Article

On the Accessibility of Newton's Method under a Hölder Condition on the First Derivative

José Antonio Ezquerro * and Miguel Ángel Hernández-Verón

Department of Mathematics and Computation, University of La Rioja, Calle Luis de Ulloa s/n, 26004 Logroño, Spain; E-Mail: mahernan@unirioja.es

* Author to whom correspondence should be addressed; E-Mail: jezquer@unirioja.es; Tel.: +34-941-299-460.

Abstract: We see how we can improve the accessibility of Newton's method for approximating a solution of a nonlinear equation in Banach spaces when a center Hölder condition on the first derivative is used to prove its semi-local convergence.

---

1. Introduction

It is often necessary in many scientific and engineering problems to find a root of a nonlinear equation. To solve this equation, we can use iterative methods. To give sufficient generality to the problem of approximating a solution of a nonlinear equation by iterative methods, we are concerned, in this work, with the problem of approximating a locally-unique solution x^* of the 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, so that many scientific and engineering problems can be written as a nonlinear equation in Banach spaces.

Newton's method is probably the best-known iterative method, and it is well known that it converges quadratically. Newton's method for solving scalar equations was extended to nonlinear equations in Banach spaces by Kantorovich in 1948 [1] in the following way:

$$x_0 \text{ given in } \Omega, \quad x_n = x_{n-1} - [F'(x_{n-1})]^{-1}F(x_{n-1}), \quad n \in \mathbb{N}, \tag{1}$$

where F' is the Fréchet derivative of the operator F. For this reason, many authors refer to it as the Newton–Kantorovich method.

Kantorovich proved the semi-local convergence of Newton's method under the hypothesis that the operator involved F is twice differentiable Fréchet with bounded second derivative:

> There exists a constant $C \geq 0$, such that $\|F''(x)\| \leq C$ for $x \in \Omega$.

It is well known that this condition can be replaced by a Lipschitz condition on the first derivative F' of the operator involved [2]:

> There exists a constant $L \geq 0$, such that $\|F'(x) - F'(y)\| \leq L\|x - y\|$ for $x, y \in \Omega$.

In view of the applications, numerous papers have recently appeared where the convergence of the method is proven under different conditions for the derivatives of the operator F. A known variant is the study of the convergence of the method under a Hölder condition on the first derivative [3–5]:

> There exist two constants $K \geq 0$ and $p \in (0, 1]$ such that $\|F'(x) - F'(y)\| \leq K\|x - y\|^p$
> for $x, y \in \Omega$.

In order to improve the accessibility of Newton's method under the last condition, we can use different strategies that change the condition. In this work, we use a similar one as that used by Argyros in [6] for the Lipschitz condition on the first derivative F', which consists of noticing that, as a consequence of the last condition being satisfied in Ω, we have, for the starting point x_0, that:

> There exist two constants $K_0 \geq 0$ and $p \in (0, 1]$, such that $\|F'(x) - F'(x_0)\| \leq K_0\|x - x_0\|^p$
> for $x \in \Omega$,

with $K_0 \leq K$. We then say that F' is center Hölder in x_0.

In this paper, we focus our attention on the analysis of the semi-local convergence of Sequence (1), which is based on demanding the last condition to the initial approximation x_0 and provides the so-called domain of parameters corresponding to the conditions required for the initial approximation that guarantee the convergence of Sequence (1) to the solution x^*.

In this work, we carry out an analysis of the domain of parameters for Newton's method under the last two conditions for F' and use a technique based on recurrence relations. As a consequence, we improve the domains of parameters associated with the semi-local convergence result given for Newton's method by Hérnández in [4].

We prove in this paper that center conditions on the first derivative of the operator involved in the solution of nonlinear equations play an important role in the semi-local convergence of Newton's method, since we can improve the accessibility of Newton's method from them.

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

2. Preliminaries

The best-known semi-local convergence result for Newton's method under a Hölder condition on the first derivative of the operator involved, when the technique of recurrence relations is used to prove it, is that given by Hernández in [4], which is established under the following conditions:

(A1) There exists $\Gamma_0 = [F'(x_0)]^{-1} \in \mathcal{L}(Y, X)$, for some $x_0 \in \Omega$, with $\|\Gamma_0\| \leq \beta$ and $\|\Gamma_0 F(x_0)\| \leq \eta$, where $\mathcal{L}(Y, X)$ is the set of bounded linear operators from Y to X.

(A2) There exist two constants $K \geq 0$ and $p \in (0, 1]$, such that $\|F'(x) - F'(y)\| \leq K\|x - y\|^p$ for $x, y \in \Omega$.

(A3) If $\xi(p)$ is the unique zero of the auxiliary function:

$$\phi(x; p) = (1 + p)^p(1 - x)^{1+p} - x^p, \quad p \in (0, 1], \tag{2}$$

in the interval $\left(0, \frac{1}{2}\right]$, it is satisfied that $\hbar = K\beta\eta^p \leq \xi(p)$ and $B(x_0, R) \subset \Omega$, where $R = \frac{(1+p)(1-\hbar)}{(1+p)-(2+p)\hbar}\eta$.

Theorem 1. *(Theorem 2.1 of [4]) Let $F : \Omega \subseteq X \longrightarrow Y$ be a continuously-differentiable operator defined on a nonempty open convex domain Ω of a Banach space X with values in a Banach space Y. Suppose that Conditions (A1)(A3) are satisfied. Then, Newton's sequence, given by (1), converges to a solution x^* of the equation $F(x) = 0$, starting at x_0, and $x_n, x^* \in \overline{B(x_0, R)}$, for all $n = 0, 1, 2, \ldots$*

Note that Condition (A1), required for the initial approximation x_0, defines the parameters β and η, and Condition (A2), required for the operator F, defines the parameter K fixed for all point of Ω. Observe that, every point $z \in \Omega$, such that the operator $[F'(z)]^{-1}$ exists with $\|[F'(z)]^{-1}\| \leq \beta$ and $\|[F'(z)]^{-1}F(z)\| \leq \eta$, has associated the pair $(K, \beta\eta^p)$ of the xy-plane, where $x = K$ and $y = \beta\eta^p$. In addition, fixed $p \in (0, 1]$, if we consider the set:

$$\mathcal{D}_{\mathrm{T}} = \left\{(x, y) \in \mathbb{R}^2 : xy \leq \xi(p)\right\},$$

we can observe that every point z, such that its associated pair $(K, \beta\eta^p)$ belongs to $\mathcal{D}_{\mathrm{T}}$, can be chosen as the starting point for Newton's method, so that the method converges to a solution x^* of the equation $F(x) = 0$ when it starts at z. The set $\mathcal{D}_{\mathrm{T}}$ is then called the domain of parameters associated with Theorem 1 and can be drawn by choosing $x = K$, $y = \beta\eta^p$ and coloring the region of the xy-plane whose points satisfy the condition $\hbar \leq \xi(p)$ (namely, $xy \leq \xi(p)$) of the Theorem 1. In Figure 1, we see the domain of parameters $\mathcal{D}_{\mathrm{T}}$ (blue region).

