A short survey on Kantorovich-like theorems for Newton's method

Grégoire Lecerf, Joelle Saadé

A short survey on Kantorovich-like theorems for Newton's method

by Grégoire Lecerf, Joelle Saadé

Laboratoire d'informatique de l'École polytechnique
LIX, UMR 7161 CNRS
Campus de l'École polytechnique
1, rue Honoré d'Estienne d'Orves
Bâtiment Alan Turing, CS35003
91120 Palaiseau, France

Preliminary version of August 31, 2015

Abstract

We survey influential quantitative results on the convergence of the Newton iterator towards simple roots of continuously differentiable maps defined over Banach spaces. We present a general statement of Kantorovich's theorem, with a concise proof from scratch, dedicated to wide audience. From it, we quickly recover known results, and gather historical notes together with pointers to recent articles.

1 Introduction

During the last decades, the Newton operator has become omnipresent in numeric and symbolic computations. On specific functions such as polynomials of degree two over real numbers, the behavior of this operator may be simple, but in general it is a difficult problem to determine whether the iterates of a given point converge to a zero or not. More precisely, let $\phi: \mathbb{R} \to \mathbb{R}$ be a real function of class $\mathcal{C}^1$, which means differentiable with ϕ' continuous. In theory, it is classical that Newton sequences $(r_k)_{k \geqslant 0}$ defined by $r_{k+1} = r_k - \frac{\phi(r_k)}{\phi'(r_k)}$ converge quadratically if their *initial value* r_0 is sufficiently close to a *simple zero* r_- of ϕ, which means $\phi'(r_-) \neq 0$. But for practice this information is not sufficient, and one needs to *quantify* what is meant by "sufficiently close".

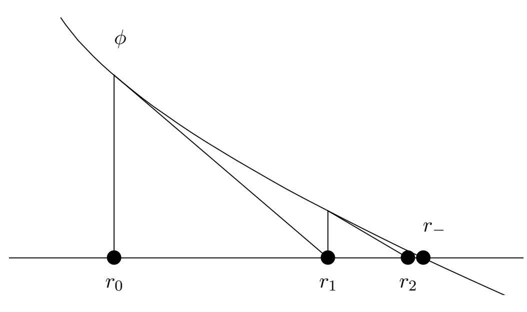

Figure 1. Graph of ϕ and the first Newton iterates of r_0.

In Figure 1, we illustrate the typical behavior of the Newton sequence in a neighborhood of a simple zero r_-. It is a classical result that if ϕ is decreasing and convex in a range $[r_0, R]$, if $\phi(r_0) > 0$, and $\phi(R) < 0$, then there exits a unique zero r_- of ϕ in $[r_0, R]$, and the Newton sequence $(r_k)_{k \geqslant 0}$ converges to r_-. In a sufficiently small neighborhood of r_- this convergence becomes quadratic, which means that the number of digits of the zero is essentially doubled at each iteration.

In general, for a complex function $f: \mathbb{C} \rightarrow \mathbb{C}$, the set of initial values leading to a sequence that converges to a prescribed zero of f are intricate fractal sets, called *Julia sets* of the meromorphic function $z \mapsto z - f(z) / f'(z)$. For practice, it is thus important to design simple criteria, with low complexity, ensuring that an initial point converges to a unique zero in its neighborhood. And we only expect necessary conditions, in the sense that if the criterion fails, then we cannot deduce whether the convergence holds or not. Several such criteria are intensively used in practice. Choosing or designing the most efficient criterion for a given purpose might be quite tedious, because one has to discover good compromises between speed and accuracy. The choice actually depends on the data structure to represent the map f, the way its derivative can be obtained, and also on the type of underlying arithmetic: hardware double precision, intervall or ball arithmetic, arbitrary precision, etc.

Our presentation begins with a standard extension of the seminal criterion due to Kantorovich. Then we show how other old and recent criteria can be recovered from it. We also propose brief comparisons and discussions on how to design other criteria offering alternative compromises. Historical notes are included at the end.

2 Kantorovich theorem

Until the end of the article, $\mathbb{X}$ and $\mathbb{Y}$ represent Banach spaces over $\mathbb{C}$ (typically $\mathbb{C}^n$ in practice) endowed with the norm written $\|\cdot\|$. The class of functions, with values in $\mathbb{Y}$, having continuous derivatives to order ℓ in an open subset $\Omega \subseteq \mathbb{X}$ is written $\mathcal{C}^{\ell}(\Omega, \mathbb{Y})$. If $f \in \mathcal{C}^{\ell}(\Omega, \mathbb{Y})$, then its l-th derivative is written $D^l f$ in general, and $f^{(l)}$ whenever $\mathbb{X}$ has dimension 1. The open ball centered at $a \in \Omega$ and of radius r is written $B(a, r) = \{x \in \mathbb{X} \mid \|x - a\| < r\}$; Its adherence is $\bar{B}(a, r) = \{x \in \mathbb{X} \mid \|x - a\| \leqslant r\}$. If A is a linear map acting on $\mathbb{X}$, then we use the same notation for the following norm: $\|A\| = \sup_{\|x\|=1} \|A\, x\|$. We begin with a very classical lemma.

Lemma 1. *Let $A: \mathbb{X} \rightarrow \mathbb{X}$ be a linear operator such that $\|A\| < 1$, and let* Id *represent the identity map on $\mathbb{X}$. Then* $\mathrm{Id} - A$ *is invertible, of inverse* $(\mathrm{Id} - A)^{-1} = \sum_{k \geqslant 0} A^k$, *and we have* $\|(\mathrm{Id} - A)^{-1}\| \leqslant (1 - \|A\|)^{-1}$.

