JOURNAL OF COMPLEXITY 18, 304–329 (2002)
doi:10.1006/jcom.2001.0582, available online at http://www.idealibrary.com on IDEAL®

Kantorovich's Theorem on Newton's Method in Riemannian Manifolds[1]

O. P. Ferreira

IME/Universidade Federal de Goiás, Campus Samambaia, Caixa Postal 131, CEP 74001-970 Goiânia, GO, Brazil
E-mail: orizon@mat.ufg.br

and

B. F. Svaiter[2]

Instituto de Matemática Pura e Aplicada, Estrada Dona Castorina, 110, Jardim Botânico, CEP 22460-320 Rio de Janeiro, RJ, Brazil
E-mail: benar@impa.br

Received June 21, 2000; published online June 26, 2001

Newton's method for finding a zero of a vectorial function is a powerful theoretical and practical tool. One of the drawbacks of the classical convergence proof is that closeness to a non-singular zero must be supposed *a priori*. Kantorovich's theorem on Newton's method has the advantage of proving existence of a solution and convergence to it under very mild conditions. This theorem holds in Banach spaces. Newton's method has been extended to the problem of finding a singularity of a vectorial field in Riemannian manifold. We extend Kantorovich's theorem on Newton's method to Riemannian manifolds. © 2001 Elsevier Science (USA)

1. INTRODUCTION

Newton's method and its variations are the most efficient methods known for solving systems of non-linear equations when they are continuously differentiable. This includes searching for a local minimizer of a

[1] This research was partially supported by IMPA/CNPq, Brazil.
[2] Partially supported by CNPq Grant 301200/93-9(RN), PRONEX-Optimization, and FAPERJ Grant E-26/150.939/99.

0885-064X/01 $35.00

C^2 function and many other applications. Besides its practical applications, Newton's method is also a powerful theoretical tool. It has been used by Nash [15], Moser [14], Shub and Smale [19], and Smale [20] and is also used in Kolmogorov–Arnold–Moser(KAM) theory (see [26] and its references).

Although quite efficient (it has quadratic convergence under suitable conditions), Newton's method may fail to converge and may even fail to generate an infinite sequence (when a singular derivative point is reached). To ensure convergence of the method, some conditions must be imposed. The classical convergence proof [3, 10, 18] requires the initial iterate to be "close enough" to a solution and the derivative (or Jacobian) of the function to be non-singular in this solution. One drawback of this result is that closeness to a solution (and so existence) must be know or given *a priori.* The advantages of Kantorovich's theorem on Newton's method in Banach spaces [11, 13] (from now on, Kantorovich's Theorem) is that it ensures convergence of Newton's method under very mild assumptions and it is also a proof of existence and local uniqueness of a solution under such assumptions. Furthermore, non-singularity of the derivative at this solution is not imposed. (For other proofs of this theorem, see [5, 16, 25]). Applications of Kantorovich's Theorem, particularly the local existence and uniqueness results, can be found in [1, 9, 12, 13, 17]. Newton's method in Riemannian manifolds has been studied by many authors [7, 8, 21, 22, 24], but an extension of Kantorovich's Theorem to this context was lacking. We present in this paper an extension of Kantorovich's Theorem for Newton's method in Riemannian manifolds.

First let us recall Newton's method for solving

$$F(x) = 0,$$

where $A \subseteq \mathbb{R}^n$ is an open set and $F: A \to \mathbb{R}^n$ is C^1:

ALGORITHM 1.1 (Newton's Method). Take $x_0 \in A$. For $k = 0, 1, \ldots$ define

$$v_k = -DF(x_k)^{-1} F(x_k), \tag{1.1}$$

$$x_{k+1} = x_k + v_k. \tag{1.2}$$

Note that $DF(x^k)^{-1}$ stands for the inverse of the linear mapping $DF(x^k) : \mathbb{R}^n \to \mathbb{R}^n$. So, $DF(x^k)$ must be non-singular.

The extension of this method to the problem of finding a singularity of a vector field X defined on a Riemannian manifold M,

$$X(p) = 0, \qquad p \in M,$$

is straightforward. The derivative of F at x_k is replaced by the covariant derivative of X at p_k,

$$\nabla_{(\cdot)} X_{p_k} : T_{p_k}(M) \to T_{p_k}(M),$$
$$v \mapsto \nabla_v X.$$