In relation to the above, we can think that the larger the size of the domain of parameters is, the more possibilities we have for choosing good starting points for Newton's method. As a consequence, we are interested in $\mathcal{D}_{\mathrm{T}}$ being as big as possible, since this fact allows us to find a greater number of good starting points for Newton's method.

3. Semi-Local Convergence of Newton's Method

To improve the semi-local convergence of Newton's method from increasing the domain of parameters $\mathcal{D}_\mathrm{T}$, we consider a procedure that consists of observing that, as a consequence of Condition (A2), once $x_0 \in \Omega$ is fixed, the condition:

$$\|F'(x) - F'(x_0)\| \leq K_0 \|x - x_0\|^p, \quad x \in \Omega, \tag{3}$$

is satisfied with $K_0 \leq K$. Then, by considering jointly the parameters K and K_0, we can relax the semi-local convergence conditions of Newton's method given in Theorem 1 and obtain a larger domain of parameters.

Now, we present a semi-local convergence result for Newton's method under Condition (A1) for the starting point x_0 and Condition (A2) for the first derivative F'. Note that Condition (3) follows from Condition (A2) for the starting point x_0. In addition, we obtain a semi-local convergence result by combining Conditions (A2) and (3), which allows increasing the domain of parameters $\mathcal{D}_\mathrm{T}$, so that the possibility of choosing good starting points for Newton's method is increased, as we can see later in the applications. In particular, we study the convergence of Newton's method to a solution of the equation $F(x) = 0$ under certain conditions for the pair (F, x_0). From some real parameters, a system of four recurrence relations is constructed in which two sequences of positive real numbers are involved. The convergence of Newton's method is then guaranteed from them.

3.1. Auxiliary Scalar Sequences

From Conditions (A1)–(A2), we define $\gamma = \varepsilon \hbar$, $b_0 = \hbar$ and $\delta = \frac{\gamma}{(1+p)(1-\gamma)}$, where $\varepsilon = \frac{K_0}{K} \in [0, 1]$. Now, we define $b_1 = \frac{\delta^p}{1-\gamma} b_0$ and:

$$a_n = \frac{b_n}{(1+p)(1-b_n)}, \quad n \geq 1, \tag{4}$$

$$b_n = \frac{b_{n-1} a_{n-1}^p}{1 - b_{n-1}}, \quad n \geq 2. \tag{5}$$

Observe that we consider the case $b_0 > 0$, since if $b_0 = 0$, a trivial problem results, as the solution of the equation $F(x) = 0$ is x_0.

Next, we prove the following four recurrence relations for Sequences (1), (4) and (5):

$$\|\Gamma_1\| = \|[F'(x_1)]^{-1}\| \leq \frac{\|\Gamma_0\|}{1 - \gamma}, \tag{6}$$

$$\|x_2 - x_1\| \leq \delta \|x_1 - x_0\|, \tag{7}$$

$$K\|\Gamma_1\|\|x_2 - x_1\|^p \leq b_1, \tag{8}$$

$$\|x_2 - x_0\| \leq (1 + \delta)\|x_1 - x_0\|, \tag{9}$$

provided that:

$$x_1 \in \Omega \quad \text{and} \quad \gamma < 1. \tag{10}$$

If $x_1 \in \Omega$, then:

$$\|I - \Gamma_0 F'(x_1)\| \leq \|\Gamma_0\|\|F'(x_0) - F'(x_1)\| \leq K\beta\eta^p = \hbar < 1.$$

Then, by the Banach lemma on invertible operators, it follows that the operator Γ_1 exists and:

$$\|\Gamma_1\| \leq \frac{\|\Gamma_0\|}{1 - \|I - \Gamma_0 F'(x_1)\|} \leq \frac{\|\Gamma_0\|}{1 - \gamma}.$$

After that, from Taylor's series and Sequence (1), we have:

$$\|F(x_1)\| = \left\| \int_0^1 (F'(x_0 + \tau(x_1 - x_0)) - F'(x_0))(x_1 - x_0) \, d\tau \right\| \leq \frac{K_0 \eta^p}{1 + p} \|x_1 - x_0\|.$$

Thus,

$$\|x_2 - x_1\| \leq \|\Gamma_1\| \|F(x_1)\| \leq \delta \|x_1 - x_0\|$$

$$\|\Gamma_1\| \|x_2 - x_1\|^p \leq \frac{K \delta^p}{1 - \gamma} \|\Gamma_0\| \|x_1 - x_0\|^p \leq b_1,$$

$$\|x_2 - x_0\| \leq \|x_2 - x_1\| + \|x_1 - x_0\| \leq (1 + \delta) \|x_1 - x_0\| < \left(1 + \frac{\delta}{1 - a_1} \right) \eta = R,$$

provided that $\delta \leq 1$ and $a_1 < 1$.

Now, we prove in the same way as above the following four recurrence relations for Sequences (1), (4) and (5):

$$\|\Gamma_2\| = \|[F'(x_2)]^{-1}\| \leq \frac{\|\Gamma_1\|}{1 - b_1}, \tag{11}$$

$$\|x_3 - x_2\| \leq a_1 \|x_2 - x_1\|, \tag{12}$$

$$K\|\Gamma_2\| \|x_3 - x_2\|^p \leq b_2, \tag{13}$$

$$\|x_3 - x_0\| \leq (1 + \delta(1 + a_1)) \|x_1 - x_0\|, \tag{14}$$

provided that:

$$x_2 \in \Omega \qquad \text{and} \qquad b_1 < 1. \tag{15}$$

In addition, we generalize the last recurrence relations to every point of Sequence (1), so that we can guarantee that (1) is a Cauchy sequence from them. For this, we analyze the scalar sequences defined in (4) and (5) in order to prove later the semi-local convergence of Sequence (1). For this, it suffices to see that (1) is a Cauchy sequence and (10) and (15) are true for all x_n and b_{n-1} with $n \geq 3$. We begin by presenting a technical lemma.

