Frontiers in Mathematics

José Antonio Ezquerro Fernández
Miguel Ángel Hernández Verón

Newton's Method: an Updated Approach of Kantorovich's Theory

Birkhäuser

Birkhäuser

Frontiers in Mathematics

More information about this series at http://www.springer.com/series/5388

José Antonio Ezquerro Fernández
Miguel Ángel Hernández Verón

Newton's Method: an Updated Approach of Kantorovich's Theory

Birkhäuser

José Antonio Ezquerro Fernández
Department of Mathematics and Computation
University of La Rioja
La Rioja, Spain

Miguel Ángel Hernández Verón
Department of Mathematics and Computation
University of La Rioja
La Rioja, Spain

ISSN 1660-8046 ISSN 1660-8054 (electronic)
Frontiers in Mathematics
ISBN 978-3-319-55975-9 ISBN 978-3-319-55976-6 (eBook)
DOI 10.1007/978-3-319-55976-6

Library of Congress Control Number: 2017944725

Mathematics Subject Classification (2010): 47H99, 65J15, 65H10, 45G10, 34B15

Printed on acid-free paper

This book is published under the trade name Birkhäuser, www.birkhauser-science.com
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Dedicated to my daughter María,
for always being there for me. JAE.

Dedicated to my dear children
Ramón, Jorge, Miguel, Diego and Mercedes,
who are the engine of my life,
for their love and patience. MAH.

Preface

One of the most common problems in mathematics is the solution of a nonlinear equation $F(x) = 0$. This problem is not always easy to solve, since we cannot frequently obtain an exact solution to the previous equation, so that we usually look for a numerical approximation to a solution. In this case, we use approximation methods, which are generally iterative. The best known iteration to solve nonlinear equations is undoubtedly Newton's method:

$$x_{n+1} = x_n - [F'(x_n)]^{-1}F(x_n), \quad n \geq 0, \quad \text{with } x_0 \text{ given.}$$

The geometric interpretation of Newton's method is well known if F is a real function. In such a case, x_{n+1} is the point where the tangential line $y - F(x_n) = F'(x_n)(x - x_n)$ of the function $F(x)$ intersects the x-axis at the point $(x_n, F(x_n))$. The geometric interpretation of the complex Newton method, $F : \mathbb{C} \longrightarrow \mathbb{C}$, is given by Yau and Ben-Israel in [84]. In the general case, $F(x)$ is approximated at point x_n as $F(x) \approx L_n(x) = F(x_n) + F'(x_n)(x - x_n)$ and the zero of $L_n(x) = 0$ defines the new approximation x_{n+1}.

In spite of its simple principle [83], depending on the domain of F, Newton's method is applicable to various types of equations such as systems of nonlinear algebraic equations including matrix eigenvalue problems, differential equations, integral equations, etc., and even to random operator equations [8]. Hence, the method fascinates many researchers. However, as it is well known, a disadvantage of the method is that the initial approximation x_0 must be chosen sufficiently close to a true solution in order to guarantee its convergence. Finding a criterion for choosing x_0 is quite difficult. In this text we try to facilitate the choice of x_0 under conditions as general as possible.

The long history of Newton's method has already been well studied, see, e.g., N. Kollerstrom [57] or T. J. Ypma [85]. According to these articles, in [19], P. Deuflhard points out the facts seem to be agreed upon among the experts. The interested reader may find more historical details in the book by H. H. Goldstine [35].

In particular, as a summary, we rewrite some notes and remarks given by Ortega and Rheinboldt in [64], which is a good survey until 1970. Point-of-attraction results date back to the 19th century. For one-dimensional equations, the quadratic convergence of Newton's method was established by Cauchy (1829). For equations in $\mathbb{R}^n$, a point-of-attraction theorem was given by Runge (1899), who also stressed the quadratic convergence. Independently, a result for $n = 2$ was given by Blutel (1910). For convergence results which do not assume the existence of a solution, Fine (1916) appears to have been the first to prove the convergence of Newton's method in n dimensions, where the derivative $F'(x)$ is assumed to be invertible on some suitable ball. In the same year, Bennet (1916) formulated related results for operators on infinite dimensional spaces but for the proofs he referred to Fine. The article of Fine appears to have been overlooked, and twenty years later Ostrowski (1936) presented, independently, new convergence theorems and also discussed error estimates. Concurrently, Willers (1938) also proved similar convergence conditions. Although both these authors observed that their results extend immediately to the case of a general n, they themselves presented them only for $n = 2$ and $n = 3$, respectively. Bussmann (1940), in an unpublished dissertation, proved these results and some extensions for general n; Bussmann's theorems are quoted by Rehbock (1942).