Proof. Since $\|A\| < 1$ the sum $B = \sum_{k \geqslant 0} A^k$ converges and has norm bounded by $\sum_{k \geqslant 0} \|A\|^k = (1 - \|A\|)^{-1}$. Then it suffices to verify that $(\mathrm{Id} - A) \sum_{k \geqslant 0} A^k$ actually converges to Id. □

For any $f \in \mathcal{C}^{\ell}(\Omega, \mathbb{Y})$, any two points a, b in Ω, and any integer $l \in \{0, \ldots, \ell\}$, we write

$$R_l(f; a, b) \;=\; f(b) - \sum_{k=0}^{l} D^k f(a) \, \frac{(b-a)^k}{k!}$$

for the remainder of the Taylor expansion of f to order l, centered at a and evaluated at b. If $l + 1 \leqslant \ell$, and if the segment $[a, b]$ is included in Ω, then it admits the integral form

$$R_l(f; a, b) \;=\; \int_{[a,b]} D^{l+1} f(z) \, \frac{(b-z)^l}{l!} \, \mathrm{d}z.$$

From now, x_0 is a point in Ω such that $D f(x_0)$ is invertible. We assume we are given a constant $\beta \geqslant \|D f(x_0)^{-1} f(x_0)\|$, and a *continuous non-negative* and *non-decreasing* function $L: [0, R] \rightarrow \mathbb{R}_{\geqslant 0}$ satisfying the following Lipschitzian condition:

$$\|D f(x_0)^{-1} (D f(b) - D f(a))\| \;\leqslant\; L(r) \, \|b - a\|, \text{ for all } r \in [0, R] \text{ and all } a, b \in \bar{B}(x_0, r) \cap \Omega. \tag{1}$$

We consider the function

$$\phi(r) \;=\; \beta - r + \int_0^r L(s) \, (r - s) \, \mathrm{d}s, \tag{2}$$

which is defined in $[0,\ R]$. In order to compute its first derivative, we take a parameter ε in a neighborhood of 0, and calculate: $\int_0^{r+\varepsilon} L(s)\,(r+\varepsilon-s)\,\mathrm{d}s - \int_0^r L(s)\,(r-s)\,\mathrm{d}s = \int_r^{r+\varepsilon} L(s)\,(r+\varepsilon-s)\,\mathrm{d}s + \int_0^r L(s)\,((r+\varepsilon-s)-(r-s))\,\mathrm{d}s = \varepsilon \int_0^r L(s)\,\mathrm{d}s + O(\varepsilon^2)$. We thus see that ϕ admits continuous derivatives to order 2 on $(0,R)$: $\phi'(r) = -1 + \int_0^r L(s)\,\mathrm{d}s$ and $\phi''(r) = L(r)$. These derivatives naturally extend continuously at 0 from right and at R from left.

Lemma 2. *Condition (1) is equivalent to: for all segment $[a,b] \subset B(x_0,R)$ such that $\|a-x_0\| + \|b-a\| \leqslant R$,*

$$\|Df(x_0)^{-1}\,(Df(b)-Df(a))\| \;\leqslant\; \int_{\|x_0-a\|}^{\|x_0-a\|+\|b-a\|} L(s)\,\mathrm{d}s. \tag{3}$$

Proof. We let $r_a = \|a-x_0\|$ and $r_b = r_a + \|b-a\|$. We divide the segment $[a,b]$ into N consecutive subsegments $[c_i,\ c_{i+1}]$ where $c_i = a + i\ \frac{b-a}{N}$. We also let $r_i = r_a + i\ \frac{r_b - r_a}{N}$, so that we have $\|c_{i+1}-c_i\| = r_{i+1} - r_i$ and $\max(\|c_{i+1}-x_0\|, \|c_i - x_0\|) \leqslant r_{i+1}$.

Assume that (1) holds, and apply it on each $[c_i, c_{i+1}]$ as follows:

$$\|Df(x_0)^{-1}\,(Df(b)-Df(a))\| \;=\; \left\| \sum_{i=0}^{N-1} Df(x_0)^{-1}\,(Df(c_{i+1})-Df(c_i)) \right\| \;\leqslant\; \sum_{i=0}^{N-1} L(r_{i+1})\,\|c_{i+1}-c_i\|.$$

The latter sum converges to $\int_{r_a}^{r_b} L(s)\,\mathrm{d}s$ when N tends to infinity, which gives the first implication.

Conversely, assume that condition (3) holds. Without loss of generality we may assume that $\|b-x_0\| \geqslant \|a-x_0\|$. Then, with N sufficiently large, precisely such that $\|b-x_0\| + \|b-a\|/N \leqslant R$, we have $\|Df(x_0)^{-1}\ (Df(b) - Df(a))\| \leqslant \sum_{i=0}^{N-1} \|Df(x_0)^{-1}\ (Df(c_{i+1}) - Df(c_i))\| \leqslant \sum_{i=0}^{N-1} \int_{\|c_i-x_0\|}^{\|c_i-x_0\|+\|c_{i+1}-c_i\|} L(s)\,\mathrm{d}s \leqslant \sum_{i=0}^{N-1} L(\|c_i-x_0\|+\|c_{i+1}-c_i\|)\,\|c_{i+1}-c_i\| \leqslant L(\|b-x_0\| + \|b-a\|/N)\,\|b-a\|$. The latter expression converges to $L(\|b-x_0\|)\,\|b-a\|$ when N tends to infinity. □

Lemma 3. *For all segment $[a,b] \subset B(x_0,R)$ such that $\|a-x_0\| + \|b-a\| \leqslant R$, we have:*

$$\|R_1(Df(x_0)^{-1}\,f;a,b)\| \;\leqslant\; R_1(\phi; \|a-x_0\|, \|a-x_0\| + \|b-a\|).$$

Proof. We let $r_a = \|a-x_0\|$, $r_b = r_a + \|b-a\|$, and use Lemma 2 as follows:

$$\begin{aligned} \|R_1(Df(x_0)^{-1}\,f;a,b)\| &= \|Df(x_0)^{-1}\,(f(b)-f(a)-Df(a)\,(b-a))\| \\ &= \left\| \int_a^b Df(x_0)^{-1}\,(Df(z)-Df(a))\,\mathrm{d}z \right\| \\ &\leqslant \int_{r_a}^{r_b}\int_{r_a}^{r} L(s)\,\mathrm{d}r\,\mathrm{d}s \;=\; \int_{r_a}^{r_b} (\phi'(r)-\phi'(r_a))\,\mathrm{d}r \;=\; R_1(\phi; r_a, r_b). \end{aligned}$$ □

Built on these lemmas, the following theorem gives necessary conditions that ensure convergence to a zero, and also uniqueness of this zero in a larger region. The central idea is the comparison of the convergence of the Newton iterates for f with the ones for ϕ.