Lemma 2. *If* $\gamma \leq \frac{1+p}{2+p}$ *and* b_1 *is such that*

$$b_1 < \frac{1+p}{2+p} \qquad \text{and} \qquad b_1 + a_1^p < 1, \tag{16}$$

then:

(a) *the sequences* $\{b_n\}$ *and* $\{a_n\}$ *are strictly decreasing,*

(b) $a_n < 1$ *and* $b_n < 1$, *for all* $n \geq 1$.

If $b_1 = 1 - a_1^p < \frac{1+p}{2+p}$, *then* $a_n = a_1 < 1$ *and* $b_n = b_1 < 1$ *for all* $n \geq 2$.

Proof. We first consider the case in which b_1 satisfies (16). Item (a) is proven by mathematical induction on n. As $b_1 + a_1^p < 1$, then $b_2 < b_1$ and $a_2 < a_1$. If we now suppose that $b_j < b_{j-1}$ and $a_j < a_{j-1}$, for all $i = 2, 3, \ldots, n$, then:

$$b_{n+1} = \frac{b_n a_n^p}{1 - b_n} < \frac{b_n a_1^p}{1 - b_1} < b_n \qquad \text{and} \qquad a_{n+1} = \frac{b_{n+1}}{(1 + p)(1 - b_{n+1})} < a_n.$$

As a result, the sequences $\{a_n\}$ and $\{b_n\}$ are strictly decreasing for all $n \geq 2$.

To see Item (b), we have $a_n < a_1 < 1$ and $b_n < b_1 < 1$, for all $n \geq 2$, by Item (a) and the conditions given in (16).

Second, if $b_1 = 1 - a_1^p$, then $b_n = b_1 = 1 - a_1^p < 1$, for all $n \geq 2$. Moreover, if $b_1 < \frac{1+p}{2+p}$, then we have $a_n = a_1 < 1$, for all $n \geq 2$. $\square$

3.2. Main Result

We now give a semi-local convergence result for Newton's method from a modification of the convergence conditions required in Theorem 1. Therefore, we consider Conditions (A1), (A2) and the modification of Condition (A3) given by:

(A3b) $\gamma \leq \dfrac{1+p}{2+p}$, b_1 satisfies (16) and $B(x_0, R) \subset \Omega$, where $R = \left(1 + \frac{\delta}{1 - a_1}\right) \eta$.

Remember that Condition (3) follows from Condition (A2) for the starting point x_0.

Theorem 3. *Let $F : \Omega \subseteq X \longrightarrow Y$ be a continuously-differentiable operator defined on a nonempty open convex domain Ω of a Banach space X with values in a Banach space Y. Suppose that Conditions (A1), (A2) and (A3b) are satisfied. Then, Newton's sequence, given by (1), converges to a solution x^* of the equation $F(x) = 0$, starting at x_0, and $x_n, x^* \in \overline{B(x_0, R)}$, for all $n = 0, 1, 2, \ldots$.*

Proof. We begin by proving the following four items for Sequences (1), (4) and (5) and $n \geq 3$:

(I) There exists $\Gamma_{n-1} = [F'(x_{n-1})]^{-1}$ and $\|\Gamma_{n-1}\| \leq \dfrac{\|\Gamma_{n-2}\|}{1 - b_{n-2}}$,

(II) $\|x_n - x_{n-1}\| \leq a_{n-2}\|x_{n-1} - x_{n-2}\|$,

(III) $K\|\Gamma_{n-1}\|\|x_n - x_{n-1}\|^p \leq b_{n-1}$,

(IV) $x_n \in \Omega$.

Observe that $x_1 \in \Omega$, since $\eta < R$. Moreover, from (6), (7), (8) and (9), it follows that $x_2 \in \Omega$. Furthermore, from (11), (12), (13) and (14), we have that Items (I), (II), (III) and (IV) hold for $n = 3$. If we now suppose that Items (I), (II) and (III) are true for some $n - 1$, it follows, by analogy to the case where $n = 3$ and induction, that Items (I), (II) and (III) also hold for n. Notice that $b_n < 1$ for all $n \geq 1$. Now, we prove (IV). Therefore,

$$
\begin{aligned}
\|x_n - x_0\| \quad &\leq \quad \|x_n - x_{n-1}\| + \|x_{n-1} - x_{n-2}\| + \cdots + \|x_1 - x_0\| \\
&\overset{\text{(II)}}{\leq} \quad \left(1 + \sum_{i=1}^{n-2}\left(\prod_{j=1}^{i} a_j\right)\right)\|x_2 - x_1\| + \|x_1 - x_0\| \\
&< \quad \left(1 + \sum_{i=1}^{n-2} a_1^i\right)\|x_2 - x_1\| + \|x_1 - x_0\| \\
&< \quad \frac{1}{1 - a_1}\|x_2 - x_1\| + \|x_1 - x_0\| \\
&\leq \quad \left(1 + \frac{\delta}{1 - a_1}\right)\|x_1 - x_0\| \\
&\leq \quad R
\end{aligned}
$$

and $x_n \in B(x_0, R)$. As $B(x_0, R) \subset \Omega$, then $x_n \in \Omega$ for all $n \geq 0$. Note that the conditions given in (15) are satisfied for all x_n and b_{n-1} with $n \geq 3$.

Next, we prove that $\{x_n\}$ is a Cauchy sequence. For this, we follow an analogous procedure to the latter. Therefore, for $m \geq 2$ and $n \geq 2$, we have:

$$
\begin{aligned}
\|x_{n+m} - x_n\| \quad &\leq \quad \|x_{n+m} - x_{n+m-1}\| + \|x_{n+m-1} - x_{n+m-2}\| + \cdots + \|x_{n+1} - x_n\| \\
&\overset{\text{(II)}}{\leq} \quad \sum_{i=n-1}^{n+m-2}\left(\prod_{j=1}^{i} a_j\right)\|x_2 - x_1\| \\
&\overset{\text{Lemma } 2\,(a)}{<} \quad \left(\sum_{i=n-1}^{n+m-2} a_1^i\right)\|x_2 - x_1\| \\
&= \quad \left(\sum_{i=0}^{m-1} a_1^{n+i-1}\right)\|x_2 - x_1\| \\
&= \quad \frac{1 - a_1^m}{1 - a_1}\, a_1^{n-1}\|x_2 - x_1\|.
\end{aligned}
$$

Thus, $\{x_n\}$ is a Cauchy sequence.

