Vol. IX, Nº 1, 2 Nov. (2001)
Matemáticas: 37–50

Matemáticas: Enseñanza Universitaria

A one-step Intermediate Newton iterative scheme for nonlinear equations

Livinus U. Uko   Antonio Ramirez Sierra

Abstract

We study an iterative scheme for the computation of a zero of the sum of two nonlinear operators. The proposed method contains Newton scheme and the Modified Newton scheme as special cases and therefore provides a unified setting for the study of both methods. It could also be viewed as an intermediate scheme between Newton's method and the modified Newton's method.

Keywords: Zeroes of nonlinear operators, modified Newton scheme, Kantorovich-type conditions.

1 Introduction

Let X and Y be Banach spaces, let D be an open convex subset of X, and let $f, g\ D \mapsto Y$ be Fréchet differentiable functions, fixed all through this paper. In the sequel, given any $u_0 \in D$ and $r > 0$, $B[u_0, r]$ will designate the set $\{x \in X\ \|x - u_0\| \leq r\}$, and $B(u_0, r)$ will designate the interior of $B[u_0, r]$.

We are interested in the iterative solution of the equation

$$f(u) + g(u) = 0. \tag{1}$$

In this paper, we study the iterative scheme

$$u_{m+1} = u_m - [f'(u_m) + g'(u_0)]^{-1}[f(u_m) + g(u_m)], \quad m = 0, 1, \ldots. \tag{2}$$

When $g = 0$, this scheme becomes the classical Newton scheme for the equation $f(u) = 0$, and when $f = 0$, it becomes the Modified Newton scheme for the equation $g(u) = 0$. Therefore the scheme (2) provides a unified setting for the study of both methods.

The scheme (2) could also be viewed as an intermediate scheme between Newton's method

$$u_{m+1} = u_m - [f'(u_m) + g'(u_m)]^{-1}[f(u_m) + g(u_m)], \quad m = 0, 1, \ldots. \tag{3}$$

and the modified Newton's method

$$u_{m+1} = u_m - [f'(u_0) + g'(u_0)]^{-1}[f(u_m) + g(u_m)], \quad m = 0, 1, \dots. \quad (4)$$

It is well-known that, in general, Newton's method converges faster than the modified Newton method. However, the Newton's method updates derivatives at each iteration step, while the modified method performs only one derivative evaluation. This led several authors (cf. [1,6,7]) to propose intermediate Newton methods which converge faster than the modified method and are easier to implement than the classical scheme. The Newton-type scheme (2) is evidently a (one-step) scheme of this kind.

The basic convergence results for Newton's method and the modified Newton method were proved by Kantorovich (cf. [3–5]), and were later improved by many authors (cf. [8–12]). In the sequel we give convergence results for the intermediate Newton scheme (2) under Kantorovich-type conditions.

2 Some preliminary results

The next two simple results will be used repeatedly in the sequel. The proofs are included for the sake of completeness.

Lemma 1. *Let $u_0 \in D$, let $F : D \mapsto Y$ be a Fréchet differentiable function, and suppose that $F'(u_0)^{-1}$ exists and there exist $K > 0$ such that for some $z \in D$ (independent of K),*

$$\|F'(u_0)^{-1}[F'(x) - F'(z)]\| \le K\|x - z\|, \quad \text{for all } x \in D.$$

Then for all $x, y \in D$, we have

$$\|F'(u_0)^{-1}[F(x) - F(y) - F'(z)(x - y)]\| \le K\|x - y\| \frac{(\|x - z\| + \|y - z\|)}{2}.$$

Proof. We have

$$F'(u_0)^{-1}[F(x) - F(y) - F'(z)(x - y)] = \int_0^1 F'(u_0)^{-1}[F'(sx + (1 - s)y) - F'(z)](x - y)\, ds.$$

Therefore, it follows from the hypotheses that

$$\begin{aligned}\|F'(u_0)^{-1}[F(x) - F(y) - F'(z)(x-y)]\| &\\ &\leq K\|x-y\| \int_0^1 (s\|x-z\| + (1-s)\|y-z\|)\, ds \\ &= K\|x-y\| \frac{(\|x-z\| + \|y-z\|)}{2}.\end{aligned}$$

Lemma 2. *Let $u_0 \in D$, with $B(u_0, r) \subset D$, let $F\, D \mapsto Y$ be twice Fréchet differentiable function, and suppose that $F'(u_0)^{-1}$ exists and $F'(u_0)^{-1}F''(u_0)$ is a bounded map. If*

$$\begin{aligned}\|F'(u_0)^{-1}[F''(x) - F''(u_0)]\| &\leq \lambda\|x - u_0\|, \quad \text{for all } x \in B(u_0, r),\\ \|F'(u_0)^{-1}F''(u_0)]\| &\leq b\end{aligned}$$

then

$$\begin{aligned}\|F'(u_0)^{-1}[F(x) - F(y) - F'(y)(x-y)]\| \leq &\\ &\|x-y\|^2 \frac{(r\lambda + b)}{2}, \quad \text{for all } x, y \in B(u_0, r).\end{aligned}$$