Theorem 4. *Let $f \in \mathcal{C}^1(\Omega, \mathbb{Y})$, and let $x_0 \in \Omega$ be such that $Df(x_0)$ is invertible. We assume we are given a constant $\beta \geqslant \|Df(x_0)^{-1}\,f(x_0)\|$, and a continuous non-negative and non-decreasing function L: $[0,\ R] \to \mathbb{R}_{\geqslant 0}$ satisfying (1) and $B(x_0,R) \subseteq \Omega$. The function ϕ, as defined in (2), is supposed to admit a unique zero r_- in $[0,R)$, and to satisfy $\phi(R) \leqslant 0$.*

Then the Newton sequence $r_0 = 0$, $r_{k+1} = r_k - \frac{\phi(r_k)}{\phi'(r_k)}$ *is well defined in* $[0, r_-]$, *and converges to* r_-. *The sequence* $x_{k+1} = x_k - Df(x_k)^{-1} f(x_k)$ *is also well defined in* $\bar{B}(x_0, r_-)$, *and converges to the unique zero* ζ *of* f *in* $B(x_0, R)$. *In addition, we have* $\|\zeta - x_k\| \leqslant r_- - r_k$ *and* $\|x_{k+1} - x_k\| \leqslant r_{k+1} - r_k$.

Proof. First, we examine the convergence of the sequence $(r_k)_{k\geqslant 0}$. Since $\phi'' \geqslant 0$, it is classical that the sequence $(r_k)_{k\geqslant 0}$ is non-decreasing, remains in $[0, r_-]$, and therefore converges to r_-, as pictured in Figure 1.

We shall prove by induction that $\|x_{k+1} - x_k\| \leqslant r_{k+1} - r_k$ holds for all $k \geqslant 0$. For $k = 0$ this is true because $\|x_1 - x_0\| = \|Df(x_0)^{-1} f(x_0)\| \leqslant \beta = r_1 - r_0$. Now assume that the inequality holds up to some $k \geqslant 0$, and let us prove that it also holds for $k+1$. In order to bound $\|x_{k+1} - x_k\| = \|Df(x_k)^{-1} f(x_k)\|$, we bound $\|Df(x_k)^{-1} Df(x_0)\|$ and $\|Df(x_0)^{-1} f(x_k)\|$ separately. As for the first expression, using the induction hypothesis, we obtain

$$\|x_k - x_0\| = \sum_{i=0}^{k-1} \|x_{i+1} - x_i\| \leqslant \sum_{i=0}^{k-1} (r_{i+1} - r_i) = r_k - r_0 = r_k \leqslant r_-,$$

so that Lemma 2 gives us $\|Df(x_0)^{-1}(Df(x_k) - Df(x_0))\| \leqslant 1 + \phi'(r_k) < 1$, and Lemma 1 implies that $Df(x_k)$ is invertible with norm

$$\|Df(x_k)^{-1} Df(x_0)\| \leqslant \frac{1}{1 - \|Df(x_0)^{-1}(Df(x_k) - Df(x_0))\|} \leqslant \frac{1}{\phi'(r_k)}. \tag{4}$$

Consequently x_{k+1} is well-defined. Then, in order to bound $\|Df(x_0)^{-1} f(x_k)\|$, we write the Taylor expansion of f at x_{k-1}, and use the definition of x_k:

$$f(x_k) = f(x_{k-1}) + Df(x_{k-1})(x_k - x_{k-1}) + R_1(f; x_{k-1}, x_k) = R_1(f; x_{k-1}, x_k).$$

Combining Lemma 3, inequality $\|x_k - x_{k-1}\| \leqslant r_k - r_{k-1}$, and the definition of r_k, we obtain:

$$\|R_1(Df(x_0)^{-1} f; x_{k-1}, x_k)\| \leqslant R_1(\phi; r_{k-1}, r_k) = \phi(r_k) - \phi(r_{k-1}) - \phi'(r_{k-1})(r_k - r_{k-1}) = \phi(r_k).$$

We thus have achieved $\|Df(x_0)^{-1} f(x_k)\| \leqslant \phi(r_k)$, which combined to (4) leads to

$$\|x_{k+1} - x_k\| \leqslant -\frac{\phi(r_k)}{\phi'(r_k)} = r_{k+1} - r_k,$$

whence the induction hypothesis at $k+1$. At this point of the proof we know that $(x_k)_{k\geqslant 0}$ is a Cauchy sequence in $\bar{B}(x_0, r_-)$. Consequently it converges to a zero ζ of f in $\bar{B}(x_0, r_-)$. It remains to show that ζ is the unique zero of f in $B(x_0, R)$.

Let ξ be a zero of f in $B(x_0, R)$, and let $\theta = \frac{\|\xi - x_0\|}{R} < 1$. We shall prove by induction that $\|\xi - x_k\| \leqslant \theta^{2^k}(R - r_k)$ holds for all $k \geqslant 0$, which will yield $\zeta = \xi$. The induction hypothesis clearly holds for $k = 0$. Assume that it holds up to some value of $k \geqslant 0$.

Writing $x_{k+1} - \xi = Df(x_k)^{-1}(f(\xi) - f(x_k) - Df(x_k)(\xi - x_k))$, we aim at bounding $\|Df(x_k)^{-1} Df(x_0)\|$ and $\|R_1(Df(x_0)^{-1} f; x_k, \xi)\|$. Using that $\phi'' = L$ is non-decreasing, the latter bound can been achieved *via* Lemma 3:

$$\begin{aligned}\|R_1(Df(x_0)^{-1} f; x_k, \xi)\| &\leqslant R_1(\phi; r_k, r_k + \|\xi - x_k\|) = \int_{r_k}^{r_k + \|\xi - x_k\|} \phi''(s)(r_k + \|\xi - x_k\| - s)\,\mathrm{d}s \\ &\leqslant \theta^{2^{k+1}} \int_{r_k}^{R} \phi''(s)(R - s)\,\mathrm{d}s = \theta^{2^{k+1}}(\phi(R) - \phi(r_k) - \phi'(r_k)(R - r_k)).\end{aligned}$$

Combined to inequality (4), we deduce $\|\xi - x_{k+1}\| \leqslant \theta^{2^{k+1}} \frac{\phi(R) - \phi(r_k) - \phi'(r_k)\,(R - r_k)}{-\phi'(r_k)}$. Since $\phi(R) \leqslant 0$, this yields $\|\xi - x_{k+1}\| \leqslant \theta^{2^{k+1}}\,(R - r_{k+1})$, whence the induction hypothesis at $k+1$. □

3 Other criteria

In this section we show how the latter theorem allows one to retrieve both Kantorovich's original theorem, subsequent variants for higher orders, and recent formulations in terms of majorant series.

