Frontiers in Mathematics

José Antonio Ezquerro Fernandez
Miguel Ángel Hernández Verón

Mild Differentiability Conditions for Newton's Method in Banach Spaces

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

Mild Differentiability Conditions for Newton's Method in Banach Spaces

Birkhäuser

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

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

ISSN 1660-8046 ISSN 1660-8054 (electronic)
Frontiers in Mathematics
ISBN 978-3-030-48701-0 ISBN 978-3-030-48702-7 (eBook)
https://doi.org/10.1007/978-3-030-48702-7

Mathematics Subject Classification: 47H99, 65J15, 65H10, 45G10, 34B15, 35J60

This book is published under the imprint Birkhäuser, www.birkhauser-science.com, by the registered company Springer Nature Switzerland AG.
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

To my daughter María,
the pillar of my life. JAE.

To my dear and always remembered parents
Carmen and Miguel. MAHV.

Preface

Many scientific and engineering problems can be written as a nonlinear equation $F(x) = 0$, where F is a nonlinear operator defined on a nonempty open convex subset Ω of a Banach space X with values in a Banach space Y. The solutions of this equation can rarely be found in closed form, so that we usually look for numerical approximations of these solutions. As a consequence, the methods for solving the previous equation are usually iterative. So, starting from one initial approximation of a solution x^* of the equation $F(x) = 0$, a sequence $\{x_n\}$ of approximations is constructed such that the sequence $\{\|x_n - x^*\|\}$ is decreasing and a better approximation to the solution x^* is then obtained at every step. Obviously, one is interested in $\lim_n x_n = x^*$.

We can then obtain a sequence of approximations $\{x_n\}$ in different ways, depending on the iterative methods that are applied. Among these, the best known and most used is Newton's method, whose algorithm is

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

It is well known that three types of studies can be carried out when we are interested in proving the convergence of Newton's sequence $\{x_n\}$ to the solution x^*: local, semilocal and global. First, the local study of the convergence is based on imposing conditions on the solution x^*, based on certain conditions on the operator F, and provides the so-called ball of convergence [25] of the sequence $\{x_n\}$, which 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 imposing conditions on the initial approximation x_0, based on certain conditions on the operator F, and provides the so-called domain of parameters [32] 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, based on 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 approaches involve conditions on the operator F. However, requirement of conditions on the solution, on the initial approximation, or on 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 underlying operators, as a consequence of absence of conditions on the initial approximation and on the solution. There is a plethora of studies devoted to the weakening and/or extension of the hypotheses made on the underlying operators. In this monograph, we focus on the analysis of the semilocal convergence of Newton's method.

Three types of conditions are required to obtain semilocal convergence results for Newton's method: conditions on the starting point x_0, conditions on the underlying operator F and conditions that the two proceeding types of conditions. An important feature of the semilocal convergence results is that conclusions about the existence and uniqueness of solution of the equation to be solved can be drawn based on the theoretical result and the initial approximation. This fact makes the choice of the starting points for Newton's method a basic aspect in semilocal convergence studies.

The generalization of Newton's method to Banach spaces is due to the Russian mathematician L. V. Kantorovich, who was the first researcher to study the semilocal convergence of Newton's method in Banach spaces by publishing some several in the mid-twenty century, [49–57], and giving an influential result, known as *the Newton–Kantorovich theorem*. This gave rise to what is now known as *Kantorovich's theory*.

In our monograph *Newton's Method: An Updated Approach of Kantorovich's Theory*, [37], we analyse Kantorovich's theory based on the well-known *majorant principle* developed by Kantorovich, which in turn is based on the concept of *majorizing sequence*. There we present an adapted approach of this theory that includes old results, refines old results, proves the most relevant results and gives alternative approaches that lead to new sufficient semilocal convergence criteria for Newton's method. As we can see in that monograph, if we pay attention to the type of conditions required for the operator F to guarantee the semilocal convergence of Newton's method, there are conditions on F', as well as conditions on F'' or even conditions on successive derivatives of F.