Proof. We have

$$\begin{aligned}F'(u_0)^{-1}&[F(x) - F(y) - F'(y)(x-y)] = \\ &\int_0^1 F'(u_0)^{-1}[F''(sx + (1-s)y) - F''(u_0)](x-y)^2(1-s)\, ds \\ &+ \frac{1}{2}F'(u_0)^{-1}F''(u_0)(x-y)^2.\end{aligned}$$

Therefore, it follows from the hypotheses that for all $x, y \in B(u_0, r)$, we have

$$\begin{aligned}\|F'(u_0)^{-1}[F(x) - F(y) - F'(y)(x-y)]\| &\leq \frac{b\|x-y\|^2}{2} + \\ \lambda\|x-y\|^2 &\int_0^1 (s\|x-u_0\| + (1-s)\|y-u_0\|)(1-s)\, ds \\ \leq \|x-y\|^2 &\left[\frac{b}{2} + \lambda r \int_0^1 (1-s)\, ds\right] \\ = \frac{\|x-y\|^2(b+\lambda r)}{2}&\end{aligned}$$

In the sequel we will use the method of majorizing sequences to analyze the convergence of the intermediate Newton scheme (2). In the following proposition we define the majorant sequence that will be used, and give its main properties.

Proposition 1 *Let $a > 0$, let M and L be non-negative numbers such that $M + L > 0$ and $2(M + L)a \leq 1$, and let*

$$t_* = (1 - \sqrt{1 - 2(M + L)a})/(M + L).$$

Let $t_0 = 0$, and for $m = 0, 1, \ldots,$ let

$$t_{m+1} = (Lt_m^2/2 - Mt_m^2/2 + a)/(1 - Mt_m) \tag{5}$$

Then the t_m are well defined and we have

$$t_{m-1} < t_m < t_*, \quad \text{for } m = 1, 2, \ldots, \tag{6}$$

and

$$\lim_{n \to \infty} t_m = t_*.$$

Proof. Since $t_0 = 0 < t_1 = a$ and $t_* - a = (M + L)t_*^2/2 > 0$, we see that (6) holds when $m = 1$. Suppose now, by induction, $m \geq 1$ and that (6) holds. If $M > 0$, then

$$Mt_m < Mt_* \leq (M + L)t_* = 1 - \sqrt{1 - 2(M + L)a} \leq 1,$$

and if $M = 0$, then $Mt_m = 0 < 1$. In either case, we see that t_{m+1} is well defined. We have the estimates

$$\begin{aligned}
t_* - t_{m+1} &= [(1 - Mt_m)t_* - a - Lt_m^2/2 + Mt_m^2/2]/(1 - Mt_m) \\
&= [t_* - a - Mt_m t_* - Lt_m^2/2 + Mt_m^2/2]/(1 - Mt_m) \\
&= [(M + L)t_*^2/2 - Mt_m t_* - Lt_m^2/2 + Mt_m^2/2]/(1 - Mt_m) \\
&= [M(t_* - t_m)^2/2 + L(t_*^2 - t_m^2)/2]/(1 - Mt_m) > 0 \\
t_{m+1} - t_m &= [a - t_m + Lt_m^2/2 + Mt_m^2/2]/(1 - Mt_m) \\
&= [(1 - Mt_{m-1})t_m + Mt_{m-1}^2/2 - Lt_{m-1}^2/2 - t_m \\
&\quad + Lt_m^2/2 + Mt_m^2/2]/(1 - Mt_m) \\
&= [-Mt_{m-1}t_m + Mt_{m-1}^2/2 - Lt_{m-1}^2/2 \\
&\quad + Lt_m^2/2 + Mt_m^2/2]/(1 - Mt_m) \\
&= [M(t_m - t_{m-1})^2/2 + L(t_m^2 - t_{m-1}^2)/2]/(1 - Mt_m) > 0
\end{aligned}$$

which show that (6) holds when m is replaced with $m + 1$ and hence, by induction, that it holds for all positive integral values of m.