3.1 Order two

In order to use Theorem 4 in practice, it is worth considering functions L, and thus ϕ, which are polynomial of low degrees. Taking ϕ of degree one would force L to be identically 0 hence $D f$ to be constant which is not of interest. Taking ϕ of degree two corresponds to the original case due to Kantorovich.

Corollary 5. *Let $f \in \mathcal{C}^1(\Omega, \mathbb{Y})$, and let $x_0 \in \Omega$ be such that $D f(x_0)$ is invertible. We assume we are given constants β, λ satisfying $\beta \geqslant \|D f(x_0)^{-1}\, f(x_0)\|$, $0 < \beta\,\lambda < 1/2$, $B(x_0, r_+) \subset \Omega$, and such that for all $a, b \in B(x_0, r_+)$,*

$$\|D f(x_0)^{-1}\,(D f(b) - D f(a))\| \leqslant \lambda\,\|b - a\|, \text{ where } r_- = \frac{2\,\beta}{1 + \sqrt{1 - 2\,\beta\,\lambda}} \text{ and } r_+ = \frac{1 + \sqrt{1 - 2\,\beta\,\lambda}}{\lambda}.$$

Then, with $\varphi(r) = \lambda\,r^2/2 - r + \beta$, the Newton sequence $r_0 = 0$, $r_{k+1} = r_k - \frac{\varphi(r_k)}{\varphi'(r_k)}$ is well defined in $[0, r_-]$, and converges to r_-. The sequence $(x_k)_{k \geqslant 0}$ defined by $x_{k+1} = x_k - D f(x_k)^{-1}\, f(x_k)$ is well defined in $\bar{B}(x_0, r_-)$ and converges to the unique zero ζ of f in $B(x_0, r_+)$.

Proof. We simply invoke Theorem 4 with $L(r) = \lambda$, $R = r_+$, so that $\phi = \varphi$. □

This criterion is clearly sharp for equations of degree two, and more precisely when $f = \varphi$. This corollary may also be completed with an explicit formula for $(r_k)_{k \geqslant 0}$, which is obtained from the auxiliary sequence $t_k = \frac{r_k - r_-}{r_k - r+}$, that satisfies $t_{k+1} = t_k^2$.

Example 6. Consider $\mathbb{X} = \mathbb{Y} = \mathbb{C}$, $f(x) = x^3/128 + x^2/4 - x + 9/10$, $x_0 = 0$, and $\beta = |f(0)/f'(0)| = 9/10$. Since $f''(x) = 3\,x/64 + 1/2$, for all candidate value for R, one necessarily takes λ larger than $3\,R/64 + 1/2$. Since the closest root to x_0 is $\zeta \simeq 1.4475$, Corollary 5 does not apply. However we shall show later that the Newton iterates of x_0 converge to ζ.

3.2 Higher orders

Now we examine what happens when $\ell \geqslant 2$. We still assume we are given a constant $\beta \geqslant \|D f(x_0)^{-1}\, f(x_0)\|$, but also additional constants $\gamma_i \geqslant \|D f(x_0)^{-1}\, D^i f(x_0)\|$ for $i \in \{2, \ldots, \ell\}$, and a continuous non-negative and non-decreasing function L_ℓ: $[0, R] \to \mathbb{R}_{\geqslant 0}$ satisfying:

$$\|D f(x_0)^{-1}\,(D^\ell f(b) - D^\ell f(a))\| \leqslant L_\ell(r)\,\|b - a\|, \text{ for all } r \in [0, R] \text{ and all } a, b \in \bar{B}(x_0, r) \cap \Omega. \tag{5}$$

We consider the function $\phi_\ell(r) = \beta - r + \gamma_2 \frac{r^2}{2!} + \cdots + \gamma_\ell \frac{r^\ell}{\ell!} + \int_0^r L_\ell(s) \frac{(r - s)^\ell}{\ell!}\,\mathrm{d}s$, defined in $[0, R]$. In order to compute its derivatives, we take a parameter ε in a neighborhood of 0, and calculate: $\int_0^{r+\varepsilon} L_\ell(s) \frac{(r + \varepsilon - s)^\ell}{\ell!}\,\mathrm{d}s - \int_0^r L_\ell(s) \frac{(r - s)^\ell}{\ell!}\,\mathrm{d}s = \int_r^{r+\varepsilon} L_\ell(s) \frac{(r + \varepsilon - s)^\ell}{\ell!}\,\mathrm{d}s + \int_0^r L_\ell(s) \left(\frac{(r + \varepsilon - s)^\ell}{\ell!} - \frac{(r - s)^\ell}{\ell!}\right) \mathrm{d}s = \varepsilon \int_0^r L_\ell(s) \frac{(r - s)^{\ell - 1}}{(\ell - 1)!}\,\mathrm{d}s + O(\varepsilon^2)$. By a straightforward induction, this shows that the derivative to order $l \leqslant \ell$ of $\int_0^r L_\ell(s) \frac{(r - s)^\ell}{\ell!}\,\mathrm{d}s$ exists and equals $\int_0^r L_\ell(s) \frac{(r - s)^{\ell - l}}{(\ell - l)!}\,\mathrm{d}s$. Consequently ϕ_ℓ is of class $\mathcal{C}^{\ell+1}([0, R], \mathbb{R})$.

Corollary 7. *Let $f \in \mathcal{C}^{\ell}(\Omega, \mathbb{Y})$, with $\ell \geqslant 2$, and let $x_0 \in \Omega$ be such that $D f(x_0)$ is invertible. We assume given $\beta, \gamma_2, \ldots, \gamma_\ell$, L_ℓ and ϕ_ℓ as defined above, satisfying (5), and such that ϕ_ℓ admits a unique zero r_- in $[0, R)$, with $B(x_0, R) \subseteq \Omega$ and $\phi(R) \leqslant 0$.*

Then the Newton sequence $r_0 = 0$, $r_{k+1} = r_k - \frac{\phi_\ell(r_k)}{\phi_\ell'(r_k)}$ is well defined in $[0, r_-]$, and converges to r_-. The Newton sequence $x_{k+1} = x_k - D f(x_k)^{-1} f(x_k)$ is also well defined in $\bar{B}(x_0, r_-)$, and converges to the unique zero ζ of f in $B(x_0, R)$. In addition, we have $\|\zeta - x_k\| \leqslant r_- - r_k$ and $\|x_{k+1} - x_k\| \leqslant r_{k+1} - r_k$.

