Kantorovich's Theorem on Newton's Method

O. P. Ferreira* B. F. Svaiter†

March 09, 2007

Abstract

In this work we present a simplifyed proof of Kantorovich's Theorem on Newton's Method. This analysis uses a technique which has already been used for obtaining new extensions of this theorem.
AMSC: 49M15, 90C30.

1 Introduction

Kantorovich's Theorem assumes semi-local conditions to ensure existence and uniqueness of a solution of a nonlinear equation $F(x) = 0$, where F is a differentiable application between Banach spaces [5, 6, 7, 12]. This theorem uses *constructively* Newton method and also guarantee convergence to a solution of this iterative procedure. Apart from the elegance of this theorem, it has many theoretical and practical applications, in [10] we can find a reviews of recent applications and in [11] an application in interior point methods. This theorem has also many extensions, some of then encompassing previously unrelated results see [1, 13]. Some of these generalizations and extensions are quite recent, because in the last few year the Kantorovich's Theorem has been the subject of intense research, see [1, 2, 3, 10, 11, 13].

The aim of this paper is to present a new technique for the analysis of the Kantorovich's Theorem. This technique, was introduced in [2] and since then it has been used for obtaining new extensions of Kantorovich's Theorem see [1, 3]. Here, it will be used to present a simplified proof of its "classical" formulation.

The main idea is to define "good" regions for Newton method, by comparing the nonlinear function F with its scalar majorant function. Once these good regions are obtained, an invariant set for Newton method is also obtained and there, Newton iteration can be repeated indefinitely.

*IME/UFG, Campus II- Caixa Postal 131, CEP 74001-970 - Goiânia, GO, Brazil (E-mail:orizon@mat.ufg.br). The author was supported in part by FUNAPE/UFG, PADCT-CNPq, PRONEX–Optimization(FAPERJ/CNPq), CNPq Grant 302618/2005-8, CNPq Grant 475647/2006-8 and IMPA.

†IMPA, Estrada Dona Castorina, 110, Jardim Botânico, CEP 22460-320 - Rio de Janeiro, RJ, Brazil (E-mail:benar@impa.br). The author was supported in part by CNPq Grant 301200/93-9(RN), CNPq Grant 475647/2006-8 and by PRONEX–Optimization(FAPERJ/CNPq).

The following notation is used throughout our presentation. Let X be a Banach space. The open and closed ball at $x \in X$ are denoted, respectively by

$$B(x,r) = \{y \in X;\ \|x-y\| < r\} \quad \text{and} \quad B[x,r] = \{y \in X;\ \|x-y\| \leqslant r\}.$$

For the Frechet derivative of a mapping F we use the notation F' and for the Dual space of X we use X^*.

First, let us recall Kantorovich's theorem on Newton's method in its classical formulation, see [4, 6, 8, 9, 11].

Theorem 1. *Let X, Y be Banach spaces, $C \subseteq X$ and $F : C \to Y$ a continuous function, continuously differentiable on* $\operatorname{int}(C)$*. Take $x_0 \in \operatorname{int}(C)$, L, $b > 0$ and suppose that*

1) $F'(x_0)$ *is non-singular,*

2) $\|F'(x_0)^{-1}[F'(y) - F'(x)]\| \le L\|x-y\|$ *for any $x, y \in C$,*

3) $\|F'(x_0)^{-1}F(x_0)\| \le b$,

4) $2bL \le 1$.