Later, the Russian mathematician L. V. Kantorovich gave his now famous convergence results for Newton's method in Banach spaces. In 1948, Kantorovich published the seminal paper [47], where he suggested an extension of Newton's method to functional spaces and established a semilocal convergence result for Newton's method in a Banach space, which is now called Kantorovich's theorem or, more specifically, *the Newton-Kantorovich theorem*, as we will call from now on. The result was also included in the survey paper [48]. Further developments of the method can be found in [49, 50, 51, 52, 53] and in the monographs [54, 55].

The main contribution of Kantorovich is the formulation of the problem in a general setting, the spaces of Banach, that uses appropriate techniques of functional analysis. This event cannot be overestimated, since Newton's method became a powerful tool in numerical analysis as well as in pure mathematics. The approach of Kantorovich guarantees the application of Newton's method to solve a large variety of functional equations: nonlinear integral equations, ordinary and partial differential equations, variational problems, etc. Various examples of such applications are presented in [54, 55]. For a list of relevant publications where Newton's method is applied to different functional equations, see [64] and the references cited there.

The Kantorovich result is a masterpiece not only by its sheer importance but by the original and powerful proof technique. The results of Kantorovich and his school initiated some very intensive research on the Newton and related methods. A great number of variants and extensions of his results emerged in the literature. Basic results on Newton's method and numerous references may be found in the books of Ostrowski [65] and Ortega and Rheinboldt [64]. More recent bibliography is available in the books of Rheinboldt [74] and Deulfhard [18], survey paper [85] and the special web site devoted to Newton's method [59]. A revision of the most important theoretical results on Newton's method concerning the convergence properties, the error estimates, the numerical stability and the computational complexity of the algorithm may be found in [33].

On the other hand, three types of studies can be done when we are interested in proving the convergence of the sequence $\{x_n\}$ given by Newton's method: local, semilocal and global. First, the local study of the convergence is based on demanding conditions to the solution x^*, from certain conditions on the operator F, and provide the so-called ball of convergence ([16]) of the sequence $\{x_n\}$, that shows the accessibility to x^* from the initial approximation x_0 belonging to the ball. Second, the semilocal study of the convergence is based on demanding conditions to the initial approximation x_0, from certain conditions on the operator F, and provide the so-called domain of parameters ([30]) corresponding to the conditions required to the initial approximation that guarantee the convergence of the sequence $\{x_n\}$ to the solution x^*. Third, the global study of the convergence guarantees, from certain conditions on the operator F, the convergence of the sequence $\{x_n\}$ to the solution x^* in a domain and independently of the initial approximation x_0. The three studies demand conditions on the operator F. However, requirement of conditions to the solution, to the initial approximation, or to none of these, determines the different types of studies.

The local study of the convergence has the disadvantage of being able to guarantee that the solution, that is unknown, can satisfy certain conditions. In general, the global study of the convergence is very specific as regards the type of operators to consider, as a consequence of absence of conditions on the initial approximation and on the solution. There is a plethora of studies on the weakness and/or extension of the hypothesis made on the underlying operators. In this textbook, we focus our attention on the analysis of the semilocal convergence of Newton's method.

This textbook is written for researchers interested in the theory of Newton's method in Banach spaces. Each chapter contains several theoretical results and interesting applications in the solution of nonlinear integral and differential equations.

Chapter 1 presents an analysis of Kantorovich's theory for Newton's method where the original theory is given along with the best-known variant, which is due to Ortega and uses the method of majorizing sequences. In addition, we include a new approach by introducing a new concept of majorant function which is different from that defined by Kantorovich. In all the results presented, we suppose that the second derivative of the operator involved is bounded in norm in the domain where the operator is defined.

In Chapter 2, we analyse the semilocal convergence of Newton's method under different modifications of the condition on the second derivative of the operator involved. We begin by presenting the result of Huang where the second derivative of the operator involved is Lipschitz continuous in the domain where the operator is defined. We pay attention to the proof given by Huang and see that the condition on the second derivative can be relaxed to a center Lipschitz condition to establish the semilocal convergence of Newton's method. This observation leads us to propose milder conditions on the second derivative of the operator involved. So, we first prove the semilocal convergence of Newton's method under a center ω-Lipschitz condition for the second derivative of the operator involved. Next, the condition of the second derivative of the operator required by Kantorovich is generalized to a condition where the second derivative is ω-bounded in the domain where the operator is defined. This generalization is very interesting because it avoids having to look for a domain where the second derivative of the operator is bounded and contains a solution of the equation to solve, which is an important problem that presents Kantorovich's theory.