Proof. We define $L(r) = \phi_\ell''(r) = \gamma_2 + \cdots + \gamma_\ell \frac{r^{\ell-2}}{(\ell-2)!} + \int_0^r L_\ell(s) \frac{(r-s)^{\ell-2}}{(\ell-2)!} \mathrm{d}s$ and $r_a = \|a - x_0\|$. We claim that $\|D f(x_0)^{-1} D^2 f(a)\| \leqslant L(r_a)$, so that L satisfies hypotheses of Theorem 4 with $\phi(r) = \beta - r + \int_0^r L(s)\,(r-s)\,\mathrm{d}s = \phi_\ell(r)$, which concludes the proof. In order to prove the latter claim, we notice that Lemma 2 applied to $D^{\ell-1} f$ yields $\|D f(x_0)^{-1} (D^\ell f(a) - D^\ell f(x_0))\| \leqslant \int_0^{r_a} L_\ell(s)\,\mathrm{d}s$, and we begin with

$$D^2 f(a) \;=\; \sum_{l=2}^{\ell} D^l f(x_0) \frac{(a - x_0)^{l-2}}{(l-2)!} + R_{\ell-2}(D^2 f; x_0, a).$$

If $\ell = 2$, then $R_{\ell-2}(D f(x_0)^{-1} D^2 f; x_0, a) = D f(x_0)^{-1} (D^2 f(a) - D^2 f(x_0))$, hence has norm at most $\int_0^{r_a} L_\ell(s)\,\mathrm{d}s = R_{\ell-2}(\phi_\ell''; 0, r_a)$. Otherwise, if $\ell \geqslant 3$, the integral form of the Taylor remainder of $D^2 f$ to order $\ell - 3$ leads to

$$\begin{aligned} R_{\ell-2}(D f(x_0)^{-1} D^2 f; x_0, a) &= R_{\ell-3}(D f(x_0)^{-1} D^2 f; x_0, a) - D^\ell f(x_0) \frac{(a - x_0)^{\ell-2}}{(\ell-2)!} \\ &= \int_{[x_0, a]} (D^\ell f(z) - D^\ell f(x_0)) \frac{(a - z)^{\ell-3}}{(\ell-3)!} \mathrm{d}z. \end{aligned}$$

It follows that

$$\|R_{\ell-2}(D f(x_0)^{-1} D^2 f; x_0, a)\| \;\leqslant\; \int_0^{r_a} \int_0^{r} L_\ell(s) \frac{(r_a - r)^{\ell-3}}{(\ell-3)!} \mathrm{d}r\,\mathrm{d}s \;=\; R_{\ell-2}(\phi_\ell''; 0, r_a),$$

whence the claimed bound $\|D f(x_0)^{-1} D^2 f(a)\| \leqslant \sum_{l=2}^{\ell} \gamma_l \frac{r_a^{l-2}}{(l-2)!} + R_{\ell-2}(\phi_\ell''; 0, r_a) = \phi_\ell''(r_a)$. □

The first case of use concerns $\ell = 2$, with $L(r)$ being a constant say ν, so that $\phi_2(r) = \beta - r + \gamma_2 \frac{r^2}{2!} + \nu \frac{r^3}{3!}$. Since $\phi_2(r)$ admits a unique negative root, it admits two distinct positive roots r_- and r_+ if, and only if, its discriminant is positive. In this case, the previous corollary applies with $R = r_+$ in a way similar to the case $\ell = 1$.

Example 8. With f as in Example 6, we may take $\ell = 2$, $\beta = |f'(0)^{-1} f(0)| = 9/10$, $\gamma_2 = 1/2$, $L(r) = \nu = 3/64$, so that $\phi_\ell(r)$ has positive discriminant. It follows that the Newton iterates of x_0 converge quadratically to the root ζ of f, where $\zeta \simeq 1.4475$.

When $\ell = 2$, this corollary is sharp for polynomials of degree 3, and we could build upon it conditions in degree 4, 5, etc. In fact using higher order Kantorovich conditions might be tempting to work for instance with low floating point precision. However computing even rough bounds on high order derivatives becomes as much expensive as the dimension of the ambient space grows up. It is therefore in general recommended to restrict to degree 2.

3.3 Majorant series

In a context of analytic and meromorphic functions, it is natural to consider generating series of norms of derivatives, thus taking poles into accounts. Considering the limit of the previous case when ℓ tends to infinity, we obtain an other corollary:

Corollary 9. *Let $f \in \mathcal{C}^\infty(\Omega, \mathbb{Y})$, and let $x_0 \in \Omega$ be such that $D f(x_0)$ is invertible. We assume we are given a constant $\beta \geqslant \|D f(x_0)^{-1} f(x_0)\|$, and a sequence of constants $\gamma_i \geqslant \|D f(x_0)^{-1} D^i f(x_0)\|$ for $i \geqslant 2$. We define the function $\phi_\infty \colon [0, R] \to \mathbb{R}$ as $\phi_\infty(r) = \beta - r + \sum_{l \geqslant 2} \gamma_l \frac{r^l}{l!}$, assuming that this sum converges, and that ϕ_∞ admits a unique zero $r_- \geqslant 0$ in $[0, R)$, such that $B(x_0, R) \subseteq \Omega$ and $\phi_\infty(R) \leqslant 0$.*

Then the Newton sequence $r_0 = 0$, $r_{k+1} = r_k - \frac{\phi_\infty(r_k)}{\phi'_\infty(r_k)}$ is well defined in $[0, r_-]$, and converges to r_-. The Newton sequence $x_{k+1} = x_k - D f(x_k)^{-1} f(x_k)$ is also well defined in $\bar{B}(x_0, r_-)$, and converges to the unique zero ζ of f in $B(x_0, R)$. In addition, we have $\|\zeta - x_k\| \leqslant r_- - r_k$ and $\|x_{k+1} - x_k\| \leqslant r_{k+1} - r_k$.