However, if we look at the algorithm of Newton's method, we see that only the first derivative F' of the operator F is involved, so one should try to prove the semilocal convergence of the method by imposing conditions only to F'. If we proceed in this way, then the technique based on majorizing sequences of Kantorovich cannot be used to prove the semilocal convergence of Newton's method in all the situations that can be encountered. So, in the present monograph, we focus our attention on the analysis of the semilocal convergenece of Newton's method under mild differentiability conditions on F' and use a technique based on recurrence relations which is different from that based on majorizing sequences and which was introduced and developed by us over the years. As a consequence, we improve the domains of parameters associated with the Newton–Kantorovich theorem and other existing semilocal convergence results for Newton's method which are obtained under mild differentiability conditions on F'. In addition, center conditions on the operator F' play an important role in the study of the semilocal convergence of Newton's method, since we can improve the domain of starting points when the technique of recurrence relations is used to prove semilocal convergence.

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

We begin the monograph with a quick overview of Newton's method in Chap. 1, presenting a brief history of the method that ends with Kantorovich's theory, where the Newton–Kantorovich theorem is remembered and illustrated with an application to a Hammerstein integral equation. Then, we define what we mean by accessibility of an iterative method, defining the three ways in which it can be seen: basin of attraction, region of accessibility and domain of parameters. We finish the chapter by introducing mild differentiability conditions on F' as generalizations of Kantorovich's condition on F', along with a technique based on recurrence relations, that is used throughout the monograph as an alternative to the majorant principle of Kantorovich, to prove the semilocal convergence of Newton's method.

In Chap. 2, we develop the technique based on recurrence relations to prove the semilocal convergence of Newton's method when F' is Lipschitz continuous in the domain of definition of the operator F and conclude with an application to a Chandrasekhar integral equation.

The first generalization of the condition that F' is Lipschitz continuous is presented in Chap. 3, where we require that F' is Hölder continuous in the domain of definition of the operator F. We do an analysis similar to that given in the previous chapter for the Lipschitz case, complete this analysis with a comparative study involving semilocal convergence results given by other authors, and finish with an application that highlights how the theoretical power of Newton's method is used to draw conclusions about the existence and uniqueness of a solution and about the region in which it is located.

Chapter 4 contains a variant of the Hölder continuity condition on F' discussed in Chap. 3 that includes the Lipschitz and Hölder cases as special ones and leads to a modification of the domain of starting points, obtained previously, coming to a greater applicability of the method.

Chapter 5 introduces what we call the ω-Lipschitz continuous operators and analyzes the semilocal convergence of Newton's method when F' is ω-Lipschitz continuous in the domain of definition of the operator F. This condition includes, besides the Lipschitz and Hölder cases, the case in which F' is a combination of operators such that F' is Lipschitz or Hölder continuous, which often occurs for some nonlinear integral equations of mixed Hammerstein type.

We show in Chaps. 6 and 7 the important role played by the previous conditions when they are centered at the starting point x_0 of Newton's method, which leads to an improvement of the domain of starting points. We complete this analysis by comparing our results with results by other authors and highlight the importance of the domain of parameters associated with a semilocal convergence result. We illustrate the conclusions given applications to conservative problems and mildly nonlinear elliptic equations.

The fact that the conditions imposed to the starting point and to the operator F are independent enables us to can choose the initial approximation inside a domain of starting points depending on the conditions that the two types of hypotheses. In another case, if the two types of conditions are connected, the domain of starting points can be significantly reduced and this is a problem. In Chap. 8, we try to solve this problem by introducing an auxiliary point, different from the starting point, which allows us to eliminate the connection between the conditions required for the starting points and those required for the operator F, and thus recover the domain of starting points.

Applications to nonlinear integral and differential equations are included to motivate the ideas presented and illustrate the results given. In particular, we consider Hammerstein integral equations, conservative problems and elliptic equations, which are solved by discretization.

We have developed all the proofs presented in the monograph for a better understanding of the ideas presented, so that the reading of the monograph follows without difficulty. All the ideas presented in the monograph have been developed by us over the recent years and references to our work as well as to works of other researchers are provided in the bibliography.