The ideas developed in Chapter 2 are extended to higher order derivatives of the operator involved in Chapter 3, where new starting points for Newton's method are located despite the new conditions imposed to the operator. Three sections are included where the semilocal convergence of Newton's method is analysed for polynomial operators, operators with ω-bounded k-th-derivative and operators with ω-Lipschitz k-th-derivative.

Chapter 4 examines the typical situation in which conditions on the first derivative of the operator are only required, instead of conditions on higher order derivatives, as a consequence of the fact that this is the only derivative of the operator involved appearing in the algorithm of the method. We analyse in detail Ortega's variant of the original Kantorovich's result from the method of majorizing sequences and pay special attention to the fact that, if we want to relax the conditions imposed to the first derivative of the operator, we need to apply another distinct technique to that of the method of majorizing sequences presented in Chapter 1 to guarantee the semilocal convergence of Newton's method. In particular, we propose a system of recurrence relations. Techniques based on recurrence relations were already used by Kantorovich in his first proofs of the semilocal convergence of Newton's method. We then analyse two cases: the cases in which the first derivative of the operator involved is ω-Lipschitz continuous and center ω-Lipschitz continuous in the domain where the operator involved is defined.

We have included a numerical example in each situation analysed theoretically to so justify its analysis. In particular, in Chapter 1, we present two types of problems that are used throughout the following chapters: nonlinear integral equations of Hammerstein type and nonlinear boundary value problems. We analyse these problems some times on a continuous form and others in a discrete way.

We emphasize the fact that we have tried to facilitate the reading of the textbook from a detailed development of the proofs of the results given. We also want to point out that the textbook presents some of our investigations into Newton's method which we have carried out over many years, including an abundant and specialized bibliography.

We ended up saying that the main aim of this textbook is to develop, expand and update the theory introduced by Kantorovich for Newton's method under different conditions on the operator involved, but always with the clear objective to improve the applicability of the method based on the location of new starting points.

Logroño, La Rioja J. A. Ezquerro
June 2016 M. A. Hernández-Verón

Contents

Preface **vii**

1 The classic theory of Kantorovich **1**
- 1.1 The Newton-Kantorovich theorem . . . 1
  - 1.1.1 Recurrence relations of Kantorovich . . . 2
  - 1.1.2 The majorant principle of Kantorovich . . . 7
  - 1.1.3 The “method of majorizing sequences” . . . 17
  - 1.1.4 New approach: a “majorant function” from an initial value problem . . . 24
- 1.2 Applications . . . 28
  - 1.2.1 Integral equations . . . 29
  - 1.2.2 Boundary value problems . . . 32

2 Convergence conditions on the second derivative of the operator **39**
- 2.1 Operators with Lipschitz-type second-derivative . . . 40
  - 2.1.1 Operators with Lipschitz second-derivative . . . 40
  - 2.1.2 Operators with Hölder second-derivative . . . 46
  - 2.1.3 Operators with ω-Lipschitz second-derivative . . . 47
    - 2.1.3.1 Existence of a majorizing sequence . . . 48
    - 2.1.3.2 Existence of a majorant function . . . 54
    - 2.1.3.3 Semilocal convergence . . . 55
    - 2.1.3.4 Uniqueness of solution and order of convergence . . . 55
    - 2.1.3.5 Applications . . . 57
    - 2.1.3.6 A little more: three particular cases . . . 63
- 2.2 Operators with ω-bounded second-derivative . . . 70
  - 2.2.1 Existence of a majorant function . . . 72
  - 2.2.2 Existence of a majorizing sequence . . . 73
  - 2.2.3 Semilocal convergence . . . 75
  - 2.2.4 Uniqueness of solution and order of convergence . . . 76
  - 2.2.5 Applications . . . 78

3 Convergence conditions on the k-th derivative of the operator **83**
- 3.1 Polynomial type equations . . . 83
  - 3.1.1 Semilocal convergence . . . 85
  - 3.1.2 Uniqueness of solution and order of convergence . . . 89
  - 3.1.3 Applications . . . 91
- 3.2 Operators with ω-bounded k-th-derivative . . . 95
  - 3.2.1 Existence of a majorizing sequence . . . 96

