D_N(\alpha)$ for computing x_i, for each $i = 0, 1, \cdots, N$.

Theorem 10.3.6. *(cf. [18], Theorem 3.1) Assume that there exists $i \in \{0, 1, 2, \cdots, N\}$ such that $\varphi(\alpha_i) \leq \frac{\delta+\varepsilon_h}{\alpha_i}$. Let the assumptions of Theorem 10.3.4 and Theorem 10.3.5 hold and let*

$$l := max\left\{i : \varphi(\alpha_i) \leq \frac{\delta+\varepsilon_h}{\alpha_i}\right\} < N,$$

$$k := max\{i : \|x_i - x_j\| \leq 4C_0\frac{\delta+\varepsilon_h}{\alpha_j}, \quad j = 0, 1, 2, \cdots, i\}.$$

Then $l \leq k$ and $\|\hat{x} - x_k\| \leq c\psi^{-1}(\delta+\varepsilon_h)$ where $c = 6C_0\mu$.

10.4. Implementation of Adaptive Choice Rule

The balancing algorithm associated with the choice of the parameter specified in Theorem 10.3.6 involves the following steps:

- Choose $\alpha_0 > 0$ such that $\delta_0 < \alpha_0$ and $\mu > 1$.
- Choose $\alpha_i := \mu^i \alpha_0, i = 0, 1, 2, \cdots, N$.

10.4.1. Algorithm

1. Set $i = 0$.
2. Choose $n_i := min\left\{n : b^{2n} \leq \frac{\delta+\varepsilon_h}{\alpha_i}\right\}$.
3. Solve $x_i := x^{h,\delta}_{n_i,\alpha_i}$ by using the iteration (10.1.3) and (10.1.4).
4. If $\|x_i - x_j\| > 4C_0\frac{\delta+\varepsilon_h}{\alpha_j}, j < i$, then take $k = i-1$ and return x_k.
5. Else set $i = i+1$ and go to 2.

10.5. Numerical Example

In this section we consider the example considered in [18] for illustrating the algorithm considered in section IV. We apply the algorithm by choosing a sequence of finite dimensional subspace (V_n) of X with $dimV_n = n+1$. Precisely we choose V_n as the linear span of $\{v_1, v_2, \cdots, v_{n+1}\}$ where $v_i, i = 1, 2, \cdots, n+1$ are the linear splines in a uniform grid of $n+1$ points in $[0, 1]$ (see [10] for details).

Example 10.5.1. *(see [18], section 4.3) Let $F : D(F) \subseteq L^2(0,1) \longrightarrow L^2(0,1)$ defined by*

$$F(u) := \int_0^1 k(t,s)u^3(s)ds,$$

where

$$k(t,s) = \begin{cases} (1-t)s, & 0 \le s \le t \le 1 \\ (1-s)t, & 0 \le t \le s \le 1 \end{cases} .$$

Then for all $x(t), y(t) : x(t) > y(t)$:

$$\langle F(x) - F(y), x - y\rangle = \int_0^1 \left[\int_0^1 k(t,s)(x^3 - y^3)(s)ds\right]$$

$$\times (x-y)(t)dt \ge 0.$$

Thus the operator F is monotone. The Fréchet derivative of F is given by

$$F'(u)w = 3\int_0^1 k(t,s)u^2(s)w(s)ds. \tag{10.5.1}$$

As in [10] one can see that F' satisfies the Assumption 10.1.2. In our computation, we take $f(t) = (t - t^{11})/110$ and $f^\delta = f + \delta$. Then the exact solution

$$\hat{x}(t) = t^3.$$

We use

$$x_0(t) = t^3 + \frac{3}{56}(t - t^8)$$

as our initial guess, so that the function $x_0 - \hat{x}$ satisfies the source condition

$$x_0 - \hat{x} = \varphi(F'(\hat{x}))1$$

where $\varphi(\lambda) = \lambda$.

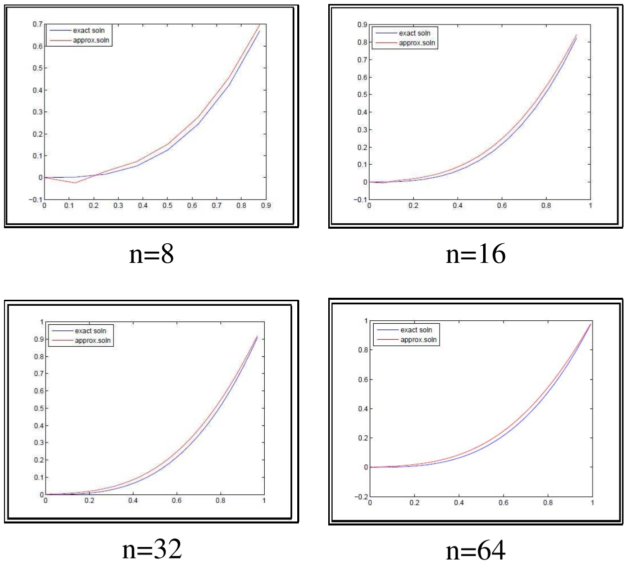
For the operator $F'(.)$ defined in (10.5.1), $\varepsilon_h = O(n^{-2})$ (cf. [11]). Thus we expect to obtain the rate of convergence $O((\delta + \varepsilon_h)^{\frac{1}{2}})$.

We choose $\alpha_0 = (1.1)(\delta + \varepsilon_h)$, $\mu = 1.1$, $\rho = 0.11$, $\gamma_\rho = 0.7818$ and $b = 0.99$. The results of the computation are presented in Table 1. The plots of the exact solution and the approximate solution obtained are given in Figures 1 and 2.

Example 10.5.2. *Let $X = Y = \mathbb{R}$, $D = [0,\infty)$, $x_0 = 1$ and define function F on D by*

$$F(x) = \frac{x^{1+\frac{1}{i}}}{1+\frac{1}{i}} + c_1 x + c_2, \tag{10.5.2}$$

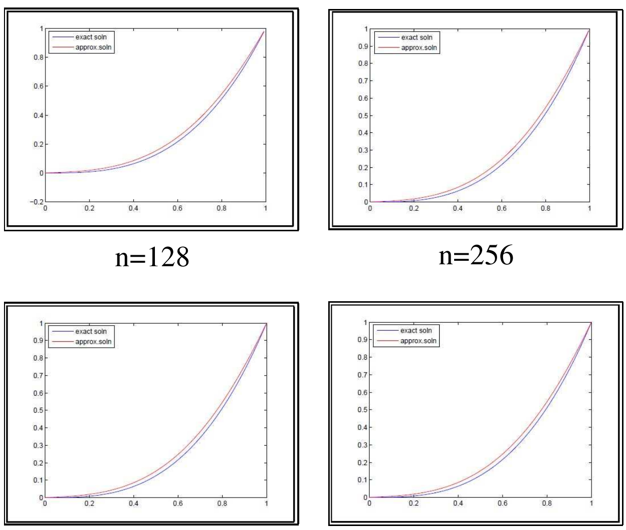
where c_1, c_2 are real parameters and $i > 2$ an integer. Then $F'(x) = x^{1/i} + c_1$ is not Lipschitz on D. However Assumption 10.1.3 holds for $K_0 = 1$.



n=8 n=16

n=32 n=64

Figure 10.5.1. Curves of the exact and approximate solutions.



n=128 n=256

n=512 n=1024

Figure 10.5.2. Curves of the exact and approximate solutions.

Indeed, we have

$$\begin{aligned} \|F'(x)-F'(x_0)\| &= |x^{1/i}-x_0^{1/i}| \\ &= \frac{|x-x_0|}{x_0^{\frac{i-1}{i}}+\cdots+x^{\frac{i-1}{i}}} \end{aligned}$$

so

$$\|F'(x)-F'(x_0)\| \leq K_0|x-x_0|.$$

Table 10.5.1. Iterations and corresponding error estimates

n	**k**	n_k	$\delta+\varepsilon_h$	α	$\|x_k-\hat{x}\|$	$\frac{\|x_k-\hat{x}\|}{(\delta+\varepsilon_h)^{1/2}}$
8	2	2	0.0134	0.0178	0.2217	1.9158
16	2	2	0.0133	0.0178	0.1835	1.5885
32	2	2	0.0133	0.0177	0.1383	1.1981
64	2	2	0.0133	0.0177	0.0998	0.8647
128	2	2	0.0133	0.0177	0.0699	0.6051
256	30	2	0.0133	0.2559	0.0470	0.4070
512	30	2	0.0133	0.2559	0.0290	0.2509
1024	30	2	0.0133	0.2559	0.0121	0.1049

Example 10.5.3. *We consider the integral equations*

$$u(s) = f(s) + \lambda \int_a^b G(s,t)u(t)^{1+1/n}dt, \quad n \in \mathbb{N}. \tag{10.5.3}$$

Here, f is a given continuous function satifying $f(s) > 0, s \in [a,b], \lambda$ is a real number, and the kernel G is continuous and positive in $[a,b]\times[a,b]$.

For example, when $G(s,t)$ is the Green kernel, the corresponding integral equation is equivalent to the boundary value problem

$$\begin{aligned} u'' &= \lambda u^{1+1/n} \\ u(a) &= f(a), u(b) = f(b). \end{aligned}$$

These type of problems have been considered in [1]- [5].

Equation of the form (10.5.3) generalize equations of the form

$$u(s) = \int_a^b G(s,t)u(t)^n dt \tag{10.5.4}$$

studied in [1]-[5]. Instead of (10.5.3) we can try to solve the equation $F(u) = 0$ where

$$F : \Omega \subseteq C[a,b] \to C[a,b], \Omega = \{u \in C[a,b] : u(s) \geq 0, s \in [a,b]\},$$

and

$$F(u)(s) = u(s) - f(s) - \lambda \int_a^b G(s,t)u(t)^{1+1/n}dt.$$

The norm we consider is the max-norm.

The derivative F' is given by

$$F'(u)v(s) = v(s) - \lambda(1+\frac{1}{n})\int_a^b G(s,t)u(t)^{1/n}v(t)dt, \quad v \in \Omega.$$

First of all, we notice that F' does not satisfy a Lipschitz-type condition in Ω. Let us consider, for instance, $[a,b]=[0,1], G(s,t)=1$ and $y(t)=0$. Then $F'(y)v(s)=v(s)$ and

$$\|F'(x)-F'(y)\| = |\lambda|(1+\frac{1}{n})\int_a^b x(t)^{1/n}dt.$$

If F' were a Lipschitz function, then

$$\|F'(x)-F'(y)\| \leq L_1\|x-y\|,$$

or, equivalently, the inequality

$$\int_0^1 x(t)^{1/n}dt \leq L_2 \max_{x\in[0,1]} x(s), \tag{10.5.5}$$

would hold for all $x\in\Omega$ and for a constant L_2. But this is not true. Consider, for example, the functions

$$x_j(t)=\frac{t}{j}, \;\; j\geq 1, \;\; t\in[0,1].$$

If these are substituted into (10.5.5)

$$\frac{1}{j^{1/n}(1+1/n)} \leq \frac{L_2}{j} \Leftrightarrow j^{1-1/n} \leq L_2(1+1/n), \;\; \forall j\geq 1.$$

This inequality is not true when $j\to\infty$.

Therefore, condition (10.5.5) is not satisfied in this case. However, Assumption 10.1.3 holds. To show this, let $x_0(t)=f(t)$ and $\gamma=\min_{s\in[a,b]} f(s), \alpha>0$ Then for $v\in\Omega$,

$$\begin{aligned}\|[F'(x)-F'(x_0)]v\| &= |\lambda|(1+\frac{1}{n})\max_{s\in[a,b]}|\int_a^b G(s,t)(x(t)^{1/n}-f(t)^{1/n})v(t)dt| \\ &\leq |\lambda|(1+\frac{1}{n})\max_{s\in[a,b]} G_n(s,t)\end{aligned}$$

where $G_n(s,t)=\frac{G(s,t)|x(t)-f(t)|}{x(t)^{(n-1)/n}+x(t)^{(n-2)/n}f(t)^{1/n}+\cdots+f(t)^{(n-1)/n}}\|v\|$.

Hence,

$$\begin{aligned}\|[F'(x)-F'(x_0)]v\| &= \frac{|\lambda|(1+1/n)}{\gamma^{(n-1)/n}}\max_{s\in[a,b]}\int_a^b G(s,t)dt\|x-x_0\| \\ &\leq K_0\|x-x_0\|,\end{aligned}$$

where $K_0=\frac{|\lambda|(1+1/n)}{\gamma^{(n-1)/n}}N$ and $N=\max_{s\in[a,b]}\int_a^b G(s,t)dt$. Then Assumption 10.1.3 holds for sufficiently small λ.

Example 10.5.4. *Let $X=D(F)=\mathbb{R}$, $x_0=0$, and define function F on $D(F)$ by*

$$F(x)=d_0x+d_1+d_2\sin e^{d_3x},$$

where d_0,d_1,d_2 and d_3 are given parameters. Then, it can easily be seen that for d_3 sufficiently large and d_1 sufficiently small, $\frac{K_0}{k_0}$ can be arbitrarily small.

References

[1] Argyros, I.K., *Convergenve and Applications of Newton-type Iterations*, Springer, New York, 2008.

[2] Argyros, I.K., Approximating solutions of equations using Newton's method with a modified Newton's method iterate as a starting point. *Rev. Anal. Numer. Theor. Approx.* 36, (2007), 123-138.

[3] Argyros, I.K., A Semilocal convergence for directional Newton methods, *Math. Comput. (AMS).* 80, (2011), 327-343.

[4] Argyros, I.K., and Hilout, S., Weaker conditions for the convergence of Newton's method, *J. Complexity*, 28, (2012), 364-387.

[5] Argyros, I.K., Cho, Y.J., and Hilout, S., Numerical methods for equations and its applications, CRC Press, Taylor and Francis, New York, 2012.

[6] George, S., Elmahdy, A.I., An analysis of Lavrentiev regularization for nonlinear ill-posed problems using an iterative regularization method, *Int. J. Comput. Appl. Math.*, 5(3) (2010),369-381.

[7] George, S., Elmahdy, A.I., An iteratively regularized projection method for nonlinear ill-posed problems", *Int. J. Contemp. Math. Sciences*, 5 (52) (2010), 2547-2565.

[8] George, S., Elmahdy, A.I., A quadratic convergence yielding iterative method for nonlinear ill-posed operator equations, *Comput. Methods Appl. Math*, 12(1) (2012), 32-45

[9] George, S., Elmahdy, A.I., An iteratively regularized projection method with quadratic convergence for nonlinear ill-posed problems, *Int. J. of Math. Analysis*, 4(45) (2010), 2211-2228.

[10] George, S., Pareth, S., An application of Newton type iterative method for Lavrentiev regularization for ill-posed equations: Finite dimensional realization, *IJAM*, 42(3) (2012), 164-170.

[11] Groetsch, C.W., King, J.T., Murio, D., Asymptotic analysis of a finite element method for Fredholm equations of the first kind, in Treatment of Integral Equations by Numerical Methods, Eds.: C.T.H. Baker and G.F. Miller, Academic Press, London, 1982, 1-11.

[12] Jaan, J., Tautenhahn, U., On Lavrentiev regularization for ill-posed problems in Hilbert scales, *Numer. Funct. Anal. Optim.*, 24(5-6) (2003), 531-555.

[13] Kelley, C.T., Iterative methods for linear and nonlinear equations, *SIAM* Philadelphia, 1995.

[14] Mathe, P. Perverzev, S.V., Geometry of linear ill-posed problems in variable Hilbert scales, *Inverse Problems*, 19(3) (2003), 789-803.

[15] Mahale, P., Nair, M. T., Iterated Lavrentiev regularization for nonlinear ill-posed problems, *ANZIAM Journal*, 51 (2009), 191-217.

[16] Pareth, S., George, S., *"Newton type methods for Lavrentiev regularization of nonlinear ill-posed operator equations"*, (2012), (submitted).

[17] Perverzev,S.V., Schock, E., On the adaptive selection of the parameter in regularization of ill-posed problems, *SIAM J.Numer.Anal.*, 43 (2005), 2060-2076.

[18] Semenova, E.V., Lavrentiev regularization and balancing principle for solving ill-posed problems with monotone operators, *Comput. Methods Appl. Math.*, 4 (2010), 444-454.

[19] Tautanhahn, U., On the method of Lavrentiev regularization for nonlinear ill-posed problems, *Inverse Problems*, 18 (2002), 191-207.

Chapter 11

New Approach to Relaxed Proximal Point Algorithms Based on $A-$Maximal

11.1. Introduction

Let X be a real Hilbert space with the norm $\|\cdot\|$ and the inner product $\langle .,. \rangle$. Here we consider the inclusion problem of the form: find a solution to

$$0 \in M(x), \tag{11.1.1}$$

where $M : X \to 2^X$ is a set-valued mapping on X.

Based on the work of Rockafellar [11] on the proximal point algorithm and its applications to certain computational methods, Eckstein and Bertsekas [3] introduced the relaxed proximal point algorithm and then they applied to the Douglas-Rachford splitting method for finding zero of the sum of two monotone operators. Furthermore, they showed that it was, in fact, a special case of the proximal point algorithm. Fukushima [6] applied the primal Douglas-Rachford splitting method for a class of monotone operators with applications to the traffic equilibrium problem.

Highly motivated by these algorithmic developments (see [1-18] and references therein), we generalize the relaxed proximal point algorithm based on the notions of $A-$maximal monotonicity (also referred to as $A-$monotonicity in literature [16]) and $(A,\eta)-$maximal monotonicity (also referred to as $(A,\eta)-$monotonicity [15]) for solving general inclusion problems in Hilbert space settings. These concepts generalize the general theory of maximal monotone set-valued mappings in a Hilbert space setting. Our approach differs significantly than the one used by Rockafellar [11], where the locally Lipschitz type condition on the mapping M^{-1} is imposed achieving the convergence rate estimate. The main ingredients for our approach consist of the more generalized framework for the relaxed proximal point algorithm based on the $A-$maximal monotonicity, and considering the convergence rate as a quadratic polynomial in terms of α_k, where $\{\alpha_k\}$ is a scalar sequence.

The notion of $A-$maximal monotonicity was introduced and studied by Verma [16] in the context of solving variational inclusion problems using the resolvent operator technique, while this work was followed by an accelerated research developments. Furthermore it generalizes the existing theory of maximal monotone operators (based on the classical resolvent), including the $H-$maximal monotonicity by Fang and Huang [4] that concerns with the generalization of the classical maximal monotonicity. Fang and Huang [4] introduced the notion of $H-$maximal monotonicity, while investigating the solvability of a general class of inclusion problems. They applied $(H,\eta)-$maximal monotonicity [5] in the context of approximating the solutions of inclusion problems using the generalized resolvent operator technique. The generalized resolvent operator technique is equally effective applying to several other problems, such as equilibria problems in economics, global optimization and control theory, operations research, mathematical finance, management and decision sciences, mathematical programming, and engineering science. For more details on the resolvent operator technique and its applications, and further developments, we refer the reader to [1- 33] and references therein.

11.2. $A-$Maximal Monotonicity and Auxiliary Results

In this section we discuss some results based on the basic properties and auxiliary results on $A-$ maximal monotonicity (also referred to as $A-$ monotonicity in literature) and its variant forms. Let $M: X \to 2^X$ be a multivalued mapping on X. We shall denote both the map M and its graph by M, that is, the set $\{(x,y): y \in M(x)\}$. This is equivalent to stating that a mapping is any subset M of $X \times X$, and $M(x) = \{y : (x,y) \in M\}$. If M is single-valued, we shall still use $M(x)$ to represent the unique y such that $(x,y) \in M$ rather than the singleton set $\{y\}$. This interpretation shall much depend on the context. The domain of a map M is defined (as its projection onto the first argument) by

$$D(M) = \{x \in X : \exists y \in X : (x,y) \in M\} = \{x \in X : M(x) \neq \emptyset\}.$$

dom(M)=X, shall denote the full domain of M, and the range of M is defined by

$$R(M) = \{y \in X : \exists x \in X : (x,y) \in M\}.$$

The inverse M^{-1} of M is $\{(y,x) : (x,y) \in M\}$. For a real number ρ and a mapping M, let $\rho M = \{(x,\rho y) : (x,y) \in M\}$. If L and M are any mappings, we define

$$L+M = \{(x,y+z) : (x,y) \in L, (x,z) \in M\}.$$

Definition 11.2.1. *Let $M: X \to 2^X$ be a multivalued mapping on X. The map M is said to be:*

(i) $(r)-$ *strongly monotone if there exists a positive constant r such that*

$$\langle u^* - v^*, u - v\rangle \geq r\|u-v\|^2 \,\forall (u,u^*),(v,v^*) \in graph(M).$$

(ii) $(m)-$*relaxed monotone if there exists a positive constant m such that*

$$\langle u^* - v^*, u - v\rangle \geq (-m)\|u-v\|^2 \,\forall (u,u^*),(v,v^*) \in graph(M).$$

Definition 11.2.2. ([16]). Let $A : X \to X$ be a single-valued mapping. The map $M : X \to 2^X$ is said to be $A-$ maximal monotone if

(i) M is $(m)-$relaxed monotone for $m > 0$,

(ii) $R(A+\rho M) = X$ for $\rho > 0$.

Example 11.2.1. Let $A : X \to X$ be an $(r)-$strongly monotone mapping on X for $r > 0$. Let $f : X \to R$ be a locally Lipschitz functional such that ∂f, the subdifferential of f, is $(m)-$relaxed monotone, where $m > 0$. Then $A+\partial f$ is $(r-m)-$strongly monotone for $r-m > 0$. Then it follows that $A+\partial f$ is pseudomonotone, which is, in fact, maximal monotone. This is equivalent to stating that ∂f is $A-$maximal monotone.

Definition 11.2.3. ([16]). Let $A : X \to X$ be an $(r)-$strongly monotone mapping and let $M : X \to 2^X$ be an $A-$maximal monotone mapping. Then the generalized resolvent operator $J^M_{\rho,A} : X \to X$ is defined by

$$J^M_{\rho,A}(u) = (A+\rho M)^{-1}(u).$$

Definition 11.2.4. ([5]). Let $H : X \to X$ be a single-valued mapping. The map $M : X \to 2^X$ is said to be to $H-$ maximal monotone if

(i) M is monotone,

(ii) $R(H+\rho M) = X$ for $\rho > 0$.

Definition 11.2.5. ([4]). Let $H : X \to X$ be an $(r)-$strongly monotone mapping and let $M : X \to 2^X$ be an $H-$ monotone mapping. Then the generalized resolvent operator $J^M_{\rho,H} : X \to X$ is defined by

$$J^M_{\rho,H}(u) = (H+\rho M)^{-1}(u).$$

Proposition 11.2.1. ([18]). Let $A : X \to X$ be an $(r)-$strongly monotone mapping and let $M : X \to 2^X$ be an $A-$ maximal monotone mapping. Then $(A+\rho M)$ is maximal monotone for $\rho > 0$.

Proposition 11.2.2. ([18]) Let $A : X \to X$ be an $(r)-$strongly monotone mapping and let $M : X \to 2^X$ be an $A-$- maximal monotone mapping. Then the operator $(A+\rho M)^{-1}$ is single-valued for $r-\rho m > 0$.

Proposition 11.2.3. ([4]) Let $H : X \to X$ be an $(r)-$strongly monotone mapping and let $M : X \to 2^X$ be an $H-$ maximal monotone mapping. Then $(H+\rho M)$ is maximal monotone for $\rho > 0$.

Proposition 11.2.4. ([4]) Let $H : X \to X$ be an $(r)-$strongly monotone mapping and let $M : X \to 2^X$ be an $H-$- maximal monotone mapping. Then the operator $(H+\rho M)^{-1}$ is single-valued.

11.3. The Generalized Relaxed Proximal Point Algorithm

This section deals with an introduction of a generalized version of the relaxed proximal point algorithm and its applications to approximation solvability of the inclusion problem (11.1.1) based on the A-maximal monotonicity.

Lemma 11.3.1. ([18]) *Let X be a real Hilbert space, let $A : X \to X$ be $(r)-$strongly monotone, and let $M : X \to 2^X$ be $A-$ maximal monotone. Then the generalized resolvent operator associated with M and defined by*

$$J^M_{\rho,A}(u) = (A+\rho M)^{-1}(u)\,\forall u \in X,$$

is $(\frac{1}{r-\rho m})-$ Lipschitz continuous for $r-\rho m > 0$.

Lemma 11.3.2. *Let X be a real Hilbert space, let $A : X \to X$ be $(r)-$strongly monotone and $(s)-$Lipschitz continuous, and let $M : X \to 2^X$ be $A-$ maximal monotone. Then the generalized resolvent operator associated with M and defined by*

$$J^M_{\rho,A}(u) = (A+\rho M)^{-1}(u)\,\forall u \in X$$

satisfies

$$\|J^M_{\rho,A}(A(u)) - J^M_{\rho,A}(A(v))\| \leq \frac{1}{r-\rho m}\|A(u)-A(v)\|, \tag{11.3.1}$$

where $r-\rho m > 0$.

Theorem 11.3.3. *Let X be a real Hilbert space, let $A : X \to X$ be $(r)-$strongly monotone, and let $M : X \to 2^X$ be $A-$ maximal monotone. Then the following statements are equivalent:*

(i) An element $u \in X$ is a solution to (11.1.1).

(ii) For an $u \in X$, we have

$$u = J^M_{\rho,A}(A(u)),$$

where

$$J^M_{\rho,A}(u) = (A+\rho M)^{-1}(u).$$

Theorem 11.3.4. *Let X be a real Hilbert space, let $H : X \to X$ be $(r)-$strongly monotone, and let $M : X \to 2^X$ be $H-$ maximal monotone. Then the following statements are equivalent:*

(i) An element $u \in X$ is a solution to (11.1.1).

(ii) For an $u \in X$, we have

$$u = J^M_{\rho,H}(H(u)),$$

where

$$J^M_{\rho,H}(u) = (H + \rho M)^{-1}(u).$$

Lemma 11.3.5. *Let X be a real Hilbert space, let $A : X \to X$ be $(r)-$strongly monotone and $(s)-$ Lipschitz continuous, and let $M : X \to 2^X$ be A- maximal monotone. Then*

$$\langle (J^M_{\rho,A} oA)(u) - (J^M_{\rho,A} oA)(v), A(u) - A(v)\rangle \leq \frac{1}{r - \rho m}\|A(u) - A(v)\|^2 \forall u, v \in X,$$

where $r - \rho m > 0$.

Lemma 11.3.6. *Let X be a real Hilbert space, let $H : X \to X$ be $(r)-$strongly monotone and $(s)-$ Lipschitz continuous, and let $M : X \to 2^X$ be H- maximal monotone. Then*

$$\langle (J^M_{\rho,H} oH)(u) - (J^M_{\rho,H} oH)(v), H(u) - H(v)\rangle \leq \frac{1}{r}\|H(u) - H(v)\|^2 \forall u, v \in X.$$

In the following theorem, we apply the generalized relaxed proximal point algorithm to approximate the solution of (11.1.1), and as a result, we succeed achieving linear convergence.

Theorem 11.3.7. *Let X be a real Hilbert space, let $A : X \to X$ be $(r)-$strongly monotone and $(s)-$Lipschitz continuous, and let $M : X \to 2^X$ be $A-$ maximal monotone. For an arbitrarily chosen initial point x^0, suppose that the sequence $\{x^k\}$ is generated by the generalized proximal point algorithm*

$$A(x^{k+1}) = (1 - \alpha_k)A(x^k) + \alpha_k y^k \;\forall\, k \geq 0, \tag{11.3.2}$$

and y^k satisfies

$$\|y^k - A(J^M_{\rho_k,A}(A(x^k)))\| \leq \delta_k \|y^k - A(x^k)\|,$$

where $J^M_{\rho_k,A} = (A + \rho_k M)^{-1}$, and

$$\{\delta_k\}, \{\alpha_k\}, \{\rho_k\} \subseteq [0, \infty)$$

are scalar sequences such that for $\gamma \in (0, \frac{1}{2})$, $\alpha_k \leq \gamma$,

$$r - \rho_k m \geq 1 + \frac{2\gamma^2(s^2 - 1)}{1 - 2\gamma + \sqrt{(1 - 2\gamma)^2 - 4\gamma^4(s^2 - 1)}}, \tag{11.3.3}$$

$$1 < s \leq \sqrt{1 + \Big(\frac{1 - 2\gamma}{2\gamma^2}\Big)^2}, \tag{11.3.4}$$

$\sum_{k=0}^{\infty} \delta_k < \infty$, $\delta_k \to 0$, *and* $\alpha = \limsup_{k\to\infty} \alpha_k$, *and* $\rho = \limsup_{k\to\infty} \rho_k$.

Then the sequence $\{x^k\}$ *converges linearly to a solution* x^* *of* (11.1.1) *with the convergence rate*

$$\theta_k = \sqrt{(1-\alpha_k)^2 + 2\alpha_k(1-\alpha_k)\frac{1}{(r-\rho_k m)} + \alpha_k^2 \frac{s^2}{(r-\rho_k m)^2}}$$
$$= \frac{1}{r-\rho_k m}\sqrt{\Big(s^2+(r-\rho_k m)^2 - 2(r-\rho_k m)\Big)\alpha_k^2 - 2\Big(1-(r-\rho_k m)\Big)\alpha_k + 1}$$
$$= \frac{1}{r-\rho_k m}\sqrt{P_k(\alpha_k)} \in (0,1),$$

where

$$P_k(\alpha_k) = \Big(s^2+(r-\rho_k m)^2 - 2(r-\rho_k m)\Big)\alpha_k^2 - 2\Big(1-(r-\rho_k m)\Big)\alpha_k + 1$$
$$= \Big(1-\alpha_k(r-\rho_k m - 1)\Big)^2 + \alpha_k^2(s^2-1).$$

Proof. Note that it follows from hypotheses (11.3.2) and (11.3.3) that $\theta_k \in (0,1)$. Suppose that x^* is a zero of M. Then from Theorem 11.3.1, it follows that any solution to (11.1.1) is a fixed point of $J^M_{\rho_k,A} oA$. For all $k \geq 0$, we express

$$A(z^{k+1}) = (1-\alpha_k)A(x^k) + \alpha_k A(J^M_{\rho_k,A}(A(x^k))).$$

Next, applying Lemma 11.3.2, we find the estimate

$$\begin{aligned}
&\|A(z^{k+1}) - A(x^*)\|^2 = \|(1-\alpha_k)A(x^k) + \alpha_k A(J^M_{\rho_k,A}(A(x^k)))\\
-\ & [(1-\alpha_k)A(x^*) + \alpha_k A(J^M_{\rho_k,A}(A(x^*)))]\|^2\\
=\ & \|(1-\alpha_k)(A(x^k)-A(x^*)) + \alpha_k(A(J^M_{\rho_k,A}(A(x^k))) - A(J^M_{\rho_k,A}(A(x^*))))\|^2\\
=\ & (1-\alpha_k)^2\|A(x^k)-A(x^*)\|^2\\
+\ & 2\alpha_k(1-\alpha_k)\langle A(x^k)-A(x^*), A(J^M_{\rho_k,A}(A(x^k))) - A(J^M_{\rho_k,A}(A(x^*)))\rangle\\
+\ & \alpha_k^2\|A(J^M_{\rho_k,A}(A(x^k))) - A(J^M_{\rho_k,A}(A(x^*)))\|^2\\
\leq\ & (1-\alpha_k)^2\|A(x^k)-A(x^*)\|^2 + 2\alpha_k(1-\alpha_k)\frac{1}{(r-\rho_k m)}\|A(x^k)-A(x^*)\|^2\\
+\ & \alpha_k^2 s^2\|J^M_{\rho_k,A}(A(x^k)) - J^M_{\rho_k,A}(A(x^*))\|^2\\
\leq\ & (1-\alpha_k)^2\|A(x^k)-A(x^*)\|^2 + 2\alpha_k(1-\alpha_k)\frac{1}{(r-\rho_k m)}\|A(x^k)-A(x^*)\|^2\\
+\ & \alpha_k^2\frac{s^2}{(r-\rho_k m)^2}\|A(x^k)A(x^*)\|^2\\
=\ & \Big[(1-\alpha_k)^2 + 2\alpha_k(1-\alpha_k)\frac{1}{(r-\rho_k m)} + \alpha_k^2\frac{s^2}{(r-\rho_k m)^2}\Big]\|A(x^k)-A(x^*)\|^2\\
=\ & \frac{1}{(r-\rho_k m)^2}\Big[\Big(s^2+(r-\rho_k m)^2 - 2(r-\rho_k m)\Big)\alpha_k^2 - 2\Big(1-(r-\rho_k m)\Big)\alpha_k + 1\Big]\\
\cdot\ & \|A(x^k)-A(x^*)\|^2\\
=\ & \theta_k^2\|A(x^k)-A(x^*)\|^2,
\end{aligned}$$

where

$$\theta_k^2 = \frac{P_k(\alpha_k)}{(r-\rho_k m)^2}.$$

Thus, we have

$$\begin{aligned} & \|A(z^{k+1}) - A(x^*\| \le \theta_k \|A(x^k) - A(x^*)\| \\ = & \frac{1}{r-\rho_k m}\sqrt{P_k(\alpha_k)}\|A(x^k) - A(x^*)\|. \end{aligned} \tag{11.3.5}$$

Since $A(x^{k+1}) = (1-\alpha_k)A(x^k) + \alpha_k y^k$, we have $A(x^{k+1}) - A(x^k) = \alpha_k(y^k - A(x^k))$. It follows that

$$\begin{aligned} & \|A(x^{k+1}) - A(z^{k+1})\| \\ = & \|(1-\alpha_k)A(x^k) + \alpha_k y^k - [(1-\alpha_k)A(x^k) + \alpha_k A(J^M_{\rho_k,A}(A(x^k)))]\| \\ = & \|\alpha_k(y^k - A(J^M_{\rho,A}(A(x^k))))\| \\ \le & \alpha_k \delta_k \|y^k - A(x^k)\|. \end{aligned}$$

Next, we estimate using the above arguments that

$$\begin{aligned} & \|A(x^{k+1}) - A(x^*)\| \\ \le & \|A(z^{k+1}) - A(x^*)\| + \|A(x^{k+1}) - A(z^{k+1})\| \\ \le & \|A(z^{k+1}) - A(x^*)\| + \alpha_k\delta_k\|y^k - A(x^k)\| \\ = & \|A(z^{k+1}) - A(x^*)\| + \delta_k\|A(x^{k+1}) - A(x^k)\| \\ \le & \|A(z^{k+1}) - A(x^*)\| + \delta_k\|A(x^{k+1}) - A(x^*)\| + \delta_k\|A(x^k) - A(x^*)\|. \end{aligned} \tag{11.3.6}$$

This implies from (11.3.6) on applying (11.3.5) that

$$\begin{aligned} & (1-\delta_k)\|A(x^{k+1}) - A(x^*)\| \\ \le & \|A(z^{k+1}) - A(x^*)\| + \delta_k\|A(x^k) - A(x^*)\| \\ \le & \theta_k\|A(x^k) - A(x^*)\| + \delta_k\|A(x^k) - A(x^*)\| \\ = & \Big(\theta_k + \delta_k\Big)\|A(x^k) - A(x^*)\|. \end{aligned} \tag{11.3.7}$$

Therefore,

$$\|A(x^{k+1}) - A(x^*)\| \le \frac{(\theta_k + \delta_k)}{1-\delta_k}\|A(x^k) - A(x^*)\|, \tag{11.3.8}$$

where

$$\begin{aligned} & \limsup \frac{(\theta_k + \delta_k)}{1-\delta_k} = \limsup \theta_k \\ = & \frac{1}{(r-\rho_k m)}\sqrt{P_k(\alpha_k)}. \end{aligned} \tag{11.3.9}$$

P_k is a quadratic polynomial for each k whose leading coefficient

$$\Big(s^2 + (r-\rho_k m)^2 - 2(r-\rho_k m)\Big) = \Big(1-(r-\rho_k m)\Big)^2 + s^2 - 1$$

is positive since $s > 1$. Hence, each P_k has a minimum which is given by

$$\begin{aligned} & \frac{\Big(s^2+(r-\rho_k m)^2-2(r-\rho_k m)\Big)-\Big(1-(r-\rho_k m)\Big)^2}{\Big(s^2+(r-\rho_k m)^2-2(r-\rho_k m)\Big)} \\ = \quad & 1-\frac{\Big(1-(r-\rho_k m)\Big)^2}{\Big(1-(r-\rho_k m)\Big)^2+s^2-1} < 1. \end{aligned}$$

Now, it follows from (11.3.8) in light of (11.3.9) that the sequence $\{A(x^k)\}$ converges to $A(x^*)$. On the other hand, A is $(r)-$strongly monotone (and hence, $\|A(x)-A(y)\| \geq r\|x-y\|$), we have that

$$\|x^k - x^*\| \leq \frac{\theta_k}{r}\|A(x^k)-A(x^*)\| \to 0, \tag{11.3.10}$$

which completes the proof. □

Corollary 11.3.8. Let X be a real Hilbert space, let $H : X \to X$ be $(r)-$strongly monotone and $(s)-$Lipschitz continuous, and let $M : X \to 2^X$ be $H-$ maximal monotone. For an arbitrarily chosen initial point x^0, suppose that the sequence $\{x^k\}$ is generated by the generalized proximal point algorithm

$$H(x^{k+1}) = (1-\alpha_k)H(x^k)+\alpha_k y^k \,\forall\, k \geq 0, \tag{11.3.11}$$

and y^k satisfies

$$\|y^k - H(J^M_{\rho_k,H}(H(x^k)))\| \leq \delta_k\|y^k - H(x^k)\|,$$

where $J^M_{\rho_k,H} = (H+\rho_k M)^{-1}$, and

$$\{\delta_k\},\{\alpha_k\},\{\rho_k\} \subseteq [0,\infty)$$

are scalar sequences such that for $\gamma \in (0,\frac{1}{2})$, $\alpha_k \leq \gamma$,

$$r \geq 1+\frac{2\gamma^2(s^2-1)}{1-2\gamma+\sqrt{(1-2\gamma)^2-4\gamma^4(s^2-1)}},$$

$$1 < s \leq \sqrt{1+\Big(\frac{1-2\gamma}{2\gamma^2}\Big)^2}.$$

Then the sequence $\{x^k\}$ converges linearly to a solution of $(11.1.1)$ with convergence rate

$$\theta_k = \frac{1}{r}\sqrt{(s^2+r^2-2r)\alpha_k^2-2(1-r)\alpha_k+1},$$

where $\sum_{k=0}^{\infty}\delta_k < \infty$, $\delta_k \to 0$, and $\alpha = \lim sup_{k\to\infty}\alpha_k$, and $\rho = \lim sup_{k\to\infty}\rho_k$.

11.4. An Application

Let X be a real Hilbert space and let $f : X \to R$ be a locally Lipschitz functional on X. We consider the inclusion problem: determine a solution to

$$0 \in \partial f(x), \tag{11.4.1}$$

where $\partial f : X \to 2^X$ is a set-valued mapping on X. Then it turns out that $A + \partial f$ is $(r-m)-$strongly monotone for $r - m > 0$, if $A : X \to X$ is $(r)-$strongly monotone, and $\partial f : X \to 2^X$ is $(m)-$relaxed monotone. This is equivalent to stating that ∂f is $A-$maximal monotone. Now all the conditions for Theorem 11. 3.3 are satisfied, one can apply Theorem 11.3.3 to the solvability of $(11.4.1)$ in the form:

Theorem 11.4.1. Let X be a real Hilbert space, and let $A : X \to X$ be $(r)-$strongly monotone and $(s)-$Lipschitz continuous. Let $f : X \to R$ be a locally Lipschitz functional on X, and let $\partial f : X \to 2^X$ be $A-$ maximal monotone. For an arbitrarily chosen initial point x^0, suppose that the sequence $\{x^k\}$ is generated by the generalized proximal point algorithm

$$A(x^{k+1}) = (1-\alpha_k)A(x^k) + \alpha_k y^k \, \forall \, k \geq 0, \tag{11.4.2}$$

and y^k satisfies

$$\|y^k - A(J^{\partial f}_{\rho_k,A}(A(x^k)))\| \leq \delta_k \|y^k - A(x^k)\|,$$

where $J^{\partial f}_{\rho_k,A} = (A + \rho_k \partial f)^{-1}$, and

$$\{\delta_k\}, \{\alpha_k\}, \{\rho_k\} \subseteq [0, \infty)$$

are scalar sequences such that for $\gamma \in (0, \frac{1}{2})$, $\alpha_k \leq \gamma$,

$$r - \rho_k m \geq 1 + \frac{2\gamma^2(s^2-1)}{1 - 2\gamma + \sqrt{(1-2\gamma)^2 - 4\gamma^4(s^2-1)}},$$

$$1 < s \leq \sqrt{1 + \left(\frac{1-2\gamma}{2\gamma^2}\right)^2}.$$

Then the sequence $\{x^k\}$ converges linearly to a solution of $(11.4.1)$ with convergence rate given in Theorem 11.3.3.

References

[1] Agarwal, R.P., Verma,R.U., General system of $(A,\eta)-$maximal relaxed monotone variational inclusion problems based on generalized hybrid algorithms, *Communications in Nonlinear Science and Numerical Simulations* 15 (2010), 238–251.

[2] Argyros, I.K., Cho, Y.J., Hilout, S., *Numerical Methods for Equations and its Applications*, CRC Press, Taylor & Francis, New York, 2012.

[21] Dhage, B.C., Verma, R.U., Second order boundary value problems of discontinuous differential inclusions, *Communications on Applied Nonlinear Analysis* **12**(3) (2005), 37-44.

[3] Eckstein, J., Bertsekas, D.P., On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators, *Mathematical Programming* 55 (1992), 293–318.

[4] Fang, Y.P., Huang, N.J., $H-$ monotone operators and system of variational inclusions, *Communications on Applied Nonlinear Analysis* 11(1) (2004), 93–101.

[5] Fang, Y.P., Huang, N.J., Thompson, H.B., A new system of variational inclusions with $(H,\eta)-$ monotone operators, *Computers and Mathematics with Applications* 49(2-3) (2005), 365–374.

[6] Fukushima, M., The primal Douglas-Rachford splitting algorithm for a class of monotone operators with applications to the traffic equilibrium problem, *Mathematical Programming* 72(1996), 1–15.

[7] Glowinski, R., Le Tellec, P., *Augmented Lagrangians and Operator-Splitting Methods in Continuum Mechanics,* SIAM, Philadelphia, PA, 1989.

[8 Lan, H.Y., Kim, J.H., Cho, Y.J., On a new class of nonlinear $A-$monotone multivalued variational inclusions, *Journal of Mathematical Analysis and Applications* 327(1) (2007), 481–493.

[9] Moudafi, A., Mixed equilibrium problems: Sensitivity analysis and algorithmic aspect,*Computers and Mathematics with Applications* 44 (2002), 1099-1108.

[10] Robinson,S.M., Composition duality and maximal monotonicity, *Mathematical Programming* 85 (1999a), 1-13.

[11] Rockafellar, R.T., Monotone operators and the proximal point algorithm, *SIAM Journal of Control and Optimization* 14 (1976), 877-898.

[12] Rockafellar, R.T., Augmented Lagrangians and applications of the proximal point algorithm in convex programming, *Mathematics of Operations Research* 1 (1976b), 97–116.

[13] Tseng, P, Alternating projection-proximal methods for convex programming and variational inequalities, *SIAM Journal of Optimization* 7 (1997), 951–965.

[14] Tseng, P., A modified forward-backward splitting method for maximal monotone mappings, *SIAM Journal of Control and Optimization* 38 (2000), 431–446.

[15] Verma, R.U., Sensitivity analysis for generalized strongly monotone variational inclusions based on the $(A,\eta)-$ resolvent operator technique, *Applied Mathematics Letters* 19 (2006), 1409–1413.

[16] Verma, R.U., $A-$ monotonicity and its role in nonlinear variational inclusions, *Journal of Optimization Theory and Applications* 129(3) (2006), 457-467.

[17] Verma, R.U., General system of $A-$ monotone nonlinear variational inclusion problems, *Journal of Optimization Theory and Applications* 131(1) (2006), 151-157.

[18] Verma, R.U., $A-$ monotone nonlinear relaxed cocoercive variational inclusions, *Central European Journal of Mathematics* 5(2) (2007), 1-11.

[19] Verma, R.U., General system of $(A,\eta)-$monotone variational inclusion problems based on generalized hybrid algorithm, *Nonlinear Analysis: Hybrid Systems* 1 (3) (2007), 326-335.

[20] Verma, R.U., Approximation solvability of a class of nonlinear set-valued inclusions involving $(A,\eta)-$ monotone mappings, *Journal of Mathematical Analysis and Applications* 337 (2008), 969-975.

[22] Verma, R.U., Auxiliary problem principle and its extension applied to variational inequalities, *Mathematical Sciences Research Hot-Line* 4(2) (2000), 55-63.

[23] Verma, R.U., On a class of nonlinear variational inequalities involving partially relaxed monotone and partially strongly monotone mappings, *Mathematical Sciences Research Hot-Line* 3 (12) (1999), 7-26.

[24] Verma, R.U., A class of generalized implicit variational inequality type algorithms and their applications, *Mathematical Sciences Research Hot-Line* 4 (3) (2000), 17-30.

[25] Verma, R.U., RKKM mappings and their applications, *Mathematical Sciences Research Hot-Line* 4(10) (2000), 23–27.

[26] Verma, R.U., A class of new minimax inequalities in generalized H-spaces, *Mathematical Sciences Research Hot-Line* 4(10) (2000), 29-32.

[27] Verma, R.U., On a system of nonlinear variational inequalities involving multivalued mappings, *Mathematical Sciences Research Hot-Line* 4(12) (2000), 21-31.

[28] Verma, R.U., A class of generalized iterative algorithms nonlinear quasivariational inequalities, *Mathematical Sciences Research Hot-Line* 4(12) (2000), 33-45.

[29] Verma, R.U., Averaging techniques and cocoercively monotone mappings, *Mathematical Sciences Research Journal* 10(3) (2006), 79-82.

[30] Verma, R.U., General class of implicit variational inclusions and graph convergence on $A-$maximal relaxed monotonicity, *Journal of Optimization Theory and Applications* 155(1) (2012), 196-214.

[31] Xu, H.K., Iterative algorithms for nonlinear operators, *Journal of London Mathematical Society* 66 (2002), 240-256.

[32] Zeidler, E., *Nonlinear Functional Analysis and its Applications* I, Springer-Verlag, New York, New York, 1986.

[33] Zeidler, E., *Nonlinear Functional Analysis and its Applications* II/A, Springer-Verlag, New York, New York, 1990.

Chapter 12

Newton-Type Iterative Methods for Nonlinear Ill-Posed Hammerstein-Type Equations

12.1. Introduction

This chapter is devoted to the study of non-linear ill-posed Hammerstein type operator equations. Recall that ([13, 14, 15, 16]) an equation of the form

$$(KF)x = y \tag{12.1.1}$$

is called a non-linear ill-posed Hammerstein type operator equation. Here $F : D(F) \subseteq X \to Z$, is a nonlinear operator, $K : Z \to Y$ is a bounded linear operator and X, Z, Y are Hilbert spaces with corresponding inner product $\langle .,. \rangle_X$, $\langle .,. \rangle_Z$, $\langle .,. \rangle_Y$, and norm $\|.\|_X$, $\|.\|_Z$, $\|.\|_Y$ respectively. A typical example of a Hammerstein type operator is the nonlinear integral operator

$$(Ax)(t) := \int_0^1 k(s,t) f(s, x(s)) ds$$

where $k(s,t) \in L^2([0,1] \times [0,1]),\ x \in L^2[0,1]$ and $t \in [0,1]$.

The above integral operator A admits a representation of the form $A = KF$ where $K : L^2[0,1] \to L^2[0,1]$ is a linear integral operator with kernel $k(t,s)$: defined as

$$Kx(t) = \int_0^1 k(t,s) x(s) ds$$

and $F : D(F) \subseteq L^2[0,1] \to L^2[0,1]$ is a nonlinear superposition operator (cf. [24]) defined as

$$Fx(s) = f(s, x(s)). \tag{12.1.2}$$

The first author and his collaborators ([13, 14, 15, 16]), studied ill-posed Hammerstein type equation extensively under some assumptions on the Fréchet derivative of F. Precisely, in [13, 15], it is assumed that $F'(x_0)^{-1}$ exists and in [16] it is assumed that $F'(x)^{-1}$ exists for all x in a ball of radius r around x_0.

Note that if the function f in (12.1.2) is differentiable with respect to the second variable and for all $x \in B_r(x_0)$, $t \in [0,1]$; $\partial_2 f(t,x(t)) \geq \kappa_1$, then $F'(u)^{-1}$ exists and is a bounded operator for all $u \in B_r(x_0)$ (see Remark 2.1 in [15]), here $\partial_2 f(t,s)$ represents the partial derivative of f with respect to the second variable.

Throughout this chapter it is assumed that the available data is y^δ with

$$\|y - y^\delta\|_Y \leq \delta$$

and hence one has to consider the equation

$$(KF)x = y^\delta \tag{12.1.3}$$

instead of (12.1.1). Observe that the solution x of (12.1.3) can be obtained by solving

$$Kz = y^\delta \tag{12.1.4}$$

for z and then solving the non-linear problem

$$F(x) = z. \tag{12.1.5}$$

In [16], for solving (12.1.5), George and Kunhanandan considered the sequence defined iteratively by

$$x^\delta_{n+1,\alpha} = x^\delta_{n,\alpha} - F'(x^\delta_{n,\alpha})^{-1}(F(x^\delta_{n,\alpha}) - z^\delta_\alpha)$$

where $x^\delta_{0,\alpha} := x_0$,

$$z^\delta_\alpha = (K^*K + \alpha I)^{-1}K^*(y^\delta - KF(x_0)) + F(x_0) \tag{12.1.6}$$

and obtained local quadratic convergence.

Recall that a sequence (x_n) in X with $\lim x_n = x^*$ is said to be convergent of order $p > 1$, if there exist positive reals c_1, c_2, such that for all $n \in N$

$$\|x_n - x^*\|_X \leq c_1 e^{-c_2 p^n}.$$

If the sequence (x_n) has the property that $\|x_n - x^*\|_X \leq c_1 q^n$, $0 < q < 1$, then (x_n) is said to be linearly convergent. For an extensive discussion of convergence rate see Kelley [23].

And in [15], George and Nair studied the modified Laverentiev regularization

$$z^\delta_\alpha = (K + \alpha I)^{-1}(y^\delta - KF(x_0))$$

for obtaining an approximate solution of (12.1.4) and introduced modified Newton's iterations,

$$x^\delta_{n,\alpha} = x^\delta_{n-1,\alpha} - F'(x_0)^{-1}(F(x^\delta_{n-1,\alpha}) - F(x_0) - z^\delta_\alpha)$$

for solving (12.1.5) and obtained local linear convergence. In fact in [15] and [16], a solution $\hat{x}$ of (12.1.1) is called an x_0-minimum norm solution if it satisfies

$$\|F(\hat{x}) - F(x_0)\|_Z := \min\{\|F(x) - F(x_0)\|_Z : KF(x) = y, x \in D(F)\}. \tag{12.1.7}$$

We also assume throughout that the solution $\hat{x}$ satisfies (12.1.7). In all these papers ([13, 14, 15, 16]), it is assumed that the ill-posedness of (12.1.1) is due to the nonclosedness of the operator K. In this chapter we consider two cases:

Case (1) $F'(x_0)^{-1}$ exists and is a bounded operator, i.e., (12.1.5) is regular.

Case (2) F is monotone ([26], [31]), $Z = X$ is a real Hilbert space and $F'(x_0)^{-1}$ does not exist, i.e., (12.1.5) is also ill-posed.

The case when F is not monotone and $F'(x_0)^{-1}$ does not exist is the subject matter of the forthcoming chapter.

One of the advantages of (approximately) solving (12.1.4) and (12.1.5) to obtain an approximate solution for (12.1.3) is that, one can use any regularization method ([8, 22]) for linear ill-posed equations, for solving (12.1.4) and any iterative method ([10, 12]) for solving (12.1.5). In fact in this chapter we consider Tikhonov regularization([11, 13, 16, 19, 20]) for approximately solving (12.1.4) and we consider a modified two step Newton method ([1, 6, 7, 9, 21, 25]) for solving (12.1.5). Note that the regularization parameter α is chosen according to the adaptive method considered by Pereverzev and Schock in ([28]) for the linear ill-posed operator equations and the same parameter α is used for solving the non-linear operator equation (12.1.5), so the choice of the regularization parameter is not depending on the non-linear operator F, this is another advantage over treating (12.1.3) as a single non-linear operator equation.

This chapter is organized as follows. Preparatory results are given in Section 12.2 and Section 12.3 comprises the proposed iterative method for case (1) and case (2). Section 12.4 deals with the algorithm for implementing the proposed method. Numerical examples are given in Section 12.5. Finally the chapter ends with a conclusion in section 12.6.

12.2. Preparatory Results

In this section we consider Tikhonov regularized solution z_α^δ defined in (12.1.6) and obtain an a priori and an a posteriori error estimate for $\|F(\hat{x}) - z_\alpha^\delta\|_Z$. The following assumption is required to obtain the error estimate .

Assumption 12.2.1. *There exists a continuous, strictly monotonically increasing function* $\varphi : (0,a] \to (0,\infty)$ *with* $a \geq \|K^*K\|_{Y\to X}$ *satisfying;*

- $\lim_{\lambda\to 0}\varphi(\lambda) = 0$
- $$\sup_{\lambda\geq 0}\frac{\alpha\varphi(\lambda)}{\lambda+\alpha} \leq \varphi(\alpha), \qquad \forall\lambda \in (0,a]$$

  and
- *there exists* $v \in X, \|v\|_X \leq 1$ *such that*

$$F(\hat{x}) - F(x_0) = \varphi(K^*K)v.$$

Theorem 12.2.2. *(see (4.3) in [16]) Let* z_α^δ *be as in (12.1.6) and Assumption 12.2.1 holds. Then*

$$\|F(\hat{x}) - z_\alpha^\delta\|_Z \leq \varphi(\alpha) + \frac{\delta}{\sqrt{\alpha}}. \tag{12.2.1}$$

12.2.1. A Priori Choice of the Parameter

Note that the estimate $\varphi(\alpha)+\frac{\delta}{\sqrt{\alpha}}$ in (12.2.1) is of optimal order for the choice $\alpha := \alpha_\delta$ which satisfies $\varphi(\alpha_\delta)=\frac{\delta}{\sqrt{\alpha_\delta}}$. Let $\psi(\lambda) := \lambda\sqrt{\varphi^{-1}(\lambda)}, 0<\lambda\leq \|K\|_Y^2$. Then we have $\delta = \sqrt{\alpha_\delta}\varphi(\alpha_\delta)=\psi(\varphi(\alpha_\delta))$ and

$$\alpha_\delta = \varphi^{-1}(\psi^{-1}(\delta)).$$

So the relation (12.2.1) leads to $\|F(\hat{x})-z_\alpha^\delta\|_Z \leq 2\psi^{-1}(\delta)$.

12.2.2. An Adaptive Choice of the Parameter

In this chapter, we propose to choose the parameter α according to the adaptive choice established by Pereverzev and Shock [28] for solving ill-posed problems. We denote by D_M the set of possible values of the parameter α

$$D_M = \{\alpha_i = \alpha_0\mu^{2i}, i = 0,1,2,....,M\}, \mu > 1.$$

Then the selection of numerical value k for the parameter α according to the adaptive choice is performed using the rule

$$k := \max\{i : \alpha_i \in D_M^+\} \tag{12.2.2}$$

where $D_M^+ = \{\alpha_i \in D_M : \|z_{\alpha_i}^\delta - z_{\alpha_j}^\delta\|_Z \leq \frac{4\delta}{\sqrt{\alpha_j}}, j = 0,1,2,....,i-1\}$. Let

$$l := \max\{i : \varphi(\alpha_i) \leq \frac{\delta}{\sqrt{\alpha_i}}\}. \tag{12.2.3}$$

We will be using the following theorem from [16] for our error analysis.

Theorem 12.2.3. *(cf. [16], Theorem 4.3) Let l be as in (12.2.3), k be as in (12.2.2) and $z_{\alpha_k}^\delta$ be as in (12.1.6) with $\alpha = \alpha_k$. Then $l \leq k$ and*

$$\|F(\hat{x}) - z_{\alpha_k}^\delta\|_Z \leq (2+\frac{4\mu}{\mu-1})\mu\psi^{-1}(\delta).$$

12.3. Convergence Analysis

Throughout this chapter we assume that the operator F possess a uniformly bounded Fr*é*chet derivative $F'(.)$ for all $x \in D(F)$. In the earlier papers [16, 17, 18] the authors used the following Assumption:

Assumption 12.3.1. *(cf.[30], Assumption 3 (A3)) There exist a constant $K_0 \geq 0$ such that for every $x,u \in B_r(x_0)\cup B_r(\hat{x}) \subseteq D(F)$ and $v \in X$ there exists an element $\Phi(x,u,v) \in X$ such that $[F'(x)-F'(u)]v = F'(u)\Phi(x,u,v)$, $\|\Phi(x,u,v)\|_X \leq K_0\|v\|_X\|x-u\|_X$.*

The hypotheses of Assumption 12.3.1 may not hold or may be very expensive or impossible to verify in general. In particular, as it is the case for well-posed nonlinear equations the computation of the Lipschitz constant K_0 even if this constant exists is very difficult. Moreover, there are classes of operators for which Assumption 12.3.1 is not satisfied but the iterative method converges.

In the present chapter, we expand the applicability of the Newton-type iterative method under less computational cost. We achieve this goal by the following weaker Assumption.

Assumption 12.3.2. *Let $x_0 \in X$ be fixed. There exists a constant k_0 such that for every $u \in B_r(x_0) \subseteq D(F)$ and $v \in X$, there exists an element $\Phi_0(x_0,u,v) \in X$ satisfying*

$$[F'(x_0) - F'(u)]v = F'(x_0)\Phi_0(x_0,u,v), \|\Phi(x_0,u,v)\|_X \leq k_0\|v\|_X\|x_0 - u\|_X.$$

Note that

$$k_0 \leq K_0$$

holds in general and $\frac{K_0}{k_0}$ can be arbitrary large. The advantages of the new approach are:

(1) Assumption 12.3.2 is weaker than Assumption 12.3.1. Notice that there are classes of operators that satisfy Assumption 12.3.2 but do not satisfy Assumption 12.3.1;

(2) The computational cost of finding the constant k_0 is less than that of constant K_0, even when $K_0 = k_0$;

(3) The sufficient convergence criteria are weaker;

(4) The computable error bounds on the distances involved (including k_0) are less costly and more precise than the old ones (including K_0);

(5) The information on the location of the solution is more precise;

and

(6) The convergence domain of the iterative method is larger.

These advantages are also very important in computational mathematics since they provide under less computational cost a wider choice of initial guesses for iterative method and the computation of fewer iterates to achieve a desired error tolerance. Numerical examples for (1)-(6) are presented in Section 4.

12.3.1. Iterative Method for Case (1)

In this subsection for an initial guess $x_0 \in X$, we consider the sequence v^δ_{n,α_k} defined iteratively by

$$v^\delta_{n,\alpha_k} = v^\delta_{n,\alpha_k} - F'(x_0)^{-1}(F(v^\delta_{n,\alpha_k}) - z^\delta_{\alpha_k})$$

where $v^\delta_{0,\alpha_k} = x_0$ for obtaining an approximation $x^\delta_{\alpha_k}$ of x such that $F(x) = z^\delta_{\alpha_k}$.

Let

$$y^\delta_{n,\alpha_k} = v^\delta_{2n-1,\alpha_k} \tag{12.3.1}$$

and

$$x^\delta_{n+1,\alpha_k} = v^\delta_{2n,\alpha_k}, \tag{12.3.2}$$

for $n > 0$. We will be using the following notations;

$$M \geq \|F'(x_0)\|_{X\to Z};$$

$$\beta := \|F'(x_0)^{-1}\|_{Z\to X};$$

$$k_0 < \frac{1}{4}\min\{1, \frac{1}{\beta}\};$$

$$\delta_0 < \frac{\sqrt{\alpha_0}}{4k_0\beta};$$

$$\rho := \frac{1}{M}(\frac{1}{4k_0\beta} - \frac{\delta_0}{\sqrt{\alpha_0}});$$

$$\gamma_\rho := \beta[M\rho + \frac{\delta_0}{\sqrt{\alpha_0}}];$$

and

$$e^\delta_{n,\alpha_k} := \|y^\delta_{n,\alpha_k} - x^\delta_{n,\alpha_k}\|_X, \qquad \forall n \geq 0. \tag{12.3.3}$$

For convenience, we use the notation x_n, y_n and e_n for x^δ_{n,α_k}, y^δ_{n,α_k} and e^δ_{n,α_k} respectively.

Further we define

$$q := k_0 r, r \in (r_1, r_2) \tag{12.3.4}$$

where

$$r_1 = \frac{1 - \sqrt{1 - 4k_0\gamma_\rho}}{2k_0}$$

and

$$r_2 = \min\{\frac{1}{k_0}, \frac{1 + \sqrt{1 - 4k_0\gamma_\rho}}{2k_0}\}.$$

Note that r is well defined because $\gamma_\rho \leq \frac{1}{4k_0}$. We will be using the relation $e_0 \leq \gamma_\rho$ for proving our results, which can be seen as follows;

$$\begin{aligned}
e_0 = \|y_0 - x_0\|_X &= \|F'(x_0)^{-1}(F(x_0) - z^\delta_{\alpha_k})\|_X \\
&\leq \|F'(x_0)^{-1}\|_{Z\to X}\|(F(x_0) - z^\delta_{\alpha_k})\|_Z \\
&\leq \beta\|F(x_0) - z_{\alpha_k} + z_{\alpha_k} - z^\delta_{\alpha_k}\|_Z \\
&\leq \beta[\|F(x_0) - F(\hat{x})\|_Z + \|z_{\alpha_k} - z^\delta_{\alpha_k}\|_Z] \\
&\leq \beta[M\rho + \frac{\delta}{\sqrt{\alpha}}] \\
&\leq \beta[M\rho + \frac{\delta_0}{\sqrt{\alpha_0}}] \\
&= \gamma_\rho.
\end{aligned}$$

Theorem 12.3.3. *Let e_n, q be as in (12.3.3), (12.3.4) respectively and x_n, y_n be as in (12.3.2), (12.3.1) respectively with $\delta \in (0, \delta_0]$. Then by Assumption 12.3.2 and Theorem 12.2.3 $x_n, y_n \in B_r(x_0)$ and the following estimates hold for all $n \geq 0$.*

(a) $\|x_{n+1} - y_n\|_X \leq q\|y_n - x_n\|_X$;

(b) $\|y_{n+1} - x_{n+1}\|_X \leq q^2\|y_n - x_n\|_X$;

(c) $e_n \leq q^{2n}\gamma_\rho, \qquad \forall n \geq 0.$

Proof. Suppose $x_n, y_n \in B_r(x_0)$. Then

$$\begin{aligned} x_{n+1} - y_n &= y_n - x_n - F'(x_0)^{-1}(F(y_n) - F(x_n)) \\ &= F'(x_0)^{-1}[F'(x_0)(y_n - x_n) - (F(y_n) - F(x_n))] \\ &= F'(x_0)^{-1}\int_0^1 [F'(x_0) - F'(x_n + t(y_n - x_n))](y_n - x_n)dt \end{aligned}$$

and hence by Assumption 12.3.2, we have

$$\|x_{n+1} - y_n\|_X \leq k_0 r\|y_n - x_n\|_X \leq q\|y_n - x_n\|_X.$$

This proves (a). To prove (b) we observe that

$$\begin{aligned} e_{n+1} = \|y_{n+1} - x_{n+1}\|_X &= \|x_{n+1} - y_n - F'(x_0)^{-1}(F(x_{n+1}) - F(y_n))\|_X \\ &= \|F'(x_0)^{-1}\int_0^1 [F'(x_0) - F'(y_n + t(x_{n+1} - y_n)] \\ &\quad dt(x_{n+1} - y_n)\|_X \\ &\leq k_0 r\|y_n - x_{n+1}\|_X \\ &\leq q^2\|x_n - y_n\|_X. \end{aligned}$$

The last but one step follows from Assumption 12.3.2 and the last step follows from (a). This completes the proof of (b) and (c) follows from (b). Now we shall show that $x_n, y_n \in B_r(x_0)$ by induction. For $n = 1$,

$$\begin{aligned} x_1 - y_0 &= y_0 - x_0 - F'(x_0)^{-1}(F(y_0) - F(x_0)) \\ &= F'(x_0)^{-1}[F'(x_0)(y_0 - x_0) - (F(y_0) - F(x_0))] \\ &= F'(x_0)^{-1}\int_0^1 [F'(x_0) - F'(x_0 + t(y_0 - x_0))](y_0 - x_0)dt \end{aligned}$$

and hence by Assumption 12.3.2, we have

$$\begin{aligned} \|x_1 - y_0\|_X &\leq \frac{k_0}{2}\|y_0 - x_0\|_X^2 \\ &\leq k_0 r e_0. \end{aligned} \tag{12.3.5}$$

So by triangle inequality and (12.3.5)

$$\begin{aligned} \|x_1 - x_0\|_X &\leq \|x_1 - y_0\|_X + \|y_0 - x_0\|_X \\ &\leq (1+q)e_0 \\ &\leq \frac{e_0}{1-q} \\ &\leq \frac{\gamma_\rho}{1-q} \\ &\leq r. \end{aligned}$$

i.e., $x_1 \in B_r(x_0)$. Observe that

$$\begin{aligned} \|y_1 - x_1\|_X &= \|x_1 - y_0 - F'(x_0)^{-1}(F(x_1) - F(y_0))\|_X \\ &\leq k_0 r\|x_1 - y_0\|_X \end{aligned}$$

and hence by (12.3.5)

$$\|y_1 - x_1\|_X \leq q^2 e_0. \tag{12.3.6}$$

Therefore by (12.3.4), (12.3.6) and triangle inequality,

$$\begin{aligned}
\|y_1 - x_0\|_X &\leq \|y_1 - x_1\|_X + \|x_1 - x_0\|_X \\
&\leq (1+q+q^2)e_0 \\
&\leq \frac{e_0}{1-q} \\
&\leq \frac{\gamma_\rho}{1-q} \\
&\leq r
\end{aligned}$$

i.e., $y_1 \in B_r(x_0)$. Suppose $x_m, y_m \in B_r(x_0)$. Then

$$\begin{aligned}
\|x_{m+1} - x_0\|_X &\leq \|x_{m+1} - x_m\|_X + \|x_m - x_{m-1}\|_X + \cdots + \|x_1 - x_0\|_X \\
&\leq (q+1)e_m + (q+1)e_{m-1} + \cdots + (q+1)e_0 \\
&\leq (q+1)(e_m + e_{m-1} + \cdots + e_0) \\
&\leq (q+1)(q^{2m} + q^{2(m-1)} + \cdots + 1)e_0 \\
&\leq (q+1)\frac{1-(q^{2m+1})}{1-q^2}e_0 \\
&\leq \frac{e_0}{1-q} \\
&\leq \frac{\gamma_\rho}{1-q} \\
&\leq r
\end{aligned}$$

i.e., $x_{m+1} \in B_r(x_0)$ and

$$\begin{aligned}
\|y_{m+1} - x_0\|_X &\leq \|y_{m+1} - x_{m+1}\|_X + \|x_{m+1} - x_0\|_X \\
&\leq q^2 e_m + (q+1)e_m + (q+1)e_{m-1} + \cdots + (q+1)e_0 \\
&\leq (q^2+q+1)e_m + (q+1)e_{m-1} + \cdots + (q+1)e_0 \\
&\leq (q^{2(m+1)} + \cdots + q^3 + q^2 + q + 1)e_0 \\
&\leq \frac{e_0}{1-q} \\
&\leq \frac{\gamma_\rho}{1-q} \\
&\leq r
\end{aligned}$$

i.e., $y_{m+1} \in B_r(x_0)$. Thus by induction $x_n, y_n \in B_r(x_0)$. This completes the proof of the Theorem.

The main result of this section is the following Theorem.

Theorem 12.3.4. *Let x_n and y_n be as in (12.3.2) and (12.3.1) respectively and assumptions of Theorem 12.3.3 hold. Then (x_n) is a Cauchy sequence in $B_r(x_0)$ and converges to $x_{\alpha_k}^{\delta} \in \overline{B_r(x_0)}$. Further $F(x_{\alpha_k}^{\delta}) = z_{\alpha_k}^{\delta}$ and*

$$\|x_n - x_{\alpha_k}^{\delta}\|_X \leq Cq^{2n}$$

where $C = \frac{\gamma_\rho}{1-q}$

Proof. Using the relation (b) and (c) of Theorem 12.3.3, we obtain

$$\begin{aligned}
\|x_{n+m} - x_n\|_X &\leq \sum_{i=0}^{i=m-1} \|x_{n+i+1} - x_{n+i}\|_X \\
&\leq \sum_{i=0}^{i=m-1} (1+q)e_{n+i} \\
&\leq \sum_{i=0}^{i=m-1} (1+q)q^{2(n+i)}e_0 \\
&= (1+q)q^{2n}e_0 + (1+q)q^{2(n+1)}e_0 + \ldots.. + (1+q)q^{2(n+m)}e_0 \\
&\leq (1+q)q^{2n}(1+q^2+q^{2(2)}+\cdots+q^{2m})e_0 \\
&\leq q^{2n}[\frac{1-(q^2)^{m+1}}{1-q}]\gamma_\rho \\
&\leq Cq^{2n}.
\end{aligned}$$

Thus x_n is a Cauchy sequence in $B_r(x_0)$ and hence it converges, say to $x_{\alpha_k}^{\delta} \in \overline{B_r(x_0)}$. Observe that

$$\begin{aligned}
\|F(x_n) - z_{\alpha_k}^{\delta}\|_Z &= \|F'(x_0)(x_n - y_n)\|_Z \\
&\leq \|F'(x_0)\|_{X\to Z}\|(x_n - y_n)\|_Z \\
&\leq Me_n \leq Mq^{2n}\gamma_\rho. \qquad (12.3.7)
\end{aligned}$$

Now by letting $n \to \infty$ in (12.3.7) we obtain $F(x_{\alpha_k}^{\delta}) = z_{\alpha_k}^{\delta}$. This completes the proof.

Hereafter we assume that

$$\|\hat{x} - x_0\|_X < \rho \leq r.$$

Theorem 12.3.5. *Suppose that the hypothesis of Assumption 12.3.2 holds. Then*

$$\|\hat{x} - x_{\alpha_k}^{\delta}\|_X \leq \frac{\beta}{1-k_0 r}\|F(\hat{x}) - z_{\alpha_k}^{\delta}\|_Z.$$

Proof. Note that $k_0 r < 1$ and by Assumption 12.3.2, we have

$$\begin{aligned}
\|\hat{x} - x_{\alpha_k}^{\delta}\|_X &\leq \|\hat{x} - x_{\alpha_k}^{\delta} + F'(x_0)^{-1}[F(x_{\alpha_k}^{\delta}) - F(\hat{x}) + F(\hat{x}) - z_{\alpha_k}^{\delta}]\|_X \\
&\leq \|F'(x_0)^{-1}[F'(x_0)(\hat{x} - x_{\alpha_k}^{\delta}) + F(x_{\alpha_k}^{\delta}) - F(\hat{x})]\|_X \\
&\quad + \|F'(x_0)^{-1}(F(\hat{x}) - z_{\alpha_k}^{\delta})\|_X \\
&\leq k_0\|x_0 - \hat{x} - t(x_{\alpha_k}^{\delta} - \hat{x})\|_X\|\hat{x} - x_{\alpha_k}^{\delta}\|_X + \beta\|F(\hat{x}) - z_{\alpha_k}^{\delta}\|_Z \\
&\leq k_0 r\|\hat{x} - x_{\alpha_k}^{\delta}\|_Z + \beta\|F(\hat{x}) - z_{\alpha_k}^{\delta}\|_Z.
\end{aligned}$$

This completes the proof. The following Theorem is a consequence of Theorem 12.3.4 and Theorem 12.3.5.

Theorem 12.3.6. *Let x_n be as in (12.3.2), assumptions in Theorem 12.3.4 and Theorem 12.3.5 hold. Then*

$$\|\hat{x}-x_n\|_X \leq Cq^{2n} + \frac{\beta}{1-k_0 r}\|F(\hat{x})-z^{\delta}_{\alpha_k}\|_Z$$

where C is as in Theorem 12.3.4.