Finally, throughout the monograph, we pursue to extend the application of Newton's method from the modification of the domain of starting points. To end, we impose various conditions on the operator involved and used a technique based on recurrence relations that allows us to study the semilocal convergence of Newton's method under mild differentiability conditions on the first derivative of the operator. In this way, Kantorovich's theory for Newton's method has been considerably broadened.

Logroño, La Rioja, Spain
October 2019

José Antonio Ezquerro Fernández
Miguel Ángel Hernández Verón

Contents

1 The Newton-Kantorovich Theorem 1
1.1 Brief History of Newton's Method 2
1.2 Kantorovich's Theory 9
1.2.1 The Newton-Kantorovich Theorem 10
1.2.2 Application to a Hammerstein Integral Equation 11
1.3 Accessibility of Newton's Method 15
1.3.1 Basin of Attraction 16
1.3.2 Region of Accessibility 17
1.3.3 Domain of Parameters 17
1.4 Mild Alternatives to Kantorovich's Conditions 19
1.5 A Technique Based on Recurrence Relations 21

2 Operators with Lipschitz Continuous First Derivative 23
2.1 Convergence Analysis 24
2.1.1 Recurrence Relations 24
2.1.2 Analysis of the Scalar Sequence 26
2.1.3 Proof of the Semilocal Convergence 27
2.2 Application to a Chandrasekhar Integral Equation 31

3 Operators with Hölder Continuous First Derivative 35
3.1 Convergence Analysis 36
3.1.1 Recurrence Relations 36
3.1.2 Analysis of the Scalar Sequence 37
3.1.3 A Semilocal Convergence Result 40
3.2 A Little More 44
3.2.1 Case $p = 0$ 44
3.2.2 Case $p = 1$ 44
3.2.3 Scalar Sequence 45

3.2.4 Starting Points 46
3.2.4.1 Region of Accessibility 46
3.2.4.2 Domain of Parameters 47
3.2.4.3 Comparative Study 47
3.3 Application to a Hammerstein Integral Equation 49
3.3.1 Existence and Uniqueness of a Solution 51
3.3.2 Localization of Solutions 52
3.3.3 Approximation of Solutions 55

4 Operators with Hölder-Type Continuous First Derivative 57
4.1 Convergence Analysis 59
4.1.1 Recurrence Relations 59
4.1.2 Analysis of the Scalar Sequence 61
4.1.3 Semilocal Convergence Result 63
4.2 A Little More 66
4.2.1 Case $p = 0$ 66
4.2.2 Particular Cases 67
4.3 Application to a Hammerstein Integral Equation 67
4.3.1 Existence and Uniqueness of a Solution 67
4.3.2 Localization of Solutions 72
4.3.3 Approximation of Solutions 73

5 Operators with ω-Lipschitz Continuous First Derivative 77
5.1 Convergence Analysis 79
5.1.1 Recurrence Relations 79
5.1.2 Analysis of the Scalar Sequences 81
5.1.3 Semilocal Convergence Result 82
5.2 A Little More 86
5.2.1 On the R-order of Convergence 86
5.2.2 Particular Cases 90
5.3 Application to a Hammerstein Integral Equation 90
5.3.1 Existence and Uniqueness of a Solution 91
5.3.2 Localization of Solutions 92
5.3.3 Approximation of Solution 93

6 Improving the Domain of Starting Points Based on Center Conditions for the First Derivative 95
6.1 Operators with Lipschitz Continuous First Derivative 97
6.1.1 Convergence Analysis 97
6.1.1.1 Recurrence Relations 97
6.1.1.2 Analysis of the Scalar Sequence 99
6.1.1.3 Semilocal Convergence Result 100