3.2.2 Existence of a majorant function . . . 102
3.2.3 Semilocal convergence . . . 103
3.2.4 Uniqueness of solution and order of convergence . . . 104
3.2.5 Application . . . 105
3.3 Operators with ω-Lipschitz k-th-derivative . . . 109
3.3.1 Existence of a majorizing sequence . . . 110
3.3.2 Existence of a majorant function . . . 116
3.3.3 Semilocal convergence . . . 118
3.3.4 Uniqueness of solution and order of convergence . . . 118
3.3.5 Application . . . 120
3.3.6 Relaxing convergence conditions . . . 122

4 Convergence conditions on the first derivative of the operator 127
4.1 Operators with Lipschitz first-derivative . . . 128
4.1.1 Existence of a majorizing sequence . . . 128
4.1.2 Existence of a majorant function . . . 132
4.1.3 Semilocal convergence . . . 132
4.2 Operators with ω-Lipschitz first-derivative . . . 134
4.2.1 Semilocal convergence using recurrence relations . . . 135
4.2.1.1 Recurrence relations . . . 135
4.2.1.2 Analysis of the scalar sequences . . . 136
4.2.1.3 Semilocal convergence . . . 137
4.2.2 Uniqueness of solution and order of convergence . . . 139
4.2.3 Application . . . 143
4.2.4 Particular cases . . . 147
4.2.4.1 Operators with Lipschitz first-derivative . . . 148
4.2.4.2 Operators with Hölder first-derivative . . . 149
4.3 Operators with center ω-Lipschitz first-derivative . . . 150
4.3.1 Semilocal convergence . . . 151
4.3.2 Particular case: operators with center Lipschitz first-derivative . . . 156
4.3.3 Application . . . 157

Bibliography 161

Chapter 1

The classic theory of Kantorovich

According to Polyak [66], Kantorovich proved in 1939 the semilocal convergence of Newton's method [46] on the basis of the contraction mapping principle of Banach, and later improved to semilocal quadratic convergence in 1948/49 (*the Newton-Kantorovich theorem*) [47, 49]. Also in 1949, Mysovskikh [61] gave a much simpler independent proof of semilocal quadratic convergence under slightly different theoretical assumptions, which are exploited in modern Newton algorithms, see [18].

Kantorovich gave two basically different proofs of the Newton-Kantorovich theorem using recurrence relations or majorant functions. The original proof given by Kantorovich uses recurrence relations [47]. A nice treatment of this theorem using recurrence relations can be found in [72]. In [50] Kantorovich gave a proof based on the concept of a majorant function.

An important feature of the Newton-Kantorovich theorem, or related results, is that it does not assume the existence of a solution, so that the theorem is not only a convergence result for Newton's method, but simultaneously a theorem of existence of solution for nonlinear equations in Banach spaces. In addition, the theoretical significance of Newton's method can be used to draw conclusions about the existence and uniqueness of a solution and about the region in which it is located, without finding the solution itself and this is sometimes more important than the actual knowledge of the solution (see [55]). The results of this section follows Kantorovich's paper [53].

After Kantorovich establishes the Newton-Kantorovich theorem, a large number of results has been published concerning convergence and error bounds for Newton's method under the assumptions of the Newton-Kantorovich theorem or under closely related ones. Among later convergence theorems, the ones due to Ortega and Rheinboldt [64] are worth mentioning. Also, there exist numerous versions of the Newton-Kantorovich theorem that differ in assumptions and results, and it would be impossible to list all relevant publications here. We then only mention some versions that may find in [11, 54, 55, 58, 62, 65].

Throughout the textbook we denote $\overline{B(x,\varrho)} = \{y \in X; \|y-x\| \leq \varrho\}$ and $B(x,\varrho) = \{y \in X; \|y-x\| < \varrho\}$.

1.1 The Newton-Kantorovich theorem

Newton's method has the form

$$x_{n+1} = N_F(x_n) = x_n - [F'(x_n)]^{-1}F(x_n), \quad n \geq 0, \quad \text{with } x_0 \text{ given}, \tag{1.1}$$

J.A. Ezquerro Fernández, M.Á. Hernández Verón, *Newton's Method: an Updated Approach of Kantorovich's Theory*, Frontiers in Mathematics, DOI 10.1007/978-3-319-55976-6_1

for the solution of the nonlinear equation $F(x) = 0$, where $F : X \longrightarrow Y$, X and Y are Banach spaces and F is twice continuously Fréchet differentiable. Below, we present some results given by Kantorovich on the convergence of Newton's method.