Observe that from section 2.2, $l \leq k$ and $\alpha_\delta \leq \alpha_{l+1} \leq \mu\alpha_l$, we have

$$\frac{\delta}{\sqrt{\alpha_k}} \leq \frac{\delta}{\sqrt{\alpha_l}} \leq \mu\frac{\delta}{\sqrt{\alpha_\delta}} = \mu\varphi(\alpha_\delta) = \mu\psi^{-1}(\delta).$$

This leads to the following theorem,

Theorem 12.3.7. *Let x_n be as in (12.3.2), assumptions in Theorem 12.2.3, Theorem 12.3.4 and Theorem 12.3.5 hold. Let*

$$n_k := \min\{n : q^{2n} \leq \frac{\delta}{\sqrt{\alpha_k}}\}.$$

Then

$$\|\hat{x}-x_{n_k}\|_X = O(\psi^{-1}(\delta)).$$

12.3.2. Iterative Method for Case (2)

F is a monotone operator (i.e., $\langle F(x)-F(y), x-y\rangle \geq 0, \quad \forall x,y \in D(F)$), $Z = X$ is a real Hilbert space and $F'(x_0)^{-1}$ does not exist. Thus the ill-posedness of (12.1.1) in this case is due to the ill-posedness of F as well as the nonclosedness of the range of the linear operator K. The following assumptions are needed in addition to the earlier assumptions for our convergence analysis.

Assumption 12.3.8. *There exists a continuous, strictly monotonically increasing function $\varphi_1 : (0,b] \to (0,\infty)$ with $b \geq \|F'(x_0)\|_{X\to X}$ satisfying;*

- $\lim_{\lambda\to 0}\varphi_1(\lambda) = 0$,
- $$\sup_{\lambda\geq 0}\frac{\alpha\varphi_1(\lambda)}{\lambda+\alpha} \leq \varphi_1(\alpha) \qquad \forall\lambda \in (0,b]$$

  and
- *there exists $v \in X$ with $\|v\|_X \leq 1$ (cf. [26]) such that*

  $$x_0 - \hat{x} = \varphi_1(F'(x_0))v.$$

Assumption 12.3.9. *For each $x \in B_{\tilde{r}}(x_0)$ there exists a bounded linear operator $G(x,x_0)$ (see [29]) such that*

$$F'(x) = F'(x_0)G(x,x_0)$$

with $\|G(x,x_0)\|_{X\to X} \leq k_2$.

The iterative method for this case

$$\tilde{v}^{\delta}_{n,\alpha_k} = \tilde{v}^{\delta}_{n,\alpha_k} - R(x_0)^{-1}[F(\tilde{v}^{\delta}_{n,\alpha_k}) - z^{\delta}_{\alpha_k} + \frac{\alpha_k}{c}(\tilde{v}^{\delta}_{n,\alpha_k} - x_0)]$$

where $\tilde{v}^{\delta}_{0,\alpha_k} := x_0$ is the initial guess and $R(x_0) := F'(x_0) + \frac{\alpha_k}{c}I$, with $c \leq \alpha_k$. Let

$$\tilde{y}^{\delta}_{n,\alpha_k} = \tilde{v}^{\delta}_{2n-1,\alpha_k} \tag{12.3.8}$$

and

$$\tilde{x}^{\delta}_{n+1,\alpha_k} = \tilde{v}^{\delta}_{2n,\alpha_k} \tag{12.3.9}$$

for $n > 0$.

First we prove that $\tilde{x}_{n,\alpha_k}$ converges to the zero x^{δ}_{c,α_k} of

$$F(x) + \frac{\alpha_k}{c}(x - x_0) = z^{\delta}_{\alpha_k} \tag{12.3.10}$$

and then we prove that x^{δ}_{c,α_k} is an approximation for $\hat{x}$.

Let

$$\tilde{e}^{\delta}_{n,\alpha_k} := \|\tilde{y}^{\delta}_{n,\alpha_k} - \tilde{x}^{\delta}_{n,\alpha_k}\|_X, \qquad \forall n \geq 0. \tag{12.3.11}$$

For the sake of simplicity, we use the notation $\tilde{x}_n$, $\tilde{y}_n$ and $\tilde{e}_n$ for $\tilde{x}^{\delta}_{n,\alpha_k}$, $\tilde{y}^{\delta}_{n,\alpha_k}$ and $\tilde{e}^{\delta}_{n,\alpha_k}$ respectively.

Hereafter we assume that $\|\hat{x} - x_0\|_X < \rho \leq \tilde{r}$ where

$$\rho < \frac{1}{M}(1 - \frac{\delta_0}{\sqrt{\alpha_0}})$$

with $\delta_0 < \sqrt{\alpha_0}$. Let

$$\tilde{\gamma}_\rho := M\rho + \frac{\delta_0}{\sqrt{\alpha_0}}.$$

and we define

$$q_1 = k_0\tilde{r}, \tilde{r} \in (\tilde{r}_1, \tilde{r}_2) \tag{12.3.12}$$

where

$$\tilde{r}_1 = \frac{1 - \sqrt{1 - 4k_0\tilde{\gamma}_\rho}}{2k_0}$$

and

$$\tilde{r}_2 = \min\{\frac{1}{k_0}, \frac{1 + \sqrt{1 - 4k_0\tilde{\gamma}_\rho}}{2k_0}\}.$$

Theorem 12.3.10. *Let $\tilde{e}_n$ and q_1 be as in equation (12.3.11) and (12.3.12) respectively, $\tilde{x}_n$ and $\tilde{y}_n$ be as in (12.3.9) and (12.3.8) respectively with $\delta \in (0, \delta_0]$ and suppose Assumption 12.3.2 holds. Then we have the following.*

(a) $\|\tilde{x}_n - \tilde{y}_{n-1}\|_X \leq q_1 \|\tilde{y}_{n-1} - \tilde{x}_{n-1}\|_X$;

(b) $\|\tilde{y}_n - \tilde{x}_n\|_X \leq q_1^2 \|\tilde{y}_{n-1} - \tilde{x}_{n-1}\|_X$,

(c) $\tilde{e}_n \leq q_1^{2n}\tilde{\gamma}_\rho, \qquad \forall n \geq 0.$

Proof. Suppose $\tilde{x}_n, \tilde{y}_n \in B_{\tilde{r}}(x_0)$, then

$$\begin{aligned}\tilde{x}_n - \tilde{y}_{n-1} &= \tilde{y}_{n-1} - \tilde{x}_{n-1} - R(x_0)^{-1}(F(\tilde{y}_{n-1}) - F(\tilde{x}_{n-1}) \\ &\quad + \frac{\alpha_k}{c}(\tilde{y}_{n-1} - \tilde{x}_{n-1})) \\ &= R(x_0)^{-1}[R(x_0)(\tilde{y}_{n-1} - \tilde{x}_{n-1}) \\ &\quad -(F(\tilde{y}_{n-1}) - F(\tilde{x}_{n-1})) - \frac{\alpha_k}{c}(\tilde{y}_{n-1} - \tilde{x}_{n-1})] \\ &= R(x_0)^{-1}\int_0^1 [F'(x_0) - (F(\tilde{y}_{n-1}) - F(\tilde{x}_{n-1}))] \\ &\quad \times (\tilde{y}_{n-1} - \tilde{x}_{n-1})dt.\end{aligned}$$

Now since $\|R(x_0)^{-1}F'(x_0)\|_{X\to X} \leq 1$, the proof of (a) follows as in Theorem 12.3.3. Again observe that

$$\begin{aligned}\tilde{e}_n &\leq \|\tilde{x}_n - \tilde{y}_{n-1} - R(x_0)^{-1}(F(\tilde{x}_n) - z_{\alpha_k}^\delta + \frac{\alpha_k}{c}(\tilde{x}_n - x_0))\|_X \\ &\quad + \|R(x_0)^{-1}(F(\tilde{y}_{n-1}) - z_{\alpha_k}^\delta + \frac{\alpha_k}{c}(\tilde{y}_{n-1} - x_0))\|_X \\ &\leq \|R(x_0)^{-1}[R(x_0)(\tilde{x}_n - \tilde{y}_{n-1}) - (F(\tilde{x}_n) - F(\tilde{y}_{n-1})) - \frac{\alpha_k}{c}(\tilde{x}_n - \tilde{y}_{n-1})]\|_X \\ &\leq \|R(x_0)^{-1}\int_0^1 [F'(x_0) - (F(\tilde{x}_n) - F(\tilde{y}_{n-1}))]dt(\tilde{x}_n - \tilde{y}_{n-1})\|_X.\end{aligned}$$

So the remaining part of the proof is analogous to the proof of Theorem 12.3.3.

Theorem 12.3.11. *Let $\tilde{y}_n$ and $\tilde{x}_n$ be as in (12.3.8) and (12.3.9) respectively and assumptions of Theorem 12.3.10 holds. Then $(\tilde{x}_n)$ is a Cauchy sequence in $B_{\tilde{r}}(x_0)$ and converges to $x_{c,\alpha_k}^\delta \in \overline{B_{\tilde{r}}(x_0)}$. Further $F(x_{c,\alpha_k}^\delta) + \frac{\alpha_k}{c}(x_{c,\alpha_k}^\delta - x_0) = z_{\alpha_k}^\delta$ and*

$$\|\tilde{x}_n - x_{c,\alpha_k}^\delta\|_X \leq \tilde{C}q_1^{2n}$$

where $\tilde{C} = \frac{\tilde{\gamma}_\rho}{1-q_1}$.

Proof. Analogous to the proof of Theorem 12.3.4, one can prove that $\tilde{x}_n$ is a Cauchy sequence in $B_{\tilde{r}}(x_0)$ and hence it converges, say to $x_{c,\alpha_k}^\delta \in \overline{B_{\tilde{r}}(x_0)}$ and

$$\begin{aligned}\|F(\tilde{x}_n) - z_{\alpha_k}^\delta + \frac{\alpha_k}{c}(\tilde{x}_n - x_0)\|_X &= \|R(x_0)(\tilde{x}_n - \tilde{y}_n)\|_X \\ &\leq \|R(x_0)\|_{X\to X}\|(\tilde{x}_n - \tilde{y}_n)\|_X \\ &\leq (\|F'(x_0)\|_{X\to X} + \frac{\alpha_k}{c})\tilde{e}_n \\ &\leq (\|F'(x_0)\|_{X\to X} + \frac{\alpha_k}{c})q_1^{2n}\tilde{e}_0 \qquad (12.3.13) \\ &\leq (\|F'(x_0)\|_{X\to X} + \frac{\alpha_k}{c})q_1^{2n}\tilde{\gamma}_\rho.\end{aligned}$$

Now by letting $n \to \infty$ in (12.3.13) we obtain $F(x^{\delta}_{c,\alpha_k}) + \frac{\alpha_k}{c}(x^{\delta}_{c,\alpha_k} - x_0) = z^{\delta}_{\alpha_k}$. This completes the proof.

Assume that $k_2 < \frac{1-k_0\tilde{r}}{1-c}$ and for the sake of simplicity assume that $\varphi_1(\alpha) \leq \varphi(\alpha)$ for $\alpha > 0$.

Theorem 12.3.12. *Suppose x^{δ}_{c,α_k} is the solution of (12.3.10) and Assumptions 12.3.2, 12.3.8 and 12.3.9 hold. Then*

$$\|\hat{x} - x^{\delta}_{c,\alpha_k}\|_X = O(\psi^{-1}(\delta)).$$

Proof. Note that $c(F(x^{\delta}_{c,\alpha_k}) - z^{\delta}_{\alpha_k}) + \alpha_k(x^{\delta}_{c,\alpha_k} - x_0) = 0$, so

$$\begin{aligned}
(F'(x_0)+\alpha_k I)(x^{\delta}_{c,\alpha_k} - \hat{x}) &= (F'(x_0)+\alpha_k I)(x^{\delta}_{c,\alpha_k} - \hat{x}) \\
&\quad -c(F(x^{\delta}_{c,\alpha_k}) - z^{\delta}_{\alpha_k}) - \alpha_k(x^{\delta}_{c\alpha} - x_0) \\
&= \alpha_k(x_0 - \hat{x}) + F'(x_0)(x^{\delta}_{c,\alpha_k} - \hat{x}) \\
&\quad -c[F(x^{\delta}_{c,\alpha_k}) - z^{\delta}_{\alpha_k}] \\
&= \alpha_k(x_0 - \hat{x}) + F'(x_0)(x^{\delta}_{c,\alpha_k} - \hat{x}) \\
&\quad -c[F(x^{\delta}_{c,\alpha_k}) - F(\hat{x}) + F(\hat{x}) - z^{\delta}_{\alpha_k}] \\
&= \alpha_k(x_0 - \hat{x}) - c(F(\hat{x}) - z^{\delta}_{\alpha_k}) + F'(x_0)(x^{\delta}_{c,\alpha_k} - \hat{x}) \\
&\quad -c[F(x^{\delta}_{c,\alpha_k}) - F(\hat{x})].
\end{aligned}$$

Thus

$$\begin{aligned}
\|x^{\delta}_{c,\alpha_k} - \hat{x}\|_X &\leq \|\alpha_k(F'(x_0+\alpha_k I)^{-1}(x_0 - \hat{x})\|_X + \|(F'(x_0)+\alpha_k I)^{-1} \\
&\quad c(F(\hat{x}) - z^{\delta}_{\alpha_k})\|_X + \|(F'(x_0)+\alpha_k I)^{-1}[F'(x_0)(x^{\delta}_{c,\alpha_k} - \hat{x}) \\
&\quad -c(F(x^{\delta}_{c,\alpha_k}) - F(\hat{x}))]\|_X \\
&\leq \|\alpha_k(F'(x_0)+\alpha_k I)^{-1}(x_0 - \hat{x})\|_X + \|F(\hat{x}) - z^{\delta}_{\alpha_k}\|_X \\
&\quad +\|(F'(x_0)+\alpha_k I)^{-1}\int_0^1 [F'(x_0) - cF'(\hat{x}+t(x^{\delta}_{c,\alpha_k} - \hat{x})] \\
&\quad (x^{\delta}_{c,\alpha_k} - \hat{x})dt\|_X \\
&\leq \|\alpha_k(F'(x_0)+\alpha_k I)^{-1}(x_0 - \hat{x})\|_X \\
&\quad +\|F(\hat{x}) - z^{\delta}_{\alpha_k}\|_X + \Gamma
\end{aligned} \tag{12.3.14}$$

where $\Gamma := \|(F'(x_0)+\alpha_k I)^{-1}\int_0^1[F'(x_0) - cF'(\hat{x}+t(x^{\delta}_{c,\alpha_k} - \hat{x})](x^{\delta}_{c,\alpha_k} - \hat{x})dt\|_X$. So by Assumption 12.3.9, we obtain

$$\begin{aligned}
\Gamma &\leq \|(F'(x_0)+\alpha_k I)^{-1}\int_0^1 [F'(x_0) - F'(\hat{x}+t(x^{\delta}_{c,\alpha_k} - \hat{x})] \\
&\quad (x^{\delta}_{c,\alpha_k} - \hat{x})dt\|_X + (1-c)\|(F'(x_0)+\alpha I)^{-1}F'(x_0) \\
&\quad \int_0^1 G(\hat{x}+t(x^{\delta}_{c,\alpha_k} - \hat{x}), x_0)(x^{\delta}_{c,\alpha_k} - \hat{x})dt\|_X \\
&\leq k_0\tilde{r}\|x^{\delta}_{c,\alpha_k} - \hat{x}\|_X + (1-c)k_2\|x^{\delta}_{c,\alpha_k} - \hat{x}\|_X
\end{aligned} \tag{12.3.15}$$

and hence by (12.3.14) and (12.3.15) we have

$$\begin{aligned}\|x^{\delta}_{c,\alpha_k}-\hat{x}\|_X &\leq \frac{\|\alpha_k(F'(x_0)+\alpha_k I)^{-1}(x_0-\hat{x})\|_X+\|F(\hat{x})-z^{\delta}_{\alpha_k}\|_X}{1-(1-c)k_2-k_0\tilde{r}}\\ &\leq \frac{\varphi_1(\alpha_k)+(2+\frac{4\mu}{\mu-1})\mu\psi^{-1}(\delta)}{1-(1-c)k_2-k_0\tilde{r}}\\ &= O(\psi^{-1}(\delta)).\end{aligned}$$

This completes the proof of the Theorem.

The following Theorem is a consequence of Theorem 12.3.11 and Theorem 12.3.12.

Theorem 12.3.13. *Let $\tilde{x}_n$ be as in (12.3.9), assumptions in Theorem 12.3.11 and Theorem 12.3.12 hold. Then*

$$\|\hat{x}-\tilde{x}_n\|_X \leq \tilde{C}q_1^{2n}+O(\psi^{-1}(\delta))$$

where $\tilde{C}$ is as in Theorem 12.3.11.

Theorem 12.3.14. *Let $\tilde{x}_n$ be as in (12.3.9), assumptions in Theorem 12.2.3, Theorem 12.3.11 and Theorem 12.3.12 hold. Let*

$$n_k := \min\{n : q_1^{2n} \leq \frac{\delta}{\sqrt{\alpha_k}}\}.$$

Then

$$\|\hat{x}-\tilde{x}_{n_k}\|_X = O(\psi^{-1}(\delta)).$$

Remark 12.3.15. *Let us denote by $\bar{r}_1$, $\bar{\gamma}_\rho$, $\bar{q}$, $\bar{\delta}_0$ the parameters using K_0 instead of k_0 for Case 1 (Similarly for Case 2). Then we have,*

$$\begin{aligned} r_1 &\leq \bar{r}_1,\\ \bar{\delta}_0 &\leq \delta_0,\\ \bar{\gamma}_\rho &\leq \gamma_\rho,\\ q &\leq \bar{q}.\end{aligned}$$

Moreover, strict inequality holds in the preceding estimates if $k_0 < K_0$. Let $h_0 = 4k_0\gamma_\rho$ and $h = 4K_0\bar{\gamma}_\rho$. We can certainly choose γ_ρ sufficiently close to $\bar{\gamma}_\rho$. Then, we have that, $h \leq 1 \Rightarrow h_0 \leq 1$ but not necessarily vice versa unless if $k_0 = K_0$ and $\gamma_\rho = \bar{\gamma}_\rho$. Finally, we have that, $\frac{h_0}{h} \to 0$ as $\frac{k_0}{K_0} \to 0$. The last estimate shows by how many times our new approach using k_0 can expand the applicability of the old approach using K_0 for these methods. Hence, all the above justify the claims made at the introduction of the chapter. Finally we note that the results obtained here are useful even if Assumptions 12.3.1 is satisfied but sufficient convergence condition $h \leq 1$ is not satisfied but $h_0 \leq 1$ is satisfied. Indeed, we can start with iterative method described in Case 1 (or Case 2) until a finite step N such that $h \leq 1$ is satisfied with $x^{\delta}_{N+1,\alpha_N}$ as a starting point for faster methods such as (12.1.6). Such an approach has already been employed in [2], [4] and [5] where the modified Newton's method is used as a predictor for Newton's method.

12.4. Algorithm

Note that for $i, j \in \{0, 1, 2, \cdots, M\}$

$$z^{\delta}_{\alpha_i} - z^{\delta}_{\alpha_j} = (\alpha_j - \alpha_i)(K^*K + \alpha_j I)^{-1}(K^*K + \alpha_i I)^{-1}[K^*(y^{\delta} - KF(x_0))].$$

The algorithm for implementing the iterative methods considered in section 3 involves the following steps.

- $\alpha_0 = \delta^2$;
- $\alpha_i = \mu^{2i}\alpha_0, \mu > 1$;
- solve for w_i : $\quad (K^*K + \alpha_i I)w_i = K^*(y^{\delta} - KF(x_0))$;
- solve for $j < i$, z_{ij} : $\quad (K^*K + \alpha_j I)z_{ij} = (\alpha_j - \alpha_i)w_i$;
- if $\|z_{ij}\|_X > \frac{4}{\mu^j}$, then take $k = i - 1$;
- otherwise, repeat with $i + 1$ in place of i.
- choose $n_k = \min\{n : q^{2n} \leq \frac{\delta}{\sqrt{\alpha_k}}\}$ in case (1) and $n_k = \min\{n : q_1^{2n} \leq \frac{\delta}{\sqrt{\alpha_k}}\}$ in case (2)
- solve x_{n_k} using the iteration (12.3.2) or $\tilde{x}_{n_k}$ using the iteration (12.3.9).

12.5. Numerical Examples

We present 5 numerical examples in this section. First, we consider two examples for illustrating the algorithm considered in the above sections. We apply the algorithm by choosing a sequence of finite dimensional subspace (V_N) of X with $\dim V_N = N + 1$. Precisely we choose V_N as the space of linear splines in a uniform grid of $N + 1$ points in $[0, 1]$. Then we present two examples where Assumption 12.3.1 is not satisfied but Assumption 12.3.2 is satisfied. In the last example we show that $\frac{k_0}{K_0}$ can be arbitrarily small.

Example 12.5.1. *In this example for Case (1), we consider the operator* $KF : D(KF) \subseteq L^2(0,1) \longrightarrow L^2(0,1)$ *with* $K : L^2(0,1) \longrightarrow L^2(0,1)$ *defined by*

$$K(x)(t) = \int_0^1 k(t,s)x(s)ds$$

where $k(t,s) = \begin{cases} (1-t)s, 0 \leq s \leq t \leq 1 \\ (1-s)t, 0 \leq t \leq s \leq 1 \end{cases}$ *and*

$$F : D(F) \subseteq L^2(0,1) \longrightarrow L^2(0,1)$$

defined by $F(u) := u^3$,

Then the Fréchet derivative of F *is given by* $F'(u)w = 3(u)^2 w$.

In our computation, we take $y(t) = \frac{837t}{6160} - \frac{t^2}{16} - \frac{t^{11}}{110} - \frac{3t^5}{80} - \frac{3t^8}{112}$ *and* $y^{\delta} = y + \delta$. *Then the exact solution*

$$\hat{x}(t) = 0.5 + t^3.$$

We use

$$x_0(t) = 0.5 + t^3 - \frac{3}{56}(t - t^8)$$

as our initial guess.

We choose $\alpha_0 = (1.3)^2(\delta)^2$, $\mu = 1.2$, $\delta = 0.0667$ *the Lipschitz constant* k_0 *equals approximately 0.23 and* $r = 1$, *so that* $q = k_0 r = 0.23$. *The iterations and corresponding error estimates are given in Table 12.5.1. The last column of the Table 12.5.1 shows that the error* $\|x_{n_k} - \hat{x}\|_X$ *is of order* $O(\delta^{\frac{1}{2}})$.

Table 12.5.1. Different errors

N	**k**	α_k	$\|x_{n_k} - \hat{x}\|_X$	$\frac{\|x_{n_k} - \hat{x}\|_X}{(\delta)^{1/2}}$
16	4	0.0231	0.5376	2.0791
32	4	0.0230	0.5301	2.0523
64	4	0.0229	0.5257	2.0359
128	4	0.0229	0.5234	2.0270
256	4	0.0229	0.5222	2.0224
512	4	0.0229	0.5216	2.0200
1024	4	0.0229	0.5213	2.0188

Example 12.5.2. *In this example for Case (2), we consider the operator* $KF : D(KF) \subseteq L^2(0,1) \longrightarrow L^2(0,1)$ *where* $K : L^2(0,1) \longrightarrow L^2(0,1)$ *defined by*

$$K(x)(t) = \int_0^1 k(t,s)x(s)ds$$

and $F : D(F) \subseteq L^2(0,1) \longrightarrow L^2(0,1)$ *defined by*

$$F(u) := \int_0^1 k(t,s)u^3(s)ds,$$

where

$$k(t,s) = \begin{cases} (1-t)s, 0 \le s \le t \le 1 \\ (1-s)t, 0 \le t \le s \le 1 \end{cases}.$$

Then for all $x(t), y(t) : x(t) > y(t)$: *(see section 4.3 in [30])*

$$\langle F(x) - F(y), x - y \rangle = \int_0^1 \left[\int_0^1 k(t,s)(x^3 - y^3)(s)ds \right] (x-y)(t)dt \ge 0.$$

Thus the operator F *is monotone. The Fréchet derivative of* F *is given by*

$$F'(u)w = 3\int_0^1 k(t,s)(u(s))^2 w(s)ds.$$

So for any $u \in B_r(x_0), x_0(s) \ge k_3 > 0, \forall s \in (0,1)$, *we have*

$$F'(u)w = F'(x_0)G(u,x_0)w,$$

where $G(u,x_0) = (\frac{u}{x_0})^2$.

In our computation, we take $y(t) = \frac{1}{110}(\frac{t^{13}}{156} - \frac{t^3}{6} + \frac{25t}{156})$ *and* $y^\delta = y + \delta$. *Then the exact solution*

$$\hat{x}(t) = t^3.$$

We use

$$x_0(t) = t^3 + \frac{3}{56}(t - t^8)$$

as our initial guess, so that the function $x_0 - \hat{x}$ *satisfies the source condition*

$$x_0 - \hat{x} = \frac{3}{56}(t - t^8) = F'(x_0)(\frac{t^6}{x_0(t)^2}) = \varphi_1(F'(x_0))(\frac{t^6}{x_0(t)^2})$$

where $\varphi_1(\lambda) = \lambda$. *Thus we expect to have an accuracy of order at least* $O(\delta^{\frac{1}{2}})$.

We choose $\alpha_0 = (1.3)\delta$, $\delta = 0.0667 =:$ *c the Lipschitz constant* k_0 *equals approximately 0.21 as in [30] and* $\tilde{r} = 1$, *so that* $q_1 = k_0\tilde{r} = 0.21$. *The results of the computation are presented in Table 12.5.2.*

Table 12.5.2. Results of computation

N	**k**	α_k	$\|\tilde{x}_{n_k} - \hat{x}\|_X$	$\frac{\|\tilde{x}_{n_k} - \hat{x}\|_X}{\delta^{1/2}}$
8	4	0.0494	0.1881	0.7200
16	4	0.0477	0.1432	0.5531
32	4	0.0473	0.1036	0.4010
64	4	0.0472	0.0726	0.2812
128	4	0.0471	0.0491	0.1900
256	4	0.0471	0.0306	0.1187
512	4	0.0471	0.0140	0.0543
1024	4	0.0471	0.0133	0.0515

In the next two cases, we present examples for nonlinear equations where Assumption 12.3.2 is satisfied but not Assumption 12.3.1.

Example 12.5.3. *Let* $X = Y = \mathbb{R}$, $D = [0,\infty), x_0 = 1$ *and define function* F *on* D *by*

$$F(x) = \frac{x^{1+\frac{1}{i}}}{1+\frac{1}{i}} + c_1 x + c_2, \tag{12.5.1}$$

where c_1, c_2 *are real parameters and* $i > 2$ *an integer. Then* $F'(x) = x^{1/i} + c_1$ *is not Lipschitz on* D. *Hence, Assumption 12.3.1 is not satisfied. However central Lipschitz condition Assumption 12.3.2 holds for* $k_0 = 1$.

Indeed, we have

$$\begin{aligned} \|F'(x) - F'(x_0)\|_X &= |x^{1/i} - x_0^{1/i}| \\ &= \frac{|x - x_0|}{x_0^{\frac{i-1}{i}} + \cdots + x^{\frac{i-1}{i}}} \end{aligned}$$

so

$$\|F'(x)-F'(x_0)\|_X \leq k_0|x-x_0|.$$

Example 12.5.4. *We consider the integral equations*

$$u(s)=f(s)+\lambda\int_a^b G(s,t)u(t)^{1+1/n}dt,\ \ n\in\mathbb{N}. \tag{12.5.2}$$

Here, f is a given continuous function satisfying $f(s)>0, s\in[a,b], \lambda$ is a real number, and the kernel G is continuous and positive in $[a,b]\times[a,b]$.

For example, when $G(s,t)$ is the Green kernel, the corresponding integral equation is equivalent to the boundary value problem

$$\begin{aligned} u'' &= \lambda u^{1+1/n} \\ u(a) &= f(a), u(b)=f(b). \end{aligned}$$

These type of problems have been considered in [1]- [5].

Equation of the form (12.5.2) generalize equations of the form

$$u(s)=\int_a^b G(s,t)u(t)^n dt \tag{12.5.3}$$

studied in [1]-[5]. Instead of (12.5.2) we can try to solve the equation $F(u)=0$ where

$$F:\Omega\subseteq C[a,b]\to C[a,b], \Omega=\{u\in C[a,b]: u(s)\geq 0, s\in[a,b]\},$$

and

$$F(u)(s)=u(s)-f(s)-\lambda\int_a^b G(s,t)u(t)^{1+1/n}dt.$$

The norm we consider is the max-norm.

The derivative F' is given by

$$F'(u)v(s)=v(s)-\lambda(1+\frac{1}{n})\int_a^b G(s,t)u(t)^{1/n}v(t)dt,\ \ v\in\Omega.$$

First of all, we notice that F' does not satisfy a Lipschitz-type condition in Ω. Let us consider, for instance, $[a,b]=[0,1], G(s,t)=1$ and $y(t)=0$. Then $F'(y)v(s)=v(s)$ and

$$\|F'(x)-F'(y)\|_{C[a,b]\to C[a,b]}=|\lambda|(1+\frac{1}{n})\int_a^b x(t)^{1/n}dt.$$

If F' were a Lipschitz function, then

$$\|F'(x)-F'(y)\|_{C[a,b]\to C[a,b]}\leq L_1\|x-y\|_{C[a,b]},$$

or, equivalently, the inequality

$$\int_0^1 x(t)^{1/n}dt\leq L_2\max_{x\in[0,1]}x(s), \tag{12.5.4}$$

would hold for all $x \in \Omega$ and for a constant L_2. But this is not true. Consider, for example, the functions

$$x_j(t) = \frac{t}{j}, \ \ j \geq 1, \ \ t \in [0,1].$$

If these are substituted into (12.5.4)

$$\frac{1}{j^{1/n}(1+1/n)} \leq \frac{L_2}{j} \Leftrightarrow j^{1-1/n} \leq L_2(1+1/n), \ \ \forall j \geq 1.$$

This inequality is not true when $j \to \infty$.

Therefore, condition (12.5.4) is not satisfied in this case. Hence Assumption 12.3.1 is not satisfied. However, condition Assumption 12.3.2 holds. To show this, let $x_0(t) = f(t)$ and $\gamma = \min_{s\in[a,b]} f(s), \alpha > 0$ Then for $v \in \Omega$,

$$\begin{aligned}
\|[F'(x) - F'(x_0)]v\|_{C[a,b]} &= |\lambda|(1+\frac{1}{n}) \max_{s\in[a,b]} |\int_a^b G(s,t)(x(t)^{1/n} - f(t)^{1/n})v(t)dt| \\
&\leq |\lambda|(1+\frac{1}{n}) \max_{s\in[a,b]} G_n(s,t)
\end{aligned}$$

where $G_n(s,t) = \frac{G(s,t)|x(t)-f(t)|}{x(t)^{(n-1)/n}+x(t)^{(n-2)/n}f(t)^{1/n}+\cdots+f(t)^{(n-1)/n}} \|v\|_{C[a,b]}$.

Hence,

$$\begin{aligned}
\|[F'(x) - F'(x_0)]v\|_{C[a,b]} &= \frac{|\lambda|(1+1/n)}{\gamma^{(n-1)/n}} \max_{s\in[a,b]} \int_a^b G(s,t)dt \|x - x_0\|_{C[a,b]} \\
&\leq k_0 \|x - x_0\|_{C[a,b]},
\end{aligned}$$

where $k_0 = \frac{|\lambda|(1+1/n)}{\gamma^{(n-1)/n}} N$ and $N = \max_{s\in[a,b]} \int_a^b G(s,t)dt$. Then Assumption 12.3.2 holds for sufficiently small λ.

Example 12.5.5. *Define the scalar function F by $F(x) = d_0 x + d_1 + d_2 \sin e^{d_3 x}, x_0 = 0$, where $d_i, i = 0,1,2,3$ are given parameters. Then, it can easily be seen that for d_3 large and d_2 sufficiently small, $\frac{k_0}{K_0}$ can be arbitrarily small.*

12.6. Conclusion

We presented an iterative method which is a combination of modified Newton method and Tikhonov regularization for obtaining an approximate solution for a nonlinear ill-posed Hammerstein type operator equation $KF(x) = y$, with the available noisy data y^δ in place of the exact data y. In fact we considered two cases, in the first case it is assumed that $F'(x_0)^{-1}$ exists and in the second case it is assumed that F is monotone but $F'(x_0)^{-1}$ does not exist. In both the cases, the derived error estimate using an a priori and balancing principle are of optimal order with respect to the general source condition. The results of the computational experiments gives the evidence of the reliability of our approach.

References

[1] Argyros, I. K., *Convergence and Application of Newton-type Iterations*, (Springer, 2008).

[2] Argyros, I. K., Approximating solutions of equations using Newton's method with a modified Newton's method iterate as a starting point. *Rev. Anal. Numer. Theor. Approx.* 36 (2007), 123-138.

[3] Argyros, I. K., A Semilocal convergence for directional Newton methods, *Math. Comput.(AMS)*. 80 (2011), 327-343.

[4] Argyros, I. K., Hilout, S. Weaker conditions for the convergence of Newton's method, *J. Complexity*, 28 (2012), 364-387.

[5] Argyros, I. K., Cho, Y. J., Hilout, S., *Numerical methods for equations and its applications*, (CRC Press, Taylor and Francis, New York, 2012).

[6] Argyros, I. K., Hilout, S., A convergence analysis for directional two-step Newton methods, *Numer. Algor.*, 55 (2010), 503-528.

[7] Bakushinskii, A. B., The problem of convergence of the iteratively regularized Gauss-Newton method, *Comput. Math. Math. Phys.*, 32 (1982), 1353-1359.

[8] Bakushinskii, A. B., Kokurin, M. Y., *Iterative Methods for Approximate Solution of Inverse Problems*, (Springer, Dordrecht, 2004).

[9] Blaschke, B., Neubauer, A., Scherzer, O., On convergence rates for the iteratively regularized Gauss-Newton method *IMA J. Numer. Anal.*, 17 (1997), 421-436.

[10] Engl, H. W., Regularization methods for the stable solution of inverse problems, *Surveys on Mathematics for Industry*, 3 (1993), 71-143.

[11] Engl, H. W., Kunisch, K., Neubauer, A., Convergence rates for Tikhonov regularization of nonlinear ill-posed problems, *Inverse Problems*, 5 (1989), 523-540.

[12] Engl, H. W., Kunisch, K., Neubauer, A., *Regularization of Inverse Problems*, (Kluwer, Dordrecht, 1996).

[13] George, S., Newton-Tikhonov regularization of ill-posed Hammerstein operator equation, *J. Inverse and Ill-Posed Problems*, 14(2) (2006), 135-146.

[14] George, S., Newton-Lavrentiev regularization of ill-posed Hammerstein operator equation, *J. Inverse and Ill-Posed Problems*, 14(6) (2006), 573-582.

[15] George, S., Nair, M.T., A modified Newton-Lavrentiev regularization for nonlinear ill-posed Hammerstein operator equations, *J. Complexity and Ill-Posed Problems*, 24 (2008), 228-240.

[16] George, S., Kunhanandan, M., An iterative regularization method for Ill-posed Hammerstein type operator equation, *J. Inv. Ill-Posed Problems* 17 (2009), 831-844.

[17] George, S., Shobha, M. E., A regularized dynamical system method for nonlinear ill-posed Hammerstein type operator equations, *J. Appl. Math. Bio*, 1(1) (2011), 65-78.

[18] Shobha, M. E., George, S., Dynamical System Method for Ill-posed Hammerstein Type Operator Equations with Monotone Operators, *International Journal of Pure and Applied Mathematics*, ISSN 1311-8080, 81(1) (2012), 129-143.

[19] George, S., Newton type iteration for Tikhonov regularization of nonlinear ill-posed problems, *Journal of Mathematics*, 2013 (2013), Article ID 439316, 9 pages, doi:10.1155/2013/439316.

[20] Groetsch, C. W., *Theory of Tikhonov regularization for Fredholm Equation of the first kind* (Pitmann Books, 1984).

[21] Kaltenbacher, B., A posteriori parameter choice strategies for some Newton-type methods for the regularization of nonlinear ill-posed problems, *Numer. Math.*, 79 (1998), 501-528.

[22] Kaltenbacher, B., Neubauer, A., Scherzer, O., *Iterative regularisation methods for nolinear ill-posed porblems* (de Gruyter, Berlin, New York 2008).

[23] Kelley, C. T., *Iterative Methods for Linear and Nonlinear Equations* (SIAM, Philadelphia 1995).

[24] Krasnoselskii, M. A., Zabreiko, P. P., Pustylnik, E. I., Sobolevskii, P. E., *Integral operators in spaces of summable functions* (Translated by T. Ando, Noordhoff International publishing, Leyden, 1976).

[25] Langer, S., Hohage, T., Convergence analysis of an inexact iteratively regularized Gauss-Newton method under general source conditions, *J. Inverse Ill-Posed Probl.*, 15 (2007), 19-35.

[26] Mahale, P., Nair, M. T., A simplified generalized Gauss-Newton method for nonlinear ill-posed problems, *Math. Comp.*, 78(265) (2009), 171-184.

[27] Nair, M.T., Ravishankar, P., Regularized versions of continuous newton's method and continuous modified newton's method under general source conditions, *Numer. Funct. Anal. Optim.* 29(9-10) (2008), 1140-1165.

[28] Pereverzev, S., Schock, E., On the adaptive selection of the parameter in regularization of ill-posed problems, *SIAM. J. Numer. Anal.*, 43(5) (2005), 2060-2076.

[29] Ramm, A. G., Smirnova, A. B., Favini, A., Continuous modified Newton's-type method for nonlinear operator equations. *Ann. Mat. Pura Appl.* 182 (2003), 37-52.

[30] Semenova, E.V., Lavrentiev regularization and balancing principle for solving ill-posed problems with monotone operators, *Comput. Methods Appl. Math.*, 4 (2010), 444-454.

[31] Tautenhahn, U., On the method of Lavrentiev regularization for nonlinear ill-posed problems, *Inverse Problems*, 18 (2002), 191-207.

Chapter 13

Enlarging the Convergence Domain of Secant-Like Methods for Equations

13.1. Introduction

Let $\mathcal{X}$, $\mathcal{Y}$ be Banach spaces and $\mathcal{D}$ be a non-empty, convex and open subset in $\mathcal{X}$. Let $U(x,r)$ and $\overline{U}(x,r)$ stand, respectively, for the open and closed ball in $\mathcal{X}$ with center x and radius $r > 0$. Denote by $\mathcal{L}(\mathcal{X},\mathcal{Y})$ the space of bounded linear operators from $\mathcal{X}$ into $\mathcal{Y}$. In the present chapter we are concerned with the problem of approximating a locally unique solution $x^\star$ of equation

$$F(x) = 0, \tag{13.1.1}$$

where F is a Fréchet continuously differentiable operator defined on $\mathcal{D}$ with values in $\mathcal{Y}$.

A lot of problems from computational sciences and other disciplines can be brought in the form of equation (13.1.1) using Mathematical Modelling [8, 10, 14]. The solution of these equations can rarely be found in closed form. That is why most solution methods for these equations are iterative. In particular, the practice of numerical analysis for finding such solutions is essentially connected to variants of Newton's method [8, 10, 14, 23, 26, 28, 33].

A very important aspect in the study of iterative procedures is the convergence domain. In general the convergence domain is small. This is why it is important to enlarge it without additional hypotheses. Then, this is our goal in this chapter.

In the present chapter we study the secant-like method defined by

$$\begin{aligned} &x_{-1}, x_0 \text{ are initial points} \\ &y_n = \lambda x_n + (1-\lambda)x_{n-1},\ \lambda \in [0,1] \\ &x_{n+1} = x_n - B_n^{-1}F(x_n), \quad B_n = [y_n, x_n; F] \qquad \text{for each} \quad n = 0,1,2,\cdots. \end{aligned} \tag{13.1.2}$$

The family of secant-like methods reduces to the secant method if $\lambda = 0$ and to Newton's method if $\lambda = 1$. It was shown in [28] (see also [7, 8, 20, 22] and the references therein) that the R–order of convergence is at least $(1+\sqrt{5})/2$ if $\lambda \in [0,1)$, the same as that of the secant method. In the real case the closer x_n and y_n are, the higher the speed of convergence.

Moreover in [19], it was shown that as λ approaches 1 the speed of convergence is close to that of Newton's method. Moreover, the advantages of using secant-like method instead of Newton's method is that the former method avoids the computation of $F'(x_n)^{-1}$ at each step. The study about convergence matter of iterative procedures is usually centered on two types: semilocal and local convergence analysis. The semilocal convergence matter is, based on the information around an initial point, to give criteria ensuring the convergence of iterative procedure; while the local one is, based on the information around a solution, to find estimates of the radii of convergence balls. There is a plethora of studies on the weakness and/or extension of the hypothesis made on the underlying operators; see for example [1]–[35] or even graphical tools to study this method [25].

The hypotheses used for the semilocal convergence of secant-like method are (see [8, 18, 19, 22]):

($\mathcal{C}_1$) There exists a divided difference of order one denoted by $[x,y;F] \in \mathcal{L}(\mathcal{X},\mathcal{Y})$ satisfying

$$[x,y;F](x-y) = F(x)-F(y) \quad \text{for all} \quad x,y \in \mathcal{D};$$

($\mathcal{C}_2$) There exist x_{-1}, x_0 in $\mathcal{D}$ and $c>0$ such that

$$\| x_0 - x_{-1} \| \leq c;$$

($\mathcal{C}_3$) There exist $x_{-1}, x_0 \in \mathcal{D}$ and $M>0$ such that $A_0^{-1} \in \mathcal{L}(\mathcal{Y},\mathcal{X})$ and

$$\| A_0^{-1}([x,y;F]-[u,v;F]) \| \leq M(\| x-u \| + \| y-v \|) \quad \text{for all} \quad x,y,u,v \in \mathcal{D};$$

($\mathcal{C}_3^\star$) There exist $x_{-1}, x_0 \in \mathcal{D}$ and $L>0$ such that $A_0^{-1} \in \mathcal{L}(\mathcal{Y},\mathcal{X})$ and

$$\| A_0^{-1}([x,y;F]-[v,y;F]) \| \leq L \, \| x-v \| \quad \text{for all} \quad x,y,v \in \mathcal{D};$$

($\mathcal{C}_3^{\star\star}$) There exist $x_{-1}, x_0 \in \mathcal{D}$ and $K>0$ such that $F(x_0)^{-1} \in \mathcal{L}(\mathcal{Y},\mathcal{X})$ and

$$\| F'(x_0)^{-1}([x,y;F]-[v,y;F]) \| \leq K \, \| x-v \| \quad \text{for all} \quad x,y,v \in \mathcal{D};$$

($\mathcal{C}_4$) There exists $\eta > 0$ such that

$$\| A_0^{-1} F(x_0) \| \leq \eta;$$

($\mathcal{C}_4^\star$) There exists $\eta > 0$ for each $\lambda \in [0,1]$ such that

$$\| B_0^{-1} F(x_0) \| \leq \eta.$$

We shall refer to ($\mathcal{C}_1$)–($\mathcal{C}_4$) as the ($\mathcal{C}$) conditions. From analyzing the semilocal convergence of the simplified secant method, it was shown [18] that the convergence criteria are milder than those of secant-like method given in [21]. Consequently, the decreasing and accessibility regions of (13.1.2) can be improved. Moreover, the semilocal convergence of (13.1.2) is guaranteed.

In the present chapter we show: an even larger convergence domain can be obtained under the same or weaker sufficient convergence criteria for method (13.1.2). In view of ($\mathcal{C}_3$) we have that

($\mathcal{C}_5$) There exists $M_0 > 0$ such that

$$\| A_0^{-1}([x,y;F]-[x_{-1},x_0;F]) \| \leq M_0 (\| x-x_{-1} \| + \| y-x_0 \|) \quad \text{for all} \quad x,y \in \mathcal{D}.$$

We shall also use the conditions

($\mathcal{C}_6$) There exist $x_0 \in \mathcal{D}$ and $M_1 > 0$ such that $F'(x_0)^{-1} \in \mathcal{L}(\mathcal{Y},\mathcal{X})$ and

$$\| F'(x_0)^{-1}([x,y;F]-F'(x_0)) \| \leq M_1 (\| x-x_0 \| + \| y-x_0 \|) \quad \text{for all} \quad x,y \in \mathcal{D};$$

($\mathcal{C}_7$) There exist $x_0 \in \mathcal{D}$ and $M_2 > 0$ such that $F'(x_0)^{-1} \in \mathcal{L}(\mathcal{Y},\mathcal{X})$ and

$$\| F'(x_0)^{-1}(F'(x)-F'(x_0)) \| \leq M_2 (\| x-x_0 \| + \| y-x_0 \|) \quad \text{for all} \quad x,y \in \mathcal{D}.$$

Note that $M_0 \leq M$, $M_2 \leq M_1$, $L \leq M$ hold in general and M/M_0, M_1/M_2, M/L can be arbitrarily large [6, 7, 8, 9, 10, 14]. We shall refer to ($\mathcal{C}_1$), ($\mathcal{C}_2$), ($\mathcal{C}_3^{\star\star}$), ($\mathcal{C}_4^{\star}$), ($\mathcal{C}_6$) as the ($\mathcal{C}^{\star}$) conditions and ($\mathcal{C}_1$), ($\mathcal{C}_2$), ($\mathcal{C}_3^{\star}$), ($\mathcal{C}_4^{\star}$), ($\mathcal{C}_5$) as the ($\mathcal{C}^{\star\star}$) conditions. Note that ($\mathcal{C}_5$) is not additional hypothesis to ($\mathcal{C}_3$), since in practice the computation of constant M requires that of M_0. Note that if ($\mathcal{C}_6$) holds, then we can set $M_2 = 2M_1$ in ($\mathcal{C}_7$).

The chapter is organized as follows. In Section 13.2. we use the ($\mathcal{C}^{\star}$) and ($\mathcal{C}^{\star\star}$) conditions instead of the ($\mathcal{C}$) conditions to provide new semilocal convergence analyses for method (13.1.2) under weaker sufficient criteria than those given in [18, 19, 22, 27, 28]. This way we obtain a larger convergence domain and a tighter convergence analysis. Two numerical examples, where we illustrate the improvement of the domain of starting points achieved with the new semilocal convergence results, are given in the Section 13.3..

13.2. Semilocal Convergence of Secant-Like Method

We present the semilocal convergence of secant-like method. First, we need some results on majorizing sequences for secant-like method.

Lemma 13.2.1. *Let $c \geq 0$, $\eta > 0$, $M_1 > 0$, $K > 0$ and $\lambda \in [0,1]$. Set $t_{-1} = 0$, $t_0 = c$ and $t_1 = c + \eta$. Define scalar sequences $\{q_n\}$, $\{t_n\}$, $\{\alpha_n\}$ for each $n = 0, 1, \cdots$ by*

$$q_n = (1-\lambda)(t_n - t_0) + (1+\lambda)(t_{n+1} - t_0),$$

$$t_{n+2} = t_{n+1} + \frac{K(t_{n+1} - t_n + (1-\lambda)(t_n - t_{n-1}))}{1 - M_1 q_n}(t_{n+1} - t_n), \tag{13.2.1}$$

$$\alpha_n = \frac{K(t_{n+1} - t_n + (1-\lambda)(t_n - t_{n-1}))}{1 - M_1 q_n}, \tag{13.2.2}$$

function $\{f_n\}$ for each $n = 1, 2, \cdots$ by

$$\begin{aligned} f_n(t) \quad = \quad & K\eta t^n + K(1-\lambda)\eta t^{n-1} + M_1\eta((1-\lambda)(1+t+\cdots+t^n)+ \\ & (1+\lambda)(1+t+\cdots+t^{n+1})) - 1 \end{aligned} \tag{13.2.3}$$

and polynomial p by

$$p(t) = M_1(1+\lambda)t^3 + (M_1(1-\lambda)+K)t^2 - K\lambda t - K(1-\lambda). \tag{13.2.4}$$

Denote by α *the smallest root of polynomial* p *in* $(0,1)$. *Suppose that*

$$0 < \alpha_0 \leq \alpha \leq 1 - 2M_1\eta. \tag{13.2.5}$$

Then, sequence $\{t_n\}$ *is non-decreasing, bounded from above by* $t^{\star\star}$ *defined by*

$$t^{\star\star} = \frac{\eta}{1-\alpha} + c \tag{13.2.6}$$

and converges to its unique least upper bound $t^{\star}$ *which satisfies*

$$c + \eta \leq t^{\star} \leq t^{\star\star}. \tag{13.2.7}$$

Moreover, the following estimates are satisfied for each $n = 0, 1, \cdots$

$$0 \leq t_{n+1} - t_n \leq \alpha^n \eta \tag{13.2.8}$$

and

$$t^{\star} - t_n \leq \frac{\alpha^n \eta}{1-\alpha}. \tag{13.2.9}$$

Proof. We shall first prove that polynomial p has roots in $(0,1)$. If $\lambda \neq 1$, $p(0) = -(1-\lambda)K < 0$ and $p(1) = 2M_1 > 0$. If $\lambda = 1$, $p(t) = t\,\overline{p}(t)$, $\overline{p}(0) = -K < 0$ and $\overline{p}(1) = 2M_1 > 0$. In either case it follows from the intermediate value theorem that there exist roots in $(0,1)$. Denote by α the minimal root of p in $(0,1)$. Note that, in particular for Newton's method (i.e. for $\lambda = 1$) and for Secant method (i.e. for $\lambda = 0$), we have, respectively by (13.2.4) that

$$\alpha = \frac{2K}{K + \sqrt{K^2 + 4M_1K}} \tag{13.2.10}$$

and

$$\alpha = \frac{2K}{K + \sqrt{K^2 + 8M_1K}}. \tag{13.2.11}$$

It follows from (13.2.1) and (13.2.2) that estimate (13.2.8) is satisfied if

$$0 \leq \alpha_n \leq \alpha. \tag{13.2.12}$$

Estimate (13.2.12) is true by (13.2.5) for $n = 0$. Then, we have by (13.2.1) that

$$\begin{aligned} & t_2 - t_1 \leq \alpha(t_1 - t_0) \Longrightarrow t_2 \leq t_1 + \alpha(t_1 - t_0) \\ & \Longrightarrow t_2 \leq \eta + t_0 + \alpha\eta = c + (1+\alpha)\eta = c + \frac{1-\alpha^2}{1-\alpha\eta} < t^{\star\star}. \end{aligned}$$

Suppose that

$$t_{k+1} - t_k \leq \alpha^k \eta \quad \text{and} \quad t_{k+1} \leq c + \frac{1-\alpha^{k+1}}{1-\alpha}\eta. \tag{13.2.13}$$

Estimate (13.2.12) shall be true for $k+1$ replacing n if

$$0 \leq \alpha_{k+1} \leq \alpha \tag{13.2.14}$$

or

$$f_k(\alpha) \leq 0. \tag{13.2.15}$$

We need a relationship between two consecutive recurrent functions f_k for each $k = 1, 2, \cdots$. It follows from (13.2.3) and (13.2.4) that

$$f_{k+1}(\alpha) = f_k(\alpha) + p(\alpha)\,\alpha^{k-1}\eta = f_k(\alpha), \tag{13.2.16}$$

since $p(\alpha) = 0$. Define function f_∞ on $(0,1)$ by

$$f_\infty(t) = \lim_{n\to\infty} f_n(t). \tag{13.2.17}$$

Then, we get from (13.2.3) and (13.2.17) that

$$\begin{aligned} f_\infty(\alpha) &= \lim_{n\to\infty} f_n(\alpha) \\ &= K\eta \lim_{n\to\infty} \alpha^n + K(1-\lambda)\eta \lim_{n\to\infty} \alpha^{n-1} + \\ &\quad M_1\eta\Big((1-\lambda)\lim_{n\to\infty}(1+\alpha+\cdots+\alpha^n) + \\ &\quad (1+\lambda)\lim_{n\to\infty}(1+\alpha+\cdots+\alpha^{n+1})\Big) - 1 \\ &= M_1\eta\left(\frac{1-\lambda}{1-\alpha} + \frac{1+\lambda}{1-\alpha}\right) - 1 = \frac{2M_1\eta}{1-\alpha} - 1, \end{aligned} \tag{13.2.18}$$

since $\alpha \in (0,1)$. In view of (13.2.15), (13.2.16) and (13.2.18) we can show instead of (13.2.15) that

$$f_\infty(\alpha) \leq 0, \tag{13.2.19}$$

which is true by (13.2.5). The induction for (13.2.8) is complete. It follows that sequence $\{t_n\}$ is non-decreasing, bounded from above by $t^{\star\star}$ given by (13.2.6) and as such it converges to $t^\star$ which satisfies (13.2.7). Estimate (13.2.9) follows from (13.2.8) by using standard majorization techniques [8, 10, 23]. The proof of Lemma 13.2.1 is complete. ☒

Lemma 13.2.2. *Let $c \geq 0$, $\eta > 0$, $M_1 > 0$, $K > 0$ and $\lambda \in [0,1]$. Set $r_{-1} = 0$, $r_0 = c$ and $r_1 = c + \eta$. Define scalar sequences $\{r_n\}$ for each $n = 1, \cdots$ by*

$$\begin{aligned} r_2 &= r_1 + \beta_1(r_1 - r_0) \\ r_{n+2} &= r_{n+1} + \beta_n(r_{n+1} - r_n), \end{aligned} \tag{13.2.20}$$

where

$$\beta_1 = \frac{M_1(r_1 - r_0 + (1-\lambda)(r_0 - r_{-1}))}{1 - M_1 q_1},$$

$$\beta_n = \frac{K(r_{n+1} - r_n + (1-\lambda)(r_n - r_{n-1}))}{1 - M_1 q_n} \quad \textit{for each} \quad n = 2, 3, \cdots$$

and function $\{g_n\}$ on $[0,1)$ for each $n = 1, 2, \cdots$ by

$$\begin{aligned} g_n(t) &= K(t + (1-\lambda))t^{n-1}(r_2 - r_1) + \\ &\quad M_1 t\left((1-\lambda)\frac{1-t^{n+1}}{1-t} + (1+\lambda)\frac{1-t^{n+2}}{1-t}\right)(r_2 - r_1) + (2M_1\eta - 1)t. \end{aligned} \tag{13.2.21}$$

Suppose that

$$0 \le \beta_1 \le \alpha \le 1 - \frac{2M_1(r_2 - r_1)}{1 - 2M_1\eta}, \tag{13.2.22}$$

where α *is defined in Lemma 13.2.1. Then, sequence* $\{r_n\}$ *is non-decreasing, bounded from above by* $r^{\star\star}$ *defined by*

$$r^{\star\star} = c + \eta + \frac{r_2 - r_1}{1 - \alpha} \tag{13.2.23}$$

and converges to its unique least upper bound $r^\star$ *which satisfies*

$$c + \eta \le r^\star \le r^{\star\star}. \tag{13.2.24}$$

Moreover, the following estimates are satisfied for each $n = 1, \cdots$

$$0 \le r_{n+2} - r_{n+1} \le \alpha^n (r_2 - r_1). \tag{13.2.25}$$

Proof. We shall use mathematical induction to show that

$$0 \le \beta_n \le \alpha. \tag{13.2.26}$$

Estimate (13.2.26) is true for $n = 0$ by (13.2.22). Then, we have by (13.2.20) that

$$\begin{aligned} 0 \le r_3 - r_2 \le \alpha(r_2 - r_1) \quad &\Longrightarrow \quad r_3 \le r_2 + \alpha(r_2 - r_1) \\ &\Longrightarrow \quad r_3 \le r_2 + (1+\alpha)(r_2 - r_1) - (r_2 - r_1) \\ &\Longrightarrow \quad r_3 \le r_1 + \frac{1-\alpha^2}{1-\alpha}(r_2 - r_1) \le r^{\star\star}. \end{aligned}$$

Suppose (13.2.26) holds for each $n \le k$, then, using (13.2.20), we obtain that

$$0 \le r_{k+2} - r_{k+1} \le \alpha^k (r_2 - r_1) \quad \text{and} \quad r_{k+2} \le r_1 + \frac{1 - \alpha^{k+1}}{1 - \alpha}(r_2 - r_1). \tag{13.2.27}$$

Estimate (13.2.26) is certainly satisfied, if

$$g_k(\alpha) \le 0, \tag{13.2.28}$$

where g_k is defined by (13.2.21). Using (13.2.21), we obtain the following relationship between two consecutive recurrent functions g_k for each $k = 1, 2, \cdots$

$$g_{k+1}(\alpha) = g_k(\alpha) + p(\alpha)\,\alpha^{k-1}(r_2 - r_1) = g_k(\alpha), \tag{13.2.29}$$

since $p(\alpha) = 0$. Define function g_∞ on $[0,1)$ by

$$g_\infty(t) = \lim_{k \to \infty} g_k(t). \tag{13.2.30}$$

Then, we get from (13.2.21) and (13.2.30) that

$$g_\infty(\alpha) = \alpha\left(\frac{2M_1(r_2 - r_1)}{1 - \alpha} + 2M_1\eta - 1\right). \tag{13.2.31}$$

In view of (13.2.28)–(13.2.31) to show (13.2.28), it suffices to have $g_\infty(\alpha) \le 0$, which true by the right hand hypothesis in (13.2.22). The induction for (13.2.26) (i.e. for (13.2.25)) is complete. The rest of the proof is omitted (as identical to the proof of Lemma 13.2.1). The proof of Lemma 13.2.2 is complete. ⊠

Remark 13.2.3. *Let us see how sufficient convergence criterion on (13.2.5) for sequence $\{t_n\}$ simplifies in the interesting case of Newton's method. That is when $c = 0$ and $\lambda = 1$. Then, (13.2.5) can be written for $L_0 = 2M_1$ and $L = 2K$ as*

$$h_0 = \frac{1}{8}(L + 4L_0 + \sqrt{L^2 + 8L_0 L})\eta \leq \frac{1}{2}. \tag{13.2.32}$$

The convergence criterion in [18] reduces to the famous for it simplicity and clarity Kantorovich hypothesis

$$h = L\eta \leq \frac{1}{2}. \tag{13.2.33}$$

Note however that $L_0 \leq L$ holds in general and L/L_0 can be arbitrarily large [6, 7, 8, 9, 10, 14]. We also have that

$$h \leq \frac{1}{2} \Longrightarrow h_0 \leq \frac{1}{2} \tag{13.2.34}$$

but not necessarily vice versa unless if $L_0 = L$ and

$$\frac{h_0}{h} \longrightarrow \frac{1}{4} \quad \text{as} \quad \frac{L}{L_0} \longrightarrow \infty. \tag{13.2.35}$$

Similarly, it can easily be seen that the sufficient convergence criterion (13.2.22) for sequence $\{r_n\}$ is given by

$$h_1 = \frac{1}{8}(4L_0 + \sqrt{L_0 L + 8L_0^2} + \sqrt{L_0 L})\eta \leq \frac{1}{2}. \tag{13.2.36}$$

We also have that

$$h_0 \leq \frac{1}{2} \Longrightarrow h_1 \leq \frac{1}{2} \tag{13.2.37}$$

and

$$\frac{h_1}{h} \longrightarrow 0, \ \frac{h_1}{h_0} \longrightarrow 0 \quad \text{as} \quad \frac{L_0}{L} \longrightarrow 0. \tag{13.2.38}$$

Note that sequence $\{r_n\}$ is tighter than $\{t_n\}$ and converges under weaker conditions. Indeed, a simple inductive argument shows that for each $n = 2, 3, \cdots$, if $M_1 < K$, then

$$r_n < t_n, \ r_{n+1} - r_n < t_{n+1} - t_n \quad \text{and} \quad r^\star \leq t^\star. \tag{13.2.39}$$

We have the following usefull and obvious extensions of Lemma 13.2.1 and Lemma 13.2.2, respectively.

Lemma 13.2.4. *Let $N = 0, 1, 2, \cdots$ be fixed. Suppose that*

$$t_1 \leq t_2 \leq \cdots \leq t_N \leq t_{N+1}, \tag{13.2.40}$$

$$\frac{1}{M_1} > (1 - \lambda)(t_N - t_0) + (1 + \lambda)(t_{N+1} - t_0) \tag{13.2.41}$$

and

$$0 \leq \alpha_N \leq \alpha \leq 1 - 2M_1(t_{N+1} - t_N). \tag{13.2.42}$$

Then, sequence $\{t_n\}$ *generated by (13.2.1) is nondecreasing, bounded from above by* $t^{\star\star}$ *and converges to* $t^\star$ *which satisfies* $t^\star \in [t_{N+1}, t^{\star\star}]$. *Moreover, the following estimates are satisfied for each* $n = 0, 1, \cdots$

$$0 \le t_{N+n+1} - t_{N+n} \le \alpha^n (t_{N+1} - t_N) \tag{13.2.43}$$

and

$$t^\star - t_{N+n} \le \frac{\alpha^n}{1-\alpha} (t_{N+1} - t_N). \tag{13.2.44}$$

Lemma 13.2.5. *Let* $N = 1, 2, \cdots$ *be fixed. Suppose that*

$$r_1 \le r_2 \le \cdots \le r_N \le r_{N+1}, \tag{13.2.45}$$

$$\frac{1}{M_1} > (1-\lambda)(r_N - r_0) + (1+\lambda)(r_{N+1} - r_0) \tag{13.2.46}$$

and

$$0 \le \beta_N \le \alpha \le 1 - \frac{2M_1 (r_{N+1} - r_N)}{1 - 2M_1 (r_N - r_{N-1})}. \tag{13.2.47}$$

Then, sequence $\{r_n\}$ *generated by (13.2.20) is nondecreasing, bounded from above by* $r^{\star\star}$ *and converges to* $r^\star$ *which satisfies* $r^\star \in [r_{N+1}, r^{\star\star}]$. *Moreover, the following estimates are satisfied for each* $n = 0, 1, \cdots$

$$0 \le r_{N+n+1} - r_{N+n} \le \alpha^n (r_{N+1} - r_N) \tag{13.2.48}$$

and

$$r^\star - r_{N+n} \le \frac{\alpha^n}{1-\alpha} (r_{N+1} - r_N). \tag{13.2.49}$$

Next, we present the following semilocal convergence result for secant-like method under the **(**$\mathcal{C}^\star$**)** conditions.

Theorem 13.2.6. *Suppose that the* **(**$\mathcal{C}^\star$**)**, *Lemma 13.2.1 (or Lemma 13.2.4) conditions and*

$$\overline{U}(x_0, t^\star) \subseteq \mathcal{D} \tag{13.2.50}$$

hold. Then, sequence $\{x_n\}$ *generated by the secant-like method is well defined, remains in* $\overline{U}(x_0, t^\star)$ *for each* $n = -1, 0, 1, \cdots$ *and converges to a solution* $x^\star \in \overline{U}(x_0, t^\star - c)$ *of equation* $F(x) = 0$. *Moreover, the following estimates are satisfied for each* $n = 0, 1, \cdots$

$$\| x_{n+1} - x_n \| \le t_{n+1} - t_n \tag{13.2.51}$$

and

$$\| x_n - x^\star \| \le t^\star - t_n. \tag{13.2.52}$$

Furthemore, if there exists $r \ge t^\star$ *such that*

$$\overline{U}(x_0, r) \subseteq \mathcal{D} \tag{13.2.53}$$

and

$$r + t^\star < \frac{1}{M_1} \quad \text{or} \quad r + t^\star < \frac{2}{M_2}, \tag{13.2.54}$$

then, the solution $x^\star$ *is unique in* $\overline{U}(x_0, r)$.

Proof. We use mathematical induction to prove that

$$\| x_{k+1}-x_k \| \leq t_{k+1}-t_k \tag{13.2.55}$$

and

$$\overline{U}(x_{k+1},t^\star-t_{k+1}) \subseteq \overline{U}(x_k,t^\star-t_k) \tag{13.2.56}$$

for each $k=-1,0,1,\cdots$. Let $z\in\overline{U}(x_0,t^\star-t_0)$. Then, we obtain that

$$\| z-x_{-1} \| \leq \| z-x_0 \| + \| x_0-x_{-1} \| \leq t^\star-t_0+c=t^\star=t^\star-t_{-1},$$

which implies $z\in\overline{U}(x_{-1},t^\star-t_{-1})$. Let also $w\in\overline{U}(x_0,t^\star-t_1)$. We get that

$$\| w-x_0 \| \leq \| w-x_1 \| + \| x_1-x_0 \| \leq t^\star-t_1+t_1-t_0=t^\star=t^\star-t_0.$$

That is $w\in\overline{U}(x_0,t^\star-t_0)$. Note that

$$\| x_{-1}-x_0 \| \leq c=t_0-t_{-1} \quad \text{and} \quad \| x_1-x_0 \| = \| B_0^{-1}F(x_0) \| \leq \eta=t_1-t_0<t^\star,$$

which implies $x_1\in\overline{U}(x_0,t^\star)\subseteq\mathcal{D}$. Hence, estimates (13.2.51) and (13.2.52) hold for $k=-1$ and $k=0$. Suppose (13.2.51) and (13.2.52) hold for all $n\leq k$. Then, we obtain that

$$\| x_{k+1}-x_0 \| \leq \sum_{i=1}^{k+1} \| x_i-x_{i-1} \| \leq \sum_{i=1}^{k+1}(t_i-t_{i-1})=t_{k+1}-t_0\leq t^\star$$

and

$$\| y_k-x_0 \| \leq \lambda \| x_k-x_0 \| +(1-\lambda) \| x_{k-1}-x_0 \| \leq \lambda t^\star+(1-\lambda)t^\star=t^\star.$$

Hence, $x_{k+1},y_k\in\overline{U}(x_0,t^\star)$. Let $E_k:=[x_{k+1},x_k;F]$ for each $k=0,1,\cdots$. Using (13.1.2), Lemma 13.2.1 and the induction hypotheses, we get that

$$\begin{aligned} &\| F'(x_0)^{-1}(B_{k+1}-F'(x_0)) \| \leq M_1(\| y_{k+1}-x_0 \| + \| x_{k+1}-x_0 \|)\\ &\leq M_1((1-\lambda) \| x_k-x_0 \| +\lambda \| x_{k+1}-x_0 \| + \| x_{k+1}-x_0 \|)\\ &\leq M_1((1-\lambda)(t_k-t_0)+(1+\lambda)(t_{k+1}-t_0))<1, \end{aligned} \tag{13.2.57}$$

since, $y_{k+1}-x_0=\lambda(x_{k+1}-x_0)+(1-\lambda)(x_k-x_0)$ and

$$\begin{aligned} &\| y_{k+1}-x_0 \| = \| \lambda(x_{k+1}-x_0)+(1-\lambda)(x_k-x_0) \|\\ &\leq \lambda \| x_{k+1}-x_0 \| +(1-\lambda) \| x_k-x_0 \| . \end{aligned}$$

It follows from (13.2.57) and the Banach lemma on invertible operators that B_{k+1}^{-1} exists and

$$\| B_{k+1}^{-1}F'(x_0) \| \leq \frac{1}{1-\Theta_k} \leq \frac{1}{1-M_1\,q_{k+1}}, \tag{13.2.58}$$

where $\Theta_k=M_1((1-\lambda) \| x_k-x_0 \| +(1+\lambda) \| x_{k+1}-x_0 \|)$. In view of (13.1.2), we obtain the identity

$$F(x_{k+1})=F(x_{k+1})-F(x_k)-B_k(x_{k+1}-x_k)=(E_k-B_k)(x_{k+1}-x_k). \tag{13.2.59}$$

Then, using the induction hypotheses, the $(\mathcal{C}^\star)$ condition and (13.2.59), we get in turn that

$$\begin{aligned}\| F'(x_0)^{-1} F(x_{k+1}) \| &= \| F'(x_0)^{-1} (E_k - B_k)(x_{k+1} - x_k) \| \\ &\leq K \| x_{k+1} - y_k \| \| x_{k+1} - x_k \| \\ &\leq K (\| x_{k+1} - x_k \| + (1-\lambda) \| x_k - x_{k-1} \|) \| x_{k+1} - x_k \| \\ &\leq K (t_{k+1} - t_k + (1-\lambda)(t_k - t_{k-1}))(t_{k+1} - t_k),\end{aligned} \tag{13.2.60}$$

since, $x_{k+1} - y_k = x_{k+1} - x_k + (1-\lambda)(x_k - x_{k-1})$ and

$$\| x_{k+1} - y_k \| \leq \| x_{k+1} - x_k \| + (1-\lambda) \| x_k - x_{k-1} \| \leq t_{k+1} - t_k + (1-\lambda)(t_k - t_{k-1}).$$

It now follows from (13.1.2), (13.2.1), (13.2.58)–(13.2.60) that

$$\begin{aligned}\| x_{k+2} - x_{k+1} \| &\leq \| B_{k+1}^{-1} F'(x_0) \| \| F'(x_0)^{-1} F(x_{k+1}) \| \\ &\leq \frac{K (t_{k+1} - t_k + (1-\lambda)(t_{k+1} - x_k))(t_{k+1} - t_k)}{1 - M_1 q_{k+1}} = t_{k+2} - t_{k+1},\end{aligned}$$

which completes the induction for (13.2.55). Furthemore, let $v \in \overline{U}(x_{k+2}, t^\star - t_{k+2})$. Then, we have that

$$\begin{aligned}\| v - x_{k+1} \| &\leq \| v - x_{k+2} \| + \| x_{k+2} - x_{k+1} \| \\ &\leq t^\star - t_{k+2} + t_{k+2} - t_{k+1} = t^\star - t_{k+1},\end{aligned}$$

which implies $v \in \overline{U}(x_{k+1}, t^\star - t_{k+1})$. The induction for (13.2.55) and (13.2.56) is complete. Lemma 13.2.1 implies that $\{t_k\}$ is a complete sequence. It follows from (13.2.55) and (13.2.56) that $\{x_k\}$ is a complete sequence in a Banach space $\mathcal{X}$ and as such it converges to some $x^\star \in \overline{U}(x_0, t^\star)$ (since $\overline{U}(x_0, t^\star)$ is a closed set). By letting $k \longrightarrow \infty$ in (13.2.60), we get that $F(x^\star) = 0$. Moreover, estimate (13.2.52) follows from (13.2.51) by using standard majorization techniques [8, 10, 23]. To show the uniqueness part, let $y^\star \in \overline{U}(x_0, r)$ be such $F(y^\star) = 0$, where r satisfies (13.2.53) and (13.2.54). We have that

$$\begin{aligned}\| F'(x_0)^{-1} ([y^\star, x^\star; F] - F'(x_0)) \| &\leq M_1 (\| y^\star - x_0 \| + \| x^\star - x_0 \|) \\ &\leq M_1 (t^\star + r) < 1.\end{aligned} \tag{13.2.61}$$

It follows by (13.2.61) and the Banach lemma on invertible operators that linear operator $[y^\star, x^\star; F]^{-1}$ exists. Then, using the identity $0 = F(y^\star) - F(x^\star) = [y^\star, x^\star; F](y^\star - x^\star)$, we deduce that $x^\star = y^\star$. The proof of Theorem 13.2.6 is complete. ⊠

In order for us to present the semilocal result for secant-like method under the $(\mathcal{C}^{\star\star})$ conditions, we first need a result on a majorizing sequence. The proof in given in Lemma 13.2.1.

Remark 13.2.7. *Clearly, (13.2.22) (or (13.2.47)), $\{r_n\}$ can replace (13.2.5) (or (13.2.42)), $\{t_n\}$, respectively in Theorem 13.2.6.*

Lemma 13.2.8. *Let $c \geq 0$, $\eta > 0$, $L > 0$, $M_0 > 0$ with $M_0 c < 1$ and $\lambda \in [0,1]$. Set*

$$s_{-1} = 0,\ s_0 = c,\ s_1 = c + \eta,\ \tilde{K} = \frac{L}{1 - M_0 c} \quad \text{and} \quad \tilde{M}_1 = \frac{M_0}{1 - M_0 c}.$$

Define scalar sequences $\{\tilde{q}_n\}$, $\{s_n\}$, $\{\tilde{\alpha}_n\}$ *for each* $n = 0, 1, \cdots$ *by*

$$\tilde{q}_n = (1-\lambda)(s_n - s_0) + (1+\lambda)(s_{n+1} - s_0),$$

$$s_{n+2} = s_{n+1} + \frac{\tilde{K}(s_{n+1} - s_n + (1-\lambda)(s_n - s_{n-1}))}{1 - \tilde{M}_1 \tilde{q}_n}(s_{n+1} - s_n),$$

$$\tilde{\alpha}_n = \frac{\tilde{K}(s_{n+1} - s_n + (1-\lambda)(s_n - s_{n-1}))}{1 - \tilde{M}_1 \tilde{q}_n},$$

function $\{\tilde{f}_n\}$ *for each* $n = 1, 2, \cdots$ *by*

$$\begin{aligned}\tilde{f}_n(t) \quad = \quad & \tilde{K}\eta t^n + \tilde{K}(1-\lambda)\eta t^{n-1} + \tilde{M}_1 \eta((1-\lambda)(1+t+\cdots+t^n)+ \\ & (1+\lambda)(1+t+\cdots+t^{n+1})) - 1\end{aligned}$$

and polynomial $\tilde{p}$ *by*

$$\tilde{p}(t) = \tilde{M}_1(1+\lambda)t^3 + (\tilde{M}_1(1-\lambda) + \tilde{K})t^2 - \tilde{K}\lambda t - \tilde{K}(1-\lambda).$$

Denote by $\tilde{\alpha}$ *the smallest root of polynomial* $\tilde{p}$ *in* $(0,1)$. *Suppose that*

$$0 \leq \tilde{\alpha}_0 \leq \tilde{\alpha} \leq 1 - 2\tilde{M}_1 \eta. \tag{13.2.62}$$

Then, sequence $\{s_n\}$ *is non-decreasing, bounded from above by* $s^{\star\star}$ *defined by*

$$s^{\star\star} = \frac{\eta}{1-\tilde{\alpha}} + c$$

and converges to its unique least upper bound $s^\star$ *which satisfies* $c + \eta \leq s^\star \leq s^{\star\star}$. *Moreover, the following estimates are satisfied for each* $n = 0, 1, \cdots$

$$0 \leq s_{n+1} - s_n \leq \tilde{\alpha}^n \eta \quad \text{and} \quad s^\star - s_n \leq \frac{\tilde{\alpha}^n \eta}{1-\tilde{\alpha}}.$$

Next, we present the semilocal convergence result for secant-like method under the $(\mathcal{C}^{\star\star})$ conditions.