6.1.2 Domain of Parameters ... 103
6.1.3 Application to a Conservative Problem ... 107
6.1.3.1 Setting up a Finite Difference Scheme ... 108
6.2 Operators with ω-Lipschitz Continuous First Derivative ... 111
6.2.1 Convergence Analysis ... 112
6.2.1.1 Recurrence Relations ... 112
6.2.1.2 Analysis of the Scalar Sequence ... 114
6.2.1.3 Semilocal Convergence Result ... 115
6.2.2 Domain of Parameters and Applications ... 117
6.2.2.1 Domain of Parameters ... 117
6.2.2.2 Application to a Conservative Problem ... 123
6.2.2.3 Particular Cases ... 125

7 Operators with Center ω-Lipschitz Continuous First Derivative ... 139
7.1 Convergence Analysis ... 143
7.2 Application to an Elliptic Problem ... 148
7.3 Particular Case ... 150

8 Using Center ω-Lipschitz Conditions for the First Derivative at Auxiliary Points ... 155
8.1 Convergence Analysis ... 158
8.2 Particular Cases and Applications to Elliptic Problems ... 162
8.2.1 The Lipschitz Case ... 163
8.2.2 Domain of Parameters ... 163
8.2.3 Region of Accessibility ... 165
8.2.4 Applications ... 167
8.2.5 The Hölder Case ... 170
8.2.6 General Case ... 171

Bibliography ... 175

1 The Newton-Kantorovich Theorem

Solving nonlinear equations is one of the mathematical problems that is frequently encountered in diverse scientific disciplines. Thus, with the notation

$$f(x) = 0,$$

we include the problem of finding unknown quantity x, which can be a real or complex number, a vector, a function, etc., from data provided by the function f, which can be, for example, a real function, a system of equations, a differential equation, an integral equation, etc. Even when f is a real function of a real variable, it is well known that in general it is not possible to solve a nonlinear equation accurately. Instead, iterative techniques are usually employed to obtain approximations of a solution. Among the iterative techniques, Newton's method is undoubtedly the most studied and used in practice. Thus, in order to approximate a solution α of a nonlinear equation $f(x) = 0$, Newton's method constructs, starting from an initial approximation x_0 of α, a sequence of the form

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad n \geq 0. \tag{1.1}$$

Under adequate conditions, the sequence (1.1) converges to the solution α.

Among researchers, it is customary to baptise their discoveries with their own names or with the name of a relevant celebrity in the matter. In the present, the name of the method is linked to the eminent British scientist Isaac Newton. His works at the end of the seventeenth century seem to be the germ of the method that currently bears his name. However, as it is shown in more detail in Sect. 1.1 and references therein, the method is the

J. A. Ezquerro Fernández, M. Á. Hernández Verón, *Mild Differentiability Conditions for Newton's Method in Banach Spaces*, Frontiers in Mathematics,
https://doi.org/10.1007/978-3-030-48702-7_1

fruit of the contributions of a great number of scientists, both before and after Newton's work. The various forms constructions that the method admits constitute another example of the plurality of backgrounds on Newton's method, see [30].

In the mid-twentieth century, the Soviet mathematician Leonid Vitaliyevich Kantorovich extended the study of Newton's method to equations defined in Banach spaces, initially what is currently known as Kantorovich's theory. Combining techniques from functional analysis and numerical analysis, Kantorovich's theory allows us to address numerous nonlinear problems such as solving integral equations, ordinary and partial differential equations, or problems of variational calculus, as it will be detailed throughout this monograph.

1.1 Brief History of Newton's Method

The "paternity" of Newton's method is attributed to Isaac Newton, who described it in several of his works published at the end of the seventeenth century. However, the idea of finding an unknown amount through successive approximations dates back many centuries before Newton. Thus, in antique Greece, techniques to approximate irrational numbers (mostly, π) by rational numbers were known. But, even earlier, 2000 years before Christ, Mesopotamians already knew techniques to approximate the square root of a number. Relevant references are abundant. For example, in [61, p. 42–43], it highlights how the famous tablet YBC 7289 (see Fig. 1.1) from the *Yale Babylonian Collection* shows a square of 30 units of side whose diagonal displays[1] the numbers 1; 24, 51, 10 and 42; 25, 35.