We adopt the notation $\mathscr{D}X(p)\, v = \nabla_v X$. Hence, $\mathscr{D}X(p)$ is a linear mapping of $T_p M$ into $T_p M$. So, in this new context, Eq. (1.1) becomes

$$v_k = -(\mathscr{D}X(p_k))^{-1}\, X(p_k),$$

and (when the above equation make sense) $v_k \in T_{p_k}(M)$. In $\mathscr{R}^n$, x_{k+1} is obtained from x_k by taking it along a straight line which passes through x_k with direction v_k (and at a distance $\|v_k\|$). In a Riemannian manifold, geodesics play the role of straight lines, so the natural generalization of (1.2) is

$$p_{k+1} = \exp_{p_k}(v_k).$$

Therefore, Newton's method in a Riemannian manifold becomes

ALGORITHM 1.2 (Newton's Method in Riemannian Manifolds). Take $p_0 \in M$. For $k = 0, 1, ...$ define

$$v_k = -\mathscr{D}X(p_k)^{-1}\, X(p_k),$$
$$p_{k+1} = \exp_{p_k}(v_k),$$

or, equivalently,

$$p_{k+1} = \exp_{p_k}(-\mathscr{D}X(p_k)^{-1}\, X(p_k)). \tag{1.3}$$

Our aim is to find conditions which guarantees the well definedness of the above method, convergence of the generated sequence to a singular point of X, and uniqueness in some region. This, piper is organized as follows. In Section 2 some useful definitions are given and auxiliary results are stated. In Section 3 Kantorovich's Theorem in Riemannian manifolds is enunciated and proved. Some final remarks are made in Section 4.

2. BASIC DEFINITIONS AND AUXILIARY RESULTS

In this section, we introduce some fundamental properties and notations of Riemannian manifolds. References for this section are [6] and (17).

Let M be a connected, complete (geodesic), and finite dimensional Riemannian manifold. For any points p' and p the *Riemannian distance* from p' to p is $d(p', p) := inf_c L(c)$, where $c : [a, b] \to M$ is a piecewise smooth curve in M from p' to p and $L(c) = \int_a^b \|c'(t)\|\, dt$ is the *arc length* of c. We denote by $B(p, r)$ and $B[p, r]$, respectively, the *open metric ball* and the *closed metric ball* at p,

$$B(p, r) = \{q \in M; d(p, q) < r\},$$

$$B[p, r] = \{q \in M; d(p, q) \leqslant r\}.$$

From the Hopf–Rinow Theorem we have that (M, d) is a complete metric space, and for any points p' and p of M there exists a geodesic γ joining p' and p with $L(\gamma) = d(p', p)$. This geodesic is called *minimizing geodesic*.

The Levi–Civita connection of M will be fundamental for our development. This is an affine connection on M which is symmetric and compatible with the metric. Existence and unicity (of such an affine connection) is established by the Levi–Civita Theorem [6, Chap. 2, Theorem 3.6]. The Levi–Civita connection will be denoted by ∇.

Let X be a C^1 vector field on M. The covariant derivative of X determined by the Levi–Civita connection ∇ defines on each $p \in M$ a linear application of $T_p M$ into $T_p M$,

$$T_p M \ni v \mapsto \nabla_v X(p).$$

We will denote this linear map by $\mathscr{D}X(p)$.

$$\begin{aligned} \mathscr{D}X(p) \colon T_p M &\to T_p M, \\ v &\mapsto \mathscr{D}X(p)\, v := \nabla_v X(p), \end{aligned} \tag{2.1}$$

and we will call $\mathscr{D}X(p)$ the covariant derivative of X at p. Observe that if $M = \mathbb{R}^N$ with the canonical inner product, then $\mathscr{D}X(p)$ is the "classical" derivative of X.

In the language of tensorial calculus [23], X is a $(1, 0)$ tensor (one index "contravariant tensor") and ∇X is a $(1, 1)$ tensor ("contravariant" in one index and "covariant" in the other), obtained by taking the "covariant" derivative of X. We identify this tensor with a field of linear operators.

Let $c : \mathbb{R} \to M$ be a C^∞ curve. The *parallel transport* along c will be denoted by P_c. So, for $a, b \in \mathbb{R}$,

$$\begin{aligned} P_{c,a,b} : T_{c(a)} M &\to T_{c(b)} M \\ v &\mapsto P_{c,a,b}(v) = V(b), \end{aligned} \tag{2.2}$$

where V is a unique vector field on c such that $\nabla_{c'(t)} V = 0$ and $V(a) = v$. Observe that $P_{c,a,b}$ is linear. Since ∇ is compatible with the metric, $P_{c,a,b}$ is an isometry. Note also that

$$P_{c,b_1,b_2} \circ P_{c,a,b_1} = P_{c,a,b_2},$$
$$P_{c,b,a} = P_{c,a,b}^{-1}.$$

To simplify the notation, we will also write $P_{c,a,b} v$ for $P_{c,a,b}(v)$.

If X is a C^1 vector field, then

$$\begin{aligned}\mathscr{D}X(c(s))\, c'(s) &= \nabla_{c'(s)} X \\ &= \lim_{h \to 0} \frac{1}{h} (P_{c,s+h,s} X(c(s+h)) - X(c(s))).\end{aligned} \tag{2.3}$$

The first equality is (2.1) with $p = c(s)$, $v = c'(s)$. The second equality is proposed as an exercise in [6, Chap. 2, Exercise 2]. To prove it, consider $W_1(t), ..., W_n(t)$ to be an orthonormal basis parallel transported along the curved $c(t)$. Using this base to express $X(c(t))$, and using also [6, Chap. 1, Prop. 2.2] and the properties of Levi–Civita connection, the conclusion follows.