Theorem 13.2.9. *Suppose that the* $(\mathcal{C}^{\star\star})$ *conditions, (13.2.62) (or Lemma 13.2.2 conditions with* $\tilde{\alpha}_n$, $\tilde{\alpha}$, $\tilde{M}_1$ *replacing, respectively,* α_n, α, M_1*) and* $\overline{U}(x_0, s^\star) \subseteq \mathcal{D}$ *hold. Then, sequence* $\{x_n\}$ *generated by the secant-like method is well defined, remains in* $\overline{U}(x_0, s^\star)$ *for each* $n = -1, 0, 1, \cdots$ *and converges to a solution* $x^\star \in \overline{U}(x_0, s^\star)$ *of equation* $F(x) = 0$. *Moreover, the following estimates are satisfied for each* $n = 0, 1, \cdots$

$$\| x_{n+1} - x_n \| \leq s_{n+1} - s_n \quad \text{and} \quad \| x_n - x^\star \| \leq s^\star - s_n.$$

Furthemore, if there exists $r \geq s^\star$ *such that* $\overline{U}(x_0, r) \subseteq \mathcal{D}$ *and* $r + s^\star + c < 1/M_0$, *then, the solution* $x^\star$ *is unique in* $\overline{U}(x_0, r)$.

Proof. The proof is analogous to Theorem 13.2.6. Simply notice that in view of ($\mathcal{C}_5$), we obtain instead of (13.2.57) that

$$\begin{aligned}&\| A_0^{-1}\,(B_{k+1}-A_0)\,\| \le M_0\,(\| y_{k+1}-x_{-1}\| + \| x_{k+1}-x_0 \|)\\ &\le M_0\,((1-\lambda)\,\| x_k-x_0\| + \lambda\,\| x_{k+1}-x_0\| + \| x_0-x_{-1}\| + \| x_{k+1}-x_0\|)\\ &\le M_0\,((1-\lambda)\,(s_k-s_0)+(1+\lambda)\,(s_{k+1}-s_0)+c) < 1,\end{aligned}$$

leading to B_{k+1}^{-1} exists and

$$\| B_{k+1}^{-1}A_0 \| \le \frac{1}{1-\Xi_k},$$

where $\Xi_k = M_0\,((1-\lambda)\,(s_k-s_0)+(1+\lambda)\,(s_{k+1}-s_0)+c)$. Moreover, using ($\mathcal{C}_3^\star$) instead of ($\mathcal{C}_3^{\star\star}$), we get that

$$\| A_0^{-1}\,F(x_{k+1})\,\| \le L\,(s_{k+1}-s_k+(1-\lambda)\,(s_k-s_{k-1}))\,(s_{k+1}-s_k).$$

Hence, we have that

$$\begin{aligned}&\| x_{k+2}-x_{k+1} \| \le \| B_{k+1}^{-1}A_0 \|\,\| A_0^{-1}\,F(x_{k+1})\,\|\\ &\le \frac{L\,(s_{k+1}-s_k+(1-\lambda)\,(s_k-s_{k-1}))\,(s_{k+1}-s_k)}{1-M_0\,((1+\lambda)\,(s_{k+1}-s_0)+(1-\lambda)\,(s_k-s_0)+c)}\\ &\le \frac{\tilde{K}\,(s_{k+1}-s_k+(1-\lambda)\,(s_k-s_{k-1}))\,(s_{k+1}-s_k)}{1-\tilde{M}_1\,((1+\lambda)\,(s_{k+1}-s_0)+(1-\lambda)\,(s_k-s_0))} = s_{k+2}-s_{k+1}.\end{aligned}$$

The uniqueness part is given in Theorem 13.2.6 with r, $s^\star$ replacing R_2 and R_0, respectively. The proof of Theorem 13.2.9 is complete. ⊠

Remark 13.2.10. *(a) Condition (13.2.50) can be replaced by*

$$\overline{U}(x_0,t^{\star\star}) \subseteq \mathcal{D}, \tag{13.2.63}$$

where $t^{\star\star}$ is given in the closed form by (13.2.55).

(b) The majorizing sequence $\{u_n\}$ essentially used in [18] is defined by

$$\begin{aligned}&u_{-1}=0,\ u_0=c,\ u_1=c+\eta\\ &u_{n+2}=u_{n+1}+\frac{M\,(u_{n+1}-u_n+(1-\lambda)\,(u_n-u_{n-1}))}{1-M\,q_n^\star}\,(u_{n+1}-u_n),\end{aligned} \tag{13.2.64}$$

where

$$q_n^\star = (1-\lambda)\,(u_n-u_0)+(1+\lambda)\,(u_{n+1}-u_0).$$

Then, if $K<M$ or $M_1<M$, a simple inductive argument shows that for each $n = 2,3,\cdots$

$$t_n < u_n, \quad t_{n+1}-t_n < u_{n+1}-u_n \quad \text{and} \quad t^\star \le u^\star = \lim_{n\to\infty} u_n. \tag{13.2.65}$$

Clearly $\{t_n\}$ converges under the ($\mathcal{C}$) conditions and conditions of Lemma 2.1. Moreover, as we already showed in Remark 13.2.3, the sufficient convergence criteria of Theorem 13.2.6 can be weaker than those of Theorem 13.2.9. Similarly if $L \le M$, $\{s_n\}$ is a tighter sequence than $\{u_n\}$. In general, we shall test the convergence criteria and use the tightest sequence to estimate the error bounds.

(c) *Clearly the conclusions of Theorem 13.2.9 hold if $\{s_n\}$, (13.2.62) are replaced by $\{\tilde{r}_n\}$, (13.2.22), where $\{\tilde{r}_n\}$ is defined as $\{r_n\}$ with M_0 replacing M_1 in the definition of β_1 (only at the numerator) and the tilda letters replacing the non-tilda letters in (13.2.22).*

13.3. Numerical Examples

Now, we check numerically with two examples that the new semilocal convergence results obtained in Theorems 13.2.6 and 13.2.9 improve the domain of starting points obtained by the following classical result given in [21].

Theorem 13.3.1. *Let X and Y be two Banach spaces and $F : \Omega \subseteq X \to Y$ be a nonlinear operator defined on a non-empty open convex domain Ω. Let $x_{-1}, x_0 \in \Omega$ and $\lambda \in [0,1]$. Suppose that there exists $[u,v;F] \in \mathcal{L}(X,Y)$, for all $u, v \in \Omega$ ($u \neq v$), and the following four conditions*

- $\|x_0 - x_{-1}\| = c \neq 0$ *with* $x_{-1}, x_0 \in \Omega$,
- *Fixed* $\lambda \in [0,1]$, *the operator* $B_0 = [y_0, x_0; F]$ *is invertible and such that* $\|B_0^{-1}\| \leq \beta$,
- $\|B_0^{-1} F(x_0)\| \leq \eta$,
- $\|[x,y;F] - [u,v;F]\| \leq Q(\|x-u\| + \|y-v\|); Q \geq 0; x,y,u,v \in \Omega; x \neq y; u \neq v$,

are satisfied. If $B(x_0,\rho) \subseteq \Omega$, where $\rho = \dfrac{1-a}{1-2a}\eta$,

$$a = \frac{\eta}{c+\eta} < \frac{3-\sqrt{5}}{2} \qquad \text{and} \qquad b = \frac{Q\beta c^2}{c+\eta} < \frac{a(1-a)^2}{1+\lambda(2a-1)}, \tag{13.3.1}$$

then the secant-like methods defined by (13.1.2)converge to a solution x^* of equation $F(x) = 0$ with R-order of convergence at least $\frac{1+\sqrt{5}}{2}$. Moreover, $x_n, x^* \in \overline{B(x_0,\rho)}$, the solution x^* is unique in $B(x_0,\tau) \cap \Omega$, where $\tau = \frac{1}{Q\beta} - \rho - (1-\lambda)\alpha$.

13.3.1. Example 1

We illustrate the above-mentioned with an application, where a system of nonlinear equations is involved. We see that Theorem 13.3.1 cannot guarantee the semilocal convergence of secant-like methods (13.1.2), but Theorem 13.2.6 can do it.

It is well known that energy is dissipated in the action of any real dynamical system, usually through some form of friction. However, in certain situations this dissipation is so slow that it can be neglected over relatively short periods of time. In such cases we assume the law of conservation of energy, namely, that the sum of the kinetic energy and the potential energy is constant. A system of this kind is said to be conservative.

If φ and ψ are arbitrary functions with the property that $\varphi(0) = 0$ and $\psi(0) = 0$, the general equation

$$\mu \frac{d^2x(t)}{dt^2} + \psi\left(\frac{dx(t)}{dt}\right) + \varphi(x(t)) = 0, \tag{13.3.2}$$

can be interpreted as the equation of motion of a mass μ under the action of a restoring force $-\varphi(x)$ and a damping force $-\psi(dx/dt)$. In general these forces are nonlinear, and equation (13.3.2) can be regarded as the basic equation of nonlinear mechanics. In this chapter we shall consider the special case of a nonlinear conservative system described by the equation

$$\mu\frac{d^2x(t)}{dt^2}+\varphi(x(t))=0,$$

in which the damping force is zero and there is consequently no dissipation of energy. Extensive discussions of (13.3.2), with applications to a variety of physical problems, can be found in classical references [4] and [32].

Now, we consider the special case of a nonlinear conservative system described by the equation

$$\frac{d^2x(t)}{dt^2}+\phi(x(t))=0 \tag{13.3.3}$$

with the boundary conditions

$$x(0)=x(1)=0. \tag{13.3.4}$$

After that, we use a process of discretization to transform problem (13.3.3)–(13.3.4) into a finite-dimensional problem and look for an approximated solution of it when a particular function ϕ is considered. So, we transform problem (13.3.3)–(13.3.4) into a system of nonlinear equations by approximating the second derivative by a standard numerical formula.

Firstly, we introduce the points $t_j = jh$, $j=0,1,\ldots,m+1$, where $h=\frac{1}{m+1}$ and m is an appropriate integer. A scheme is then designed for the determination of numbers x_j, it is hoped, approximate the values $x(t_j)$ of the true solution at the points t_j. A standard approximation for the second derivative at these points is

$$x_j''\approx\frac{x_{j-1}-2x_j+x_{j+1}}{h^2},\quad j=1,2,\ldots,m.$$

A natural way to obtain such a scheme is to demand that the x_j satisfy at each interior mesh point t_j the difference equation

$$x_{j-1}-2x_j+x_{j+1}+h^2\phi(x_j)=0. \tag{13.3.5}$$

Since x_0 and x_{m+1} are determined by the boundary conditions, the unknowns are $x_1,x_2,\ldots,x_m$.

A further discussion is simplified by the use of matrix and vector notation. Introducing the vectors

$$\mathbf{x}=\begin{pmatrix}x_1\\x_2\\\vdots\\x_m\end{pmatrix},\qquad v_{\mathbf{x}}=\begin{pmatrix}\phi(x_1)\\\phi(x_2)\\\vdots\\\phi(x_m)\end{pmatrix}$$

and the matrix

$$A=\begin{pmatrix}-2&1&0&\cdots&0\\1&-2&1&\cdots&0\\0&1&-2&\cdots&0\\\vdots&\vdots&\vdots&\ddots&\vdots\\0&0&0&\cdots&-2\end{pmatrix},$$

the system of equations, arising from demanding that (13.3.5) holds for $j = 1,2,\ldots,m$, can be written compactly in the form

$$F(\mathbf{x}) \equiv A\mathbf{x} + h^2 v_{\mathbf{x}} = 0, \tag{13.3.6}$$

where F is a function from $\mathbb{R}^m$ into $\mathbb{R}^m$.

From now on, the focus of our attention is to solve a particular system of form (13.3.6). We choose $m = 8$ and the infinity norm.

The steady temperature distribution is known in a homogeneous rod of length 1 in which, as a consequence of a chemical reaction or some such heat-producing process, heat is generated at a rate $\phi(x(t))$ per unit time per unit length, $\phi(x(t))$ being a given function of the excess temperature x of the rod over the temperature of the surroundings. If the ends of the rod, $t = 0$ and $t = 1$, are kept at given temperatures, we are to solve the boundary value problem given by (13.3.3)–(13.3.4), measured along the axis of the rod. For an example we choose an exponential law $\phi(x(t)) = \exp(x(t))$ for the heat generation.

Taking into account that the solution of (13.3.3)–(13.3.4) with $\phi(x(t)) = \exp(x(t))$ is of the form

$$x(s) = \int_0^1 G(s,t)\exp(x(t))\,dt,$$

where $G(s,t)$ is the Green function in $[0,1]\times[0,1]$, we can locate the solution $x^*(s)$ in some domain. So, we have

$$\|x^*(s)\| - \frac{1}{8}\exp(\|x^*(s)\|) \le 0,$$

so that $\|x^*(s)\| \in [0,\rho_1]\cup[\rho_2,+\infty]$, where $\rho_1 = 0.1444$ and $\rho_2 = 3.2616$ are the two positive real roots of the scalar equation $8t - \exp(t) = 0$.

Observing the semilocal convergence results presented in this chapter, we can only guarantee the semilocal convergence to a solution $x^*(s)$ such that $\|x^*(s)\| \in [0,\rho_1]$. For this, we can consider the domain

$$\Omega = \{x(s) \in C^2[0,1]\,;\ \|x(s)\| < \log(7/4), s \in [0,1]\},$$

since $\rho_1 < \log\left(\frac{7}{4}\right) < \rho_2$.

In view of what the domain Ω is for equation (13.3.3), we then consider (13.3.6) with $F : \widetilde{\Omega} \subset \mathbb{R}^8 \to \mathbb{R}^8$ and

$$\widetilde{\Omega} = \{\mathbf{x} \in \mathbb{R}^8;\ \|\mathbf{x}\| < \log(7/4)\}.$$

According to the above-mentioned, $v_{\mathbf{x}} = (\exp(x_1),\exp(x_2),\ldots,\exp(x_8))^t$ if $\phi(x(t)) = \exp(x(t))$. Consequently, the first derivative of the function F defined in (13.3.6) is given by

$$F'(\mathbf{x}) = A + h^2\mathrm{diag}(v_{\mathbf{x}}).$$

Moreover,

$$F'(\mathbf{x}) - F'(\mathbf{y}) = h^2\mathrm{diag}(\mathbf{z}),$$

where $\mathbf{y} = (y_1,y_2,\ldots,y_8)^t$ and $\mathbf{z} = (\exp(x_1)-\exp(y_1),\exp(x_2)-\exp(y_2),\ldots,\exp(x_8)-\exp(y_8))$. In addition,

$$\|F'(\mathbf{x}) - F'(\mathbf{y})\| \le h^2 \max_{1\le i\le 8} |\exp(\ell_i)|\,\|\mathbf{x}-\mathbf{y}\|,$$

where $\ell = (\ell_1, \ell_2, \ldots, \ell_8)^t \in \widetilde{\Omega}$ and $h = \frac{1}{9}$, so that

$$\|F'(\mathbf{x}) - F'(\mathbf{y})\| \leq \frac{7}{4} h^2 \|\mathbf{x} - \mathbf{y}\|. \tag{13.3.7}$$

Considering (see [28])

$$[\mathbf{x}, \mathbf{y}; F] = \int_0^1 F'(\tau\mathbf{x} + (1-\tau)\mathbf{y})\, d\tau,$$

taking into account

$$\int_0^1 \|\tau(\mathbf{x} - \mathbf{u}) + (1-\tau)(\mathbf{y} - \mathbf{v})\| d\tau \leq \frac{1}{2}\left(\|\mathbf{x} - \mathbf{u}\| + \|\mathbf{y} - \mathbf{v}\|\right),$$

and (13.3.7), we have

$$\begin{aligned}
\|[\mathbf{x}, \mathbf{y}; F] - [\mathbf{u}, \mathbf{v}; F]\| &\leq \int_0^1 \|F'(\tau\mathbf{x} + (1-\tau)\mathbf{y}) - F'(\tau\mathbf{u} + (1-\tau)\mathbf{v})\| \, d\tau \\
&\leq \frac{7}{4} h^2 \int_0^1 (\tau\|\mathbf{x} - \mathbf{u}\| + (1-\tau)\|\mathbf{y} - \mathbf{v}\|)\, d\tau \\
&= \frac{7}{8} h^2 \left(\|\mathbf{x} - \mathbf{u}\| + \|\mathbf{y} - \mathbf{v}\|\right).
\end{aligned}$$

From the last, we have $L = \frac{7}{648}$ and $M_1 = \frac{7}{648}\|[F'(x_0)]^{-1}\|$.

If we choose $\lambda = \frac{1}{2}$ and the starting points $\mathbf{x_{-1}} = (\frac{1}{10}, \frac{1}{10}, \ldots, \frac{1}{10})^t$ and $\mathbf{x_0} = (0, 0, \ldots, 0)^t$, we obtain $c = \frac{1}{10}$, $\beta = 11.202658\ldots$ and $\eta = 0.138304\ldots$, so that (13.3.1) of Theorem 13.3.1 is not satisfied, since

$$a = \frac{\eta}{c + \eta} = 0.580368\ldots > \frac{3 - \sqrt{5}}{2} = 0.381966\ldots$$

Thus, according to Theorem 13.3.1, we cannot guarantee the convergence of secant-like method (13.1.2) with $\lambda = \frac{1}{2}$ for approximating a solution of (13.3.6) with $\phi(s) = \exp(s)$.

However, we can do it by Theorem 13.2.6, since all the inequalities which appear in (2.5) are satisfied:

$$0 < \alpha_0 = 0.023303\ldots \leq \alpha = 0.577350\ldots \leq 1 - 2M_1\eta = 0.966625\ldots,$$

where $\|[F'(x_0)]^{-1}\| = 11.169433\ldots$, $M_1 = 0.120657\ldots$ and

$$p(t) = (0.180986\ldots)t^3 + (0.180986\ldots)t^2 - (0.060328\ldots)t - (0.060328\ldots).$$

Then, we can use secant-like method (13.1.2) with $\lambda = \frac{1}{2}$ to approximate a solution of (13.3.6) with $\phi(u) = \exp(u)$, the approximation given by the vector $\mathbf{x}^* = (x_1^*, x_2^*, \ldots, x_8^*)^t$ shown in Table 13.3.1 and reached after four iterations with a tolerance 10^{-16}. In Table 13.3.2 we show the errors $\|\mathbf{x_n} - \mathbf{x}^*\|$ using the stopping criterion $\|\mathbf{x_n} - \mathbf{x_{n-1}}\| < 10^{-16}$. Notice that the vector shown in Table 13.3.1 is a good approximation of the solution of (13.3.6) with $\phi(u) = \exp(u)$, since $\|F(\mathbf{x}^*)\| \leq C \times 10^{-16}$. See the sequence $\{\|F(\mathbf{x_n})\|\}$ in Table 13.3.2.

Table 13.3.1. Approximation of the solution $\mathbf{x}^*$ of (13.3.6) with $\phi(u) = \exp(u)$

n	x_i^*	n	x_i^*	n	x_i^*	n	x_i^*
1	$0.05481058\ldots$	3	$0.12475178\ldots$	5	$0.13893761\ldots$	7	$0.09657993\ldots$
2	$0.09657993\ldots$	4	$0.13893761\ldots$	6	$0.12475178\ldots$	8	$0.05481058\ldots$

Table 13.3.2. Absolute errors obtained by secant-like method (13.1.2) with $\lambda = \frac{1}{2}$ and $\{\|F(\mathbf{x_n})\|\}$

n	$\|\mathbf{x_n} - \mathbf{x}^*\|$	$\|F(\mathbf{x_n})\|$
-1	$1.3893\ldots \times 10^{-1}$	$8.6355\ldots \times 10^{-2}$
0	$4.5189\ldots \times 10^{-2}$	$1.2345\ldots \times 10^{-2}$
1	$1.43051\ldots \times 10^{-4}$	$2.3416\ldots \times 10^{-5}$
2	$1.14121\ldots \times 10^{-7}$	$1.9681\ldots \times 10^{-8}$
3	$4.30239\ldots \times 10^{-13}$	$5.7941\ldots \times 10^{-14}$

13.3.2. Example 2

Consider the following nonlinear boundary value problem

$$\begin{cases} u'' = -u^3 - \dfrac{1}{4}\,u^2 \\ u(0) = 0, \quad u(1) = 1. \end{cases}$$

It is well known that this problem can be formulated as the integral equation

$$u(s) = s + \int_0^1 Q(s,t)\,(u^3(t) + \frac{1}{4}\,u^2(t))\,dt \tag{13.3.8}$$

where, Q is the Green function:

$$Q(s,t) = \begin{cases} t\,(1-s), & t \le s \\ s\,(1-t), & s < t. \end{cases}$$

We observe that

$$\max_{0 \le s \le 1} \int_0^1 |Q(s,t)|\,dt = \frac{1}{8}.$$

Then problem (13.3.8) is in the form (13.1.1), where, F is defined as

$$[F(x)]\,(s) = x(s) - s - \int_0^1 Q(s,t)\,(x^3(t) + \frac{1}{4}\,x^2(t))\,dt.$$

The Fréchet derivative of the operator F is given by

$$[F'(x)y]\,(s) = y(s) - 3\int_0^1 Q(s,t)x^2(t)y(t)dt - \frac{1}{2}\int_0^1 Q(s,t)x(t)y(t)dt.$$

Choosing $x_0(s) = s$ and $R = 1$ we have that $\|F(x_0)\| \leq \frac{1+\frac{1}{4}}{8} = \frac{5}{32}$. Define the divided difference defined by

$$[x,y;F] = \int_0^1 F'(\tau x + (1-\tau)y)d\tau.$$

Taking into account that

$$\begin{aligned}
\|[\mathbf{x},\mathbf{y};F] - [\mathbf{v},\mathbf{y};F]\| &\leq \int_0^1 \|F'(\tau\mathbf{x} + (1-\tau)\mathbf{y}) - F'(\tau\mathbf{v} + (1-\tau)\mathbf{y})\| \, d\tau \\
&\leq \frac{1}{8}\int_0^1 \left(3\tau^2\|x^2 - v^2\| + 2\tau(1-\tau)\|y\|\|x-v\| + \frac{\tau}{2}\|x-v\|\right) d\tau \\
&\leq \frac{1}{8}\left(\|\mathbf{x}^2 - \mathbf{v}^2\| + \left(\|y\| + \frac{1}{4}\right)\right)\|\mathbf{x}-\mathbf{v}\| \\
&\leq \frac{1}{8}\left(\|x+v\| + \|y\| + \frac{1}{4}\right)\|\mathbf{x}-\mathbf{v}\| \\
&\leq \frac{25}{32}\|\mathbf{x}-\mathbf{v}\|
\end{aligned}$$

Choosing $x_{-1}(s) = \frac{9s}{10}$, we find that

$$\begin{aligned}
\|1 - A_0\| &\leq \int_0^1 \|F'(\tau\mathbf{x_0} + (1-\tau)\mathbf{x_{-1}})\| \, d\tau \\
&\leq \frac{1}{8}\int_0^1 \left(3\left(\tau + (1-\tau)\frac{9}{10}\right)^2 + \frac{1}{2}\left(\tau + (1-\tau)\frac{9}{10}\right)\right) d\tau \\
&\leq 0.409375\ldots
\end{aligned}$$

Using the Banach Lemma on invertible operators we obtain

$$\|A_0^{-1}\| \leq 1.69312\ldots$$

and so

$$L \geq \frac{25}{32}\|A_0^{-1}\| = 1.32275\ldots$$

.

In an analogous way, choosing $\lambda = 0.8$ we obtain

$$M_0 = 0.899471\ldots,$$

$$\|B_0^{-1}\| = 1.75262\ldots$$

and

$$\eta = 0.273847\ldots.$$

Notice that we can not guarantee the convergence of the secant method by Theorem 13.3.1 since the first condition of (3.1) is not satisfied:

$$a = \frac{\eta}{c+\eta} = 0.732511\ldots > \frac{3-\sqrt{5}}{2} = 0.381966\ldots$$

On the other hand, observe that

$$\tilde{M}_1 = 0.0988372\ldots,$$

$$\tilde{K} = 1.45349\ldots,$$

$$\alpha_0 = 0.434072\ldots,$$

$$\alpha = 0.907324\ldots$$

and

$$1 - 2\tilde{M}_1\eta = 0.945868\ldots.$$

And condition (2.62) $0 < \alpha_0 \leq \alpha \leq 1 - 2\tilde{M}_1\eta$ is satisfied and as a consequence we can ensure the convergence of the secant method by Theorem 13.2.9.

Conclusion

We presented a new semilocal convergence analysis of the secant-like method for approximating a locally unique solution of an equation in a Banach space. Using a combination of Lipschitz and center-Lipschitz conditions, instead of only Lipschitz conditions invested in [18], we provided a finer analysis with larger convergence domain and weaker sufficient convergence conditions than in [15, 18, 19, 22, 27, 28]. Numerical examples validate our theoretical results.

References

[1] Amat, S., Busquier, S., Negra, M., Adaptive approximation of nonlinear operators, *Numer. Funct. Anal. Optim.*, 25 (2004) 397-405.

[2] Amat, S., Busquier, S., Gutiérrez, J. M., On the local convergence of secant-type methods, *International Journal of Computer Mathematics*, 81 (8) (2004), 1153-1161.

[3] Amat, S., Bermúdez, C., Busquier, S., Gretay, J. O., Convergence by nondiscrete mathematical induction of a two step secant's method, *Rocky Mountain Journal of Mathematics*, 37 (2) (2007), 359-369.

[4] Andronow, A.A., Chaikin, C.E., *Theory of oscillations*, Princenton University Press, New Jersey, 1949.

[5] Argyros, I.K., *Polynomial Operator Equations in Abstract Spaces and Applications*, St.Lucie/CRC/Lewis Publ. Mathematics series, 1998, Boca Raton, Florida, U.S.A.

[6] Argyros, I.K., A unifying local-semilocal convergence analysis and applications for two-point Newton-like methods in Banach space, *J. Math. Anal. Appl.*, 298 (2004), 374-397.

[7] Argyros, I.K., New sufficient convergence conditions for the secant method *Chechoslovak Math. J.*, 55 (2005), 175-187.

[8] Argyros, I.K., *Convergence and Applications of Newton–type Iterations*, Springer–Verlag Publ., New–York, 2008.

[9] Argyros, I.K., A semilocal convergence analysis for directional Newton methods, *Math. Comput.*, 80 (2011), 327-343.

[10] Argyros, I.K., Cho, Y.J., Hilout, S., *Numerical Methods for Equations and its Applications*, CRC Press/Taylor and Francis, Boca Raton, Florida, USA, 2012

[11] Argyros, I.K., Hilout, S., Convergence conditions for secant–type methods, *Chechoslovak Math. J.*, 60 (2010), 253-272.

[12] Argyros, I.K., Hilout, S., Weaker conditions for the convergence of Newton's method, *J. Complexity*, 28 (2012), 364-387.

[13] Argyros, I.K., Hilout, S., Estimating upper bounds on the limit points of majorizing sequences for Newton's method, *Numer. Algorithms*, 62 (1) (2013), 115-132.

[14] Argyros, I.K., Hilout, S., Numerical methods in nonlinear analysis, World Scientific Publ. Comp., New Jersey, 2013

[15] Argyros, I.K., Ezquerro, J.A., Hernández, M.Á. Hilout, S., Romero, N., Velasco, A.I., Expanding the applicability of secant-like methods for solving nonlinear equations, *Carp. J. Math.* 31 (1) (2015), 11-30.

[16] Dennis, J.E., Toward a unified convergence theory for Newton–like methods, in Nonlinear Functional Analysis and Applications (L.B. Rall, ed.), Academic Press, New York, (1971), 425-472.

[17] Ezquerro, J.A., Gutiérrez, J.M., Hernández, M.A., Romero, N., Rubio, M.J., The Newton method: from Newton to Kantorovich. (Spanish), *Gac. R. Soc. Mat. Esp.*, 13 (2010), 53-76.

[18] Ezquerro, J.M., Hernández, M.A., Romero, N., A.I. Velasco,Improving the domain of starting point for secant-like methods,*App. Math. Comp.*, 219 (8) (2012), 3677–3692.

[19] Ezquerro, J.A., Rubio, M.J., A uniparametric family of iterative processes for solvingnondifferentiable equations, *J. Math. Anal. Appl*, 275 (2002), 821-834.

[20] A. Fraile, E. Larrodé, Á. A. Magreñán, J. A. Sicilia. Decision model for siting transport and logistic facilities in urban environments: A methodological approach. *J. Comp. Appl. Math.*, 291 (2016), 478-487.

[21] Hernández, M.A., Rubio, M.J., Ezquerro, J.A., secant-like methods for solving nonlinear integral equations of the Hammerstein type. Proceedings of the 8th International Congress on Computational and Applied Mathematics, ICCAM-98 (Leuven), *J. Comput. Appl. Math.*, 115 (2000), 245-254.

[22] Hernández, M.A., Rubio, M.J., Ezquerro, J.A., Solving a special case of conservative problems by secant-like methods, *App. Math. Comp.*, 169 (2005), 926-942.

[23] Kantorovich, L.V., Akilov, G.P., *Functional Analysis*, Pergamon Press, Oxford, 1982.

[24] Laasonen, P., Ein überquadratisch konvergenter iterativer algorithmus, Ann. Acad. Sci. Fenn. Ser I, 450 (1969), 1–10.

[25] Magreñán, Á. A., A new tool to study real dynamics: The convergence plane, *App. Math. Comp.*, 248 (2014), 215-224.

[26] Ortega, J.M., Rheinboldt, W.C., *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970.

[27] Potra, F.A., Sharp error bounds for a class of Newton–like methods, *Libertas Mathematica*, 5 (1985), 71-84.

[28] Potra, F.A., Pták, V., *Nondiscrete Induction and Iterative Processes*, Pitman, New York, 1984.

[29] Proinov, P.D., General local convergence theory for a class of iterative processes and its applications to Newton's method, *J. Complexity*, 25 (2009), 38-62.

[30] Proinov, P.D., New general convergence theory for iterative processes and its applications to Newton–Kantorovich type theorems, *J. Complexity*, 26 (2010), 3-42.

[31] Schmidt, J.W., Untere Fehlerschranken fur Regula-Falsi Verhafren, *Period. Hungar.*, 9 (1978), 241-247.

[32] Stoker, J. J., *Nonlinear vibrations*, Interscience-Wiley, New York, 1950.

[33] Traub, J.F., *Iterative Methods for the Solution of Equations*, Prentice-Hall, Englewood Cliffs, New Jersay, 1964.

[34] Yamamoto, T., A convergence theorem for Newton–like methods in Banach spaces, *Numer. Math.*, 51 (1987), 545-557.

[35] Wolfe, M.A., Extended iterative methods for the solution of operator equations, *Numer. Math.*, 31 (1978), 153-174.

Chapter 14

Solving Nonlinear Equations System via an Efficient Genetic Algorithm with Symmetric and Harmonious Individuals

14.1. Introduction

In this chapter, we introduce genetic algorithms as a general tool for solving optimum problems. As a special case we use these algorithms to find the solution of the system of nonlinear equations

$$\begin{cases} f_1(x_1, x_2, \ldots, x_n) = 0, \\ f_2(x_1, x_2, \ldots, x_n) = 0, \\ \vdots \\ f_n(x_1, x_2, \ldots, x_n) = 0, \end{cases} \tag{14.1.1}$$

where $f = (f_1, f_2, \ldots, f_n) : D = [a_1, b_1] \times [a_2, b_2] \cdots \times [a_n, b_n] \subseteq R^n \to R^n$ is continuous.

Genetic algorithms (GA) were first introduced in the 1970s by John Holland at University of Michigan [8]. Since then, a great deal of developments on GA have been obtained, see [5, 6] and references therein. GA are used as an adaptive machine study approach in their former period of development, and they have been successfully applied in numerous areas such as artificial intelligence, self-adaption control, systematic engineering, image processing, combinatorial optimization and financial system. GA show us very extensive application prospect. Genetic algorithms are search algorithms based on the mechanics of natural genetics.

In a genetic algorithm, a population of candidate solutions (called individuals or phenotypes) to an optimization problem is evolved towards better solutions. Each candidate solution has a set of properties (its chromosomes or genotype) which can be mutated and altered; traditionally, solutions are represented in binary as strings of 0s and 1s, but other encodings are also possible [5]. The evolution usually starts from a population of randomly generated individuals and happens in generations. In each generation, the fitness of every individual in the population is evaluated, the more fit individuals are stochastically selected from the current population, and each individual's genome is modified (recombined and possibly randomly mutated) to form a new population. The new population is then used in the next iteration of the algorithm. Commonly, the algorithm terminates when either a maximum number of generations has been produced, or a satisfactory fitness level has been reached for the population.

In the use of genetic algorithms to solve the practical problems, the premature phenomenon often appears, which limits the search performance of genetic algorithm. The reason for causing premature phenomenon is that highly similar exists between individuals after some generations, and the opportunity of the generation of new individuals by further genetic manipulation has been reduced greatly. Many ideas have been proposed to avoid the premature phenomenon [6, 15, 16, 18]. [18] introduces two special individuals at each generation of genetic algorithm to make the population maintains diversity in the problem of finding the minimized distance between surfaces, and the computational efficiency has been improved. We introduces other two special individuals at each generation of genetic algorithm in the same problem in [15], and the computational efficiency has been further improved. Furthermore, we suggest to put some symmetric and harmonious individuals at each generation of genetic algorithm applied to a general optimal problem in [16], and good computational efficiency has been obtained. Some application of our methods in reservoir mid-ong hydraulic power operation has been given in [17].

Recently many authors use GA to solve nonlinear equations system, see [9, 3, 11, 10, 14]. These works give us impression that GA are effective methods for solving nonlinear equations system. However, efforts are still needed so as to solve nonlinear equations system more effectively. In this chapter, we present a new genetic algorithm to solve Eq. (14.1.1).

The chapter is organized as follows: We convert the equation problem (14.1.1) to an optimal problem in Section 14.2, in Section 14.3 we present our new genetic algorithm with symmetric and harmonious individuals for the corresponding optimal problem, in Section 14.4 we give a mixed method by our method with Newton's method, whereas in Section 14.5 we provide some numerical examples to show that our new methods are very effective. Some remarks and conclusions are given in the concluding section 14.6.

14.2. Convert (14.1.1) to an Optimal Problem

Let us define function $F : D = [a_1,b_1]\times[a_2,b_2]\cdots\times[a_n,b_n]\subseteq R^n \to R$ as follows

$$F(x_1,x_2,\ldots,x_n) = \frac{1}{n}\sum_{i=1}^{n}|f_i(x_1,x_2,\ldots,x_n)|. \tag{14.2.1}$$

Then, we convert problem (14.1.1) to the following optimal problem

$$\begin{aligned}&\min\ F(x_1,x_2,\ldots,x_n)\\ &s.t\ \ (x_1,x_2,\ldots,x_n)\in D.\end{aligned} \tag{14.2.2}$$

Suppose $x^\star \in D$ is a solution of Eq. (14.1.1), then we have $F(x^\star) = 0$ from the definition (14.2.1) of F. Since $F(x) \geq 0$ holds for all $x \in D$, we deduce that $x^\star$ is the solution of problem (14.2.2). On the other hand, assume $x^\star \in D$ is a solution of problem (14.2.2). Then for all $x \in D$, we have $F(x) \geq F(x^\star)$. Now we suppose (14.1.1) has at least one solution denoted as $y^\star$. Then, we have $0 \leq F(x^\star) \leq F(y^\star) = 0$, that is $F(x^\star) = 0$ is true and $x^\star$ is a solution of Eq. (14.1.1). Hence, Eq. (14.1.1) is equivalent to problem (14.2.2) if Eq. (14.1.1) has at least one solution.

From now on, we always suppose (14.1.1) has at least one solution, and try to find its solution by finding a solution of (14.2.2) via a genetic algorithm.

14.3. New Genetic Algorithm: SHEGA

Since the simple genetic algorithm (called standard genetic algorithm) is not very efficient in practical computation, many varieties have been given [6]. In this chapter, we give a new genetic algorithm. Our main idea is to put pairs of symmetric and harmonious individuals in generations.

14.3.1. Coding

We use the binary code in our method, as used in the simple genetic algorithm [6]. The binary code is the most used code in GA, which represents the candidate solution by a string of 0s and 1s. The length of the string is relation to the degree of accuracy needed for the solution, and satisfies the following inequality

$$L_i \geq \log_2 \frac{b_i - a_i}{\varepsilon_i}, \tag{14.3.1}$$

where, L_i is the length of the string standing for the i-component of an individual, and ε_i is the degree of accuracy needed for x_i. In fact, we should choose the minimal positive integer to satisfy (14.3.1) for L_i. Usually, all ε_i are equal to one another, so it can be denoted by ε_x in this status. For example, let $a_1 = 0$, $b_1 = 1$ and $\varepsilon_x = 10^{-6}$, then we can choose $L_1 = 20$, since $L_1 \geq \log_2 10^6 \approx 19.93156857$. Variable $x_1 \in [a_1, b_1]$ is in the form of real numbers, it can be represented by 20 digits of the binary code:00000000000000000000 stands for 0, 00000000000000000001 stands for $\frac{1}{2^{L_i}}$, 00000000000000000010 stands for $\frac{2}{2^{L_i}}$,..., 11111111111111111111 stands for 1.

14.3.2. Fitness Function

Since the simple genetic algorithm is used to find a solution of a maximal problem, one should make some change for the fitness function. In this chapter, we define the fitness function as follows

$$g(x_1, x_2, \ldots, x_n) = \frac{1}{1 + F(x_1, x_2, \ldots, x_n)}. \tag{14.3.2}$$

For this function g, it satisfies: (1) The fitness value is always positive, which is needed in the following genetic operators; (2) The fitness value will be bigger if the point $(x_1, x_2, \ldots, x_n)$ is closer to a solution $x^\star$ of problem (14.2.2).

14.3.3. Selection

We use Roulette Wheel Section (called proportional election operator) in our method, as used in the simple genetic algorithm. The probability of individual is selected and the fitness function is proportional to the value. Suppose the size of population is N, the fitness of individual i is g_i. The individual i was selected to the next generation with probability p_i given by

$$p_i = \frac{g_i}{\sum_{i=1}^{N} g_i} \, (i = 1, 2, \cdots, N). \tag{14.3.3}$$

14.3.4. Symmetric and Harmonious Individuals

We first give the definition of symmetric and harmonious individuals.

Definition 14.3.1. *Suppose individuals* $M_1 = (x_1, x_2, \ldots, x_n)$ *and* $M_2 = (y_1, y_2, \ldots, y_n)$ *can be represented in the binary code by* $M_1' = (x_{1_1}x_{1_2}\ldots x_{1_{L_1}}x_{2_1}x_{2_2}\ldots x_{2_{L_2}}\ldots x_{n_1}x_{n_2}$ $\ldots x_{n_{L_n}})$ *and* $M_2' = (y_{1_1}y_{1_2}\ldots y_{1_{L_1}}y_{2_1}y_{2_2}\ldots y_{2_{L_2}}\ldots y_{n_1}y_{n_2}\ldots y_{n_{L_n}})$*, respectively. They are called symmetric and harmonious individuals if and only if* $x_{i_j} = 1 - y_{i_j}$ *holds for any* $i = 1, 2, \ldots, n$ *and* $j = 1, 2, \ldots, L_i$*.*

In Definition 14.3.1 individuals M_1' and M_2' are complements in the binary sense.

In order to avoid the premature phenomenon of genetic algorithm, we introduce some pair of symmetric and harmonious individuals in generation. We don't use the fixed symmetric and harmonious individuals as in [15] and [18]. Contrarily, we generate pair of symmetric and harmonious individuals randomly. On one hand, these pair of symmetry of individuals continue to enrich the diversity of the population. On the other hand, they continue to explore the space even if they haven't been selected to participate in genetic manipulation of exchange or mutation.

Suppose the size of population is N, and $\lambda \in [0, 0.5)$ is a parameter. Let $\lfloor r \rfloor$ be the biggest integer equal to or less than r. We introduce $\lfloor \lambda * N \rfloor$ pairs of symmetric and harmonious individuals in current generation provided that the quantity between the best fitness value of the last generation to the one of the current generation is less than a preset precision denoted by ε_1. Here, we call λ as symmetry and harmonious factor.

14.3.5. Crossover and Mutation

We use one-point crossover operator in our genetic algorithm, just as used in the simple genetic algorithm. That is, a single crossover point on both parents' organism strings is selected. All data beyond that point in either organism string is swapped between the two parent organisms. The resulting organisms are the children. An example is shown as follows:

$$\begin{array}{l} A: 10110111 \,|\, 001 \\ B: 00011100 \,|\, 110 \end{array} \Rightarrow \begin{array}{l} A': 10110111 \,|\, 110 \\ B': 00011100 \,|\, 001. \end{array}$$

We use bit string mutation, just as used in the simple genetic algorithm. That is, the mutation of bit strings ensue through bit flips at random positions. The following example is provided to show this:

$$A: 1010\,1\,0101010 \;\Rightarrow\; A': 1010\,0\,0101010.$$

14.3.6. Elitist Model

It is well-known that the global convergence of the simple genetic algorithm cannot be assured [6]. In order to guarantee the global convergence of the genetic algorithm, we use an elitist model (or the optimal preservation strategy) in this chapter. That is, at each generation we reserves the individual with the maximum fitness value to the next generation.

14.3.7. Procedure

For the convenience of discussion, we call our genetic algorithm with symmetric and harmonious individuals as SHEGA, and call the simple genetic algorithm with the elitist model as EGA. We can give the procedure of SHEGA as follows:

Step 1. Assignment of parameters of genetic algorithm: The size N of population, the number n of variables of (14.2.2), the lengthes $L_1, L_2, \ldots, L_n$ (computed from (14.3.1)) of the binary string of the components of an individual, symmetry and harmonious factor λ, controlled precision ε_1 in subsection 14.3.4 to introduce the symmetric and harmonious individuals, the probability p_c of the crossover operator, the probability p_m of the mutation operator, and the largest genetic generation G.

Step 2. Generate the initial population randomly.

Step 3. Calculate the fitness value of each individual of the contemporary population, and reserve the optimal individual of the contemporary population to the next generation.

Step 4. If the distance between the best fitness value of the last generation to that of the current generation is less than a preset precision ε_1, we generate $N - 2 * \lfloor \lambda * N \rfloor - 1$ individuals using Roulette Wheel Section and $\lfloor \lambda * N \rfloor$ pairs of symmetry and harmonious individuals randomly. Otherwise we generate $N - 1$ individuals using Roulette Wheel Section directly. The population is then divided into two parts: one is the seed subpopulation constituted by symmetry and harmonious individuals, and the other is a subpopulation ready to be bred and constituted by the residual individuals.

Step 5. Take the crossover operator between each individual in the seed subpopulation to one individual selected from the other subpopulation randomly. Take the crossover operator each other using two two paired method with probability p_c and take the mutation operator with probability p_m for each residual individual in the subpopulation ready to be bred.

Step 6. Repeat Step 3-Step 5 until the maximum genetic generation G is reached.

14.4. Mixed Algorithm: SHEGA-Newton Method

In order to improve further the efficiency of SHEGA, we can apply it by mixed with a classical iterative method such as Newton's method [2, 13]

$$x^{(k+1)} = x^{(k)} - f'(x^{(k)})^{-1} f(x^{(k)}) \quad (k \geq 0) \quad (x^{(0)} \in D). \tag{14.4.1}$$

Here, $f'(x)$ denotes the Fréchet derivative of function f.

Local as well as semilocal convergence results for Newton method (14.4.1) under various assumptions have been given by many authors [1, 2, 4, 7, 12, 13]. It is well-known that Newton's method converges to the solution quadratically provided that some necessary conditions are satisfied. However, there are two deficiencies to limit the application of Newton's method. First, function f must be differentiable, which will not be satisfied in practical application. Second, a good initial point for beginning the iterative is key to ensure the convergence of the iterative sequence, but it is a difficult task to choose the initial point in advance. In fact, choosing good initial points to begin the corresponding iteration is the common question for all the classical iterative methods used to solve equation (14.1.1)[2, 13].

Here, we use Newton's method as an example. In fact, one can develop other methods by mixing SHEGA and other iterative methods. We can state SHEGA-Newton method as follows:

Step 1. Given the maximal iterative step S and the precision accuracy ε_y. Set $s = 1$.

Step 2. Find an initial guess $x^{(0)} \in D$ by using SHEGA given in Section 3.

Step 3. Compute $f_i(x_1^{(s)}, x_2^{(s)}, \ldots, x_n^{(s)})(i = 1, 2, \ldots, n)$. If $F(x_1^{(s)}, x_2^{(s)}, \ldots, x_n^{(s)}) \leq \varepsilon_y$, report that the approximation solution $x^{(s)} = (x_1^{(s)}, x_2^{(s)}, \ldots, x_n^{(s)})$ is found and exit from the circulation, where F is defined in (2.1).

Step 4. Compute the Jacobian $J_s = f'(x^{(s)})$ and solve the linear equations

$$J_s u^{(s)} = f(x^{(s)}). \tag{14.4.2}$$

Set $x^{(s+1)} = x^{(s)} + u^{(s)}$.

Step 5. If $s \leq S$, set $s = s + 1$ and goto Step 3. Otherwise, report that the approximation solution cannot be found.

14.5. Numerical Examples

In this section, we will provide some examples to show the efficiency of our new method.

Example 14.5.1. *Let f be defined in $D = [-3.5, 2.5] \times [-3.5, 2.5]$ by*

$$\begin{cases} f_1(x_1, x_2) = x_1^2 + x_2^2 + x_1 + x_2 - 8 = 0, \\ f_2(x_1, x_2) = x_1 x_2 + x_1 + x_2 - 5 = 0. \end{cases} \tag{14.5.1}$$

Let us choose parameters as follows:

$$N = 40, \quad p_c = 0.9, \quad p_m = 0.005, \quad \varepsilon_x = 0.001, \quad \varepsilon_1 = 0.001. \tag{14.5.2}$$

Since the genetic algorithms are random algorithms, we run each method 30 *times, and compare convergence number of times under various maximal generation G for EGA and SHEGA. The comparison results are given in Table 14.5.1. We also give the comparison results of the average of the best function value F under the fixed maximal generation* $G = 300$ *in Table 14.5.2. Here, we say the corresponding genetic algorithm is convergent if the function value $F(x_1, x_2, \ldots, x_n)$ is less than a fixed precision ε_y. We set* $\varepsilon_y = 0.05$ *for this example. Tables 1 and 2 show us that SHEGA with proper symmetry and harmonious factor λ performs better than EGA.*

Table 14.5.1. The comparison results of convergence number of times for Example 1

	$G = 50$	$G = 100$	$G = 150$	$G = 200$	$G = 250$	$G = 300$
EGA	11	11	11	11	12	12
SHEGA($\lambda = 0.05$)	13	15	18	20	20	21
SHEGA($\lambda = 0.10$)	8	20	23	26	29	29
SHEGA($\lambda = 0.15$)	16	21	27	28	28	29
SHEGA($\lambda = 0.20$)	16	26	29	30	30	30
SHEGA($\lambda = 0.25$)	20	28	29	29	30	30
SHEGA($\lambda = 0.30$)	22	28	30	30	30	30
SHEGA($\lambda = 0.35$)	17	27	29	30	30	30
SHEGA($\lambda = 0.40$)	17	29	30	30	30	30
SHEGA($\lambda = 0.45$)	19	30	30	30	30	30

Table 14.5.2. The comparison results of the average of the best function value F for Example 1

	$G = 300$
EGA	0.129795093953972
SHEGA($\lambda = 0.05$)	0.053101422547384
SHEGA($\lambda = 0.10$)	0.041427669194903
SHEGA($\lambda = 0.15$)	0.034877317449825
SHEGA($\lambda = 0.20$)	0.035701604675096
SHEGA($\lambda = 0.25$)	0.038051665705034
SHEGA($\lambda = 0.30$)	0.039332883168632
SHEGA($\lambda = 0.35$)	0.035780879206619
SHEGA($\lambda = 0.40$)	0.034509424138501
SHEGA($\lambda = 0.45$)	0.037425326021257

Example 14.5.2. *Let f be defined in $D = [-5,5] \times [-1,3] \times [-5,5]$ by*

$$\begin{cases} f_1(x_1,x_2,x_3) = 3x_1^2 + \sin(x_1x_2) - x_3^2 + 2 = 0, \\ f_2(x_1,x_2,x_3) = 2x_1^3 - x_2^2 - x_3 + 3 = 0, \\ f_3(x_1,x_2,x_3) = \sin(2x_1) + \cos(x_2x_3) + x_2 - 1 = 0. \end{cases} \tag{14.5.3}$$

Let us choose parameters as follows:

$$N = 50, \quad p_c = 0.8, \quad p_m = 0.05, \quad \varepsilon_x = 0.0001, \quad \varepsilon_1 = 0.001. \tag{14.5.4}$$

We run each method 20 times, and compare convergence number of times under various maximal generation G for EGA and SHEGA. The comparison results are given in Table 14.5.3. We also give the comparison results of the average of the best function value F under *the fixed maximal generation $G = 500$ in Table 14.5.4. Here, we say the corresponding genetic algorithm is convergent if the function value $F(x_1, x_2, \ldots, x_n)$ is less than a fixed precision ε_y. We set $\varepsilon_y = 0.02$ for this example. Tables 14.5.3 and 14.5.4 show us that SHEGA with proper symmetry and harmonious factor λ performs better than EGA.*

Table 14.5.3. The comparison results of convergence number of times for Example 2

	$G = 100$	$G = 200$	$G = 300$	$G = 400$	$G = 500$
EGA	3	4	4	4	4
SHEGA($\lambda = 0.05$)	2	6	8	9	11
SHEGA($\lambda = 0.10$)	2	7	11	14	15
SHEGA($\lambda = 0.15$)	5	11	14	14	16
SHEGA($\lambda = 0.20$)	3	11	14	17	18
SHEGA($\lambda = 0.25$)	8	11	14	17	17
SHEGA($\lambda = 0.30$)	8	16	16	17	17
SHEGA($\lambda = 0.35$)	6	14	18	19	20
SHEGA($\lambda = 0.40$)	5	14	19	19	20
SHEGA($\lambda = 0.45$)	6	17	19	19	20

Table 14.5.4. The comparison results of the average of the best function value F for Example 2

	$G = 500$
EGA	0.1668487275323187
SHEGA($\lambda = 0.05$)	0.0752619855962109
SHEGA($\lambda = 0.10$)	0.0500864062405815
SHEGA($\lambda = 0.15$)	0.0358268275921585
SHEGA($\lambda = 0.20$)	0.0257859494269335
SHEGA($\lambda = 0.25$)	0.0239622084932336
SHEGA($\lambda = 0.30$)	0.0247106452514721
SHEGA($\lambda = 0.35$)	0.0171980128114993
SHEGA($\lambda = 0.40$)	0.0179659124369376
SHEGA($\lambda = 0.45$)	0.0158999282064303

Next, we will provide an example when f is non-differentiable. Newton's method (14.4.1) and SHEGA-Newton cannot be used to solve this problem since f is non-differentiable. However, SHEGA can apply. Moreover, with the same idea used for SHEGA-Newton method, one can develop some mixed methods which don't require the differentiability of function f to solve the problem.

Example 14.5.3. *Let f be defined in $D = [-2,2] \times [1,6]$ by*

$$\begin{cases} f_1(x_1,x_2) = x_1^2 - x_2 + 1 + \frac{1}{9}|x_1 - 1| = 0, \\ f_2(x_1,x_2) = x_2^2 + x_1 - 7 + \frac{1}{9}|x_2| = 0. \end{cases} \tag{14.5.5}$$

Let us choose parameters as follows:

$$N = 40, \quad p_c = 0.85, \quad p_m = 0.008, \quad \varepsilon_x = 0.001, \quad \varepsilon_1 = 0.005. \tag{14.5.6}$$

We run each method 20 times, and compare convergence number of times under various maximal generation G for EGA and SHEGA. The comparison results are given in Table 14.5.5. We also give the comparison results of the average of the best function value F under the fixed maximal generation $G = 300$ in Table 14.5.6. Here, we say the corresponding genetic algorithm is convergent if the function value $F(x_1,x_2,\ldots,x_n)$ is less than a fixed precision ε_y. We set $\varepsilon_y = 0.003$ for this example. Tables 14.5.5 and 14.5.6 show us that SHEGA with proper symmetry and harmonious factor λ performs better than EGA.

Table 14.5.5. The comparison results of convergence number of times for Example 3

	$G = 100$	$G = 150$	$G = 200$	$G = 250$	$G = 300$
EGA	6	9	10	10	10
SHEGA($\lambda = 0.05$)	10	15	16	17	19
SHEGA($\lambda = 0.10$)	12	15	17	18	20
SHEGA($\lambda = 0.15$)	15	18	19	20	20
SHEGA($\lambda = 0.20$)	15	20	20	20	20
SHEGA($\lambda = 0.25$)	11	14	17	17	18
SHEGA($\lambda = 0.30$)	10	14	16	17	19
SHEGA($\lambda = 0.35$)	16	17	18	19	20
SHEGA($\lambda = 0.40$)	10	12	18	18	20
SHEGA($\lambda = 0.45$)	9	12	15	15	17

14.6. Conclusion

We presented a genetic algorithm as a general tool for solving optimum problems. Note that in the special case of approximating solutions of systems of nonlinear equations there are many deficiencies that limit the application of the usually employed methods. For example in the case of Newton's method function f must be differentiable and a good initial point must be found. To avoid these problems we have introduced some pairs of symmetric and harmonious individuals for the generation of a genetic algorithm. The population diversity is preserved this way and the method guarantees convergence to a solution of the system. Numerical examples are illustrating the efficiency of the new algorithm.

Table 14.5.6. The comparison results of the average of the best function value F for Example 3

	$G = 300$
EGA	0.0081958187235953
SHEGA($\lambda = 0.05$)	0.0021219384775699
SHEGA($\lambda = 0.10$)	0.0019286950245614
SHEGA($\lambda = 0.15$)	0.0018367719544782
SHEGA($\lambda = 0.20$)	0.0022816080103967
SHEGA($\lambda = 0.25$)	0.0023297925904943
SHEGA($\lambda = 0.30$)	0.0023318357433983
SHEGA($\lambda = 0.35$)	0.0021392510790106
SHEGA($\lambda = 0.40$)	0.0022381534380744
SHEGA($\lambda = 0.45$)	0.0025798012930550

References

[1] Amat, S., Busquier, S., M. Negra, Adaptive approximation of nonlinear operators, *Numer. Funct. Anal. Optim.*, 25 (2004), 397-405.

[2] Argyros, I.K., Computational Theory of Iterative Methods, *Series: Studies in Computational Mathematics* 15, Editors: C.K. Chui and L. Wuytack, Elsevier Publ. Co., New York, U.S.A., 2007.

[3] El-Emary, I.M., El-Kareem, M.A., Towards using genetic algorithm for solving nonlinear equation systems, *World Applied Sciences Journal* 5 (2008), 282-289.

[4] Ezquerro, J.A., Hernández, M.A., On an application of Newton's method to nonlinear equations with $w-$ condition second derivative, *BIT* 42 (2002), 519-532.

[5] Fogel, D.B., *Evolutionary Computation: Toward a New Philosophy of Machine Intelligence*, IEEE Press, New York, 2000.

[6] Goldberg, D.E., *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley, Second Edition, 1989.

[7] Gutiérrez, J.M., A new semilocal convergence theorem for Newton's method, *J. Comput. Appl. Math.* 79 (1997), 131-145.

[8] Holland, J.H., Adaptation in Natural and Artificial System, Michigan Univ Press, Ann Arbor, 1975.

[9] Kuri-Morales, A. F., No, R. H., Solution of simultaneous non-linear equations using genetic algorithms, *WSEAS Transactions on SYSTEMS*, 2 (2003), 44-51.

[10] Mastorakis, N.E., Solving non-linear equations via genetic algorithms, Proceedings of the 6th WSEAS Int. Conf. on Evolutionary Computing, Lisbon, Portugal, June 16-18, 2005, 24-28.

[11] Nasira, G.N., Devi, D., Solving nonlinear equations through jacobian sparsity pattern using genetic algorithm, *International Journal of Communications and Engineering*, 5 (2012), 78-82.

[12] Neta, B., A new iterative method for the solution of system of nonlinear equations, *Approx. Th. and Applic* (Z. Ziegler, ed.), Academic Press, 249-263 1981.

[13] Ortega, J.M., Rheinbolt, W.C., *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970.

[14] Mhetre, P.H., Genetic algorithm for linear and nonlinear equation, *International Journal of Advanced Engineering Technology*, 3 (2012), 114-118.

[15] Ren, H., Bi, W., Wu, Q., Calculation of minimum distance between free-form surfaces by a type of new improved genetic algorithm, *Computer Engineering and Application*, 23 (2004), 62-64 (in Chinese).

[16] Ren, H., Wu, Q., Bi, W., A genetic algorithm with symmetric and harmonious individuals, *Computer Engineering and Application*, 5(87) (2005) 24-26, 87 (in Chinese).

[17] Wan, X, Zhou, J., Application of genetic algorithm for self adaption, symmetry and congruity in reservoir mid-ong hydraulic power operation, *Advances in Water Science*, 18 (2007), 598-603 (in Chinese).

[18] Xi, G., Cai, Y., Finding the minimized distance between surfaces, *Journal of CAD and Graphics*, 14 (2002), 209–213 (in Chinese).

Chapter 15

On the Semilocal Convergence of Modified Newton-Tikhonov Regularization Method for Nonlinear Ill-Posed Problems

15.1. Introduction

In this chapter we are concerned with the problem of approximately solving the nonlinear ill-posed operator equation

$$F(x) = f, \tag{15.1.1}$$

where $F : D(F) \subseteq X \to Y$ is a nonlinear operator between the Hilbert spaces X and Y. Here and below $\langle .,. \rangle$ denote the inner product and $\|.\|$ denote the corresponding norm. We assume throughout that $f^\delta \in Y$ are the available data with

$$\|f - f^\delta\| \leq \delta$$

and (15.1.1) has a solution $\hat{x}$ (which need not be unique). Then the problem of recovery of $\hat{x}$ from noisy equation $F(x) = f^\delta$ is ill-posed, in the sense that a small perturbation in the data can cause large deviation in the solution.

Further it is assumed that F possesses a locally uniformly bounded Fréchet derivative $F'(.)$ in the domain $D(F)$ of F. A large number of problems in mathematical physics and engineering are solved by finding the solutions of equations in a form like (15.1.1). If one works with such problems, the measurement data will be distorted by some measurement error. Therefore, one has to consider appropriate regularization techniques for approximately solving (15.1.1).

Iterative regularization methods are used for approximately solving (15.1.1). Recall ([20]) that, an iterative method with iterations defined by

$$x_{k+1}^\delta = \Phi(x_0^\delta, x_1^\delta, \cdots, x_k^\delta; y^\delta),$$

where $x_0^\delta := x_0 \in D(F)$ is a known initial approximation of $\hat{x}$, for a known function Φ together with a stopping rule which determines a stopping index $k_\delta \in \mathbb{N}$ is called an iterative regularization method if $\|x_{k_\delta}^\delta - \hat{x}\| \to 0$ as $\delta \to 0$.

The Levenberg-Marquardt method([18], [21], [9], [10], [11], [14], [24], [6]) and iteratively regularized Gauss-Newton method (IRGNA) ([3], [5]) are the well-known iterative regularization methods. In Levenberg-Marquardt method, the iterations are defined by,

$$x_{k+1}^\delta = x_k^\delta - (A_{k,\delta}^* A_{k,\delta} + \alpha_k I)^{-1} A_{k,\delta}^* (F(x_k^\delta) - y^\delta), \tag{15.1.2}$$

where $A_{k,\delta}^* := F'(x_k^\delta)^*$ is as usual the adjoint of $A_{k,\delta} := F'(x_k^\delta)$ and (α_k) is a positive sequence of regularization parameter ([5]). In Gauss-Newton method, the iterations are defined by

$$x_{k+1}^\delta = x_k^\delta - (A_{k,\delta}^* A_{k,\delta} + \alpha_k I)^{-1} [A_{k,\delta}^* (F(x_k^\delta) - y^\delta) + \alpha_k (x_k^\delta - x_0)] \tag{15.1.3}$$

where $x_0^\delta := x_0$ and (α_k) is as in (15.1.2).

In [3], Bakushinskii obtained local convergence of the method (15.1.3), under the smoothness assumption

$$\hat{x} - x_0 = (F'(\hat{x})^* F'(\hat{x}))^\nu w, \qquad w \in N(F'(\hat{x}))^\perp \tag{15.1.4}$$

with $\nu \geq 1, w \neq 0$ and $F'(.)$ is Lipschitz continuous; $N(F'(\hat{x}))$ denotes the nullspace of $F'(\hat{x})$. For noise free case Bakushinskii ([3]) obtained the rate

$$\|x_k^\delta - \hat{x}\| = O(\alpha_k),$$

and Blaschke et.al.([5]) obtained the rate

$$\|x_k^\delta - \hat{x}\| = O(\alpha_k^\nu), \tag{15.1.5}$$

for $\frac{1}{2} \leq \nu < 1$.

It is proved in [5], that the rate (15.1.5) can be obtained for $0 \leq \nu < \frac{1}{2}$ provided $F'(.)$ satifies the following conditions:

$$F'(\bar{x}) = R(\bar{x}, x) F'(x) + Q(\bar{x}, x)$$

$$\|I - R(\bar{x}, x)\| \leq C_R \qquad \bar{x}, x \in B_{2\rho}(x_0)$$

$$\|Q(\bar{x}, x)\| \leq C_Q \|F'(\hat{x})(\bar{x} - x)\|$$

with ρ, C_R and C_Q sufficiently small. In fact in [5], Blaschke et.al. obtained the rate

$$\|x_k^\delta - \hat{x}\| = o(\alpha_k^{\frac{2\nu}{2\nu+1}}), \qquad 0 \leq \nu < \frac{1}{2}$$

by choosing the stopping index k_δ according to the discrepancy principle

$$\|F(x_{k_\delta}^\delta) - y^\delta\| \leq \tau\delta < \|F(x_k^\delta) - y^\delta\|, \qquad 0 \leq k < k_\delta$$

with $\tau > 1$ chosen sufficiently large. Subsequently, many authors extended, modified, and generalized Bakushinskii's work to obtain error bounds under various contexts(see [4], [12], [13], [15], [16], [17], [7]).

In [20], Mahale and Nair considered a method in which the iterations are defined by

$$x_{k+1}^{\delta} = x_0 - g_{\alpha_k}(A_0^*A_0)A_0^*[F(x_k^{\delta}) - y^{\delta} - A_0(x_k^{\delta} - x_0)], \qquad x_0^{\delta} := x_0 \tag{15.1.6}$$

where $A_0 := F'(x_0), (\alpha_k)$ is a sequence of regularization parameters which satisfies,

$$\alpha_k > 0, \qquad 1 \leq \frac{\alpha_k}{\alpha_{k+1}} \leq \mu_1, \qquad \lim_{k\to 0} \alpha_k = 0 \tag{15.1.7}$$

for some constant $\mu_1 > 1$ and each g_α, for $\alpha > 0$ is a positive real-valued piecewise continuous function defined on $[0, M]$ with $M \geq \|A_0\|^2$. They choose the stopping index k_δ for this iteration as the positive integer which satisfies

$$\max\{\|F(x_{k_\delta-1}^{\delta}) - y^{\delta}\|, \tilde{\beta}_{k_\delta}\} \leq \tau\delta < max\{\|F(x_{k-1}^{\delta}) - y^{\delta}\|, \tilde{\beta}_k\} \qquad 1 \leq k < k_\delta$$

where $\tau > 1$ is a sufficiently large constant not depending on δ, and

$$\tilde{\beta}_k := \|F(x_{k-1}^{\delta}) - y^{\delta} + A_0(x_k^{\delta} - x_{k-1}^{\delta})\|.$$

In fact, Mahle and Nair obtained an order optimal error estimate, in the sense that an improved order estimate which is applicable for the case of linear ill-posed problems as well is not possible, under the following new source condition on $x_0 - \hat{x}$.

Assumption 15.1.1. *There exists a continuous, stricly monotonically increasing function* $\varphi : (0, M] \to (0, \infty)$ *satisfying* $\lim_{\lambda\to 0} \varphi(\lambda) = 0$ *and* $\rho_0 > 0$ *such that*

$$x_0 - \hat{x} = [\varphi(A_0^*A_0)]^{1/2} w \tag{15.1.8}$$

for some $w \in X$ *with* $\|w\| \leq \rho_0$.

In [7], the author considered a particular case of this method, namely, regularized modified Newton's method defined iteratively by

$$x_{k+1}^{\delta} = x_k^{\delta} - (A_0^*A_0 + \alpha I)^{-1}[A_0^*(F(x_k^{\delta}) - y^{\delta}) + \alpha(x_k^{\delta} - x_0)], \qquad x_0^{\delta} := x_0 \tag{15.1.9}$$

for approximately solving (15.1.1). Using a suitably constructed majorizing sequence (see, [1], page 28), it is proved that the sequence(x_k^{δ}) converges linearly to a solution x_α^{δ} of the equation

$$A_0^*F(x_\alpha^{\delta}) + \alpha(x_\alpha^{\delta} - x_0) = A_0^*y^{\delta} \tag{15.1.10}$$

and that x_α^{δ} is an approximation of $\hat{x}$. The error estimate in this chapter was obtained under the following source condition on $x_0 - \hat{x}$

Assumption 15.1.2. *There exists a continuous, stricly monotonically increasing function* $\varphi : (0, a_1] \to (0, \infty)$ *with* $a_1 \geq \|F'(\hat{x})\|^2$ *satisfying*

1. $\lim_{\lambda\to 0} \varphi(\lambda) = 0$

2. *for* $\alpha \leq 1, \varphi(\alpha) \geq \alpha$

3. $\sup_{\lambda\geq 0} \frac{\alpha\varphi(\lambda)}{\lambda+\alpha} \leq c_\varphi \varphi(\alpha), \qquad \forall \lambda \in (0, a_1]$

4. *there exists $w \in X$ such that*

$$x_0 - \hat{x} = \varphi(F'(\hat{x})^* F'(\hat{x}))w. \tag{15.1.11}$$

Later in [8], using a two step Newton method (see, [2]), the author proved that the sequence (x_k^δ) in (15.1.9) converges linarly to the solution x_α^δ of (15.1.10). The error estimate in [8] was based on the following source condition

$$x_0 - \hat{x} = \varphi(A_0^* A_0)w,$$

where φ is as in Assumption 15.1.1 with $a_1 \geq \|A_0\|^2$. In the present chapter we improve the semilocal convergence by modifying the method (15.1.9).

15.1.1. The New Method

In this chapter we define a new iteration procedure

$$x_{n+1,\alpha}^\delta = x_{n,\alpha}^\delta - (A_0^* A_n + \alpha I)^{-1}[A_0^*(F(x_{n,\alpha}^\delta) - y^\delta) + \alpha(x_{n,\alpha}^\delta - x_0)], \qquad x_{0,\alpha}^\delta := x_0 \tag{15.1.12}$$

where $A_n := F'(x_{n,\alpha}^\delta)$ and $\alpha > 0$ is the regularization parameter. Using an assumption on the Fréchet derivative of F we prove that the iteration in (15.1.12) converges quadratically to the solution x_α^δ of (15.1.10).

Recall ([22]) that, a sequence (x_n) is said to converge quadratically to x^* if there exists positive reals β, γ such that

$$\|x_{n+1} - x^*\| \leq \beta e^{-\gamma 2^n}$$

for all $n \in \mathbb{N}$. And the convergence of (x_n) to x^* is said to be linear if there exists a positive number $M_0 \in (0,1)$, such that

$$\|x_{n+1} - x^*\| \leq M_0 \|x_n - x^*\|.$$

Quadratically convergent sequence will always eventually converge faster than a linear convergent sequence.

We choose the regularization parameter α from some finite set

$$\{\alpha_0 < \alpha_1 < \cdots < \alpha_N\}$$

using the balancing principle considered by Perverzev and Schock in [23].

The rest of this chapter is organized in the following way. In Section 15.2 we provide the convergence analysis of the proposed method and in Section 15.3 we provide the error analysis. Finally in Section 15.4 we provide the details for implementing the method and the algorithm.

15.2. Convergence Analysis of (15.1.12)

The following assumption is used extensively for proving the results in this chapter.

Assumption 15.2.1. *There exists a constant* $k_0 > 0, r > 0$ *such that for every* $x, u \in B(x_0, r) \cup B(\hat{x}, r) \subset D(F)$ *and* $v \in X$, *there exists an element* $\Phi(x, u, v) \in X$ *such that*

$$[F'(x) - F'(u)]v = F'(u)\Phi(x, u, v), \quad \|\Phi(x, u, v)\| \leq k_0\|v\|\|x - u\|.$$

In view of Assumption 15.2.1 there exists an element $\Phi_0(x, x_0, v) \in X$ such that

$$[F'(x) - F'(x_0)]v = F'(x_0)\Phi_0(x, x_0, v), \quad \|\Phi_0(x, x_0, v)\| \leq l_0\|v\|\|x - x_0\|.$$

Note that

$$l_0 \leq k_0$$

holds in general and $\frac{k_0}{l_0}$ can be arbitrarily large [1], [2]. Let $\delta_0 < \sqrt{\alpha_0}$,

$$\rho < \frac{\sqrt{1 + 2l_0(1 - \frac{\delta_0}{\sqrt{\alpha_0}})} - 1}{l_0},$$

and

$$\gamma_\rho := \frac{l_0}{2}\rho^2 + \rho + \frac{\delta_0}{\sqrt{\alpha_0}}.$$

For $r \leq \frac{2-3k_0}{(2+3l_0)k_0}, k_0 \leq \frac{2}{3}$ let $g : (0,1) \to (0,1)$ be the function defined by

$$g(t) := \frac{3(1 + l_0 r)k_0}{2(1 - l_0 r)}t \qquad \forall t \in (0,1).$$

Lemma 15.2.2. *Let* $l_0 r < 1$ *and* $u \in B_r(x_0)$. *Then* $(A_0^* A_u + \alpha I)$ *is invertible:*

(i)

$$(A_0^* A_u + \alpha I)^{-1} = [I + (A_0^* A_0 + \alpha I)^{-1} A_0^*(A_u - A_0)]^{-1}(A_0^* A_0 + \alpha I)^{-1}$$

and

(ii)

$$\|(A_0^* A_u + \alpha I)^{-1} A_0^* A_u\| \leq \frac{1 + l_0 r}{1 - l_0 r},$$

where $A_u := F'(u)$.

Proof. Note that by Assumption 15.2.1, we have

$$\begin{aligned}
\|(A_0^* A_0 + \alpha I)^{-1} A_0^*(A_u - A_0)\| &= \sup_{\|v\| \leq 1} \|(A_0^* A_0 + \alpha I)^{-1} A_0^*(A_u - A_0)v\| \\
&= \sup_{\|v\| \leq 1} \|(A_0^* A_0 + \alpha I)^{-1} A_0^* A_0 \Phi_0(u, x_0, v)\| \\
&\leq l_0\|u - x_0\| \leq l_0 r < 1.
\end{aligned}$$

So $I + (A_0^* A_0 + \alpha I)^{-1} A_0^*(A_u - A_0)$ is invertible. Now (i) follows from the following relation

$$A_0^* A_u + \alpha I = (A_0^* A_0 + \alpha I)[I + (A_0^* A_0 + \alpha I)^{-1} A_0^*(A_u - A_0)].$$

To prove (ii), observe that by Assumption 15.2.1 and (i), we have

$$\begin{aligned}
\|(A_0^*A_u+\alpha I)^{-1}A_0^*A_u\| &= \sup_{\|v\|\leq 1}\|(A_0^*A_u+\alpha I)^{-1}A_0^*A_u v\| \\
&= \sup_{\|v\|\leq 1}\|(A_0^*A_u+\alpha I)^{-1}A_0^*(A_u-A_0+A_0)v\| \\
&= \sup_{\|v\|\leq 1}\|[I+(A_0^*A_0+\alpha I)^{-1}A_0^*(A_u-A_0)]^{-1} \\
&\quad (A_0^*A_0+\alpha I)^{-1}A_0^*(A_u-A_0+A_0)v\| \\
&\leq \frac{1}{1-k_0 r}[\|(A_0^*A_0+\alpha I)^{-1}A_0^*A_0\Phi_0(u,x_0,v)\| \\
&\quad +\|(A_0^*A_0+\alpha I)^{-1}A_0^*A_0 v\|] \\
&\leq \frac{1+l_0 r}{1-l_0 r}.
\end{aligned}$$

This completes the proof.