Define

$$t_* := \frac{1-\sqrt{1-2bL}}{L}, \qquad t_{**} := \frac{1+\sqrt{1-2bL}}{L}. \tag{1}$$

If

$$B[x_0, t_*] \subset C,$$

then the sequences $\{x_k\}$ generated by Newton's Method for solving $F(x) = 0$ with starting point x_0,

$$x_{k+1} = x_k - F'(x_k)^{-1}F(x_k), \qquad k = 0, 1, \cdots, \tag{2}$$

is well defined, is contained in $B(x_0, t_*)$, converges to a point $x_* \in B[x_0, t_*]$ which is the unique zero of F in $B[x_0, t_*]$ and

$$\|x_* - x_{k+1}\| \le \frac{1}{2}\|x_* - x_k\|, \qquad k = 0, 1, \cdots. \tag{3}$$

Moreover, if assumption **4** *holds as an strict inequality, i.e. $2bL < 1$, then*

$$\|x_* - x_{k+1}\| \le \frac{1-\theta^{2^k}}{1+\theta^{2^k}} \frac{L}{2\sqrt{1-2bL}}\|x_* - x_k\|^2 \le \frac{L}{2\sqrt{1-2bL}}\|x_* - x_k\|^2, \quad k = 0, 1, \cdots, \tag{4}$$

where $\theta := t_*/t_{**} < 1$, and x_* is the unique zero of F in $B[x_0, \rho]$ for any ρ such that

$$t_* \le \rho < t_{**}, \qquad B[x_0, \rho] \subset C.$$

Note that under assumption **1-4**, convergence of $\{x_k\}$ is Q-linear, according to (3). The additional assumption $2bL < 1$ guarantee Q-quadratic convergence, according to (4). This additional assumption also guarantee that x_* is the unique zero of F in $B(x_0, t_{**})$, whenever $B(x_0, t_{**}) \subset C$.

¿From now on, we assume that the hypotheses of Theorem 1 hold, with the exception of $2bL < 1$ which will be considered to hold only when explicitly stated.

2 Kantorovich's Theorem for a scalar quadratic function

In this section we analyze Newton method applied to solve the scalar equation $f(t) = 0$, for $f : \mathbb{R} \to \mathbb{R}$

$$f(t) = \frac{L}{2}t^2 - t + b. \tag{5}$$

The analysis to be performed can also be viewed as Kantorovich's theorem for function f. This function and the sequence generated by Newton method for solving $f(t) = 0$ with starting point t_0,

$$t_0 := 0, \tag{6}$$

both will play an important rule in the analysis of Theorem 1.

Note that the assumptions of Theorem 1 are satisfied in the *very* particular case $F = f$, $X = Y = C = \mathbb{R}$, $x_0 = t_0$. The roots of f are t_* and t_{**}, as defined in (1). As b, $L > 0$,

$$0 < t_* \leq t_{**},$$

with strict inequality between t_* and t_{**} if and only if $2bL < 1$. Hence

- t_* is the unique root of f in $B[t_0, t_*]$,
- if $2bL < 1$, then t_* is the unique root of f in $B(t_0, t_{**})$.

So, the existence and uniqueness part of Theorem 1 for zeros of f holds.

Proposition 2. *The scalar function f has a smallest nonnegative root $t_* \in (0, 1/L]$. Moreover, for any $t \in [0, t_*)$*

$$f(t) > 0, \qquad f'(t) \leq L(t - t_*) < 0.$$

Proof. For the first statement, it remains to prove that $t_* \leq 1/L$, which is a trivial consequence of the assumptions on b and L.

As $f(0) > 0$, f shall be strictly positive in $[0, t_*)$. For the last inequalities, use the inequality $t_* \leq 1/L$ and (5) to obtain

$$f'(t) = Lt - 1 = L(t - 1/L) \leq L(t - t_*).$$

Now, the last inequality follows directly from the assumption $t < t_*$. □

According to Proposition 2, $f'(t) \neq 0$ for all $t \in [0, t_*)$. Therefore, Newton iteration is well defined in $[0, t_*)$. Let us call it n_f,

$$\begin{array}{rcl} n_f : [0, t_*) & \to & \mathbb{R} \\ t & \mapsto & t - f(t)/f'(t). \end{array} \tag{7}$$

Note that, up to now, only *one single iteration* of newton method is well defined in $[0, t_*)$. In principle, Newton iteration could map some $t \in [0, t_*)$ in to $1/L$. In such a case, the second iterate for t would be not defined.