It follows that t_m is monotone increasing and bounded above by t_*. Hence it converges, as m tends to infinity, to a real number t, with the property that $0 \leq t \leq t_*$. On the other hand, it follows from (5) that

$$Lt_m^2/2 - Mt_m^2/2 + a - t_{m+1}(1 - Mt_m) = 0,$$

and, on letting m tend to infinity, we see that

$$(M + L)t^2/2 - t + a = 0.$$

Since t_* is the smallest solution of this equation we conclude that $t_* \leq t \leq t_*$, and hence that (7) holds..

3 Convergence results

We first prove the convergence of the intermediate scheme (2) under Kantorovich-type conditions.

Theorem 1 *Let a, M and L be non-negative numbers such that $M+L > 0$ and $2(M + L)a \leq 1$, let $T_* = (1 + \sqrt{1 - 2(M + L)a})/(M + L)$ and let t_* and t_m be defined as in Proposition 1. Let $u_0 \in D$ and let $f, g\, D \mapsto Y$ be differentiable functions such that $J_0 = f'(u_0) + g'(u_0)$ is invertible, $\|J_0^{-1}[f(u_0) + g(u_0)]\| \leq a$, $B[u_0, t_*] \subset D$, and*

$$\begin{aligned} \|J_0^{-1}[f'(x) - f'(y)]\| &\leq M\|x - y\|, \\ \|J_0^{-1}[g'(x) - g'(y)]\| &\leq L\|x - y\|, \ \forall x, y \in D. \end{aligned}$$

Then the intermediate Newton iterates in (2) are all well defined and converge to the unique solution u of equation (1) in $B[u_0, t_*] \cup [D \cap B(u_0, T_*)]$, with error estimates

$$\begin{aligned} \|u_m - u_{m-1}\| &\leq t_m - t_{m-1}, \\ \|u_m - u_0\| &\leq t_m, \\ \|u - u_m\| &\leq t_* - t_m. \end{aligned}$$

Proof. If $a = 0$, then $u = u_0$ solves equation (1) and, since $u_m = u_0$ and $t_m = t_0$ for all m, the estimates (10) - (12) hold trivially. In the rest of the proof we assume $a > 0$. Since $\|u_1 - u_0\| \leq a = t_1 - t_0$, we see that (10) - (11) hold when $m = 1$. Suppose now, by induction, that $m \geq 1$ and that the u_m are well defined and satisfy (10) - (11). Then, on letting $J_m \equiv f'(u_m) + g'(u_0) = J_0(I + A)$ or, equivalently, $A = J_0^{-1}[f'(u_m) - f'(u_0)]$, and applying (8), we see that $\|A\| \leq Mt_m < 1$. Therefore $(I + A)^{-1}$ exists, with

$$\|(I + A)^{-1}\| \leq 1/(1 - Mt_m),$$

and it follows that J_m is invertible, and that $J_m^{-1}J_0 = (I+A)^{-1}$. Hence $\|J_m^{-1}J_0\| \le 1/(1-Mt_m)$, and it follows from the induction hypotheses that

$$\|u_{m+1}-u_m\| = \|J_m^{-1}[f(u_m)+g(u_m)]\| \le \|J_m^{-1}J_0\|\|J_0^{-1}[f(u_m)+g(u_m)]\|.$$

But $f(u_m)+g(u_m) = [f(u_m)-f(u_{m-1})-f'(u_{m-1})(u_m-u_{m-1})]+[g(u_m)-g(u_{m-1})-g'(u_0)(u_m-u_{m-1})]$. Hence it follows from (8), (9) and Lemma 1 that

$$\begin{aligned}
&\|J_0^{-1}[f(u_m)+g(u_m)]\| \\
&\qquad \le \|J_0^{-1}[f(u_m)-f(u_{m-1})-f'(u_{m-1})(u_m-u_{m-1})]\| \\
&\qquad\quad + \|J_0^{-1}[g(u_m)-g(u_{m-1})-g'(u_0)(u_m-u_{m-1})]\| \\
&\qquad \le M\frac{\|u_m-u_{m-1}\|^2}{2} + L\|u_m-u_{m-1}\|[\|u_m-u_0\| \\
&\qquad\quad + \|u_{m-1}-u_0\|]/2 \\
&\qquad \le M(t_m-t_{m-1})^2/2 + L(t_m-t_{m-1})(t_m+t_{m-1}))/2 \\
&\qquad = M(t_m-t_{m-1})^2/2 + L(t_m^2-t_{m-1}^2)/2
\end{aligned}$$

and

$$\begin{aligned}
\|u_{m+1}-u_m\| &\le [M(t_m-t_{m-1})^2/2 + L(t_m^2-t_{m-1}^2)/2]/(1-Mt_m) \\
&= t_{m+1}-t_m, \\
\|u_{m+1}-u_0\| &\le \|u_{m+1}-u_m\| + \|u_m-u_0\| \\
&\le t_{m+1}-t_m+t_m \\
&= t_{m+1}
\end{aligned}$$