Theorem 15.2.3. *Suppose Assumption 15.2.1 holds. Let* $\frac{\gamma_\rho}{1-g(\gamma_\rho)}\leq r\leq \frac{2-3k_0}{(2+3l_0)k_0}$, $\delta\in(0,\delta_0]$. *Then the sequence* $(x_{n,\alpha}^\delta)$ *defined in (15.1.12) is a Cauchy sequence in* $B_r(x_0)$ *and hence converges to* $x_\alpha^\delta\in\overline{B_r(x_0)}$. *Further* x_α^δ *satisfies (15.1.10) and the following estimate holds for all* $n\geq 0$:

$$\|x_{n,\alpha}^\delta - x_\alpha^\delta\|\leq re^{-\gamma 2^n} \tag{15.2.1}$$

where $\gamma=-ln(g(\gamma_\rho))$.

Proof. Suppose $x_{n,\alpha}^\delta\in B_r(x_0),\forall n\geq 0$. Then

$$\begin{aligned}
x_{n+1,\alpha}^\delta - x_{n,\alpha}^\delta &= (A_0^*A_n+\alpha I)^{-1}[A_0^*A_n(x_{n,\alpha}^\delta - x_{n-1,\alpha}^\delta) \qquad (15.2.2)\\
&\quad -A_0^*(F(x_{n,\alpha}^\delta)-F(x_{n-1,\alpha}^\delta))]+(A_0^*A_n+\alpha I)^{-1}A_0^*(A_n-A_{n-1}) \\
&\quad (A_0^*A_{n-1}+\alpha I)^{-1}[A_0^*(F(x_{n-1,\alpha}^\delta)-y^\delta)+\alpha(x_{n-1,\alpha}^\delta - x_0)] \\
&= (A_0^*A_n+\alpha I)^{-1}A_0^*[A_n(x_{n,\alpha}^\delta - x_{n-1,\alpha}^\delta)-(F(x_{n,\alpha}^\delta)-F(x_{n-1,\alpha}^\delta))] \\
&\quad +(A_0^*A_n+\alpha I)^{-1}A_0^*(A_n-A_{n-1})(x_{n,\alpha}^\delta - x_{n-1,\alpha}^\delta) \\
&:= \zeta_1+\zeta_2 \qquad (15.2.3)
\end{aligned}$$

where $\zeta_1=(A_0^*A_n+\alpha I)^{-1}A_0^*[A_n(x_{n,\alpha}^\delta - x_{n-1,\alpha}^\delta)-(F(x_{n,\alpha}^\delta)-F(x_{n-1,\alpha}^\delta))]$ and $\zeta_2=(A_0^*A_n+\alpha I)^{-1}A_0^*(A_n-A_{n-1})(x_{n,\alpha}^\delta - x_{n-1,\alpha}^\delta)$. So by Fundamental Theorem of Integration, $\zeta_1 = (A_0^*A_n+\alpha I)^{-1}A_0^*[\int_0^1(A_n-F'(x_{n-1,\alpha}^\delta+t(x_{n,\alpha}^\delta - x_{n-1,\alpha}^\delta)dt](x_{n,\alpha}^\delta - x_{n-1,\alpha}^\delta)$ and hence by Assumption 15.2.1 and Lemma 15.2.2,

$$\begin{aligned}
\|\zeta_1\| &\leq \|(A_0^*A_n+\alpha I)^{-1}A_0^*A_n\int_0^1\Phi(x_{n-1,\alpha}^\delta+t(x_{n,\alpha}^\delta - x_{n-1,\alpha}^\delta),x_{n,\alpha}^\delta,x_{n-1,\alpha}^\delta - x_{n,\alpha}^\delta)dt\| \\
&\leq \frac{1+l_0 r}{1-l_0 r}\int_0^1\Phi(x_{n-1,\alpha}^\delta+t(x_{n,\alpha}^\delta - x_{n-1,\alpha}^\delta),x_{n,\alpha}^\delta,x_{n-1,\alpha}^\delta - x_{n,\alpha}^\delta)dt\| \\
&\leq \frac{(l_0 r+1)k_0}{2(1-l_0 r)}\|x_{n,\alpha}^\delta - x_{n-1,\alpha}^\delta\|^2. \qquad (15.2.4)
\end{aligned}$$

Similarly,

$$\begin{aligned}
\|\zeta_2\| &\leq \|(A_0^*A_n+\alpha I)^{-1}A_0^*(A_n-A_{n-1})(x_{n-1,\alpha}^\delta - x_{n,\alpha}^\delta)\| \\
&\leq \|(A_0^*A_n+\alpha I)^{-1}A_0^*A_n\Phi(x_{n,\alpha}^\delta, x_{n-1,\alpha}^\delta, x_{n-1,\alpha}^\delta - x_{n,\alpha}^\delta)\| \\
&\leq \frac{(1+l_0r)k_0}{1-l_0r}\|x_{n,\alpha}^\delta - x_{n-1,\alpha}^\delta\|^2. \qquad (15.2.5)
\end{aligned}$$

So by (15.2.3), (15.2.4)and (15.2.5), we have

$$\begin{aligned}
\|x_{n+1,\alpha}^\delta - x_{n,\alpha}^\delta\| &\leq \frac{3(1+l_0r)k_0)}{2(1-l_0r)}\|x_{n,\alpha}^\delta - x_{n-1,\alpha}^\delta\|^2 \\
&\leq g(e_n)e_n, \qquad (15.2.6)
\end{aligned}$$

where

$$e_n := \|x_{n,\alpha}^\delta - x_{n-1,\alpha}^\delta\|, n = 1,2,\cdots.$$

Now using induction we shall prove that $x_{n,\alpha}^\delta \in B_r(x_0)$. Note that

$$\begin{aligned}
e_1 &= \|x_{1,\alpha}^\delta - x_0\| \\
&= \|(A_0^*A_0+\alpha I)^{-1}A_0^*(F(x_0)-y^\delta)\| \\
&= \|(A_0^*A_0+\alpha I)^{-1}A_0^*(F(x_0)-F(\hat{x})-F'(x_0)(x_0-\hat{x}) \\
&\quad +F'(x_0)(x_0-\hat{x})+F(\hat{x})-y^\delta)\| \\
&\leq \|(A_0^*A_0+\alpha I)^{-1}A_0^*(\int_0^1[F'(\hat{x}+t(x_0-\hat{x}))-F'(x_0)](x_0-\hat{x})dt \\
&\quad +F'(x_0)(x_0-\hat{x})+F(\hat{x})-y^\delta)\| \\
&\leq \|(A_0^*A_0+\alpha I)^{-1}A_0^*A_0(\int_0^1\Phi(x_0,\hat{x}+t(x_0-\hat{x}),x_0-\hat{x})\| \\
&\quad +\|(A_0^*A_0+\alpha I)^{-1}A_0^*F'(x_0)(x_0-\hat{x})\| \\
&\quad +\|(A_0^*A_0+\alpha I)^{-1}A_0^*(F(\hat{x})-y^\delta)\| \\
&\leq \frac{l_0}{2}\rho^2+\rho+\frac{\delta}{\sqrt{\alpha}} \leq \gamma_\rho \leq r \qquad (15.2.7)
\end{aligned}$$

i.e.,$x_{1,\alpha}^\delta \in B_r(x_0)$.

Now since $\gamma_\rho < 1$, by (15.2.7), $e_1 < 1$. Therefore by (15.2.6) and the fact that $g(\mu t) \leq \mu g(t)$, for all $t \in (0,1)$, we have that $e_n < 1, \forall n \geq 1$ and

$$g(e_1)^{2^n-1}e_1.$$

Now suppose $x_{k,\alpha}^\delta \in B_r(x_0)$ for some k. Then

$$\begin{aligned}
\|x_{k+1,\alpha}^\delta - x_0\| &\leq \|x_{k+1,\alpha}^\delta - x_{k,\alpha}^\delta\| + \|x_{k,\alpha}^\delta - x_{k-1,\alpha}^\delta\| + \cdots + \|x_{1,\alpha}^\delta - x_0\| \\
&\leq (g(e_1)^{2^k-1}+g(e_1)^{2^{k-1}-1}+\cdots+1)e_1 \\
&\leq \frac{e_1}{1-g(e_1)} \leq \frac{\gamma_\rho}{1-g(\gamma_\rho)} \leq r.
\end{aligned}$$

Thus by induction $x_{n,\alpha}^{\delta} \in B_r(x_0), \forall n \geq 0$.

Next we shall prove that $(x_{k+1,\alpha}^{\delta})$ is a Cauchy sequence in $B_r(x_0)$.

$$\begin{aligned}
\|x_{n+m,\alpha}^{\delta} - x_{n,\alpha}^{\delta}\| &\leq \sum_{i=0}^{m} \|x_{n+i+1,\alpha}^{\delta} - x_{n+i,\alpha}^{\delta}\| \qquad (15.2.8)\\
&\leq \sum_{i=0}^{m} g(e_1)^{2^{n+i}-1} e_1 \\
&\leq g(e_1)^{2^n-1} e_1 (1 + g(e_1)^2 + \cdots + g(e_1)^{2^m}) \\
&\leq \frac{g(e_1)^{2^n-1} e_1}{1-g(e_1)} \leq \frac{g(\gamma_\rho)^{2^n-1}\gamma_\rho}{1-g(\gamma_\rho)} \leq re^{-\gamma 2^n}. \qquad (15.2.9)
\end{aligned}$$

Thus $(x_{n,\alpha}^{\delta})$ is a Cauchy sequence in $B_r(x_0)$ and hence converges, say to $x_\alpha^{\delta} \in \overline{B_r(x_0)}$. Further by letting $n \to \infty$ in (15.1.12) we obtain

$$F'(x_0)^*(F(x_\alpha^{\delta}) - y^{\delta}) + \alpha(x_\alpha^{\delta} - x_0) = 0.$$

The estimate in (15.2.1) follows by letting m tends to ∞ in (15.2.9).

Remark 15.2.4. *Note that if $r \in (r_1, r_2)$ where*

$$r_1 := \frac{2 + (2l_0 - 3k_0)\gamma_\rho - \sqrt{(4l_0^2 + 9k_0^2 - 36k_0l_0)\gamma_\rho^2 - (12k_0 + 8l_0)\gamma_\rho + 4}}{2l_0(2 + 3k_0\gamma_\rho)}$$

and

$$r_2 := \min\{\frac{2 + (2l_0 - 3k_0)\gamma_\rho + \sqrt{(4l_0^2 + 9k_0^2 - 36k_0l_0)\gamma_\rho^2 - (12k_0 + 8l_0)\gamma_\rho + 4}}{2l_0(2 + 3k_0\gamma_\rho)}, \frac{2 - 3k_0}{(2 + 3l_0)k_0}\},$$

with $\gamma_\rho \leq c_{l_0k_0} := \min\{1, \frac{\sqrt{(8l_0-12k_0)^2 + 16(36k_0l_0 - 9k_0 - 4l_0)} - (8l_0 + 12k_0)}{2(36k_0l_0 - 9k_0^2 - 4l_0^2)}\}$ *then* $\frac{\gamma_\rho}{1-g(\gamma_\rho)} \leq r$ *and* $l_0 r < 1$.

15.3. Error Analysis

We use the following assumption to obtain an error estimate for $\|x_\alpha^{\delta} - \hat{x}\|$.

Assumption 15.3.1. *There exists a continuous, strictly monotonically increasing function* $\varphi : (0, a] \to (0, \infty)$ *with* $a \geq \|F'(x_0)\|^2$ *satisfying;*

- $\lim_{\lambda \to 0} \varphi(\lambda) = 0$
- $$\sup_{\lambda \geq 0} \frac{\alpha\varphi(\lambda)}{\lambda + \alpha} \leq \varphi(\alpha), \qquad \forall \lambda \in (0, a].$$
- *there exists $v \in X$ such that*
$$x_0 - \hat{x} = \varphi(A_0^* A_0)v.$$

Theorem 15.3.2. *Let x_α^δ be as in (15.1.10). Then*

$$\|x_\alpha^\delta - \hat{x}\| \leq \frac{\max\{1, \|v\|\}}{1-q}(\frac{\delta}{\sqrt{\alpha}} + \varphi(\alpha))$$

where $q = l_0 r$.

Proof. Let $M = \int_0^1 F'(\hat{x} + t(x_\alpha^\delta - \hat{x}))dt$. Then

$$F(x_\alpha^\delta) - F(\hat{x}) = M(x_\alpha^\delta - \hat{x})$$

and hence by (15.1.10), we have $(A_0^* M + \alpha I)(x_\alpha^\delta - \hat{x}) = A_0^*(y^\delta - y) + \alpha(x_0 - \hat{x})$. Thus

$$\begin{aligned} x_\alpha^\delta - \hat{x} &= (A_0^* A_0 + \alpha I)^{-1}[A_0^*(y^\delta - y) + \alpha(x_0 - \hat{x}) + A_0^*(A_0 - M)(x_\alpha^\delta - \hat{x})] \\ &= s_1 + s_2 + s_3 \end{aligned} \tag{15.3.1}$$

where $s_1 := (A_0^* A_0 + \alpha I)^{-1} A_0^*(y^\delta - y), s_2 := (A_0^* A_0 + \alpha I)^{-1}\alpha(x_0 - \hat{x})$ and $s_3 := (A_0^* A_0 + \alpha I)^{-1} A_0^*(A_0 - M)(x_\alpha^\delta - \hat{x})$. Note that

$$\|s_1\| \leq \frac{\delta}{\sqrt{\alpha}}, \tag{15.3.2}$$

by Assumption 15.3.1

$$\|s_2\| \leq \varphi(\alpha)\|v\| \tag{15.3.3}$$

and by Assumption 15.2.1

$$\|s_3\| \leq l_0 r \|x_\alpha^\delta - \hat{x}\|. \tag{15.3.4}$$

The result now follows from (15.3.1), (15.3.2), (15.3.3) and (15.3.4).

15.3.1. Error Bounds under Source Conditions

Combining the estimates in Theorem 15.2.3 and Theorem 15.3.2 we obtain the following.

Theorem 15.3.3. *Let the assumptions in Theorem 15.2.3 and Theorem 15.3.2 hold and let $x_{n,\alpha}^\delta$ be as in (15.1.12). Then*

$$\|x_{n,\alpha}^\delta - \hat{x}\| \leq r e^{-\gamma 2^n} + \frac{\max\{1, \|v\|\}}{1-q}(\frac{\delta}{\sqrt{\alpha}} + \varphi(\alpha)).$$

Further if $n_\delta := \min\{n : e^{-\gamma 2^n} < \frac{\delta}{\sqrt{\alpha}}\}$, then

$$\|x_{n_\delta,\alpha}^\delta - \hat{x}\| \leq \tilde{C}(\frac{\delta}{\sqrt{\alpha}} + \varphi(\alpha))$$

where $\tilde{C} := r + \frac{\max\{1, \|v\|\}}{1-q}$.

15.3.2. A Priori Choice of the Parameter

Observe that the estimate $\frac{\delta}{\sqrt{\alpha}} + \varphi(\alpha)$ in Theorem 15.3.3 is of optimal order for the choice $\alpha := \alpha_\delta$ which satisfies $\frac{\delta}{\sqrt{\alpha_\delta}} = \varphi(\alpha)$. Now, using the function $\psi(\lambda) := \lambda\sqrt{\varphi^{-1}(\lambda)}, 0 < \lambda \leq a$, we have $\delta = \sqrt{\alpha}\varphi(\alpha) = \psi(\varphi(\alpha))$ so that $\alpha_\delta = \varphi^{-1}[\psi^{-1}(\delta)]$.

Theorem 15.3.4. *Let* $\psi(\lambda) = \lambda\sqrt{\varphi^{-1}(\lambda)}, 0 < \lambda \leq a$ *and assumptions in Theorem 15.3.3 holds. For* $\delta > 0$, *let* $\alpha_\delta = \varphi^{-1}[\psi^{-1}(\delta)]$ *and let* n_δ *be as in Theorem 15.3.3. Then*

$$\|x^\delta_{n_\delta,\alpha} - \hat{x}\| = O(\psi^{-1}(\delta)).$$

15.3.3. Adaptive Choice of the Parameter

In the balancing principle considered by Pereverzev and Schock in [23], the regularization parameter $\alpha = \alpha_i$ are selected from some finite set

$$D_N := \{\alpha_i : 0 < \alpha_0 < \alpha_1 < \cdots < \alpha_N\}.$$

Let

$$n_i = \min\{n : e^{-\gamma 2^n} \leq \frac{\delta}{\sqrt{\alpha_i}}\}$$

and let $x^\delta_{\alpha_i} := x^\delta_{n_i,\alpha_i}$ where $x^\delta_{n_i,\alpha_i}$ be as in (15.1.12) with $\alpha = \alpha_i$ and $n = n_i$. Then from Theorem 15.3.3, we have

$$\|x^\delta_{\alpha_i} - \hat{x}\| \leq \tilde{C}(\frac{\delta}{\sqrt{\alpha_i}} + \varphi(\alpha_i)), \forall i = 1, 2, \cdots N.$$

Precisely we choose the regularization parameter $\alpha = \alpha_k$ from the set D_N defined by

$$D_N := \{\alpha_i = \mu^i \alpha_0, i = 1, 2, \cdots N\}$$

where $\mu > 1$.

To obtain a conclusion from this parameter choice we considered all possible functions φ satisfying Assumption 15.2.1 and $\varphi(\alpha_i) \leq \frac{\delta}{\sqrt{\alpha_i}}$. Any of such functions is called admissible for $\hat{x}$ and it can be used as a measure for the convergence of $x^\delta_\alpha \to \hat{x}$ (see [19]).

The main result of this section is the following theorem, proof of which is analogous to the proof of Theorem 4.4 in [7].

Theorem 15.3.5. *Assume that there exists* $i \in \{0, 1, \cdots, N\}$ *such that* $\varphi(\alpha_i) \leq \frac{\delta}{\sqrt{\alpha_i}}$. *Let assumptions of Theorem 15.3.3 be satisfied and let*

$$l := \max\{i : \varphi(\alpha_i) \leq \frac{\delta}{\sqrt{\alpha_i}}\} < N,$$

$$k = \max\{i : \forall j = 1, 2, \cdots, i; \|x^\delta_{\alpha_i} - x^\delta_{\alpha_j}\| \leq 4\tilde{C}\frac{\delta}{\sqrt{\alpha_j}}\}$$

where $\tilde{C}$ *is as in Theorem 15.3.3. Then* $l \leq k$ *and*

$$\|x^\delta_{\alpha_k} - \hat{x}\| \leq 6\tilde{C}\mu\psi^{-1}(\delta).$$

15.4. Implementation of the Method

Finally the balancing algorithm associated with the choice of the parameter specified in Theorem 15.3.5 involves the following steps:

- Choose $\alpha_0 > 0$ such that $\delta_0 < c_{k_0 l_0}\sqrt{\alpha_0}$ and $\mu > 1$.
- Choose N big enough but not too large and $\alpha_i := \mu^i \alpha_0, i = 0, 1, 2, \cdots, N$.
- Choose $\rho \leq \frac{\sqrt{1+2l_0(c_{k_0 l_0} - \frac{\delta_0}{\sqrt{\alpha_0}})} - 1}{l_0}$ where $c_{k_0 l_0}$ is as in Remark 15.2.4.
- Choose $r \in (r_1, r_2)$.

15.4.1. Algorithm

1. Set $i = 0$.
2. Choose $n_i = min\{n : e^{-\gamma 2^n} \leq \frac{\delta}{\sqrt{\alpha_i}}\}$.
3. Solve $x^\delta_{n_i,\alpha_i} = x^\delta_{\alpha_i}$ by using the iteration (15.1.12) with $n = n_i$ and $\alpha = \alpha_i$.
4. If $\|x^\delta_{\alpha_i} - x^\delta_{\alpha_j}\| > 4\tilde{C}\frac{\delta}{\sqrt{\alpha_j}}, j < i$, then take $k = i - 1$ and return $x^\delta_{\alpha_k}$.
5. Else set $i = i + 1$ and return to Step 2.

References

[1] Argyros, I.K., *Convergence and Application of Newton-type Iterations*, Springer, 2008.

[2] Argyros, I.K., Hilout, S., A convergence analysis for directional two-step Newton methods, *Numer. Algor.*, 55 (2010), 503-528.

[3] Bakushinskii, A.B., The problem of the convergence of the iteratively regularized Gauss-Newton method, *Comput. Math. Phy.*, 32 (1992), 1353-1359.

[4] Bakushinskii, A.B., Iterative methods without saturation for solving degenerate nonlinear operator equations, *Dokl. Akad. Nauk*, 344 (1995), 7-8, MR1361018.

[5] Blaschke, B., Neubauer, A., Scherzer, O., On convergence rates for the iteratively regularized Gauss-Newton method, *IMA Journal on Numerical Analysis*, 17 (1997), 421- 436.

[6] Bockmann, C., Kammanee, A., Braunb, A., Logarithmic convergence of Levenberg-Marquardt method with application to an inverse potential problem, *J. Inv. Ill-Posed Problems*, 19 (2011), 345-367.

[7] George, S., On convergence of regularized modified Newton's method for nonlinear ill-posed problems, *J. Inv. Ill-Posed Problems*, 18 (2010), 133-146.

[8] George, S., "Newton type iteration for Tikhonov regularization of nonlinear ill-posed problems," *J. Math.*, 2013 (2013), Article ID 439316, 9 pages.

[9] Hanke, M., A regularizing Levenberg-Marquardt scheme, with applications to inverse groundwater filtration problems, *Inverse Problems*, 13 (1997), 79-95.

[10] Kanke, M., The regularizing Levenberg-Marquardt scheme is of optimal order, *J. Integral Equations Appl.* 22 (2010), 259-283.

[11] Hochbruck, M., Honig, M., On the convergence of a regularizing Levenberg-Marquardt scheme for nonlinear ill-posed problems, *Numer. Math.* 115 (2010), 71-79.

[12] Hohage, T., Logarithmic convergence rate of the iteratively regularized Gauss-Newton method for an inverse potential and an inverse scattering problem, *Inverse Problems*, 13 (1997), 1279-1299.

[13] Hohage, T., Regularization of exponentially ill-posed problems, *Numer. Funct. Anal&Optim.*, 21 (2000), 439-464.

[14] Jin, Q., On a regularized Levenberg-Marquardt method for solving nonlinear inverse problems, *Numer. Math.* 16 (2010), 229-259.

[15] Kaltenbacher, B., A posteriori parameter choice strategies for some Newton-type methods for the regularization of nonlinear ill-posed problems, *Numerische Mathematik*, 79 (1998), 501-528.

[16] Kaltenbacher, B., A note on logarithmic convergence rates for nonlinear Tikhonov regularization, *J. Inv. Ill-Posed Problems*, 16 (2008), 79-88.

[17] Langer, S., Hohage T., Convergence analysis of an inexact iteratively regularized Gauss-Newton method under general source conditions, *J. of Inverse & Ill-Posed Problems*, 15 (2007), 19-35.

[18] Levenberg, K., A Method for the solution of certain problems in least squares, *Quart. Appl. Math.* 2 (1944), 164-168.

[19] Lu, S., Pereverzev, S.V., Sparsity reconstruction by the standard Tikhonov method, *RICAM-Report* No. 2008-17.

[20] Mahale, P., Nair, M.T., A Simplified generalized Gauss-Newton method for nonlinear ill-posed problems, *Mathematics of Computation*, 78(265) (2009), 171-184.

[21] Marquardt, D., An Algorithm for least-squares estimation of nonlinear parameters, *SIAM J. Appl.Math.*, 11(1963), 431-441.

[22] Ortega, J.M., Rheinboldt, W.C., Iterative solution of nonlinear equations in general variables, Academic Press, New York and London (1970).

[23] Perverzev, S.V., Schock, E., "On the adaptive selection of the parameter in regularization of ill-posed problems", *SIAM J. Numer. Anal.* 43 (2005), 2060-2076.

[24] Pornsawad, P., Bockmann, C., Convergence rate analysis of the first stage Runga-Kutta-type regularizations, *Inverse Problems*, 26 (2010), 035005.

Chapter 16

Local Convergence Analysis of Proximal Gauss-Newton Method for Penalized Nonlinear Least Squares Problems

16.1. Introduction

Let $\mathcal{X}$ and Y be Hilbert spaces. Let $\mathcal{D} \subseteq \mathcal{X}$ be open set and $F : \mathcal{D} \longrightarrow \mathcal{Y}$ be continuously Fréchet-differentiable. Moreover, let $J : \mathcal{D} \rightarrow \mathbb{R} \cup \{\infty\}$ be proper, convex and lower semi-continuous. In this study we are concerned with the problem of approximating a locally unique solution $x^\star$ of the penalized nonlinear least squares problem

$$\min_{x \in \mathcal{D}} \| F(x) \|^2 + J(x). \tag{16.1.1}$$

A solution $x^\star \in \mathcal{D}$ of (16.1.1) is also called a least squares solution of the equation $F(x) = 0$.

Many problems from computational sciences and other disciplines can be brought in a form similar to equation (16.1.1) using Mathematical Modelling [3, 6, 14, 16]. For example in data fitting, we have $\mathcal{X} = \mathbb{R}^i$, $\mathcal{Y} = \mathbb{R}^j$, i is the number of parameters and j is the number of observations.

The solution of (16.1.1) can rarely be found in closed form. That is why the solution methods for these equations are usually iterative. In particular, the practice of numerical analysis for finding such solutions is essentially connected to Newton-type methods [1, 2, 3, 5, 4, 6, 7, 14, 17]. The study about convergence matter of iterative procedures is usually centered on two types: semilocal and local convergence analysis. The semilocal convergence matter is, based on the information around an initial point, to give criteria ensuring the convergence of iterative procedures; while the local one is, based on the information around a solution, to find estimates of the radii of convergence balls. A plethora of sufficient conditions for the local as well as the semilocal convergence of Newton-type methods as well as an error analysis for such methods can be found in [1]–[20].

If $J = 0$, we obtain the well known Gauss-Newton method defined by

$$x_{n+1} = x_n - F'(x_n)^+ F(x_n), \qquad \text{for each} \quad n = 0, 1, 2, \ldots, \tag{16.1.2}$$

where $x_0 \in \mathcal{D}$ is an initial point [12] and $F'(x_n)^+$ is the Moore-Penrose inverse of the linear operator $F'(x_n)$. In the present paper we use the proximal Gauss-Newton method (to be precised in Section 16.2, see (16.2.6)) for solving penalized nonlinear least squares problem (16.1.1). Notice that if $J = 0$, $x^\star$ is a solution of (16.1.1), $F(x^\star) = 0$ and $F'(x^\star)$ is invertible, then the theories of Gauss-Newton methods merge into those of Newton method. A survey of convergence results under various Lipschitz-type conditions for Gauss-Newton-type methods can be found in [2, 6] (see also [5, 9, 10, 12, 15, 18]). The convergence of these methods requires among other hypotheses that F' satisfies a Lipschitz condition or F'' is bounded in $\mathcal{D}$. Several authors have relaxed these hypotheses [4, 8, 9, 10, 15]. In particular, Ferreira et al. [1, 9, 10] have used the majorant condition in the local as well as semilocal convergence of Newton-type method. Argyros and Hilout [3, 4, 5, 6, 7] have also used the majorant condition to provide a tighter convergence analysis and weaker convergence criteria for Newton-type method. The local convergence of inexact Gauss-Newton method was examined by Ferreira et al. [9] using the majorant condition. It was shown that this condition is better that Wang's condition [15], [20] in some sence. A certain relationship between the majorant function and operator F was established that unifies two previously unrelated results pertaining to inexact Gauss-Newton methods, which are the result for analytical functions and the one for operators with Lipschitz derivative.

In [7] motivated by the elegant work in [10] and optimization considerations we presented a new local convergence analysis for inexact Gauss-Newton-like methods by using a majorant and center majorant function (which is a special case of the majorant function) instead of just a majorant function with the following advantages: larger radius of convergence; tighter error estimates on the distances $\| x_n - x^\star \|$ for each $n = 0, 1, \cdots$ and a clearer relationship between the majorant function and the associated least squares problems (16.1.1). Moreover, these advantages are obtained under the same computational cost, since as we will see in Section 16.3. and Section 16.4., the computation of the majorant function requires the computation of the center-majorant function. Furthermore, these advantages are very important in computational mathematics, since we have a wider choice of initial guesses x_0 and fewer computations to obtain a desired error tolerance on the distances $\| x_n - x^\star \|$ for each $n = 0, 1, \cdots$. In the present paper, we obtain the same advantages over the work by Allende and Gonçalves [1] but using the proximal Gauss-Newton method [6, 18].

The paper is organized as follows. In order to make the paper as self contained as possible, we provide the necessary background in Section 16.2.. Section 16.3. contains the local convergence analysis of inexact Gauss-Newton-like methods. Some proofs are abbreviated to avoid repetitions with the corresponding ones in [18]. Special cases and applications are given in the concluding Section 16.4..

16.2. Background

Let $U(x, r)$ and $\overline{U}(x, r)$ stand, respectively, for the open and closed ball in $\mathcal{X}$ with center $x \in \mathcal{D}$ and radius $r > 0$. Let $A : \mathcal{X} \longrightarrow \mathcal{Y}$ be continuous linear and injective with closed image, the Moore-Penrose inverse [3] $A^+ : \mathcal{Y} \longrightarrow \mathcal{X}$ is defined by $A^+ = (A^\star A)^{-1} A^\star$. I denotes the identity operator on $\mathcal{X}$ (or $\mathcal{Y}$). Let $\mathcal{L}(\mathcal{X}, \mathcal{Y})$ be the space of bounded linear operators from $\mathcal{X}$ into $\mathcal{Y}$. Let $M \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$, the $Ker(M)$ and $Im(M)$ denote the Kernel and image of M, respectively and M^* its adjoint operator. Let $M \in \mathcal{L}(\mathcal{X},\mathcal{Y})$ with a closed image. Recall that the Moore-Pentose inverse of M is the linear operator $M^+ \in \mathcal{L}(\mathcal{Y},\mathcal{X})$ which satisfies

$$MM^+M = M, M^+MM^+ = M, (MM^+)^* = MM^+, (M^+M)^* = M^+M. \tag{16.2.1}$$

It follows from (16.2.1) that if $\prod_S$ denotes the projection of X onto subspace S, then

$$M^+M = I_X - \prod_{Ker(M)}, MM^+ = \prod_{Im(M)}. \tag{16.2.2}$$

Moreover, if M is injective, then

$$M^+ = (M^*M)^{-1}M^*, M^+M = I_X, \|M^+\|^2 = \|(M^*M)^{-1}\|. \tag{16.2.3}$$

Lemma 16.2.1. *[3, 6, 14] (Banach's Lemma) Let $A : \mathcal{X} \longrightarrow \mathcal{X}$ be a continuous linear operator. If $\| A - I \| < 1$ then $A^{-1} \in \mathcal{L}(\mathcal{X},\mathcal{X})$ and $\| A^{-1} \| \leq 1/(1- \| A - I \|)$.*

Lemma 16.2.2. *[1, 3, 6, 10] Let $A, E : \mathcal{X} \longrightarrow \mathcal{Y}$ be two continuous linear operators with closed images. Suppose $B = A + E$, A is injective and $\| EA^+ \| < 1$. Then, B is injective.*

Lemma 16.2.3. *[1, 3, 6, 10] Let $A, E : \mathcal{X} \longrightarrow \mathcal{Y}$ be two continuous linear operators with closed images. Suppose $B = A + E$ and $\| A^+ \| \| E \| < 1$. Then, the following estimates hold*

$$\| B^+ \| \leq \frac{\| A^+ \|}{1- \| A^+ \| \| E \|} \quad \textit{and} \quad \| B^+ - A^+ \| \leq \frac{\sqrt{2} \| A^+ \|^2 \| E \|}{1- \| A^+ \| \| E \|}.$$

The semilocal convergence of proximal Gauss-Newton method using Wang's condition was introduced in [18]. Next, in order to make the paper as self contained, as possible, we briefly illustrate how this method is defined. Let $Q : \mathcal{X} \to \mathcal{X}$ be continuous, positive, self adjoint and bounded from below. It follows that $Q^{-1} \in \mathcal{L}(\mathcal{X},\mathcal{X})$. Define a scalar product on X by $< u, v > = < u, Q_v >$. Then, the corresponding induced norm $\| . \|_Q$ is equivalent to the given norm on X, since $\frac{1}{\|Q^{-1}\|} \| x \| \leq \| x \|_Q^2 \leq \| Q \| \| x \|^2$. The Moreau approximation of J [18] with respect to the scalar product induced by Q in the functional $\Gamma : \mathcal{X} \to \mathbb{R}$ defined by

$$\Gamma(y) = \inf_{x \in X} \left\{ J(x) + \frac{1}{2} \| x - y \|_Q^2 \right\} \tag{16.2.4}$$

It follows from the properties of J that the infimum in (16.2.4) is obtained at a unique point. Let us denote by $\text{prox}_J^Q(y)$ the proximity operator:

$$\begin{array}{rccl} \text{prox}_J^Q : & \mathcal{X} & \to & \mathcal{X} \\ & y & \to & \Gamma(y) = \text{argmin}_{x \in X} \left\{ J(x) + \frac{1}{2} \| x - y \|_Q^2 \right\} \end{array} \tag{16.2.5}$$

The first optimality condition for (16.2.4) leads to

$$\begin{aligned} z = \text{prox}_J^Q(y) \quad &\Leftrightarrow \quad 0 \in \partial J(z) + Q(z - y) \\ &\Leftrightarrow \quad Q(z) \in (\partial I + Q)(z), \end{aligned}$$

which leads to

$$\operatorname{prox}_J^Q(y) = (\partial I + Q)^{-1}(Q(y))$$

by using the minimum in (16.2.4). In view, of the above, we can define the proximal Gauss-Newton method by

$$x_{n+1} = \operatorname{prox}_J^{H(x_n)}(x_n - F'(x_n)^+ F(x_n)) \text{ for each } n = 0,1,2,\ldots \tag{16.2.6}$$

where x_0 is an initial point, $H(x_n) = F'(x_n)^* F'(x_n)$ and $\operatorname{prox}_J^{H(x_n)}$ is defined in (16.2.5).

Next, we present some auxiliary results.

Lemma 16.2.4. *[18] Let Q_1 and Q_2 be continuous, positive self adjoint operators and bounded from below on X. Then, the following hold*

$$\begin{aligned} \| \operatorname{prox}_J^{Q_1}(y_1) - \operatorname{prox}_J^{Q_2}(y_2) \| \quad &\leq \quad \sqrt{\| Q_1 \| \| Q_1^{-1} \|} \, \| y_1 - y_2 \| \\ &+ \| Q_1^{-1} \| \| (Q_1 - Q_2)(y_2 - \operatorname{prox}_J^{Q_2}(y_2)) \| . \end{aligned} \tag{16.2.7}$$

for each $y_1,\ y_2 \in X$.

Lemma 16.2.5. *[18] Given $x_n \in X$, if $F'(x_n)$ is injective with closed image, then x_{n+1} satisfies*

$$x_{n+1} = \operatorname{argmin}_{x \in X} \frac{1}{2} \| F(x_n) + F'(x_n)(x - x_n) \|^2 + J(x). \tag{16.2.8}$$

Lemma 16.2.6. *[18] Suppose: $x^* \in \mathcal{D}$ satisfies $-F'(x^*)^* F(x^*) \in \partial J(x^*)$; $F'(x^*)$ is injective and $Im(F'(x^*))$ is closed. Then x^* satisfies*

$$x^* = \operatorname{prox}_J^{H(x^*)}\left(x^* - F'(x_n)^+ F(x^*)\right). \tag{16.2.9}$$

Proposition 16.2.7. *[10] Let $R > 0$. Suppose $g : [0,R) \longrightarrow \mathbb{R}$ is convex. Then, the following holds*

$$D^+ g(0) = \lim_{u \to 0^+} \frac{g(u) - g(0)}{u} = \inf_{u > 0} \frac{g(u) - g(0)}{u}.$$

Proposition 16.2.8. *[10] Let $R > 0$ and $\theta \in [0,1]$. Suppose $g : [0,R) \longrightarrow \mathbb{R}$ is convex. Then, $h : (0,R) \longrightarrow \mathbb{R}$ defined by $h(t) = (g(t) - g(\theta t))/t$ is increasing.*

16.3. Local Convergence Analysis of the Proximal Gauss-Newton Method

We shall prove the main local convergence results for the proximal Gauss-Newton method (16.2.6) for solving the penalized nonlinear least squares problem (16.1.1) under the (H) conditions given as follows:

(H_0) Let $D \subseteq X$ be open; $J : D \to \mathbb{R} \cup \{+\infty\}$ be proper, convex and lower semicontinuously Fréchet-differentiable such that F' has a closed image in D;

(H_1) Let $x^* \in D, R > 0, \alpha := \| F(x^*) \|, \beta := \| F'(x^*)^+ \|, \gamma := \beta \| F'(x^*) \|$ and $\delta := \sup\{t \in [0,R) : U(x^*,t) \subset \mathcal{D}\}$. Operator $-F'(x^*)^* F(x^*) \in \partial J(x^*)$, $F'(x)$ is injective and there exist $f_0,\ f : [0,R) \to \mathbb{R}$ continuously differentiable such that for each $x \in U(x^*,\delta)$, $\theta \in [0,1]$ and $\lambda(x) = \| x - x^* \|$:

$$\beta \| F'(x) - F'(x^*) \| \leq f_0'(\lambda(x)) - f_0'(0) \tag{16.3.1}$$

and

$$\beta \| F'(x) - F'(x^* + \theta(x - x^*)) \| \leq f_0'(\lambda(x)) - f_0'(\theta\lambda(x)); \tag{16.3.2}$$

(H_2) $f_0(0) = f(0) = 0$ and $f_0'(0) = f'(0) = -1$;

(H_3) f_0', f' are strictly increasing and for each $t \in [0,R)$

$$f_0(t) \leq f(t) \text{ and } f_0'(t) \leq f'(t);$$

(H_4) $\left[\left(1+\sqrt{2}\right)\gamma + 1\right] \alpha\beta\mathcal{D}^+ f_0'(0) < 1$;

Let positive constants $\nu,\ \rho,\ r$ and function Ψ be defined by

$$\nu := \sup\left\{t \in [0,R) : f_0'(t) < 0\right\},$$

$$\rho := \sup\{t \in [0,\nu) : \Psi(t) < 1\},$$

$$r := \min\{\nu,\ \rho\}$$

and

$$\Psi(t) := \frac{(f_0'(t) + 1 + \gamma)\left[t f'(t) - f(t) + \alpha\beta\left(1+\sqrt{2}\right)(f_0'(t)+1)\right] + \alpha\beta(f_0'(t)+1)}{t\left[f_0'(t)\right]^2}.$$

Remark 16.3.1. *In the literature, with the exception of our works [2, 3, 4, 5, 6, 7] only* (16.3.2) *is used. However, notice that* (16.3.2) *always implies* (16.3.1). *That is* (16.3.1) *is not an additional to* (16.3.2) *hypothesis. Moreover,*

$$f_0'(t) \leq f'(t) \text{ for each } t \in [0,R) \tag{16.3.3}$$

holds in general and $\frac{f'(t)}{f_0'(t)}$ *can be arbitrarily large [3, 6]. Using more precise* (16.3.1) *instead of* (16.3.2) *for the computation of the upper bounds on the norms* $\| F'(x)^+ \|$ *and* $\| F'(x)^+ - F'(x^*)^+ \|$ *leads to a tighter error estimates on* $\| x_n - x^* \|$ *and a larger radius of convergence (if* $f_0'(t) < f'(t)$*) that if only* (16.3.2) *was used (see also Remark 3.11, the last Section and the numerical example).*

Theorem 16.3.2. *Under the* (H) *hypotheses, let* $x_0 \in U(x^*,r) \setminus \{x^*\}$*. Then, sequence* $\{x_n\}$ *generated by proximal Gauss-Newton method* (16.2.6) *for solving penalized nonlinear least squares problem* (16.1.1) *is well defined, remains in* $U(x^*,r)$ *and converges to* x^**. Moreover, the following estimates hold for each* $n = 0,1,2,\ldots$*:*

$$\lambda_{n+1} = \lambda(x_{n+1}) \leq \varphi_{n+1} := \varphi(\lambda_0, \lambda_n, f, f', f_0'), \tag{16.3.4}$$

where

$$\begin{aligned}
\varphi(\lambda_0,\lambda_n,f,f',f_0') &= \frac{(f_0'(\lambda_0)+1+\gamma)[f'(\lambda_0)\lambda_0-f(\lambda_0)]}{(\lambda_0 f_0'(\lambda_0))^2}\lambda_n^2 \\
&+\frac{\left(1+\sqrt{2}\right)\alpha\beta(f_0'(\lambda_0)+1)^2}{(\lambda_0 f_0'(\lambda_0))^2}\lambda_n^2 \\
&+\frac{\alpha\beta\left[\left(1+\sqrt{2}\right)\gamma+1\right][f_0'(\lambda_0)+1]}{\lambda_0(f_0'(\lambda_0))^2}\lambda_n.
\end{aligned}$$

In order for us to prove Theorem 16.3.2 we shall need several auxiliarity results. The proofs of the next four Lemmas are omitted, since they have been given, respectively in Lemmas 16.3.1-16.3.4 in [7]. From now on we assume that hypotheses (H) are satisfied.

Lemma 16.3.3. *The following hold,* $\nu > 0$, *and* $f_0'(t) < 0$ *for each* $t \in [0,\nu)$.

Lemma 16.3.4. *The function* g_i, $i = 1,2,\ldots,7$ *defined by*

$$g_1(t) = -\frac{1}{f_0'(t)},$$

$$g_2(t) = -\frac{f_0'(t)+1+\gamma}{f_0'(t)},$$

$$g_3(t) = \frac{tf'(t)-f(t)}{t^2},$$

$$g_4(t) = \frac{f_0'(t)+1}{t},$$

$$g_5(t) = \frac{(f_0'(t)+1+\gamma)(tf'(t)-f(t))}{(tf_0'(t))^2},$$

$$g_6(t) = \frac{(f_0'(t)+1)^2}{(tf_0'(t))^2}$$

and

$$g_7(t) = \frac{f_0'(t)+1}{t(f_0'(t))^2}$$

for each $t \in [0,\nu)$ *are positive and increasing.*

Lemma 16.3.5. *The following hold,* $\rho > 0$, *and* $0 \le \psi(t) < 1$ *for each* $t \in [0,\rho)$, *where function* ψ *is defined in the* (H) *hypotheses.*

Lemma 16.3.6. *Let* $x \in \mathcal{D}$. *Suppose that* $\lambda(x) < \min\{\nu,\rho\}$ *and the* (H) *hypotheses hold excluding* (16.3.2). *Then, the following items hold:*

$$\| F'(x)^+ \| \le -\frac{\beta}{f_0'(\lambda(x))},$$

$$\| F'(x)^+ - F'(x^*)^+ \| \le -\frac{\sqrt{2}\beta(f_0'(\lambda(x))+1)}{f_0'(\lambda(x))}$$

and

$$H(x) = F'(x^*)F'(x) \text{ is invertible on } U(x^*,r).$$

Remark 16.3.7. *It is worth noticing (see also Remark 16.3.2 that the estimates in Lemma 16.3.6 hold with f_0 replaced by f (i.e. using* (16.3.2) *instead of* (16.3.1)*). However, in this case these estimates are less tight.*

Lemma 16.3.8. *Let $x \in \mathcal{D}$. Suppose that $\lambda(x) < \min\{\nu, \delta\}$ and the (H) hypotheses excluding* (16.3.2) *hold. Then, the following items hold for each $x \in \mathcal{D}$:*

(a) $\| H(x) \|^{\frac{1}{2}} \leq \frac{f_0'(\lambda(x))+1+\gamma}{\beta}$;

(b) $\| H(x)^{-1} \|^{\frac{1}{2}} \leq -\frac{\beta}{f_0'(\lambda(x))}$

and

(c) $\beta \| (H(x) - H(x^*))F'(x^*)^+ \| \leq (f_0'(\lambda(x)) + 2 + \gamma)(f_0'(\lambda(x)) + 1)$.

Proof.

(a) It follows from (16.3.1) that

$$\begin{aligned}
\beta \| F'(x) \| &= \| F'(x^*)^+ \| \| F'(x) \| \\
&\leq \beta(\| F'(x) - F'(x^*) \| + \| F'(x^*) \|) \\
&\leq f_0'(\lambda(x)) + 1 + \gamma.
\end{aligned}$$

Then (a) follows from the preceding estimate and

$$\| H(x) \|^{\frac{1}{2}} = F'(x) \|^{\frac{1}{2}} = \| F'(x) \| .$$

(b) Use Lemma 16.3.6, the definition of H and the last property in (16.2.3).

(c) We use (16.2.2) , (b) and (16.3.1) to obtain in turn that

$$\begin{aligned}
\beta \| (H(x) - H(x^*))F'(x^*)^+ \| &= \beta \| F'(x)^*(F'(x) - F'(x^*))F'(x^*)^+ \\
&\quad +(F'(x) - F'(x^*))^* \textstyle\prod_{Im(F'(x^*))} \| \\
&\leq (\| F'(x) \| \| F'(x^*)^+ \| + 1)\beta \| F'(x) - F'(x^*) \|) \\
&\leq (f_0'(\lambda(x)) + 2 + \gamma)(f_0'(\lambda(x)) + 1).
\end{aligned}$$

The proof of the Lemma is complete. ⊠

As in [1, 7, 10, 18] we define the linearization error at a point in D by

$$E_F(x,y) := F(y) - [F(x) + F'(x)(y - x)] \text{ for each } x, y \in \mathcal{D}.$$

Then using (16.3.2) we bound this error by majorant function

$$e_f(t,u) = f(u) - [f(t) + f'(t)(u - t)] \text{ for each } t, u \in [0, R).$$

In particular we have (see Lemma 16.3.5 in [7] for the proof).

Lemma 16.3.9. *Let $x \in \mathcal{D}$. Suppose that $\lambda(x) < \delta$, then the following items hold:*

$$\beta \parallel E_F(x,x^*) \parallel \leq e_f(\lambda(x),0).$$

Remark 16.3.10. **(a)** *Using* (16.3.2) *only according to Lemma 16.3.9 we have that*

$$\beta \parallel E(x,x^*) \parallel \leq e_{f_0}(\lambda(x),0) + 2(f_0(\lambda(x)) + \lambda(x)).$$

(b) *Let us denote by G the proximal Gauss-Newton iteration operator by*

$$\begin{array}{rlcl} G: & U(x^*,r) & \to & \mathcal{X} \\ & x & \to & \operatorname{prox}_J^{H(x)}(G(x)), \end{array}$$

where

$$G(x) = x - F'(x)^+ F(x).$$

Notice that according to Lemma 16.3.8 $H(x)$ is invertible in $U(x^*,r)$. Hence, $F'(x)^+$ and $\operatorname{prox}_J^{H(x)}$ *are well defined in $U(x^*,r)$.*

Next, we provide the proof of the Theorem 16.3.2.

Proof. Let $x \in \mathcal{D}$. Suppose that $\lambda(x) < r$. Then, we shall first show that operator G is well defined and

$$\parallel G(x) - x^* \parallel \leq \varphi(\lambda(x), \lambda(x), f, f', f_0'), \tag{16.3.5}$$

where function φ was defined in Theorem 16.3.2. Using Lemma 16.2.6 as $-F'(x^*)^* F(x^*) \in \partial J(x^*)$ and $F'(x)$ is injective we have that $x^* = \operatorname{prox}_J^{H(x)}(G_F(x^*))$. Then, according to Lemma 2.4 we have in turn that

$$\begin{aligned} \parallel G(x) - x^* \parallel &= \parallel \operatorname{prox}_J^{H(x)}(G_F(x) - \operatorname{prox}_J^{H(x^*)}(G_F(x^*))) \\ &\leq (\parallel H(x) \parallel \parallel H(x)^{-1} \parallel)^{\frac{1}{2}} \parallel G(x) - G(x^*) \parallel \\ &\quad + \parallel H(x)^{-1} \parallel \parallel (H(x) - H(x^*))(G(x^*) - \operatorname{prox}_J^{H(x^*)}(G(x^*))) \parallel \\ &\leq P_1(x,x^*) + P_2(x,x^*), \end{aligned} \tag{16.3.6}$$

where for simplicity we set

$$P_1(x,x^*) = (\parallel H(x) \parallel \parallel H(x)^{-1} \parallel)^{\frac{1}{2}} \parallel G(x) - G(x^*) \parallel$$

and

$$P_2(x,x^*) = \parallel H(x)^{-1} \parallel \parallel (H(x) - H(x^*)) F'(x^*)^+ \parallel \parallel F(x^*) \parallel .$$

Using the definition of P_2 and items (b) and (c) of Lemma 16.3.8 we get that

$$P_2(x,x^*) \leq \frac{\alpha\beta}{(f_0'(\lambda(x)))^2} (f_0'(\lambda(x)) + 2 + \gamma)(f_0'(\lambda(x)) + 1). \tag{16.3.7}$$

Then, to find an upper bound on P_2, we first need to find an upper bound on $\| G(x)-G(x^*) \|$. Indeed, we have in turn that

$$\begin{aligned}
\| G(x)-G(x^*) \| &= \| F'(x)^+ [F'(x)(x-x^*)-F(x)+F(x^*)] + (F'(x^*)^+ - F'(x)^+)F(x^*) \| \\
&\leq \| F'(x)^+ \| \| E_F(x,x^*) \| + \| F'(x^*)^+ - F'(x)^+ \| \| F(x^*) \| \\
&\leq -\frac{e_f(\lambda(x),0)}{f_0'(\lambda(x))} - \frac{\sqrt{2}\alpha\beta(f_0'(\lambda(x))+1)}{f_0'(\lambda(x))} \\
&\leq \frac{f_0'(\lambda(x))+1+\gamma}{f_0'(\lambda(x))^2}\Big(e_f(\lambda(x),0)+\sqrt{2}\alpha\beta(f_0'(\lambda(x))+1)\Big),
\end{aligned} \tag{16.3.8}$$

where we used Lemma 16.3.6, Lemma 16.3.8 (a) and (b) and Lemma 16.3.9. Then, (16.3.5) follows from (16.3.6) by summing up (16.3.7) we have that

$$\| G(x)-x^* \| \leq q(x)\lambda(x), \tag{16.3.9}$$

where

$$\begin{aligned}
q(x) &= \frac{(f_0'(\lambda(x))+1+\gamma)\left[\lambda(x)f'(\lambda(x))-f(\lambda(x))+\alpha\beta\left(1+\sqrt{2}\right)(f_0'(\lambda(x))+1)\right]}{\lambda(x)\,[f_0(\lambda(x))]^2} \\
&\quad + \frac{+\alpha\beta(f_0'(\lambda(x))+1)}{\lambda(x)\,[f_0(\lambda(x))]^2}.
\end{aligned}$$

But $q(x) \in [0,1)$, by Lemma 16.3.5, since $x \in U(x^*,r)\setminus\{x^*\}$, so that $0<\lambda(x)<r<\rho$. That is we have

$$\| G(x)-x^* \| < \| x-x^* \| . \tag{16.3.10}$$

In particular $x_0 \in U(x^*,r)\setminus\{x^*\}$. That is $0<\lambda(x_0)<r$. Then, using mathematical induction, Lemma 16.3.6 and (16.3.10) for $x=x_0$ we get that $\lambda(x_1) = \| x_1 - x^* \| < \| x_0 - x^* \| = \lambda(x_0) < r$. Similarly, we get as in (16.3.9) that

$$\begin{aligned}
\| x_{k+1}-x^* \| &\leq q(x_0) \| x_k - x^* \| \\
&< \| x_k - x^* \| \\
&< r
\end{aligned}$$

from which it follows that $\lim_{k\to\infty} x_k = x^*$ and sequence $\{x_k\}$ remains in $U(x^*,r)\setminus\{x^*\}$. ⊠

Remark 16.3.11. *If $f_0 = f$, then the results of this Section reduce to the corresponding ones in [1] (see also [9]). Otherwise, i.e. if strict inequality holds in (16.3.3), then: our sufficient convergence condition (H_4) is weaker than the one in [1] using f' instead of f_0' (i.e. the applicability of the method is extended in cases that cannot be covered before); our convergence ball is larger and the estimates on the distances $\|x_n - x^*\|$ more precise, which imply that we have a wider choice of the initial guesses and less iterates are required to obtain a given error tolerance. Notice also that these advantages are obtained under the same computational cost as in [1, 9], since in practice the computation of the function f requires the computation of f_0 as a special case. Therefore, these developments are very important in computational mathematics.*

16.4. Special Cases and Numerical Examples

We present a special case of Theorem 16.3.2. This case is based on the center-Lipschitz and Lipschitz conditions [2, 3, 4, 5, 7]. We refer the reader to [3, 6] for another case based on Smale's alpha theory [19].

Remark 16.4.1. *Let us define functions $f_0, f : [0,\gamma] \to \mathbb{R}$ by*

$$f_0(t) = \frac{L_0}{2}t^2 - t \text{ and } f(t) = \frac{L}{2}t^2 - t, \tag{16.4.1}$$

where $0 < L_0 < L$ are the center-Lipschitz and Lipschitz constants, respectively. We have that $f_0(0) = f(0) = 0$ and $f_0'(0) = f'(0) = -1$. Notice that (16.3.3) holds as a strict inequality in this case. Then, one can specialize Theorem 16.3.2 using the above choices. Clearly, the results improve the corresponding ones (with advantages as already stated in the introduction of this study and in Remark 16.3.11) using only (16.2.2) *(i.e. if $f_0 = f$).*

Since such results as far as we know are not available, let us at least consider the case $\alpha = 0$. That is we consider the case of zero-residual problems. Then, Theorem 16.3.2 specializes to:

Corollary 16.4.2. *Let $\mathcal{D} \subseteq X$ be open, $J : \mathcal{D} \to \mathbb{R} \cup \{+\infty\}$ be proper, convex and lower semi-continuous and $F : \mathcal{D} \to \mathcal{Y}$ be continuously Fréchet-differentiable and F' be with closed image in $\mathcal{D}$. Let $x^* \in \mathcal{D}$, $R > 0$, $\beta = \| F'(x^*)^+ \|$, $\gamma = \beta \| F'(x^*) \|$ and $\delta = \sup\{t \in [0,R) : U(x^*,t) \subset \mathcal{D}\}$. Suppose that $F(x^*) = 0$, $0 \in \partial J(x^*)$, $F'(x^*)$ is injective and there exists L_0 and L such that for each $x \in U(x^*,\delta)$, $\theta \in [0,1]$:*

$$\beta \| F'(x) - F'(x^*) \| \leq L_0 \| x - x^* \|$$

and

$$\beta \| F'(x) - F'(x^* + \theta(x - x^*)) \| \leq L(1-\theta) \| x - x^* \| .$$

Let

$$r := \min\left\{ \frac{4+\gamma - \sqrt{(4+\gamma)^2 - 8}}{2L_0}, \delta \right\}.$$

Then, sequence $\{x_n\}$ generated by proximal Gauss-Newton method (16.2.6) *for solving penalized nonlinear least squares problem* (16.1.1) *is well defined, remains in $U(x^*,r)$ and converges to x^* provided that $x_0 \in U(x^*,r) \setminus \{x^*\}$. Moreover, the following estimates hold*

$$\| x_{k+1} - x^* \| \leq \frac{L(\gamma + 2L_0 \| x_0 - x^* \|)}{2(1 - L_0 \| x_0 - x^* \|)} \| x_n - x^* \|^2 \text{ for each } n = 0,1,2,\ldots.$$

The preceding results improve earlier ones [1, 8, 9, 10, 12, 15, 18] when $L_0 < L$ (see also Remark 16.3.11). Next, we present an example where $L_0 < L$. More example, where $L_0 < L$ in the Lipschitz case or in Smale's alpha theory can be found in [3, 4, 5, 6, 7].

Example 16.4.3. *Let $X = Y = \mathbb{R}^3$, $D = U(0,1)$, $x^* = (0,0,0)$ and define function F on D by*

$$F(x,y,z) = (e^x - 1, \frac{e-1}{2}y^2 + y, z). \tag{16.4.2}$$

For simplicity we consider the nonlinear equation $F(x) = 0$ *instead of* (16.1.1)*. We have that for* $u = (x, y, z)$

$$F'(u) = \begin{pmatrix} e^x & 0 & 0 \\ 0 & (e-1)y+1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{16.4.3}$$

Using the norm of the maximum of the rows and (16.4.2)–(16.4.3) we see that since $F'(x^*) = diag\{1,1,1\}$*, we can define parameters* L_0 *and* L *by*

$$L_0 = e - 1 < L = e.$$

References

[1] Allende, G.B., Gonçalves, M.L.N., Local convergence analysis of a proximal Gauss-Newton method under a majorant condition, preprint http://arxiv.org/abs/1304.6461

[2] Amat, S., Busquier, S., Gutiérrez, J.M., Geometric constructions of iterative functions to solve nonlinear equations, *J. Comput. Appl. Math.*, 157 (2003), 197-205.

[3] Argyros, I.K., *Convergence and applications of Newton–type iterations*, Springer–Verlag Publ., New York, 2008.

[4] Argyros, I.K., Hilout, S., Extending the applicability of the Gauss-Newton method under average Lipschitz-type conditions, *Numer. Algorithms*, 58 (2011), 23-52.

[5] Argyros, I.K., Hilout, S., Improved local convergence of Newton's method under weak majorant condition, *J. Comput. Appl. Math.*, 236 (2012), 1892-1902.

[6] Argyros, I.K., Hilout, S., *Computational methods in Nonlinear Analysis*, World Scientific Publ. Comp. New Jersey, 2013.

[7] Argyros, I.K., Hilout, S., Improved local convergence analysis of inexact Gauss-Newton like methods under the majorizing condition in Banach space, *J. Franklin Institute*, 350 (2013), 1531-1544.

[8] Chen, J., Li, W., Local convergence results of Gauss-Newton's like method in weak conditions, *J. Math. Anal. Appl.*, 324 (2006), 1381–1394.

[9] Ferreira, O.P., Gonçalves, M.L.N, Oliveira, P.R., Local convergence analysis of Gauss–Newton like methods under majorant condition, *J. Complexity*, 27 (2011), 111–125.

[10] Ferreira, O.P., Gonçalves, M.L.N, Oliveira, P.R., Local convergence analysis of inexact Gauss–Newton like methods under majorant condition, *J. Comput. Appl. Math.*, 236 (2012), 2487–2498.

[11] Gutiérrez, J.M., Hernández, M.A., Newton's method under weak Kantorovich conditions, *IMA J. Numer. Anal.*, 20 (2000), 521–532.

[12] Häubler, W.M., A Kantorovich–type convergence analysis for the Gauss–Newton–method, *Numer. Math.*, 48 (1986), 119–125.

[13] Huang, Z.A., Convergence of inexact Newton method, *J. Zhejiang Univ. Sci. Ed.*, 30 (2003), 393–396.

[14] Kantorovich, L.V., Akilov, G.P., Functional Analysis, Pergamon Press, Oxford, 1982.

[15] Li, C., Hu, N., Wang, J., Convergence bahavior of Gauss–Newton's method and extensions to the Smale point estimate theory, *J. Complexity*, 26 (2010), 268–295.

[16] Potra, F.A., Sharp error bounds for a class of Newton–like methods, *Libertas Mathematica*, 5 (1985), 71–84.

[17] Proinov, P.D., General local convergence theory for a class of iterative processes and its applications to Newton's method, *J. Complexity*, 25 (2009), 38–62.

[18] Salzo, S., Villa, S., Convergence analysis of a proximal Gauss-Newton method, *Comput. Optim. Appl.*, 53 (2012), 557–589.

[19] Smale, S., Newton's method estimates from data at one point. The merging of disciplines: new directions in pure, applied, and computational mathematics (Laramie, Wyo., 1985), 185-196, Springer, New York, 1986.

[20] Wang, X.H., Convergence of Newton's method and uniqueness of the solution of equations in Banach spaces, *IMA J. Numer. Anal.*, 20 (2000), 123-134.

Chapter 17

On the Convergence of a Damped Newton Method with Modified Right-Hand Side Vector

17.1. Introduction

In this chapter we are concerned with the problem of approximating a locally unique solution x^* of the nonlinear equation

$$F(x) = 0, \tag{17.1.1}$$

where F is a Fréchet-differentiable operator defined on a open convex subset D of a Banach space $\mathbb{X}$ with values in a Banach space $\mathbb{Y}$.

Many problems from Computational Sciences and other disciplines can be brought in a form similar to equation (17.1.1) using Mathematical Modeling [2, 6, 10]. For example in data fitting, we have $\mathbb{X} = \mathbb{Y} = \mathbb{R}^i$, i is number of parameters and i is number of observations.

The solution of (17.1.1) can rarely be found in closed form. That is why the solution methods for these equations are usually iterative. In particular, the practice of Numerical Analysis for finding such solutions is essentially connected to Newton-type methods [1]–[15]. The study about convergence matter of iterative procedures is usually centered on two types: semilocal and local convergence analysis. The semilocal convergence matter is, based on the information around an initial point, to give criteria ensuring the convergence of iteration procedures; while the local one is, based on the information around a solution, to find estimates of the radii of the convergence balls.

In the present chapter, we study the convergence of the Damped Newton method defined by

$$x_{n+1} = x_n - A^{-1}\left(I - \alpha_n\left(F'(x_n) - A\right)\right)F(x_n), \ \text{ for each } n = 0, 1, 2, \ldots, \tag{17.1.2}$$

where $A \in \mathcal{L}(\mathbb{X}, \mathbb{Y})$ the space of bounded linear operators from $\mathbb{X}$ into $\mathbb{Y}$, $A^{-1} \in \mathcal{L}(\mathbb{Y}, \mathbb{X})$, α_n is a sequence of real numbers chosen to force convergence of sequence x_n and x_0 is an initial point. If $A = F'(x_0)$ and $\alpha_n = 0$ for each $n = 0, 1, 2, \ldots$, we obtain the modified Newton's method

$$y_{n+1} = y_n - F'(x_0)^{-1}F(y_n),\ y_0 = x_0, \ \text{ for each } n = 0, 1, 2, \ldots, \tag{17.1.3}$$

which converges linearly [2, 10].

The local convergence of Newton-like method (17.1.2) was studied by Krejić and Lužanin [13] (see also [11]) in the case when $\mathbb{X} = \mathbb{Y} = \mathbb{R}^i$.

Newton's method

$$z_{n+1} = z_n - F'(z_n)F(z_n), \text{ for each } n = 0, 1, 2, \ldots, \tag{17.1.4}$$

converges quadratically provided that the iteration starts close enough to the solution. However, the cost of a Newton iterate may be very expensive, since all the elements of the Jacobian matrix involved must be computed, as well as the need for an exact slowdown of a system of linear equations using a new matrix for every iterate. As noted in [13] Newton-like method (17.1.2) uses a modification of the right hand side vector, which is cheaper than the Newton and faster than the modified Newton method. One step of the method requires the solution of a linear system, but the system matrix is the same in all iterations.

We present a new local and semilocal convergence analysis for Newton-like method. In contrast to the work in [11, 13], in the local case the radius of convergence can be computed as well as the error bounds on the distances $\|x_n - x^*\|$ for each $n = 0, 1, 2, \ldots$. In the semilocal case, we present estimates on the smallness of $\|F(x_0)\|$ as well as computable estimates for $\|x_n - x^*\|$ (not given in [11, 13] in terms of the Lipschitz constants and other initial data).

We denote by $U(\nu, \mu)$ the open ball centered at $\nu \in \mathbb{X}$ and of radius $\mu > 0$. Moreover, by $\overline{U(\nu, \mu)}$ we denote the closure of $U(\nu, \mu)$.

The chapter is organized as follows. Sections 17.2. and 17.3. contain the semilocal and local convergence analysis of Newton-like method (17.1.2), respectively. The numerical examples are presented in the concluding Section 17.4..

17.2. Semilocal Convergence

In this section we present the semilocal convergence of Damped Newton method (17.1.2). We shall use the following conditions:

C_0 $F : D \subseteq \mathbb{X} \to \mathbb{Y}$ is Fréchet-differentiable and there exists $A \in \mathcal{L}(\mathbb{X}, \mathbb{Y})$ such that $A^{-1} \in \mathcal{L}(\mathbb{Y}, \mathbb{X})$ with $\|A^{-1}\| \leq a$;

C_1 There exists $L > 0$ such that for each $x, y \in D$ the Lipschitz condition

$$\|F'(x) - F'(y)\| \leq L\|x - y\| \tag{17.2.1}$$

holds;

C_2 There exist $L_0 > 0$ such that for each $x \in D$ the center-Lipschitz condition

$$\|F'(x) - F'(x_0)\| \leq L_0\|x - x_0\| \tag{17.2.2}$$

holds;

C$_3$ There exist $x_0 \in D$, $\alpha \geq 0$, $a_0 \geq 0$, $a_1 \geq 0$, and $q \in (0,1)$ such that for $\|A^{-1}(F'(x_0)-A)\| \leq a_0$, $\|F'(x_0)-A\| \leq a_1$ the following hold

$$|\alpha_n| \leq \alpha,\ a+\alpha\left(\frac{aL_0q}{1-q}\|F(x_0)\|+a_0\right) \leq q \tag{17.2.3}$$

and

$$\frac{Lq^2}{2}\|F(x_0)\|+\left(\frac{L_0q}{1-q}\|F(x_0)\|+a_1\right)(\alpha+q) \leq q; \tag{17.2.4}$$

C$_4$ There exist $x_0 \in D$, $\alpha \geq 0$, $a_0 \geq 0$, $a_1 \geq 0$, and $q \in (0,1)$ such that for $\|A^{-1}(F'(x_0)-A)\| \leq a_0$, $\|F'(x_0)-A\| \leq a_1$ the inequality (17.2.3) and

$$\left(\frac{2}{1-q}+\frac{1}{2}\right)L_0q\|F(x_0)\|+\left(\frac{L_0q}{1-q}\|F(x_0)\|+a_1\right)(\alpha+q) \leq q \tag{17.2.5}$$

hold;

C$_5$ $\overline{U(x_0,r)} \subseteq D$ with $r = \frac{q\|F(x_0)\|}{1-q}$.

Notice that (1) implies (2),

$$L_0 \leq L \tag{17.2.6}$$

holds in general and $\frac{L}{L_0}$ can be arbitrarily large [2, 3, 6]. The conditions involving $\|F(x_0)\|$ and q in (3) and (4) can be solved for $\|F(x_0)\|$ and q. However, these representations are very long and unattractive. That is why we decided to leave these conditions as uncluttered as possible. Notice also that these conditions determine the smallness of $\|F(x_0)\|$ and q. From now on we shall denote (0), (1), (2), (3), (5) and (0),(2), (4), (5) as the (C) and (C^0) conditions, respectively. Next, we present the semilocal convergence of the Damped Newton-like method (17.1.2) first under the (C) conditions.

Theorem 17.2.1. *Suppose that the (C) conditions hold. Then sequence* $\{x_n\}$ *generated by the Damped Newton method* (17.1.2) *is well defined, remains in* $\overline{U(x_0,r)}$ *for each* $n = 0,1,2,\ldots$, *and converges to a solution* $x^* \in \overline{U(x_0,r)}$ *of equation* (17.1.1). *Moveover, the following estimates hold for each* $n = 0,1,2,\ldots$,

$$\|x_{n+1}-x_n\| \leq q\|F(x_n)\| \leq q^{n+1}\|F(x_0)\|, \tag{17.2.7}$$

and

$$\|F(x_{n+1})\| \leq q\|F(x_n)\| \leq q^{n+1}\|F(x_0)\|, \tag{17.2.8}$$

where q is defined in (3) *and r in* (5).

Proof. We have by (17.1.2) and $A^{-1} \in \mathcal{L}(\mathbb{Y},\mathbb{X})$ that sequence $\{x_n\}$ is well defined. Then, we shall show that $x_1 \in U(x_0,r)$, $\|x_1-x_0\| \leq q\|F(x_0)\|$ and $\|F(x_1)\| \leq q\|F(x_0)\|$. Indeed, we have by (17.1.2) for $n=0$ and the second condition in (3) that

$$\begin{aligned}
\|x_1-x_0\| &= \|A^{-1}\left(I-\alpha_0\left(F'(x_0)-A\right)\right)F(x_0)\| \\
&\leq \left[\|A^{-1}\|+|\alpha_0|\|A^{-1}\left(F'(x_0)-A\right)\right]\|F(x_0)\| \\
&\leq \left[\|A^{-1}\|+\alpha\|A^{-1}\left(F'(x_0)-A\right)\right]\|F(x_0)\| \\
&\leq q\|F(x_0)\| < r.
\end{aligned}$$

Hence, $x_1 \in U(x_0,r)$ and (17.2.7) holds for $n=0$. Using (17.1.2) it can easily be seen that the Ostrowski-type approximation

$$\begin{aligned} F(x_{n+1}) = &\int_0^1 \left[F'(x_n+\theta(x_{n+1}-x_n)) - F'(x_n)\right](x_{n+1}-x_n)\,\mathrm{d}\theta \\ &+ \left(F'(x_n)-A\right)(\alpha_n F(x_n) + (x_{n+1}-x_n)) \end{aligned} \tag{17.2.9}$$

holds. Using (17.2.9), and the (C) conditions, for $n=0$ we get in turn that

$$\begin{aligned} \|F(x_1)\| &= \Big\| \int_0^1 \left[F'(x_0+\theta(x_1-x_0)) - F'(x_0)\right](x_1-x_0)\,\mathrm{d}\theta \\ &\quad + \left(F'(x_0)-A\right)(\alpha_0 F(x_0) + (x_1-x_0)) \Big\| \\ &\le \frac{L_0}{2}\|x_1-x_0\|^2 + \left\|F'(x_0)-A\right\| \left(|\alpha_0|\|F'(x_0)\| + \|x_1-x_0\|\right) \\ &\le \frac{L}{2}q^2\|F(x_0)\|^2 + \left\|F'(x_0)-A\right\| (\alpha\|F(x_0)\| + q\|F(x_0)\|) \\ &\le \left[\frac{L}{2}q^2\|F(x_0)\| + \left\|F'(x_0)-A\right\|(\alpha+q)\right]\|F(x_0) \\ &\le q\|F(x_0)\|. \end{aligned}$$

That is (17.2.8) holds for $n=0$. It follows from the existence of $x_1 \in U(x_0,r)$ and $A^{-1} \in \mathcal{L}(\mathbb{X},\mathbb{Y})$ that x_2 is well defined. Using (17.1.2) for $n=1$, we get by (0), (2), (3) that

$$\begin{aligned} \|x_2-x_1\| &= \|A^{-1}\left(I-\alpha_1\left(F'(x_1)-A\right)\right)F(x_1)\| \\ &\le \left[\|A^{-1}\| + \alpha\|A^{-1}\left(\left(F'(x_1)-F'(x_0)\right) + \left(F'(x_0)-A\right)\right)\right]\|F(x_1)\| \\ &\le \left[\|A^{-1}\| + \alpha\left(\|A^{-1}\|L_0\|x_1-x_0\| + \left\|A^{-1}\left(F'(x_0)-A\right)\right\|\right)\right]\|F(x_1)\| \\ &\le q\|F(x_1)\| \le q^2\|F(x_0)\|. \end{aligned}$$

We also have that

$$\begin{aligned} \|x_2-x_0\| &\le \|x_2-x_1\| + \|x_1-x_0\| \\ &\le q^2\|F(x_0)\| + q\|F(x_0)\| \\ &= q\|F(x_0)\|(1+q)\| \\ &= q\|F(x_0)\|\frac{1-q^2}{1-q} \\ &< \frac{q\|F(x_0)\|}{1-q} = r. \end{aligned} \tag{17.2.10}$$

That is, $x_2 \in U(x_0, r)$. Then, using (17.2.9) for $n = 1$, as above we get in turn that

$$\begin{aligned}
\|F(x_2)\| &\le \frac{L}{2}\|x_2 - x_1\|^2 \\
&\quad + \left(L_0\|x_1 - x_0\| + \left\|F'(x_0) - A\right\|\right)\left(|\alpha_0|\|F'(x_1)\| + \|x_2 - x_1\|\right) \\
&\le \frac{L}{2}q^2\|F'(x_1)\|^2 \\
&\quad + \left(L_0 q\|F(x_0)\| + \left\|F'(x_0) - A\right\|\right)(\alpha\|F(x_1)\| + q\|F(x_1)\|) \\
&\le \left[\frac{L}{2}q^2\|F'(x_1)\| + \left(L_0 q\|F(x_0)\| + \left\|F'(x_0) - A\right\|\right)(\alpha + q)\right]\|F(x_1)\| \\
&\le q\|F(x_1)\| \le q^2\|F(x_0)\|.
\end{aligned}$$

Similarly, we have using (17.1.2) that

$$\begin{aligned}
\|x_3 - x_2\| &\le \left[\|A^{-1}\| + \alpha\|A^{-1}\left(\left(F'(x_2) - F'(x_0)\right) + \left(F'(x_0) - A\right)\right)\right]\|F(x_2)\| \\
&\le \left[\|A^{-1}\| + \alpha\left(\|A^{-1}\|L_0\|x_2 - x_0\| + \left\|A^{-1}\left(F'(x_0) - A\right)\right\|\right)\right]\|F(x_2)\| \\
&\le q\|F(x_2)\| \le q^3\|F(x_0)\|.
\end{aligned}$$

We also have that

$$\begin{aligned}
\|x_3 - x_0\| &\le \|x_3 - x_2\| + \|x_2 - x_1\| + \|x_1 - x_0\| \\
&\le (q^3 + q^2 + q)\|F(x_0)\| \\
&= q\|F(x_0)\|\frac{1 - q^3}{1 - q} < r,
\end{aligned}$$

and

$$\begin{aligned}
\|F(x_3)\| &\le \frac{L}{2}\|x_3 - x_2\|^2 \\
&\quad + \left(L_0\|x_2 - x_0\| + \left\|F'(x_0) - A\right\|\right)\left(|\alpha_0|\|F'(x_1)\| + \|x_3 - x_2\|\right) \\
&\le \frac{L}{2}q^2\|F(x_2)\|^2 \\
&\quad + \left(L_0\frac{q\|F(x_0)\|}{1 - q} + \left\|F'(x_0) - A\right\|\right)(\alpha\|F(x_2)\| + q\|F(x_2)\|) \\
&\le \left[\frac{L}{2}q^2\|F(x_2)\| + \left(L_0\frac{q\|F(x_0)\|}{1 - q} + \left\|F'(x_0) - A\right\|\right)(\alpha + q)\right]\|F(x_2)\| \\
&\le q\|F(x_2)\| \le q^3\|F(x_0)\|.
\end{aligned}$$

The rest follows in analogous way using induction (simply replace x_2, x_3 by x_n, x_{n+1} in the above estimates). By letting $n \to \infty$ in (17.2.7) we obtain $F(x^*) = 0$. □

Condition (1) may not be satisfied but weaker condition (2) may be satisfied. In this case (1) can be dropped. Then, using instead of approximation (17.2.9) the approximation

$$\begin{aligned}
F(x_{n+1}) &= \int_0^1 \left[F'(x_n + \theta(x_{n+1} - x_0)) - F'(x_0)\right](x_{n+1} - x_n)\,\mathrm{d}\theta \\
&\quad + \left(F'(x_0) - F'(x_n)\right)(x_{n+1} - x_n) \\
&\quad + \left[\left(F'(x_n) - F'(x_0)\right) + \left(F'(x_0) - A\right)\right]\left(\alpha_n F(x_n) + (x_{n+1} - x_n)\right),
\end{aligned} \tag{17.2.11}$$

we arrive in an analogous way to Theorem 17.2.1 at the following semilocal convergence result for the Damped Newton method (17.1.2) under the (C^0) conditions.