Now, we shall prove that Newton iteration can be repeated indefinitely at any starting point in $[0, t_*)$.

Proposition 3. *For any $t \in [0, t_*)$*

$$t_* - n_f(t) = -\frac{L}{2f'(t)}(t_* - t)^2, \qquad t < n_f(t) < t_*.$$

In particular, n_f maps $[0, t^*)$ in $[0, t^*)$.

Proof. Take $t \in [0, t_*)$. As f is a second-degree polynomial and $f(t_*) = 0$,

$$0 = f(t) + f'(t)(t_* - t) + \frac{L}{2}(t_* - t)^2.$$

Dividing by $f'(t)$ we obtain, after direct rearranging

$$t_* - t + f(t)/f'(t) = -\frac{L}{2f'(t)}(t_* - t)^2.$$

Note that, by (7), the left hand side of the above equation is $t_* - n_f(t)$, which proves the first equality.

Using Proposition 2 we have $f(t) > 0$ and $f'(t) < 0$. Combining these inequalities with definition (7) and the first equality in the proposition, respectively, we obtain $t < n_f(t) < t_*$. The last statement of the Proposition follows directly from these inequalities. □

Proposition 3 shows, in particular, that for any $t \in [0, t_*)$, the sequence $\{n_f^k(t)\}$,

$$n_f^0(t) = t, \quad n_f^{k+1}(t) = n_f\left(n_f^k(t)\right), \qquad k = 0, 1, \cdots.$$

is well defined, strictly increasing, remains in $[0, t_*)$ and so, is convergent. Therefore, Newton method for solving $f(t) = 0$ with starting point $t_0 = 0$ (see (6)) generates an infinite sequence $\{t_k = n_f^k(t_0)\}$, which can be also defined as

$$t_0 = 0, \quad t_{k+1} = n_f(t_k), \qquad k = 0, 1, \cdots. \tag{8}$$

As we already observed, this sequence is strictly increasing, remains in $[0, t_*)$ and converges.

Corollary 4. *The sequence $\{t_k\}$ is well defined, strictly increasing and is contained in $[0, t_*)$. Moreover, it converges Q-linearly to t_*, as follows*

$$t_* - t_{k+1} = \frac{L}{-2f'(t_k)}(t_* - t_k)^2 \le \frac{1}{2}(t_* - t_k), \qquad k = 0, 1, \cdots. \tag{9}$$

If $2bL < 1$, then the sequence $\{t_k\}$ converge Q-quadratically as follows

$$t_* - t_{k+1} = \frac{1 - \theta^{2^k}}{1 + \theta^{2^k}} \frac{L}{2\sqrt{1 - 2bL}}(t_* - t_k)^2 \le \frac{L}{2\sqrt{1 - 2bL}}(t_* - t_k)^2, \qquad k = 0, 1, \cdots, \tag{10}$$

where $\theta = t_*/t_{**} < 1$.

Proof. The first statement of the corollary have already been proved.

Using Proposition 3, we have for any k

$$t_* - n_f(t_k) = \frac{L}{-2f'(t_k)}(t_* - t_k)^2,$$

which combined with (8), yields the equality on (9). As $t_k \in [0, t_*)$, using Proposition 2 we have

$$f'(t_k) \le L(t_k - t_*) < 0.$$

The multiplication of the first above inequality by $(t_* - t_k)/(2f'(t_k)) < 0$ yields the inequality in (9).