1.1.1 Recurrence relations of Kantorovich

Kantorovich established a semilocal convergence theorem for Newton's method in a Banach space in 1948 under the following conditions for the operator F and the starting point x_0 [47]:

(W1) For $x_0 \in X$, there exists $\Gamma_0 = [F'(x_0)]^{-1} \in \mathcal{L}(Y, X)$, where $\mathcal{L}(Y, X)$ is the set of bounded linear operators from Y to X, such that $\|\Gamma_0\| \leq \beta$,

(W2) $\|\Gamma_0 F(x_0)\| \leq \eta$,

(W3) there exists $R > 0$ such that $\|F''(x)\| \leq M$, for $x \in B(x_0, R)$,

(W4) $h = M\beta\eta \leq \frac{1}{2}$.

Theorem 1.1. (The Newton-Kantorovich theorem) *Let $F : \Omega \subseteq X \longrightarrow Y$ be a twice continuously Fréchet differentiable operator defined on a non-empty open convex domain Ω of a Banach space X with values in a Banach space Y. Suppose that conditions (W1)-(W2)-(W3)-(W4) are satisfied and $B(x_0, \rho^*) \subset B(x_0, R)$ with $\rho^* = \frac{1-\sqrt{1-2h}}{h}\eta$. Then, Newton's sequence defined in (1.1) and starting at x_0 converges to a solution x^* of the equation $F(x) = 0$ and the solution x^* and the iterates x_n belong to $\overline{B(x_0, \rho^*)}$, for all $n \geq 0$ Moreover, if $h < \frac{1}{2}$, the solution x^* is unique in $B(x_0, \rho^{**}) \cap B(x_0, R)$, where $\rho^{**} = \frac{1+\sqrt{1-2h}}{h}\eta$, and, if $h = \frac{1}{2}$, x^* is unique in $\overline{B(x_0, \rho^*)}$. Furthermore, we have the following error estimates:*

$$\|x^* - x_n\| \leq \frac{1}{2^{n-1}}(2h)^{2^n-1}\eta, \quad n = 0, 1, 2, \ldots \tag{1.2}$$

Proof. First of all, we define $\beta_0 = \beta$, $\eta_0 = \eta$ and $h_0 = h$. After that, we observe that x_1 is well-defined, since the operator $\Gamma_0 = [F'(x_0)]^{-1}$ exists by the hypotheses. Moreover, $\|x_1 - x_0\| \leq \eta_0$, so that $x_1 \in B(x_0, \rho^*)$.

Next, taking into account that

$$\begin{aligned}
\|I - \Gamma_0 F'(x_1)\| &\leq \|\Gamma_0\|\|F'(x_0) - F'(x_1)\| \\
&\leq \beta_0 \left\| \int_0^1 F''(x_0 + \tau(x_1 - x_0))\, d\tau (x_1 - x_0) \right\| \\
&\leq M\beta_0 \|x_1 - x_0\| \\
&\leq M\beta_0\eta_0 \\
&= h_0 \\
&< 1,
\end{aligned}$$

it follows, by the Banach lemma on invertible operators [72], that the operator $\Gamma_1 = [F'(x_1)]^{-1}$ exists and

$$\|\Gamma_1\| \leq \frac{\beta_0}{1 - h_0} = \beta_1.$$

Hence, x_2 is well-defined.

In addition, as

$$\begin{aligned} F(x_1) &= F(x_0) + F'(x_0)(x_1 - x_0) + \int_{x_0}^{x_1} F''(z)(x_1 - z)\, dz \\ &= \int_{x_0}^{x_1} F''(z)(x_1 - z)\, dz \\ &= \int_0^1 F''(x_0 + \tau(x_1 - x_0))(x_1 - x_0)^2 (1-\tau)\, d\tau, \end{aligned}$$

we have

$$\begin{aligned} \|F(x_1)\| &\leq \int_0^1 \|F''(x_0 + \tau(x_1 - x_0))\| \,(1-\tau)\, d\tau \|x_1 - x_0\|^2 \\ &\leq \frac{M}{2}\|x_1 - x_0\|^2 \\ &\leq \frac{M}{2}\eta_0^2 \end{aligned}$$

and

$$\|x_2 - x_1\| = \|\Gamma_1 F(x_1)\| \leq \|\Gamma_1\|\|F(x_1)\| \leq \frac{\beta_0}{1-h_0}\frac{M}{2}\eta_0^2 = \frac{h_0\eta_0}{2(1-h_0)} = \eta_1 \leq h_0\eta_0.$$

Moreover,

$$h_1 = M\beta_1\eta_1 = \frac{h_0^2}{2(1-h_0)^2} \leq 2h_0^2 \leq \frac{1}{2}.$$