Proof. We define $L(r) = \phi''_\infty(r) = \sum_{l \geqslant 2} \gamma_l \frac{r^{l-2}}{(l-2)!}$. By considering the Taylor expansion of $D^2 f$ at x_0, we obtain that $\|D f(x_0)^{-1} D^2 f(a)\| \leqslant L(\|a - x_0\|)$ holds for all $a \in B(x_0, R)$, so that L satisfies the hypothesis of Theorem 4 with $\phi(r) = \beta - r + \int_0^r L(s)\,(r - s)\,\mathrm{d}s = \phi_\infty(r)$. □

The first case of practical interest is for when ϕ_∞ is a rational function with a numerator of degree 2 and a denominator of degree 1. We thus assume given a constant $\gamma \geqslant \left\| D f(x_0)^{-1} \frac{D^l f(x_0)}{l!} \right\|^{\frac{1}{l-1}}$ for all $l \geqslant 2$, and take $\gamma_l = l!\, \gamma^{l-1}$ so that $\phi_\infty(r) = \beta - r + \frac{\gamma r^2}{1 - \gamma r} = \frac{\beta - (\alpha + 1) r + 2 \gamma r^2}{1 - \gamma r}$, where $\alpha = \beta \gamma$. Conditions of Corollary 9 rewrite into $R < 1/\gamma$, and $\alpha < 3 - 2\sqrt{2}$. This special case is known as the *α-Theorem*. This case of Kantorovich theorem has the advantage to fix the parameter R in terms of γ. It therefore turns out to be useful for analyzing the complexity of numerical algorithms. In practice some specific class of functions might benefit from it, such as algebraic and holonomic functions, where one might expect to rely on external machinery to compute candidate values for γ.

In terms of α, β, and γ, the sequence $(r_k)_{k \geqslant 0}$ may be computed explicitly by introducing the auxiliary sequence $t_k = \frac{r_k - r_-}{r_k - r+}$:

$$t_{k+1} = \frac{r_k - r_- - \frac{\phi_\infty(r_k)}{\phi'_\infty(r_k)}}{r_k - r_+ - \frac{\phi_\infty(r_k)}{\phi'_\infty(r_k)}} = t_k \frac{\frac{\phi'_\infty(r_k)}{\phi_\infty(r_k)} - \frac{1}{r_k - r_-}}{\frac{\phi'_\infty(r_k)}{\phi_\infty(r_k)} - \frac{1}{r_k - r_+}} = t_k \frac{\frac{1}{r_k - r_+} + \frac{\gamma}{1 - \gamma r_k}}{\frac{1}{r_k - r_-} + \frac{\gamma}{1 - \gamma_{r_k}}} = t_k^2 \frac{1 - \gamma r_+}{1 - \gamma r_-}.$$

4 Historical notes

Corollary 5 essentially corresponds to the first occurrence of Kantorovich's theorem in the literature, published in [38, Глава IV, p. 170], which was requiring $f \in \mathcal{C}^2(\Omega, \mathbb{X})$ and was using separate bounds on $\|D f(x_0)^{-1}\|$ and $\|D^2 f(x_0)\|$. Then variants and improvements have been proposed by various authors [50, 52, 60], before being merged by Gragg and Tapia, who introduced the Lipschitzian condition of Corollary 5, and detailed the limit case $r_- = r_+$, where the quadratic convergence does not hold anymore [33]. Gragg and Tapia also provided the sharp *a priori* convergence bound from the explicit formula for r_k that was borrowed from [51, Appendix F]. Our Section 3.1 is actually inspired from [33].

The extension to degree three (given as an example of Corollary 7) has been first presented in [37]. Explicit convergence bounds have then been given in [35]. Theorem 4 first appeared in [65], under assumption (3), which is shown to be equivalent to the more classical Lipschitzian condition (1) in our Lemma 2. Corollary 7 is inspired from [25, Theorem 1], which admits variants in [15, 24], that had been developed independently of [65]. In fact our Section 3.2 highlights the fact that higher order assumptions are essentially specializations of the general case handled by Theorem 4.

The α-theorem first appeared in an article by Smale [59] with the non-optimal condition $\alpha < 0.130707$. At the same time a one dimensional version was also designed by Kim [40, 41]. Subsequent improvements of the latter constant are due to Wang and Han [64] (see also [63]). Wang also made explicit the relationship between the α-theorem and Kantorovich's theorem [65]. The systematic treatment in terms of majorant series emerged in [32], from which Corollary 9 is extracted. Important applications to complexity of numeric polynomial system solving started in [56, 57, 58].

Classical books for Kantorovich's theorem and historical notes are [8, 13, 23, 39, 49]. For the α-theorem and its applications to polynomial system solving by homotopy methods, we refer the reader to [18, 21]. Let us also mention the survey [53], and the article [20] for detailled recent proofs of Kantorovich's original theorem with slight variants.

A plethora of literature is dedicated to variations of assumptions on f and its derivatives: Other kinds of Lipschitzian conditions (centered, or in balls or annuli) and comparisons between them [3, 4, 6, 26, 28, 30, 61, 66]; Mixed centered Lipschitzian conditions extending the α-theorem [14]; Weak continuity of the derivative [12]; Hölder conditions [19]. Convergence rate and error bounds have been refined in [5, 47]; *A posteriori* bounds can be found in [54, 69, 70, 71, 72].

Finally, let us mention that Kantorovich's technique has been successfully applied and extended to other Newton-like operators in wider contexts: Robust variant [67]; Modified Newton method [22, 44]; Inexact Newton method [7, 17, 31, 34, 48, 55, 68]; Gauss–Newton method [11, 36, 42, 43]; Halley's method (extension of Newton operator to order 3) [2, 9, 16, 16, 27, 45, 46]; Extensions to differential vector fields on Riemannian manifolds [1, 10, 29, 62] (Theorem 4 is for instance extended in [1]).

Bibliography

[1] F. Alvarez, J. Bolte, and J. Munier. A unifying local convergence result for Newton's method in Riemannian manifolds. *Found. Comput. Math.*, 8(2):197–226, 2008.

[2] S. Amat, C. Bermúdez, S. Busquier, and S. Plaza. On a third-order Newton-type method free of bilinear operators. *Numer. Linear Algebra Appl.*, 17(4):639–653, 2010.

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

[4] I. K. Argyros. A new semilocal convergence theorem for Newton's method in Banach space using hypotheses on the second Fréchet-derivative. *J. Comput. Appl. Math.*, 130(1-2):369–373, 2001.

[5] I. K. Argyros and D. González. Extending the applicability of Newton's method by improving a local result due to Dennis and Schnabel. *S$\vec{e}$MA J.*, 63:53–63, 2014.