After that, we prove that x^* is a solution of equation $F(x) = 0$. As $\|\Gamma_n F(x_n)\| \to 0$ when $n \to \infty$, if we take into account that:

$$
\|F(x_n)\| \leq \|F'(x_n)\|\|\Gamma_n F(x_n)\|
$$

and $\{\|F'(x_n)\|\}$ is bounded, since:

$$
\|F'(x_n)\| \leq \|F'(x_0)\| + K_0\|x_n - x_0\|^p \leq \|F'(x_0)\| + K_0 R^p,
$$

it follows that $\|F(x_n)\| \to 0$ when $n \to \infty$. As a consequence, we obtain $F(x^*) = 0$ by the continuity of F in $\overline{B(x_0, R)}$. $\square$

4. Accessibility of Newton's Method

The accessibility of an iterative method is analyzed from the set of possible starting points that guarantee the convergence of the iterative method when it starts at them. As we have indicated, the

set of starting points that guarantees the convergence of the iterative method is related to the domain of parameters associated with a result of semi-local convergence of the iterative method.

Next, we study the domain of parameters associated with Theorem 1 and compare it with that associated with Theorem 3. To guarantee the convergence of Newton's method from Theorem 3, the following three conditions must be satisfied:

$$\gamma \leq \frac{1+p}{2+p}, \qquad b_1 < \frac{1+p}{2+p} \qquad \text{and} \qquad b_1 + \frac{b_1^p}{(1+p)^p(1-b_1)^p} < 1, \tag{17}$$

where $p \in (0, 1]$.

From the auxiliary function given in (2), the third condition of (17) can be written as:

$$\phi(b_1; p) > 0. \tag{18}$$

Observe that $\phi(x; p)$ is a non-increasing and convex function and $\phi(0; p) = (1+p)^p > 0$ and $\phi\left(\frac{1}{2}; p\right) \leq 0$, for all $p \in (0, 1]$. Besides, if $p = 1$, the unique zero of $\phi(x; 1)$ in the interval $\left(0, \frac{1}{2}\right]$ is $\frac{1}{2}$. If we now denote, for a fixed $p \in (0, 1]$, the unique zero of $\phi(x; p)$ in $\left(0, \frac{1}{2}\right]$ by $\xi(p)$ and demand $b_1 < \xi(p)$, then Condition (18) holds. Moreover, since $\xi(p) \leq \frac{1}{2}$, the second condition of (17) is satisfied. Now, as:

$$b_1 = \frac{\delta^p b_0}{1-\gamma} = \frac{\gamma^p b_0}{(1+p)^p(1-\gamma)^{1+p}},$$

the second and third conditions of (17) are satisfied, provided that:

$$b_0 < \xi(p)\frac{(1+p)^p(1-\gamma)^{1+p}}{\gamma^p}.$$

After that, we write the first condition of (17) as:

$$\hbar \leq \frac{1+p}{\varepsilon(2+p)} \tag{19}$$

and take into account that the second and third conditions of (17) are satisfied if:

$$\hbar < \xi(p)\frac{(1+p)^p(1-\varepsilon\hbar)^{1+p}}{\varepsilon^p\hbar^p}, \tag{20}$$

since $b_0 = \hbar$ and $\gamma = \varepsilon\hbar$. In addition, Condition (20) is equivalent to:

$$\varpi(\hbar) = \varepsilon^p\hbar^{1+p} - \xi(1+p)^p(1-\varepsilon\hbar)^{1+p} < 0.$$

Furthermore, $\varpi'(\hbar) \geq 0$, so that $\varpi(\hbar)$ is a nondecreasing function for all $\hbar \geq 0$, and $\varpi(0) \leq 0$ and $\varpi\left(\frac{1+p}{\varepsilon(2+p)}\right) > 0$. Therefore, Condition (19) is satisfied if Condition (20) holds, and we can then give the following result.

Corollary 4. *Let $F : \Omega \subseteq X \longrightarrow Y$ be a continuously-differentiable operator defined on a nonempty open convex domain Ω of a Banach space X with values in a Banach space Y. Suppose that (A1)–(A2) are satisfied. If Condition (20), where $\xi(p)$ is the unique zero of function (2) in the interval $\left(0, \frac{1}{2}\right]$, is satisfied and $B(x_0, R) \subset \Omega$, where $R = \left(1 + \frac{\upsilon}{1-\nu}\right)\eta$, $\upsilon = \frac{\varepsilon\hbar}{(1+p)(1-\varepsilon\hbar)}$, $\nu = \frac{\vartheta}{(1+p)(1-\vartheta)}$ and $\vartheta = \frac{\hbar\upsilon^p}{1-\varepsilon\hbar}$, then Newton's sequence, given by (1), converges to a solution x^* of the equation $F(x) = 0$, starting at x_0, and $x_n, x^* \in \overline{B(x_0, R)}$, for all $n = 0, 1, 2, \ldots$*

From the last result, we can define the following domains of parameters:

$$\mathcal{D}_{\mathrm{C}}^{\varepsilon} = \left\{ (x, y) \in \mathbb{R}^2 : xy \leq \xi(p) \frac{(1+p)^p (1 - \varepsilon xy)^{1+p}}{\varepsilon^p x^p y^p} \right\},$$

where $\varepsilon = \frac{K_0}{K} \in [0, 1]$, $p \in (0, 1]$ and $\xi(p)$ is the unique zero of Function (2) in the interval $\left(0, \frac{1}{2}\right]$.