Therefore, the hypotheses of the Newton-Kantorovich theorem are also satisfied when we substitute β_1, η_1 and h_1 for β_0, η_0 and h_0, respectively. This allows us to continue the successive determination of the elements x_n and the numbers connected with them, β_n, η_n and h_n, so that if we assume

$$\|\Gamma_n\| \leq \beta_n = \frac{\beta_{n-1}}{1-h_{n-1}}, \tag{1.3}$$

$$\|x_{n+1} - x_n\| \leq \frac{M}{2}\beta_n\eta_{n-1}^2 = \frac{h_{n-1}\eta_{n-1}}{2(1-h_{n-1})} = \eta_n \leq h_{n-1}\eta_{n-1} \leq \frac{1}{2^n}(2h_0)^{2^n-1}\eta_0, \tag{1.4}$$

$$h_n = M\beta_n\eta_n = \frac{h_{n-1}^2}{2(1-h_{n-1})^2} \leq 2h_{n-1}^2 \leq \frac{1}{2}, \tag{1.5}$$

where the operator $\Gamma_n = [F'(x_n)]^{-1}$ exists, it follows the following.

As

$$\begin{aligned} \|I - \Gamma_n F'(x_{n+1})\| &\leq \|\Gamma_n\|\|F'(x_n) - F'(x_{n+1})\| \\ &\leq \beta_n \left\| \int_0^1 F''(x_n + \tau(x_{n+1} - x_n))\, d\tau (x_{n+1} - x_n) \right\| \\ &\leq M\beta_n\|x_{n+1} - x_n\| \\ &\leq M\beta_n\eta_n \\ &= h_n \\ &< 1, \end{aligned}$$

we have, by the Banach lemma on invertible operators, that the operator Γ_{n+1} exists and

$$\|\Gamma_{n+1}\| \leq \frac{\beta_n}{1-h_n} = \beta_{n+1}.$$

Hence, x_{n+2} is well-defined.

Besides, as

$$\begin{aligned} F(x_{n+1}) &= F(x_n) + F'(x_n)(x_{n+1}-x_n) + \int_{x_n}^{x_{n+1}} F''(z)(x_{n+1}-z)\,dz \\ &= \int_{x_n}^{x_{n+1}} F''(z)(x_{n+1}-z)\,dz \\ &= \int_0^1 F''(x_n+\tau(x_{n+1}-x_n))(x_{n+1}-x_n)^2(1-\tau)\,d\tau \end{aligned}$$

we have

$$\begin{aligned} \|F(x_{n+1})\| &\leq \int_0^1 \|F''(x_n+\tau(x_{n+1}-x_n))\|\,(1-\tau)\,d\tau\|x_{n+1}-x_n\|^2 \\ &\leq \frac{M}{2}\|x_{n+1}-x_n\|^2 \\ &\leq \frac{M}{2}\eta_n^2 \end{aligned}$$

and

$$\|x_{n+2}-x_{n+1}\| = \|\Gamma_{n+1}F(x_{n+1})\| \leq \|\Gamma_{n+1}\|\|F(x_{n+1})\| \leq \frac{M}{2}\beta_{n+1}\eta_n^2 = \frac{h_n\eta_n}{2(1-h_n)} = \eta_{n+1}.$$

Moreover, since $h_n \leq \frac{1}{2}$, we also have

$$\eta_{n+1} \leq h_n\eta_n \leq \cdots \leq h_nh_{n-1}\cdots h_0\eta_0 \leq \frac{1}{2^{n+1}}(2h_0)^{2^{n+1}-1}\eta_0$$

and

$$h_{n+1} = M\beta_{n+1}\eta_{n+1} = \frac{h_n^2}{2(1-h_n)^2} \leq 2h_n^2 \leq \cdots \leq \frac{1}{2}(2h_0)^{2^{n+1}} \leq \frac{1}{2}.$$

As a consequence, (1.3), (1.4) and (1.5) are true for all positive integers n by mathematical induction.