[6] I. K. Argyros and S. K. Khattri. Weaker Kantorovich type criteria for inexact Newton methods. *J. Comput. Appl. Math.*, 261:103–117, 2014.

[7] Ioannis K. Argyros. Concerning the convergence of inexact Newton methods. *J. Comput. Appl. Math.*, 79(2):235–247, 1997.

[8] Ioannis K. Argyros. *Convergence and applications of Newton-type iterations*. Springer, New York, 2008.

[9] Ioannis K. Argyros, Yeol Je Cho, and Saïd Hilout. On the semilocal convergence of the Halley method using recurrent functions. *J. Appl. Math. Comput.*, 37(1-2):221–246, 2011.

[10] Ioannis K. Argyros and Saïd Hilout. Newton's method for approximating zeros of vector fields on Riemannian manifolds. *J. Appl. Math. Comput.*, 29(1-2):417–427, 2009.

[11] Ioannis K. Argyros and Saïd Hilout. Extending the applicability of the Gauss-Newton method under average Lipschitz-type conditions. *Numer. Algorithms*, 58(1):23–52, 2011.

[12] Ioannis K. Argyros and Saïd Hilout. Newton-Kantorovich approximations under weak continuity conditions. *J. Appl. Math. Comput.*, 37(1-2):361–375, 2011.

[13] Ioannis K. Argyros and Saïd Hilout. *Computational methods in nonlinear analysis. Efficient algorithms, fixed point theory and applications*. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2013.

[14] Ioannis K. Argyros, Saïd Hilout, and Sanjay K. Khattri. Expanding the applicability of Newton's method using Smale's α-theory. *J. Comput. Appl. Math.*, 261:183–200, 2014.

[15] Ioannis K. Argyros and Ángel Alberto Magreñán Ruiz. General convergence conditions of Newton's method for m-Fréchet differentiable operators. *J. Appl. Math. Comput.*, 43(1-2):491–506, 2013.

[16] Ioannis K. Argyros and Hongmin Ren. Ball convergence theorems for Halley's method in Banach space. *J. Appl. Math. Comput.*, 38(1-2):453–465, 2012.

[17] Ioannis K. Argyros and Livinus U. Uko. A semilocal convergence analysis of an inexact Newton method using recurrent relations. *Punjab Univ. J. Math. (Lahore)*, 45:25–32, 2013.

[18] Lenore Blum, Felipe Cucker, Michael Shub, and Steve Smale. *Complexity and real computation*. Springer-Verlag, New York, 1998.

[19] Filomena Cianciaruso. A further journey in the "terra incognita" of the Newton-Kantorovich method. *Nonlinear Funct. Anal. Appl.*, 15(2):173–183, 2010.

[20] Philippe G. Ciarlet and Cristinel Mardare. On the Newton-Kantorovich theorem. *Anal. Appl. (Singap.)*, 10(3):249–269, 2012.

[21] Jean-Pierre Dedieu. *Points fixes, zéros et la méthode de Newton*, volume 54 of *Mathématiques & Applications (Berlin)*. Springer, Berlin, 2006.

[22] J. E. Dennis, Jr. On the Kantorovich hypothesis for Newton's method. *SIAM J. Numer. Anal.*, 6:493–507, 1969.

[23] J. E. Dennis, Jr. and Robert B. Schnabel. *Numerical methods for unconstrained optimization and nonlinear equations*, volume 16 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1996. Corrected reprint of the 1983 original.

[24] J. A. Ezquerro, D. González, and M. A. Hernández. A modification of the classic conditions of Newton-Kantorovich for Newton's method. *Math. Comput. Modelling*, 57(3-4):584–594, 2013.

[25] J. A. Ezquerro, D. González, and M. A. Hernández. On the local convergence of Newton's method under generalized conditions of Kantorovich. *Appl. Math. Lett.*, 26(5):566–570, 2013.

[26] J. A. Ezquerro, D. González, and M. A. Hernández-Verón. A semilocal convergence result for Newton's method under generalized conditions of Kantorovich. *J. Complexity*, 30(3):309–324, 2014.

[27] J. A. Ezquerro, J. M. Gutiérrez, and M. A. Hernández. A construction procedure of iterative methods with cubical convergence. II. Another convergence approach. *Appl. Math. Comput.*, 92(1):59–68, 1998.

[28] José Antonio Ezquerro, Daniel González, and Miguel Ángel Hernández. A general semilocal convergence result for Newton's method under centered conditions for the second derivative. *ESAIM Math. Model. Numer. Anal.*, 47(1):149–167, 2013.

[29] O. P. Ferreira and B. F. Svaiter. Kantorovich's theorem on Newton's method in Riemannian manifolds. *J. Complexity*, 18(1):304–329, 2002.

[30] O. P. Ferreira and B. F. Svaiter. Kantorovich's majorants principle for Newton's method. *Comput. Optim. Appl.*, 42(2):213–229, 2009.

[31] O. P. Ferreira and B. F. Svaiter. A robust Kantorovich's theorem on the inexact Newton method with relative residual error tolerance. *J. Complexity*, 28(3):346–363, 2012.

[32] M. Giusti, G. Lecerf, B. Salvy, and J.-C. Yakoubsohn. On location and approximation of clusters of zeros: case of embedding dimension one. *Found. Comput. Math.*, 7(1):1–49, 2007.

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

[34] Xue-Ping Guo and Iain S. Duff. Semilocal and global convergence of the Newton-HSS method for systems of nonlinear equations. *Numer. Linear Algebra Appl.*, 18(3):299–315, 2011.

[35] José M. Gutiérrez. A new semilocal convergence theorem for Newton's method. *J. Comput. Appl. Math.*, 79(1):131–145, 1997.

[36] W. M. Häussler. A Kantorovich-type convergence analysis for the Gauss-Newton-method. *Numer. Math.*, 48(1):119–125, 1986.

[37] Zheng Da Huang. A note on the Kantorovich theorem for Newton iteration. *J. Comput. Appl. Math.*, 47(2):211–217, 1993.

[38] L. V. Kantorovič. Functional analysis and applied mathematics. *Uspehi Matem. Nauk (N.S.)*, 3(6(28)):89–185, 1948.

[39] L. V. Kantorovich and G. P. Akilov. *Functional analysis in normed spaces*. Translated from the Russian by D. E. Brown. Edited by A. P. Robertson. International Series of Monographs in Pure and Applied Mathematics, Vol. 46. The Macmillan Co., New York, 1964.