Theorem 17.2.2. *Suppose that the (C^0) conditions hold. Then sequence $\{x_n\}$ generated by the Damped Newton method* (17.1.2) *is well defined, remains in $\overline{U(x_0,r)}$ for each $n = 0,1,2,\ldots$, and converges to a solution $x^* \in \overline{U(x_0,r)}$ of equation* (17.1.1). *Moveover, the following estimates hold for each $n = 0,1,2,\ldots$,*

$$\|x_{n+1} - x_n\| \leq q\|F(x_n)\| \leq q^{n+1}\|F(x_0)\|,$$

and

$$\|F(x_{n+1})\| \leq q\|F(x_n)\| \leq q^{n+1}\|F(x_0)\|,$$

where q is defined in (4) *and r in* (5).

Concerning the uniqueness of the solution x^* in $\overline{U(x_0,r)}$ we have the following result.

Proposition 17.2.3. *Suppose that the (C) or (C^0) conditions hold. Moveover, suppose that there exist $x_0 \in D$ and $r_1 \geq r$ such that $F'(x_0)^{-1} \in \mathcal{L}(\mathbb{X},\mathbb{Y})$ and*

$$F'(x_0)^{-1}L_0(r_1 + r) < 2. \tag{17.2.12}$$

Then the solution x^* is the only solution of equation (17.1.1) *in $\overline{U(x_0,r_1)}$, where r is defined in* (5).

Proof. The existence of the solution x^* is guaranteed by conditions (C) or (C^0). To show uniqueness, let $y^* \in \overline{U(x_0,r_1)}$ with $F(y^*) = 0$. Define $\mathcal{M} = \int_0^1 F'(x^* + \theta(y^* - x^*))\, d\theta$. Then, using (2) and (17.2.12) we obtain in turn that

$$\begin{aligned}
\|F'(x_0)^{-1}\|\|\mathcal{M} - F'(x_0)\| &\leq \|F'(x_0)^{-1}\|L_0 \int_0^1 \|(x^* - x_0) + \theta(y^* - x^*)\|\, d\theta \\
&\leq \|F'(x_0)^{-1}\|L_0 \int_0^1 \|(1-\theta)(x^* - x_0) + \theta(y^* - x_0)\|\, d\theta \\
&\leq \|F'(x_0)^{-1}\|\frac{L_0}{2}(r + r_1) < 1. \qquad (17.2.13)
\end{aligned}$$

It follows from (17.2.13) and the Banach lemma on invertible operator [10] that $\mathcal{M}^{-1} \in \mathcal{L}(\mathbb{Y},\mathbb{X})$. Moreover, we have that $0 = F(y^*) - F(x^*) = \mathcal{M}(y^* - x^*)$, which implies $x^* = y^*$. □

17.3. Local Convergence

In this section we present the local convergence of Newton-like method(17.1.2). We shall use the following conditions:

- C_0 $F : D \subseteq \mathbb{X} \to \mathbb{Y}$ is Fréchet-differentiable and there exists $A \in \mathcal{L}(\mathbb{X},\mathbb{Y})$, $x^* \in D$ such that $A^{-1} \in \mathcal{L}(\mathbb{Y},\mathbb{X})$, $F(x^*) = 0$ with $\|A^{-1}\| \leq a$ and $\|F'(x^*)\| \leq \beta$;
- C_1 There exist $L > 0$ such that for each $x,y \in D$ the Lipschitz condition (17.2.1) holds;

C_2 There exists $l_0 > 0$ such that for each $x \in D$ the center-Lipschitz condition(17.2.2)

$$\|F'(x) - F'(x^*)\| \leq l_0 \|x - x^*\|$$

holds;

C_3 Let $\|A^{-1}(F'(x^*) - A)\| \leq \beta_1$.

$$|\alpha_n| \leq \alpha, \beta_1(1 + \alpha\beta) < 1$$

Denote by R_1 the positive root of quadratic polynomial

$$p_1(t) = \frac{\alpha l_0^2}{2} a t^2 + \left(\frac{La}{2} + \frac{\alpha l_0 \beta_1}{2} + l_0 a + \alpha l_0 a \beta\right) t + \beta_1(1 + \alpha\beta) - 1; \qquad (17.3.1)$$

Moreover, denote by R_2 the positive root of quadratic polynomial

$$p_2(t) = \frac{a\alpha l_0^2}{2} t^2 + \left(\frac{3al_0}{2} + \frac{\alpha l_0 \beta_1}{2} + l_0 a + \alpha l_0 a \beta\right) t + \beta_1(1 + \alpha\beta) - 1; \qquad (17.3.2)$$

C_4 $\overline{U(x^*, R)} \subseteq D$, where R is R_1 or R_2.

Notice that (1) implies (2),

$$l_0 \leq L \qquad (17.3.3)$$

holds in general and $\frac{L}{l_0}$ can be arbitrarily large [2, 3, 6]. The quadratic polynomials in (3) and (4) have a positive root by the second hypothesis in (3) or (4) and since the coefficients of t and t^2 are positive. From now on we shall denote (0), (1), (2), (3), (4) and (0),(2), (4) as the (H) and (H^0) conditions, respectively. Next, we present the local convergence of Newton-like method (17.1.2) first under the (H) conditions. In view of (17.1.2) and $F(x^*) = 0$, we can have the following identity

$$\begin{aligned} x_{n+1} - x^* = -A^{-1}\Big\{ &\int_0^1 \left[F'(x^* + \theta(x_n - x^*)) - F'(x_n)\right] d\theta \\ &- \left((A - F'(x^*)) + (F'(x^*) - F'(x_n))\right) \Big[(I - \alpha_n F'(x^*)) \\ &- \alpha_n \int_0^1 \left[F'(x^* + \theta(x_n - x^*)) - F'(x^*)\right]\Big]\Big\} (x_n - x^*) \end{aligned} \qquad (17.3.4)$$

Then, using (17.2.9), and the (H) conditions, it is standard to arrive at [2, 3, 4, 5, 6, 8, 9, 10, 14, 15]:

Theorem 17.3.1. *Suppose that the (H) conditions hold. Then sequence $\{x_n\}$ generated by the Damped Newton method* (17.1.2) *is well defined, remains in $\overline{U(x^*, R_1)}$ for each $n = 0, 1, 2, \ldots$, and converges to x^* provided that $x_0 \in \overline{U(x^*, R_1)}$. Moveover, the following estimates hold for each $n = 0, 1, 2, \ldots$,*

$$\|x_{n+1} - x^*\| \leq e_n \|x_n - x^*\| < \|x_n - x^*\| < R_1, \qquad (17.3.5)$$

where

$$\begin{aligned} e_n &= \frac{La}{2}\|x_n - x^*\| + \beta_1 + \alpha\beta_1\beta + \frac{\beta_1\alpha l_0}{2}\|x_n - x^*\| + l_0 a\|x_n - x^*\| \\ &\quad + \alpha l_0 a\beta\|x_n - x^*\| + \frac{\alpha l_0^2 a}{2}\|x_n - x^*\|^2 \\ &< p_1(R_1) + 1 < 1. \end{aligned}$$

In cases (1) cannot be verified by (2) holds, we can present the local convergence of the Damped Newton method (17.1.2) under the (H^0) conditions using the following modification of the Ostrowski representation (17.3.4) given by

$$\begin{aligned} x_{n+1} - x^* = -A^{-1}\Big\{ &\int_0^1 \left[F'\left(x^* + \theta(x_n - x^*)\right) - F'(x^*)\right] \mathrm{d}\theta \\ &+ [F'(x^*) - F'(x_n)] \\ &- \big((A - F'(x^*)) + (F'(x^*) - F'(x_n))\big)\Big[(I - \alpha_n F'(x^*) \\ &- \alpha_n \int_0^1 \left[F'\left(x^* + \theta(x_n - x^*)\right) - F'(x^*)\right]\Big]\Big\}(x_n - x^*) \end{aligned} \tag{17.3.6}$$

Theorem 17.3.2. *Suppose that the (H^0) conditions hold. Then sequence $\{x_n\}$ generated by the Damped Newton method* (17.1.2) *is well defined, remains in $\overline{U(x^*,R_2)}$ for each $n = 0,1,2,\ldots$, and converges to x^* provided that $x_0 \in \overline{U(x^*,R_2)}$. Moveover, the following estimates hold for each $n = 0,1,2,\ldots$,*

$$\|x_{n+1} - x^*\| \le e_n^0\|x_n - x^*\| < \|x_n - x^*\| < R_2, \tag{17.3.7}$$

where

$$\begin{aligned} e_n^0 &= \frac{3l_0 a}{2}\|x_n - x^*\| + \beta_1 + \alpha\beta_1\beta + \frac{\beta_1\alpha l_0}{2}\|x_n - x^*\| + l_0 a\|x_n - x^*\| \\ &\quad + \alpha l_0 a\beta\|x_n - x^*\| + \frac{\alpha l_0^2 a}{2}\|x_n - x^*\|^2 \\ &< p_2(R_2) + 1 < 1. \end{aligned}$$

17.4. Numerical Examples

Example 17.4.1. *In this example we present an application of the previous analysis to the Chandrasekhar equation:*

$$x(s) = 1 + \frac{s}{4}x(s)\int_0^1 \frac{x(t)}{s+t}dt, \quad s \in [0,1], \tag{17.4.1}$$

which arises in the theory of radiative transfer [7]; $x(s)$ is the unknown function which is sought in $C[0,1]$. The physical background of this equation is fairly elaborate. It was developed by Chandraseckhar [7] to solve the problem of determination of the angular distribution of the radiant flux emerging from a plane radiation field. This radiation field *must be isotropic at a point, that is the distribution in independent of direction at that point. Explicit definitions of these terms may be found in the literature [7]. It is considered to be the prototype of the equation,*

$$x(s) = 1 + \lambda s x(s) \int_0^1 \frac{\varphi(s)}{s+t} x(t)\, dt, \quad s \in [0,1],$$

for more general laws of scattering, where $\varphi(s)$ is an even polynomial in s with

$$\int_0^1 \varphi(s)\, ds \leq \frac{1}{2}.$$

Integral equations of the above form also arise in the other studies [7]. We determine where a solution is located, along with its region of uniqueness.

Note that solving (17.4.1) *is equivalent to solve $F(x) = 0$, where $F : C[0,1] \to C[0,1]$ and*

$$[F(x)](s) = x(s) - 1 - \frac{s}{4} x(s) \int_0^1 \frac{x(t)}{s+t} dt, \quad s \in [0,1]. \tag{17.4.2}$$

To obtain a numerical solution of (17.4.1)*, we first discretize the problem and approach the integral by a Gauss-Legendre numerical quadrature with eight nodes,*

$$\int_0^1 f(t)\, dt \approx \sum_{j=1}^{8} w_j f(t_j),$$

where

$$\begin{aligned}
&t_1 = 0.019855072, \quad t_2 = 0.101666761, \quad t_3 = 0.237233795, \quad t_4 = 0.408282679,\\
&t_5 = 0.591717321, \quad t_6 = 0.762766205, \quad t_7 = 0.898333239, \quad t_8 = 0.980144928,\\
&w_1 = 0.050614268, \quad w_2 = 0.111190517, \quad w_3 = 0.156853323, \quad w_4 = 0.181341892,\\
&w_5 = 0.181341892, \quad w_6 = 0.156853323, \quad w_7 = 0.111190517, \quad w_8 = 0.050614268.
\end{aligned}$$

If we denote $x_i = x(t_i)$, $i = 1,2,\ldots,8$, equation (3.7) is transformed into the following non-linear system:

$$x_i = 1 + \frac{x_i}{4} \sum_{j=1}^{8} a_{ij} x_j, \quad i = 1,2,\ldots,8,$$

where, $a_{ij} = \dfrac{t_i w_j}{t_i + t_j}$.

Denote now $\overline{x} = (x_1, x_2, \ldots, x_8)^T$, $\overline{1} = (1,1,\ldots,1)^T$, $A = (a_{ij})$ and write the last nonlinear system in the matrix form:

$$\overline{x} = \overline{1} + \frac{1}{4} \overline{x} \odot (A\overline{x}), \tag{17.4.3}$$

where $\odot$ represents the inner product. Set $G(x) = x$. If we choose $\overline{x}_0 = (1,1,\ldots,1)^T$ and $\overline{x}_{-1} = (0,0,\ldots,0)^T$. Assume sequence $\{\overline{x}_n\}$ is generated with different choices of α_n and $A = F'(x_0)$. The computational order of convergence (COC) is shown in Table 17.4.1 for various methods. Here (COC) is defined in [12] by

$$\rho \approx \ln\left(\frac{\|\overline{x}_{n+1} - \overline{x}^\star\|_\infty}{\|\overline{x}_n - \overline{x}^\star\|_\infty}\right) / \ln\left(\frac{\|\overline{x}_n - \overline{x}^\star\|_\infty}{\|\overline{x}_{n-1} - \overline{x}^\star\|_\infty}\right), \qquad n \in \mathbb{N},$$

The Table 17.4.1 shows the (COC).

Table 17.4.1. The comparison results of the COC for Example 1 using various α_n

n	$\alpha_n = 0$	$\alpha_n = 0.0001$	$\alpha_n = 0.001$	$\alpha_n = 0.01$	$\alpha_n = 0.1$	$\alpha_n = 1$
ρ	1.0183391	1.0645848	1.0645952	1.0646989	1.0657398	1.0764689

Example 17.4.2. *In this example, we consider the Singular Broyden [13] problem defined as*

$$F_1(x) = ((3 - hx_1)x_1 - 2x_2 + 1)^2,$$
$$F_i(x) = ((3 - hx_i)x_i - x_{i-1} - 2x_{i+1} + 1)^2,$$
$$F_n(x) = ((3 - hx_n)x_n - x_{n-1} + 1)^2,$$

Taking as starting approximation $x_0 = (-1, \ldots, -1)^T$ and $h = 2$. The Table 17.4.2 shows the (COC) computed as in previous example.

Table 17.4.2. The comparison results of the COC for Example 2 using various α_n

n	$\alpha_n = 0$	$\alpha_n = 0.01$	$\alpha_n = 0.02$	$\alpha_n = 0.03$	$\alpha_n = 0.04$	$\alpha_n = 05$
ρ	1.7039443	1.7041146	1.7048251	1.7178472	1.5650132	1.6619946

Example 17.4.3. *Let $X = Y = \mathbb{R}^2$, $D = \overline{U}(1,1)$ and $x_0 = (1, 0.5)$. Define function F on D for $w = (x,y)$ by*

$$F(w) = (x^3 - 3xy^2 - 1, 3x^2y - y^3). \tag{17.4.4}$$

Then, the Fréchet derivative of F is given by

$$F'(w) = \begin{pmatrix} 3\left(x^2 - y^2\right) & -6xy \\ 6xy & 3x^2 \end{pmatrix}$$

Moreover we see in Figure 17.4.1 the number of iterations needed to arrive at the solution with 300 digits, starting in $x_0 = \{1, 0.5\}$

Example 17.4.4. *Let $X = Y = \mathbb{R}^3$, $D = \overline{U}(0,1)$ and $x^* = (0,0,0)$. Define function F on D for $w = (x,y,z)$ by*

$$F(w) = (e^x - 1, \frac{e-1}{2} y^2 + y, z). \tag{17.4.5}$$

Then, the Fréchet derivative of F is given by

$$F'(w) = \begin{pmatrix} e^x & 0 & 0 \\ 0 & (e-1)y + 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Notice that we have $F(x^*) = 0$, $F'(x^*) = F'(x^*)^{-1} = \operatorname{diag}\{1,1,1\}$

Moreover we see in Figure 17.4.2 the number of iterations needed to arrive at the solution with 300 digits, starting

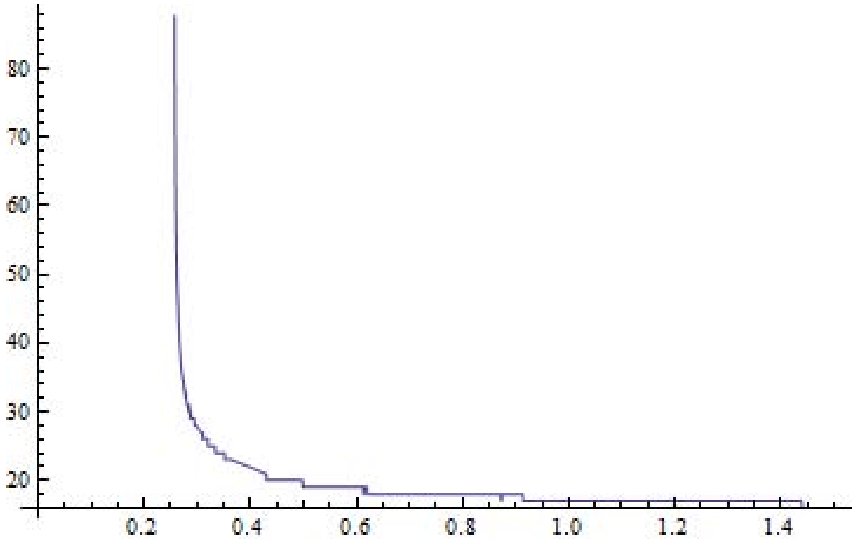


Figure 17.4.1. Number of iterations needed.

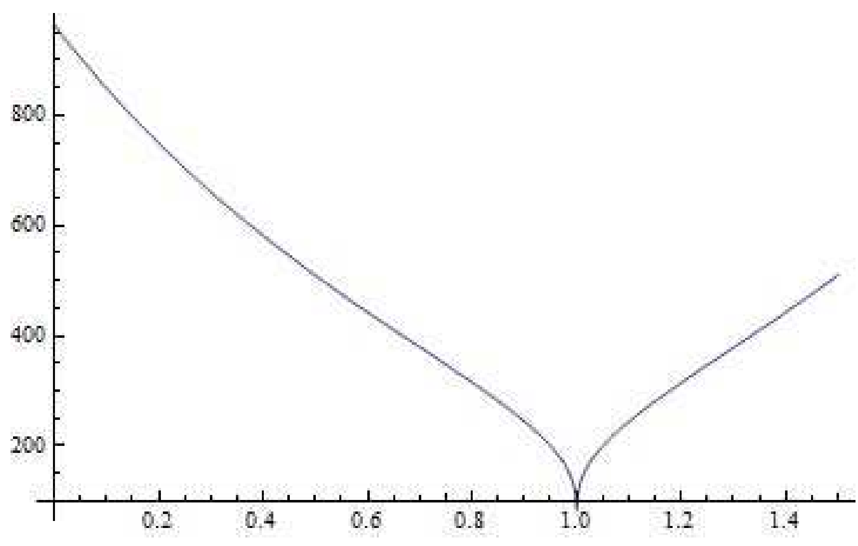


Figure 17.4.2. Number of iterations needed.

References

[1] Amat, S., Busquier, S., Gutiérrez, J. M., Geometric constructions of iterative functions to solve nonlinear equations, *J. Comput. Appl. Math.*, 157 (2003), 197-205.

[2] Argyros, I.K., *Convergence and applications of Newton–type iterations*, Springer–Verlag Publ., New York, 2008.

[3] Argyros, I. K., Chen, J., Improved results on estimating and extending the radius of an attraction ball, *Appl. Math. Lett.*, 23 (2010), 404-408.

[4] Argyros, I. K., Chen, J., On local convergence of a Newton-type method in Banach space, *Int. J. Comput. Math.*, 86 (2009), 1366-1374.

[5] Argyros, I.K., Hilout, S., Improved local convergence of Newton's method under weak majorant condition, *J. Comput. Appl. Math.*, 236 (2012), 1892–1902.

[6] Argyros, I.K., S. Hilout, *Computational methods in Nonlinear Analysis*, World Scientific Publ. Comp. New Jersey, 2013.

[7] Chandrasekhar, S., *Radiative transfer*, Dover Publ., New York, 1960.

[8] Chen, J., Sun, Q., The convergence ball of Newton-like methods in Banach space and applications, *Taiwanese J. Math.*, 11 (2007), 383-397.

[9] Chen, J, Li, W., Convergence behaviour of inexact Newton methods under weak Lipschitz condition, *J. Comput. Appl. Math.*, 191 (2006), 143-164.

[10] Kantorovich, L.V., Akilov, G.P., *Functional Analysis*, Pergamon Press, Oxford, 1982.

[11] Herceg, D., Krejić, N., Lužanin, Z., Quasi-Newton's method with correction, *Novi Sad J. Math.*, 26 (1996), 115-127.

[12] Grau, M., Noguera, M., Gutiérrez, J. M., On some computational orders of convergence. *Appl. Math. Let.*, 23 (2010), 472-478.

[13] Krejić, N., Lužanin, Z., Newton-like method with modification of the right-hand-side vector, *Math. Comp.* 71 (2002), 237-250

[14] Potra, F.A., Sharp error bounds for a class of Newton–like methods, *Libertas Mathematica*, 5 (1985), 71-84.

[15] Proinov, P.D., General local convergence theory for a class of iterative processes and its applications to Newton's method, *J. Complexity*, 25 (2009), 38–62.

Chapter 18

Local Convergence of Inexact Newton-Like Method under Weak Lipschitz Conditions

18.1. Introduction

Let $\mathcal{X}$, $\mathcal{Y}$ be Banach spaces and $\mathcal{D}$ be a non-empty, convex and open subset in $\mathcal{X}$. Let $U(x,r)$ and $\overline{U}(x,r)$ stand, respectively, for the open and closed ball in $\mathcal{X}$ with center x and radius $r > 0$. Denote by $\mathcal{L}(\mathcal{X},\mathcal{Y})$ the space of bounded linear operators from $\mathcal{X}$ into $\mathcal{Y}$. In this chapter, we are concerned with the problem of approximating a solution $x^{\star}$ of equation

$$F(x) = 0 \tag{18.1.1}$$

where F is a Fréchet continuously differentiable operator defined on $\mathcal{D}$ with values in $\mathcal{Y}$. Many problems from computational sciences and other disciplines can be brought in the form of equation 18.1.1 using Mathematical Modelling [1, 3, 6, 7, 9, 12]. The solution of these equations can rarely be found in closed form. That is why the solution methods for these equations are iterative. In particular, the practice of numerical analysis for finding such solutions is essentially connected to variants of Newton's method [1]-[14]. The study about convergence matter of iterative procedures is usually centered on two types: semilocal and local convergence analysis. The semilocal convergence matter is, based on the information around an initial point, to give criteria ensuring the convergence of iterative procedure; while the local one is, based on the information around a solution, to find estimates of the radii of convergence balls. There is a plethora of studies on the weakness and/or extension of the hypothesis made on the underlying operators; see for example [1]-[14].

Undoubtedly the most popular iterative method, for generating a sequence approximating $x^{\star}$, is the Newton's method (NM) which is defined as

$$x_{n+1} = x_n - F'(x_n)^{-1}F(x_n) \quad \text{for each} \quad n = 0,1,2,\ldots \tag{18.1.2}$$

where x_0 is an initial point. There are two difficulties with the implementation of (NM). The first is to evaluate F' and the second difficulty is to exactly solve the following Newton equation

$$F'(x_n)(x_{n+1}-x_n) = -F(x_n) \quad \text{for each} \quad n = 0,1,2,\ldots. \tag{18.1.3}$$

It is well-known that evaluating F' and solving equation (18.1.3) may be computationally expensive [1, 5, 6, 8, 12, 13, 14]. That is why inexact Newton method (INLM) has been used [1, 2, 8, 9, 12, 13, 14]:

For $n = 0$ step 1 until convergence do
Find the step Δ_n which satisfies

$$B_n\Delta_n = -F(x_n) + r_n, \quad \text{where} \quad \frac{\|P_n r_n\|}{\|P_n F(x_n)\|} \le \eta_n \tag{18.1.4}$$

Set $x_{n+1} = x_n + \Delta_n$ where P_n is an invertible operator for each $n = 0,1,2,\cdots$. Here, $\{r_n\}$ is a null-sequence in the Banach space $\mathcal{Y}$. Clearly, the convergence behavior of (INLM) depends on the residual controls of $\{r_n\}$ and hypotheses on F'. In particular, Lipschitz continuity conditions on F' have been used and residual controls of the form

$$\begin{aligned}
\|r_n\| &\le \eta_n \|F(x_n)\|,\\
\|F'(x^\star)^{-1} r_n\| &\le \eta_n \|F'(x^\star)^{-1} F(x_n)\|,\\
\|F'(x^\star)^{-1} r_n\| &\le \eta_n \|F'(x^\star)^{-1} F(x_n)\|^{1+\theta},\\
\|P_n r_n\| &\le \theta_n \|P_n F(x_n)\|^{1+\theta},
\end{aligned} \tag{18.1.5}$$

for some $\theta \in [0,1]$ and for each $n = 0,1,2,\ldots$, have been employed. Here, $\{\eta_n\}$, $\{\theta_n\}$ are sequences in $[0,1]$, $\{P_n\}$ is a sequence in $\mathcal{L}(\mathcal{Y},\mathcal{X})$ and $F'(x^\star)^{-1}F'$ satisfies a Lipschitz or Hölder condition on $U(x^\star, r)$ [1]-[6], [8, 9, 10, 13, 14].

In this chapter, we are motivated by the works of Argyros et al.[1, 2], Chen et al.[5] and Zhang et al.[13] and optimization considerations. We suppose that F has a continuous Fréchet-derivative in $\overline{U}(x^\star, r)$, $F(x^\star) = 0$, $F'(x^\star)^{-1}F'$ exists and $F'(x^\star)^{-1}F'$ satisfies the Lipschitz with $L-$average radius condition

$$\|F'(x^\star)^{-1}(F'(x) - F'(x^\tau))\| \le \int_{\tau\rho(x)}^{\rho(x)} \mathcal{L}(u)\,du \tag{18.1.6}$$

for each $x \in U(x^\star, r)$. Here, $\rho(x) = \|x - x^\star\|$, $x^\tau = x^\star + \tau(x - x^\star)$, $\tau \in [0,1]$ and $\mathcal{L}$ is a monotone function on $[0,r]$. Condition (18.1.6) was inagurated by Wang in [14].

In view of (18.1.6) there exists a monotone function $\mathcal{L}_0$ on $[0,r]$ such that the center Lipschitz with $\mathcal{L}_0-$average condition

$$\|F'(x^\star)^{-1}(F'(x) - F'(x^\star))\| \le \int_0^{\rho(x)} \mathcal{L}_0(u)\,du \tag{18.1.7}$$

holds for each $x \in U(x^\star, r)$. Clearly, we have

$$\mathcal{L}_0(u) \le \mathcal{L}(u) \tag{18.1.8}$$

for each $u \in [0,r]$ and $\mathcal{L}/\mathcal{L}_0$ can be arbitrarily large [1, 2, 4] (see also the numerical example at the end of the chapter). It is worth noticing that (18.1.7) is not an additional to

(18.1.6) hypothesis, since in practice the convergence of (18.1.6) requires the computation of (18.1.7).

In the computation of $\|F(x)^{-1}F'(x^\star)\|$ we use the condition (18.1.7) which is tighter than 18.1.6, and the Banach lemma on invertible operators [7], to obtain the perturbation bound

$$\|F'(x)^{-1}F'(x^\star)\| \leq \left(1-\int_0^{\rho(x)} \mathcal{L}_0(u)\,du\right)^{-1} \quad \text{for each} \quad x \in U(x^\star, r), \tag{18.1.9}$$

instead of using 18.1.6 to obtain

$$\|F'(x)^{-1}F'(x^\star)\| \leq \left(1-\int_0^{\rho(x)} \mathcal{L}(u)\,du\right)^{-1} \quad \text{for each} \quad x \in U(x^\star, r). \tag{18.1.10}$$

Notice that (18.1.6) and (18.1.10) have been used in [5], [13], [14]. It turns out that using (18.1.9) instead of (18.1.10), in the case when $\mathcal{L}_0(u) < \mathcal{L}(u)$ for each $u \in [0, r]$, leads to tighter majorizing sequences for (INLM). This observation in turn leads to the following advantages over the earlier works (for $\eta_n = 0$ for each $n = 0, 1, 2, \ldots$ or not and $\mathcal{L}$ being a constant or not):

1. Larger radius of convergence.
2. Tighter error estimates on the distances $\|x_{n+1} - x_n\|$, $\|x_n - x^\star\|$ for each $n = 0, 1, 2, \ldots$.
3. Fewer iteration to achieve a desired error tolerance.

The rest of the chapter is organized as follows. In Section 18.2 we present some auxiliary results. Section 18.3 contains the local convergence analysis of (INLM). In Section 18.4, we present special cases. The numerical example appears in Section 18.5 and the conclusion in Section 18.6.

18.2. Background

In this section we present three auxiliary results. The first two are Banach-type perturbation lemmas.

Lemma 18.2.1. *Suppose that F is such that F' is continuously Fréchet- differentiable in $U(x^\star, r)$, $F'(x^\star)^{-1} \in L(\mathcal{Y}, \mathcal{X})$ and $F'(x^\star)^{-1}F'$ satisfies the center-Lipschitz condition with $\mathcal{L}_0-$average. Let r satisfy*

$$\int_0^r \mathcal{L}_0(u)du \leq 1. \tag{18.2.1}$$

Then, for each $x \in U(x^\star, r)$, $F'(x)$ is invertible and

$$\|F'(x)^{-1}F'(x^\star)\| \leq \frac{1}{1-\int_0^{\rho(x)} \mathcal{L}_0(u)du}. \tag{18.2.2}$$

Proof. Let $x \in U(x^\star, r)$. Using (18.1.7) and (18.2.1) we get in turn that

$$\|F'(x^\star)^{-1}(F'(x)-F'(x^\star))\| \le \int_0^{\rho(x)} \mathcal{L}_0(u)du < \int_0^{r} \mathcal{L}_0(u)du \le 1. \tag{18.2.3}$$

It follows from (18.2.3) and the Banach Lemma on invertible operators [7] that $F'(x)^{-1} \in L(\mathcal{Y},\mathcal{X})$ and (18.2.2) holds.

Lemma 18.2.2. *Suppose that F is such that F' is continuously Fréchet-differentiable in $U(x^\star,r)$, $F'(x^\star)^{-1} \in L(\mathcal{Y},\mathcal{X})$ and $F'(x^\star)^{-1}F'$ satisfies the radius Lipschitz condition with $\mathcal{L}-$average and the center-Lipschitz condition with $\mathcal{L}_0-$average. Then, we have*

$$\begin{aligned}\|F'(x)^{-1}F(x)\| &\le \rho(x) + \frac{\int_0^{\rho(x)} \mathcal{L}(u)udu - \int_0^{\rho(x)}(\mathcal{L}(u)-\mathcal{L}_0(u))\rho(x)du}{1-\int_0^{\rho(x)}\mathcal{L}_0(u)du} && (18.2.4)\\ &\le \rho(x) + \frac{\int_0^{\rho(x)} \mathcal{L}(u)udu}{1-\int_0^{\rho(x)}\mathcal{L}_0(u)du}. && (18.2.5)\end{aligned}$$

If $F'(x^\star)^{-1}F'$ satisfies the center-Lipschitz condition, then we have

$$\|F'(y)^{-1}F(x)\| \le \frac{\rho(x)+\int_0^{\rho(x)}\mathcal{L}_0(u)(\rho(x)-u)\rho(x)du}{1-\int_0^{\rho(y)}\mathcal{L}_0(u)du} \tag{18.2.6}$$

Proof. Let $x \in U(x^\star, r)$. We have that

$$\|F'(x)^{-1}F(x)\| \le \|F'(x)^{-1}F(x^\star)\|\|F'(x^\star)^{-1}F(x)\|. \tag{18.2.7}$$

But in view of (18.2.2) and the estimate

$$\|F'(x)^{-1}F(x)\| \le \rho(x) + \int_0^{\rho(x)} \mathcal{L}(u)(u-\rho(x))du \tag{18.2.8}$$

shown in [5, Lemma 2.1, 1.3], we obtain that

$$\|F'(x)^{-1}F(x)\| \le \frac{\rho(x)+\int_0^{\rho(x)}\mathcal{L}(u)(u-\rho(x))\rho(x)du}{1-\int_0^{\rho(x)}\mathcal{L}_0(u)du}$$

which implies (18.2.4) and since $\mathcal{L}_0(u) \le \mathcal{L}(u)$ (18.2.4) implies (18.2.5). Estimate (18.2.6) is shown in [5, Lemma 2.2, 1.3].

Remark 18.2.3. *If $\mathcal{L}_0 = \mathcal{L}$, then our two preceeding results are reduced to the corresponding ones in [5, 13]. Otherwise, i.e., if strict inequality holds in (18.1.8), then our estimates are more precise, since*

$$\frac{1}{1-\int_0^{\rho(x)}\mathcal{L}_0(u)du} < \frac{1}{1-\int_0^{\rho(x)}\mathcal{L}(u)du} \tag{18.2.9}$$

and

$$\rho(x) + \frac{\int_0^{\rho(x)}\mathcal{L}(u)udu}{1-\int_0^{\rho(x)}\mathcal{L}_0(u)du} < \rho(x) + \frac{\int_0^{\rho(x)}\mathcal{L}(u)udu}{1-\int_0^{\rho(x)}\mathcal{L}(u)du}. \tag{18.2.10}$$

Notice that the right hand sides of (18.2.9) and (18.2.10) are the upper bounds of the norms $\|F'(x)^{-1}F(x^\star)\|$, $\|F'(x)^{-1}F(x)\|$, *respectively obtained in the corresponding Lemmas in [5], [13].*

It turns out that in view of estimates (18.2.9) and (18.2.10), we obtain the advantages already mentioned in the introduction of this chapter of our approach over the corresponding ones in [5, 13, 14].

Next, we present another auxiliary result due to Wang [14, Lemma 2.2].

Lemma 18.2.4. *Suppose that the function* $\mathcal{L}_\alpha$ *defined by*

$$\mathcal{L}_\alpha(t) := t^{1-\alpha}\mathcal{L}(t) \tag{18.2.11}$$

is nondecreasing for some α *with* $\alpha \in [0,1]$, *where* $\mathcal{L}$ *is a positive integrable function. Then, for each* $\beta \geq 0$, *the function* $\varphi_{\beta,\alpha}$ *defined by*

$$\varphi_{\beta,\alpha} = \frac{1}{t^{\alpha+\beta}}\int_0^t u^\beta \mathcal{L}(u)du \tag{18.2.12}$$

is also nondecreasing.

18.3. Local Convergence

In this section we present the local convergence of inexact Newton method using (18.1.6) and (18.1.7). We shall first consider the case $B_n = F'(x_n)$ for each $n = 0,1,2,\cdots$.

Theorem 18.3.1. *Suppose* $x^\star$ *satisfies (18.1.1),* F *has a continuous Fréchet derivative in* $U(x^\star,r)$, $F'(x^\star)^{-1}$ *exists and* $F'(x^\star)F'$ *satisfies the radius Lipschitz condition (18.1.6) and the center-Lipschitz condition (18.1.7). Assume* $B_n = F'(x_n)$, *for each n in (18.1.3),* $v_n = \theta_n\|(P_nF'(x_n))^{-1}\|\|P_nF'(x_n)\| = \theta_n Cond(P_nF'(x_n))$ *with* $v_n \leq v < 1$. *Let* $r > 0$ *satisfy*

$$\int_0^r \mathcal{L}_0(u)du \leq \frac{1-v}{2}. \tag{18.3.1}$$

Then (INLM) (for $B_n = F'(x_n)$*) is convergent for all* $x_0 \in U(x^\star,r)$ *and*

$$\|x_{n+1}-x^\star\| \leq \left((1+v)\frac{\int_0^{\rho(x_0)}\mathcal{L}(u)du}{1-\int_0^{\rho(x_0)}\mathcal{L}_0(u)du}+v\right)\|x_n-x^\star\|, \tag{18.3.2}$$

where

$$q = (1+v)\frac{\int_0^{\rho(x_0)}\mathcal{L}(u)du}{1-\int_0^{\rho(x_0)}\mathcal{L}_0(u)du}+v \tag{18.3.3}$$

is less than 1. Further, suppose that the function $\mathcal{L}_\alpha$ *defined in (18.2.11) is nondecreasing for some* α *with* $0 < \alpha \leq 1$. *Let* $\tilde{r}$ *satisfy*

$$\frac{(1+v)\int_0^{\tilde{r}}\mathcal{L}(u)udu}{1-\int_0^{\tilde{r}}\mathcal{L}_0(u)du}+v \leq 1. \tag{18.3.4}$$

Then (INLM) (for $B_n = F'(x_n)$) is convergent for all $x_0 \in U(x^\star, \tilde{r})$ and

$$\|x_{n+1} - x^\star\| \leq \left((1+v) \frac{\int_0^{\rho(x_0)} \mathcal{L}(u)du}{\rho(x_0)^{1+\alpha}(1 - \int_0^{\rho(x_0)} \mathcal{L}_0(u)du)} \rho(x_n)^\alpha + v \right) \|x_n - x^\star\|, \quad (18.3.5)$$

where

$$\tilde{q} = (1+v) \frac{\int_0^{\rho(x_0)} \mathcal{L}(u)du}{\rho(x_0)(1 - \int_0^{\rho(x_0)} \mathcal{L}_0(u)du)} + v \quad (18.3.6)$$

is less than 1.

Proof. Let $x_0 \in B(x^\star, r)$, where r satisfies (18.3.1), then q given by (18.3.3) is such that $q \in (0,1)$. Indeed, by the positivity of $\mathcal{L}$, we have

$$\begin{aligned} q &= (1+v) \frac{\int_0^{\rho(x_0)} \mathcal{L}(u)du}{1 - \int_0^{\rho(x_0)} \mathcal{L}_0(u)du} + v \\ &< (1+v) \frac{\int_0^{r} \mathcal{L}(u)du}{1 - \int_0^{r} \mathcal{L}_0(u)du} + v \leq 1. \end{aligned}$$

Suppose that (notice that $x_0 \in U(x^\star, r)$) $x_n \in U(x^\star, r)$, we have by (18.1.3)

$$\begin{aligned} x_{n+1} - x^\star &= x_n - x^\star - F'(x_n)^{-1}(F(x_n) - F(x^\star)) + F'(x_n)^{-1} r_n \\ &= F'(x_n)^{-1} F(x^\star) \int_0^1 F'(x^\star)^{-1}(F'(x_n) - F'(x^\theta))(x_n - x^\star) d\theta + F'(x_n) P_n^{-1} P_n r_n \end{aligned}$$

where $x^\theta = x^\star + \theta(x_n - x^\star)$. It follows, by Lemma 18.2.1 and 18.2.2 and conditions (18.1.6) and (18.1.7) that we can obtain in turn

$$\begin{aligned} \|x_{n+1} - x^\star\| &= \|F'(x_n)^{-1} F(x^\star)\| \int_0^1 \|F'(x^\star)^{-1}(F'(x_n) - F'(x^\theta))\| \|(x_n - x^\star)\| d\theta \\ &\quad + \theta_n \|F'(x_n) P_n^{-1}\| \|P_n F(x_n)\| \\ &\leq \frac{1}{1 - \int_0^{\rho(x)} \mathcal{L}_0(u)du} \int_0^1 \int_{\theta\rho(x)}^{\rho(x)} \mathcal{L}(u)du \rho(x) d\theta \\ &\quad + \theta_n \|(P_n F'(x_n))^{-1}\| \|P_n F'(x_n) F'(x_n)^{-1} F(x_n)\| \\ &\leq \frac{\int_0^{\rho(x)} \mathcal{L}(u)udu}{1 - \int_0^{\rho(x)} \mathcal{L}_0(u)du} \int_0^1 \int_{\theta\rho(x)}^{\rho(x)} \mathcal{L}(u)du \rho(x) d\theta \\ &\quad + \theta_n Cond(P_n F'(x_n)) \left(\|x_n - x^\star\| + \frac{\int_0^{\rho(x_n)} \mathcal{L}(u)du}{1 - \int_0^{\rho(x_n)} \mathcal{L}_0(u)du} \right) \\ &\leq (1+v_n) \frac{\int_0^{\rho(x_n)} \mathcal{L}(u)udu}{1 - \int_0^{\rho(x_n)} \mathcal{L}_0(u)du} + v_n \rho(x_n) \\ &\leq \left((1+v_n) \frac{\int_0^{\rho(x_n)} \mathcal{L}(u)du}{1 - \int_0^{\rho(x_n)} \mathcal{L}_0(u)du} + v_n \right) \rho(x_n). \end{aligned} \quad (18.3.7)$$

In particular, if $n=0$ in (18.3.7), we obtain $\|x_1-x^\star\| \le q\|x_0-x^\star\|$. Hence $x_1 \in U(x^\star, r)$, this shows that (INLM) can be continued an infinite number of times. By mathematical induction, all $x_n \in U(x^\star, r)$ and $\rho(x_n)=\|x_n-x^\star\|$ decreases monotonically. Consequently, we have for each $n=0,1,2,\cdots$

$$\begin{aligned}\|x_{n+1}-x^\star\| &\le \left((1+v_n)\frac{\int_0^{\rho(x_n)}\mathcal{L}(u)du}{1-\int_0^{\rho(x_n)}\mathcal{L}_0(u)du}+v_n\right)\|x_n-x^\star\| \\ &\le \left((1+v)\frac{\int_0^{\rho(x_0)}\mathcal{L}(u)du}{1-\int_0^{\rho(x_0)}\mathcal{L}_0(u)du}+v\right)\|x_n-x^\star\|.\end{aligned}$$

Hence, we showed (18.3.3). Moreover, if $\tilde r$ satisfies (18.3.4) and $\mathcal{L}_\alpha$ defined by (18.2.11) is nondecreasing for some α with $0<\alpha\le 1$, then we get

$$\begin{aligned}\tilde q &= (1+v)\frac{\int_0^{\rho(x_0)}\mathcal{L}(u)udu}{\rho(x_0)^{1+\alpha}(1-\int_0^{\rho(x_0)}\mathcal{L}_0(u)du)}\rho(x_0)^\alpha+v \\ &< (1+v)\frac{\int_0^{\tilde r}\mathcal{L}(u)udu}{\tilde r^{1+\alpha}(1-\int_0^{\tilde r}\mathcal{L}_0(u)du)}\tilde r^\alpha+v\le 1.\end{aligned}$$

If, $n=0$ in (18.3.1), we get $\|x_1-x^\star\|\le\tilde q\|x_0-x^\star\|<\|x_0-x^\star\|$. Hence, $x_1\in U(x^\star,\tilde r)$. That is (INM) can be continued an infinite number of times. It follows by mathematical induction that, all x_n belongs to $U(x^\star,\tilde r)$ and $\rho(x_n)=\|x_n-x^\star\|$ decreases monotonically. Therefore, for all $k\ge 0$, from (18.3.7) and lemma 18.2.4 we get in turn that

$$\begin{aligned}\|x_n-x^\star\| &\le (1+v_n)\frac{\int_0^{\rho(x_n)}\mathcal{L}(u)udu}{1-\int_0^{\rho(x_n)}\mathcal{L}_0(u)du}+v_n\rho(x_n) \\ &= (1+v_n)\frac{\varphi_{1,\alpha}(\rho(x_n))}{1-\int_0^{\rho(x_n)}\mathcal{L}_0(u)du}\rho(x_n)^{1+\alpha}+v_n\rho(x_n) \\ &\le (1+v_n)\frac{\varphi_{1,\alpha}(\rho(x_0))}{1-\int_0^{\rho(x_0)}\mathcal{L}_0(u)du}\rho(x_n)^{1+\alpha}+v_n\rho(x_n) \\ &\le (1+v)\frac{\varphi_{1,\alpha}(\rho(x_0))}{1-\int_0^{\rho(x_0)}\mathcal{L}_0(u)du}\rho(x_n)^{1+\alpha}+v\rho(x_n).\end{aligned}$$

□

Remark 18.3.2. *If $\mathcal{L}_0=\mathcal{L}$ our Theorem 18.3.1 reduces to Theorem 18.3.1 in [13] (see also [5]). Otherwise, i.e., if $\mathcal{L}_0<\mathcal{L}$, then our Theorem 18.3.1 constitutes an improvement. In particular, for $v=0$, the estimate for the radii of convergence ball for Newton's method are given by*

$$\int_0^r \mathcal{L}_0(u)du\le\frac{1}{2}$$

and

$$\frac{1}{\tilde r}\int_0^{\tilde r}(\mathcal{L}_0(u)\tilde r+\mathcal{L}(u)u)du\le 1,$$

which reduce to the ones in [14] if $\mathcal{L}_0 = \mathcal{L}$. Then, we can conclude that vanishing residual, Theorem 18.3.1 merges into the theory of the Newton method. Besides, if the function $\mathcal{L}_\alpha$ defined by (18.2.11) is nondecreasing for $\alpha = 1$, we improve the result in [5].

Next, we present a result analogous to Theorem 18.3.1 can also be proven for inexact Newton-like method, where $B_n = B(x_n)$ approximates $F'(x_n)$.

Theorem 18.3.3. *Suppose $x^\star$ satisfies (18.1.1), F has a continuous derivative in $U(x^\star, r)$, $F'(x^\star)^{-1}$ exists and $F'(x^\star)F'$ satisfies the radius Lipschitz condition (18.1.6) and the center Lipschitz condition (18.1.7). Let $B(x)$ be an approximation to the $F'(x)$ for all $x \in U(x^\star, r)$, $B(x)$ is invertible and*

$$\|B(x)^{-1}F'(x)\| \leq \omega_1, \ \|B(x)^{-1}F'(x) - I\| \leq \omega_2, \tag{18.3.8}$$

where $v_n = \theta_n \|(P_nF'(x_n))^{-1}\|\|P_nF'(x_n)\| = \theta_n Cond(P_nF'(x_n))\|$ with $v_n \leq v < 1$. Let $r > 0$ satisfy

$$\int_0^r \mathcal{L}_0(u)du < \frac{1-\omega_2-\omega_1 v}{1+\omega_1-\omega_2}. \tag{18.3.9}$$

Then the (INLM) method is convergent for all $x_0 \in U(x^\star, r)$ and

$$\|x_{n+1} - x^\star\| \leq \left((1+v)\frac{\omega_1 \int_0^{\rho(x_0)} \mathcal{L}(u)du}{1 - \int_0^{\rho(x_0)} \mathcal{L}_0(u)du} + \omega_2 + \omega_1 v \right) \|x_n - x^\star\|, \tag{18.3.10}$$

where

$$q = (1+v)\frac{\omega_1 \int_0^{\rho(x_0)} \mathcal{L}(u)du}{1 - \int_0^{\rho(x_0)} \mathcal{L}_0(u)du} + \omega_2 + \omega_1 v \tag{18.3.11}$$

is less than 1. Further, suppose that the function $\mathcal{L}_\alpha$ defined by (18.2.11) is nondecreasing for some α with $0 < \alpha \leq 1$. Ler $\tilde{r}$ satisfy

$$(1+v)\frac{\omega_1 \int_0^{\tilde{r}} \mathcal{L}(u)du}{1 - \int_0^{\tilde{r}} \mathcal{L}_0(u)du} + \omega_2 + \omega_1 v \leq 1. \tag{18.3.12}$$

Then (INLM) is convergent for all $x_0 \in U(x^\star, \tilde{r})$ and

$$\|x_{n+1} - x^\star\| \leq (1+v)\frac{\omega_1 \int_0^{\rho(x_0)} \mathcal{L}(u)du}{\rho(x_0)^{1+\alpha}(1 - \int_0^{\rho(x_0)} \mathcal{L}_0(u)du)}\rho(x_n)^{1+\alpha} + (\omega_2 + \omega_1 v)\rho(x_n), \tag{18.3.13}$$

where

$$\tilde{q} = (1+v)\frac{\omega_1 \int_0^{\rho(x_0)} \mathcal{L}(u)du}{1 - \int_0^{\rho(x_0)} \mathcal{L}_0(u)du} + \omega_2 + \omega_1 v \tag{18.3.14}$$

is less than 1.

Proof. Let $x_0 \in U(x^\star, r)$, where r satisfies (18.3.9), then q given by (18.3.11) is such that $q \in (0,1)$. Indeed, by the positivity of $\mathcal{L}$, we have

$$\begin{aligned} q &= (1+v)\frac{\omega_1 \int_0^{\rho(x_0)} \mathcal{L}(u)du}{1 - \int_0^{\rho(x_0)} \mathcal{L}_0(u)du} + \omega_2 + \omega_1 v \\ &= (1+v)\frac{\omega_1 \int_0^{r} \mathcal{L}(u)du}{1 - \int_0^{r} \mathcal{L}_0(u)du} + \omega_2 + \omega_1 v \leq 1. \end{aligned}$$

Moreover, if $x_n \in U(x^\star, r)$, we have by (18.1.3) in turn that

$$\begin{aligned} x_{n+1}-x^\star &= x_n - x^\star - B_n^{-1}(F(x_n)-F(x^\star)) + B_n^{-1}r_n \\ &= x_n - x^\star - \int_0^1 B_n^{-1}F'(x^\theta)d\theta(x_n-x^\star) + B_n^{-1}P_n^{-1}P_n r_n \\ &= -B_n^{-1}F'(x_n)\int_0^1 F'(x_n)^{-1}F'(x^\star)F'(x^\star)^{-1}(F'(x^\star)-F'(x^\theta))(x_n-x^\star)d\theta \\ &\quad +B_n^{-1}(F'(x_n)-B_n)(x_n-x^\star)+B_n^{-1}P_n^{-1}P_n r_n, \end{aligned}$$

where $x^\theta = x^\star + \theta(x_n - x^\star)$. Using, Lemma 18.2.1 and 18.2.2 and condition (18.3.8) we obtain

$$\begin{aligned} \|x_{n+1}-x^\star\| &\le \|B_n^{-1}F'(x_n)\|\int_0^1 \|F'(x_n)^{-1}F'(x^\star)\|\|F'(x^\star)^{-1}(F'(x^\star)-F'(x^\theta))\| \\ &\quad \|x_n-x^\star\|d\theta + \|B_n^{-1}(F'(x_n)-B_n)\|\|x_n-x^\star\| + \theta_n\|B_n^{-1}P_n^{-1}\|\|B_nF(x_n)\| \\ &\le \frac{\omega_1}{1-\int_0^{\rho(x_n)}\mathcal{L}_0(u)du}\int_0^1\int_{\theta\rho(x_n)}^{\rho(x_n)}\mathcal{L}(u)du\rho(x_n)d\theta \\ &\quad +\theta_n\|P_n^{-1}F'(x_n)\|\|(P_nF'(x_n))^{-1}\|\|P_nF'(x_n)\|\|F'(x_n)^{-1}F(x_n)\| \\ &\le \frac{\omega_1\int_0^{\rho(x_n)}\mathcal{L}(u)udu}{1-\int_0^{\rho(x_n)}\mathcal{L}_0(u)du} + \omega_2\rho(x_n) + \omega_1 v_n\left(\rho(x_n) + \frac{\int_0^{\rho(x_n)}\mathcal{L}(u)udu}{1-\int_0^{\rho(x_n)}\mathcal{L}_0(u)du}\right) \\ &\le (1+v_n)\frac{\omega_1\int_0^{\rho(x_n)}\mathcal{L}(u)udu}{1-\int_0^{\rho(x_n)}\mathcal{L}_0(u)du} + (\omega_2+\omega_1 v_n)\rho(x_n) \qquad (18.3.15) \\ &\le \left((1+v_n)\frac{\omega_1\int_0^{\rho(x_n)}\mathcal{L}(u)udu}{1-\int_0^{\rho(x_n)}\mathcal{L}_0(u)du} + \omega_2 + \omega_1 v_n\right)\rho(x_n). \end{aligned}$$

If $n = 0$,in (18.3.15), we obtain $\|x_1 - x^\star\| \le q\|x_0 - x^\star\| < \|x_0 - x^\star\|$. Hence $x_1 \in U(x^\star, r)$, this shows that the iteration can be continued an infinite number of times. By mathematical induction, $x_n \in U(x^\star, r)$ and $\rho(x_n) = \|x_n - x^\star\|$ decreases monotonically. Therefore, for all $n \ge 0$, we have in turn that

$$\begin{aligned} \|x_{n+1}-x^\star\| &\le \left((1+v_n)\frac{\omega_1\int_0^{\rho(x_n)}\mathcal{L}(u)udu}{1-\int_0^{\rho(x_n)}\mathcal{L}_0(u)du} + \omega_2 + \omega_1 v_n\right)\rho(x_n) \\ &\le \left((1+v)\frac{\omega_1\int_0^{\rho(x_0)}\mathcal{L}(u)udu}{1-\int_0^{\rho(x_0)}\mathcal{L}_0(u)du} + \omega_2 + \omega_1 v\right)\rho(x_n), \end{aligned}$$

which implies (18.3.10). Furthermore, if $\tilde{r}$ satisfies (18.3.12) and $\mathcal{L}_\alpha$ defined by (18.2.11) is nondecreasing for some α with $0 < \alpha \le 1$, then we get

$$\begin{aligned} \tilde{q} &= (1+v)\frac{\omega_1\int_0^{\rho(x_0)}\mathcal{L}(u)udu}{\rho(x_0)^{1+\alpha}(1-\int_0^{\rho(x_0)}\mathcal{L}_0(u)du)}\rho(x_n)^\alpha + \omega_2 + \omega_1 v \\ &< (1+v)\frac{\omega_1\int_0^{\tilde{r}}\mathcal{L}(u)udu}{\tilde{r}^{1+\alpha}(1-\int_0^{\tilde{r}}\mathcal{L}_0(u)du)}\tilde{r}^\alpha + \omega_2 + \omega_1 v \le 1. \end{aligned}$$

If, $n=0$ in (18.3.15), we obtain $\|x_1-x^\star\| \leq \tilde{q}\|x_0-x^\star\| < \|x_0-x^\star\|$. Hence, $x_1 \in U(x^\star,\tilde{r})$, this shows that (18.1.4) can be continued infinite number of times. By mathematical induction, $x_n \in U(x^\star,\tilde{r})$ and $\rho(x_n)=\|x_n-x^\star\|$ decreases monotonically. Therefore, for all $n\geq 0$, we have

$$\begin{aligned}
\|x_{n+1}-x^\star\| &\leq (1+v_n)\frac{\omega_1\int_0^{\rho(x_n)}\mathcal{L}(u)udu}{1-\int_0^{\rho(x_n)}\mathcal{L}_0(u)du}+(\omega_2+\omega_1 v_n)\rho(x_n)\\
&\leq (1+v)\frac{\omega_1\varphi_{1,\alpha}(\rho(x_n))}{1-\int_0^{\rho(x_0)}\mathcal{L}_0(u)du}\rho(x_n)^{1+\alpha}+(\omega_2+\omega_1 v)\rho(x_n)\\
&\leq (1+v)\frac{\omega_1\varphi_{1,\alpha}(\rho(x_0))}{1-\int_0^{\rho(x_0)}\mathcal{L}_0(u)du}\rho(x_n)^{1+\alpha}+(\omega_2+\omega_1 v)\rho(x_n).
\end{aligned}$$

□

Remark 18.3.4. *If $\mathcal{L}_0=\mathcal{L}$ our Theorem 18.3.3 reduces to Theorem 18.3.2 in [13] (see also [5]). Otherwise, i.e., if $\mathcal{L}_0<\mathcal{L}$, then our Theorem 18.3.3 constitutes an improvement. In in Theorem 18.3.2, the function $Ł_\alpha$ defined by (18.2.11) is nondecreasing for $\alpha=1$, we improve the result of [5]. In particular, for $v=0$, we can get the radii of converence ball for the Newton-like method [14].*

18.4. Special Cases

In this section, we consider the following special cases of Theorem 18.3.1 and Theorem 18.3.3:

Corollary 18.4.1. *Suppose $x^\star$ satisfies (18.1.1), F has a continuous derivative in $U(x^\star,r)$, $F'(x^\star)^{-1}$ exist, $F'(x^\star)F'$ satisfies the radius Lipschitz condition with*

$$\mathcal{L}(u)=c\alpha u^{\alpha-1}\|F'(x^\star)^{-1}(F'(x)-F'(x^{)})\|\leq c(1-\theta^\alpha)\|x-x^\star\|^\alpha \tag{18.4.1}$$

for each $x\in U(x^\star,r)$, $0\leq\theta\leq 1$, where $x^\theta=x^\star+\theta(x-x^\star)$ and the center- radius Lipschitz condition with

$$\mathcal{L}_0(u)=c_0\alpha u^{\alpha-1}\|F'(x^\star)^{-1}(F'(x)-F'(x^\star))\|\leq c_0\|x-x^\star\|^\alpha \tag{18.4.2}$$

for each $x\in U(x^\star,r)$, $0\leq\theta\leq 1$ for some $c_0\leq c$. Assume $B_n=F'(x_n)$, for each n in (18.1.3), $v_n=\theta_n\|(P_nF'(x_n))^{-1}\|\|P_nF'(x_n)\|=\theta_n Cond(P_nF'(x_n))$ with $v_n\leq v<1$. Let $\tilde{r}>0$ satisfy

$$\tilde{r}=\left(\frac{(1-v)(1+\alpha)}{c(1+v)\alpha+c_0(1-v)(1+\alpha)}\right)^{\frac{1}{\alpha}}.$$

Then the inexact Newton method is convergent for all $x_0\in U(x^\star,\tilde{r})$ and

$$\|x_{n+1}-x^\star\|\leq\left(\frac{c\alpha(1+v)}{(1+\alpha)(1-c_0\|x_0-x^\star\|^\alpha)}\|x_0-x^\star\|^\alpha+v\right)\|x_n-x^\star\|,$$

where

$$q=\frac{c\alpha(1+v)}{(1+\alpha)(1-c_0\|x_0-x^\star\|^\alpha)}\|x_0-x^\star\|^\alpha+v$$

is less than 1.

Corollary 18.4.2. *Suppose $x^\star$ satisfies (18.1.1), F has a continuous derivative in $U(x^\star,r)$, $F'(x^\star)^{-1}$ exists and $F'(x^\star)F'$ satisfies the radius Lipschitz condition (18.4.1) and the center Lipschitz condition (18.4.2). Let $B(x)$ be an approximation to the $F'(x)$ for all $x \in B(x^\star,r)$, $B(x)$ is invertible and satisfies condition (18.3.8), $v_n = \theta_n\|(P_nF'(x_n))^{-1}\|\|P_nF'(x_n)\| = \theta_n Cond(P_nF'(x_n))$ with $v_k \le v < 1$. Let $\tilde{r} > 0$ satisfy*

$$\tilde{r} = \left(\frac{(1+\alpha)(1-\omega_2-\omega_1 v)}{c(1+v)\omega_1\alpha + c_0(1+\alpha)(1-\omega_2-\omega_1 v)}\right)^{\frac{1}{\alpha}}.$$

Then the inexact Newton method is convergent for all $x_0 \in U(x^\star,\tilde{r})$ and

$$\|x_{n+1}-x^\star\| \le \left(\frac{c\alpha(1+v)\omega_1}{(1+\alpha)(1-c_0\|x_0-x^\star\|^\alpha)}\|x_0-x^\star\| + \omega_2 + \omega_1 v\right)\|x_n-x^\star\|,$$

where

$$q = \frac{c\alpha(1+v)\omega_1}{(1+\alpha)(1-c_0\|x_0-x^\star\|^\alpha)}\|x_0-x^\star\|^\alpha + \omega_2 + \omega_1 v$$

is less than 1.

Remark 18.4.3. *(a) If, $v = 0$ in Corollary 18.4.1, the estimate for the radius of convergence ball for Newton's method is given by*

$$\tilde{r} = \left(\frac{1+\alpha}{c\alpha + c_0(1+\alpha)}\right)^{\frac{1}{\alpha}},$$

which improves the result in [5, 13] for $c_0 < c$. Moreover, if $\alpha = 1$, our radius reduces to $\tilde{r} = \frac{2}{2c_0+c}$, which is larger than the one obtained by Rheinholdt and Traub [11, 12] given by $\tilde{r} = \frac{2}{3c}$ if $c_0 < c$ (see also the numerical examples at the end of the chapter).

(b) The results in section 18.5 of [5, 13] using only center-Lipschitz condition can be improved, if rewritten using $\mathcal{L}_0$ instead of $\mathcal{L}$.

18.5. Examples

Finally, we provide an example where $\mathcal{L}_0 < \mathcal{L}$.

Example 18.5.1. *Let $\mathcal{X} = \mathcal{Y} = \mathbb{R}^3$, $\mathcal{D} = \overline{U}(0,1)$ and $x^\star = (0,0,0)$. Define function F on $\mathcal{D}$ for $w = (x,y,z)$ by*

$$F(w) = (e^x - 1, \frac{e-1}{2}y^2 + y, z). \tag{18.5.1}$$

Then, the Fréchet derivative of F is given by

$$F'(w) = \begin{pmatrix} e^x & 0 & 0 \\ 0 & (e-1)y+1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Notice that we have $F(x^\star) = 0$, $F'(x^\star) = F'(x^\star)^{-1} = \mathrm{diag}\{1,1,1\}$ and $\mathcal{L}_0 = e-1 < \mathcal{L} = e$.

More examples where $\mathcal{L}_0 < \mathcal{L}$ can be found in [1, 2, 3].

18.6. Conclusion

Under the hypothesis that $F'(x^{\star})F'$ satisfies the center Lipschitz condition (18.1.7) and the radius Lipschitz condition (18.1.6), we presented a more precise local convergence analysis for the enexact Newton method under the same computational cost as in earlier studies such as Chen and Li [5], Zhang, Li and Xie [13]. Numerical examples are provided to show that the center Lipschitz function can be smaller than the radius Lipschitz function.

References

[1] Argyros, I.K., Cho, Y.J, Hilout, S., Numerical methods for equations and its applications, CRC Press, Taylor and Francis, New York, 2012.

[2] Argyros, I.K., Hilout, S., Weak convergence conditions for inexact Newton-type methods, *App. Math. Comp.*, 218 (2011), 2800-2809.

[3] Argyros, I.K., Hilout, S., On the semilocal convergence of a Damped Newton's method, *Appl.Math. Comput.*, 219, 5(2012), 2808-2824.

[4] Argyros, I.K., Hilout, S., Weaker conditions for the convergence of Newton's method, *J. Complexity*, 28 (2012), 364-387.

[5] Chen, J., Li, W., Convergence behaviour of inexact Newton methods under weaker Lipschitz condition, *J. Comput. Appl. Math.*, 191 (2006), 143-164.

[6] Dembo, R.S., Eisenstat, S.C., Steihaug, T., Inexact Newton methods, *SIAM J. Numer. Anal.*, 19 (1982), 400-408.

[7] Kantorovich, L.V., Akilov, G.P., Functional Analysis, 2nd ed., Pergamon Press, Oxford, 1982.

[8] Morini, B., Convergence behaviour of inexact Newton method, *Math. Comp.* 61 (1999), 1605-1613.

[9] Ortega, J.M., Rheinholdt, W.C., *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970.

[10] Proinov, P.D., General local convergence theory for a class of iterative processes and its applications to Newton's method, *J. Complexity*, 25 (2009), 38- 62.

[11] Rheinholdt, W.C., An adaptive continuation process for solving systems of nonlinear equations, Polish Academy of Science, *Banach Ctr. Publ.* 3 (1977), 129-142.

[12] Traub, J.F., Iterative methods for the solution of equations, Prentice- Hall Series in *Automatic Computation*, Englewood Cliffs, N. J., 1964.

[13] Zhang, H., Li, W., Xie, L., Convergence of inexact Newton methods of nonlinear operator equations in Banach spaces.

[14] Wang, X., Convergence of Newton's method and uniqueness of the solution of equations in Banach space, *IMA J. Numer. Anal.*, 20 (2000), 123-134.

Chapter 19

Expanding the Applicability of Secant Method with Applications

19.1. Introduction

In this chapter we are concerned with the problem of approximating a locally unique solution $x^\star$ of equation

$$F(x) = 0, \tag{19.1.1}$$

where F is a Fréchet–differentiable operator defined on a convex subset $\mathcal{D}$ of a Banach space $\mathcal{X}$ with values in a Banach space $\mathcal{Y}$.

A vast number of problems from Applied Science including engineering can be solved by means of finding the solutions equations in a form like (19.1.1) using mathematical modelling [7, 10, 15, 18]. For example, dynamic systems are mathematically modeled by difference or differential equations, and their solutions usually represent the states of the systems. Except in special cases, the solutions of these equations cannot be found in closed form. This is the main reason why the most commonly used solution methods are iterative. Iteration methods are also applied for solving optimization problems. In such cases, the iteration sequences converge to an optimal solution of the problem at hand. Since all of these methods have the same recursive structure, they can be introduced and discussed in a general framework. The convergence analysis of iterative methods is usually divided into two categories: semilocal and local convergence analysis. In the semilocal convergence analysis one derives convergence criteria from the information around an initial point whereas in the local analysis one finds estimates of the radii of convergence balls from the information around a solution.

We consider the Secant method in the form

$$x_{n+1} = x_n - \delta F(x_{n-1}, x_n)^{-1}\, F(x_n) \quad (n \geq 0), \quad (x_{-1}, x_0 \in \mathcal{D}) \tag{19.1.2}$$

where $\delta F(x,y) \in \mathcal{L}(\mathcal{X},\mathcal{Y})$ $(x, y \in \mathcal{D})$ the space of bounded linear operators from $\mathcal{X}$ into $\mathcal{Y}$. of the Fréchet–derivative of F [15, 18].

The semilocal convergence matter is, based on the information around an initial point, to give criteria ensuring the convergence of iteration procedures. A very important problem in the study of iterative procedures is the convergence domain. In general the convergence domain is small. Therefore, it is important to enlarge the convergence domain without additional hypotheses. Another important problem is to find more precise error estimates on the distances $\|x_{n+1}-x_n\|$, $\|x_n-x^\star\|$. These are our objectives in this chapter.

The secant method, also known under the name of Regula Falsi or the method of chords, is one of the most used iterative procedures for solving nonlinear equations. According to A. N. Ostrowski [19], this method is known from the time of early Italian algebraists. In the case of equations defined on the real line, the Secant method is better than Newton's method from the point of view of the efficiency index [7]. The Secant method was extended for the solution of nonlinear equations in Banach Spaces by A. S. Sergeev [24] and J. W. Schmidt [23].

The simplified Secant method

$$x_{n+1}=x_n-\delta F(x_{-1},x_0)^{-1}F(x_n) \qquad (n\geq 0), \quad (x_{-1},x_0\in D)$$

was first studied by S. Ulm [25]. The first semilocal convergence analysis was given by P. Laasonen [21]. His results was improved by F. A. Potra and V. Pták [20, 21, 22]. A semilocal convergence analysis for general secant-type methods was given in general by J. E. Dennis [13]. Bosarge and Falb [9], Dennis [10], Potra [20, 21, 22], Argyros [5, 6, 7, 8], Hernández et al. [13] and others [14], [18], [26], have provided sufficient convergence conditions for the Secant method based on Lipschitz–type conditions on δF. Moreover, there exist new graphical tools to study this kind of methods [17].

The conditions usually associated with the semilocal convergence of Secant method (19.1.2) are:

- F is a nonlinear operator defined on a convex subset $\mathcal{D}$ of a Banach space $\mathcal{X}$ with values in a Banach space $\mathcal{Y}$;
- x_{-1} and x_0 are two points belonging to the interior $\mathcal{D}^0$ of $\mathcal{D}$ and satisfying the inequality
$$\| x_0-x_{-1} \|\leq c;$$
- F is Fréchet–differentiable on $\mathcal{D}^0$, and there exists an operator $\delta F:\ \mathcal{D}^0\times\mathcal{D}^0\to\mathcal{L}(\mathcal{X},\mathcal{Y})$ such that:

  the linear operator $A=\delta F(x_{-1},x_0)$ is invertible, its inverse A^{-1} is bounded, and:
$$\| A^{-1}\ F(x_0) \|\leq \eta;$$
$$\| A\ [\delta F(x,y)-F'(z)] \|\leq \ell\ (\| x-z \|+\| y-z \|);$$
  for all $x,y,z\in\mathcal{D}$;

$$\ell\, c+2\,\sqrt{\ell\,\eta}\leq 1. \tag{19.1.3}$$

The sufficient convergence condition(19.1.3) is easily violated (see the Numerical Examples). Hence, there is no guarantee in these cases that equation (19.1.1) under the information (ℓ,c,η) has a solution that can be found using Secant method (19.1.2). In this chapter we are motivated by optimization considerations, and the above observation.

The use of Lipschitz and center–Lipschitz conditions is one way used to enlarge the convergence domain of different methods. This technique consist on using both conditions together instead of using only the Lipschitz one which allow us to find a finer majorizing sequence, that is, a larger convergence domain. It has been used in order to find weaker convergence criteria for Newton's method by Argyros in [8]. Gutiérrez et al in [12] give sufficient conditions for Newton's method using both Lipschitz and center-Lipschitz conditions, for the damped Newton's methods and Amat et al in [3, 4] or García-Olivo [11] for other methods.

Here using Lipschitz and center–Lipschitz conditions, we provide a new semilocal convergence analysis for (19.1.2). It turns out that our new convergence criteria can always be weaker than the old ones given in earlier studies such as [2, 14, 16, 18, 20, 21, 22, 23, 26, 27]. The chapter is organized as follows: The semilocal convergence analysis of the secant method is presented in Section 19.2. Numerical examples are provided in Section 19.3.

19.2. Semilocal Convergence Analysis of the Secant Method

In this Section, we present the semilocal convergence analysis of the secant-method (19.1.2). First, we present two auxiliary results concerning convergence criteria and majorizing sequences.

Lemma 19.2.1. *Let $\ell_0 > 0$, $\ell > 0$, $c > 0$ and $\eta > 0$ be constants with $\ell_0 \leq \ell$. Then, the following items hold*

(i)

$$0 < \frac{\ell(c+\eta)}{1-\ell_0(c+\eta)} \leq \frac{2\ell}{\ell+\sqrt{\ell^2+4\ell_0\ell}} < \frac{1-\ell_0(c+\eta)}{1-\ell_0 c} \Leftrightarrow c+\eta \leq \frac{4\ell^2}{\left(\ell+\sqrt{\ell^2+4\ell_0\ell}\right)^2}; \tag{19.2.1}$$

(ii)

$$\ell c \leq \frac{3-\sqrt{1+4\frac{\ell_0}{\ell}}}{1+\sqrt{1+4\frac{\ell}{\ell_0}}} \Leftrightarrow \frac{(1-\ell c)^2}{4} \leq b^2-\ell c; \tag{19.2.2}$$

(iii)

$$\ell c \geq \frac{3-\sqrt{1+4\frac{\ell_0}{\ell}}}{1+\sqrt{1+4\frac{\ell}{\ell_0}}} \Leftrightarrow \frac{(1-\ell c)^2}{4} \geq b^2-\ell c; \tag{19.2.3}$$

(iv)

$$\ell c \leq \frac{3-\sqrt{1+4\frac{\ell_0}{\ell}}}{1+\sqrt{1+4\frac{\ell}{\ell_0}}} \textit{ and } \ell c+\sqrt{\ell\eta} \leq 1 \Rightarrow c+\eta \leq \frac{4\ell}{\left(\ell+\sqrt{\ell^2+4\ell_0\ell}\right)^2}c; \tag{19.2.4}$$

(v)

$$\ell c \geq \frac{3-\sqrt{1+4\frac{\ell_0}{\ell}}}{1+\sqrt{1+4\frac{\ell}{\ell_0}}} \text{ and } c+\eta \leq \frac{4\ell}{\left(\ell+\sqrt{\ell^2+4\ell_0\ell}\right)^2} \Rightarrow \ell c+\sqrt{\ell\eta} \leq 1. \quad (19.2.5)$$

Proof. Let $x = 1-\ell c$, $y = \ell\eta$, $a = \dfrac{\ell_0}{\ell}$ and $b = \dfrac{2}{1+\sqrt{1+4a}}$. Then, we have that $ab^2+b-1 = 0$ and $ab+1 = \dfrac{1}{b}$.