Next, we compare the conditions required for the semi-local convergence of Newton's method in Theorem 1 and Corollary 4. In Figure 1, we see that $\mathcal{D}_{\mathrm{T}} \subset \mathcal{D}_{\mathrm{C}}^{\varepsilon}$ for $p = \frac{1}{10}, \frac{1}{5}, \frac{2}{5}, \frac{4}{5}$, so that we can guess that the smaller the quantity $\varepsilon = \frac{K_0}{K} \in [0, 1]$ is, the larger the domain of parameters is: orange for $\varepsilon = 0.1$, green for $\varepsilon = 0.2$, red for $\varepsilon = 0.4$ and yellow for $\varepsilon = 0.8$. Note that, if $\varepsilon \to 1$, the domain of parameters associated with Corollary 4, $\mathcal{D}_{\mathrm{C}}^{\varepsilon}$, tends to be that obtained by Theorem 1 (blue region). As a consequence,

$$\mathcal{D}_{\mathrm{T}} = \mathcal{D}_{\mathrm{C}}^{1} = \mathcal{D}_{\mathrm{C}}^{\varepsilon_j} \subset \mathcal{D}_{\mathrm{C}}^{\varepsilon_{j-1}} \subset \cdots \subset \mathcal{D}_{\mathrm{C}}^{\varepsilon_0} \qquad \text{for} \qquad \varepsilon_0 < \cdots < \varepsilon_{j-1} < \varepsilon_j = 1.$$
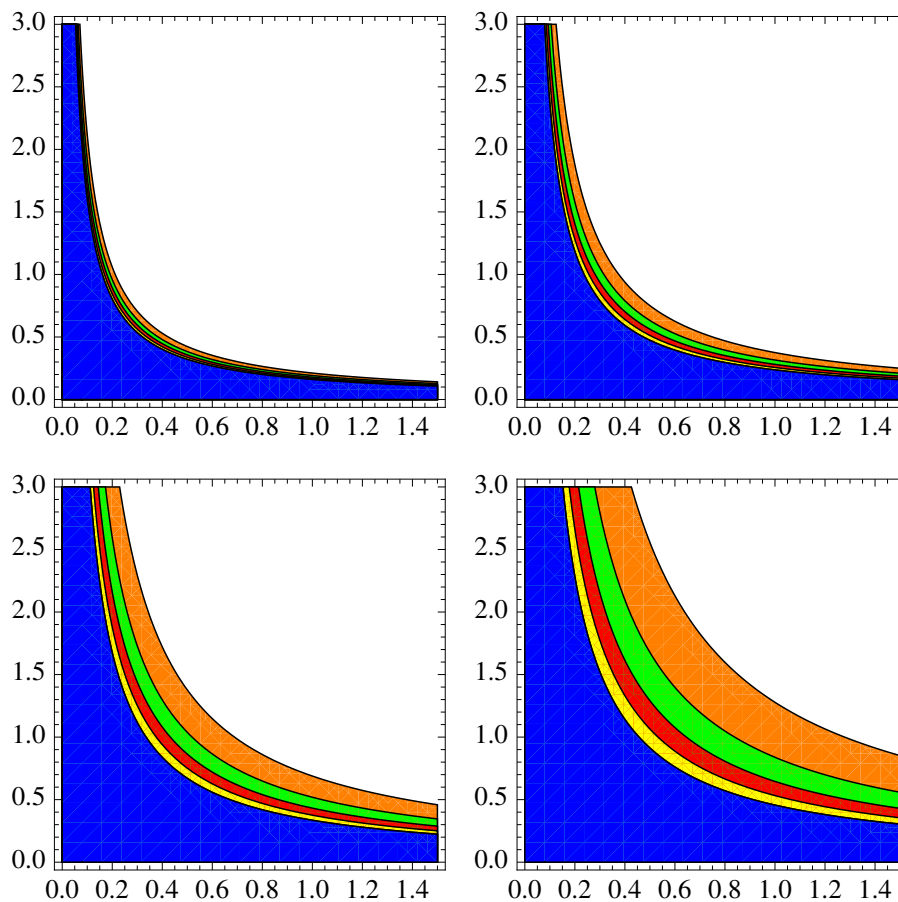

Figure 1. Domains of parameters of Newton's method associated with Theorem 1 (blue region) and Corollary 4 (orange for $\varepsilon = 0.1$, green for $\varepsilon = 0.2$, red for $\varepsilon = 0.4$ and yellow for $\varepsilon = 0.8$) for $p = \frac{1}{10}, \frac{1}{5}, \frac{2}{5}, \frac{4}{5}$.

On the other hand, in Figure 2, we observe the relationship between the domains of parameters associated with Theorem 1 (gray region) and Corollary 4 (magenta region) from the variability of ε and for four different values of p: $p = \frac{1}{10}, \frac{1}{5}, \frac{2}{5}, \frac{4}{5}$. As we can see, the domain associated with Corollary 4 is always larger, for all $\varepsilon \in [0, 1]$, than that associated with Theorem 1.
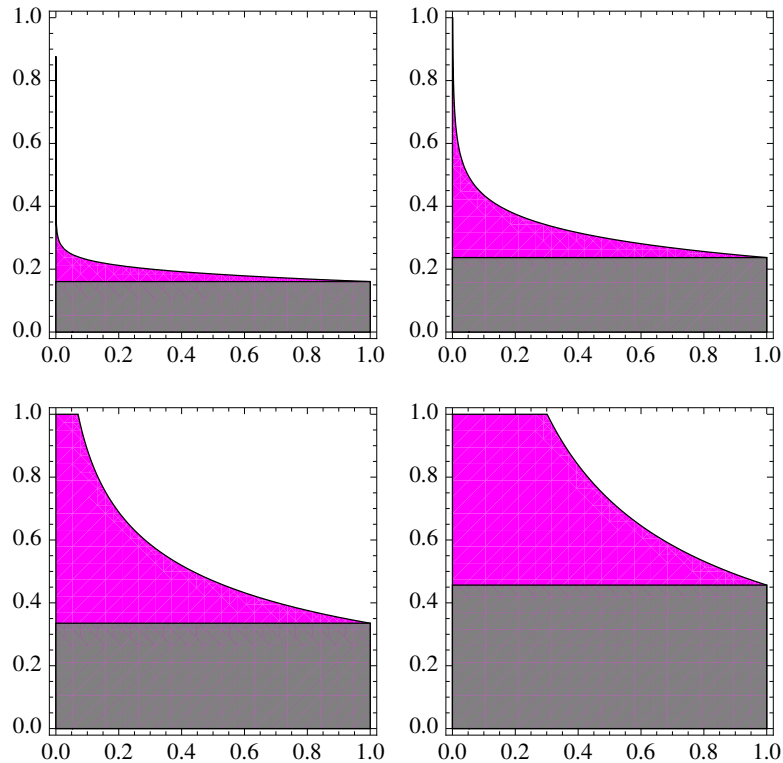
Figure 2. Domains of parameters of Newton's method associated with Corollary 4 (magenta region) and Theorem 1 (gray region) for $p = \frac{1}{10}, \frac{1}{5}, \frac{2}{5}, \frac{4}{5}$.