Now suppose that $2bL < 1$ or equivalently $t_* < t_{**}$. A closed expression for t_k is available (see, e.g. [[9], Appendix F], [4]) see the Appendix A. In this case

$$t_k = t_* - \frac{\theta^{2^k}}{1-\theta^{2^k}} \frac{2\sqrt{1-2bL}}{L}, \qquad k = 0, 1, \cdots .$$

From above equation we have that

$$f'(t_k) = -\frac{1+\theta^{2^k}}{1-\theta^{2^k}} \frac{\sqrt{1-2bL}}{a}, \qquad k = 0, 1, \cdots .$$

Therefore, to obtain the equality in (10) combine the equality in (9) and latter equality. As $(1-\theta^{2^k})/(1+\theta^{2^k}) \le 1$ the inequality in (10) follows. □

3 Simplifying assumption and convergence

Newton method is invariant under (non-singular) linear transformations. This fact will be used to simplify our analysis. We claim that it is enough to prove Theorem 1 for the case $X = Y$ and $F'(x_0) = I$. Indeed, if $F'(x_0) \ne I$, define

$$G = F'(x_0)^{-1} F.$$

Then, the domain, the roots, the domain of the derivative and the points where the derivative is non-singular are the same for F and G. Moreover, Newton method applied to $F(x) = 0$ is equivalent to Newton methods applied to $G(x) = 0$, i.e., at the points where $F'(x)$ is nonsingular,

$$G'(x)^{-1}G(x) = F'(x)^{-1}F(x), \qquad x - G'(x)^{-1}G(x) = x - F'(x)^{-1}F(x).$$

Finally, G will satisfy the *same* assumptions wich F satisfy. So, from now one we assume

$$X = Y, \qquad F'(x_0) = I. \tag{11}$$

Note that this assumption simplifies conditions **2** and **3** of Theorem 1.

Proposition 5. *If $0 \le t < t_*$ and $x \in B(x_0, t)$, then $F'(x)$ is non-singular and*

$$\left\| F'(x)^{-1} \right\| \le 1/\left| f'(t) \right|.$$

Proof. Recall that $t_* \le 1/L$. Hence, $0 \le t < 1/L$. Using (11) and assumption **2**, with x and x_0 we have

$$\|F'(x) - I\| \le Lt < 1.$$

Hence, using Banach's Lemma, we conclude that $F'(x)$ is non-singular and

$$\left\|F'(x)^{-1}\right\| \le 1/(1 - Lt).$$

To end the proof, use (5) to obtain $|f'(t)| = 1 - Lt$ for $0 \le t < t_*$. □

The error in the first order approximation of F at point $x \in \operatorname{int}(C)$ can be estimated in any $y \in C$, whenever the line segment with extreme points x, y lays in C. Since balls are convex, we have:

Proposition 6. *If $x \in B(x_0, R)$ and $y \in B[x_0, R] \subset C$, then*

$$\left\|F(y) - \left[F(x) + F'(x)(y - x)\right]\right\| \le \frac{L}{2}\|y - x\|^2.$$

Proof. Define, for $\theta \in [0, 1]$,

$$y(\theta) = x + \theta(y - x), \qquad R(\theta) = F(y(\theta)) - \left[F(x) + F'(x)(y(\theta) - x)\right].$$

We shall estimate $\|R(1)\|$. From Hahn-Banach Theorem, there exists $\xi \in X^*$ such that

$$\|\xi\| = 1, \qquad \xi(R(1)) = \|R(1)\|.$$

Define, for $\theta \in [0, 1]$,

$$g(\theta) = \xi(R(\theta)).$$

Direct calculation yields, for $\theta \in [0, 1)$

$$\frac{dg}{d\theta}(\theta) = \xi\Big(F'(y(\theta)) - F'(x)\Big).$$

In particular, g is C^1 on $[0, 1)$. Using assumption **2**, we have

$$\frac{dg}{d\theta}(\theta) \le L\theta\,\|y - x\|.$$

To end the prove, note that $\xi(R(0)) = 0$ and perform direct integration on the above inequality. □

Proposition 5 guarantee non-singularity of F', and so well definedness of Newton iteration map for solving $F(x) = 0$ in $B(x_0, t_*)$. Let us call N_F the Newton iteration map (for $F(x) = 0$) in that region

$$\begin{array}{rcl} N_F : B(x_0, t_*) & \to & X \\ x & \mapsto & x - F'(x)^{-1}F(x). \end{array} \tag{12}$$

One can apply a *single* Newton iteration on any $x \in B(x_0, t_*)$ to obtain $N_F(x)$ which may not belong to $B(x_0, t_*)$, or even may not belong to the domain of F. To ensure that Newton iterations may be repeated indefinitely from x_0, we need some additional results.