It follows that (10) and (11) also hold when m is replaced with $m+1$ and hence, by induction, that they hold for all positive integral values of m. This implies that

$$\|u_{m+q}-u_m\| \le \sum_{k=m+1}^{m+q}\|u_k-u_{k-1}\| \le \sum_{k=m+1}^{m+q}(t_k-t_{k-1}) = t_{m+q}-t_m.$$

Since t_m is a Cauchy sequence, it follows that u_m is also a Cauchy sequence converging to some $u \in B[u_0,t_*]$. On letting q tend to infinity we see that (12) holds. It follows from (2) that $[f'(u_m)+g'(u_0)](u_{m+1}-u_m)+f(u_m)+g(u_m)=0$ and on letting m tend to infinity we see that u solves equation (1).

Finally, to prove the uniqueness assumption, we suppose that v is another solution of (1) in $B[u_0, t_*] \cup [D \cap B(u_0, T_*)]$. Then we have

$$\begin{aligned} u_0 - v &= J_0^{-1}[f(v) - f(u_0) - f'(u_0)(v - u_0)] \\ &\quad + J_0^{-1}[g(v) - g(u_0) - g'(u_0)(v - u_0)] + J_0^{-1}[f(u_0) + g(u_0)] \\ \|u_0 - v\| &\leq \|J_0^{-1}[f(v) - f(u_0) - f'(u_0)(v - u_0)]\| \\ &\quad + \|J_0^{-1}[g(v) - g(u_0) - g'(u_0)(v - u_0)]\| + \|J_0^{-1}[f(u_0) + g(u_0)]\| \\ &\leq (M + L)\|v - u_0\|^2/2 + a, \end{aligned}$$

and we conclude that

$$\|v - u_0\| \leq (M + L)\|v - u_0\|^2/2 + a.$$

By solving this inequality, we see that either $\|v - u_0\| \geq T_*$ or $\|v - u_0\| \leq t_*$. Since $v \in B[u_0, t_*] \cup [D \cap B(u_0, T_*)]$, we conclude that $\|v - u_0\| \leq t_*$.

We show, by induction that the relation $\|v - u_m\| \leq t_* - t_m$ holds for all positive integral values of m. It follows from the previous paragraph that it holds when $m = 0$. If we suppose that it holds for m, then we have

$$\begin{aligned} v - u_{m+1} = &- [f'(u_m) + g'(u_0)]^{-1}[f(v) - f(u_m) \\ &- f'(u_m)(v - u_m) + g(v) - g(u_m) - g'(u_0)(v - u_m)], \end{aligned}$$

and it follows from (8), (9) and Lemma 1 that

$$\begin{aligned} \|v - u_{m+1}\| &\leq \|J_m^{-1}[f(v) - f(u_m) - f'(u_m)(v - u_m)]\| \\ &\quad + \|J_m^{-1}[g(v) - g(u_m) - g'(u_0)(v - u_m)]\| \\ &\leq \|J_m^{-1}J_0\|[M\|v - u_m\|^2/2 \\ &\quad + L\|v - u_m\|(\|v - u_0\| + \|u_m - u_0\|)/2] \\ &\leq [M(t_* - t_m)^2/2 \\ &\quad + L(t_* - t_m)(t_* + t_m)/2]/(1 - Mt_m) \\ &= [M(t_* - t_m)^2/2 + L(t_*^2 - t_m^2)/2]/(1 - Mt_m) \\ &= [(M + L)t_*^2/2 - Mt_mt_* - Lt_m^2/2 + Mt_m^2/2]/(1 - Mt_m) \\ &= [t_* - a - Mt_mt_* - Lt_m^2/2 + Mt_m^2/2]/(1 - Mt_m) \\ &= t_* - [Lt_m^2/2 - Mt_m^2/2 + a]/(1 - Mt_m) \\ &= t_* - t_{m+1}. \end{aligned}$$

which shows that the induction hypotheses is true when m is replaced with $m + 1$. We conclude that the inequality $\|v - u_m\| \leq t_* - t_m$ holds for all m. By letting m tend to infinity we see that $u = v$. We therefore conclude that every solution of (1) in $B[u_0, t_*] \cup [D \cap B(u_0, T_*)]$ coincides with u. That completes the proof of the theorem.