On the other hand, if we note the identity

$$\eta_n\varphi(h_n) - \eta_{n+1}\varphi(h_{n+1}) = \eta_n,$$

where $\varphi(t) = \frac{1-\sqrt{1-2t}}{t}$, which is verificable directly, since

$$\eta_{n+1}\varphi(h_{n+1}) = \eta_{n+1}\frac{1-\sqrt{1-2h_{n+1}}}{h_{n+1}} = \eta_n\frac{1-h_n-\sqrt{1-2h_n}}{h_n} = \eta_n\varphi(h_n)-\eta_n,$$

it follows, by (1.5), for $m \geq 1$ and $n \geq 1$, that

$$\begin{aligned} \|x_{n+m}-x_n\| &\leq \|x_{n+m}-x_{n+m-1}\| + \|x_{n+m-1}-x_{n+m-2}\| + \cdots + \|x_{n+1}-x_n\| \\ &\leq \eta_{n+m-1} + \eta_{n+m-2} + \cdots + \eta_n \\ &= \eta_n\varphi(h_n) - \eta_{n+m}\varphi(h_{n+m}) \\ &\leq \eta_n\varphi(h_n) \qquad (1.6) \\ &\leq 2\eta_n \\ &\leq \frac{1}{2^{n-1}}(2h_0)^{2^n-1}\eta_0, \qquad (1.7) \end{aligned}$$

so that $\{x_n\}$ is a Cauchy sequence and then convergent. In addition, passing to the limit in (1.7) as $m \to +\infty$, we obtain (1.2).

If we now do $n = 0$ in (1.6), then

$$\|x_m - x_0\| \ \le \ \eta_0 \varphi(h_0) = \rho^*,$$

so that $x_m \in \overline{B(x_0, \rho^*)}$, for all $m \in \mathbb{N}$. Moreover, $\lim_n x_n = x^* \in \overline{B(x_0, \rho^*)}$. Furthermore, x^* is a solution of $F(x) = 0$, since $\|\Gamma_n F(x_n)\| = \|x_{n+1} - x_n\| \to 0$, when $n \to +\infty$, $\|F(x_n)\| \le \|F'(x_n)\|\|x_{n+1} - x_n\|$, the sequence $\{\|F'(x_n)\|\}$ is bounded, since

$$\begin{aligned}
\|F'(x_n)\| &\le \|F'(x_0)\| + \|F'(x_n) - F'(x_0)\| \\
&\le \|F'(x_0)\| + M\|x_n - x_0\| \\
&\le \|F'(x_0)\| + M\eta_0 f(h_0),
\end{aligned}$$

and $\|F(x_n)\| \to 0$ as $n \to +\infty$. Therefore, by the continuity of F in $\overline{B(x_0, \rho^*)}$, we obtain $F(x^*) = 0$.

Finally, we prove the uniqueness of the solution x^*. We first analyse the case $h < \frac{1}{2}$. Suppose that there exists a solution $y^* \in B(x_0, \rho^{**}) \cap B(x_0, R)$ of $F(x) = 0$ and such that $y^* \neq x^*$. Then, we have $\|y^* - x_0\| \le \theta\rho^{**} = \theta\eta_0\psi(h_0)$, where $\theta \in (0, 1)$ and $\psi(t) = \frac{1+\sqrt{1-2t}}{t}$. Next, we suppose $\|y^* - x_j\| \le \theta^{2^j}\eta_j\,\psi(h_j)$, for $j = 0, 1 \dots, n$, and prove $\|y^* - x_{n+1}\| \le \theta^{2^{n+1}}\eta_{n+1}\,\psi(h_{n+1})$.

Indeed, from $F(y^*) = 0$ and $x_{n+1} = x_n - \Gamma_n F(x_n)$, it follows

$$\begin{aligned}
y^* - x_{n+1} &= y^* - x_n + \Gamma_n F(x_n) \\
&= -\Gamma_n \left(F(y^*) - F(x_n) - F'(x_n)(y^* - x_n)\right) \\
&= -\Gamma_n \int_{x_n}^{y^*} F''(z)(y^* - z)\,dz \\
&= -\Gamma_n \int_0^1 F''(x_n + \tau(y^* - x_n))(y^* - x_n)^2(1 - \tau)\,d\tau
\end{aligned} \tag{1.8}$$

and

$$\begin{aligned}
\|y^* - x_{n+1}\| &\le \frac{M}{2}\|\Gamma_n\|\|y^* - x_n\|^2 \\
&\le \frac{M}{2}\beta_n\left(\theta^{2^n}\eta_n\,\psi(h_n)\right)^2 \\
&= \theta^{2^{n+1}}\eta_{n+1}\,\psi(h_{n+1}).
\end{aligned}$$

Then, by mathematical induction, we have proved that $\|y^* - x_j\| \le \theta^{2^j}\eta_j\,\psi(h_j)$ are true for all positive integers j.