In addition, we prove analytically in the following what we have just seen graphically. First, we prove that $\mathcal{D}_{\mathrm{T}} \subset \mathcal{D}_{\mathrm{C}}^{\varepsilon}$, for each $p \in (0,1]$ and $\varepsilon \in [0,1]$. For this, if $(K, \beta\eta^p) \in \mathcal{D}_{\mathrm{T}}$, then $\hbar \leq \xi(p) \leq \frac{1}{2} \leq \frac{1}{2\varepsilon} \leq \frac{1+p}{\varepsilon(2+p)}$, since $p \in (0,1]$ and $\varepsilon \in [0,1]$. Moreover, $\hbar \leq \xi(p) \leq \xi(p)\frac{(1+p)^p(1-\varepsilon\hbar)^{1+p}}{\varepsilon^p\hbar^p}$, since $\phi(\varepsilon\hbar;p) \geq 0$ when $\hbar \leq \xi(p)$ and $\varepsilon\hbar \leq \xi(p)$. Therefore, $(K, \beta\eta^p) \in \mathcal{D}_{\mathrm{C}}^{\varepsilon}$ for each $p \in (0,1]$ and $\varepsilon \in [0,1]$.

Finally, we see that $\mathcal{D}_{\mathrm{C}}^{\varepsilon_2} \subset \mathcal{D}_{\mathrm{C}}^{\varepsilon_1}$ if $\varepsilon_1 < \varepsilon_2$ with $\varepsilon_1, \varepsilon_2 \in [0,1]$. Indeed, from $\psi(\varepsilon) = \frac{1+p}{(2+p)\varepsilon}$ and $\varphi(\varepsilon) = \xi(p)\frac{(1+p)^p(1-\varepsilon\hbar)^{1+p}}{\varepsilon^p\hbar^p}$, we have $\psi'(\varepsilon) = \frac{-(1+p)}{(2+p)\varepsilon^2} \leq 0$ and $\varphi'(\varepsilon) = -\frac{(1+p)^p\xi(p)(1-\varepsilon\hbar)^p(p+\varepsilon\hbar)}{\varepsilon^{1+p}\hbar^p} \leq 0$, since $1 - \varepsilon\hbar \geq \frac{1}{2+p} \geq 0$, so that $\psi(\varepsilon_2) \leq \psi(\varepsilon_1)$ and $\varphi(\varepsilon_2) \leq \varphi(\varepsilon_1)$, and therefore $\mathcal{D}_{\mathrm{C}}^{\varepsilon_2} \subset \mathcal{D}_{\mathrm{C}}^{\varepsilon_1}$.

To that end, we have proven the improvement obtained for the domain of parameters of Newton's method with the help of the conditions of type (3) that we have just shown in Figure 2.

5. Application

Now, we illustrate all of that mentioned above with the following mildly nonlinear elliptic equation:

$$u_{xx} + u_{yy} = u^{5/3}. \tag{21}$$

This type of equation is of interest in the theory of gas dynamics [7]. An associated Dirichlet problem can be formulated as follows. Suppose that the equation is satisfied in the interior of the square $0 \leq x, y \leq 1$ in $\mathbb{R}^2$ and that $u(x,y) > 0$ is given and continuous on the boundary of the square ([8]):

$$u(x,0) = 2x^2 - x + 1, \quad u(x,1) = 2, \quad 0 \leq x \leq 1,$$

$$u(0,y) = 2y^2 - y + 1, \quad u(1,y) = 2, \quad 0 \leq y \leq 1. \tag{22}$$

Our discussion focuses on the formulation of the finite difference equation for the elliptic boundary value Problems (21)–(22). The method of finite differences applied to this problem yields a finite system of equations. For general use, iterative techniques often represent the best approach to the solution of such finite systems of equations.

Specifically, central difference approximations for (21) are used, so that Problems (21)–(22) are reduced to the problem of finding a real zero of a function $F : \Omega \subseteq \mathbb{R}^m \longrightarrow \mathbb{R}^m$, namely a real solution $\mathbf{x}^*$ of a nonlinear system $F(\mathbf{x}) = 0$ with m equations and m unknowns. The common technique used to approximate $\mathbf{x}^*$ is the application of iterative methods. In this case, Newton's method goes on being the most used iterative method for approximating $\mathbf{x}^*$, since this method is one of the most efficient.

For Problems (21)–(22) in $\mathbb{R}^2$, Equation (21) can be approximated using central difference approximations for the spacial derivatives. Consider a grid with step size $h = \frac{1}{N+1}$ in x and $k = \frac{1}{M+1}$ in y defined over the region D, so that D is partitioned into a grid consisting of $(N+1) \times (M+1)$ rectangles with sides $h = \frac{1}{N+1}$ and $k = \frac{1}{M+1}$. The mesh points (x_i, y_j) are given by:

$$x_i = ih, \quad y_i = jk, \quad i = 0, 1, \ldots, N+1, \quad j = 0, 1, \ldots, M+1.$$

Considering the following finite difference expressions to approximate the partial differentials:

$$u_{xx}(x_i, y_j) = \frac{u(x_{i-1}, y_j) - 2u(x_i, y_j) + u(x_{i+1}, y_j)}{h^2} + \mathcal{O}(h^2),$$

$$u_{yy}(x_i, y_j) = \frac{u(x_i, y_{j-1}) - 2u(x_i, y_j) + u(x_i, y_{j+1})}{k^2} + \mathcal{O}(k^2),$$

Equation (21) is approximated at each interior grid point (x_i, y_j) by the difference equation:

$$\frac{u(x_{i+1}, y_j) - 2u(x_i, y_j) + u(x_{i-1}, y_j)}{h^2} + \frac{u(x_i, y_{j+1}) - 2u(x_i, y_j) + u(x_i, y_{j-1})}{k^2} = u(x_i, y_j)^{5/3},$$

for $i = 1, 2, \ldots, N$ and $j = 1, 2, \ldots, M$. The boundary conditions are:

$$u(x_i, y_0) = 2x_i^2 - x_i + 1, \quad u(x_i, y_M) = 2, \quad i = 1, 2, \ldots, N,$$

$$u(x_0, y_j) = 2y_j^2 - y_j + 1, \quad u(x_N, y_j) = 2, \quad j = 0, 1, \ldots, M+1.$$