One could also analyze the convergence rate of the scheme (2) with the method of nondiscrete mathematical induction developed by Pták (cf. [8–10]). In this method a central role is played by the concept of a rate function, defined as mapping of the form $rT \mapsto T$ where $T = (0, \infty)$ or $T = (0, a]$ for some positive real number a, with the property that the sum function $s(t) = \sum_{m=0}^{\infty} r^{(m)}(t)$ is convergent for all $t \in T$, where $r^{(0)}$ is the identity and $r^{(m+1)} \equiv r \circ r^{(m)}$. The basic nondiscrete induction theorem is the following result from [10].

Theorem 2. *Let W be a complete metric space and let w be a mapping from a subset of W into W. Suppose that there exists a point $u_0 \in W$, a rate function r on an interval T, and subsets $Z(t)$ of W (one for each $t \in T$) such that the following conditions are satisfied:*

$$u_0 \in Z(a) \text{ for a certain } a \in T,$$

$$(t, x) \in T \times Z(t) \text{ implies } w(x) \in Z(r(t)) \text{ and } d(w(x), x) \leq t.$$

Then the iterates generated from the successive approximation scheme

$$u_{m+1} = w(u_m), \quad m = 0, 1, \ldots$$

converge to an element $u \in W$ and satisfy the estimates

$$\begin{aligned} d(u_m, u) &\leq s(r^{(m)}(a)), \\ d(u_m, u_{m+1}) &\leq r^{(m)}(a). \end{aligned}$$

In the following result we deduce the convergence of the intermediate Newton-type scheme (2) from this basic nondiscrete induction Theorem.

Theorem 3. *Suppose that the hypotheses of Theorem 1 hold. Then the Newton-type iterates in (2) converge to a solution u of (1), at the rate*

$$\begin{aligned} \|u_m - u\| &\leq s(r^{(m)}(a)), \\ \|u_{m+1} - u_m\| &\leq r^{(m)}(a), \end{aligned}$$

and the rate and sum functions given, respectively, by the expressions

$$\begin{aligned} r(t) &= \frac{t[2L(1 - z(t)) + t(M^2 + L^2)]}{2L + 2M(z(t) - Lt)} \\ s(t) &= [Mt - \sqrt{1 - 2(M + L)a} + z(t)]/(M + L), \end{aligned}$$

with $z(t) = \sqrt{1-2(M+L)a+2Lt+M^2t^2}$, *defined for all* $t \in T \equiv (0, a]$.

Proof. All we have to do is verify the hypotheses of Theorem 2. We observe that $z(a) = 1 - Ma$ and $s(a) = (1-\sqrt{1-2(M+L)a})/(M+L) = t_*$.

Given any $t \in T$, it follows the hypotheses that $2t \le 2a \le 1/(M+L)$, and hence that $1-2(M+L)t \ge 0$ and

$$z(t) = \sqrt{(1-Mt)^2 - 2(M+L)(a-t)} \le 1 - Mt.$$

Also, since

$$|z(t)|^2 - t^2(M+2L)^2 = 1 - 2(M+L)a + 2Lt[1-2(M+L)t] \ge 0,$$

we see that $Mt + 2Lt \le z(t) \le 1 - Mt, \quad \forall t \in T$, which implies that $r(t)$ is well-defined and positive. Also, since $z(t) \ge Mt + 2Lt \ge (M + L)t/2$, it follows from (17) that $r(t) \le t$, and consequently, that r maps T into itself. Since $s(t) = t + s(r(t))$ and $\lim_{t\downarrow 0} s(t) = 0$, it follows from Proposition 1 of [9] that $r(t)$ is a rate function, with $s(t)$ as its corresponding sum function.

For any $x \in D$, let $w(x) = [f'(x)+g'(u_0)]^{-1}[f'(x)x+g'(u_0)x-f(x)-g(x)]$, whenever it exists. Then it is evident that the iterative scheme (2) coincides with the successive approximation scheme $u_{m+1} = w(u_m)$ and that u solves equation (1) if and only if it is a fixed point of w.

We define, for every $t \in T$, the set $Z(t)$ consisting of elements of D for which $f'(x) + g'(u_0)$ is invertible, with the further properties that $\|x - u_0\| \le s(a) - s(t)$ and $\|[f'(x) + g'(u_0)]^{-1}[f(x) + g(x)]\| \le t$.

The hypotheses show that J_0 is invertible, that $w(u_0)$ is well defined, and $\|w(u_0) - u_0\| = \|J_0^{-1}[f(u_0) + g(u_0)]\| \le a$. It follows that $u_0 \in Z(a)$. That completes the verification of (13).