Conversion to the decimal system of the first number is $1.4142129629\ldots$, which matches $\sqrt{2} = 1.4142135623\ldots$ up to the fifth decimal digit. The second number is the product of 30 by the first and is, therefore, the length of the diagonal of the square. So, it seems clear that the Babylonians knew an approximate value for $\sqrt{2}$ and used it in calculations.

Another indication of the Babylonians knew how to approximate irrational numbers appears in tablet VAT6598, which is preserved in the Berlin Museum and is dated in 2000–1700 BC, where the problem of finding the diagonal of a rectangle of height 40 and side 10 is stated among others. In the current notation, the problem amounts to finding

$$\sqrt{40^2 + 10^2} = \sqrt{1700}.$$

[1] The Babylonians used a system of cuneiform numbering with sexadecimal base. Currently, experts in the field write the Babylonian numbers by using a mixture of our notation in base 10 and their notation in base 60. The Babylonian equivalent of the decimal comma is denoted by a point and coma. The rest of the digits are separated by commas. So, the number 5, 51, 13; 2, 30 means $5 \times 60^2 + 51 \times 60 + 13 + 2 \times 1/60 + 30 \times 1/60^2 \simeq 21073.0416$.

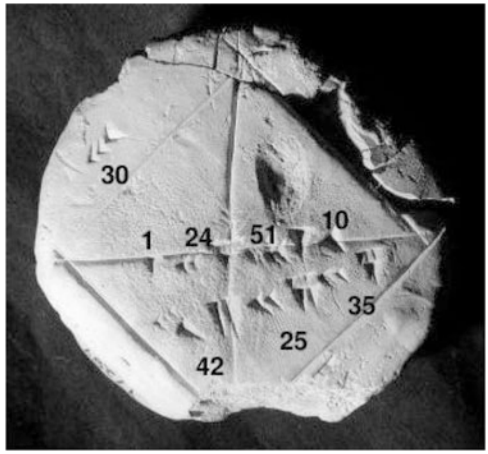

Fig. 1.1 Tablet YBC 7289 of the *Yale Babylonian Collection* (photograph of Bill Casselman)

In the same tablet, the number 41; 15 = 41 + 15/60 is proposed as an approximation. It is not known how this number was obtained or if there is evidence of the use of an iterative method, but some authors [19] mention the fact that this number coincides with the known approximation for a square root

$$\sqrt{h^2 + l^2} \simeq h + \frac{l^2}{2h}$$

for $h = 40$ and $l = 10$.

The proceeding approximation is known as the *formula of Heron* for the calculation of square roots, in which, starting from an initial approximation a of $\sqrt{A}$, the value $(a + A/a)/2$ is proposed as a new approximation. Indeed, for $A = h^2 + l^2$ and $a = h$, the approximation given in the Babylonian tablet coincides with Heron's. Although there are people who attributed the formula of Heron to the Pythagorean Archytas of Tarentum (428–347 BC) or even to Archimedes (282–212 BC), the method appears in the first volume of the *Metrica* that Heron published in the first century. This book, discovered by H. Schöne in 1896 (see [19] for details) shows how Heron estimated the area of a triangle of sides 7, 8 and 9 units, namely $\sqrt{720}$. In the same book, Heron mentions explicitly that a given approximation can be chosen as starting point to obtain best approximations. It seems clear, therefore, that this book contains the first reference of the use of an iterative method.

Now then, was Heron's method original in his time? or was it a technique already known and used by previous civilizations? The answer is in the air, although the majority of researchers of this part of the history of mathematics seem to lean towards the second option, since there is evidence of the use of Babylonian texts by mathematicians and astronomers contemporary with Heron. For example, in the work known as the *Almagest*, Claudius Ptolemy (100–170 AC) cites astronomical data of the time of the Assyrian King Nabonassar (747 BC).