First, define some subsets of $B(x_0, t_*)$ in which, as we shall prove, Newton iteration (12) is "well behaved".

$$K(t) := \{x \in B[x_0, t] \, : \, \|F(x)\| \leq f(t)\}, \qquad t \in [0, t_*), \tag{13}$$

$$K := \bigcup_{t \in [0, t_*)} K(t). \tag{14}$$

Lemma 7. *For any $t \in [0, t_*)$ and $x \in K(t)$,*

1. $\left\|F'(x)^{-1}F(x)\right\| \leq -f(t)/f'(t)$,
2. $\|x_0 - N_F(x)\| \leq n_f(t)$,
3. $\|F(N_F(x))\| \leq f(n_f(t))$.

In particular,

$$N_F(K(t)) \subset K(n_f(t)), \qquad \forall\, t \in [0, t_*)$$

and N_F maps K in K, i.e., $N_F(K) \subset K$.

Proof. Take $x \in K(t)$. Using Proposition 5 and (13) we conclude that $F'(x)$ is non-singular,

$$\left\|F'(x)^{-1}\right\| \leq 1/\left|f'(t)\right|, \qquad \|F(x)\| \leq f(t).$$

Hence,

$$\left\|F'(x)^{-1}F(x)\right\| \leq \left\|F'(x)^{-1}\right\| \, \|F(x)\| \leq f(t)/\left|f'(t)\right|,$$

which combined with the inequality $f'(t) < 0$ yields item 1.

To prove item 2 use item 1, triangular inequality and definition (12) to obtain

$$\|x_0 - N_F(x)\| \leq \|x_0 - x\| + \left\|F'(x)^{-1}F(x)\right\| \leq t - f(t)/f'(t).$$

To end the prove of item 2, combine the above equation with definition (7).

From item 2 and Proposition 3, $N_F(x) \in B(x_0, t_*)$. So, Proposition 6 implies

$$\left\|F(N_F(x)) - \left[F(x) + F'(x)\,(N_F(x) - x)\right]\right\| \leq \frac{L}{2}\left\|F'(x)^{-1}F(x)\right\|^2.$$

Note that by (12)

$$F(x) + F'(x)\,(N_F(x) - x) = 0.$$

Combining last two equations, item 1 and identity $f(n_f(t)) = L(f(t)/f'(t))^2/2$ (which follows from (5) and (7)) we conclude that item 3 also holds.

Since $t < n_f(t) < t_*$ (Proposition 3), using also items 2 and 3 we have that

$$N_F(x) \in K(n_f(t)).$$

As x is an arbitrary element of $K(t)$, we have $N_F(K(t)) \subset K(n_f(t))$.

To prove the last inclusion, take $x \in K$. Then $x \in K(t)$ for some $t \in [0, t_*)$, which readily implies $N_F(x) \in K(n_f(t)) \subset K$. □

The last inclusion in Lemma 7 shows that for any $x \in K$, the sequence $\{N_F^k(x)\}$,

$$N_F^0(x) = x, \quad N_F^{k+1}(x) = N_F\left(N_F^k(x)\right), \qquad k = 0, 1, \cdots,$$

is well defined and remains in K. The assumptions of Theorem 1 guarantee

$$x_0 \in K(0) \subset K. \tag{15}$$

Therefore, the sequence $\{x_k = N_F^k(x_0)\}$ is well defined and remains in K. This sequence can be also defined as

$$x_0 = 0, \quad x_{k+1} = N_F(x_k), \qquad k = 0, 1, \cdots. \tag{16}$$

which happens to be the same sequence specified in (2), Theorem 1.