[40] Myong-Hi Kim. *Computational Complexity of the Euler Type Algorithms for the Roots of Complex Polynomials*. PhD thesis, City University of New York, 1985.

[41] Myong-Hi Kim. On approximate zeros and rootfinding algorithms for a complex polynomial. *Math. Comp.*, 51(184):707–719, 1988.

[42] S. Kim. A Kantorovich-type convergence analysis for the quasi-Gauss-Newton method. *J. Korean Math. Soc.*, 33(4):865–878, 1996.

[43] Chong Li, Nuchun Hu, and Jinhua Wang. Convergence behavior of Gauss-Newton's method and extensions of the Smale point estimate theory. *J. Complexity*, 26(3):268–295, 2010.

[44] Rongfei Lin, Yueqing Zhao, Zdeněk Šmarda, Yasir Khan, and Qingbiao Wu. Newton-Kantorovich and Smale uniform type convergence theorem for a deformed Newton method in Banach spaces. *Abstr. Appl. Anal.*, pages Art. ID 923898, 8, 2013.

[45] Rongfei Lin, Yueqing Zhao, Zdeněk Šmarda, Qingbiao Wu, and Yasir Khan. Newton-Kantorovich convergence theorem of a new modified Halley's method family in a Banach space. *Adv. Difference Equ.*, pages 2013:325, 11, 2013.

[46] Rongfei Lin, Yueqing Zhao, Qingbiao Wu, and Jueliang Hu. Convergence theorem for a family of new modified Halley's method in Banach space. *J. Appl. Math.*, pages Art. ID 468694, 8, 2014.

[47] G. J. Miel. The Kantorovich theorem with optimal error bounds. *Amer. Math. Monthly*, 86(3):212–215, 1979.

[48] Igor Moret. A Kantorovich-type theorem for inexact Newton methods. *Numer. Funct. Anal. Optim.*, 10(3-4):351–365, 1989.

[49] J. M. Ortega and W. C. Rheinboldt. *Iterative solution of nonlinear equations in several variables*, volume 30 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2000. Reprint of the 1970 original.

[50] James M. Ortega. The Newton-Kantorovich theorem. *Amer. Math. Monthly*, 75(6):658–660, 1968.

[51] A. M. Ostrowski. *Solution of equations and systems of equations*. Second edition. Pure and Applied Mathematics, Vol. 9. Academic Press, New York-London, 1966.

[52] Alexandre Ostrowski. La méthode de Newton dans les espaces de Banach. *C. R. Acad. Sci. Paris Sér. A-B*, 272:A1251–A1253, 1971.

[53] B. T. Polyak. Newton-Kantorovich method and its global convergence. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)*, 312(Teor. Predst. Din. Sist. Komb. i Algoritm. Metody. 11):256–274, 316, 2004.

[54] Florian-A. Potra and Vlastimil Pták. Sharp error bounds for Newton's process. *Numer. Math.*, 34(1):63–72, 1980.

[55] Weiping Shen and Chong Li. Kantorovich-type convergence criterion for inexact Newton methods. *Appl. Numer. Math.*, 59(7):1599–1611, 2009.

[56] M. Shub and S. Smale. Computational complexity: on the geometry of polynomials and a theory of cost. II. *SIAM J. Comput.*, 15(1):145–161, 1986.

[57] Mike Shub and Steven Smale. Computational complexity. On the geometry of polynomials and a theory of cost. I. *Ann. Sci. École Norm. Sup. (4)*, 18(1):107–142, 1985.

[58] Steve Smale. The fundamental theorem of algebra and complexity theory. *Bull. Amer. Math. Soc. (N.S.)*, 4(1):1–36, 1981.

[59] Steve Smale. Newton's method estimates from data at one point. In *The merging of disciplines: new directions in pure, applied, and computational mathematics (Laramie, Wyo., 1985)*, pages 185–196. Springer, New York, 1986.

[60] R. A. Tapia. Classroom Notes: The Kantorovich Theorem for Newton's Method. *Amer. Math. Monthly*, 78(4):389–392, 1971.

[61] Livinus U. Uko and Ioannis K. Argyros. A generalized Kantorovich theorem on the solvability of nonlinear equations. *Aequationes Math.*, 77(1-2):99–105, 2009.

[62] J. H. Wang. Convergence of Newton's method for sections on Riemannian manifolds. *J. Optim. Theory Appl.*, 148(1):125–145, 2011.

[63] Xing Hua Wang. Some results relevant to Smale's reports. In *From Topology to Computation: Proceedings of the Smalefest (Berkeley, CA, 1990)*, pages 456–465. Springer, New York, 1993.

[64] Xing Hua Wang and Dan Fu Han. On dominating sequence method in the point estimate and Smale theorem. *Sci. China Ser. A*, 33(2):135–144, 1990.

[65] Xinghua Wang. Convergence of Newton's method and inverse function theorem in Banach space. *Math. Comp.*, 68(225):169–186, 1999.

[66] Xinghua Wang, Chong Li, and Ming-Jun Lai. A unified convergence theory for Newton-type methods for zeros of nonlinear operators in Banach spaces. *BIT*, 42(1):206–213, 2002.

[67] Zhengyu Wang and Xinyuan Wu. A semi-local convergence theorem for a robust revised Newton's method. *Comput. Math. Appl.*, 58(7):1320–1327, 2009.

[68] Xiubin Xu, Yuan Xiao, and Tao Liu. Semilocal convergence analysis for inexact Newton method under weak condition. *Abstr. Appl. Anal.*, pages Art. ID 982925, 13, 2012.

[69] T. Yamamoto. A unified derivation of several error bounds for Newton's process. In *Proceedings of the international conference on computational and applied mathematics (Leuven, 1984)*, volume 12/13, pages 179–191, 1985.

[70] Tetsuro Yamamoto. A method for finding sharp error bounds for Newton's method under the Kantorovich assumptions. *Numer. Math.*, 49(2-3):203–220, 1986.

[71] Tetsuro Yamamoto. A convergence theorem for Newton-like methods in Banach spaces. *Numer. Math.*, 51(5):545–557, 1987.

[72] Tetsuro Yamamoto and Xiao Jun Chen. Ball-convergence theorems and error estimates for certain iterative methods for nonlinear equations. *Japan J. Appl. Math.*, 7(1):131–143, 1990.