From Heron's formula on techniques for calculating the square root of a number (and, in general, n-th roots) were transmitted and/or rediscovered over centuries and civilizations the seventeenth century. Although there is not much written evidence of what took place during that long period of time, we can find some references on methods for the calculation of n-th roots [19]. We can mention, for example, the Chinese mathematics book par excellence, the *Jiuzhang suanshu*, which translates as *Nine chapters of mathematical art*. There exists a third century version, with reviews of Liu Hui (220–280 AC, approximately), which contains a collection of problems that require the calculation of square and cubic roots. Later, in the fourth century, Theon of Alexandria (335–405 AC, approximately), father of Hypatia, developed a completely geometric method to approximate square roots. In the works of the Persian mathematician Sharaf Al-Din Al-Tusi (1135–1213) one finds the solutions, both algebraic and numerical, of some cubic equations. It seems that Al-Tusi was also the first to calculate the derivative of a third-degree polynomial.

In the work *Raf al-Hijab* of the Arab mathematician Al-Marrakushi Ibn Al-Banna (1256–1321), which one can translate by *Lifting the Veil*, it is shown how to calculate square roots by using series and continued fractions. It seems that Al-Banna was a great collector of the mathematical knowledge of his time, and he shows its versions of the works of earlier Arab mathematicians in his writings.

The problem of finding the n-th root of a number continued to evolve towards the more general problem of finding the roots of a polynomial equation and, even of a transcendental equation (for example, Kepler's equation). Starting with the fifteenth century, the problem bifurcated into several lines (algebraic solutions of polynomial equations, approximate solutions by using fixed-point iterations, approximations by continuous fractions, etc.). A detailed analysis of the historical development of these problems is beyond the scope of this monograph, so we refer the interested reader to one of the specialized textbooks, such as [19], or the paper [79].

Focusing on the birth of Newton's method, we can highlight the antecedent work of the French mathematician François Viète (1540–1603), who developed an ambitious project aimed at positive solutions of polynomial equations of degree 2 to 6 of generic form. Viète was the first to represent the parameters of an equation by letters, not only the unknowns. The method employed by Viète ("specious logistic" or "art of the calculation on symbols") was rooted in the Greek geometric tradition. The method of Viète, written in an archaic language and with tedious notations, did not have continuation, soon become ignored and was displaced by the Cartesian geometry. However, Viète was the first to understand the relationship between roots and coefficients of a polynomial and to try to use algebra.

It seems that the work of Viète was what inspired Isaac Newton (1643–1727) to develop his method of solving equations. The first written reference to Newton's method is found in *De analysi per aequationes numero terminorum infinitas*, in a letter written to his colleagues Barrow and Collins in 1669, which however was not published til 1711. Two years after writing this letter, in 1671, Newton developed his method in *De metodis fluxionum et serierum infinitarum*. Again, the publication of this work was delayed and it was not til 1736 that a translation was published under the title *Method of Fluxions*.

To get an idea of how Newton worked, we can illustrate his method with the same example that he considered, the equation $x^3 - 2x - 5 = 0$. Newton argued as follows:

> By estimation, we see that the solution is near 2. Taking $x = 2 + \varepsilon$ and substituting in the equation, we obtain:
>
> $$\varepsilon^3 + 6\varepsilon^2 + 10\varepsilon - 1 = 0. \tag{1.2}$$
>
> Ignoring the terms $\varepsilon^3 + 6\varepsilon^2$ because ε is small, we have $10\varepsilon - 1 \simeq 0$, or $\varepsilon = 0.1$. Then, $x = 2.1$ is a better approximation of the solution than the initial one. Doing now $\varepsilon = 0.1 + \nu$ and substituting in (1.2), we get
>
> $$\nu^3 + 6.3\nu^2 + 11.23\nu + 0.061 = 0.$$
>
> Ignoring again the terms in ν of degree greater than or equal to two, we have $\nu \simeq -0.054$ and, therefore, $x = 2.046$ is an approximation that improves the previous ones. Newton indicated that the process can be repeated as many times as necessary.