(i) The triple inequality in (19.2.1) holds, if

$$\frac{\ell c+\ell\eta}{1-a\ell(c+\eta)} \leq \frac{2\ell}{\ell+\sqrt{\ell^2+4a\ell^2}} = b, \quad (19.2.6)$$

$$b < \frac{1-a\ell(c+\eta)}{1-a\ell c} \quad (19.2.7)$$

and

$$\ell(c+\eta) < \frac{1}{a} \quad (19.2.8)$$

or, if

$$y \leq b^2-(1-x), \quad (19.2.9)$$

$$y < \frac{1-b}{a}-(1-b)(1-x) = b^2-(1-b)(1-x), \quad (19.2.10)$$

and

$$y \leq \frac{1}{a}-(1-x), \quad (19.2.11)$$

respectively. We have that $ab^2 = 1-b < 1$ by the definition of a and b. It follows that

$$b^2-(1-x) < \frac{1}{a}-(1-x) \quad (19.2.12)$$

and from $(1-b)(1-x) < (1-x)$ we get that

$$b^2-(1-x) < b^2-(1-b)(1-x). \quad (19.2.13)$$

Hence, it follows from (19.2.12) and (19.2.13) that (19.2.6)–(19.2.8) are satisfied if (19.2.9) holds. But (19.2.9) is equivalent to the right hand side inequality in (19.2.1). Conversely, if the right hand side inequality in (19.2.1) holds, then (19.2.9), (19.2.12) and (19.2.13) imply (19.2.10) and (19.2.11) imply (19.2.6)–(19.2.8) which imply the triple inequality in (19.2.1).

(ii)

$$\begin{aligned}\ell c &\leq \frac{3-\sqrt{1+4\frac{\ell_0}{\ell}}}{1+\sqrt{1+4\frac{\ell}{\ell_0}}} \Leftrightarrow 2(1-b) < x < 2(1+b) \Leftrightarrow x^2-4x+4(1-b^2) \leq 0 \\ &\Leftrightarrow \frac{x^2}{4} \leq b^2-(1-x) \Leftrightarrow \frac{(\ell\eta)^2}{4} \leq b^2-\ell c.\end{aligned}$$

(iii)

$$\frac{(\ell\eta)^2}{4} \geq b^2 - \ell c \Leftrightarrow \frac{x^2}{4} \geq b^2 - (1-x) \Leftrightarrow x^2 - 4x + 4(1-b^2) \geq 0 \Rightarrow x \leq 2(1-b)$$

$$\Leftrightarrow \ell c \geq \frac{3 - \sqrt{1 + 4\frac{\ell_0}{\ell}}}{1 + \sqrt{1 + 4\frac{\ell}{\ell_0}}}$$

(since $x \geq 2(1+b)$ cannot hold).

(iv) The hypotheses in (19.2.4) and (19.2.2) imply $\ell\eta \leq b^2 - \ell c$ which is

$$c + \eta \leq \frac{4\ell}{\left(\ell + \sqrt{\ell^2 + 4\ell_0\ell}\right)^2}.$$

(v) The hypothesis in (19.2.5) and (19.2.3) imply

$$\ell c + \sqrt{\ell\eta} \leq 1.$$

■

We need the following result on majorizing sequences for the Secant method (19.1.2).

Lemma 19.2.2. *Let* $\ell_0 > 0$*,* $\ell > 0$*,* $c > 0$*, and* $\eta > 0$ *be constants with* $\ell_0 \leq \ell$*.*

Suppose:

$$c + \eta \leq \frac{4\ell^2}{\ell + \sqrt{\ell^2 + 4\ell_0\ell}}. \tag{19.2.14}$$

Then, scalar sequence $\{t_n\}$ *(*$n \geq -1$*) given by*

$$t_{-1} = 0,\ t_0 = c,\ t_1 = c + \eta,\ t_{n+2} = t_{n+1} + \frac{\ell\,(t_{n+1} - t_{n-1})\,(t_{n+1} - t_n)}{1 - \ell_0\,(t_{n+1} - t_0 + t_n)} \tag{19.2.15}$$

is increasing, bounded from above by

$$t^{\star\star} = \frac{\eta}{1-b} + c, \tag{19.2.16}$$

and

converges to its unique least upper bound $t^\star$ *such that*

$$c + \eta \leq t^\star \leq t^{\star\star}, \tag{19.2.17}$$

Moreover, the following estimates hold for all $n \geq 0$*:*

$$0 \leq t_{n+2} - t_{n+1} \leq b\,(t_{n+1} - t_n) \leq b^{n+1}\,\eta, \tag{19.2.18}$$

where b is given in Lemma 19.2.1.

Proof. We shall show using induction on $k \geq 0$ that

$$0 \leq t_{k+2} - t_{k+1} \leq b\,(t_{k+1} - t_k). \tag{19.2.19}$$

Using (19.2.15) for $k = 0$, we must show

$$0 < \frac{\ell\,(t_1 - t_{-1})}{1 - \ell_0\, t_1} \leq b$$

or

$$0 < \frac{\ell\,(c + \eta)}{1 - \ell_0\,(c + \eta)} \leq b,$$

which is true by (19.2.1) and (19.2.14). Let assume that (19.2.19) holds for $k \leq n + 1$.

It then follows from the induction hypotheses that

$$\begin{aligned} t_{k+2} &\leq t_{k+1} + b\,(t_{k+1} - t_k) \\ &\leq t_k + b\,(t_k - t_{k-1}) + b\,(t_{k+1} - t_k) \\ &\leq t_1 + b\,(t_1 - t_0) + \cdots + b\,(t_{k+1} - t_k) \\ &\leq c + \eta + b\,\eta + \cdots + b^{k+1}\,\eta \\ &= c + \frac{1 - b^{k+2}}{1 - b}\,\eta < \frac{\eta}{1 - b} + c = t^{\star\star}. \end{aligned} \tag{19.2.20}$$

Moreover, we can have:

$$\begin{aligned} &\ell\,(t_{k+2} - t_{k+1}) + b\,\ell_0\,(t_{k+2} - t_0 + t_{k+1}) \\ &\leq \ell\left((t_{k+2} - t_{k+1}) + (t_{k+1} - t_k)\right) + b\,\ell_0\left(\frac{1 - b^{k+2}}{1 - b} + \frac{1 - b^{k+1}}{1 - b}\right)\eta + b\,\ell_0\,c \\ &\leq \ell\,(b^k + b^{k+1})\,\eta + \frac{b\,\ell_0}{1 - b}\,(2 - b^{k+1} - b^{k+2})\,\eta + b\,\ell_0\,c. \end{aligned} \tag{19.2.21}$$

In view of (19.2.21), inequality (19.2.19) holds, if

$$\ell\,(b^k + b^{k+1})\,\eta + \frac{b\,\ell_0}{1 - b}\,(2 - b^{k+1} - b^{k+2})\,\eta + b\,\ell_0\,c \leq b \tag{19.2.22}$$

or

$$\ell\,(b^{k-1} + b^k)\,\eta + \ell_0\left((1 + b + \cdots + b^k) + (1 + b + \cdots + b^{k+1})\right)\eta + \ell_0\,c - 1 \leq 0. \tag{19.2.23}$$

In view of (19.2.23), we are motivated to define recurrent functions for $k \geq 1$ on $[0, 1)$ by

$$f_k(t) = \ell\,(t^{k-1} + t^k)\,\eta + \ell_0\left(2\,(1 + t + \cdots + t^k) + t^{k+1}\right)\eta + \ell_0\,c - 1. \tag{19.2.24}$$

We need the relationship between two consecutive functions f_k. Using (19.2.24), we obtain

$$\begin{aligned} f_{k+1}(t) &= \ell\,(t^k+t^{k+1})\,\eta+\ell_0\left(2\,(1+t+\cdots+t^{k+1})+t^{k+2}\right)\eta+\ell_0\,c-1 \\ &= \ell\,(t^{k-1}+t^k)\,\eta+\ell\,(t^k+t^{k+1})\,\eta-\ell\,(t^{k-1}+t^k)\,\eta \\ &\quad +\ell_0\left(2\,(1+t+\cdots+t^k)+t^{k+1}\right)\eta+\ell_0\,(2\,t^{k+1}+t^{k+2})\,\eta \\ &\quad -\ell_0\,t^{k+1}\,\eta+\ell_0\,c-1 \\ &= f_k(t)+\ell\,(t^{k+1}-t^{k-1})\,\eta+\ell_0\,(t^{k+1}+t^{k+2})\,\eta \\ &= p(t)\,t^{k-1}\,\eta+f_k(t), \end{aligned} \tag{19.2.25}$$

where $p(t)=\ell_0t^3+(\ell_0+\ell)t^2-\ell$. Notice that by Descarte's rule of signs, b is the only positive root of polynomial p. We can show instead of (19.2.23)

$$f_k(b)\leq 0 \qquad k\geq 1. \tag{19.2.26}$$

Define functions f_∞ on interval $[0,1)$ by $f_\infty(t)=\lim_{k\to\infty} f_k(t)$. Then, in view of (19.2.24) we get that

$$f_\infty(t)=\frac{2\ell_0\eta}{1-t}+\ell_0c-1. \tag{19.2.27}$$

We have that $f_k(b)=f_{k+1}(b)=f_\infty(b)$. Hence, we can show instead of (19.2.26) that $f_\infty(b)\leq 0$, which is true by (19.2.1), (19.2.14) and (19.2.27). Hence, we showed sequence $\{t_n\}$ $(n\geq -1)$ is increasing and bounded from above by $t^{\star\star}$, so that (19.2.18) holds. It follows that there exists $t^\star\in[c+\eta,t^{\star\star}]$, so that $\lim_{n\longrightarrow\infty} t_n=t^\star$. ■

We denote by $U(z,\rho)$ the open ball centered ar $z\in\mathcal{X}$ and of radius $\rho>0$. We also denote by $\bar{U}(z,\rho)$ the closure of $U(z,\rho)$. We shall study the Secant method (19.1.2) for triplets (F,x_{-1},x_0) belonging to the class $\mathcal{C}(\ell,\ell_0,\eta,c)$ defined as follows:

Definition 19.2.3. *Let ℓ, ℓ_0, η, c be positive constants satisfying the hypotheses of Lemma 19.2.2.*

We say that a triplet (F,x_{-1},x_0) belongs to the class $\mathcal{C}(\ell,\ell_0,\eta,c)$ if:

(c_1) F is a nonlinear operator defined on a convex subset $\mathcal{D}$ of a Banach space $\mathcal{X}$ with values in a Banach space $\mathcal{Y}$;

(c_2) x_{-1} and x_0 are two points belonging to the interior $\mathcal{D}^0$ of $\mathcal{D}$ and satisfying the inequality

$$\parallel x_0-x_{-1}\parallel\leq c;$$

(c_3) F is Fréchet–differentiable on $\mathcal{D}^0$, and there exists an operator $\delta F:\ \mathcal{D}^0\times\mathcal{D}^0\to\mathcal{L}(\mathcal{X},\mathcal{Y})$ such that:

the linear operator $A=\delta F(x_{-1},x_0)$ is invertible, its inverse A^{-1} is bounded and:

$$\begin{aligned} \parallel A^{-1}\,F(x_0)\parallel &\leq \eta; \\ \parallel A\,[\delta F(x,y)-F'(z)]\parallel &\leq \ell\,(\parallel x-z\parallel+\parallel y-z\parallel); \\ \parallel A\,[\delta F(x,y)-F'(x_0)]\parallel &\leq \ell_0\,(\parallel x-x_0\parallel+\parallel y-x_0\parallel) \end{aligned}$$

for all $x,y,z\in\mathcal{D}$.

(c_4) the set $\mathcal{D}_c = \{x \in \mathcal{D};\ F$ is continuous at $x\}$ contains the closed ball $\overline{U}(x_0, t^\star - t_0)$, where $t^\star$ is given in Lemma 19.2.2.

We present the following semilocal convergence theorem for Secant method (19.1.2).

Theorem 19.2.4. *If $(F, x_{-1}, x_0) \in \mathcal{C}(\ell, \ell_0, \eta, c)$, then sequence $\{x_n\}$ $(n \geq -1)$ generated by Secant method (19.1.2) is well defined, remains in $\overline{U}(x_0, t^\star - t_0)$ for all $n \geq 0$ and converges to a unique solution $x^\star \in \overline{U}(x_0, t^\star - t_0)$ of equation $F(x) = 0$. Moreover the following estimates hold for all $n \geq 0$*

$$\| x_{n+2} - x_{n+1} \| \leq t_{n+2} - t_{n+1}, \tag{19.2.28}$$

and

$$\| x_n - x^\star \| \leq t^\star - t_n \tag{19.2.29}$$

where the sequence $\{t_n\}$ $(n \geq 0)$ given by (19.2.15). Furthermore, if there exists $R \geq t^\star - t_0$, such that

$$\ell_0\ (c + \frac{\eta}{1-b} + R) \leq 1, \tag{19.2.30}$$

and

$$U(x_0, R) \subseteq \mathcal{D}, \tag{19.2.31}$$

then, the solution $x^\star$ is unique in $\overline{U}(x_0, R)$.

Proof. We first show operator $L = \delta F(u, v)$ is invertible for $u, v \in \overline{U}(x_0, t^\star - t_0)$. It follows from (19.2.1), (c_2) and (c_3) that:

$$\begin{aligned} \| I - A^{-1}\, L \| = \| A^{-1}\, (L - A) \| &\leq \| A^{-1}(L - F'(x_0)) \| + \| A^{-1}(F'(x_0) - A) \| \\ &\leq \ell_0\ (\| u - x_0 \| + \| v - x_0 \| + \| x_0 - x_{-1} \|) \\ &\leq \ell_0\ (t^\star - t_0 + t^\star - t_0 + c) \\ &\leq \ell_0 \left(2 \left(\frac{\eta}{1-b} + c \right) - c. \right) < 1 \end{aligned} \tag{19.2.32}$$

According to the Banach Lemma on invertible operators [8], [15], and (19.2.32), L is invertible and

$$\| L^{-1}\, A \| \leq \left(1 - \ell_0\ (\| x_k - x_0 \| + \| x_{k+1} - x_0 \| + c) \right)^{-1}. \tag{19.2.33}$$

The second condition in (c_3) implies the Lipschitz condition for F'

$$\| A^{-1}\ (F'(u) - F'(v)) \| \leq 2\, \ell\ \| u - v \|, \quad u, v \in \mathcal{D}^0. \tag{19.2.34}$$

By the identity,

$$F(x) - F(y) = \int_0^1 F'(y + t(x-y))\ dt\ (x - y) \tag{19.2.35}$$

we get

$$\| A_0^{-1}\ [F(x) - F(y) - F'(u)(x-y)] \| \leq \ell\ (\| x - u \| + \| y - u \|)\ \| x - y \| \tag{19.2.36}$$

and

$$\| A_0^{-1} \, [F(x) - F(y) - \delta F(u,v) \, (x-y)] \, \| \leq \ell \, (\| \, x - v \, \| + \| \, y - v \, \| + \| \, u - v \, \|) \, \| \, x - y \, \| \tag{19.2.37}$$

for all $x, y, u, v \in \mathcal{D}^0$. By a continuity argument (19.2.34)–(19.2.37) remain valid if x and/or y belong to $\mathcal{D}_c$. We first show (19.2.28). If (19.2.28) holds for all $n \leq k$ and if $\{x_n\}$ $(n \geq 0)$ is well defined for $n = 0, 1, 2, \cdots, k$ then

$$\| \, x_0 - x_n \, \| \leq t_n - t_0 < t^\star - t_0, \quad n \leq k. \tag{19.2.38}$$

That is (19.1.2) is well defined for $n = k+1$. For $n = -1$, and $n = 0$, (19.2.28) reduces to $\| \, x_{-1} - x_0 \, \| \leq c$, and $\| \, x_0 - x_1 \, \| \leq \eta$. Suppose (19.2.28) holds for $n = -1, 0, 1, \cdots, k$ $(k \geq 0)$. Using (19.2.33), (19.2.37) and

$$F(x_{k+1}) = F(x_{k+1}) - F(x_k) - \delta F(x_{k-1}, x_k) \, (x_{k+1} - x_k) \tag{19.2.39}$$

we obtain in turn:

$$\begin{aligned} \| \, A^{-1} F(x_{k+1}) \, \| &= \ell(\| \, x_{k+1} - x_k \, \| + \| \, x_k - x_{k-1} \, \|) \, \| \, x_{k+1} - x_k \, \| \\ &= \ell(t_{k+1} - t_k + t_k - t_{k-1})(t_{k+1} - t_k) \\ &= \ell(t_{k+1} - t_{k-1})(t_{k+1} - t_k) \end{aligned} \tag{19.2.40}$$

and

$$\begin{aligned} \| \, x_{k+2} - x_{k+1} \, \| &= \| \, \delta F(x_k, x_{k+1})^{-1} \, F(x_{k+1}) \, \| \\ &\leq \| \, \delta F(x_k, x_{k+1})^{-1} \, A \, \| \, \| \, A^{-1} \, F(x_{k+1}) \, \| \\ &\leq \frac{\ell \, (t_{k+1} - t_k + t_k - t_{k-1})}{1 - \ell_0 \, (t_{k+1} - t_0 + t_k - t_0 + t_0 - t_{-1})} \, (t_{k+1} - t_k) \\ &= t_{k+2} - t_{k+1}. \end{aligned} \tag{19.2.41}$$

The induction for (19.2.28) is completed. It follows from (19.2.28) and Lemma 19.2.2 that sequence $\{x_n\}$ $(n \geq -1)$ is complete in a Banach space $\mathcal{X}$, and as such it converges to some $x^\star \in \overline{U}(x_0, t^\star - t_0)$ (since $\overline{U}(x_0, t^\star - t_0)$ is a closed set). By letting $k \to \infty$ in (19.2.41), we obtain $F(x^\star) = 0$. Estimate (19.2.29) follows from (19.2.28) by using standard majoration techniques [7, 15, 18, 22]. We shall first show uniqueness in $\overline{U}(x_0, t^\star - t_0)$. Let $y^\star \in \overline{U}(x_0, t^\star - t_0)$ be a solution of equation (19.1.1).

Set

$$\mathcal{M} = \int_0^1 F'(y^\star + t \, (y^\star - x^\star)) \, dt.$$

It then by (c_3):

$$\begin{aligned} \| \, A^{-1} \, (A - \mathcal{M}) \, \| &= \ell_0 \, (\| \, y^\star - x_0 \, \| + \| \, x^\star - x_0 \, \| + \| \, x_0 - x_{-1} \, \|) \\ &\leq \ell_0 \, ((t^\star - t_0) + (t^\star - t_0) + t_0) \\ &\leq \ell_0 \left(2 \left(\frac{\eta}{1-b} + c \right) - c \right) \\ &= \ell_0 \left(\frac{2 \, \eta}{1-b} + c \right) < 1. \end{aligned} \tag{19.2.42}$$

It follows from (19.2.1), and the Banach lemma on invertible operators that $\mathcal{M}^{-1}$ exists on $\overline{U}(x_0, t^\star - t_0)$. Using the identity:

$$F(x^\star) - F(y^\star) = \mathcal{M}\ (x^\star - y^\star) \tag{19.2.43}$$

we deduce $x^\star = y^\star$. Finally, we shall show uniqueness in $U(x_0, R)$. As in (19.2.42), we arrive at

$$\| A^{-1}\ (A - \mathcal{M})\ \| < \ell_0 \left(\frac{\eta}{1-b} + c + R \right) \leq 1,$$

by (19.2.30). ∎

Remark 19.2.5. **(a)** *Let us define the majoring sequence $\{w_n\}$ used in earlier studies such as [2, 14, 16, 18, 20, 21, 22, 23, 26, 27] (under condition (19.1.3)):*

$$w_{-1} = 0,\ w_0 = c,\ w_1 = c + \eta,\ w_{n+2} = w_{n+1} + \frac{\ell\ (w_{n+1} - w_{n-1})\ (w_{n+1} - w_n)}{1 - \ell\ (w_{n+1} - w_0 + w_n)}. \tag{19.2.44}$$

Note that in general

$$\ell_0 \leq \ell \tag{19.2.45}$$

holds, and $\dfrac{\ell}{\ell_0}$ can be arbitrarily large [5, 6, 7, 8]. In the case $\ell_0 = \ell$, then $t_n = w_n$ $(n \geq -1)$. Otherwise:

$$t_{n+1} - t_n \leq w_{n+1} - w_n, \tag{19.2.46}$$

$$0 \leq t^\star - t_n \leq w^\star - w_n, \quad w^\star = \lim_{n \longrightarrow \infty} w_n. \tag{19.2.47}$$

Note also that strict inequality holds in (19.2.46) for $n \geq 1$, if $\ell_0 < \ell$. It is worth noticing that the center-Lipschitz condition is not an additional hypothesis to the Lipschitz condition, since in practice the computation of constant ℓ requires the computation of ℓ_0. It follows from the proof of Theorem 19.2.4 that sequence $\{s_n\}$ defined by

$$s_{-1} = 0, \qquad s_0 = c, \qquad s_1 = c + \eta, \qquad s_2 = s_1 + \frac{\ell_0 (s_1 - s_{-1})(s_1 - s_0)}{1 - \ell_0 s_1}$$

$$s_{n+2} = s_{n+1} + \frac{\ell (s_{n+1} - s_{n-1})(s_{n+1} - s_n)}{1 - \ell_0 (s_{n+1} - s_0 + s_n)} \qquad \text{for } n = 1, 2, \ldots.$$

is also a majorizing sequence for $\{x_n\}$ which is tighter than $\{t_n\}$.

(b) *In practice constant c depends on initial guesses x_{-1} and x_0 which can be chosen to be as close to each other as we wish. Therefore, in particular, we can always choose*

$$\ell c < \frac{3 - \sqrt{1 + 4\dfrac{\ell_0}{\ell}}}{1 + \sqrt{1 + 4\dfrac{\ell_0}{\ell}}},$$

which according to (iv) in Lemma 19.2.1 implies that the new sufficient convergence criterion (19.2.14) is weaker than the old one given by (19.1.3).

19.3. Numerical Examples

Example 19.3.1. *Let $\mathcal{X} = \mathcal{Y} = \mathcal{C}[0,1]$, equipped with the max-norm. Consider the following nonlinear boundary value problem*

$$\begin{cases} u'' = -u^3 - \gamma\, u^2 \\ u(0) = 0, \quad u(1) = 1. \end{cases}$$

It is well known that this problem can be formulated as the integral equation

$$u(s) = s + \int_0^1 Q(s,t)\,(u^3(t) + \gamma\, u^2(t))\,dt \tag{19.3.1}$$

where, Q is the Green function:

$$Q(s,t) = \begin{cases} t\,(1-s), & t \le s \\ s\,(1-t), & s < t. \end{cases}$$

We observe that

$$\max_{0\le s\le 1} \int_0^1 |Q(s,t)|\,dt = \frac{1}{8}.$$

Then problem (19.3.1) is in the form (19.1.1), where, $F\,:\,\mathcal{D} \longrightarrow \mathcal{Y}$ is defined as

$$[F(x)]\,(s) = x(s) - s - \int_0^1 Q(s,t)\,(x^3(t) + \gamma\, x^2(t))\,dt.$$

The Fréchet derivative of the operator F is given by

$$[F'(x)y]\,(s) = y(s) - 3\int_0^1 Q(s,t)x^2(t)y(t)dt - 2\gamma\int_0^1 Q(s,t)x(t)y(t)dt.$$

Then, we have that

$$[(I - F'(x_0))(y)](s) = 3\int_0^1 Q(s,t)x_0^2(t)y(t)dt + 2\gamma\int_0^1 Q(s,t)x_0(t)y(t)dt.$$

Hence, if $2\gamma < 5$, then

$$\|I - F'(x_0)\| \le 2(\gamma - 2) < 1.$$

It follows that $F'(x_0)^{-1}$ exists and

$$\|F'(x_0)^{-1}\| \le \frac{1}{5-2\gamma}.$$

We also have that $\|F(x_0)\| \le 1+\gamma$. Define the divided difference defined by

$$\delta F(x,y) = \int_0^1 F'(y + t(x-y))dt.$$

Choosing $x_{-1}(s)$ such that $\|x_{-1} - x_0\| \le c$ and $k_0 c < 1$. Then, we have

$$\|\delta F(x_{-1},x_0)^{-1}F(x_0)\| \le \|\delta F(x_{-1},x_0)^{-1}F'(x_0)\|\|F'(x_0)F(x_0)\|$$

and

$$\|\delta F(x_{-1},x_0)^{-1}F'(x_0)\| \leq \frac{1}{(1-k_0c)},$$

where k_0 is such that

$$\|F'(x_0)^{-1}(F'(x_0)-A_0)\| \leq k_0c,$$

Set $u_0(s)=s$ and $\mathcal{D}=U(u_0,R)$. It is easy to verify that $U(u_0,R)\subset U(0,R+1)$ since $\| u_0 \|=1$. If $2\,\gamma<5$, and $k_0c<1$ the operator F' satisfies conditions of Theorem 19.2.6, with

$$\eta=\frac{1+\gamma}{(1-k_0c)(5-2\,\gamma)},\quad l=\frac{\gamma+6\,R+3}{8(5-2\,\gamma)(1-k_0c)},\quad l_0=\frac{2\,\gamma+3\,R+6}{16(5-2\,\gamma)(1-k_0c)}.$$

Choosing $R_0=0.9$, $\gamma=0.5$ and $c=1$ we obtain that

$$k_0=0.1938137822\ldots,$$

$$\eta=0.465153\ldots,$$

$$l=0.344989\ldots$$

and

$$l_0=0.187999\ldots.$$

Then, criterion (19.1.3) is not satisfied since $lc+2\sqrt{l\eta}=1.14617\ldots>1$, but criterion (19.2.14) is satisfied since

$$\eta+c=1.46515\ldots\leq\frac{4l}{(l^2+\sqrt{l^2+4l_0l})^2}=1.49682\ldots.$$

As a consequence the convergence of the secant-method is guaranteed by Theorem 19.2.4.

Example 19.3.2. *Let $\mathcal{X}=\mathcal{Y}=\mathbb{R}$ and let consider the real functions*

$$F(x)=x^3-k$$

where $k\in\mathbb{R}$ and we are going to apply secant-method to find the solution of $F(x)=0$. We take the starting point $x_0=1$ we consider the domain $\Omega=B(x_0,1)$ and we let x_{-1} free in order to find a relation between k and x_{-1} for which criterion (19.1.3) is not satisfied but new criterion (19.2.14) is satisfied. In this case, we obtain

$$\eta=|(1-k)(1+x_{-1}+x_{-1}^2)|,$$

$$l=\frac{6}{|1+x_{-1}+x_{-1}^2|},$$

$$l_0=\frac{9}{2|1+x_{-1}+x_{-1}^2|},$$

Taking all this data into account we obtain the following criteria:

(i) *If* $55/54 < k \leq 25/24$ *and*

$$\alpha < x_{-1} \leq \frac{2-27k}{2(-29+27k)} - \frac{1}{2}\sqrt{3}\sqrt{-\frac{2164-3024k+729k^2}{(-29+27k)^2}},$$

where α *is the smallest positive root of*

$$p(t)-73+24k+(22+48k)t+(-111+72k)t^2+(-38+48k)t^3+(-25+24k)t^4.$$

(ii) *If* $25/24 < k < 29/27$ *and*

$$1 < x_1 \leq \frac{2-27k}{2(-29+27k)} - \frac{1}{2}\sqrt{3}\sqrt{-\frac{2164-3024k+729k^2}{(-29+27k)^2}}.$$

(iii) *If* $55/54 < k < 25/24$ *and*

$$\frac{56-27k}{2(-29+27k)} + \frac{1}{2}\sqrt{3}\sqrt{-\frac{-968-108k+729k^2}{(-29+27k)^2}} \leq x_{-1} < \alpha,$$

where α *is the greatest positive root of*

$$p(t) = -49+24k+(22+48k)t+(-111+72k)t^2+(-62+48k)t^3+(-25+24k)t^4.$$

(iv) *If* $25/24 \leq k < 29/27$ *and*

$$\frac{56-27k}{2(-29+27k)} + \frac{1}{2}\sqrt{3}\sqrt{-\frac{-968-108k+729k^2}{(-29+27k)^2}} \leq x_{-1} < 1.$$

(v) *If* $25/27 < k < 23/24$ *and*

$$1 \leq x_{-1} < \frac{52-27k}{2(-25+27k)} - \frac{1}{2}\sqrt{3}\sqrt{-\frac{-968+108k+729k^2}{(-25+27k)^2}}.$$

(vi) *If* $23/24 \leq k < 53/54$ *and*

$$\alpha \leq x_{-1} < \frac{52-27k}{2(-25+27k)} - \frac{1}{2}\sqrt{3}\sqrt{-\frac{-968+108k+729k^2}{(-25+27k)^2}},$$

where α *is the smallest positive root of*

$$p(t) = 25+24k+(-118+48k)t+(-33+72k)t^2+(-58+48k)t^3+(-23+24k)t^4.$$

(vii) *If* $25/27 < k \leq 23/24$ *and*

$$\frac{-2-27k}{2(-25+27k)} + \frac{1}{2}\sqrt{3}\sqrt{-\frac{1732-2808k+729k^2}{(-25+27k)^2}} \leq x_{-1} < 1.$$

(viii) *If* $23/24 < k < 53/54$ *and*

$$\frac{-2-27k}{2(-25+27k)} + \frac{1}{2}\sqrt{3}\sqrt{-\frac{1732-2808k+729k^2}{(-25+27k)^2}} \leq x_{-1} < \alpha,$$

where α *is the greatest positive root of*

$$p(t) = 1+24k+(-118+48k)t+(-33+72k)t^2+(-34+48k)t^3+(-23+24k)t^4.$$

Now we consider a case in which both criteria (19.1.3) and (19.2.14) are satisfied to compare the majorizing sequences. We choose $k = 0.99$ *and* $x_{-1} = 1.2$ *and we obtain*

$$c = 0.2, \qquad \eta = 0.0364\ldots, \qquad l = 1.64835, \qquad l_0 = 1.23626.$$

Moreover, criterion (19.1.3)

$$lc + 2\sqrt{l\eta} = 0.819568 < 1,$$

is satisfied and criterion (19.2.14)

$$c+\eta = 0.2364\ldots \leq 0.26963\ldots = \frac{4l}{(l^2+\sqrt{l^2+4l_0 l})^2},$$

is also satisfied. In Table 19.3.1 it is shown that $\{s_n\}$*,* $\{t_n\}$ *and* $\{w_n\}$ *are majorizing sequences and it is shown also that the tighter sequence is* $\{s_n\}$*.*

Table 19.3.1. Comparison between the sequences $\{s_n\}$**,** $\{t_n\}$ **and** $\{w_n\}$

n	$\|s_{n+1}-s_n\|$	$\|t_{n+1}-t_n\|$	$\|w_{n+1}-w_n\|$
1	$0.0150308\ldots$	$0.0200411\ldots$	$0.0232399\ldots$
2	$0.00197814\ldots$	$0.00292257\ldots$	$0.00446203\ldots$
3	$0.0000890021\ldots$	$0.000181477\ldots$	$0.000339709\ldots$
4	4.88677×10^{-7}	1.53289×10^{-6}	4.52784×10^{-6}
5	1.16179×10^{-10}	7.63675×10^{-10}	4.32958×10^{-9}
6	1.66533×10^{-16}	3.16414×10^{-15}	5.45120×10^{-14}

Conclusion

We present a new semilocal convergence analysis for the secant method in order to approximate a locally unique solution of a nonlinear equation in a Banach space setting. We showed that the new convergence criteria can be always weaker than the corresponding ones in earlier studies such as [2, 14, 16, 18, 20, 21, 22, 23, 26, 27]. Numerical examples where the old results cannot guarantee the convergence but our new convergence criteria can are also provided in this chapter.

References

[1] Amat, S., Busquier, S., Negra, M., Adaptive approximation of nonlinear operators, *Numer. Funct. Anal. Optim.*, 25 (2004), 397-405.

[2] Amat, S., Busquier, S., Gutiérrez, J. M., On the local convergence of secant-type methods, *Int. J. Comp. Math.*, 81(8) (2004), 1153-1161.

[3] Amat, S., Busquier, S., Magreñán, Á.A., Reducing chaos and bifurcations in Newton-type methods, *Abst. Appl. Anal.*, 2013 (2013), Article ID 726701, 10 pages, http://dx.doi.org/10.1155/2013/726701.

[4] Amat, S., Magreñán, Á.A., Romero, N., On a family of two step Newton-type methods, *App. Math. Comp.*, 219(4) (2013), 11341-11347.

[5] Argyros, I.K., A unifying local–semilocal convergence analysis and applications for two–point Newton–like methods in Banach space, *J. Math. Anal. Appl.*, 298 (2004), 374-397.

[6] Argyros, I.K., New sufficient convergence conditions for the Secant method, *Chechoslovak Math. J.*, 55 (2005), 175-187.

[7] Argyros, I.K., *Convergence and applications of Newton–type iterations*, Springer–Verlag Publ., New–York, 2008.

[8] Argyros, I.K., Weaker conditions for the convergence of Newton's method, *J. Complexity*, 28 (2012), 364-387.

[9] Bosarge, W.E., Falb, P.L., A multipoint method of third order, *J. Optimiz. Th. Appl.*, 4 (1969), 156-166.

[10] Dennis, J.E., Toward a unified convergence theory for Newton–like methods, in *Nonl. Funct. Anal. Appl.* (L.B. Rall, ed.), Academic Press, New York, (1971), 425-472.

[11] García-Olivo, M., *El método de Chebyshev para el cálculo de las raíces de ecuaciones no lineales (PhD Thesis)*, Servicio de Publicaciones, Universidad de La Rioja, 2013. http://dialnet.unirioja.es/descarga/tesis/37844.pdf

[12] Gutiérrez, J. M., Magreñán, Á.A., N. Romero, On the semilocal convergence of Newton-Kantorovich method under center-Lipschitz conditions, *App. Math. Comp.*, 221 (2013), 79-88.

[13] Hernández, M.A., Rubio, M.J., Ezquerro, J.A., Solving a special case of conservative problems by Secant–like method, *Appl. Math. Comp.*, 169 (2005), 926-942.

[14] Hernández, M.A., Rubio, M.J., Ezquerro, J.A., Secant–like methods for solving nonlinear integral equations of the Hammerstein type, *J. Comp. Appl. Math.*, 115 (2000), 245-254.

[15] Kantorovich, L.V., Akilov, G.P., *Functional Analysis*, Pergamon Press, Oxford, 1982.

[16] Laasonen, P., Ein überquadratisch konvergenter iterativer algorithmus, *Ann. Acad. Sci. Fenn. Ser* I, 450 (1969), 1–10.

[17] Magreñán, Á.A., A new tool to study real dynamics: The convergence plane, *App. Math. Comp.*, 248 (2014), 215-224.

[18] Ortega, J.M., Rheinboldt, W.C., *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970.

[19] Ostrowski, A.M., *Solution of equations in Euclidian and Banach Spaces*, Academic Press, New York, 1972.

[20] Potra, F.A., An error analysis for the secant method, *Numer. Math.*, 38 (1982), 427-445.

[21] Potra, F.A., Sharp error bounds for a class of Newton–like methods, *Libertas Mathematica*, 5 (1985), 71-84.

[22] Potra, F.A., Pták, V., *Nondiscrete Induction and Iterative Processes*, Pitman, New York, 1984.

[23] Schmidt, J.W., Untere Fehlerschranken fur Regula–Falsi Verhafren, *Period. Hungar.*, 9 (1978), 241-247.

[24] Sergeev, A.S., On the method of Chords (in Russian), Sibirsk, *Math. J.*, 11, (1961), 282–289.

[25] Ulm, S., Majorant principle and the method of Chords (in Russian), *Izv. Akad. Nauk Eston. SSR*, Ser. Fiz.-Mat., 13 (1964), 217-227.

[26] Yamamoto, T., A convergence theorem for Newton–like methods in Banach spaces, *Numer. Math.*, 51 (1987), 545-557.

[27] Wolfe, M.A., Extended iterative methods for the solution of operator equations, *Numer. Math.*, 31 (1978), 153-174.

Chapter 20

Expanding the Convergence Domain for Chun-Stanica-Neta Family of Third Order Methods in Banach Spaces

20.1. Introduction

In this study we are concerned with the problem of approximating a locally unique solution x^* of the equation

$$F(x) = 0, \tag{20.1.1}$$

where F is a Fréchet-differentiable operator defined on a convex subset D of a Banach space X with values in a Banach space Y.

Many problems in computational mathematics and other disciplines can be brought in a form like (20.1.1) using mathematical modelling [1, 3, 11, 15, 18, 19]. The solutions of these equations can rarely be found in closed form. That is why most solution methods for these equations are usually iterative. In particular the practice of Numerical Functional Analysis for finding such solutions is essentially connected to Newton-like methods [1, 3, 15, 17, 18, 19]. The study about convergence of iterative procedures is normally centered on two types: semilocal and local convergence analysis. The semilocal convergence matter is, based on the information around an initial point, to give criteria ensuring the convergence of the iterative procedures. While the local analysis is based on the information around a solution, to find estimates of the radii of convergence balls. There exist many studies which deal with the local and the semilocal convergence analysis of Newton-like methods such as [1]-[20].

Majorizing sequences in connection to the Kantorovich theorem have been used extensively for studying the convergence of these methods [1, 2, 3, 4, 11, 15, 10]. Rall [19] suggested a different approach for the convergence of these methods, based on recurrent relations. Candela and Marquina [5, 6], Parida[16], Parida and Gupta [17], Ezquerro and Hernández [7], Gutiérrez and Hernández [8, 9]. Argyros [1, 2, 3] used this idea for several high-order methods. In particular, Kou and Li [12] introduced a third order family of methods for solving equation (20.1.1), when $X = Y = \mathbb{R}$ defined by

$$\begin{aligned} y_n &= x_n - \theta F'(x_n)^{-1}F(x_n), \text{ for each } n = 0,1,2,\cdots \\ x_{n+1} &= x_n - \frac{\theta^2+\theta-1}{\theta^2}F'(x_n)^{-1}F(x_n) - \frac{1}{\theta^2}F'(x_n)^{-1}F(y_n), \end{aligned} \tag{20.1.2}$$

where x_0 is an initial point and $\theta \in \mathbb{R} - \{0\}$. This family uses two evaluations of F and one evaluation of F'. Third order methods requiring one evaluation of F and two evaluation of F' can be found in [1, 3, 12, 18]. It is well known that the convergence domain of high order methods is in general very small. This fact limits the applicability of these methods. In the present study we are motivated by this fact and recent work by Chun, Stanica and Neta [4] who provided a semilocal convergence analysis of the third order method (20.1.2) in a Banach space setting. Their semilocal convergence analysis is based on recurrent relations. In Section 20.2 we show convergence of the third order method (20.1.2) using more precise recurrent relations under less computational cost and weaker convergence criterion. Moreover, the error estimates on the distances $\|x_{n+1} - x_n\|$, $\|x_n - x^*\|$ are more precise and the information on the location of the solution at least as precise. In Section 20.3 using our technique of recurrent functions we present a semilocal convergence analysis using majorizing sequence. The convergence criterion can be weaker than the older convergence criteria or the criteria of Section 20.2. Numerical examples are presented in Section 20.4 that show the advantages of our work over the older works.

20.2. Semilocal Convergence I

Let $U(w,\rho)$, $\overline{U}(w,\rho)$ stand for the open and closed ball, respectively, with center $w \in X$ and of radius $\rho > 0$. Let also $L(X,Y)$ denote the space of bounded linear operators from X into Y.

The semilocal convergence analysis of third order method (20.1.2),given by Chun, Stanica and Neta [4] is based on the following conditions. Suppose:
$(\mathcal{C})$:

(1) There exists $\|F'(x) - F'(y)\| \le K\|x - y\|$ for each x and $y \in D$;

(2)

$$\|F''(x)\| \le M, \text{ for each } x \in D;$$

(3)

$$\|F'(x_0)^{-1}\| \le \beta;$$

(4)

$$\|F'(x_0)^{-1}F(x_0)\| \le \eta.$$

They defined certain parameters and sequences by

$$
\begin{aligned}
a &= K\beta\eta, \\
\alpha &= \frac{|\theta^2+\theta-1|+|1-\theta|}{\theta^2}, \\
\gamma &= \frac{M}{2}\beta\eta, \\
a_0 = b_0 = 1, &\qquad d_0 = \alpha+\gamma,\ b_{-1} = 0, \\
a_{n+1} &= \frac{a_n}{1-aa_nd_n}, \\
b_{n+1} &= a_{n+1}\beta\eta c_n, \\
k_n &= \frac{|1+\theta|(\theta-1)^2+|1-\theta|}{\theta^2}b_n+\frac{M}{2}a_n\beta b_n^2\eta, \\
c_n &= \frac{M}{2}k_n^2+K|\theta|b_nk_n+\frac{M}{2}|\theta^2-1|b_n^2
\end{aligned}
$$

and

$$d_{n+1} = \alpha b_{n+1}+\gamma a_{n+1}b_{n+1}^2.$$

We suppose $(\mathcal{C}^0)$:

(1)

$$\|F'(x_0)^{-1}(F'(x)-F'(y))\| \le K\|x-y\|, \text{ for each } x,y \in D;$$

(2)

$$\|F'(x_0)^{-1}(F'(x)-F'(x_0))\| \le K_0\|x-x_0\|, \text{ for each } x \in D;$$

(3)

$$\|F'(x_0)^{-1}F(x_0)\| \le \eta.$$

Notice that the new conditions are given in affine invariant form and the condition on the second Fréchet-derivative has been dropped. The advantages of presenting results in affine invariant form instead of non-affine invariant form are well known [1, 3, 11, 15, 18]. If operator F is twice Fréchet differentiable, then (1) in $(\mathcal{C}^0)$ implies (2) in $(\mathcal{C})$.

In order for us to compare the old approach with the new, let us rewrite the conditions $(\mathcal{C})$ in affine invariant form. We shall call these conditions again $(\mathcal{C})$.

$(\mathcal{C}_1)$ $\|F'(x_0)^{-1}(F'(x)-F'(y))\| \le K\|x-y\|$ for each x and $y \in D$;

$(\mathcal{C}_2)$

$$\|F'(x_0)^{-1}F''(x)\| \le M, \text{ for each } x \in D;$$

$(\mathcal{C}_4)$

$$\|F'(x_0)^{-1}F(x_0)\| \le \eta.$$

The parameters and sequences are defined as before but $\beta = 1$. Then, we can certainly set $K = M$. Define parameters

$$\begin{aligned}
a^0 &= K\eta, \\
\alpha^0 &= \alpha, \\
\gamma^0 &= \frac{K}{2}\eta, \\
a_0^0 &= b_0^0 = 1,\ d_0^0 = \alpha^0 + \gamma^0,\ b_{-1}^0 = 0, \\
a_{n+1}^0 &= \frac{1}{1 - K_0(d_n^0 + d_{n-1}^0 + \cdots + d_0^0)}, \\
b_{n+1}^0 &= a_{n+1}^0 \eta c_n^0, \\
c_n^0 &= K[\frac{(k_n^0)^2}{2} + |\theta| b_n^0 k_n^0 + \frac{|\theta^2 - 1|}{2}(b_n^0)^2], \\
k_n^0 &= \frac{|\theta+1|(\theta-1)^2 + |1-\theta|}{\theta^2} b_n^0 + \frac{K}{2} a_n^0 (b_n^0)^2 \eta
\end{aligned}$$

and

$$d_{n+1}^0 = \alpha^0 b_{n+1}^0 + \gamma^0 a_{n+1}^0 (b_{n+1}^0)^2.$$

We have that

$$K_0 \le K \tag{20.2.1}$$

holds in general and $\frac{K}{K_0}$ can be arbitrarily large [1]-[3]. Notice that the center Lipschitz condition is not an additional condition to the Lipschitz condition, since in practice the computation of K involves the computation of K_0 as a special case. We have by the definition of a_{n+1} in turn that

$$\begin{aligned}
a_{n+1} &= \frac{a_n}{1 - K\eta a_n d_n} \\
&= \frac{a_n}{1 - K\eta d_n \frac{a_{n-1}}{1 - K\eta a_{n-1} d_{n-1}}} \\
&= \frac{a_n (1 - K\eta a_{n-1} d_{n-1})}{1 - K\eta a_{n-1}(d_n + d_{n-1})} \\
&= \frac{\frac{a_{n-1}}{1 - K\eta a_{n-1} d_{n-1}} (1 - K\eta a_{n-1} d_{n-1})}{1 - K\eta a_{n-1}(d_n + d_{n-1})} \\
&= \frac{a_{n-1}}{1 - K\eta a_{n-1}(d_n + d_{n-1})} \\
&\ \vdots \\
&= \frac{a_0}{1 - K\eta a_{n-1}(d_n + d_{n-1} + \cdots + d_0)} \\
&= \frac{1}{1 - K\eta(d_n + d_{n-1} + \cdots + d_0)}.
\end{aligned}$$

Hence, we deduce that

$$a_{n+1}^0 \le a_{n+1} \text{ for each } n = 0, 1, 2, \cdots. \tag{20.2.2}$$

Moreover, strict inequality holds in (20.2.2) if $K_0 < K$. Hence, using a simple inductive argument we also have that

$$\begin{aligned} b^0_{n+1} &\leq b_{n+1}, &(20.2.3)\\ c^0_n &\leq c_n, &(20.2.4)\\ k^0_n &\leq k_n &(20.2.5)\end{aligned}$$

and

$$d^0_{n+1} \leq d_{n+1}. \quad (20.2.6)$$

Lemma 20.2.1. *Under the ($\mathcal{C}^0$) conditions the following hold*

$$\begin{aligned} \|F'(x_n)^{-1}F'(x_0)\| &\leq a^0_n,\\ \|F'(x_n)^{-1}F(x_n)\| &\leq b^0_n\eta,\\ \|x_{n+1}-x_n\| &\leq d^0_n\eta,\\ \|x_{n+1}-y_n\| &\leq (d^0_n+2k^0_{n-1}+\theta b^0_n)\eta.\end{aligned}$$

Moreover, under the ($\mathcal{C}$) conditions the following hold

$$\begin{aligned} \|F'(x_n)^{-1}F'(x_0)\| &\leq a^0_n \leq a_n,\\ \|F'(x_n)^{-1}F(x_n)\| &\leq b^0_n\eta \leq b_n\eta,\\ \|x_{n+1}-x_n\| &\leq d^0_n\eta \leq d_n\eta,\\ \|x_{n+1}-y_n\| &\leq (d^0_n+2k^0_{n-1}+\theta b^0_n)\eta \leq (d_n+2k_{n-1}+\theta b_n)\eta.\end{aligned}$$

Proof. It follows from the proof of Lemma 1 in [4] by simply noticing: the expressions involving

(i) the second Fréchet-derivative

$$\int_0^1 F''(x_n+t(y_n-x_n))(1-t)(y_n-x_n)^2dt$$

and

$$\int_0^1 F''(y_n+t(x_{n+1}-y_n))(1-t)(x_{n+1}-y_n)^2dt$$

are not needed and can be replaced, respectively, by

$$\int_0^1 [F'(y_n+t(x_n-y_n)-F'(x_n)](y_n-x_n)dt$$

and

$$\int_0^1 [F'(y_n+t(x_{n+1}-y_n)-F'(y_n)](x_{n+1}-y_n)dt.$$

Hence, condition (2) in ($\mathcal{C}$) is not needed and can be replaced by condition (1) in ($\mathcal{C}^0$) to produce the same bounds as in [4] (for $K = M$) (see also the proof of Theorem 20.3.2 that follows)

(ii) The computation of the upper bounds on $\|F'(x_n)^{-1}F'(x_0)\|$ in [4] uses condition (1) in $(\mathcal{C})$ and the estimate

$$\|F'(x_n)^{-1}(F'(x_n)-F'(x_{n+1}))\| \leq \|F'(x_n)^{-1}F'(x_0)\|K\|x_n-x_{n+1}\|$$

to arrive at

$$\|F'(x_n)^{-1}F'(x_0)\| \leq a_{n+1}, \tag{20.2.7}$$

whereas we use (2) in $(\mathcal{C}^0)$ and estimate

$$\begin{aligned}\|F'(x_0)^{-1}(F'(x_n)-F'(x_{n+1}))\| &\leq K_0\|x_{n+1}-x_n\| \\ &\leq K_0(\|x_{n+1}-x_n\|+\cdots+\|x_1-x_0\|) \\ &\leq K_0(d_n^0+d_{n-1}^0+\cdots+d_0^0)\end{aligned}$$

to arrive at the estimate

$$\|F'(x_n)^{-1}F'(x_0)\| \leq a_{n+1}^0, \tag{20.2.8}$$

which is more precise (see also (20.2.2)).

□

Lemma 20.2.2. *Suppose that*

$$a_1^0 b_1^0 < 1. \tag{20.2.9}$$

Then, sequence $\{p_n^0\}$ *defined by* $p_n^0 = a_n^0 b_n^0$ *is decreasingly convergent to* 0 *such that*

$$p_{n+1}^0 \leq \xi_1^{2^{n+1}} \frac{1}{\xi_1},\ \xi_1 := a_1^0 b_1^0$$

and

$$d_n^0 \leq (\alpha^0+\gamma^0)\xi_1^{2^n}\frac{1}{\xi_1}.$$

Moreover, if

$$a_1 b_1 < 1, \tag{20.2.10}$$

then, sequence $\{p_n\}$ *defined by* $p_n = a_n b_n$ *is also decreasingly convergent to* 0 *such that*

$$p_{n+1}^0 \leq p_{n+1} \leq \xi^{2^{n+1}}\frac{1}{\xi},\ \xi = a_1 b_1,$$

$$d_n^0 \leq d_n \leq (\alpha+\gamma)\xi^{2^n}\frac{1}{\xi},$$

and

$$\xi_1 \leq \xi.$$

Proof. It follows from the proof of Lemma 3 in [4] by simply using $\{p_n^0\}$, a_1^0, b_1^0, ξ_1 instead of $\{p_n\}$, a_1, b_1, ξ, respectively. □

Next, we present the main semilocal convergence result for the third order method (20.1.2) under the $(\mathcal{C}^0)$ conditions, (20.2.9) and the convergence criterion

$$a(\alpha+\gamma) < 1. \tag{20.2.11}$$

The proof follows from the proof of Theorem 5 in [4] (with the exception of the uniqueness of the solution part) by simply replacing the $(\mathcal{C})$ conditions and (20.2.10) by the $(\mathcal{C}^0)$ conditions and (20.2.9) respectively.

Theorem 20.2.3. *Suppose that conditions ($\mathcal{C}^0$), (20.2.9) and (20.2.11) hold. Moreover, suppose that*

$$U_0^0 = \overline{U}(x_0, r^0\eta) \subset D, \tag{20.2.12}$$

where

$$r_0 = \sum_{n=0}^{\infty} d_n^0. \tag{20.2.13}$$

Then, sequences $\{x_n\}$ generated by the third order method (20.1.2) is well defined, remains in U_0^0 for each $n = 0, 1, 2, \cdots$ and converges to a unique solution x^* of equation $F(x) = 0$ in $U(x_0, \frac{2}{K_0} - r_0\eta) \bigcap D$. Moreover, the following estimates hold

$$\|x_{n+1} - x^*\| \leq \sum_{k=n+1}^{\infty} d_k^0 \eta \leq \frac{\alpha+\gamma}{\xi_1}\eta \sum_{k=n+1}^{\infty} \xi_1^{2^k}. \tag{20.2.14}$$

Proof. As already noted above, we only need to show the uniqueness part. Let $y^* \in U(x_0, \frac{2}{K_0} - r_0\eta)$ be such that $F(y^*) = 0$. Define $Q = \int_0^1 F'(x^* + t(y^* - x^*))dt$. Using condition (2) in ($\mathcal{C}^0$) we get in turn that

$$\begin{aligned}
\|F'(x_0)^{-1}(F'(x_0) - Q)\| &\leq K_0 \int_0^1 \|x^* + t(y^* - x^*) - x_0\| dt \\
&\leq K_0 \int_0^1 [(1-t)\|x^* - x_0\| + t\|y^* - x_0\|] dt \\
&< \frac{K_0}{2}[r_0\eta + \frac{2}{K_0} - r_0\eta] = 1.
\end{aligned} \tag{20.2.15}$$

It follows from (20.2.15) and the Banach lemma on invertible operators [1, 3, 11, 15, 18] that $Q^{-1} \in L(Y, X)$. Then, using the identity

$$0 = F(x^*) - F(y^*) = Q(x^* - y^*),$$

we deduce that $x^* = y^*$. □

Remark 20.2.4. *If $K_0 = K$, and operator F is twice Fréchet differentiable then Lemma 20.2.1, Lemma 20.2.2 and Theorem 20.2.3 reduce to Lemma 1, Lemma 3 and Theorem 5 in [4], respectively. Otherwise i.e., if $K_0 < K$ or if the twice Fréchet differentiability of operator F is not assumed, then our results constitute an improvement. It is worth noticing that if $K_0 < K$, then (20.2.10) implies (20.2.9) (but not necessarily vice versa) and $\xi_1 < \xi$.*

20.3. Semilocal Convergence II

We need to introduce some scalar sequences that shall be shown to be majorizing for the third order methods (20.1.2) in Theorem 20.3.2.

Let $K_0 > 0, K > 0, \eta > 0$ and $\theta \in \mathbb{R} - \{0\}$. Set $t_0 = 0$ and $s_0 = |\theta|\eta$. Define polynomials f and g by

$$\begin{aligned}
f(t) &= (\frac{K|\theta|}{2} + K_0)t^3 + \frac{|\theta|}{2}Kt^2 + K(\frac{|\theta^2 - 1|}{2|\theta|} - |\theta|)t \\
&\quad -\frac{K}{2}\frac{|\theta^2 - 1|}{|\theta|}
\end{aligned} \tag{20.3.1}$$

and

$$\begin{aligned} g(t) &= K_0t^4+\frac{K}{2\theta^2}[1+|1-\theta|(1+|1-\theta^2|)]t^3 \\ &+\frac{K}{2\theta^2}[|1-\theta|(1+|1-\theta^2|)-1]t^2 \\ &+\frac{K}{\theta^2}|1-\theta|(1+|1-\theta^2|)(\frac{|\theta^2-1|}{2\theta^2}-1)t \\ &-\frac{K}{2\theta^4}|1-\theta||1-\theta^2|(1+|1-\theta^2|). \end{aligned} \tag{20.3.2}$$

We have $f(0)=-\frac{K}{2}\frac{|\theta^2-1|}{|\theta|}<0$ for $\theta\neq\pm1$ and $f(1)=K_0>0$ for $K_0\neq0$. It follows from the intermediate value theorem that polynomial f has roots in $(0,1)$. Denote by δ_f the smallest root of f in $(0,1)$. Similarly, we have $g(0)=-\frac{K}{2\theta^4}|1-\theta||\theta^2-1|(1+|1-\theta^2|)<0$ for $\theta\neq\pm1$ and $g(1)=K_0+\frac{K}{2\theta^2}>0$. Denote by δ_g the smallest root of g in $(0,1)$. Set

$$\delta=\min\{\delta_f,\delta_g\}. \tag{20.3.3}$$

Moreover, suppose that δ satisfies

$$\left|\frac{1-\theta}{\theta^3}\right|(1+|1-\theta^2|)+\frac{K\eta}{2\theta}\leq\delta, \tag{20.3.4}$$

$$0<\frac{K|\theta|}{1-K_0(1+\delta)s_0}\left[\frac{|\theta^2-1|}{2\theta^2}+\frac{\delta^2}{2}+\delta\right](s_0-t_0)\leq\delta \tag{20.3.5}$$

and

$$\begin{aligned} 0 &< \frac{K}{\theta^2(1-K_0(1+\delta)s_0)}\{|1-\theta|(1+|1-\theta^2|) \\ &\left[\frac{|\theta^2-1|}{2\theta^2}+\frac{\delta^2}{2}+\delta\right]+\frac{\delta^2}{2}\}(s_0-t_0)\leq\delta^2. \end{aligned} \tag{20.3.6}$$

We shall assume from now on that δ satisfies conditions (20.3.3)-(20.3.6). These conditions shall be referred to as the $(\triangle)$ conditions. Moreover, define scalar sequences $\{t_n\}$, $\{s_n\}$ by

$$\begin{aligned} t_0=0, \qquad & s_0=t_0+\theta\eta, \\ t_1 &= s_0+\left[\frac{|1-\theta|}{|\theta^3|}(1+|1-\theta^2|)+\frac{(s_0-t_0)K}{2\theta^2}\right](s_0-t_0) \end{aligned}$$

for each $n=0,1,2,\cdots$.

$$\begin{aligned} s_{n+1} &= t_{n+1}+\frac{K|\theta|}{1-K_0t_{n+1}} \\ &\left[\frac{|1-\theta^2|}{2\theta^2}(s_n-t_n)^2+\frac{(t_{n+1}-s_n)^2}{2}+(s_n-t_n)(t_{n+1}-s_n)\right] \end{aligned} \tag{20.3.7}$$

$$\begin{aligned} t_{n+2} &= s_{n+1} + \frac{K}{\theta^2(1-K_0 t_{n+1})}\{|1-\theta|(1+|1-\theta^2|) \\ &\left[\frac{|1-\theta^2|}{2\theta^2}(s_n-t_n)^2 + \frac{(t_{n+1}-s_n)^2}{2} + (s_n-t_n)(t_{n+1}-s_n)\right] \\ &+\frac{1}{2}(s_{n+1}-t_{n+1})^2\}. \end{aligned} \tag{20.3.8}$$

Then, we can show the following auxiliary result for majorizing sequences $\{t_n\}$, $\{s_n\}$ under the ($\triangle$) conditions.

Lemma 20.3.1. *Suppose that the ($\triangle$) conditions hold. Then, sequence $\{t_n\}$, $\{s_n\}$ defined by (20.3.7) and (20.3.8) are increasingly convergent to their unique least upper bound denoted by t^* which satisfies*

$$\theta\eta \le t^* \le t^{**} := \frac{\theta\eta}{1-\delta}. \tag{20.3.9}$$

Moreover, the following estimates hold for each $n = 0, 1, 2, \cdots$.

$$0 < s_n - t_n \le \delta^n \theta\eta \tag{20.3.10}$$

and

$$0 < t_{n+1} - s_n \le \delta^{n+1}\theta\eta. \tag{20.3.11}$$

Proof. We shall show estimates (20.3.10) and (20.3.11) using induction. If $n = 0$, (20.3.10) holds by the definition of t_0 and s_0, whereas (20.3.11) holds by (20.3.4). We then have that

$$t_1 \le s_0 + \delta s_0 = (1+\delta)s_0 = \frac{1-\delta^2}{1-\delta}s_0 < t^{**}. \tag{20.3.12}$$

If $n = 1$, estimates (20.3.10) and (20.3.11) hold by (20.3.5), (20.3.6), (20.3.12) and (20.3.10), (20.3.11) for $n = 0$. Suppose that (20.3.10) and (20.3.11) hold for all $m \le n$. Then, we have that

$$\begin{aligned} t_{m+1} &\le s_m + \delta^{m+1}(s_0-t_0) \le t_m + \delta^m(s_0-t_0) \\ &\delta^{m+1}(s_0-t_0) \le \cdots \le t_0 + (s_0-t_0) + \delta(s_0-t_0) \\ &+\cdots+\delta^{m+1}(s_0-t_0) = \frac{1-\delta^{m+2}}{1-\delta}(s_0-t_0) < t^{**}. \end{aligned} \tag{20.3.13}$$

Next, we shall show (20.3.10) for $m+1$ replacing n. We have by the induction hypotheses and (20.3.13) that

$$\begin{aligned} s_{m+1} - t_{m+1} &\le \frac{K|\theta|}{1-K_0\frac{1-\delta^{m+2}}{1-\delta}} \\ &\left[\frac{|\theta^2-1|}{\theta^2}(\delta^m(s_0-t_0))^2 + \frac{(\delta^m(s_0-t_0))^2}{2} + \delta^{2m+1}(s_0-t_0)^2\right] \end{aligned}$$

must be smaller or equal to $\delta^{m+1}(s_0-t_0)$, or

$$\frac{K|\theta|}{1-K_0\frac{1-\delta^{m+2}}{1-\delta}}\left[\frac{|\theta^2-1|}{\theta^2}\delta^m + \frac{\delta^{m+2}}{2} + \delta^{m+1}\right](s_0-t_0) \le \delta. \tag{20.3.14}$$

Estimate (20.3.14) motivates us to define recurrent polynomials f_m on $(0,1)$ by

$$\begin{aligned} f_m(t) &= K\left[\frac{|\theta|}{2}t^{m+2}+|\theta|t^{m+1}+\frac{|\theta^2-1|}{2|\theta|}t^m\right](s_0-t_0) \\ &\quad +K_0t(1+t+\cdots+t^{m+1})(s_0-t_0)-t. \end{aligned} \tag{20.3.15}$$

We need a relationship between two consecutive polynomials f_m. Using (20.3.15) and (20.3.1) by direct algebraic manipulation we get that

$$f_{m+1}(t) = f_m(t)+f(t)t^{m-1}(s_0-t_0). \tag{20.3.16}$$

Evidently, condition (20.3.14) is satisfied, if

$$f_m(\delta) \leq 0. \tag{20.3.17}$$

We also have from (20.3.17) that

$$f_{m+1}(\delta) \leq f_m(\delta), \tag{20.3.18}$$

since $f(\delta) \leq 0$. It then, follows from (20.3.17) and (20.3.18) that (20.3.17) holds, if

$$f_0(\delta) \leq 0, \tag{20.3.19}$$

which is true by (20.3.5). Hence, we showed (20.3.10) for $m+1$ replacing n. Next, we shall show (20.3.11) for $m+1$ replacing n. We have in turn that

$$\begin{aligned} s_{m+2}-s_{m+1} &\leq \frac{K}{\theta^2(1-K_0\frac{1-\delta^{m+2}}{1-\delta})}\{|1-\theta|(1+|\theta^2-1|) \\ &\quad \left[\frac{|\theta^2-1|}{2\theta^2}(\delta^m(s_0-t_0))^2+\frac{(\delta^{m+1}(s_0-t_0))^2}{2}+\delta^{2m+1}(s_0-t_0)^2\right] \\ &\quad +(\delta^{m+1}(s_0-t_0))^2\} \end{aligned}$$

must be smaller or equal to $\delta^{m+2}(s_0-t_0)$. As in the preceding case we are motivated to define polynomials g_m on $[0,1]$ by

$$\begin{aligned} g_m(t) &= K\{\frac{|1-\theta|(1+|\theta^2-1|)}{\theta^2}\left[\frac{|\theta^2-1|}{\theta^2}t^m+\frac{t^{m+2}}{2}+t^{m+1}\right]+\frac{t^{m+2}}{2\theta^2}\} \\ &\quad \times(s_0-t_0)+t^2K_0(1+t+\cdots+t^{m+1})(s_0-t_0)-t^2. \end{aligned} \tag{20.3.20}$$

Using (20.3.20) and (20.3.2) by direct algebraic manipulation we get that

$$g_{m+1}(t) = g_m(t)+g(t)t^m(s_0-t_0). \tag{20.3.21}$$

Condition (20.3.11) is satisfied, if

$$g_m(\delta) \leq 0. \tag{20.3.22}$$

We also have from (20.3.21) and ($\triangle$) that

$$g_{m+1}(\delta) \leq g_m(\delta), \tag{20.3.23}$$

since $g(\delta) \leq 0$. Hence, (20.3.22) is satisfied, if

$$g_0(\delta) \leq 0, \tag{20.3.24}$$

which is true by (20.3.6). The induction for (20.3.11) is completed. It then, follows that

$$t_{m+2} \leq \frac{1-\delta^{m+3}}{1-\delta} s_0 < t^{**}. \tag{20.3.25}$$

Hence, sequences $\{t_n\}$, $\{s_n\}$ are increasing, bounded above by t^{**} and as such they converge to their unique least upper bound t^* which satisfies (20.3.9). □

We can show the main semilocal convergence result for the third order method (20.1.2) under the $(\mathcal{C}^0)$ and $(\triangle)$ conditions using $\{t_n\}$ and $\{s_n\}$ as majorizing sequences.

Theorem 20.3.2. *Suppose that*

$$\overline{U}(x_0, t^*) \subset D, \tag{20.3.26}$$

the $(\mathcal{C}^0)$ and $(\triangle)$ conditions hold. Then, sequences $\{x_n\}$, $\{y_n\}$ generated by the third order method (20.1.2) are well defined, remain in $\overline{U}(x_0,t^*)$ for each $n = 0,1,2,\cdots$ and converge to a unique solution x^* of equation $F(x) = 0$ in $\overline{U}(x_0,t^*) \cap D$. Moreover the following estimates hold for each $n = 0,1,2,\cdots$.

$$\begin{aligned} \|y_n - x_n\| &\leq s_n - t_n, & (20.3.27)\\ \|x_{n+1} - y_n\| &\leq t_{n+1} - s_n & (20.3.28)\\ \|x_{n+1} - x_n\| &\leq t_{n+1} - t_n & (20.3.29)\end{aligned}$$

and

$$\|x_n - x^*\| \leq t^* - t_n. \tag{20.3.30}$$

Furthermore, if there exists $R > t^*$ such that

$$K_0(t^* + R) < 2, \tag{20.3.31}$$

then, the point x^* is the only solution of equation $F(x) = 0$ in $U(x_0,R)$.

Proof. We shall first show (20.3.27) and (20.3.28) using induction. We have by (20.1.2) and (20.3.7) that

$$\|y_0 - x_0\| = |\theta| \|F'(x_0)^{-1} F(x_0)\| \leq |\theta|\eta = s_0 = s_0 - t_0.$$

Hence, (20.3.27) holds for $n = 0$. It follows from the first substep of (20.1.2) that

$$\begin{aligned} F(y_0) &= F(y_0) - \theta F(x_0) - F'(x_0)(y_0 - x_0)\\ &= (1-\theta)F(x_0)\\ &\quad + \int_0^1 [F'(x_0 + t(y_0 - x_0)) - F'(x_0)](y_0 - x_0)dt. \qquad (20.3.32)\end{aligned}$$

Composing (20.3.32) by $F'(x_0)^{-1}$ and using (2), (3) in $(\mathcal{C}^0)$ and (20.3.7)

$$\begin{aligned}\|F'(x_0)^{-1}F(y_0)\| &\leq |1-\theta|\|\|F'(x_0)^{-1}F(x_0)\| \\ &\quad +\|\int_0^1 [F'(x_0+t(y_0-x_0))-F'(x_0)](y_0-x_0)dt \\ &\leq \frac{|1-\theta|}{|\theta|}(s_0-t_0)+\frac{K_0}{2}\|y_0-x_0\|^2 \\ &\leq (\frac{|1-\theta|}{|\theta|}+\frac{K_0}{2}(s_0-t_0))(s_0-t_0). \qquad (20.3.33)\end{aligned}$$

Subtracting the first from the second substep in (20.1.2) we get that

$$x_1-y_0=-\frac{(\theta+1)(\theta-1)^2}{\theta^2}F'(x_0)^{-1}F(x_0)-\frac{1}{\theta^2}F'(x_0)^{-1}F(y_0) \qquad (20.3.34)$$

Hence, using (20.3.33) and (20.3.34), we get that

$$\begin{aligned}\|x_1-y_0\| &= \frac{|\theta+1||\theta-1|^2}{\theta^2}\|F'(x_0)^{-1}F(x_0)\|+\frac{1}{\theta^2}\|F'(x_0)^{-1}F(y_0)\| \\ &\leq \frac{|\theta+1||\theta-1|^2}{\theta^2}(s_0-t_0)+\frac{1}{\theta^2}(\frac{|1-\theta|}{|\theta|}+\frac{K}{2}(s_0-t_0))(s_0-t_0) \\ &= t_1-s_0, \qquad (20.3.35)\end{aligned}$$

which shows (20.3.28) for $n=0$. Then, (20.3.29) holds for $n=0$, since

$$\|x_1-x_0\| \leq \|x_1-y_0\|+\|y_0-x_0\| \leq t_1-s_0+s_0-t_0=t_1-t_0 \leq t^*.$$

Then, we have $x_1 \in \overline{U}(x_0,t^*)$. Notice that $K_0t^*<1$ from the proof of Lemma 20.3.1. Let us suppose $x \in \overline{U}(x_0,t^*)$. Then, using (2) in $(\mathcal{C}^0)$ we have that

$$\|F'(x_0)^{-1}(F'(x)-F'(x_0))\| \leq K_0\|x-x_0\| \leq K_0t^*<1. \qquad (20.3.36)$$

It follows from (20.3.36) and the Banach lemma that $F'(x)^{-1} \in L(Y,X)$ and

$$\|F'(x_1)^{-1}F'(x_0)\| \leq \frac{1}{1-K_0\|x_1-x_0\|} \leq \frac{1}{1-K_0t_1}. \qquad (20.3.37)$$

Suppose that (20.3.27)-(20.3.29) hold for all $m \leq n$ and $x_m \in \overline{U}(x_0,t^*)$. Using the first step in (20.1.2) we get that

$$\begin{aligned}F(y_m) &= F(y_m)-\theta F(x_m)-F'(x_m)(y_m-x_m) \\ &= (1-\theta)F(x_m) \\ &\quad +\int_0^1 [F'(x_m+t(y_m-x_m))-F'(x_m)](y_m-x_m)dt. \qquad (20.3.38)\end{aligned}$$

Subtracting the first step in (20.1.2) from the second step to obtain

$$F'(x_m)(x_{m+1}-y_m)=\frac{\theta^3-\theta^2-\theta+1}{\theta^2}F(x_m)-\frac{1}{\theta^2}F(y_m). \qquad (20.3.39)$$

We also have by (20.3.38) that

$$\begin{aligned}
F(x_{m+1}) &= F'(x_m)(x_{m+1}-y_m)+F(y_m)+[F'(y_m)-F'(x_m)](x_{m+1}-y_m) \\
&\quad +F(x_{m+1})-F(y_m)-F'(y_m)(x_{m+1}-y_m) \\
&= \frac{1-\theta}{\theta^2}F(x_m)-\frac{1}{\theta^2}F(y_m) \\
&\quad +\int_0^1 [F'(x_m+t(y_m-x_m))-F'(x_m)](y_m-x_m)dt \\
&\quad +\int_0^1 [F'(y_m+t(x_{m+1}-y_m))-F'(y_m)](x_{m+1}-y_m)dt \\
&\quad +[F'(y_m)-F'(x_m)](x_{m+1}-y_m). \qquad (20.3.40)
\end{aligned}$$

Hence, we get by (20.3.40) that

$$\begin{aligned}
\|F'(x_0)^{-1}F(x_{m+1})\| &\leq K\Bigg[\frac{|\theta^2-1|}{2\theta^2}\|y_m-x_m\|^2 \\
&\quad +\frac{\|x_{m+1}-y_m\|^2}{2}+\|y_m-x_m\|\|x_{m+1}-y_m\|\Bigg] \\
&\leq K\Bigg[\frac{|\theta^2-1|}{2\theta^2}(s_m-t_m)^2 \\
&\quad +\frac{(t_{m+1}-s_m)^2}{2}+(s_m-t_m)(t_{m+1}-s_m)\Bigg]. \qquad (20.3.41)
\end{aligned}$$

Then, we get that

$$\begin{aligned}
\|y_{m+1}-x_{m+1}\| &\leq \|F'(x_{m+1})^{-1}F'(x_0)\|\|F'(x_0)^{-1}F(x_{m+1})\| \\
&\leq \frac{K}{1-K_0t_{m+1}}\Bigg[\frac{|\theta^2-1|}{2\theta^2}(s_m-t_m)^2 \\
&\quad +\frac{(t_{m+1}-s_m)^2}{2}+(s_m-t_m)(t_{m+1}-s_m)\Bigg]=s_{m+1}-t_{m+1},
\end{aligned}$$

where, we used (20.3.37) for $x=x_{m+1}$ and

$$\|x_{m+1}-x_0\|\leq\|x_{m+1}-x_m\|+\cdots+\|x_1-x_0\|\leq t_{m+1}-t_m+\cdots+t_1-t_0=t_{m+1}.$$

Hence, we showed (20.3.27). Then, we have by (20.3.39) that

$$x_{m+1}-y_m=\frac{\theta^3-\theta^2-\theta+1}{\theta^2}F'(x_m)^{-1}F(x_m)-\frac{1}{\theta^2}F'(x_m)^{-1}F(y_m). \qquad (20.3.42)$$

It follows from (20.3.42) that

$$\begin{aligned}
\|x_{m+2}-y_{m+1}\| &\leq \frac{K}{1-K_0t_{m+1}}[\frac{|1+\theta|(\theta-1)^2}{\theta^2}\|F'(x_0)^{-1}F(x_{m+1})\| \\
&\quad +\frac{1}{\theta^2}\|F'(x_0)^{-1}F(y_{m+1})\|] \\
&\leq \frac{K}{\theta^2(1-K_0t_{m+1})}\Big[|1+\theta|(\theta-1)^2(\frac{|\theta^2-1|}{2\theta^2}(s_m-t_m)^2 \\
&\quad +\frac{(t_{m+1}-s_m)^2}{2}+(s_m-t_m)(t_{m+1}-s_m)) \\
&\quad +|1-\theta|(\frac{|\theta^2-1|}{2\theta^2}(s_m-t_m)^2 \\
&\quad +\frac{(t_{m+1}-s_m)^2}{2}+(s_m-t_m)(t_{m+1}-s_m))) \\
&\quad +\frac{(s_{m+1}-t_{m+1})^2}{2}\Big] \\
&= t_{m+2}-s_{m+1}.
\end{aligned}$$

Hence, we showed (20.3.28). Then, we have that

$$\begin{aligned}
\|x_{m+2}-x_{m+1}\| &\leq \|x_{m+2}-y_{m+1}\|+\|y_{m+1}-x_{m+1}\| \\
&\leq t_{m+2}-s_{m+1}+s_{m+1}-t_{m+1} \\
&= t_{m+2}-t_{m+1},
\end{aligned}$$

which shows (20.3.29). We also have that

$$\begin{aligned}
\|x_{m+2}-x_0\| &\leq \|x_{m+2}-x_{m+1}\|+\|x_{m+1}-x_m\|+\cdots+\|x_1-x_0\| \\
&\leq t_{m+2}-t_{m+1}+t_{m+1}-t_m+\cdots+t_1-t_0 \\
&= t_{m+2}<t^*.
\end{aligned}$$

Hence, we get $x_{m+2}\in\overline{U}(x_0,t^*)$.