Next, we observe that if $x \in Z(t)$ and $t \in T$, then $w(x)$ is well defined, and

$$\|w(x) - u_0\| \le \|w(x) - x\| + \|x - u_0\| \le t + s(a) - s(t) = s(a) - s(r(t)).$$

This verifies one part of (14). To verify the second part, observe that $M \stackrel{\bullet}{<} (M+L)(1+Mt)$, and hence, that

$$\begin{aligned}
\|J_0^{-1}[f'(w(x)) + g'(u_0) - f'(u_0)] - g'(u_0)\| \\
&\le M\|w(x) - u_0\| \\
&\le M[s(a) - s(r(t))] \\
&= M(1 + Lt - z(t))/(M+L) \\
&\le M(1 - Mt - Lt)/(M+L) \\
&< 1.
\end{aligned}$$

Therefore, it follows from Banach's Lemma that $[f'(w(x)) + g'(u_0)]^{-1} J_0$ exists, and that

$$\|[f'(w(x)) + g'(u_0)]^{-1} J_0\| \le \frac{1}{1 - M[s(a) - s(r(t))]} = \frac{(M+L)}{L + M(z(t) - Lt)}.$$

Hence $w(w(x))$ is well defined and,

$$\begin{aligned}
&\|J_0^{-1}[f(w(x)) + g(w(x))]\| \\
&\qquad = \|J_0^{-1}[f(w(x)) - f(x) - (f'(x) \\
&\qquad\quad + g'(u_0))(w(x) - x) + g(w(x)) - g(x)]\| \\
&\qquad \le \|J_0^{-1}[f(w(x)) - f(x) - f'(x)(w(x) - x)\| \\
&\qquad\quad + \|J_0^{-1}[g(w(x)) - g(x) - g'(u_0)(w(x) - x)]\| \\
&\qquad \le M\|w(x) - x)\|^2/2 + L(\|w(x) - u_0\| \\
&\qquad\quad + \|x - u_0\|)\|w(x) - x) \,|/2 \\
&\qquad \le t[Mt + L(s(a) - s(r(t)) + s(a) - s(t))]/2 \\
&\qquad = t[Mt(M+L) + L(2 - 2z(t) \\
&\qquad\quad + Lt(L - M))]/[2(M+L)] \\
&\qquad = t[2L(1 - z(t)) + t(M^2 + L^2))]/[2(M+L)].
\end{aligned}$$

And

$$\begin{aligned}
\|w(w(x)) - w(x)\| &= \|[f'(w(x)) + g'(u_0)]^{-1}[f(w(x)) + g(w(x))]\| \\
&\le \|[f'(w(x)) + g'(u_0)]^{-1} J_0\| \|J_0^{-1}[f(w(x)) + g(w(x))]\| \\
&\le t[2L(1 - z(t)) + t(M^2 + L^2)]/[2(L + M(z(t) - Lt))] \\
&= r(t)
\end{aligned}$$

This shows that $w(x) \in Z(w(x))$ and verifies the second part of (14). Therefore the hypotheses of Theorem 2 are satisfied. We conclude that the sequence u_m converges to a vector u in $B[u_0, t_*]$ and that the error bounds (15)–(16) hold. It follows from (2) that $[f'(u_m) + g'(u_0)](u_{m+1} - u_m) + f(u_m) + g(u_m) = 0$ and on letting m tend to infinity we see that u solves equation (1).

Remark 1. The identity

$$s(r(t)) = [Ms(t)^2/2 - Ls(t)^2/2 + Ls_0 s(t)]/[1 - Ms_0 + Ms(t)]$$

shows that the error bounds $s_m \equiv s(r^{(m)}(a))$ in (15) satisfy the recurrence relations

$$s_{m+1} = [Ms_m^2/2 - Ls_m^2/2 + Ls_0 s_m]/[1 - Ms_0 + Ms_m], \quad m = 0, 1, \ldots.$$

Since $t_* - t_0 = t_* = s_0$ and

$$\begin{aligned} t_* - t_{m+1} &= [M(t_* - t_m)^2/2 + L(t_* - t_m)(t_* + t_m)/2]/(1 - Mt_m) \\ &= [M(t_* - t_m)^2/2 - L(t_* - t_m)^2 + Lt_*(t_* - t_m)]/(1 - Mt_m) \end{aligned}$$

we see, from a straightforward induction argument, that the error bounds (12) and (15) are related in the form $s_m = t_* - t_m$.