Thus Newton's idea consists of adding a correcting term to a given initial approximation. To obtain this approximation, we truncate Newton's binomial at the second term in expressions of the type

$$(a + \varepsilon)^n \simeq a^n + na^{n-1}\varepsilon.$$

So, to obtain the approximate value of ε, we only have to solve a linear equation.

Writing the problem in modern notation and denoting $p(x) = x^3 - 2x - 5$, we see that the new approximation is

$$2 - \frac{p(2)}{p'(2)} = 2 + \frac{1}{10} = 2.1,$$

which corresponds to the well-known formulation of Newton's method (1.1) when $f(x)$ is the polynomial $p(x)$. However, there is no evidence that Newton used differential calculus or that he expressed the process as an iterative method in the sense that one approximation can be considered as the starting point of the next approximation. Furthermore, Newton used "his method" only to solve polynomial equations. Therefore, Newton's idea of his method is far from what we have today.

The idea of iteration is attributed to Joseph Raphson (1648–1715) (see, for example, [19, 79]), who also simplified the operational aspect of Newton's technique. In 1690, Raphson published the treatise *Analysis aequationum universalis*, in which he gave explicit formulas for the corrector term for some particular cases of equations. In particular, he calculated the corrector terms for the equations $x^3 - r = 0$ and $x^3 - px - q = 0$ and found that they are

$$\frac{r - x_0^3}{3x_0^2} \quad \text{and} \quad \frac{q + px_0 - x_0^3}{3x_0^2 - p},$$

where x_0 is the initial approximation. Notice that Raphson published his work 46 years before *Newton's Method of Fluxions*. However, Raphson was the first to recognize that Newton's method was already known in the scientific circles of that time and that his method was an improved version.

The contribution of Raphson has been historically recognized and many authors call the method the Newton-Raphson method. However, in the works of Raphson, we cannot appreciate the connection existing between the corrector term, the function that defines the equation, and its derivative.

The incorporation of the differential calculus is due to Thomas Simpson (1710–1761). As we can see in [79], Simpson, in his work *Essays on Mathematics*, published in 1740, was the one who established the method as it is currently known, except for the notational aspects (Simpson explained in a rhetoric form how to obtain the successive approximations). In addition, Simpson extended the process to arbitrary function's, not only polynomials.

On the occasion of certain observations that use infinite series, Newton seemed to be concerned with the concept of convergence, but he did not provide any solution to this problem. The first time the convergence of Newton's method is discussed in the 1768 *Traité de la résolution des équations en general* of Jean Raymond Mourraille (1720–1808). Despite the fact that it contained novel ideas, most of Mourraille's work went unnoticed.

Contrary to Newton and Raphson, Mourraille emphasized the geometric aspect of Newton's method, justifying why this method is also known as the tangent method. Mourraille used the geometric representation of Newton's method to explain the behavior of the iterative sequences it generates. Besides, Mourraille observes by first time that, depending on the starting point chosen, the sequence generated by the method can converge to any of the roots of the equation, oscillate, approach to infinity or a limit that is not a solution of the equation. Finally, Mourraille also showed that the convergence can be more or less fast, but he only indicated this in quantitative form.

Later, Joseph-Louis Lagrange (1736–1813), in his *Traité de la résolution des équations numériques de tous les degrés*, published in 1808 [40], says that the method attributed to Newton is usually employed to solve numerical equations. However, he warns that this method can be only used for equations that are already "almost solved", in the sense that a good approximation of the solution is reached. Moreover, he raises questions about the accuracy of each new iteration and observes that the method may run into difficulties in the case of multiple roots or roots that are very close to one another.

Jean Baptiste Joseph Fourier (1768–1830) was the first to analyze the rate of convergence of Newton's method in a work entitled *Question d'analyse algébraique* (1818), [40]. In this work, Fourier expressed the method in the current notation and baptized it as *la méthode newtonienne*, making explicit reference to the works of Newton, Raphson and Lagrange. Perhaps, Fourier is the "originator" of the lack of recognition for Simpson's work.

The next important mathematician to study Newton's method was Augustin Louis Cauchy (1789–1857), who started to work on it in 1821, but did not give a satisfactory