We showed in Lemma 20.3.1 that sequences $\{t_n\}$, $\{s_n\}$ are complete. Hence, it follows from (20.3.27)-(20.3.29) that sequences $\{x_n\}$, $\{y_n\}$ are complete in a Banach space X and as such they converge to some $x^*\in\overline{U}(x_0,t^*)$ (since $\overline{U}(x_0,t^*)$ is a closed set.) By letting $m\to\infty$ in (20.3.41), we obtain $F(x^*)=0$. Estimate (20.3.30) follows from (20.3.29) by using standard majorization techniques [1, 3, 11, 15, 18, 19]. Let us show uniqueness, first in $\overline{U}(x_0,t^*)\cap D$. Let $y^*\in\overline{U}(x_0,t^*)$ be such that $F(y^*)=0$. Set $Q=\int_0^1F'(x^*+t(y^*-x^*))dt$. Then, using (2) in $(\mathcal{C}^0)$ we get that

$$\begin{aligned}
\|F'(x_0)^{-1}(F'(x_0)-Q)\| &\leq K_0\int_0^1[(1-t)\|x^*+t(y^*-x^*)-x_0\|dt \\
&\leq K_0\int_0^1[(1-t)\|x^*-x_0\|+t\|y^*-x^*)-x_0\|dt \\
&\leq K_0t^*<1.
\end{aligned}$$

It follows that Q^{-1} exists. Then, from the identity $0 = F(x^*) - F(y^*) = Q(x^* - y^*)$ we deduce that $x^* = y^*$. Similarly, if $F(y^*) = 0$ and $y^* \in U(x_0, R)$, we have that

$$\|F'(x_0)^{-1}(F'(x_0) - Q)\| \le \frac{K_0}{2}(R + t^*) < 1,$$

by (20.3.31). Hence, again we deduce that $x^* = y^*$. □

Remark 20.3.3. *(a) It follows from the proof of Theorem 20.3.2 that sequences $\{\bar{t}_n\}, \{\bar{s}_n\}$ defined by*

$$\begin{aligned}
\bar{t}_0 = 0, \bar{s}_0 &= \bar{t}_0 + \theta\eta, \\
\bar{t}_1 &= \bar{s}_0 + [\frac{|1-\theta}{|\theta^3|}(1 + |1-\theta^2|) + \frac{(\bar{s}_0 - \bar{t}_0)K_0}{2\theta^2}](\bar{s}_0 - \bar{t}_0), \\
\bar{s}_1 &= \bar{t}_1 + \frac{|\theta|}{1 - K_0\bar{t}_1}[\frac{K}{2}\frac{|\theta^2 - 1|}{\theta^2}(\bar{s}_0 - \bar{t}_0)^2 \\
&\quad \frac{K}{2}(\bar{t}_1 - \bar{s}_0)^2 + K_0(\bar{s}_0 - \bar{t}_0)(\bar{t}_1 - \bar{s}_0)], \\
\bar{s}_{n+1} &= \bar{t}_{n+1} + \frac{|\theta|}{1 - K_0\bar{t}_{n+1}}[\frac{|\theta^2 - 1|}{2\theta^2}(\bar{s}_n - \bar{t}_n)^2 \\
&\quad \frac{(\bar{t}_{n+1} - \bar{s}_n)^2}{2} + (\bar{s}_n - \bar{t}_n)(\bar{t}_{n+1} - \bar{s}_n)], \\
\bar{t}_{n+2} &= \bar{s}_{n+1} + \frac{K}{\theta^2(1 - K_0\bar{t}_{n+1})}\{|1-\theta|(1 + |1-\theta^2|)[\frac{|\theta^2 - 1|}{2\theta^2}(\bar{s}_n - \bar{t}_n)^2 \\
&\quad \frac{(\bar{t}_{n+1} - \bar{s}_n)^2}{2} + (\bar{s}_n - \bar{t}_n)(\bar{t}_{n+1} - \bar{s}_n)] \\
&\quad \frac{1}{2}(\bar{s}_{n+1} - \bar{t}_{n+1})^2\} \text{ for each } n = 0, 1, 2, \cdots.
\end{aligned}$$

Then, a simple induction argument shows that

$$\begin{aligned}
\bar{s}_n &\le s_n, \\
\bar{t}_n &\le t_n, \\
\bar{s}_n - \bar{t}_n &\le s_n - t_n, \\
\bar{t}_{n+1} - \bar{s}_n &\le t_{n+1} - s_n
\end{aligned}$$

and

$$\bar{t}^* = \lim_{n\to\infty} \bar{t}_n \le t^*.$$

Clearly, $\{\bar{t}_n\}$, $\{\bar{s}_n\}$, $\bar{t}^*$ can replace $\{t_n\}$, $\{s_n\}$, t^* in Theorem 20.3.2.

(b) The limit point t^* can be replaced by t^{**} given in closed form by (20.3.9).

(c) Criteria (△) or (20.2.9) and (20.2.11) are sufficient for the convergence of the third order method (20.1.2). However, these criteria are not also necessary. In practice, we shall test to see which of these criteria are satisfied (if any) and then use the best possible error bounds and uniqueness results (see also the numerical examples in the next section).

20.4. Numerical Examples

Example 20.4.1. *Let $x \in D, X = Y = \mathbb{R}$, $x_0 = 1$ and $D = \overline{U}(1,1)$. Define function F on D by*

$$F(x) = x^3 - 0.49. \tag{20.4.1}$$

Then, we get that

$$\beta = \frac{1}{3} \qquad \eta = 0.17, \qquad M = 12.$$

Now choosing $\theta = 1.15$ we obtain that

$$a = 0.68, \qquad \alpha = 0.68, \qquad \gamma = 0.34$$

and as a consequence $a_1 b_1 = 134.091 \leq 1$ condition (20.2.9) is violated. Hence, there is no guarantee under the conditions given in [4] that sequence $\{x_n\}$ converges to x^*. Calculating now δ_f and δ_g, the smallest solutions of the polynomials $f(t)$ and $g(t)$ given in (20.3.1) and (20.3.2) respectively between 0 and 1, we obtain that

$$\delta = \min\{\delta_f, \delta_g\} = .4104586\ldots$$

Moreover, we observe that the Δ conditions are satisfied since

$$\left|\frac{1-\theta}{\theta^3}\right|(1+|1-\theta^2|) + \frac{K\eta}{2\theta} = .278261\ldots \leq \delta,$$

$$0 < \frac{K|\theta|}{1-K_0(1+\delta)s_0}\left[\frac{|\theta^2-1|}{2\theta^2} + \frac{\delta^2}{2} + \delta\right](s_0 - t_0) = .360324\ldots \leq \delta$$

and

$$0 < \frac{K}{\theta^2(1-K_0(1+\delta)s_0)}\{|1-\theta|(1+|1-\theta^2|)\left[\frac{|\theta^2-1|}{2\theta^2} + \frac{\delta^2}{2} + \delta\right] + \frac{\delta^2}{2}\}(s_0 - t_0)$$

$$= .136162\ldots \leq .168476\ldots = \delta^2.$$

Consequently, convergence to the solution is guaranteed by Theorem 20.3.2. Moreover, the computational order of convergence (COC) is shown in Table 20.4.1. Here (COC) is defined by

$$\rho \approx \ln\left(\frac{\|\overline{x}_{n+1} - \overline{x}^{\star}\|_\infty}{\|\overline{x}_n - \overline{x}^{\star}\|_\infty}\right) / \ln\left(\frac{\|\overline{x}_n - \overline{x}^{\star}\|_\infty}{\|\overline{x}_{n-1} - \overline{x}^{\star}\|_\infty}\right), \qquad n \in \mathbb{N},$$

The Table 20.4.1 shows the (COC).

Example 20.4.2. *Let $X = Y = \mathcal{C}[0,1]$, the space of continuous functions defined in $[0,1]$ equipped with the max-norm. Let $\Omega = \{x \in \mathcal{C}[0,1]; \|x\| \leq R\}$, such that $R > 1$ and F defined on Ω and given by*

$$F(x)(s) = x(s) - f(s) - \lambda \int_0^1 G(s,t)x(t)^3\,dt, \quad x \in C[0,1],\ s \in [0,1],$$

Table 20.4.1. COC for Example 1 using $\theta = 1.15$

n	*COC*
1	2.73851
2	2.99157
3	2.99999
4	3.00000
5	3.00000

$\rho = 3.00000$

where $f \in \mathcal{C}[0,1]$ is a given function, λ is a real constant and the kernel G is the Green function

$$G(s,t) = \begin{cases} (1-s)t, & t \le s, \\ s(1-t), & s \le t. \end{cases}$$

In this case, for each $x \in \Omega$, $F'(x)$ is a linear operator defined on Ω by the following expression:

$$[F'(x)(v)](s) = v(s) - 3\lambda \int_0^1 G(s,t)x(t)^2 v(t)\, dt, \quad v \in C[0,1],\ s \in [0,1].$$

If we choose $x_0(s) = f(s) = 1$, it follows $\|I - F'(x_0)\| \le 3|\lambda|/8$. Thus, if $|\lambda| < 8/3$, $F'(x_0)^{-1}$ is defined and

$$\|F'(x_0)^{-1}\| \le \frac{8}{8 - 3|\lambda|}.$$

Moreover,

$$\|F(x_0)\| \le \frac{|\lambda|}{8},$$

$$\|F'(x_0)^{-1}F(x_0)\| \le \frac{|\lambda|}{8 - 3|\lambda|}.$$

On the other hand, for $x, y \in \Omega$ we have

$$[(F'(x) - F'(y))v](s) = 3\lambda \int_0^1 G(s,t)(x(t)^2 - y^2(t))v(t)\, dt$$

and for $x \in \Omega$ we get in turn that

$$\|F''(x)\| \le \frac{6|\lambda|}{8}.$$

Consequently,

$$\|F'(x) - F'(y)\| \le \|x - y\| \frac{3|\lambda|(\|x\| + \|y\|)}{8} \le \|x - y\| \frac{6R|\lambda|}{8},$$

$$\|F'(x) - F'(1)\| \le \|x - 1\| \frac{1 + 3|\lambda|(\|x\| + 1)}{8} \le \|x - 1\| \frac{1 + 3(1 + R)|\lambda|}{8}.$$

Choosing $\lambda = 1.5$*,* $R = 4.4$ *and* $\theta = 1.1$ *we have*

$$\beta = 0.677966\ldots,$$

$$\eta = 0.127119\ldots,$$

$$M = 4.95,$$

$$a = 0.426602\ldots,$$

$$\alpha = 1.16529\ldots,$$

and

$$\gamma = 0.213301\ldots$$

So, as $a_1 b_1 = 1.25402 \leq 1$*, condition (20.2.9) is violated. Hence, there is no guarantee under the conditions given in [4] that sequence* $\{x_n\}$ *converges to* x^**. Calculating now* δ_f *and* δ_g*, the smallest solutions of the polynomials* $f(t)$ *and* $g(t)$ *given in (20.3.1) and (20.3.2) respectively between 0 and 1, we obtain that*

$$\delta = \min\{\delta_f, \delta_g\} = 0.370693\ldots$$

Moreover, we observe that the Δ *conditions are satisfied since*

$$\left|\frac{1-\theta}{\theta^3}\right|(1+|1-\theta^2|) + \frac{K\eta}{2\theta} = 0.284819\ldots \leq \delta,$$

$$0 < \frac{K|\theta|}{1-K_0(1+\delta)s_0}\left[\frac{|\theta^2-1|}{2\theta^2} + \frac{\delta^2}{2} + \delta\right](s_0 - t_0) = 0.334767\ldots \leq \delta$$

and

$$0 < \frac{K}{\theta^2(1-K_0(1+\delta)s_0)}\{|1-\theta|(1+|1-\theta^2|)\left[\frac{|\theta^2-1|}{2\theta^2} + \frac{\delta^2}{2} + \delta\right] + \frac{\delta^2}{2}\}(s_0 - t_0)$$

$$= 0.0871515\ldots \leq 0.137413\ldots = \delta^2.$$

Consequently, convergence to the solution is guaranteed by Theorem 20.3.2.

References

[1] Argyros, I.K., Computational theory of iterative methods, *Series: Studies in Computational Mathematics*, 15, Editors: C.K. Chui and L. Wuytack, Elsevier Publ. Co. New York, U.S.A, 2007.

[2] Argyros, I.K., Hilout, S., Weaker conditions for the convergence of Newton's method. *J. Complexity*, 28 (2012), 364–387.

[3] Argyros, I.K., Hilout, S., Computational methods in nonlinear analysis, World Scientific Publ. Comp., New Jersey, USA 2013.

[4] Chun, C., Stanica, P., Neta, B., Third order family of methods in Banach spaces, *Comp Math. Appl.*, 61 (2011), 1665-1675.

[5] Candela, V., Marquina, A., Recurrence relations for rational cubic methods I: The Halley method, *Computing*, 44 (1990), 169-184.

[6] Candela, V., Marquina, A., Recurrence relations for rational cubic methods II: The Chebyshev method, *Computing*, 45 (1990), 355-367.

[7] Ezquerro, J. A., Hernández, M.A., Recurrence relations for Chebyshev-type methods, *Appl. Math. Optim.*, 41 (2000), 227-236.

[8] Gutiérrez, J.M., Hernández, M.A., Recurrence relations for the super-Halley method, *Computers Math. Applic.*, 36 (1998), 1-8.

[9] Gutiérrez, J.M., Hernández, M.A., Third-order iterative methods for operators with bounded second derivative, *J. Comp. Appl. Math.*, 82 (1997), 171-183.

[10] Gutiérrez, J.M., Magreñán, Á.A., Romero, N., On the semilocal convergence of Newton-Kantorovich method under center-Lipschitz conditions, *App. Math. Comp.*, 221 (2013), 79-88.

[11] Kantorovich, L.V., G.P. Akilov, *Functional Analysis*, Pergamon Press, Oxford, 1982.

[12] Kou, J.S., Li, T., Modified Chebyshev's method free from second derivative for nonlinear equations, *App. Math. Comp.*, 187 (2007), 1027-1032.

[13] Kou, J.S., Li, T., Wang, X.H., A modification of Newton method with third-order convergence, *App. Math. Comp.*, 181 (2006), 1106-1111.

[14] Magreñán, Á. A., Different anomalies in a Jarratt family of iterative root-finding methods, *App. Math. Comp.*, 233 (2014), 29–38.

[15] Ortega, J.M., Rheinboldt, W.C., *Iterative Solution of Nonlinear Equations in Several Variables*, Academic press, New York, 1970.

[16] Parida, P.K., Study of some third order methods for nonlinear equations in Banach spaces, Ph.D. Dessertation, Indian Institute of Technology, Department of Mathematics, Kharagpur, India, 2007.

[17] Parida, P.K., Gupta, D.K., Recurrence relations for semilocal convergence of a Newton-like method in Banach spaces, *J. Math. Anal. Applic.*, 345 (2008), 350-361.

[18] Potra, F.A., Ptră, V., Nondiscrete induction and iterative processes, in *Research Notes in Mathematics*, 103, Pitman, Boston, 1984.

[19] Rall, L.B., *Computational solution of Nonlinear operator equations*, Robert E. Krieger, New York, 1979.

[20] Wu, Q., Zhao, Y., Third order convergence theorem by using majorizing function for a modified Newton method in Banach space, *App. Math. Comp.*, 175 (2006), 1515-1524.

Chapter 21

Local Convergence of Modified Halley-Like Methods with Less Computation of Inversion

21.1. Introduction

In this chapter we are concerned with the problem of approximating a solution x^* of the nonlinear equation

$$F(x) = 0, \tag{21.1.1}$$

where F is a Fréchet-differentiable operator defined on a subset D of a Banach space X with values in a Banach space Y.

Many problems in computational sciences and other disciplines can be brought in a form like (21.1.1) using mathematical modeling [3]. The solutions of equation (21.1.1) can rarely be found in closed form. That is why most solution methods for these equations are usually iterative. In particular, the practice of Numerical Functional Analysis for finding such solution is essentially connected to Newton-like methods [1]-[20]. The study about convergence matter of iterative procedures is usually based on two types: semilocal and local convergence analyses. The semilocal convergence matter is, based on the information around an initial point, to give conditions ensuring the convergence of the iterative procedure; while the local one is, based on the information around a solution, to find estimates of the radii of convergence balls. There exist many studies which deal with the local and semilocal convergence analyses of Newton-like methods such as [1]-[20].

We present a local convergence analysis for the modified Halley-Like Method [30] defined for each $n = 0, 1, 2, \cdots$ by

$$\begin{aligned} y_n &= x_n - F'(x_n)^{-1}F(x_n), \\ u_n &= x_n - \theta F'(x_n)^{-1}F(x_n), \\ &= y_n + (1-\theta)F'(x_n)^{-1}F(x_n), \\ z_n &= y_n - \gamma A_{\theta,n}F'(x_n)^{-1}F(x_n), \\ x_{n+1} &= z_n - \alpha B_{\theta,n}F'(x_n)^{-1}F(z_n), \end{aligned} \tag{21.1.2}$$

where x_0 is an initial point, $\alpha,\gamma,\theta \in (-\infty,\infty) - \{0\}$ are given parameters, $H_{\theta,n} = \frac{1}{\theta}F'(x_n)^{-1}(F'(u_n) - F'(x_n))$, $A_{\theta,n} = I - \frac{1}{2}H_{\theta,n}(I - \frac{1}{2}H_{\theta,n})$ and $B_{\theta,n} = I - H_{1,n} + H^2_{\theta,n}$. The semilocal convergence of method (21.1.2) was studied in [30] in the special case when $\alpha = \gamma = 1$ and $\theta \in [0,1]$. Moreover, if $\gamma = 1$, $\alpha = 0$ and $\theta \in (0,1]$, the semilocal convergence of the resulting method (21.1.2) was given in [30].

The semilocal convergence results in [30] were given in a non-affine invariant form. However, the results obtained in our chapter are given in affine invariant form. The sufficient semilocal convergence conditions (given in affine invariant form) used in [30] are $(\mathcal{C})$:

$(\mathcal{C}_1)$ There exists $F'(x_0)^{-1} \in L(Y,X)$ and $\|F'(x_0)^{-1}\| \le \beta$;

$(\mathcal{C}_2)$

$$\|F'(x_0)^{-1}F(x_0)\| \le \beta_1;$$

$(\mathcal{C}_3)$

$$\|F'(x_0)^{-1}F''(x)\| \le \beta_2 \quad \text{for each } x \in D;$$

$(\mathcal{C}_4)$

$$\begin{aligned} \|F'(x_0)^{-1}(F''(x) - F''(y))\| &\le \beta_3\|x-y\|^q \\ &\quad \text{for each } x,y \in D, \text{ and some } q \in [0,1]. \end{aligned}$$

Under the $(\mathcal{C})$ conditions for $\alpha = \gamma = 1$ and $\theta \in (0,1]$ the convergence order was shown to be $3+2q$ in [30]. Moreover, for $\gamma = 1$, $\alpha = 0$ and $\theta \in (0,1]$ the convergence order was shown to be $2+q$ in [10].

Similar conditions have been used by several authors on other high convergence order methods [1]-[20]. The corresponding conditions for the local convergence analysis are given by simply replacing x_0 by x^* in the preceding $(\mathcal{C})$ conditions. These conditions however are very restrictive. As a motivational example, let us define function f on $D = [-\frac{1}{2}, \frac{5}{2}]$ by

$$f(x) = \begin{cases} x^3 \ln x^2 + x^5 - x^4, & x \neq 0 \\ 0, & x = 0 \end{cases}$$

Choose $x^* = 1$. We have that

$$\begin{aligned} f'(x) &= 3x^2\ln x^2 + 5x^4 - 4x^3 + 2x^2, \ F'(1) = 3, \\ f''(x) &= 6x\ln x^2 + 20x^3 - 12x^2 + 10x \\ f'''(x) &= 6\ln x^2 + 60x^2 - 24x + 22. \end{aligned}$$

Then, e.g, function f cannot satisfy condition $(\mathcal{C}_4)$, say for $q = 1$, since function f''' is unbounded on D. In the present chapter we only use hypotheses on the first Fréchet derivative (see conditions (21.2.12)-(21.2.15)). Notice that they used $\theta \in (0,1]$, whereas in this chapter θ can belong in a wider than $(0,1]$ interval and $\gamma = \alpha = 1$ in [30]. This way we expand the applicability of method (21.1.2).

The chapter is organized as follows. The local convergence of method (21.1.2) is given in Section 21.2, whereas the numerical examples are given in Section 21.3. Finally, some remarks are given in the concluding Section 21.4.

21.2. Local Convergence Analysis

We present the local convergence analysis of method (21.1.2) in this section. Denote by $U(v,\rho), \bar{U}(v,\rho)$ the open and closed balls, respectively, in X of center $v \in X$ and of radius $\rho > 0$.

Let $L_0 > 0, L > 0, \theta \in (-\infty,\infty) - \{0\}, \alpha, \gamma \in (-\infty,\infty)$ and $M > 0$ be given parameters. Define functions on the interval $[0, \frac{1}{L_0})$ by

$$\begin{aligned}
g_1(r) &= \frac{Lr}{2(1-L_0r)},\\
g_2(r) &= g_1(r) + \frac{M|1-\theta|}{1-L_0r},\\
g_3(r) &= \frac{L_0(1+g_2(r))}{2|\theta|(1-L_0r)},\\
g_4(r) &= 1+g_3(r)r+g_3^2(r)r^2,\\
g_5(r) &= g_1(r)+\frac{|\gamma|Mg_4(r)}{1-L_0r},\\
g_6(r) &= 1+2g_{1,3}(r)r+4g_3^2(r)r^2,\\
g_{1,3}(r) &= \frac{L_0(1+g_1(r))}{2(1-L_0r)}
\end{aligned}$$

and

$$g_7(r) = [1+\frac{|\alpha|Mg_6(r)}{1-L_0r}]g_5(r).$$

Moreover, define parameter

$$r_2 = \frac{2(1-M|1-\theta|)}{2L_0+L}.$$

Suppose

$$M|1-\theta| < 1.$$

Then, it follows from the definition of functions g_1 and g_2 that

$$0 < g_1(r) < 1, \text{ and } 0 < g_2(r) < 1, \text{ for each } r \in (0,r_2).$$

Evidently, $g_5(r) \in (0,1)$, if for each $r \in (0,r_5)$ and $r_5 < \frac{1}{L_0}$ to be determined, we have that

$$0 < g_1(r) + \frac{|\gamma|g_4(r)M}{1-L_0r} < 1 \text{ for each } r \in (0,r_5).$$

Define function p_5 on the interval $[0, \frac{1}{L_0}]$ by

$$p_5(r) = |\gamma|Mg_4(r) - (1-L_0r)(1-g_1(r)).$$

We have that

$$p_5((\frac{1}{L_0})^-) = |\gamma|Mg_4((\frac{1}{L_0})^-) > 0.$$

Suppose that
$$|\gamma|M < 1.$$

Then, we have that
$$p_5(0) = M|\gamma| - 1 < 0.$$

It follows from the intermediate value theorem that function p_5 has zeros in the interval $(0, \frac{1}{L_0})$. Denote by r_5 the smallest such zero. Then, we have that

$$p_5(r) < 0 \Rightarrow 0 < g_5(r) < 1 \text{ for each } r \in (0, r_5).$$

Similarly, function $g_7 \in (0,1)$ for each $r \in (0, r_7)$ and $r_7 < \frac{1}{L_0}$ to be determined, if function $p_7(r) \in (0,1)$ for each $r \in [0, r_7]$, where

$$p_7(r) = (1 - L_0 r + |\alpha| M g_6(r)) g_5(r) - (1 - L_0 r).$$

We get that
$$p_7((\frac{1}{L_0})^-) = |\gamma| M g_6((\frac{1}{L_0})^-) g_5((\frac{1}{L_0})^-) > 0.$$

and
$$p_7(0) = (1 + |\alpha| M g_6(0)) |\gamma| g_5(0) - 1 = (1 + |\alpha| M) |\gamma| M - 1.$$

Suppose that
$$(1 + |\alpha| M) |\gamma| M < 1.$$

Then, we have $p_7(0) < 0$. It follows that function p_7 has zeros in the interval $(0, \frac{1}{L_0})$. Denote by r_7 the smallest such zero. Then, we obtain that

$$p_7(0) < 0 \Rightarrow 0 < g_7(r) < 1, \text{ for each } r \in (0, r_7).$$

Set
$$r^* = \min\{r_2, r_5, r_7\}. \tag{21.2.1}$$

Then, we have that

$$0 < g_1(r) < 1, \tag{21.2.2}$$
$$0 < g_2(r) < 1 \tag{21.2.3}$$
$$0 < g_3(r) \tag{21.2.4}$$
$$0 < g_4(r) \tag{21.2.5}$$
$$0 < g_5(r) < 1 \tag{21.2.6}$$
$$0 < g_6(r) \tag{21.2.7}$$

and
$$0 < g_7(r) < 1, \text{ for each } r \in (0, r^*). \tag{21.2.8}$$

Next, we present the local convergence analysis of method (21.1.2).

Theorem 21.2.1. *Let $F : D \subseteq X \to Y$ be a Fréchet-differentiable operator. Suppose that there exist $x^* \in D$, parameters $L_0 > 0, L > 0, M > 0, \theta \in (-\infty,\infty) - \{0\}$ and $\alpha, \gamma \in (-\infty,\infty)$ such that for each $x \in D$*

$$M|1-\theta| < 1, \tag{21.2.9}$$

$$M|\gamma| < 1, \tag{21.2.10}$$

$$(1+|\alpha|M)|\gamma|M < 1, \tag{21.2.11}$$

$$F(x^*) = 0,\ F'(x^*)^{-1} \in L(Y,X), \tag{21.2.12}$$

$$\|F'(x^*)^{-1}(F'(x) - F'(x^*))\| \leq L_0\|x - x^*\|, \tag{21.2.13}$$

$$\|F'(x^*)^{-1}(F(x) - F(x^*) - F'(x)(x - x^*)\| \leq \frac{L}{2}\|x - x^*\|^2, \tag{21.2.14}$$

$$\|F'(x^*)^{-1}F'(x)\| \leq M \tag{21.2.15}$$

and

$$\bar{U}(x^*, r^*) \subseteq D, \tag{21.2.16}$$

where r^* is given in (21.2.1). Then, the sequence $\{x_n\}$ generated by method (21.1.2) for $x_0 \in U(x^*, r^*)$ is well defined, remains in $U(x^*, r^*)$ for each $n = 0, 1, 2, \cdots$ and converges to x^*. Moreover, the following estimates hold for each $n = 0, 1, 2, \cdots$,

$$\|y_n - x^*\| \leq g_1(\|x_n - x^*\|)\|x_n - x^*\| < \|x_n - x^*\| < r^*, \tag{21.2.17}$$

$$\|u_n - x^*\| \leq g_2(\|x_n - x^*\|)\|x_n - x^*\| < \|x_n - x^*\|, \tag{21.2.18}$$

$$\|H_{\theta,n}\| \leq 2g_3(\|x_n - x^*\|)\|x_n - x^*\|, \tag{21.2.19}$$

$$\|A_{\theta,n}\| \leq g_4(\|x_n - x^*\|) \tag{21.2.20}$$

$$\|z_n - x^*\| \leq g_5(\|x_n - x^*\|)\|x_n - x^*\| < \|x_n - x^*\|, \tag{21.2.21}$$

$$\|B_{\theta,n}\| \leq g_6(\|x_n - x^*\|) \tag{21.2.22}$$

and

$$\|x_{n+1} - x^*\| \leq g_7(\|x_n - x^*\|)\|x_n - x^*\| < \|x_n - x^*\|. \tag{21.2.23}$$

where the "g" functions are defined above Theorem 21.2.1.

Proof. Using (21.2.13), the definition of r^* and the hypothesis $x_0 \in U(x^*, r^*)$ we get that

$$\|F'(x^*)^{-1}(F'(x_0) - F'(x^*))\| \leq L_0\|x_0 - x^*\| < L_0 r^* < 1. \tag{21.2.24}$$

It follows from (21.2.24) and the Banach Lemma on invertible operators [3, **?**] that $F'(x_0)^{-1} \in L(Y,X)$ and

$$\|F'(x_0)^{-1}F'(x^*)\| \leq \frac{1}{1 - L_0\|x_0 - x^*\|} < \frac{1}{1 - L_0 r^*}. \tag{21.2.25}$$

Hence, y_0 and u_0 are well defined. Using the first substep in method (21.1.2) for $n = 0$, (21.2.2), (21.2.14), (21.2.25) and the definition of function g_1 we obtain in turn that

$$\begin{aligned} y_0 - x^* &= x_0 - x^* - F'(x_0)^{-1}F(x_0) \\ &= -F'(x_0)^{-1}F'(x^*)F'(x^*)^{-1}[F(x_0) - F(x^*) - F'(x_0)(x_0 - x^*)] \end{aligned}$$

so,

$$\begin{aligned}\|y_0-x^*\| &\leq \|F'(x_0)^{-1}F'(x^*)\|\|F'(x^*)^{-1}[F(x_0)-F(x^*)-F'(x_0)(x_0-x^*)]\|\\ &\leq \frac{L\|x_0-x^*\|^2}{2(1-L_0\|x_0-x^*\|)}\\ &= g_1(\|x_0-x^*\|)\|x_0-x^*\| < \|x_0-x^*\| < r^*,\end{aligned}$$

which shows (21.2.17) for $n=0$. We also have from the second substep of method (21.1.2) for $n=0$, (21.2.9), (21.2.15), (21.2.17) and the definition of functions g_1 and g_2 that

$$\begin{aligned}\|u_0-x^*\| &\leq \|y_0-x^*\|+|1-\theta|\|F'(x_0)^{-1}F'(x^*)\|\\ &\quad\times\|\int_0^1 F'(x^*+t(x_0-x^*)dt\|\|x_0-x^*\|\\ &\leq [g_1(\|x_0-x^*\|)+\frac{M|1-\theta|}{1-L_0\|x_0-x^*\|}]\|x_0-x^*\|\\ &= g_2(\|x_0-x^*\|)\|x_0-x^*\| < \|x_0-x^*\| < r^*, \qquad (21.2.26)\end{aligned}$$

which shows (21.2.18) for $n=0$.

Next, we need an estimate on $\frac{1}{2}\|H_{\theta,0}\|$. We have from (21.2.4), (21.2.13), (21.2.25), (21.2.26) and the definition of functions g_2 and g_3 that

$$\begin{aligned}\frac{1}{2}\|H_{\theta,0}\| &\leq \frac{1}{2|\theta|}\|F'(x_0)^{-1}F'(x^*)\|(\|F'(x^*)^{-1}(F'(u_0)-F'(x^*))\|\\ &\quad+\|F'(x^*)^{-1}(F'(x_0)-F'(x^*))\|)\\ &\leq \frac{L_0(\|u_0-x^*\|+\|x_0-x^*\|)}{2|\theta|(1-L_0\|x_0-x^*\|)}\\ &\leq \frac{L_0(\|x_0-x^*\|+g_2(\|x_0-x^*\|)\|x_0-x^*\|)}{2|\theta|(1-L_0\|x_0-x^*\|)}\\ &\leq \frac{L_0(1+g_2(\|x_0-x^*\|))\|x_0-x^*\|}{2|\theta|(1-L_0\|x_0-x^*\|)}\\ &= g_3(\|x_0-x^*\|)\|x_0-x^*\|, \qquad (21.2.27)\end{aligned}$$

which shows (21.2.19) for $n=0$. We also need an estimate on $\|A_{\theta,0}\|$. It follows from (21.2.27) and the definition of $A_{\theta,0}$, g_3, g_4 that

$$\begin{aligned}\|A_{\theta,0}\| &\leq 1+\frac{1}{2}\|H_{\theta,0}\|+\frac{1}{4}\|H_{\theta,0}\|^2\\ &\leq 1+g_3(\|x_0-x^*\|)\|x_0-x^*\|+g_3^2(\|x_0-x^*\|)\|x_0-x^*\|^2\\ &= g_4(\|x_0-x^*\|), \qquad (21.2.28)\end{aligned}$$

which shows (21.2.20) for $n=0$. Then, from the third substep of method (21.1.2) for $n=0$, (21.2.19), (21.2.20), (21.2.28) the definition of functions g_1, g_5 and radius r^*, we have that

$$\begin{aligned}\|z_0-x^*\| &\leq \|y_0-x^*\|+|\gamma|\|A_{\theta,0}\|\|F'(x_0)^{-1}F'(x^*)\| \\ &\quad \|\int_0^1 F'(x^*)^{-1}F'(x^*+t(x_0-x^*)dt\|\|x_0-x^*\| \\ &\leq [g_1(\|x_0-x^*\|)+\frac{M|\gamma|g_4(\|x_0-x^*\|)}{1-L_0\|x_0-x^*\|}]\|x_0-x^*\| \\ &= g_5(\|x_0-x^*\|)\|x_0-x^*\|<\|x_0-x^*\|<r^*, \end{aligned} \tag{21.2.29}$$

which shows (21.2.21) for $n=0$. Next, we need an estimate on $\|B_{\theta,0}\|$. We have by the definition of operator $B_{\theta,0}$ and functions $g_{1,3}$, g_3, g_6 that

$$\|B_{\theta,0}\|\leq 1+2g_{1,3}(\|x_0-x^*\|)\|x_0-x^*\|+4g_3^2(\|x_0-x^*\|)\|x_0-x^*\|^2=g_6(\|x_0-x^*\|), \tag{21.2.30}$$

which shows (21.2.22) for $n=0$. Using the fourth substep in method (21.1.2) for $n=0$, (21.2.3), (21.2.15), (21.2.21), (21.2.22), (21.2.29) the definition of functions g_5, g_6, g_7 and radius r^*, we obtain that

$$\begin{aligned}\|x_1-x^*\| &\leq \|z_0-x^*\|+|\alpha|\|B_{\theta,0}\|\|F'(x_0)^{-1}F'(x^*)\| \\ &\quad \|\int_0^1 F'(x^*)^{-1}F'(x^*+t(z_0-x^*)dt\|\|z_0-x^*\| \\ &\leq (1+\frac{M|\alpha|g_6(\|x_0-x^*\|)}{(1-L_0\|x_0-x^*\|)})\|z_0-x^*\| \\ &= (1+\frac{M|\alpha|g_6(\|x_0-x^*\|)}{(1-L_0\|x_0-x^*\|)})g_5(\|x_0-x^*\|)\|x_0-x^*\|, \end{aligned} \tag{21.2.31}$$

which shows (21.2.23) for $n=0$. By simply replacing y_0,u_0,z_0,x_1 by y_k,u_k,z_k,x_{k+1} in the preceding estimates we arrive at estimates (21.2.17)-(21.2.23). Finally, from the estimate $\|x_{k+1}-x^*\|<\|x_k-x^*\|$, we deduce that $\lim_{k\to\infty}x_k=x^*$.

□

Remark 21.2.2. *1. In view of (21.2.13) and the estimate*

$$\begin{aligned}\|F'(x^*)^{-1}F'(x)\| &= \|F'(x^*)^{-1}(F'(x)-F'(x^*))+I\| \\ &\leq 1+\|F'(x^*)^{-1}(F'(x)-F'(x^*))\|\leq 1+L_0\|x-x^*\|\end{aligned}$$

condition (21.2.15) can be dropped and M can be replaced by

$$M(r)=1+L_0r.$$

Moreover, condition (21.2.14) can be replaced by the popular but stronger conditions

$$\|F'(x^*)^{-1}(F'(x)-F'(y))\|\leq L\|x-y\| \text{ for each } x,y\in D \tag{21.2.32}$$

or

$$\begin{aligned}\|F'(x^*)^{-1}(F'(x^*+t(x-x^*))-F'(x))\| &\leq L(1-t)\|x-x^*\| \text{ for each} \\ &\qquad x,y\in D \text{ and } t\in[0,1].\end{aligned}$$

2. *The results obtained here can be used for operators F satisfying autonomous differential equations [3] of the form*
$$F'(x) = P(F(x))$$
where P is a continuous operator. Then, since $F'(x^*) = P(F(x^*)) = P(0)$, we can apply the results without actually knowing x^*. For example, let $F(x) = e^x - 1$. Then, we can choose: $P(x) = x+1$.

3. *The local results obtained here can be used for projection methods such as the Arnoldi's method, the generalized minimum residual method (GMRES), the generalized conjugate method(GCR) for combined Newton/finite projection methods and in connection to the mesh independence principle can be used to develop the cheapest and most efficient mesh refinement strategies [3, 4].*

4. *The radius r_A given by*
$$r \leq r_A = \frac{1}{L_0 + \frac{L}{2}}. \tag{21.2.33}$$
was shown by us to be the convergence radius of Newton's method [3, 4]
$$x_{n+1} = x_n - F'(x_n)^{-1}F(x_n) \text{ for each } n = 0,1,2,\cdots \tag{21.2.34}$$
under the conditions (21.2.13) and (21.2.32). It follows from (21.2.1) and (21.2.33) that the convergence radius r^* of the method (21.1.2) cannot be larger than the convergence radius r_A of the second order Newton's method (21.2.33). As already noted in [3, 4] r_A is at least as large as the convergence ball given by Rheinboldt [3, 4]
$$r_R = \frac{2}{3L}. \tag{21.2.35}$$
In particular, for $L_0 < L$ we have that
$$r_R < r_A$$
and
$$\frac{r_R}{r_A} \to \frac{1}{3} \; as \; \frac{L_0}{L} \to 0.$$
That is our convergence ball r_A is at most three times larger than Rheinboldt's. The same value for r_R was given by Traub [3, 4].

5. *It is worth noticing that method (21.1.2) is not changing when we use the conditions of Theorem 21.2.1 instead of the stronger (C) conditions used in [30]. Moreover, we can compute the computational order of convergence (COC) defined by*
$$\xi = \ln\left(\frac{\|x_{n+1} - x^*\|}{\|x_n - x^*\|}\right) / \ln\left(\frac{\|x_n - x^*\|}{\|x_{n-1} - x^*\|}\right)$$
or the approximate computational order of convergence
$$\xi_1 = \ln\left(\frac{\|x_{n+1} - x_n\|}{\|x_n - x_{n-1}\|}\right) / \ln\left(\frac{\|x_n - x_{n-1}\|}{\|x_{n-1} - x_{n-2}\|}\right).$$
This way we obtain in practice the order of convergence in a way that avoids the bounds given in [30] involving estimates up to the second Fréchet derivative of operator F.

21.3. Numerical Examples

We present numerical examples in this section.

Example 21.3.1. *Let* $X = Y = \mathbb{R}^2, D = \bar{U}(0,1), x^* = 0$ *and define function* F *on* D *by*

$$F(x) = (\sin x, \frac{1}{3}(e^x + 2x - 1)). \tag{21.3.1}$$

Then, using (21.2.9)-(21.2.15), we get $L_0 = L = 1, M = \frac{1}{3}(e+2), \theta = \frac{3}{4}, \gamma = \frac{3}{5}, \alpha = \frac{3}{100}$. *Then, by (21.2.1) we obtain*

$$r^* = 0.3161 < r_R = r_A = 0.6667$$

Example 21.3.2. *Let* $X = Y = \mathbb{R}^3$, $D = \overline{U}(0,1)$. *Define* F *on* D *for* $v = x, y, z)$ *by*

$$F(v) = (e^x - 1, \frac{e-1}{2}y^2 + y, z). \tag{21.3.2}$$

Then, the Fréchet-derivative is given by

$$F'(v) = \begin{bmatrix} e^x & 0 & 0 \\ 0 & (e-1)y+1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Notice that $x^* = (0,0,0), F'(x^*) = F'(x^*)^{-1} = diag\{1,1,1\}, L_0 = e-1 < L = e, M = e, \theta = \frac{3}{4}, \gamma = \frac{3}{10}, \alpha = \frac{3}{100}$. *Then, by (21.2.1) we obtain*

$$r^* = 0.2136 < r_R = 0.2453 < r_A = 0.3249.$$

Example 21.3.3. *Returning back to the motivational example at the introduction of this chapter, we see that conditions (21.2.12)–(21.2.15) are satisfied for* $x^* = 1, f'(x^*) = 3, f(1) = 0, L_0 = L = 146.6629073$ *and* $M = 101.5578008$. *Hence, the results of Theorem 2.1 can apply but not the ones in [30]. In particular, for* $\theta = 0.9902, \alpha = 0.008$ *and* $\gamma = 0.005$ *hypotheses (21.2.9)-(21.2.15) are satisfied. Moreover, we obtain*

$$r^* = 0.0032 < r_R = 0.0045 \leq r_A = 0.0045.$$

21.4. Conclusion

We present a local convergence analysis of Modified Halley-Like Methods with less computation of inversion in order to approximate a solution of an equation in a Banach space setting. Earlier convergence analysis is based on Lipschitz and Holder-type hypotheses up to the second Fréchet-derivative [1]–[20]. In this chapter the local convergence analysis is based only on Lipschitz hypotheses of the first Fréchet-derivative. Hence, the applicability of these methods is expanded under less computational cost of the constants involved in the convergence analysis.

References

[1] Ahmad, F., Hussain, S., Mir, N.A., A. Rafiq, New sixth order Jarratt method for solving nonlinear equations, *Int. J. Appl. Math. Mech.*, 5(5) (2009), 27-35.

[2] Amat, S., Hernández, M.A., Romero, N., A modified Chebyshev's iterative method with at least sixth order of convergence, *App. Math. Comp.*, 206(1) (2008), 164-174.

[3] Argyros, I.K., *Convergence and Application of Newton-type Iterations*, Springer, 2008.

[4] Argyros, I. K., Hilout, S., A convergence analysis for directional two-step Newton methods, *Numer. Algor.*, 55 (2010), 503-528.

[5] Argyros, I. K., Magreñán, Á.A., On the convergence of an optimal fourth-order family of methods and its dynamics. *App. Math. Comp.*, 252 (2015), 336-346.

[6] Bruns, D.D., Bailey, J.E., Nonlinear feedback control for operating a nonisothermal CSTR near an unstable steady state, *Chem. Eng. Sci.*, 32 (1977), 257-264.

[7] Candela, V., Marquina, A., Recurrence relations for rational cubic methods I: The Halley method, *Computing*, 44 (1990), 169-184.

[8] Candela, V., Marquina, A., Recurrence relations for rational cubic methods II: The Chebyshev method, *Computing*, 45 (1990), 355-367.

[9] Chun, C., Some improvements of Jarratt's method with sixth-order convergence, *App. Math. Comp.*, 190(2) (2007), 1432-1437.

[10] Ezquerro, J.A., Hernández, M.A., A uniparametric Halley-type iteration with free second derivative, *Int. J.Pure and Appl. Math.*, 6 (1) (2003), 99-110.

[11] Ezquerro, J.A., Hernández, M.A., New iterations of R-order four with reduced computational cost. *BIT Numer. Math.*, 49 (2009), 325-342.

[12] Ezquerro, J.A., Hernández, M.A., On the R-order of the Halley method, *J. Math. Anal. Appl.*, 303 (2005), 591-601.

[13] Gutiérrez, J.M., Hernández, M.A., Recurrence relations for the super-Halley method, *Computers Math. Applic.* 36(7) (1998), 1-8.

[14] Ganesh, M., Joshi, M.C., Numerical solvability of Hammerstein integral equations of mixed type, *IMA J. Numer. Anal.*, 11 (1991), 21-31.

[15] Gutiérrez, J.M., Magreñán, Á.A., Romero, N., On the semilocal convergence of Newton-Kantorovich method under center-Lipschitz conditions, *App. Math. Comp.*, 221 (2013), 79-88.

[16] Hernández, M.A., Chebyshev's approximation algorithms and applications, *Comput. Math. Appl.*, 41 (3-4) (2001), 433-455.

[17] Hernández, M.A., M.A. Salanova, Sufficient conditions for semilocal convergence of a fourth order multipoint iterative method for solving equations in Banach spaces. *Southwest J. Pure Appl. Math*, (1) (1999), 29-40.

[18] Jarratt, P., Some fourth order multipoint methods for solving equations, *Math. Comp.* 20(95) (1966), 434-437.

[19] Kou, J., Li, Y., An improvement of the Jarratt method, *App. Math. Comp.*, 189 (2007), 1816-1821.

[20] Kou, J., Wang, X., Semilocal convergence of a modified multi-point Jarratt method in Banach spaces under general continuity conditions, *Numer. Algorithms*, 60 (2012), 369-390.

[21] Magreñán, Á.A., Different anomalies in a Jarratt family of iterative root-finding methods, *App. Math. Comp.*, 233 (2014), 29–38.

[22] Parhi, S.K., Gupta, D.K., Semilocal convergence of a Stirling-like method in Banach spaces, *Int. J. Comput. Methods*, 7(02) (2010), 215-228.

[23] Parhi, S.K., Gupta, D.K., Recurrence relations for a Newton-like method in Banach spaces, *J. Comput. Appl. Math.*, 206(2) (2007), 873-887.

[24] Rall, L.B., Computational solution of nonlinear operator equations, Robert E. Krieger, New York (1979).

[25] Ren, H., Wu, Q., Bi, W., New variants of Jarratt method with sixth-order convergence, *Numer. Algorithms* 52(4) (2009), 585-603.

[26] Wang, X., Kou, J., Li, Y., Modified Jarratt method with sixth order convergence, *Appl. Math. Let.*, 22 (2009), 1798-1802.

[27] Ye, X., Li, C., Convergence of the family of the deformed Euler-Halley iterations under the Hölder condition of the second derivative, *J. Comp. Appl. Math.*, 194(2) (2006), 294-308.

[28] Ye, X., Li, C., Shen, W., Convergence of the variants of the Chebyshev-Halley iteration family under the Hölder condition of the first derivative, *J. Comput. Appl. Math.*, 203(1) (2007), 279-288.

[29] Zhao, Y., Wu, Q., Newton-Kantorovich theorem for a family of modified Halley's method under Hölder continuity condition in Banach spaces, *App. Math. Comp.*, 202(1) (2008), 243-251.

[30] Wang, X., Kou, J., Convergence for modified Halley-like methods with less computation of inversion, *J. Diff. Eq. and Appl.*, 19(9) (2013), 1483-1500.

Chapter 22

Local Convergence for an Improved Jarratt-Type Method in Banach Space

22.1. Introduction

In this chapter we are concerned with the problem of approximating a solution x^* of the equation

$$F(x) = 0, \tag{22.1.1}$$

where F is a Fréchet-differentiable operator defined on a convex subset D of a Banach space X with values in a Banach space Y.

Many problems in computational sciences and other disciplines can be brought in a form like (22.1.1) using mathematical modelling [11, 12, 29, 31]. The solutions of these equations can rarely be found in closed form. That is why most solution methods for these equations are iterative. The study about convergence matter of iterative procedures is usually based on two types: semilocal and local convergence analysis. The semilocal convergence matter is, based on the information around an initial point, to give conditions ensuring the convergence of the iterative procedure; while the local one is, based on the information around a solution, to find estimates of the radii of convergence balls. In particular, the practice of Numerical Functional Analysis for finding solution x^* of equation (22.1.1) is essentially connected to variants of Newton's method. This method converges quadratically to x^* if the initial guess is close enough to the solution. Iterative methods of convergence order higher than two such as Chebyshev-Halley-type methods [5, 6, 11, 14, 23, 30, 12, 19, 20, 21, 22, 24, 25, 26, 27, 28, 31, 33] require the evaluation of the second Fréchet-derivative, which is very expensive in general. However, there are integral equations, where the second Fréchet-derivative is diagonal by blocks and inexpensive or for quadratic equations the second Fréchet-derivative is constant. Moreover, in some applications involving stiff systems, high order methods are usefull. That is why in a unified way we study the local convergence of the improved Jarratt-type method (IJTM) defined for each $n = 0,1,2,\ldots$ by

$$\begin{aligned}
u_n &= x_n - F'(x_n)^{-1}F(x_n),\\
y_n &= x_n + \frac{2}{3}(u_n - x_n),\\
J_n &= \left(6F'(y_n) - 2F'(x_n)\right)^{-1}\left(3F'(y_n) + F'(x_n)\right),\\
z_n &= x_n - J_nF'(x_n)^{-1}F(x_n),\\
x_{n+1} &= z_n - (2J_n - I)F'(x_n)^{-1}F(z_n),
\end{aligned} \tag{22.1.2}$$

where x_0 is an initial point and I is the identity operator. If we set $H_n = F'(x_n)^{-1}(F'(y_n) - F'(x_n))$, then using some algebraic manipulation we obtain that

$$J_n = \frac{1}{2}\left(I + \left(I + \frac{3}{2}H_n\right)^{-1}\right) = I - \frac{3}{4}\left(I + \frac{3}{2}H_n\right)^{-1}H_n. \tag{22.1.3}$$

This method has been shown to be of convergence order between 5 and 6 [29, 33]. The usual conditions for the semilocal convergence of these methods are $(\mathcal{C})$:

$(\mathcal{C}_1)$ There exists $\Gamma_0 = F'(x_0)^{-1}$ and $\|\Gamma_0\| \le \beta$, $\beta > 0$;

$(\mathcal{C}_2)$ $\|\Gamma_0 F(x_0)\| \le \eta$, $\eta \ge 0$;

$(\mathcal{C}_3)$ $\|F''(x)\| \le \beta_1$ for each $x \in D$, $\beta_1 \ge 0$;

$(\mathcal{C}_4)$ $\|F'''(x)\| \le \beta_2$ for each $x \in D$, $\beta_2 \ge 0$

or

$(\mathcal{C}'_4)$ $\|F'''(x_0)\| \le \overline{\beta}_2$ for each $x \in D$, $\overline{\beta}_2 \ge 0$ and some $x_0 \in D$;

$(\mathcal{C}_5)$ $\|F'''(x) - F'''(y)\| \le \beta_3\|x - y\|$ for each $x, y \in D$

or $\|F'''(x) - F'''(y)\| \le \varphi(\|x - y\|))$ for each $x, y \in D$, where $\varphi : [0,+\infty) \to [0,+\infty)$ is a non-decreasing function.

The local convergence conditions are similar but x_0 is x^* in $(\mathcal{C}_1)$ and $(\mathcal{C}_2)$. There is a plethora of local and semilocal convergence results under the $(\mathcal{C})$ conditions [1]–[33]. These conditions restrict the applicability of these methods. That is why, in our chapter we assume the conditions $(\mathcal{A})$:

$(\mathcal{A}_1)$ $F : D \to Y$ is Fréchet-differentiable and there exists $x^* \in D$ such that $F(x^*) = 0$ and $F'(x^*)^{-1} \in L(Y,X)$;

$(\mathcal{A}_2)$ $\|F'(x^*)^{-1}(F'(x) - F'(x^*))\| \le L_0\|x - x^*\|$ for each $x \in D$;

$(\mathcal{A}_3)$ $\|F'(x^*)^{-1}(F'(x) - F'(y))\| \le L\|x - y\|$ for each $x, y \in D$;

and

$(\mathcal{A}_4)$ $\|F'(x^*)^{-1}F'(x)\| \le K$ for each $x \in D$, $k > 0$.

Notice that the $(\mathcal{A})$ conditions are weaker than the $(\mathcal{C})$ conditions. Hence, the applicability of (IJTM) is expanded under the $(\mathcal{A})$ conditions.

As a motivational example, let us define function f on $D = \overline{U}\left(1, \frac{3}{2}\right)$ by

$$f(x) = \begin{cases} x^3 \ln x^2 + x^5 - x^4, & x \neq 0 \\ 0, & x = 0. \end{cases}$$

Choose $x^* = 1$. We have that

$$f'(x) = 3x^2 \ln x^2 + 5x^4 - 4x^3 + 2x^2,$$

$$f''(x) = 6x \ln x^2 + 20x^3 + 12x^2 + 10x$$

and

$$f'''(x) = 6 \ln x^2 + 60x^2 - 24x + 22.$$

Notice that $f'''(x)$ is bounded on D. That is condition $(\mathcal{C}_4)$ is not satisfied. Hence, the results depending on $(\mathcal{C}_4)$ cannot apply in this case. However, we have $f'(x^*) = 3$ and $f(1) = 0$. That is, conditions $(\mathcal{A}_1)$ is satisfied. Moreover, conditions $(\mathcal{A}_2), (\mathcal{A}_3)$ are satisfied for $L_0 = L = 146.6629073\ldots$ and $K = 101.5578008\ldots$. Then, condition $(\mathcal{A}_4)$ is also satisfied. Hence, the results of our Theorem 22.2.1 that follows can apply to solve equation $f(x) = 0$ using IJTM. Hence, the applicability of IJTM is expanded under the conditions $(\mathcal{A})$.

The chapter is organized as follows: In Section 22.2. we present the local convergence of these methods. The numerical examples are given in the concluding Section 22.3..

In the rest of this chapter, $U(w,q)$ and $\overline{U}(w,q)$ stand, respectively, for the open and closed ball in X with center $w \in X$ and of radius $q > 0$.

22.2. Local Convergence

In this section we present the local convergence of IJTM under the $(\mathcal{A})$ conditions. It is convenient for the local convergence of IJTM to introduce some funcitons and parameters.

Let $L_0 > 0, L > 0$ and $K > 0$ be given constants. Define parameters r_A and r_0 by

$$r_A = \frac{2}{2L_0 + L} \tag{22.2.1}$$

and

$$r_0 = \frac{\sqrt{2}}{\sqrt{2}L_0 + L}. \tag{22.2.2}$$

Notice that

$$r_0 < r_A < \frac{1}{L_0}. \tag{22.2.3}$$

Define functions f_1 and f_2 on the interval $\left[0, \frac{1}{L_0}\right)$ by

$$f_1(t) = \frac{Lt}{2(1 - L_0 t)} \tag{22.2.4}$$

and

$$f_2(t) = \frac{1}{3}\left(1 + \frac{Lt}{1 - L_0 t}\right). \tag{22.2.5}$$

Then, we have by the choice of r_A that

$$f_1(t) \le 1 \quad \text{for each} \quad t \in [0, r_A] \tag{22.2.6}$$

and

$$f_2(t) \le 1 \quad \text{for each} \quad t \in [0, r_A]. \tag{22.2.7}$$

Define function f_3 on the interval $\left[0, \frac{1}{L_0}\right)$ by

$$f_3(t) = \frac{(Lt)^2}{2(1 - L_0 t)^2}. \tag{22.2.8}$$

Then, we have that

$$f_3(t) \le 1 \quad \text{for each} \quad t \in [0, r_0] \tag{22.2.9}$$

and

$$f_3(t) < 1 \quad \text{for each} \quad t \in [0, r_0). \tag{22.2.10}$$

Moreover, define functions f_4 and f_5 on the interval $[0, r_0)$ by

$$f_4(t) = \frac{Lt^2}{2(1 - L_0 t)}\left[1 + \frac{L^2 K t}{2(1 - L_0 t)^2 - L^2 t^2}\right] \tag{22.2.11}$$

and

$$f_5(t) = \left[1 + \frac{2K}{2(1 - L_0 t)^2 - L^2 t^2}\right] f_4(t). \tag{22.2.12}$$

Furthermore, define functions $\overline{f}_4$ and $\overline{f}_5$ on the interval $[0, r_0)$ by

$$\overline{f}_4(t) = f_4(t) - 1 \tag{22.2.13}$$

and

$$\overline{f}_5(t) = f_5(t) - 1. \tag{22.2.14}$$

We have that $\overline{f}_4(0) = \overline{f}_5(0) = -1 < 0$ and $\overline{f}_4(t) \to +\infty$, $\overline{f}_5(t) \to +\infty$ as $t \to r_0$. It follows by intermediate value theorem that functions $\overline{f}_4$ and $\overline{f}_5$ has zeros in $(0, r_0)$. Denote by r_4 and r_5 the minimal zeros of functions $\overline{f}_4$ and $\overline{f}_5$ on the interval $(0, r_0)$, respectively. Finally, define

$$r = \min\{r_4, r_5\}. \tag{22.2.15}$$

Then, we have by the choice of r that

$$f_1(t) < 1, \tag{22.2.16}$$

$$f_2(t) < 1, \tag{22.2.17}$$

$$f_3(t) < 1, \tag{22.2.18}$$

$$f_4(t) < 1, \tag{22.2.19}$$

and

$$f_5(t) < 1 \quad \text{for each} \quad t \in [0, r). \tag{22.2.20}$$

Next, we present the main local convergence for IJTM under the $(\mathcal{A})$ conditions.

Theorem 22.2.1. *Suppose that the $(\mathcal{A})$ conditions and $\overline{U}(x^*,r) \subseteq D$, where r is given by (22.2.15). Then, sequence $\{x_n\}$ generated by IJTM (22.1.2) for any $x_0 \in U(x^*,r)$ is well defined, remains in $U(x^*,r)$ for each $n = 0,1,2,\ldots$ and converges to x^*. Moreover, the following estimates hold for each $n = 0,1,2,\ldots$*

$$\|x_{n+1} - x^*\| \leq f_5(\|x_n - x^*\|)\|x_n - x^*\| < \|x_n - x^*\| < r, \tag{22.2.21}$$

where function f_5 is defined by (22.2.12).

Proof. We shall use induction to show that estimates (22.2.20) hold for each $n = 0,1,2,\ldots$ Using $(\mathcal{A}_2)$ and the hypothesis $x_0 \in U(x^*,r)$, we have that

$$\|F'(x^*)^{-1}(F'(x_0) - F'(x^*))\| \leq L_0\|x_0 - x^*\| < L_0 r < 1, \tag{22.2.22}$$

by the choice of r. It follows from (22.2.22) and the Banach lemma on invertible operators that [11, 12, 28] $F'(x_0)^{-1} \in L(Y,X)$ and

$$\|F'(x_0)^{-1}F'(x^*)\| \leq \frac{1}{1 - L_0\|x_0 - x^*\|} < \frac{1}{1 - L_0 r}. \tag{22.2.23}$$

Using the first substep of IJTM for $n = 0$, $F(x^*) = 0$, $(\mathcal{A}_1)$, $(\mathcal{A}_2)$, (22.2.22) and the choice of r we get that

$$\begin{aligned} u_0 - x^* &= x_0 - x^* - F'(x_0)^{-1}F(x_0) \\ &= -(F'(x_0)^{-1}F'(x^*))\Bigg[F'(x^*)^{-1} \\ &\quad \times \int_0^1 (F'(x^* + \theta(x_0 - x^*)) - F'(x_0))\, d\theta(x_0 - x^*)\Bigg], \end{aligned} \tag{22.2.24}$$

so

$$\begin{aligned} \|u_0 - x^*\| &\leq \|F'(x_0)^{-1}F'(x^*)\| \\ &\quad \times \left\|F'(x^*)^{-1}\int_0^1 (F'(x^* + \theta(x_0 - x^*)) - F'(x_0))\, d\theta\right\| \|x_0 - x^*\| \\ &\leq \frac{L_0\|x_0 - x^*\|^2}{2(1 - L_0\|x_0 - x^*\|)} \leq \frac{L\|x_0 - x^*\|^2}{2(1 - L_0\|x_0 - x^*\|)} \\ &\leq f_1(r)\|x_0 - x^*\| < \|x_0 - x^*\| < r, \end{aligned} \tag{22.2.25}$$

which shows $u_0 \in U(x^*,r)$. Using the second substep of IJTM, we get by (22.2.25) and (22.2.17) that

$$\begin{aligned} y_0 - x^* &= x_0 - x^* + \frac{2}{3}(u_0 - x_0) \\ &= x_0 - x^* + \frac{2}{3}(u_0 - x^*) + \frac{2}{3}(x^* - x_0) \\ &= \frac{1}{3}(x_0 - x^*) + \frac{2}{3}(u_0 - x^*) \end{aligned}$$

so,

$$\|y_0 - x^*\| \leq \frac{1}{3}\|x_0 - x^*\| + \frac{2}{3}\|u_0 - x^*\| \leq f_2(r)\|x_0 - x^*\| < r,$$

which shows that $y_0 \in U(x^*, r)$.

Next, we shall find upper bounds on $\|H_0\|$ and $\|J_0\|$. Using $(\mathcal{A}_1)$, (22.2.24), (22.2.18) that

$$\begin{aligned}
\frac{3}{2}\|H_0\| &\leq \frac{3}{2}\|F'(x_0)^{-1}F'(x^*)\|\|F'(x^*)^{-1}(F'(y_0) - F'(x_0))\| \\
&\leq \frac{3}{2}\frac{L\|y_0 - x_0\|}{1 - L_0\|x_0 - x^*\|} \leq \frac{3}{2}\cdot\frac{2}{3}\frac{L\|u_0 - x_0\|}{1 - L_0\|x_0 - x^*\|} \\
&\leq \frac{L^2\|x_0 - x^*\|^2}{2(1 - L_0\|x_0 - x^*\|^2)} < \left(\frac{Lr}{\sqrt{2}(1 - L_0 r)}\right)^2 \\
&= (f_3(r))^2 < 1. \qquad (22.2.26)
\end{aligned}$$

It follows from (22.2.25) and the Banach lemma on invertible operators that $\left(I + \frac{3}{2}H_0\right)^{-1} \in L(Y, X)$ and

$$\begin{aligned}
\left\|\left(I + \frac{3}{2}H_0\right)^{-1}\right\| &\leq \frac{1}{1 - \frac{L^2\|x_0 - x^*\|^2}{2(1 - L_0\|x_0 - x^*\|)^2}} \\
&< \frac{1}{1 - \frac{L^2 r^2}{2(1 - L_0 r)^2}}. \qquad (22.2.27)
\end{aligned}$$

It then follows from the definition of J_0, (22.2.26) and (22.2.27) that

$$\begin{aligned}
\|J_0\| &\leq 1 + \frac{3}{4}\frac{\frac{L^2\|x_0 - x^*\|^2}{3(1 - L_0\|x_0 - x^*\|)}}{1 - \frac{L^2\|x_0 - x^*\|^2}{2(1 - L_0\|x_0 - x^*\|)^2}} \\
&= 1 + \frac{1}{2}\cdot\frac{(1 - L_0\|x_0 - x^*\|)L^2\|x_0 - x^*\|^2}{[2(1 - L_0\|x_0 - x^*\|)^2 - L^2\|x_0 - x^*\|^2]}. \qquad (22.2.28)
\end{aligned}$$

Then, from the fourth substep of IJTM for $n = 0$, (22.2.25), (22.2.26), (22.2.27), (22.2.19) and $(\mathcal{A}_4)$

$$z_0 = x_0 - F'(x_0)^{-1}F(x_0) + \frac{3}{4}\left(I + \frac{3}{2}H_0\right)^{-1} H_0 F'(x_0)^{-1}F(x_0)$$

so,

$$\begin{aligned}
\|z_0-x^*\| &\leq \|x_0-x^*-F'(x_0)^{-1}F(x_0)\| \\
&\quad +\frac{3}{4}\left\|\left(I+\frac{3}{2}H_0\right)^{-1}\right\| \|H_0\| \|F'(x_0)^{-1}F'(x^*)\| \\
&\quad \times \left\|F'(x^*)^{-1}\int_0^1 F'(x^*+\theta(x_0-x^*))(x_0-x^*)\,d\theta\right\| \\
&\leq \frac{L\|x_0-x^*\|^2}{2(1-L_0\|x_0-x^*\|)}+\frac{3}{4}\frac{1}{1-\frac{L^2\|x_0-x^*\|}{2(1-L_0\|x_0-x^*\|)}} \\
&= \frac{2}{3}\frac{L^2\|x_0-x^*\|}{2(1-L_0\|x_0-x^*\|)^2}\frac{K\|x_0-x^*\|}{1-L_0\|x_0-x^*\|)} \\
&= f_4(\|x_0-x^*\|)\|x_0-x^*\| < \|x_0-x^*\| < r, \qquad (22.2.29)
\end{aligned}$$

which shows $z_0 \in U(x^*, r)$.

Notice that we used

$$F(x_0) = F(x_0)-F(x^*) = \int_0^1 F'(x^*+\theta(x_0-x^*))(x_0-x^*)\,d\theta$$

so

$$\|F'(x^*)^{-1}F(x_0)\| \leq K\|x_0-x^*\| \quad \text{by} \quad (\mathcal{A}_4). \qquad (22.2.30)$$

Next, using the last substep in IJTM for $n=0$, (22.2.23), (22.2.27), (22.2.19) and (22.2.30) (for x_0 replaced by z_0) we get in turn that

$$\begin{aligned}
\|x_1-x^*\| &\leq \|z_0-x^*\|+\frac{1}{1-\frac{L^2\|x_0-x^*\|^2}{2(1-L_0\|x_0-x^*\|)^2}}\frac{K\|z_0-x^*\|}{1-L_0\|x_0-x^*\|} \\
&\quad \times\left[1+\frac{2K(1-L_0\|x_0-x^*\|)}{2(1-L_0\|x_0-x^*\|)^2-L^2\|x_0-x^*\|^2}\right]\|z_0-x^*\| \\
&\leq f_5(\|x_0-x^*\|)\|x_0-x^*\| \leq f_5(r)\|x_0-x^*\| \\
&< \|x_0-x^*\|, \qquad (22.2.31)
\end{aligned}$$

which shows (22.2.21) for $n=0$.

To complete the induction, simple replace in all preceding estimates x_0, u_0, y_0, z_0, x_1 by $x_k, u_k, y_k, z_k, x_{k+1}$, respectively to arrive at (22.2.21), which complete the induction.

Finally it follows from (22.2.21) that $\lim_{k\to+\infty} x_k = x^*$ ■

Remark 22.2.2. *(a) Condition ($\mathcal{A}_2$) can be dropped, since this condition follows from ($\mathcal{A}_3$). Notice, however that*

$$L_0 \leq L \qquad (22.2.32)$$

holds in general and $\frac{L}{L_0}$ can be arbitrarily large [2, 3, 4, 5, 6].

(b) *In view of condition ($\mathcal{A}_2$) and the estimate*

$$\begin{aligned}\|F'(x^*)^{-1}F'(x)\| &= \|F'(x^*)^{-1}[F'(x)-F'(x^*)]+I\| \\ &\leq 1+\|F'(x^*)^{-1}(F'(x)-F'(x^*))\| \\ &\leq 1+L_0\|x-x^*\|,\end{aligned}$$

condition ($\mathcal{A}_4$) can be dropped and K can be replaced by

$$K(r)=1+L_0r. \tag{22.2.33}$$

(c) *It is worth noticing that r is such that*

$$r<r_A \text{ for } \alpha\neq 0. \tag{22.2.34}$$

The convergence ball of radius r_A was given by us in [2, 3, 5] for Newton's method under conditions ($\mathcal{A}_1$)-($\mathcal{A}_3$). Estimate (22.2.22) shows that the convergence ball of higher than two IJTM methods is smaller than the convergence ball of the quadratically convergent Newton's method. The convergence ball given by Rheinboldt [31] for Newton's method is

$$r_R=\frac{2}{3L}<r_A \tag{22.2.35}$$

if $L_0<L$ and $\frac{r_R}{r_A}\to\frac{1}{3}$ as $\frac{L_0}{L}\to 0$. Hence, we do not expect r to be larger than r_A no matter how we choose L_0, L and K. Finally note that if $\alpha=0$, then IJTM reduces to Newton's method and $r=r_A$.

(d) *The local results can be used for projection methods such as Arnoldi's method, the generalized minimum residual method (GMREM), the generalized conjugate method (GCM) for combined Newton/finite projection methods and in connection to the mesh independence principle in order to develop the cheapest and most efficient mesh refinement strategy [11, 12, 31].*

(e) *The results can also be used to solve equations where the operator F' satisfies the autonomous differential equation [11, 12, 29, 31]:*

$$F'(x)=T(F(x)), \tag{22.2.36}$$

where T is a known continuous operator. Since $F'(x^*)=T(F(x^*))=T(0)$, we can apply the results without actually knowing the solution x^*. Let as an example $F(x)=e^x-1$. Then, we can choose $T(x)=x+1$ and $x^*=0$.

(f) *It is worth noticing that IJTM is not changing if we use the ($\mathcal{A}$) instead of the ($\mathcal{C}$) conditions. Moreover for the error bounds in practice we can use the computational order of convergence (COC) [1, 2, 3, 4, 11, 12, 15] using*

$$\xi=\sup\frac{\ln\left(\frac{\|x_{n+2}-x_{n+1}\|}{\|x_{n+1}-x_n\|}\right)}{\ln\left(\frac{\|x_{n+1}-x_n\|}{\|x_n-x_{n-1}\|}\right)} \quad \textit{for each} \quad n=1,2,\ldots$$

or the approximate computational order of convergence (ACOC)

$$\xi^* = \sup \frac{\ln\left(\frac{\|x_{n+2}-x^*\|}{\|x_{n+1}-x^*\|}\right)}{\ln\left(\frac{\|x_{n+1}-x^*\|}{\|x_n-x^*\|}\right)} \quad \textit{for each} \quad n = 0,1,2,\ldots$$

instead of the error bounds obtained in Theorem 22.2.1.

22.3. Numerical Examples

We present numerical examples where we compute the radii of the convergence balls.

Example 22.3.1. *Let* $X = Y = \mathbb{R}^3$, $D = \overline{U}(0,1)$. *Define* F *on* D *for* $v = (x,y,z)$ *by*

$$F(v) = \left(e^x - 1, \frac{e-1}{2}y^2 + y, z\right). \tag{22.3.1}$$

Then, the Fréchet-derivative is given by

$$F'(v) = \begin{bmatrix} e^x & 0 & 0 \\ 0 & (e-1)y+1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Notice that $x^* = (0,0,0)$, $F'(x^*) = F'(x^*)^{-1} = diag\{1,1,1\}$, $L_0 = e-1 < L = K = e$, $r_0 = 0.274695\ldots < r_A = 0.324967\ldots < 1/L_0 = 0.581977\ldots$, $r = 0.144926\ldots$.

Example 22.3.2. *Let* $X = Y = C([0,1])$, *the space of continuous functions defined on* $[0,1]$ *be and equipped with the max norm. Let* $D = \overline{U}(0,1)$. *Define function* F *on* D *by*

$$F(\varphi)(x) = \varphi(x) - 5\int_0^1 x\theta\varphi(\theta)^3 d\theta. \tag{22.3.2}$$

We have that

$$F'(\varphi(\xi))(x) = \xi(x) - 15\int_0^1 x\theta\varphi(\theta)^2\xi(\theta)d\theta, \text{ for each } \xi \in D.$$

Then, we get that $x^* = 0$, $L_0 = 7.5$, $L = 15$ *and* $K = K(t) = 1 + 7.5t$, $r_0 = 0.055228\ldots < r_A = 0.066666\ldots < 1/L_0 = 0.133333\ldots, r = 0.0370972\ldots$.

Example 22.3.3. *Returning to the motivational example at the Introduction of this chapter, let the function* f *on* $D = \overline{U} = \left(1, \frac{3}{2}\right)$ *defined by*

$$f(x) = \begin{cases} x^3 \ln x^2 + x^5 - x^4, & x \neq 0 \\ 0, & x = 0. \end{cases}$$

Then, $L_0 = L = 146.662907\ldots$, $K = 101.557800\ldots$, $r_0 = 0.003984\ldots < r_A = 0.004545\ldots < 1/L_0 = 0.006818\ldots$ *and* $r = 0.000442389\ldots$.