Proposition 8. *The sequence $\{x_k\}$ is well defined, is contained in $B(x_0, t_*)$ and*

$$x_k \in K(t_k), \qquad k = 0, 1, \cdots. \tag{17}$$

Moreover, $\{x_k\}$ converges to a point $x_* \in B[x_0, t_*]$,

$$\|x_* - x_k\| \leq t_* - t_k, \qquad k = 0, 1, \cdots,$$

and $F(x_*) = 0$.

Proof. Well definedness of the sequence $\{x_k\}$ was already proved. We also conclude that this sequence remains in K. As $K \subset B(x_0, t_*)$ (see (13) and (14)), $\{x_k\}$ also remains in $B(x_0, t_*)$.

As $t_0 = 0$, the first inclusion in (15) can also be written as $x_0 \in K(t_0)$. So, (17) holds for $k = 0$. To complete the proof of (17) use induction in k, (16), Proposition 7 and equation (8).

Combining (17) with item 1 of Lemma 7, (16) and (8) we obtain

$$\|x_{k+1} - x_k\| \leq t_{k+1} - t_k, \qquad k = 0, 1, \cdots. \tag{18}$$

As $\{t_k\}$ converges and $\sum_{k=0}^{\infty} t_{k+1} - t_k < \infty$ we conclude that $\{x_k\}$ is a Cauchy sequence. So, $\{x_k\}$ converges to some $x_* \in B[x_0, t_*]$. Moreover, (18) implies

$$\|x_* - x_k\| \leq \sum_{j=k}^{\infty} t_{j+1} - t_j = t_* - t_k, \qquad k = 0, 1, \cdots. \tag{19}$$

Note that

$$F(x_k) = F'(x_{k-1})[x_{k-1} - x_k].$$

As $\|F'(x)\|$ is bounded by $1 + Lt_*$ in $B(x_0, t_*)$ last equation implies that

$$\lim_{k \to \infty} F(x_k) = 0.$$

Now, using the continuity of F in $B[x_0, t_*]$ we have that $F(x_*) = 0$. □

4 Uniqueness and convergence rate

To prove uniqueness and estimate the convergence rate, another auxiliary result will be needed.

Proposition 9. *Take* $x, y \in X$, $t, v \geq 0$. *If*

$$\|x - x_0\| \leq t < t_*, \quad \|y - x_0\| \leq R, \quad F(y) = 0, \quad f(v) \leq 0$$

and $B[x_0, R] \subset C$, *then*

$$\|y - N_F(x)\| \leq [v - n_f(t)] \frac{\|y - x\|^2}{(v - t)^2}.$$

Proof. Note from (12)

$$y - N_F(x) = F'(x)^{-1}[F(x) + F'(x)(y - x)].$$

As $F(y) = 0$, using also Proposition 6 we obtain

$$\left\|F'(x)^{-1}[F(x) + F'(x)(y - x)]\right\| \leq \frac{L}{2}\|y - x\|^2,$$

and from Proposition 5

$$\left\|F'(x)^{-1}\right\| \leq 1/\left|f'(t)\right|.$$

Combining these equations we have

$$\|y - N_F(x)\| \leq \frac{L}{2\left|f'(t)\right|}(v - t)^2 \frac{\|y - x\|^2}{(v - t)^2}.$$

As $f'(t) < 0$ and $f(v) \leq 0$, using also (7) we have

$$\begin{aligned} v - n_f(t) &= \frac{1}{-f'(t)}[-f(t) - f'(t)(v - t)] \\ &\geq \frac{1}{\left|f'(t)\right|}[f(v) - f(t) - f'(t)(v - t)] = \frac{L}{2\left|f'(t)\right|}(v - t)^2. \end{aligned}$$

Combining the two above inequalities we obtain the desired result. □

Corollary 10. *If* $y \in B[x_0, t_*]$ *and* $F(y) = 0$, *then*

$$\|y - x_{k+1}\| \leq \frac{t_* - t_{k+1}}{(t_* - t_k)^2}\|y - x_k\|^2, \qquad \|y - x_k\| \leq t_* - t_k, \qquad k = 0, 1, \cdots.$$

In particular, x_* *is the unique zero of* F *in* $B[x_0, t_*]$.