Remark 2. An error estimate – analogous to (12) – satisfied by the Newton scheme (3) under the hypotheses of Theorem 1 is

$$\|u - u_m\| \leq t_* - w_m,$$

with $w_0 = 0$ and $w_{m+1} = w_m + [(M+L)w_m^2/2 - w_m + a]/[1-(M+L)w_m]$. A bound for the modified Newton scheme (4) under the same conditions is

$$\|u - u_m\| \leq t_* - k_m,$$

with $k_0 = 0$ and $k_{m+1} = a + (M+L)k_m^2/2$. It is not difficult to see that for $m = 0, 1, \ldots,$ we have

$$k_m \leq t_m \leq w_m.$$

The inequality holds trivially when $m = 0$. Suppose, by induction that (21) holds for a certain m. Then

$$\begin{aligned} k_{m+1} &= a + (M+L)k_m^2/2 \leq a + (M+L)t_m^2/2 + Mt_m(t_{m+1} - t_m) \\ &= a - Mt_m^2/2 + Lt_m^2/2 + Mt_m t_{m+1} = t_{m+1} \\ &= [a - Mt_m^2/2 + Lt_m^2/2 - Lt_m t_{m+1}]/[1 - (M+L)t_m] \\ &\leq [a - (M+L)t_m^2/2]/[1 - (M+L)t_m] \\ &\leq [a - (M+L)w_m^2/2]/[1 - (M+L)w_m] = w_{m+1}. \end{aligned}$$

This shows, by induction, that (21) holds for all m, and suggests that (2) is an intermediate scheme between the Newton scheme (3) and the modified Newton scheme (4). However, under the hypotheses of Theorem 1, $J_0^{-1}[f'(x) + g'(x)]$ will have a Lipschitz constant K that is smaller than $M+L$. This implies that the error estimates (19) and (20) are too coarse, and hence that the relations (21) do not prove irrefutably that (1) is an intermediate scheme between the Newton scheme (3) and the modified Newton scheme (4). However, all numerical examples that one cares to do will confirm that that is indeed the case (under the hypotheses of Theorem 1).

Remark 3. It is well known (cf. Example 1.1 in [2]) the Kantorovich hypotheses are not necessary for the convergence of Newton's method. In

[2] Argyros introduced a different set of sufficient conditions – depending on second derivatives – for the convergence of the method. In the next result we prove the convergence of the intermediate scheme (2) under Argyros-type conditions.

Theorem 4. *Let $c > 0$ and let a, b, L and d be non-negative constants such that $M = \max\{c, b+2ad\}$ and $2(M+L)a \leq 1$. Let t_*, T_* and t_m be defined as in Proposition 1 and Theorem 1. Let $u_0 \in D$ and let $f, g\ D \mapsto Y$ be differentiable functions such that $J_0 = f'(u_0) + g'(u_0)$ is invertible, $J_0^{-1}f''(u_0)$ is bounded, $\|J_0^{-1}[f(u_0) + g(u_0)]\| \leq a$, $\|J_0^{-1}f''(u_0)\| \leq b$, $B(u_0, 1/c) \subset D$, and*

$$\begin{aligned} \|J_0^{-1}[f'(x) - f'(u_0)]\| &\leq c\|x - u_0\|, \\ \|J_0^{-1}[g'(x) - g'(u_0)]\| &\leq L\|x - u_0\|, \\ \|J_0^{-1}[f''(x) - f''(u_0)]\| &\leq d\|x - u_0\|, \ \forall x \in D. \end{aligned}$$

Then the intermediate Newton iterates in (2) are all well defined and converge to the unique solution u of equation (1) in $B[u_0, t_*] \cup [D \cap B(u_0, T_*)]$, with error estimates

$$\begin{aligned} \|u_m - u_{m-1}\| &\leq t_m - t_{m-1}, \\ \|u_m - u_0\| &\leq t_m, \\ \|u - u_m\| &\leq t_* - t_m. \end{aligned}$$

Proof. If $a = 0$, then $u = u_0$ solves equation (1) and, since $u_m = u_0$ and $t_m = t_0$ for all m, the estimates (25) – (27) hold trivially. In the rest of the proof we assume $a > 0$.

Since $\|u_1 - u_0\| = a = t_1 - t_0$, we see that (25) – (26) hold when $m = 1$.

Suppose now, by induction, that $m \geq 1$ and that the u_m are well defined and satisfy (25) – (26). Then, on letting $J_m \equiv f'(u_m) + g'(u_0) = J_0(I + A)$ or, equivalently, $A = J_0^{-1}[f'(u_m) - f'(u_0)]$, and applying (22), we see that

$$\|A\| = \|J_0^{-1}[f'(u_m) - f'(u_0)]\| \leq c\|u_m - u_0\| \leq ct_m \leq Mt_m < 1.$$

Therefore $(I + A)^{-1}$ exists, with

$$\|(I + A)^{-1}\| \leq 1/(1 - Mt_m),$$

and it follows that J_m is invertible, and that $J_m^{-1}J_0 = (I + A)^{-1}$. Hence $\|J_m^{-1}J_0\| \leq 1/(1 - Mt_m)$, and it follows that

$$\|u_{m+1} - u_m\| = \|J_m^{-1}[f(u_m) + g(u_m)]\| \leq \|J_m^{-1}J_0\|\|J_0^{-1}[f(u_m) + g(u_m)]\|.$$