References

[1] Amat, S., Busquier, S., Gutiérrez, J. M.,. Geometric constructions of iterative functions to solve nonlinear equations, *J. Comput. Appl. Math.*, 157 (2003), 197-205.

[2] Amat, S., Busquier, S., Plaza, S., Review of some iterative root-finding methods from a dynamical point of view, *Scientia Series A: Mathematical Sciences*, 10 (2004), 3-35.

[3] Amat, S., Busquier, S., Plaza, S., Dynamics of the King and Jarratt iteration, *Aequationes Math.*, 69, (2005), 3, 212-223.

[4] Amat, S., Busquier, S., Plaza, S., Chaotic dynamics of a third-order Newton-type method, *J. Math. Anal. Appl.*, 366(1) (2010), 24-32.

[5] Argyros, I. K., On the convergence of Chebyshev-Halley-type method under Newton-Kantorovich hypotheses, *Appl. Math. Let.*, 6(5) (1993), 71-74.

[6] Argyros, I. K., A note on the Halley method in Banach spaces, *App. Math. Comp.*, 58 (1993), 215-224.

[7] Argyros, I. K., The Jarratt method in a Banach space setting, *J. Comp. Appl. Math.*, 51 (1994), 103-106.

[8] Argyros, I. K., A multi-point Jarratt-Newton-type approximation algorithm for solving nonlinear operator equations in Banach spaces, *Func. Approx. Comment. Math.* XXIII (1994), 97-108.

[9] Argyros, I. K.,. A new convergence theorem for the Jarratt method in Banach spaces, *Computers and Mathematics with applications*, 36 (8) (1998), 13-18.

[10] Argyros, I. K., Chen, D.. An inverse-free Jarratt type approximation in a Banach space, *J. Approx. Th. Appl.* 12(1) (1996), 19-30.

[11] Argyros, I. K., Computational theory of iterative methods. *Series: Studies in Computational Mathematics*, 15, Editors: C.K. Chui and L. Wuytack, Elsevier Publ. Co. New York, U.S.A, 2007.

[12] Argyros, I. K., Hilout, S., *Numerical methods in Nonlinear Analysis*, World Scientific Publ. Comp. New Jersey, 2013.

[13] Argyros, I. K., and Hilout, S., Weaker conditions for the convergence of Newton's method, *J. Complexity*, 28 (2012), 364-387.

[14] Argyros, I. K., Magreñán, Á.A., On the convergence of an optimal fourth-order family of methods and its dynamics. *App. Math. Comp.*, 252 (2015), 336-346.

[15] Chicharro,F., Cordero, A., Gutiérrez, J. M., Torregrosa, J. R., Dynamics of derivative-free methods for nonlinear equations, *App. Math. Comp.*, 219 (2013), 7023-7035.

[16] Chun, C., Lee, M. Y., Neta, B., Dzunić, J., On optimal fourth-order iterative methods free from second derivative and their dynamics, *App. Math. Comp.*, 218 (2012), 6427-6438.

[17] Candela, V., Marquina, A., Recurrence relations for rational cubic methods I: The Halley method, *Computing*, 44 (1990), 169-184.

[18] Candela, V., Marquina, A., Recurrence relations for rational cubic methods II: The Chebyshev method, *Computing*, 45 (1990), 355-367.

[19] Ezquerro, J.A., Hernández, M. A., Avoiding the computation of the second Fréchet-derivative in the convex acceleration of Newton's method, *J. Comp. Appl. Math.*, 96 (1998), 1-12.

[20] Ezquerro, J.A., Hernández, M. A., On Halley-type iterations with free second derivative, *J. Comp. Appl. Math.*, 170 (2004), 455-459.

[21] Gutiérrez, J. M., Hernández, M. A., Recurrence relations for the super-Halley method, *Computers Math. Applic.* 36 (1998), 1-8.

[22] Gutiérrez, J. M., Hernández, M. A., Third-order iterative methods for operators with bounded second derivative, *J. Comp. Appl. Math.*, 82 (1997), 171-183.

[23] Gutiérrez, J.M., Magreñán, Á.A., Romero, N., On the semilocal convergence of Newton-Kantorovich method under center-Lipschitz conditions, *App. Math. Comp.*, 221 (2013), 79–88.

[24] Hernández, M. A., Salanova, M. A., Modification of the Kantorovich assumptions for semilocal convergence of the Chebyshev method, *J. Comp. Appl. Math.*, 126 (2000), 131-143.

[25] Hernández, M. A., Chebyshev's approximation algorithms and applications, *Computers Math. Applic.*, 41 (2001), 433-455.

[26] Hernández, M. A., Reduced recurrence relations for the Chebyshev method, *J. Optim. Theory App.*, 98 (1998), 385-397.

[27] Jarratt, P., Some fourth order multipoint iterative methods for solving equations, *Mathematics and Computation*, 20(95) (1996), 434-437.

[28] Kantorovich, L.V., Akilov, G. P., *Functional Analysis*, Pergamon Press, Oxford, 1982.

[29] Kou, J., Li, Y., *An improvement of the Jarratt method*, *App. Math. Comp.*, 189, 2, (2007), 1816-1821.

[30] Magreñán, Á.A., Different anomalies in a Jarratt family of iterative root-finding methods, *App. Math. Comp.*, 233 (2014), 29–38.

[31] Ortega, J.M., Rheinboldt, W.C., *Iterative Solution of Nonlinear Equations in Several Variables*, Academic press, New York, 1970.

[32] Parida, P.K., Gupta, D.K., Semilocal convergence of a family of third order methods in Banach spaces under Hölder continuous second derivative, *Nonl. Anal.*, 69 (2008), 4163-4173.

[33] Ren, H., Wo, Q., Bi, W., New variants of Jarratt method with sixth-order convergence, *Numer. Algorithms*, 52(4) (2009), 585-603.

Chapter 23

Enlarging the Convergence Domain of Secant-Like Methods for Equations

23.1. Introduction

Let $\mathcal{X}$, $\mathcal{Y}$ be Banach spaces and $\mathcal{D}$ be a non-empty, convex and open subset in $\mathcal{X}$. Let $U(x,r)$ and $\overline{U}(x,r)$ stand, respectively, for the open and closed ball in $\mathcal{X}$ with center x and radius $r > 0$. Denote by $\mathcal{L}(\mathcal{X},\mathcal{Y})$ the space of bounded linear operators from $\mathcal{X}$ into $\mathcal{Y}$. In the present chapter we are concerned with the problem of approximating a locally unique solution $x^\star$ of equation

$$F(x) = 0, \tag{23.1.1}$$

where F is a Fréchet continuously differentiable operator defined on $\mathcal{D}$ with values in $\mathcal{Y}$.

A lot of problems from computational sciences and other disciplines can be brought in the form of equation (23.1.1) using Mathematical Modelling [8, 10, 14]. The solution of these equations can rarely be found in closed form. That is why most solution methods for these equations are iterative. In particular, the practice of numerical analysis for finding such solutions is essentially connected to variants of Newton's method [8, 10, 14, 22, 25, 27, 32].

A very important aspect in the study of iterative procedures is the convergence domain. In general the convergence domain is small. This is why it is important to enlarge it without additional hypotheses. Then, this is our goal in this chapter.

In the present chapter we study the secant-like method defined by

$$\begin{aligned} &x_{-1}, x_0 \text{ are initial points} \\ &y_n = \lambda x_n + (1-\lambda)x_{n-1},\ \lambda \in [0,1] \\ &x_{n+1} = x_n - B_n^{-1}F(x_n), \quad B_n = [y_n, x_n; F] \qquad \text{for each} \quad n = 0,1,2,\cdots. \end{aligned} \tag{23.1.2}$$

The family of secant-like methods reduces to the secant method if $\lambda = 0$ and to Newton's method if $\lambda = 1$. It was shown in [27] (see also [7, 8, 21] and the references therein) that the R–order of convergence is at least $(1+\sqrt{5})/2$ if $\lambda \in [0,1)$, the same as that of the secant method. In the real case the closer x_n and y_n are, the higher the speed of convergence.

Moreover in [19], it was shown that as λ approaches 1 the speed of convergence is close to that of Newton's method. Moreover, there exist new graphical tools [24]. Furthermore, the advantages of using secant-like method instead of Newton's method is that the former method avoids the computation of $F'(x_n)^{-1}$ at each step. The study about convergence matter of iterative procedures is usually centered on two types: semilocal and local convergence analysis. The semilocal convergence matter is, based on the information around an initial point, to give criteria ensuring the convergence of iterative procedure; while the local one is, based on the information around a solution, to find estimates of the radii of convergence balls. There is a plethora of studies on the weakness and/or extension of the hypothesis made on the underlying operators; see for example [1]–[34].

The hypotheses used for the semilocal convergence of secant-like method are (see [8, 18, 19, 21]):

($\mathcal{C}_1$) There exists a divided difference of order one denoted by $[x,y;F] \in \mathcal{L}(\mathcal{X},\mathcal{Y})$ satisfying

$$[x,y;F](x-y) = F(x) - F(y) \quad \text{for all} \quad x,y \in \mathcal{D};$$

($\mathcal{C}_2$) There exist x_{-1}, x_0 in $\mathcal{D}$ and $c > 0$ such that

$$\| x_0 - x_{-1} \| \leq c;$$

($\mathcal{C}_3$) There exist $x_{-1}, x_0 \in \mathcal{D}$ and $M > 0$ such that $A_0^{-1} \in \mathcal{L}(\mathcal{Y},\mathcal{X})$ and

$$\| A_0^{-1}([x,y;F] - [u,v;F]) \| \leq M(\| x-u \| + \| y-v \|) \quad \text{for all} \quad x,y,u,v \in \mathcal{D};$$

($\mathcal{C}_3^{\star}$) There exist $x_{-1}, x_0 \in \mathcal{D}$ and $L > 0$ such that $A_0^{-1} \in \mathcal{L}(\mathcal{Y},\mathcal{X})$ and

$$\| A_0^{-1}([x,y;F] - [v,y;F]) \| \leq L \| x-v \| \quad \text{for all} \quad x,y,v \in \mathcal{D};$$

($\mathcal{C}_3^{\star\star}$) There exist $x_{-1}, x_0 \in \mathcal{D}$ and $K > 0$ such that $F(x_0)^{-1} \in \mathcal{L}(\mathcal{Y},\mathcal{X})$ and

$$\| F'(x_0)^{-1}([x,y;F] - [v,y;F]) \| \leq K \| x-v \| \quad \text{for all} \quad x,y,v \in \mathcal{D};$$

($\mathcal{C}_4$) There exists $\eta > 0$ such that

$$\| A_0^{-1} F(x_0) \| \leq \eta;$$

($\mathcal{C}_4^{\star}$) There exists $\eta > 0$ for each $\lambda \in [0,1]$ such that

$$\| B_0^{-1} F(x_0) \| \leq \eta.$$

We shall refer to ($\mathcal{C}_1$)–($\mathcal{C}_4$) as the ($\mathcal{C}$) conditions. From analyzing the semilocal convergence of the simplified secant method, it was shown [18] that the convergence criteria are milder than those of secant-like method given in [20]. Consequently, the decreasing and accessibility regions of (23.1.2) can be improved. Moreover, the semilocal convergence of (23.1.2) is guaranteed.

In the present chapter we show: an even larger convergence domain can be obtained under the same or weaker sufficient convergence criteria for method (23.1.2). In view of ($\mathcal{C}_3$) we have that

($\mathcal{C}_5$) There exists $M_0 > 0$ such that

$$\| A_0^{-1}([x,y;F]-[x_{-1},x_0;F]) \| \leq M_0 (\| x-x_{-1} \| + \| y-x_0 \|) \quad \text{for all} \quad x,y \in \mathcal{D}.$$

We shall also use the conditions

($\mathcal{C}_6$) There exist $x_0 \in \mathcal{D}$ and $M_1 > 0$ such that $F'(x_0)^{-1} \in \mathcal{L}(\mathcal{Y},\mathcal{X})$ and

$$\| F'(x_0)^{-1}([x,y;F]-F'(x_0)) \| \leq M_1 (\| x-x_0 \| + \| y-x_0 \|) \quad \text{for all} \quad x,y \in \mathcal{D};$$

($\mathcal{C}_7$) There exist $x_0 \in \mathcal{D}$ and $M_2 > 0$ such that $F'(x_0)^{-1} \in \mathcal{L}(\mathcal{Y},\mathcal{X})$ and

$$\| F'(x_0)^{-1}(F'(x)-F'(x_0)) \| \leq M_2 (\| x-x_0 \| + \| y-x_0 \|) \quad \text{for all} \quad x,y \in \mathcal{D}.$$

Note that $M_0 \leq M$, $M_2 \leq M_1$, $L \leq M$ hold in general and M/M_0, M_1/M_2, M/L can be arbitrarily large [6, 7, 8, 9, 10, 14]. We shall refer to ($\mathcal{C}_1$), ($\mathcal{C}_2$), ($\mathcal{C}_3^{\star\star}$), ($\mathcal{C}_4^{\star}$), ($\mathcal{C}_6$) as the ($\mathcal{C}^{\star}$) conditions and ($\mathcal{C}_1$), ($\mathcal{C}_2$), ($\mathcal{C}_3^{\star}$), ($\mathcal{C}_4^{\star}$), ($\mathcal{C}_5$) as the ($\mathcal{C}^{\star\star}$) conditions. Note that ($\mathcal{C}_5$) is not additional hypothesis to ($\mathcal{C}_3$), since in practice the computation of constant M requires that of M_0. Note that if ($\mathcal{C}_6$) holds, then we can set $M_2 = 2M_1$ in ($\mathcal{C}_7$).

The chapter is organized as follows. In Section 23.2. we use the ($\mathcal{C}^{\star}$) and ($\mathcal{C}^{\star\star}$) conditions instead of the ($\mathcal{C}$) conditions to provide new semilocal convergence analyses for method (23.1.2) under weaker sufficient criteria than those given in [18, 19, 21, 26, 27]. This way we obtain a larger convergence domain and a tighter convergence analysis. Two numerical examples, where we illustrate the improvement of the domain of starting points achieved with the new semilocal convergence results, are given in the Section 23.3..

23.2. Semilocal Convergence of Secant-Like Method

We present the semilocal convergence of secant-like method. First, we need some results on majorizing sequences for secant-like method.

Lemma 23.2.1. *Let $c \geq 0$, $\eta > 0$, $M_1 > 0$, $K > 0$ and $\lambda \in [0,1]$. Set $t_{-1} = 0$, $t_0 = c$ and $t_1 = c+\eta$. Define scalar sequences $\{q_n\}$, $\{t_n\}$, $\{\alpha_n\}$ for each $n = 0,1,\cdots$ by*

$$q_n = (1-\lambda)(t_n - t_0) + (1+\lambda)(t_{n+1} - t_0),$$

$$t_{n+2} = t_{n+1} + \frac{K(t_{n+1} - t_n + (1-\lambda)(t_n - t_{n-1}))}{1 - M_1 q_n}(t_{n+1} - t_n), \tag{23.2.1}$$

$$\alpha_n = \frac{K(t_{n+1} - t_n + (1-\lambda)(t_n - t_{n-1}))}{1 - M_1 q_n}, \tag{23.2.2}$$

function $\{f_n\}$ for each $n = 1,2,\cdots$ by

$$\begin{aligned} f_n(t) \;\; = \;\; & K\eta t^n + K(1-\lambda)\eta t^{n-1} + M_1\eta((1-\lambda)(1+t+\cdots+t^n)+ \\ & (1+\lambda)(1+t+\cdots+t^{n+1})) - 1 \end{aligned} \tag{23.2.3}$$

and polynomial p by

$$p(t) = M_1(1+\lambda)t^3 + (M_1(1-\lambda)+K)t^2 - K\lambda t - K(1-\lambda). \tag{23.2.4}$$

Denote by α *the smallest root of polynomial* p *in* $(0,1)$*. Suppose that*

$$0 < \alpha_0 \leq \alpha \leq 1 - 2M_1\eta. \tag{23.2.5}$$

Then, sequence $\{t_n\}$ *is non-decreasing, bounded from above by* $t^{\star\star}$ *defined by*

$$t^{\star\star} = \frac{\eta}{1-\alpha} + c \tag{23.2.6}$$

and converges to its unique least upper bound $t^{\star}$ *which satisfies*

$$c + \eta \leq t^{\star} \leq t^{\star\star}. \tag{23.2.7}$$

Moreover, the following estimates are satisfied for each $n = 0, 1, \cdots$

$$0 \leq t_{n+1} - t_n \leq \alpha^n \eta \tag{23.2.8}$$

and

$$t^{\star} - t_n \leq \frac{\alpha^n \eta}{1-\alpha}. \tag{23.2.9}$$

Proof. We shall first prove that polynomial p has roots in $(0,1)$. If $\lambda \neq 1$, $p(0) = -(1-\lambda)K < 0$ and $p(1) = 2M_1 > 0$. If $\lambda = 1$, $p(t) = t\overline{p}(t)$, $\overline{p}(0) = -K < 0$ and $\overline{p}(1) = 2M_1 > 0$. In either case it follows from the intermediate value theorem that there exist roots in $(0,1)$. Denote by α the minimal root of p in $(0,1)$. Note that, in particular for Newton's method (i.e. for $\lambda = 1$) and for Secant method (i.e. for $\lambda = 0$), we have, respectively by (23.2.4) that

$$\alpha = \frac{2K}{K + \sqrt{K^2 + 4M_1 K}} \tag{23.2.10}$$

and

$$\alpha = \frac{2K}{K + \sqrt{K^2 + 8M_1 K}}. \tag{23.2.11}$$

It follows from (23.2.1) and (23.2.2) that estimate (23.2.8) is satisfied if

$$0 \leq \alpha_n \leq \alpha. \tag{23.2.12}$$

Estimate (23.2.12) is true by (23.2.5) for $n = 0$. Then, we have by (23.2.1) that

$$\begin{aligned} & t_2 - t_1 \leq \alpha(t_1 - t_0) \Longrightarrow t_2 \leq t_1 + \alpha(t_1 - t_0) \\ & \Longrightarrow t_2 \leq \eta + t_0 + \alpha\eta = c + (1+\alpha)\eta = c + \frac{1-\alpha^2}{1-\alpha\eta} < t^{\star\star}. \end{aligned}$$

Suppose that

$$t_{k+1} - t_k \leq \alpha^k \eta \quad \text{and} \quad t_{k+1} \leq c + \frac{1-\alpha^{k+1}}{1-\alpha}\eta. \tag{23.2.13}$$

Estimate (23.2.12) shall be true for $k+1$ replacing n if

$$0 \leq \alpha_{k+1} \leq \alpha \tag{23.2.14}$$

or

$$f_k(\alpha) \le 0. \tag{23.2.15}$$

We need a relationship between two consecutive recurrent functions f_k for each $k = 1, 2, \cdots$. It follows from (23.2.3) and (23.2.4) that

$$f_{k+1}(\alpha) = f_k(\alpha) + p(\alpha)\,\alpha^{k-1}\eta = f_k(\alpha), \tag{23.2.16}$$

since $p(\alpha) = 0$. Define function f_∞ on $(0,1)$ by

$$f_\infty(t) = \lim_{n\to\infty} f_n(t). \tag{23.2.17}$$

Then, we get from (23.2.3) and (23.2.17) that

$$\begin{aligned} f_\infty(\alpha) &= \lim_{n\to\infty} f_n(\alpha) \\ &= K\eta \lim_{n\to\infty}\alpha^n + K(1-\lambda)\eta \lim_{n\to\infty}\alpha^{n-1} + \\ &\quad M_1\eta\Bigg((1-\lambda)\lim_{n\to\infty}(1+\alpha+\cdots+\alpha^n) + \\ &\quad (1+\lambda)\lim_{n\to\infty}(1+\alpha+\cdots+\alpha^{n+1})\Bigg) - 1 \\ &= M_1\eta\left(\frac{1-\lambda}{1-\alpha} + \frac{1+\lambda}{1-\alpha}\right) - 1 = \frac{2M_1\eta}{1-\alpha} - 1, \end{aligned} \tag{23.2.18}$$

since $\alpha \in (0,1)$. In view of (23.2.15), (23.2.16) and (23.2.18) we can show instead of (23.2.15) that

$$f_\infty(\alpha) \le 0, \tag{23.2.19}$$

which is true by (23.2.5). The induction for (23.2.8) is complete. It follows that sequence $\{t_n\}$ is non-decreasing, bounded from above by $t^{\star\star}$ given by (23.2.6) and as such it converges to $t^\star$ which satisfies (23.2.7). Estimate (23.2.9) follows from (23.2.8) by using standard majorization techniques [8, 10, 22]. The proof of Lemma 23.2.1 is complete. ⊠

Lemma 23.2.2. *Let $c \ge 0$, $\eta > 0$, $M_1 > 0$, $K > 0$ and $\lambda \in [0,1]$. Set $r_{-1} = 0$, $r_0 = c$ and $r_1 = c + \eta$. Define scalar sequences $\{r_n\}$ for each $n = 1, \cdots$ by*

$$\begin{aligned} r_2 &= r_1 + \beta_1\,(r_1 - r_0) \\ r_{n+2} &= r_{n+1} + \beta_n\,(r_{n+1} - r_n), \end{aligned} \tag{23.2.20}$$

where

$$\beta_1 = \frac{M_1\,(r_1 - r_0 + (1-\lambda)\,(r_0 - r_{-1}))}{1 - M_1\,q_1},$$

$$\beta_n = \frac{K\,(r_{n+1} - r_n + (1-\lambda)\,(r_n - r_{n-1}))}{1 - M_1\,q_n} \quad \textit{for each} \quad n = 2, 3, \cdots$$

and function $\{g_n\}$ on $[0,1)$ for each $n = 1, 2, \cdots$ by

$$\begin{aligned} g_n(t) &= K\,(t + (1-\lambda))\,t^{n-1}\,(r_2 - r_1) + \\ &\quad M_1\,t\left((1-\lambda)\,\frac{1-t^{n+1}}{1-t} + (1+\lambda)\,\frac{1-t^{n+2}}{1-t}\right)(r_2 - r_1) + (2M_1\eta - 1)\,t. \end{aligned} \tag{23.2.21}$$

Suppose that

$$0 \leq \beta_1 \leq \alpha \leq 1 - \frac{2M_1 (r_2 - r_1)}{1 - 2M_1 \eta}, \tag{23.2.22}$$

where α *is defined in Lemma 23.2.1. Then, sequence* $\{r_n\}$ *is non-decreasing, bounded from above by* $r^{\star\star}$ *defined by*

$$r^{\star\star} = c + \eta + \frac{r_2 - r_1}{1 - \alpha} \tag{23.2.23}$$

and converges to its unique least upper bound $r^\star$ *which satisfies*

$$c + \eta \leq r^\star \leq r^{\star\star}. \tag{23.2.24}$$

Moreover, the following estimates are satisfied for each $n = 1, \cdots$

$$0 \leq r_{n+2} - r_{n+1} \leq \alpha^n (r_2 - r_1). \tag{23.2.25}$$

Proof. We shall use mathematical induction to show that

$$0 \leq \beta_n \leq \alpha. \tag{23.2.26}$$

Estimate (23.2.26) is true for $n = 0$ by (23.2.22). Then, we have by (23.2.20) that

$$\begin{aligned} 0 \leq r_3 - r_2 \leq \alpha (r_2 - r_1) \quad &\Longrightarrow \quad r_3 \leq r_2 + \alpha (r_2 - r_1) \\ &\Longrightarrow \quad r_3 \leq r_2 + (1 + \alpha)(r_2 - r_1) - (r_2 - r_1) \\ &\Longrightarrow \quad r_3 \leq r_1 + \frac{1 - \alpha^2}{1 - \alpha} (r_2 - r_1) \leq r^{\star\star}. \end{aligned}$$

Suppose (23.2.26) holds for each $n \leq k$, then, using (23.2.20), we obtain that

$$0 \leq r_{k+2} - r_{k+1} \leq \alpha^k (r_2 - r_1) \quad \text{and} \quad r_{k+2} \leq r_1 + \frac{1 - \alpha^{k+1}}{1 - \alpha} (r_2 - r_1). \tag{23.2.27}$$

Estimate (23.2.26) is certainly satisfied, if

$$g_k(\alpha) \leq 0, \tag{23.2.28}$$

where g_k is defined by (23.2.21). Using (23.2.21), we obtain the following relationship between two consecutive recurrent functions g_k for each $k = 1, 2, \cdots$

$$g_{k+1}(\alpha) = g_k(\alpha) + p(\alpha)\, \alpha^{k-1} (r_2 - r_1) = g_k(\alpha), \tag{23.2.29}$$

since $p(\alpha) = 0$. Define function g_∞ on $[0, 1)$ by

$$g_\infty(t) = \lim_{k \to \infty} g_k(t). \tag{23.2.30}$$

Then, we get from (23.2.21) and (23.2.30) that

$$g_\infty(\alpha) = \alpha \left(\frac{2M_1 (r_2 - r_1)}{1 - \alpha} + 2M_1 \eta - 1 \right). \tag{23.2.31}$$

In view of (23.2.28)–(23.2.31) to show (23.2.28), it suffices to have $g_\infty(\alpha) \leq 0$, which true by the right hand hypothesis in (23.2.22). The induction for (23.2.26) (i.e. for (23.2.25)) is complete. The rest of the proof is omitted (as identical to the proof of Lemma 23.2.1). The proof of Lemma 23.2.2 is complete. ⊠

Remark 23.2.3. *Let us see how sufficient convergence criterion on (23.2.5) for sequence $\{t_n\}$ simplifies in the interesting case of Newton's method. That is when $c = 0$ and $\lambda = 1$. Then, (23.2.5) can be written for $L_0 = 2M_1$ and $L = 2K$ as*

$$h_0 = \frac{1}{8}(L + 4L_0 + \sqrt{L^2 + 8L_0 L})\eta \leq \frac{1}{2}. \tag{23.2.32}$$

The convergence criterion in [18] reduces to the famous for it simplicity and clarity Kantorovich hypothesis

$$h = L\eta \leq \frac{1}{2}. \tag{23.2.33}$$

Note however that $L_0 \leq L$ holds in general and L/L_0 can be arbitrarily large [6, 7, 8, 9, 10, 14]. We also have that

$$h \leq \frac{1}{2} \Longrightarrow h_0 \leq \frac{1}{2} \tag{23.2.34}$$

but not necessarily vice versa unless if $L_0 = L$ and

$$\frac{h_0}{h} \longrightarrow \frac{1}{4} \quad \text{as} \quad \frac{L}{L_0} \longrightarrow \infty. \tag{23.2.35}$$

Similarly, it can easily be seen that the sufficient convergence criterion (23.2.22) for sequence $\{r_n\}$ is given by

$$h_1 = \frac{1}{8}(4L_0 + \sqrt{L_0 L + 8L_0^2} + \sqrt{L_0 L})\eta \leq \frac{1}{2}. \tag{23.2.36}$$

We also have that

$$h_0 \leq \frac{1}{2} \Longrightarrow h_1 \leq \frac{1}{2} \tag{23.2.37}$$

and

$$\frac{h_1}{h} \longrightarrow 0, \ \frac{h_1}{h_0} \longrightarrow 0 \quad \text{as} \quad \frac{L_0}{L} \longrightarrow 0. \tag{23.2.38}$$

Note that sequence $\{r_n\}$ is tighter than $\{t_n\}$ and converges under weaker conditions. Indeed, a simple inductive argument shows that for each $n = 2, 3, \cdots$, if $M_1 < K$, then

$$r_n < t_n, \ r_{n+1} - r_n < t_{n+1} - t_n \quad \text{and} \quad r^\star \leq t^\star. \tag{23.2.39}$$

We have the following usefull and obvious extensions of Lemma 23.2.1 and Lemma 23.2.2, respectively.

Lemma 23.2.4. *Let $N = 0, 1, 2, \cdots$ be fixed. Suppose that*

$$t_1 \leq t_2 \leq \cdots \leq t_N \leq t_{N+1}, \tag{23.2.40}$$

$$\frac{1}{M_1} > (1 - \lambda)(t_N - t_0) + (1 + \lambda)(t_{N+1} - t_0) \tag{23.2.41}$$

and

$$0 \leq \alpha_N \leq \alpha \leq 1 - 2M_1(t_{N+1} - t_N). \tag{23.2.42}$$

Then, sequence $\{t_n\}$ generated by (23.2.1) is nondecreasing, bounded from above by $t^{\star\star}$ and converges to $t^\star$ which satisfies $t^\star \in [t_{N+1}, t^{\star\star}]$. Moreover, the following estimates are satisfied for each $n = 0, 1, \cdots$

$$0 \leq t_{N+n+1} - t_{N+n} \leq \alpha^n (t_{N+1} - t_N) \tag{23.2.43}$$

and

$$t^\star - t_{N+n} \leq \frac{\alpha^n}{1-\alpha} (t_{N+1} - t_N). \tag{23.2.44}$$

Lemma 23.2.5. *Let $N = 1, 2, \cdots$ be fixed. Suppose that*

$$r_1 \leq r_2 \leq \cdots \leq r_N \leq r_{N+1}, \tag{23.2.45}$$

$$\frac{1}{M_1} > (1-\lambda)(r_N - r_0) + (1+\lambda)(r_{N+1} - r_0) \tag{23.2.46}$$

and

$$0 \leq \beta_N \leq \alpha \leq 1 - \frac{2M_1 (r_{N+1} - r_N)}{1 - 2M_1 (r_N - r_{N-1})}. \tag{23.2.47}$$

Then, sequence $\{r_n\}$ generated by (23.2.20) is nondecreasing, bounded from above by $r^{\star\star}$ and converges to $r^\star$ which satisfies $r^\star \in [r_{N+1}, r^{\star\star}]$. Moreover, the following estimates are satisfied for each $n = 0, 1, \cdots$

$$0 \leq r_{N+n+1} - r_{N+n} \leq \alpha^n (r_{N+1} - r_N) \tag{23.2.48}$$

and

$$r^\star - r_{N+n} \leq \frac{\alpha^n}{1-\alpha} (r_{N+1} - r_N). \tag{23.2.49}$$

Next, we present the following semilocal convergence result for secant-like method under the **($\mathcal{C}^\star$)** conditions.

Theorem 23.2.6. *Suppose that the **($\mathcal{C}^\star$)**, Lemma 23.2.1 (or Lemma 23.2.4) conditions and*

$$\overline{U}(x_0, t^\star) \subseteq \mathcal{D} \tag{23.2.50}$$

hold. Then, sequence $\{x_n\}$ generated by the secant-like method is well defined, remains in $\overline{U}(x_0, t^\star)$ for each $n = -1, 0, 1, \cdots$ and converges to a solution $x^\star \in \overline{U}(x_0, t^\star - c)$ of equation $F(x) = 0$. Moreover, the following estimates are satisfied for each $n = 0, 1, \cdots$

$$\| x_{n+1} - x_n \| \leq t_{n+1} - t_n \tag{23.2.51}$$

and

$$\| x_n - x^\star \| \leq t^\star - t_n. \tag{23.2.52}$$

Furthemore, if there exists $r \geq t^\star$ such that

$$\overline{U}(x_0, r) \subseteq \mathcal{D} \tag{23.2.53}$$

and

$$r + t^\star < \frac{1}{M_1} \quad \text{or} \quad r + t^\star < \frac{2}{M_2}, \tag{23.2.54}$$

then, the solution $x^\star$ is unique in $\overline{U}(x_0, r)$.

Proof. We use mathematical induction to prove that

$$\| x_{k+1} - x_k \| \leq t_{k+1} - t_k \tag{23.2.55}$$

and

$$\overline{U}(x_{k+1}, t^\star - t_{k+1}) \subseteq \overline{U}(x_k, t^\star - t_k) \tag{23.2.56}$$

for each $k = -1, 0, 1, \cdots$. Let $z \in \overline{U}(x_0, t^\star - t_0)$. Then, we obtain that

$$\| z - x_{-1} \| \leq \| z - x_0 \| + \| x_0 - x_{-1} \| \leq t^\star - t_0 + c = t^\star = t^\star - t_{-1},$$

which implies $z \in \overline{U}(x_{-1}, t^\star - t_{-1})$. Let also $w \in \overline{U}(x_0, t^\star - t_1)$. We get that

$$\| w - x_0 \| \leq \| w - x_1 \| + \| x_1 - x_0 \| \leq t^\star - t_1 + t_1 - t_0 = t^\star = t^\star - t_0.$$

That is $w \in \overline{U}(x_0, t^\star - t_0)$. Note that

$$\| x_{-1} - x_0 \| \leq c = t_0 - t_{-1} \quad \text{and} \quad \| x_1 - x_0 \| = \| B_0^{-1} F(x_0) \| \leq \eta = t_1 - t_0 < t^\star,$$

which implies $x_1 \in \overline{U}(x_0, t^\star) \subseteq \mathcal{D}$. Hence, estimates (23.2.51) and (23.2.52) hold for $k = -1$ and $k = 0$. Suppose (23.2.51) and (23.2.52) hold for all $n \leq k$. Then, we obtain that

$$\| x_{k+1} - x_0 \| \leq \sum_{i=1}^{k+1} \| x_i - x_{i-1} \| \leq \sum_{i=1}^{k+1} (t_i - t_{i-1}) = t_{k+1} - t_0 \leq t^\star$$

and

$$\| y_k - x_0 \| \leq \lambda \| x_k - x_0 \| + (1-\lambda) \| x_{k-1} - x_0 \| \leq \lambda t^\star + (1-\lambda) t^\star = t^\star.$$

Hence, $x_{k+1}, y_k \in \overline{U}(x_0, t^\star)$. Let $E_k := [x_{k+1}, x_k; F]$ for each $k = 0, 1, \cdots$. Using (23.1.2), Lemma 23.2.1 and the induction hypotheses, we get that

$$\begin{aligned} &\| F'(x_0)^{-1} (B_{k+1} - F'(x_0)) \| \leq M_1 (\| y_{k+1} - x_0 \| + \| x_{k+1} - x_0 \|) \\ &\leq M_1 ((1-\lambda) \| x_k - x_0 \| + \lambda \| x_{k+1} - x_0 \| + \| x_{k+1} - x_0 \|) \\ &\leq M_1 ((1-\lambda)(t_k - t_0) + (1+\lambda)(t_{k+1} - t_0)) < 1, \end{aligned} \tag{23.2.57}$$

since, $y_{k+1} - x_0 = \lambda (x_{k+1} - x_0) + (1-\lambda)(x_k - x_0)$ and

$$\begin{aligned} &\| y_{k+1} - x_0 \| = \| \lambda (x_{k+1} - x_0) + (1-\lambda)(x_k - x_0) \| \\ &\leq \lambda \| x_{k+1} - x_0 \| + (1-\lambda) \| x_k - x_0 \| . \end{aligned}$$

It follows from (23.2.57) and the Banach lemma on invertible operators that B_{k+1}^{-1} exists and

$$\| B_{k+1}^{-1} F'(x_0) \| \leq \frac{1}{1 - \Theta_k} \leq \frac{1}{1 - M_1 q_{k+1}}, \tag{23.2.58}$$

where $\Theta_k = M_1 ((1-\lambda) \| x_k - x_0 \| + (1+\lambda) \| x_{k+1} - x_0 \|)$. In view of (23.1.2), we obtain the identity

$$F(x_{k+1}) = F(x_{k+1}) - F(x_k) - B_k (x_{k+1} - x_k) = (E_k - B_k)(x_{k+1} - x_k). \tag{23.2.59}$$

Then, using the induction hypotheses, the $(\mathcal{C}^\star)$ condition and (23.2.59), we get in turn that

$$\begin{aligned}\| F'(x_0)^{-1} F(x_{k+1}) \| &= \| F'(x_0)^{-1} (E_k - B_k)(x_{k+1} - x_k) \| \\ &\leq K \| x_{k+1} - y_k \| \| x_{k+1} - x_k \| \\ &\leq K (\| x_{k+1} - x_k \| + (1-\lambda) \| x_k - x_{k-1} \|) \| x_{k+1} - x_k \| \\ &\leq K (t_{k+1} - t_k + (1-\lambda)(t_k - t_{k-1}))(t_{k+1} - t_k),\end{aligned} \tag{23.2.60}$$

since, $x_{k+1} - y_k = x_{k+1} - x_k + (1-\lambda)(x_k - x_{k-1})$ and

$$\| x_{k+1} - y_k \| \leq \| x_{k+1} - x_k \| + (1-\lambda) \| x_k - x_{k-1} \| \leq t_{k+1} - t_k + (1-\lambda)(t_k - t_{k-1}).$$

It now follows from (23.1.2), (23.2.1), (23.2.58)–(23.2.60) that

$$\begin{aligned}\| x_{k+2} - x_{k+1} \| &\leq \| B_{k+1}^{-1} F'(x_0) \| \| F'(x_0)^{-1} F(x_{k+1}) \| \\ &\leq \frac{K (t_{k+1} - t_k + (1-\lambda)(t_{k+1} - x_k))(t_{k+1} - t_k)}{1 - M_1 q_{k+1}} = t_{k+2} - t_{k+1},\end{aligned}$$

which completes the induction for (23.2.55). Furthemore, let $v \in \overline{U}(x_{k+2}, t^\star - t_{k+2})$. Then, we have that

$$\begin{aligned}\| v - x_{k+1} \| &\leq \| v - x_{k+2} \| + \| x_{k+2} - x_{k+1} \| \\ &\leq t^\star - t_{k+2} + t_{k+2} - t_{k+1} = t^\star - t_{k+1},\end{aligned}$$

which implies $v \in \overline{U}(x_{k+1}, t^\star - t_{k+1})$. The induction for (23.2.55) and (23.2.56) is complete. Lemma 23.2.1 implies that $\{t_k\}$ is a complete sequence. It follows from (23.2.55) and (23.2.56) that $\{x_k\}$ is a complete sequence in a Banach space $\mathcal{X}$ and as such it converges to some $x^\star \in \overline{U}(x_0, t^\star)$ (since $\overline{U}(x_0, t^\star)$ is a closed set). By letting $k \longrightarrow \infty$ in (23.2.60), we get that $F(x^\star) = 0$. Moreover, estimate (23.2.52) follows from (23.2.51) by using standard majorization techniques [8, 10, 22]. To show the uniqueness part, let $y^\star \in \overline{U}(x_0, r)$ be such $F(y^\star) = 0$, where r satisfies (23.2.53) and (23.2.54). We have that

$$\begin{aligned}\| F'(x_0)^{-1} ([y^\star, x^\star; F] - F'(x_0)) \| &\leq M_1 (\| y^\star - x_0 \| + \| x^\star - x_0 \|) \\ &\leq M_1 (t^\star + r) < 1.\end{aligned} \tag{23.2.61}$$

It follows by (23.2.61) and the Banach lemma on invertible operators that linear operator $[y^\star, x^\star; F]^{-1}$ exists. Then, using the identity $0 = F(y^\star) - F(x^\star) = [y^\star, x^\star; F](y^\star - x^\star)$, we deduce that $x^\star = y^\star$. The proof of Theorem 23.2.6 is complete. ⊠

In order for us to present the semilocal result for secant-like method under the $(\mathcal{C}^{\star\star})$ conditions, we first need a result on a majorizing sequence. The proof in given in Lemma 23.2.1.

Remark 23.2.7. *Clearly, (23.2.22) (or (23.2.47)), $\{r_n\}$ can replace (23.2.5) (or (23.2.42)), $\{t_n\}$, respectively in Theorem 23.2.6.*

Lemma 23.2.8. *Let $c \geq 0$, $\eta > 0$, $L > 0$, $M_0 > 0$ with $M_0 c < 1$ and $\lambda \in [0,1]$. Set*

$$s_{-1} = 0,\ s_0 = c,\ s_1 = c + \eta,\ \tilde{K} = \frac{L}{1 - M_0 c} \quad \text{and} \quad \tilde{M}_1 = \frac{M_0}{1 - M_0 c}.$$

Define scalar sequences $\{\tilde{q}_n\}$, $\{s_n\}$, $\{\tilde{\alpha}_n\}$ *for each* $n = 0, 1, \cdots$ *by*

$$\tilde{q}_n = (1-\lambda)(s_n - s_0) + (1+\lambda)(s_{n+1} - s_0),$$

$$s_{n+2} = s_{n+1} + \frac{\tilde{K}(s_{n+1} - s_n + (1-\lambda)(s_n - s_{n-1}))}{1 - \tilde{M}_1 \tilde{q}_n}(s_{n+1} - s_n),$$

$$\tilde{\alpha}_n = \frac{\tilde{K}(s_{n+1} - s_n + (1-\lambda)(s_n - s_{n-1}))}{1 - \tilde{M}_1 \tilde{q}_n},$$

function $\{\tilde{f}_n\}$ *for each* $n = 1, 2, \cdots$ *by*

$$\begin{aligned}\tilde{f}_n(t) \quad = \quad & \tilde{K}\eta t^n + \tilde{K}(1-\lambda)\eta t^{n-1} + \tilde{M}_1 \eta ((1-\lambda)(1 + t + \cdots + t^n) + \\ & (1+\lambda)(1 + t + \cdots + t^{n+1})) - 1\end{aligned}$$

and polynomial $\tilde{p}$ *by*

$$\tilde{p}(t) = \tilde{M}_1 (1+\lambda) t^3 + (\tilde{M}_1 (1-\lambda) + \tilde{K}) t^2 - \tilde{K}\lambda t - \tilde{K}(1-\lambda).$$

Denote by $\tilde{\alpha}$ *the smallest root of polynomial* $\tilde{p}$ *in* $(0, 1)$. *Suppose that*

$$0 \leq \tilde{\alpha}_0 \leq \tilde{\alpha} \leq 1 - 2\tilde{M}_1 \eta. \tag{23.2.62}$$

Then, sequence $\{s_n\}$ *is non-decreasing, bounded from above by* $s^{\star\star}$ *defined by*

$$s^{\star\star} = \frac{\eta}{1 - \tilde{\alpha}} + c$$

and converges to its unique least upper bound $s^\star$ *which satisfies* $c + \eta \leq s^\star \leq s^{\star\star}$. *Moreover, the following estimates are satisfied for each* $n = 0, 1, \cdots$

$$0 \leq s_{n+1} - s_n \leq \tilde{\alpha}^n \eta \quad \text{and} \quad s^\star - s_n \leq \frac{\tilde{\alpha}^n \eta}{1 - \tilde{\alpha}}.$$

Next, we present the semilocal convergence result for secant-like method under the $(\mathcal{C}^{\star\star})$ conditions.

Theorem 23.2.9. *Suppose that the* $(\mathcal{C}^{\star\star})$ *conditions, (23.2.62) (or Lemma 23.2.2 conditions with* $\tilde{\alpha}_n$, $\tilde{\alpha}$, $\tilde{M}_1$ *replacing, respectively,* α_n, α, M_1*) and* $\overline{U}(x_0, s^\star) \subseteq \mathcal{D}$ *hold. Then, sequence* $\{x_n\}$ *generated by the secant-like method is well defined, remains in* $\overline{U}(x_0, s^\star)$ *for each* $n = -1, 0, 1, \cdots$ *and converges to a solution* $x^\star \in \overline{U}(x_0, s^\star)$ *of equation* $F(x) = 0$. *Moreover, the following estimates are satisfied for each* $n = 0, 1, \cdots$

$$\| x_{n+1} - x_n \| \leq s_{n+1} - s_n \quad \text{and} \quad \| x_n - x^\star \| \leq s^\star - s_n.$$

Furthemore, if there exists $r \geq s^\star$ *such that* $\overline{U}(x_0, r) \subseteq \mathcal{D}$ *and* $r + s^\star + c < 1/M_0$, *then, the solution* $x^\star$ *is unique in* $\overline{U}(x_0, r)$.

Proof. The proof is analogous to Theorem 23.2.6. Simply notice that in view of ($\mathcal{C}_5$), we obtain instead of (23.2.57) that

$$\begin{aligned}&\parallel A_0^{-1}(B_{k+1}-A_0)\parallel\leq M_0(\parallel y_{k+1}-x_{-1}\parallel+\parallel x_{k+1}-x_0\parallel)\\&\leq M_0((1-\lambda)\parallel x_k-x_0\parallel+\lambda\parallel x_{k+1}-x_0\parallel+\parallel x_0-x_{-1}\parallel+\parallel x_{k+1}-x_0\parallel)\\&\leq M_0((1-\lambda)(s_k-s_0)+(1+\lambda)(s_{k+1}-s_0)+c)<1,\end{aligned}$$

leading to B_{k+1}^{-1} exists and

$$\parallel B_{k+1}^{-1}A_0\parallel\leq\frac{1}{1-\Xi_k},$$

where $\Xi_k=M_0((1-\lambda)(s_k-s_0)+(1+\lambda)(s_{k+1}-s_0)+c)$. Moreover, using ($\mathcal{C}_3^\star$) instead of ($\mathcal{C}_3^{\star\star}$), we get that

$$\parallel A_0^{-1}F(x_{k+1})\parallel\leq L(s_{k+1}-s_k+(1-\lambda)(s_k-s_{k-1}))(s_{k+1}-s_k).$$

Hence, we have that

$$\begin{aligned}&\parallel x_{k+2}-x_{k+1}\parallel\leq\parallel B_{k+1}^{-1}A_0\parallel\parallel A_0^{-1}F(x_{k+1})\parallel\\&\leq\frac{L(s_{k+1}-s_k+(1-\lambda)(s_k-s_{k-1}))(s_{k+1}-s_k)}{1-M_0((1+\lambda)(s_{k+1}-s_0)+(1-\lambda)(s_k-s_0)+c)}\\&\leq\frac{\tilde{K}(s_{k+1}-s_k+(1-\lambda)(s_k-s_{k-1}))(s_{k+1}-s_k)}{1-\tilde{M}_1((1+\lambda)(s_{k+1}-s_0)+(1-\lambda)(s_k-s_0))}=s_{k+2}-s_{k+1}.\end{aligned}$$

The uniqueness part is given in Theorem 23.2.6 with r, $s^\star$ replacing R_2 and R_0, respectively. The proof of Theorem 23.2.9 is complete. ⊠

Remark 23.2.10. *(a) Condition (23.2.50) can be replaced by*

$$\overline{U}(x_0,t^{\star\star})\subseteq\mathcal{D},\tag{23.2.63}$$

where $t^{\star\star}$ is given in the closed form by (23.2.55).

(b) The majorizing sequence $\{u_n\}$ essentially used in [18] is defined by

$$\begin{aligned}&u_{-1}=0,\ u_0=c,\ u_1=c+\eta\\&u_{n+2}=u_{n+1}+\frac{M(u_{n+1}-u_n+(1-\lambda)(u_n-u_{n-1}))}{1-M\,q_n^\star}(u_{n+1}-u_n),\end{aligned}\tag{23.2.64}$$

where

$$q_n^\star=(1-\lambda)(u_n-u_0)+(1+\lambda)(u_{n+1}-u_0).$$

Then, if $K<M$ or $M_1<M$, a simple inductive argument shows that for each $n=2,3,\cdots$

$$t_n<u_n,\quad t_{n+1}-t_n<u_{n+1}-u_n\quad\text{and}\quad t^\star\leq u^\star=\lim_{n\to\infty}u_n.\tag{23.2.65}$$

Clearly $\{t_n\}$ converges under the ($\mathcal{C}$) conditions and conditions of Lemma 2.1. Moreover, as we already showed in Remark 23.2.3, the sufficient convergence criteria of Theorem 23.2.6 can be weaker than those of Theorem 23.2.9. Similarly if $L\leq M$, $\{s_n\}$ is a tighter sequence than $\{u_n\}$. In general, we shall test the convergence criteria and use the tightest sequence to estimate the error bounds.

(c) *Clearly the conclusions of Theorem 23.2.9 hold if* $\{s_n\}$*, (23.2.62) are replaced by* $\{\tilde{r}_n\}$*, (23.2.22), where* $\{\tilde{r}_n\}$ *is defined as* $\{r_n\}$ *with* M_0 *replacing* M_1 *in the definition of* β_1 *(only at the numerator) and the tilda letters replacing the non-tilda letters in (23.2.22).*

23.3. Numerical Examples

Now, we check numerically with two examples that the new semilocal convergence results obtained in Theorems 23.2.6 and 23.2.9 improve the domain of starting points obtained by the following classical result given in [20].

Theorem 23.3.1. *Let* X *and* Y *be two Banach spaces and* $F:\Omega\subseteq X\to Y$ *be a nonlinear operator defined on a non-empty open convex domain* Ω*. Let* $x_{-1},x_0\in\Omega$ *and* $\lambda\in[0,1]$*. Suppose that there exists* $[u,v;F]\in\mathcal{L}(X,Y)$*, for all* $u,v\in\Omega$ $(u\neq v)$*, and the following four conditions*

- $\|x_0-x_{-1}\|=c\neq 0$ *with* $x_{-1},x_0\in\Omega$,
- *Fixed* $\lambda\in[0,1]$*, the operator* $B_0=[y_0,x_0;F]$ *is invertible and such that* $\|B_0^{-1}\|\leq\beta$,
- $\|B_0^{-1}F(x_0)\|\leq\eta$,
- $\|[x,y;F]-[u,v;F]\|\leq Q(\|x-u\|+\|y-v\|);Q\geq 0;x,y,u,v\in\Omega;x\neq y;u\neq v$,

are satisfied. If $B(x_0,\rho)\subseteq\Omega$*, where* $\rho=\dfrac{1-a}{1-2a}\eta$,

$$a=\frac{\eta}{c+\eta}<\frac{3-\sqrt{5}}{2}\qquad\text{and}\qquad b=\frac{Q\beta c^2}{c+\eta}<\frac{a(1-a)^2}{1+\lambda(2a-1)},\tag{23.3.1}$$

then the secant-like methods defined by (23.1.2) converge to a solution x^* *of equation* $F(x)=0$ *with R-order of convergence at least* $\frac{1+\sqrt{5}}{2}$*. Moreover,* $x_n,x^*\in\overline{B(x_0,\rho)}$*, the solution* x^* *is unique in* $B(x_0,\tau)\cap\Omega$*, where* $\tau=\frac{1}{Q\beta}-\rho-(1-\lambda)\alpha$*.*

23.3.1. Example 1

We illustrate the above-mentioned with an application, where a system of nonlinear equations is involved. We see that Theorem 23.3.1 cannot guarantee the semilocal convergence of secant-like methods (23.1.2) , but Theorem 23.2.6 can do it.

It is well known that energy is dissipated in the action of any real dynamical system, usually through some form of friction. However, in certain situations this dissipation is so slow that it can be neglected over relatively short periods of time. In such cases we assume the law of conservation of energy, namely, that the sum of the kinetic energy and the potential energy is constant. A system of this kind is said to be conservative.

If φ and ψ are arbitrary functions with the property that $\varphi(0)=0$ and $\psi(0)=0$, the general equation

$$\mu\frac{d^2x(t)}{dt^2}+\psi\left(\frac{dx(t)}{dt}\right)+\varphi(x(t))=0,\tag{23.3.2}$$

can be interpreted as the equation of motion of a mass μ under the action of a restoring force $-\varphi(x)$ and a damping force $-\psi(dx/dt)$. In general these forces are nonlinear, and equation (23.3.2) can be regarded as the basic equation of nonlinear mechanics. In this chapter we shall consider the special case of a nonlinear conservative system described by the equation

$$\mu \frac{d^2x(t)}{dt^2} + \varphi(x(t)) = 0,$$

in which the damping force is zero and there is consequently no dissipation of energy. Extensive discussions of (23.3.2), with applications to a variety of physical problems, can be found in classical references [4] and [31].

Now, we consider the special case of a nonlinear conservative system described by the equation

$$\frac{d^2x(t)}{dt^2} + \phi(x(t)) = 0 \tag{23.3.3}$$

with the boundary conditions

$$x(0) = x(1) = 0. \tag{23.3.4}$$

After that, we use a process of discretization to transform problem (23.3.3)–(23.3.4) into a finite-dimensional problem and look for an approximated solution of it when a particular function ϕ is considered. So, we transform problem (23.3.3)–(23.3.4) into a system of nonlinear equations by approximating the second derivative by a standard numerical formula.

Firstly, we introduce the points $t_j = jh$, $j = 0, 1, \ldots, m+1$, where $h = \frac{1}{m+1}$ and m is an appropriate integer. A scheme is then designed for the determination of numbers x_j, it is hoped, approximate the values $x(t_j)$ of the true solution at the points t_j. A standard approximation for the second derivative at these points is

$$x_j'' \approx \frac{x_{j-1} - 2x_j + x_{j+1}}{h^2}, \quad j = 1, 2, \ldots, m.$$

A natural way to obtain such a scheme is to demand that the x_j satisfy at each interior mesh point t_j the difference equation

$$x_{j-1} - 2x_j + x_{j+1} + h^2\phi(x_j) = 0. \tag{23.3.5}$$

Since x_0 and x_{m+1} are determined by the boundary conditions, the unknowns are $x_1, x_2, \ldots, x_m$.

A further discussion is simplified by the use of matrix and vector notation. Introducing the vectors

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}, \qquad v_{\mathbf{x}} = \begin{pmatrix} \phi(x_1) \\ \phi(x_2) \\ \vdots \\ \phi(x_m) \end{pmatrix}$$

and the matrix

$$A = \begin{pmatrix} -2 & 1 & 0 & \cdots & 0 \\ 1 & -2 & 1 & \cdots & 0 \\ 0 & 1 & -2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -2 \end{pmatrix},$$

the system of equations, arising from demanding that (23.3.5) holds for $j = 1, 2, \ldots, m$, can be written compactly in the form

$$F(\mathbf{x}) \equiv A\mathbf{x} + h^2 v_{\mathbf{x}} = 0, \tag{23.3.6}$$

where F is a function from $\mathbb{R}^m$ into $\mathbb{R}^m$.

From now on, the focus of our attention is to solve a particular system of form (23.3.6). We choose $m = 8$ and the infinity norm.

The steady temperature distribution is known in a homogeneous rod of length 1 in which, as a consequence of a chemical reaction or some such heat-producing process, heat is generated at a rate $\phi(x(t))$ per unit time per unit length, $\phi(x(t))$ being a given function of the excess temperature x of the rod over the temperature of the surroundings. If the ends of the rod, $t = 0$ and $t = 1$, are kept at given temperatures, we are to solve the boundary value problem given by (23.3.3)–(23.3.4), measured along the axis of the rod. For an example we choose an exponential law $\phi(x(t)) = \exp(x(t))$ for the heat generation.

Taking into account that the solution of (23.3.3)–(23.3.4) with $\phi(x(t)) = \exp(x(t))$ is of the form

$$x(s) = \int_0^1 G(s,t)\exp(x(t))\,dt,$$

where $G(s,t)$ is the Green function in $[0,1]\times[0,1]$, we can locate the solution $x^*(s)$ in some domain. So, we have

$$\|x^*(s)\| - \frac{1}{8}\exp(\|x^*(s)\|) \le 0,$$

so that $\|x^*(s)\| \in [0,\rho_1]\cup[\rho_2,+\infty]$, where $\rho_1 = 0.1444$ and $\rho_2 = 3.2616$ are the two positive real roots of the scalar equation $8t - \exp(t) = 0$.

Observing the semilocal convergence results presented in this chapter, we can only guarantee the semilocal convergence to a solution $x^*(s)$ such that $\|x^*(s)\| \in [0,\rho_1]$. For this, we can consider the domain

$$\Omega = \{x(s) \in C^2[0,1];\ \|x(s)\| < \log(7/4),\ s \in [0,1]\},$$

since $\rho_1 < \log\left(\frac{7}{4}\right) < \rho_2$.

In view of what the domain Ω is for equation (23.3.3), we then consider (23.3.6) with $F : \widetilde{\Omega} \subset \mathbb{R}^8 \to \mathbb{R}^8$ and

$$\widetilde{\Omega} = \{\mathbf{x} \in \mathbb{R}^8;\ \|\mathbf{x}\| < \log(7/4)\}.$$

According to the above-mentioned, $v_{\mathbf{x}} = (\exp(x_1), \exp(x_2), \ldots, \exp(x_8))^t$ if $\phi(x(t)) = \exp(x(t))$. Consequently, the first derivative of the function F defined in (23.3.6) is given by

$$F'(\mathbf{x}) = A + h^2 \operatorname{diag}(v_{\mathbf{x}}).$$

Moreover,

$$F'(\mathbf{x}) - F'(\mathbf{y}) = h^2 \operatorname{diag}(\mathbf{z}),$$

where $\mathbf{y} = (y_1, y_2, \ldots, y_8)^t$ and $\mathbf{z} = (\exp(x_1) - \exp(y_1), \exp(x_2) - \exp(y_2), \ldots, \exp(x_8) - \exp(y_8))$. In addition,

$$\|F'(\mathbf{x}) - F'(\mathbf{y})\| \le h^2 \max_{1\le i\le 8} |\exp(\ell_i)|\, \|\mathbf{x} - \mathbf{y}\|,$$

where $\ell = (\ell_1, \ell_2, \ldots, \ell_8)^t \in \widetilde{\Omega}$ and $h = \frac{1}{9}$, so that

$$\|F'(\mathbf{x}) - F'(\mathbf{y})\| \leq \frac{7}{4} h^2 \|\mathbf{x} - \mathbf{y}\|. \tag{23.3.7}$$

Considering (see [27])

$$[\mathbf{x}, \mathbf{y}; F] = \int_0^1 F'(\tau\mathbf{x} + (1-\tau)\mathbf{y})\, d\tau,$$

taking into account

$$\int_0^1 \|\tau(\mathbf{x} - \mathbf{u}) + (1-\tau)(\mathbf{y} - \mathbf{v})\| d\tau \leq \frac{1}{2}\left(\|\mathbf{x} - \mathbf{u}\| + \|\mathbf{y} - \mathbf{v}\|\right),$$

and (23.3.7), we have

$$\begin{aligned}
\|[\mathbf{x}, \mathbf{y}; F] - [\mathbf{u}, \mathbf{v}; F]\| &\leq \int_0^1 \|F'(\tau\mathbf{x} + (1-\tau)\mathbf{y}) - F'(\tau\mathbf{u} + (1-\tau)\mathbf{v})\| \, d\tau \\
&\leq \frac{7}{4} h^2 \int_0^1 (\tau\|\mathbf{x} - \mathbf{u}\| + (1-\tau)\|\mathbf{y} - \mathbf{v}\|)\, d\tau \\
&= \frac{7}{8} h^2 \left(\|\mathbf{x} - \mathbf{u}\| + \|\mathbf{y} - \mathbf{v}\|\right).
\end{aligned}$$

From the last, we have $L = \frac{7}{648}$ and $M_1 = \frac{7}{648}\|[F'(x_0)]^{-1}\|$.

If we choose $\lambda = \frac{1}{2}$ and the starting points $\mathbf{x_{-1}} = (\frac{1}{10}, \frac{1}{10}, \ldots, \frac{1}{10})^t$ and $\mathbf{x_0} = (0, 0, \ldots, 0)^t$, we obtain $c = \frac{1}{10}$, $\beta = 11.202658\ldots$ and $\eta = 0.138304\ldots$, so that (23.3.1) of Theorem 23.3.1 is not satisfied, since

$$a = \frac{\eta}{c + \eta} = 0.580368\ldots > \frac{3 - \sqrt{5}}{2} = 0.381966\ldots$$

Thus, according to Theorem 23.3.1, we cannot guarantee the convergence of secant-like method (23.1.2) with $\lambda = \frac{1}{2}$ for approximating a solution of (23.3.6) with $\phi(s) = \exp(s)$.

However, we can do it by Theorem 23.2.6, since all the inequalities which appear in (23.2.5) are satisfied:

$$0 < \alpha_0 = 0.023303\ldots \leq \alpha = 0.577350\ldots \leq 1 - 2M_1\eta = 0.966625\ldots,$$

where $\|[F'(x_0)]^{-1}\| = 11.169433\ldots$, $M_1 = 0.120657\ldots$ and

$$p(t) = (0.180986\ldots)t^3 + (0.180986\ldots)t^2 - (0.060328\ldots)t - (0.060328\ldots).$$

Then, we can use secant-like method (23.1.2) with $\lambda = \frac{1}{2}$ to approximate a solution of (23.3.6) with $\phi(u) = \exp(u)$, the approximation given by the vector $\mathbf{x}^* = (x_1^*, x_2^*, \ldots, x_8^*)^t$ shown in Table 23.3.1 and reached after four iterations with a tolerance 10^{-16}. In Table 23.3.2 we show the errors $\|\mathbf{x_n} - \mathbf{x}^*\|$ using the stopping criterion $\|\mathbf{x_n} - \mathbf{x_{n-1}}\| < 10^{-16}$. Notice that the vector shown in Table 23.3.1 is a good approximation of the solution of (23.3.6) with $\phi(u) = \exp(u)$, since $\|F(\mathbf{x}^*)\| \leq C \times 10^{-16}$. See the sequence $\{\|F(\mathbf{x_n})\|\}$ in Table 23.3.2.

Table 23.3.1. Approximation of the solution $\mathbf{x}^*$ of (23.3.6) with $\phi(u) = \exp(u)$

n	x_i^*	n	x_i^*	n	x_i^*	n	x_i^*
1	0.05481058...	3	0.12475178...	5	0.13893761...	7	0.09657993...
2	0.09657993...	4	0.13893761...	6	0.12475178...	8	0.05481058...

Table 23.3.2. Absolute errors obtained by secant-like method (23.1.2) with $\lambda = \frac{1}{2}$ and $\{\|F(\mathbf{x_n})\|\}$

n	$\|\mathbf{x_n} - \mathbf{x}^*\|$	$\|F(\mathbf{x_n})\|$
-1	$1.3893\ldots \times 10^{-1}$	$8.6355\ldots \times 10^{-2}$
0	$4.5189\ldots \times 10^{-2}$	$1.2345\ldots \times 10^{-2}$
1	$1.43051\ldots \times 10^{-4}$	$2.3416\ldots \times 10^{-5}$
2	$1.14121\ldots \times 10^{-7}$	$1.9681\ldots \times 10^{-8}$
3	$4.30239\ldots \times 10^{-13}$	$5.7941\ldots \times 10^{-14}$

23.3.2. Example 2

Consider the following nonlinear boundary value problem

$$\begin{cases} u'' = -u^3 - \dfrac{1}{4}\, u^2 \\ u(0) = 0, \quad u(1) = 1. \end{cases}$$

It is well known that this problem can be formulated as the integral equation

$$u(s) = s + \int_0^1 Q(s,t)\,(u^3(t) + \frac{1}{4}\,u^2(t))\,dt \tag{23.3.8}$$

where, Q is the Green function:

$$Q(s,t) = \begin{cases} t\,(1-s), & t \le s \\ s\,(1-t), & s < t. \end{cases}$$

We observe that

$$\max_{0 \le s \le 1} \int_0^1 |Q(s,t)|\,dt = \frac{1}{8}.$$

Then problem (23.3.8) is in the form (23.1.1), where, F is defined as

$$[F(x)]\,(s) = x(s) - s - \int_0^1 Q(s,t)\,(x^3(t) + \frac{1}{4}\,x^2(t))\,dt.$$

The Fréchet derivative of the operator F is given by

$$[F'(x)y]\,(s) = y(s) - 3\int_0^1 Q(s,t)x^2(t)y(t)dt - \frac{1}{2}\int_0^1 Q(s,t)x(t)y(t)dt.$$

Choosing $x_0(s) = s$ and $R = 1$ we have that $\|F(x_0)\| \leq \dfrac{1+\frac{1}{4}}{8} = \dfrac{5}{32}$. Define the divided difference defined by

$$[x,y;F] = \int_0^1 F'(\tau x + (1-\tau)y)d\tau.$$

Taking into account that

$$\begin{aligned}
\|[\mathbf{x},\mathbf{y};F] - [\mathbf{v},\mathbf{y};F]\| &\leq \int_0^1 \|F'(\tau\mathbf{x} + (1-\tau)\mathbf{y}) - F'(\tau\mathbf{v} + (1-\tau)\mathbf{y})\| \, d\tau \\
&\leq \frac{1}{8}\int_0^1 \left(3\tau^2\|x^2 - v^2\| + 2\tau(1-\tau)\|y\|\|x-v\| + \frac{\tau}{2}\|x-v\|\right) d\tau \\
&\leq \frac{1}{8}\left(\|\mathbf{x}^2 - \mathbf{v}^2\| + \left(\|y\| + \frac{1}{4}\right)\right)\|\mathbf{x}-\mathbf{v}\| \\
&\leq \frac{1}{8}\left(\|x+v\| + \|y\| + \frac{1}{4}\right)\|\mathbf{x}-\mathbf{v}\| \\
&\leq \frac{25}{32}\|\mathbf{x}-\mathbf{v}\|
\end{aligned}$$

Choosing $x_{-1}(s) = \dfrac{9s}{10}$, we find that

$$\begin{aligned}
\|1 - A_0\| &\leq \int_0^1 \|F'(\tau\mathbf{x_0} + (1-\tau)\mathbf{x_{-1}})\| \, d\tau \\
&\leq \frac{1}{8}\int_0^1 \left(3\left(\tau + (1-\tau)\frac{9}{10}\right)^2 + \frac{1}{2}\left(\tau + (1-\tau)\frac{9}{10}\right)\right) d\tau \\
&\leq 0.409375\ldots
\end{aligned}$$

Using the Banach Lemma on invertible operators we obtain

$$\|A_0^{-1}\| \leq 1.69312\ldots$$

and so

$$L \geq \frac{25}{32}\|A_0^{-1}\| = 1.32275\ldots$$

.

In an analogous way, choosing $\lambda = 0.8$ we obtain

$$M_0 = 0.899471\ldots,$$

$$\|B_0^{-1}\| = 1.75262\ldots$$

and

$$\eta = 0.273847\ldots.$$

Notice that we can not guarantee the convergence of the secant method by Theorem 3.1 since the first condition of (3.1) is not satisfied:

$$a = \frac{\eta}{c+\eta} = 0.732511\ldots > \frac{3-\sqrt{5}}{2} = 0.381966\ldots$$

On the other hand, observe that

$$\tilde{M}_1 = 0.0988372\ldots,$$

$$\tilde{K} = 1.45349\ldots,$$

$$\alpha_0 = 0.434072\ldots,$$

$$\alpha = 0.907324\ldots$$

and

$$1 - 2\tilde{M}_1\eta = 0.945868\ldots.$$

And condition (2.62) $0 < \alpha_0 \leq \alpha \leq 1 - 2\tilde{M}_1\eta$ is satisfied and as a consequence we can ensure the convergence of the secant method by Theorem 23.2.9.

Conclusion

We presented a new semilocal convergence analysis of the secant-like method for approximating a locally unique solution of an equation in a Banach space. Using a combination of Lipschitz and center-Lipschitz conditions, instead of only Lipschitz conditions invested in [18], we provided a finer analysis with larger convergence domain and weaker sufficient convergence conditions than in [15, 18, 19, 21, 26, 27]. Numerical examples validate our theoretical results.

References

[1] Amat, S., Busquier, S., Negra, M., Adaptive approximation of nonlinear operators, *Numer. Funct. Anal. Optim.*, 25 (2004) 397-405.

[2] Amat, S., Busquier, S., Gutiérrez, J. M., On the local convergence of secant-type methods, *Int. J. Comp. Math.*, 81(8) (2004), 1153-1161.

[3] Amat, S., Bermúdez, C., Busquier, S., Gretay, J. O., Convergence by nondiscrete mathematical induction of a two step secant's method, *Rocky Mountain J. Math.*, 37(2) (2007), 359-369.

[4] Andronow, A.A., Chaikin, C. E., *Theory of oscillations*, Princenton University Press, New Jersey, 1949.

[5] Argyros, I.K., Polynomial Operator Equations in Abstract Spaces and Applications, St.Lucie/CRC/Lewis Publ. *Mathematics series*, 1998, Boca Raton, Florida, U.S.A.

[6] Argyros, I.K., A unifying local-semilocal convergence analysis and applications for two-point Newton-like methods in Banach space, *J. Math. Anal. Appl.*, 298 (2004), 374-397.

[7] Argyros, I.K., New sufficient convergence conditions for the secant method, *Chechoslovak Math. J.*, 55 (2005), 175-187.

[8] Argyros, I.K., *Convergence and Applications of Newton–type Iterations*, Springer–Verlag Publ., New–York, 2008.

[9] Argyros, I.K., A semilocal convergence analysis for directional Newton methods, *Math. Comput.*, 80 (2011), 327-343.

[10] Argyros, I.K., Cho, Y.J., Hilout, S., *Numerical Methods for Equations and its Applications*, CRC Press/Taylor and Francis, Boca Raton, Florida, USA, 2012

[11] Argyros, I.K., Hilout, S., Convergence conditions for secant–type methods, *Chechoslovak Math. J.*, 60 (2010), 253–272.

[12] Argyros, I.K., Hilout, S., Weaker conditions for the convergence of Newton's method, *J. Complexity*, 28 (2012), 364-387.

[13] Argyros, I.K., Hilout, S., Estimating upper bounds on the limit points of majorizing sequences for Newton's method, *Numer. Algorithms*, 62(1) (2013), 115-132.

[14] Argyros, I.K., Hilout, S., *Numerical methods in nonlinear analysis*, World Scientific Publ. Comp., New Jersey, 2013.

[15] Argyros, I.K., Ezquerro, J.A., Hernández, M.Á. Hilout, S., Romero, N., Velasco, A.I., Expanding the applicability of secant-like methods for solving nonlinear equations, *Carp. J. Math.* 31 (1) (2015), 11-30.

[16] Dennis, J.E., Toward a unified convergence theory for Newton–like methods, in *Nonlinear Functional Analysis and Applications* (L.B. Rall, ed.), Academic Press, New York, (1971), 425–472.

[17] Ezquerro, J.A., Gutiérrez, J.M., Hernández, M.A., Romero, N., Rubio, M.J., The Newton method: from Newton to Kantorovich. (Spanish), *Gac. R. Soc. Mat. Esp.*, 13 (2010), 53–76.

[18] Ezquerro, J.M., Hernández, M.A., Romero, N., A.I. Velasco,Improving the domain of starting point for secant-like methods,*App. Math. Comp.*, 219 (8) (2012), 3677–3692.

[19] Ezquerro, J.A., Rubio, M.J., A uniparametric family of iterative processes for solving nondifferentiable equations, *J. Math. Anal. Appl.*, 275 (2002), 821–834.

[20] Hernández, M.A., Rubio, M.J., Ezquerro, J.A., secant-like methods for solving nonlinear integral equations of the Hammerstein type. Proceedings of the 8th International Congress on Computational and Applied Mathematics, ICCAM-98 (Leuven), *J. Comput. Appl. Math.*, 115 (2000), 245–254.

[21] Hernández, M.A., Rubio, M.J., Ezquerro, J.A., Solving a special case of conservative problems by secant-like methods, *App. Math. Comp.*, 169 (2005), 926–942.

[22] Kantorovich, L.V., Akilov, G.P., *Functional Analysis*, Pergamon Press, Oxford, 1982.

[23] Laasonen, P., Ein überquadratisch konvergenter iterativer algorithmus, *Ann. Acad. Sci. Fenn. Ser* I, 450 (1969), 1–10.

[24] Magreñán, Á.A., A new tool to study real dynamics: The convergence plane, *App. Math. Comp.*, 248 (2014), 215–224.

[25] Ortega, J.M., Rheinboldt, W.C., *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970.

[26] Potra, F.A., Sharp error bounds for a class of Newton–like methods, *Libertas Mathematica*, 5 (1985), 71–84.

[27] Potra, F.A., Pták, V., *Nondiscrete Induction and Iterative Processes*, Pitman, New York, 1984.

[28] Proinov, P.D., General local convergence theory for a class of iterative processes and its applications to Newton's method, *J. Complexity*, 25 (2009), 38–62.

[29] Proinov, P.D., New general convergence theory for iterative processes and its applications to Newton–Kantorovich type theorems, *J. Complexity*, 26 (2010), 3–42.

[30] Schmidt, J.W., Untere Fehlerschranken fur Regula-Falsi Verhafren, *Period. Hungar.*, 9 (1978), 241–247.

[31] Stoker, J.J., *Nonlinear vibrations*, Interscience-Wiley, New York, 1950.

[32] Traub, J.F., *Iterative Methods for the Solution of Equations*, Prentice-Hall, Englewood Cliffs, New Jersay, 1964.

[33] Yamamoto, T., A convergence theorem for Newton–like methods in Banach spaces, *Numer. Math.*, 51 (1987), 545–557.

[34] Wolfe, M.A., Extended iterative methods for the solution of operator equations, *Numer. Math.*, 31 (1978), 153–174.

Author Contact Information

Ioannis K. Argyros
Professor
Department of Mathematical Sciences
Cameron University
Lawton, OK, US
Tel: (580) 581-2908
Email: iargyros@cameron.edu

Á. Alberto Magreñán
Professor
Universidad Internacional de La Rioja
Departmento de Matematicas
Logroño, La Rioja, Spain
Tel: (+34) 679-257459
Email: alberto.magrenan@unir.net

Index

A

B

C

D

F

G

H

I

T

V

W