Proof. Take an arbitrary k. From Proposition 8 we have $x_k \in K(t_k)$. So, $\|x_k - x_0\| \leq t_k$ and we can apply Proposition 9 with $x = x_k$, $t = t_k$ and $v = t_*$, to obtain

$$\|y - N_F(x_k)\| \leq [t_* - n_f(t_k)] \frac{\|y - x_k\|^2}{(t_* - t_k)^2}.$$

The first inequality now follows from the above inequality, (16) and (8).

We will prove the second inequality by induction. For $k = 0$ this inequality holds, because $y \in B[x_0, t_*]$ and $t_0 = 0$. Now, assume that the inequality holds for some k,

$$\|y - x_k\| \leq t_* - t_k.$$

Combining the above inequality with the first inequality of the corollary, we have that $\|y - x_{k+1}\| \leq t_* - t_{k+1}$, wich concludes the induction.

We already know that $x_* \in B[x_0, t_*]$ and $F(x_*) = 0$. Since $\{x_k\}$ converges to x_* and $\{t_k\}$ converges to t_*, using the second inequality of the corollary we conclude $y = x_*$. Therefore, x_* is the unique zero of F in $B[x_0, t_*]$. □

Corollary 11. *The sequences $\{x_k\}$ and $\{t_k\}$ satisfy*

$$\|x_* - x_{k+1}\| \leq \frac{t_* - t_{k+1}}{(t_* - t_k)^2} \|x_* - x_k\|^2, \qquad k = 0, 1, \cdots. \tag{20}$$

In particular,

$$\|x_* - x_{k+1}\| \leqslant \frac{1}{2}\|x_* - x_k\|, \qquad k = 0, 1, \cdots. \tag{21}$$

Additionally, if $2bL < 1$ then

$$\|x_* - x_{k+1}\| \leq \frac{1 - \theta^{2^k}}{1 + \theta^{2^k}} \frac{L}{2\sqrt{1 - 2bL}} \|x_* - x_k\|^2 \leq \frac{L}{2\sqrt{1 - 2bL}} \|x_* - x_k\|^2, \quad k = 0, 1, \cdots. \tag{22}$$

Proof. According to Proposition 8, $x_* \in B[x_0, t_*]$ and $F(x_*) = 0$. To prove equation (20) apply Corollary 10 with $y = x_*$.

Note that, by (9) in Corollary 4, and Proposition 8, for any k

$$(t_* - t_{k+1})/(t_* - t_k) \leq 1/2 \quad \text{and} \quad \|x_* - x_k\|/(t_* - t_k) \leq 1.$$

Combining these inequalities with (20) we have (21). Now, assume that $bL < 1/2$ holds. Then, (10) in Corollary 4 and (20) imply (22) and the corollary is proved. □

Corollary 12. *If $2bL < 1$, $t_* \leq \rho < t_{**}$ and $B[x_0, \rho] \subseteq C$ then x_* is the unique zero of F in $B[x_0, \rho]$.*

Proof. Assume that there exists $y_* \in C$ such that $\|y_* - x_0\| < \rho$ and $F(y_*) = 0$. Using Proposition 6 with $x = x_0$ and $y = y_*$ (recall that $F'(x_0) = I$) we obtain that

$$\|F(x_0) + y - x_0\| \leqslant \frac{L}{2} \|y - x_0\|^2 .$$

Triangle inequality and assumption **3** of Theorem 1 yield

$$\|F(x_0) + y - x_0\| \geq \|y - x_0\| - \|F(x_0)\| \geq \|y - x_0\| - b.$$

Combining the above inequalities we obtain

$$\frac{L}{2}\|y - x_0\|^2 \geq \|y_* - x_0\| - b,$$

which is equivalent to $f(\|y_* - x_0\|) \geq 0$. As $\|y_* - x_0\| \leq \rho < t_{**}$ last inequality implies that $\|y_* - x_0\| \leq t_*$. Therefore, from Corollary 10 and assumption $F(y_*) = 0$, we conclude that $y_* = x_*$. ☐

Therefore, it follows from Proposition 8, Corollary 10, Corollary 11 and Corollary 12 that all statements in Theorem 1 are valid.