But $f(u_m)+g(u_m) = [f(u_m)-f(u_{m-1})-f'(u_{m-1})(u_m-u_{m-1})]+[g(u_m)-g(u_{m-1})-g'(u_0)(u_m-u_{m-1})]$. Hence, since $2a \leq 1/(M+L) \leq 1/M \leq 1/c$ and $\|u_m - u_0\| \leq t_m < t_* \leq 2a$ and $\|u_{m-1} - u_0\| \leq t_{m-1} < t_* \leq 2a$, it follows from (23), (24) and Lemma 2 (with $r = 2a$) that

$$\begin{aligned}
\|J_0^{-1}[f(u_m) + g(u_m)]\| &\leq \\
&\leq \|J_0^{-1}[f(u_m) - f(u_{m-1}) - f'(u_{m-1})(u_m - u_{m-1})]\| \\
&\quad + \|J_0^{-1}[g(u_m) - g(u_{m-1}) - g'(u_0)(u_m - u_{m-1})]\| \\
&\leq (2ad + b)\|u_m - u_{m-1}\|^2/2 \\
&\qquad + L\|u_m - u_{m-1}\|[\|u_m - u_0\| + \|u_{m-1} - u_0\|]/2 \\
&\leq M\|u_m - u_{m-1}\|^2/2 + L\|u_m - u_{m-1}\|[\|u_m - u_0\| \\
&\qquad + \|u_{m-1} - u_0\|]/2 \\
&\leq M(t_m - t_{m-1})^2/2 + L(t_m - t_{m-1})(t_m + t_{m-1}))/2 \\
&= M(t_m - t_{m-1})^2/2 + L(t_m^2 - t_{m-1}^2)/2 \\
\|u_{m+1} - u_m\| &\leq [M(t_m - t_{m-1})^2/2 \\
&\qquad + L(t_m^2 - t_{m-1}^2)/2]/(1 - Mt_m) \\
&= t_{m+1} - t_m, \\
\|u_{m+1} - u_0\| &\leq \|u_{m+1} - u_m\| + \|u_m - u_0\| \\
&\leq t_{m+1} - t_m + t_m \\
&= t_{m+1}
\end{aligned}$$

It follows that (25) and (26) also hold when m is replaced with $m + 1$ and hence, by induction, that they hold for all positive integral values of m. The rest of the proof proceeds exactly as in the proof of Theorem 1.

References

[1] Appell, J., De Pascale E., Evkhuta N.A., Zabrejko, P.P. (1995): *On the two-step Newton method for the solution of nonlinear operator equations. Math.* Nachr. **172** 5–14.

[2] Argyros, I.K. (1999),*A new Kantorovich-type Theorem for Newton's method*, Appl. Math. **26** 151–157.

[3] Kantorovich, L.V. (1948):*On Newton's method for functional equations.* Dokl. Akad. Nauk SSSR **59**, 1237–1240.

[4] Kantorovich, L.V. (1951): *The majorant principle and Newton's method. Dokl.* Akad. Nauk SSSR **76**, 17–20.

[5] Kantorovich, L.V.,Akilov, G.P. (1982): *Functional Analysis.* Pergamon Press, New York.

[6] Ortega, L.M., Rheinboldt, W.C. (1970): *Iterative solution of Nonlinear Equations in several variables.* Academic Press, New York.

[7] Ostrowski, A.M. (1973): *Solution of equations in Euclidean and Banach spaces.* Academic Press, New York.

[8] Pták, V. (1976): *The rate of convergence of Newton's process.* Numer. Math. **25**, 279–285.

[9] Potra, F.-A. (1984): *On the a posteriori error estimates for Newton's method.* Beiträge zur Numer. Math. **12**, 125–138.

[10] Potra, F.-A., Pták, V. (1980): *Sharp error bounds for Newton's process.* Numer. Math. **34**, 63–72. Newton method. Math. Programming **73**, 251–268.

[11] Yamamoto, T. (1985): *Error bounds for Newton's process derived from the Kantorovich conditions.* Japan J. Appl. Math. **2**, 285–292.

[12] Yamamoto, T. (1986):*A method for finding sharp error bounds for Newton's method under Kantorovich assumptions.* Numerische Math. **49**, 203–220.

Dirección de los autores: Livinus U. Uko — Antonio Ramirez Sierra Departamento de Matemáticas, Facultad de Ciencias Exactas y Naturales, Universidad de Antioquia, Medellín, Colombia.