If we now denote the approximate value of $u(x_i, y_j)$ as $u_{i,j}$, we obtain the difference equation:

$$2\left[\left(\frac{h}{k}\right)^2 + 1\right] u_{i,j} - (u_{i-1,j} + u_{i+1,j}) - \left(\frac{h}{k}\right)^2 (u_{i,j-1} + u_{i,j+1}) = -h^2 u_{i,j}^{5/3},$$

for $i = 1, 2, \ldots, N$ and $j = 1, 2, \ldots, M$, with:

$$u_{i,0} = 2x_i^2 - x_i + 1, \quad u_{i,M+1} = 2, \quad i = 1, 2, \ldots, N,$$

$$u_{0,j} = 2y_j^2 - y_j + 1, \quad u_{N+1,j} = 2, \quad j = 0, 1, \ldots, M+1.$$

Equation (21) with the boundary conditions given in (22) forms an $NM \times NM$ nonlinear system of equations. To set up the nonlinear system, the $NM = m$ interior grid points are labeled row-by-row from x_1 to x_m starting from the left-bottom corner point. The resulting system is:

$$A\mathbf{x} + h^2 q(\mathbf{x}) = \mathbf{v},$$

where:

$$A = \begin{pmatrix} B & C & 0 & \cdots & \cdots & 0 \\ C & B & C & \ddots & & \vdots \\ 0 & C & B & C & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & \ddots & C \\ 0 & \cdots & \cdots & 0 & C & B \end{pmatrix}_{M \times M},$$

$$B = \begin{pmatrix} 2(\lambda+1) & -1 & 0 & \cdots & \cdots & 0 \\ -1 & 2(\lambda+1) & -1 & \ddots & & \vdots \\ 0 & -1 & 2(\lambda+1) & -1 & \ddots & \vdots \\ \vdots & \ddots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & \ddots & -1 \\ 0 & \cdots & & 0 & -1 & 2(\lambda+1) \end{pmatrix}_{N \times N},$$

$C = -\lambda I$, $\lambda = \left(\frac{h}{k}\right)^2$, I is the identity matrix in $\mathbb{R}^N$, $\mathbf{x} = (x_1, x_2, \ldots, x_m)^t$, $q(\mathbf{x}) = \left(x_1^{5/3}, x_2^{5/3}, \ldots, x_m^{5/3}\right)^t$ and $\mathbf{v}$ is a vector formed from the boundary conditions (systems of this type are so-called mildly nonlinear systems).

If we denote the previous system by $F(\mathbf{x}) = 0$, where:

$$F(\mathbf{x}) = A\mathbf{x} + h^2 q(\mathbf{x}) - \mathbf{v} \qquad \text{and} \qquad F : \mathbb{R}^m \longrightarrow \mathbb{R}^m, \quad m = NM, \tag{23}$$

then $F'(\mathbf{x})$ is a linear operator, which is given by the matrix:

$$F'(\mathbf{x}) = A + h^2 Q(\mathbf{x}), \qquad Q(\mathbf{x}) = \frac{5}{3} \text{diag}\left\{x_1^{2/3}, x_2^{2/3}, \ldots, x_m^{2/3}\right\}. \tag{24}$$

We now choose $N = M = 4$ and the infinity norm. In addition, the number of equations is $m = 16$, $h = k = \frac{1}{5}$ and $\lambda = 1$. Besides, $q(\mathbf{x}) = \left(x_1^{5/3}, x_2^{5/3}, \ldots, x_{16}^{5/3}\right)^t$ and:

$$\mathbf{v} = \left(\frac{44}{25}, \frac{23}{25}, \frac{28}{25}, \frac{87}{25}, \frac{23}{25}, 0, 0, 2, \frac{28}{25}, 0, 0, 2, \frac{87}{25}, 2, 2, 4\right)^t.$$

In this case, we observe that a solution $\mathbf{x}^*$ of the system $F(x) = 0$ with F defined in (23) satisfies:

$$\|\mathbf{x}^*\| \leq \|A^{-1}\| \left(\|\mathbf{v}\| + h^2\|q(\mathbf{x})\|\right) \qquad \Rightarrow \qquad \|\mathbf{x}^*\| - \frac{5}{3}\left(4 + \frac{1}{25}\|\mathbf{x}^*\|^{5/3}\right) \leq 0,$$

where $\|A^{-1}\| = \frac{5}{3}$ and $\|\mathbf{v}\| = 4$, so that $\|\mathbf{x}^*\| \in [0, \varrho_1] \cup [\varrho_2, +\infty]$, where $\varrho_1 = 9.5149\ldots$ and $\varrho_2 = 45.9125\ldots$ are the two positive real roots of the scalar equation $t - \frac{5}{3}\left(4 + \frac{t^{5/3}}{25}\right) = 0$. Then, we can consider:

$$F : \Omega \subset \mathbb{R}^{16} \longrightarrow \mathbb{R}^{16} \qquad \text{with} \qquad \Omega = \left\{\mathbf{x} \in \mathbb{R}^{16} : \|\mathbf{x}\| < 10\right\},$$

since $\varrho_1 < 10 < \varrho_2$.

Moreover, $F'(\mathbf{x})$ is the linear operator given by the matrix:

$$A + \frac{1}{15} \operatorname{diag} \left\{ x_1^{2/3}, x_2^{2/3}, \ldots, x_{16}^{2/3} \right\}$$

and:

$$F'(\mathbf{x}) - F'(\mathbf{y}) = \frac{1}{15} \operatorname{diag} \left\{ x_1^{2/3} - y_1^{2/3}, x_2^{2/3} - y_2^{2/3}, \ldots, x_{16}^{2/3} - y_{16}^{2/3} \right\},$$

where $\mathbf{y} = (y_1, y_2, \ldots, y_{16})^t$. In addition,

$$\|F'(\mathbf{x}) - F'(\mathbf{y})\| \le \frac{1}{15} \left\| \mathbf{x}^{1/3} + \mathbf{y}^{1/3} \right\| \left\| \mathbf{x}^{1/3} - \mathbf{y}^{1/3} \right\| \le \frac{2}{15} \sqrt[3]{10} \, \|\mathbf{x} - \mathbf{y}\|^{1/3},$$

$$\|F'(\mathbf{x}) - F'(\mathbf{x_0})\| \le \frac{1}{15} \left(\|\mathbf{x}\|^{1/3} + \|\mathbf{x_0}\|^{1/3} \right) \|\mathbf{x} - \mathbf{x_0}\|^{1/3} \le \frac{1}{15} \left(\sqrt[3]{10} + \|\mathbf{x_0}\|^{1/3} \right) \|\mathbf{x} - \mathbf{x_0}\|^{1/3}.$$

Thus, $K = 0.2872\ldots$, $K_0 = 0.2199\ldots$ and $p = \frac{1}{3}$.