Now, as

$$\eta_n\,\psi(h_n) = \eta_n\frac{1 + \sqrt{1 - 2h_n}}{h_n} \le \frac{2\eta_n}{h_n} = \frac{2}{M\beta_n}$$

and $\beta_0 < \beta_n$, for $n \ge 0$, we have

$$\|y^* - x_n\| \le \theta^{2^n}\frac{2}{M\beta_n} < \theta^{2^n}\frac{2}{M\beta_0}$$

and therefore $\|y^* - x_n\| \to 0$ as $n \to +\infty$, so that $y^* = x^*$, since $x^* = \lim_n x_n$.

For the case $h = \frac{1}{2}$, we suppose that y^* is a solution of $F(x) = 0$ in $\overline{B(x_0, \rho^*)}$ and such that $y^* \neq x^*$. Moreover, $\|y^* - x_0\| \leq \rho^* = 2\eta_0$. We now suppose $\|y^* - x_j\| \leq \frac{\eta_0}{2^{j-1}}$, for $j = 0, 1 \ldots, n$, and prove $\|y^* - x_{n+1}\| \leq \frac{\eta_0}{2^n}$. Indeed, from (1.8), it follows

$$\|y^* - x_{n+1}\| \leq \frac{M}{2}\|\Gamma_n\|\|y^* - x_n\|^2 \leq \frac{M}{2}\beta_n(2\eta_n)^2 = 2h_n\eta_n \leq \eta_n \leq \frac{\eta_0}{2^n}.$$

Then, by mathematical induction on j, we conclude that $\|y^* - x_j\| \leq \frac{\eta_0}{2^{j-1}}$ are true for all positive integers j. As a consequence, $\|y^* - x_n\| \to 0$ as $n \to +\infty$ and therefore $y^* = x^*$. ∎

Note that condition (W4) of the Newton-Kantorovich theorem, which is often called *the Kantorovich condition*, is critical, since it means that, at the initial approximation x_0, the value $\|F(x_0)\|$ should be small enough, that is, x_0 should be close to a solution.

According to Galantai [33], if conditions (W1)-(W2)-(W3)-(W4) of the Newton-Kantorovich theorem are satisfied, then not only the Newton sequence $\{x_n\}$ exists and converges to a solution x^* but $[F'(x^*)]^{-1}$ exists in this case. Rall proves in [71] that the existence of $[F'(x^*)]^{-1}$ conversely guarantees that the hypotheses of the Newton-Kantorovich theorem with $h < \frac{1}{2}$ are satisfied at each point of an open ball with center x^*.

Notice that the Newton iterates x_n are invariant under any affine transformation $F \longrightarrow G = AF$, where A denotes any bounded and bijective linear mapping from Y to any Banach space Z [33]. This property is easily verified, since $[G'(x)]^{-1}G(x) = [F'(x)]^{-1}A^{-1}AF(x) = [F'(x)]^{-1}F(x)$. The affine invariance property is clearly reflected in the Newton-Kantorovich theorem. For other affine invariant theorems, we may see Deuflhard and Heindl [17].

Remark 1.2. The speed of convergence of an iterative method is usually measured by the order of convergence of the method. The first definition of order of convergence was given in 1870 by Schröder [76], but a very commonly measure of speed of convergence in Banach spaces is the R-order of convergence [69], which is defined as follows:

> Let $\{x_n\}$ a sequence of points of a Banach space X converging to a point $x^* \in X$ and let $\sigma \geq 1$ and
>
> $$e_n(\sigma) = \begin{cases} n & \text{if} \quad \sigma = 1, \\ \sigma^n & \text{if} \quad \sigma > 1, \end{cases} \qquad n \geq 0.$$
>
> (a) We say that σ is an R-order of convergence of the sequence $\{x_n\}$ if there are two constants $b \in (0, 1)$ and $B \in (0, +\infty)$ such that
>
> $$\|x_n - x^*\| \leq Bb^{e_n(\sigma)}.$$
>
> (b) We say that σ is the exact R-order of convergence of the sequence $\{x_n\}$ if there are four constants $a, b \in (0, 1)$ and $A, B \in (0, +\infty)$ such that
>
> $$Aa^{e_n(\sigma)} \leq \|x_n - x^*\| \leq Bb^{e_n(\sigma)}, \quad n \geq 0.$$

In general, check double inequalities of (b) is complicated, so that normally only seek upper inequalities as (a). Therefore, if we find an R-order of convergence σ of sequence $\{x_n\}$, we then say that sequence $\{x_n\}$ has order of convergence at least σ. So, according to this, estimates (1.2) guarantee that Newton's method has R-order of convergence ([69]) at least two if $h < \frac{1}{2}$ and at least one if $h = \frac{1}{2}$.