4.1 Appendix: A closed formula for t_k

Note that $f(t) = (L/2)(t - t_*)(t - t_{**})$, and $f'(t) = (L/2)[(t - t*) + (t - t_{**})]$. Using the above equations and (8),

$$t_{k+1} - t_* = (t_k - t_*) - \frac{(t_k - t_*)(t_k - t_{**})}{(t_k - t*) + (t_k - t_{**})} = \frac{(t_k - t_*)^2}{(t_k - t*) + (t_k - t_{**})}.$$

By similar manipulations, we have

$$t_{k+1} - t_{**} = \frac{(t_k - t_{**})^2}{(t_k - t*) + (t_k - t_{**})}.$$

Combing two latter equality we obtain that

$$\frac{t_{k+1} - t_*}{t_{k+1} - t_{**}} = \left(\frac{t_k - t_*}{t_k - t_{**}}\right)^2.$$

Suppose that $2bL < 1$. In this case, $t_* < t_{**}$. Hence, using the definition $\theta := t_*/t_{**} < 1$, and induction in k we have

$$\frac{t_k - t_*}{t_k - t_{**}} = \theta^{2^k}.$$

After some algebraic manipulation in above equality we obtain hat

$$t_k = \frac{t_{**}\theta^{2^k} - t_*}{\theta^{2^k} - 1} = t_* - \frac{\theta^{2^k}}{1 - \theta^{2^k}} \frac{2\sqrt{1 - 2bL}}{L}, \qquad k = 0, 1, \cdots.$$

References

[1] Alvarez, F., Botle, J. and Munier, J., *A Unifying Local Convergence Result for Newton's Method in Riemannian Manifolds* INRIA, Rapport de recherche, N. 5381, (2004).

[2] Ferreira, O. P. and Svaiter, B. F., *Kantorovich's Theorem on Newton's method in Riemannian Manifolds* Journal of Complexity, 18, (2002), 304–329.

[3] Ferreira, O. P. and Svaiter, B. F., *Kantorovich's Majorants Principle for Newton's Method*, to appear in Optimization Methods and software (2006).

[4] Gragg, W. B. and Tapia, R. A. *Optimal Error Bound for Newton-Kantorovich Theorem*, SIAM J. Numer. Anal., 11, 1 (1974), 10-13.

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

[6] Kantorovich, L. V., and Akilov, G. P., *Functional analysis in normed spaces*, Oxford, Pergamon (1964).

[7] Ortega, J. M., *The Newton-Kantorovich Theorem*, The American Mathematical Monthly, 75, 6 (1968), 658–660.

[8] Ortega, J. M., and Rheimboldt, W. C., *Interactive solution of nonlinear equations in several variables*, New York: Academic Press (1970).

[9] Ostrowski, A. M., *Solution of equations and systems of equations*, Academic Press, New York (1996).

[10] Polyak, B. T. *Newton-Kantorovich method and its global convergence*, Journal of Mathematical Science, 133, 4 (2006), 1513-1523.

[11] Potra, Florian A., *The Kantorovich Theorem and interior point methods* Mathematical Programming, 102, 1 (2005), 47–70.

[12] Tapia, R. A.,*The Kantorovich Theorem for Newton's Method*, The American Mathematical Monthly, 78, 4 (1971), 389–392.

[13] Wang, X., *Convergence of Newton's method and inverse function theorem in Banach space*, Math. Comp. **68**, 225 (1999), pp.169-186.