If we choose the starting point $\mathbf{x_0} = \left(\frac{3}{2}, \frac{3}{2}, \ldots, \frac{3}{2} \right)^t$, we obtain $\beta = 1.4862\ldots$ and $\eta = 0.5185\ldots$, so that the condition $\hbar = K\beta\eta^p \le \xi(p)$ of Theorem 1 is not satisfied, since $\hbar = 0.3429\ldots > \xi(p) = \xi\left(\frac{1}{3}\right) = 0.3071\ldots$, where $\xi\left(\frac{1}{3}\right)$ is the unique zero of the corresponding auxiliary function given by (2), $\phi\left(x; \frac{1}{3}\right) = \frac{1}{3}\left(6^{2/3}(1 - t)^{4/3} - 3\sqrt[3]{t} \right)$, in the interval $\left(0, \frac{1}{2}\right]$. As a consequence, we cannot use Theorem 1 to guarantee the convergence of Newton's method for approximating a solution of the system $F(\mathbf{x}) = 0$, where F is defined in (23).

However, we can guarantee the convergence of Newton's method from Corollary 4, since Condition (20) is satisfied: $\hbar = 0.3429\ldots < \xi(p)\frac{(1+p)^p(1 - \varepsilon\hbar)^{1+p}}{\varepsilon^p \hbar^p} = 0.3517\ldots$ with $\varepsilon = 0.7656\ldots$. Therefore, we can then apply Newton's method for approximating a solution of the system $F(\mathbf{x}) = 0$ with F defined in (23) and obtain the approximation given by the vector $\mathbf{x}^* = (x_1^*, x_2^*, \ldots, x_{16}^*)^t$ that is shown in Table 1, reached after four iterations with a tolerance 10^{-16}. In Table 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 1 is a good approximation of the solution of the system, since $\|F(\mathbf{x}^*)\| \le C \times 10^{-16}$. See the sequence $\{\|F(\mathbf{x_n})\|\}$ in Table 2.

Table 1. Approximation of the solution $\mathbf{x}^*$ of $F(\mathbf{x}) = 0$ with F given in (23).

i	x_i^*	i	x_i^*	i	x_i^*	i	x_i^*
1	$0.979069\ldots$	5	$1.097445\ldots$	9	$1.291651\ldots$	13	$1.587503\ldots$
2	$1.097445\ldots$	6	$1.245767\ldots$	10	$1.422935\ldots$	14	$1.664776\ldots$
3	$1.291651\ldots$	7	$1.422935\ldots$	11	$1.561551\ldots$	15	$1.742203\ldots$
4	$1.587503\ldots$	8	$1.664776\ldots$	12	$1.742203\ldots$	16	$1.843388\ldots$

Table 2. Absolute errors obtained by Newton's method and $\{\|F(\mathbf{x_n})\|\}$.

n	$\|\mathbf{x_n} - \mathbf{x}^*\|$	$\|F(\mathbf{x_n})\|$
0	$5.2093\ldots \times 10^{-1}$	$1.318622\ldots$
1	$2.4028\ldots \times 10^{-3}$	$5.4477\ldots \times 10^{-3}$
2	$8.4485\ldots \times 10^{-8}$	$1.2913\ldots \times 10^{-7}$
3	$1.2409\ldots \times 10^{-16}$	$1.4741\ldots \times 10^{-16}$

Finally, we note that if we interpolate the points of Table 1 and take into account that the solution satisfies the boundary conditions given in (22), we obtain the approximation of the numerical solution shown in Figure 3.
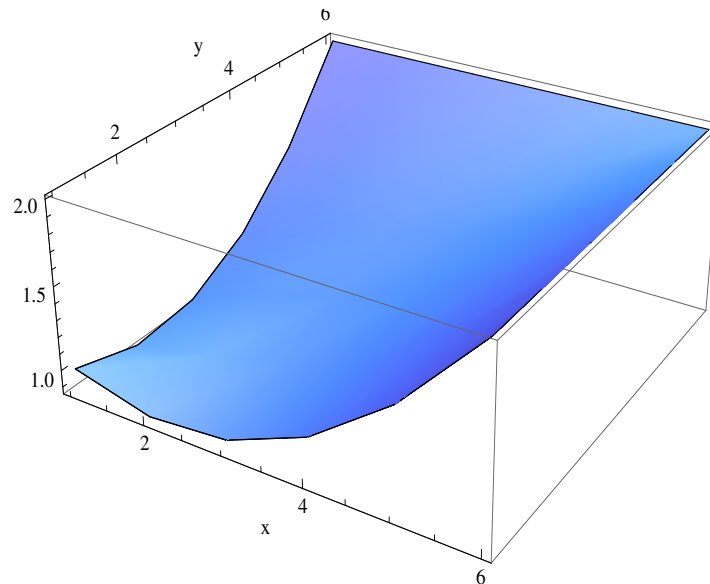

Figure 3. Approximated solution of Problems (21)–(22).

Acknowledgments

Author Contributions

The contributions of the two authors have been similar. Both authors have worked together to develop the present manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Kantorovich, L.V.; Akilov, G.P. *Functional Analysis*; Pergamon Press: Oxford, UK, 1982.
2. Fenyö, I. Über die lösung der in Banachschen raume definierten nichtlinearen gleichungen. *Acta Math. Acad. Sci. Hung.* **1954**, *5*, 85–93.
3. Argyros, I.K. Remarks on the convergence of Newton's method under Hölder continuity conditions. *Tamkang J. Math.* **1992**, *23*, 269–277.
4. Hernández, M.A. The Newton method for operators with Hölder continuous first derivative. *J. Optim. Theory Appl.* **2001**, *109*, 631–648.

5. Rokne, J. Newton's method under mild differentiability conditions with error analysis. *Numer. Math.* **1972**, *18*, 401–412.

6. Argyros, I.K. On the Newton-Kantorovich hypothesis for solving equations. *J. Comput. Appl. Math.* **2004**, *169*, 315–332.

7. Greenspan, D. *Introductory Numerical Analysis of Elliptic Boundary Value Problems*; Harper and Row: New York, NY, USA, 1965.

8. Rall, L.B. *Computational Solution of Nonlinear Operator Equations*; Robert E. Krieger Publishing Company: Malabar, FL, USA, 1979.