Now, the elementary "fundamental theorem of calculus" becomes

$$\begin{aligned}P_{c,t,0}(X(c(t)) &= X(c(0)) + \int_0^t P_{c,s,0}(\nabla_{c'(s)} X_{c(s)})\, ds \\ &= X(c(0)) + \int_0^t P_{c,s,0}(\mathscr{D}X(c(s))\, c'(s))\, ds.\end{aligned} \tag{2.4}$$

Indeed, consider the curve $\eta(s) = P_{c,s,0} X(c(s))$ in $T_{c(0)}M$. Direct calculation gives

$$\begin{aligned}\eta'(s) &= \lim_{h \to 0} \frac{\eta(s+h) - \eta(s)}{h} \\ &= \lim_{h \to 0} \frac{1}{h} (P_{c,s+h,0} X(c(s+h)) - P_{c,s,0} X(c(s))) \\ &= \lim_{h \to 0} \frac{1}{h} (P_{c,s,0} \circ P_{c,s+h,s} X(c(s+h)) - P_{c,s,0} X(c(s))).\end{aligned}$$

Since $P_{c,s,0}$ is linear, we have

$$\eta'(s) = \lim_{h\to 0} \frac{1}{h} P_{c,s,0}(P_{c,s+h,s} X(c(s+h)) - X(c(s)))$$
$$= \lim_{h\to 0} P_{c,s,0}\left(\frac{1}{h}(P_{c,s+h,s} X(c(s+h)) - X(c(s)))\right)$$
$$= P_{c,s,0}\left(\lim_{h\to 0} \frac{1}{h}(P_{c,s+h,s}(X(c(s+h))) - X(c(s)))\right).$$

Using now (2.3) we obtain

$$\eta'(s) = P_{c,s,0}(\nabla_{c'(s)} X(c(s))) = P_{c,s,0}(\mathscr{D}X(c(s))\, c'(s)),$$

which implies (2.4).

Instead of working with Frobenius norm of rank-two tensors, we will use the "operator norm" for linear transformations in each tangent space.

Definition 2.1. Let $A : T_pM \to T_pM$ be a linear operator. Define

$$\|A\|_{\mathrm{op}} = \sup\{\, \|Av\| \mid v \in T_pM,\ \|v\| = 1\}.$$

Definition 2.2. Let Ω be an open subset of M and let X be a C^1 vector field defined on Ω. The covariant derivative $\mathscr{D}X$ is Lipschitz with constant $L > 0$ if for any geodesic γ and $a, b \in \mathbb{R}$ satisfying $\gamma([a, b]) \subseteq \Omega$ it holds that

$$\|P_{\gamma,b,a} \mathscr{D}X(\gamma(b))\, P_{\gamma,a,b} - \mathscr{D}X(\gamma(a))\|_{\mathrm{op}} \leqslant L \int_a^b \|\gamma'(t)\|\, dt.$$

We use the notation $\mathscr{D}X \in \mathrm{Lip}_L(\Omega)$.

Note that if M is the Euclidean space (with the canonical inner product), the above definition coincides with the usual definition of Lipschitz continuity of the derivative of a vectorial function. A similar concept was presented for the first time in [4] for defining a Lipschitz continuous vector field in a Riemannian manifold.

Let $p \in M$ and $v \in T_pM$. There exist a unique geodesic γ such that $\gamma(0) = v$ and $\gamma'(0) = v$. The point $\gamma(1)$ is the image of v by the exponential map (at p), i.e., $\exp_p(v) = \gamma(1)$. Hence, for any $t \in \mathbb{R}$, $\gamma(t) = \exp_p(tv)$. The exponential map has many important properties (see [6, Chap. 3].) We will use the exponential map mainly as short notation for a geodesic with a given "starting point" p and "initial velocity" $v \in T_pM$. A basic property of geodesics is that $\gamma'(t)$ is parallel along $\gamma(t)$. Since we are using the Levi–Civita connection, this implies that $\|\gamma'(t)\|$ is constant.

LEMMA 2.3. *Let Ω be an open subset of M. Let X be a continuous vector field defined on $\bar{\Omega}$ and C^1 in Ω with $\mathscr{D}X \in \mathrm{Lip}_L(\Omega)$. Take $p \in \Omega$, $v \in T_pM$, and define*

$$\gamma(t) = \exp_p(tv).$$

If $\gamma([0, t)) \subseteq \Omega$ then

$$P_{\gamma,t,0}X(\gamma(t)) = X(p) + t\,\mathscr{D}X(p)\,v + R(t)$$

with

$$\|R(t)\| \leqslant \frac{L}{2}\,t^2\,\|v\|^2.$$

Proof. From (2.4) it follows that

$$P_{\gamma,t,0}X(\gamma(t)) - X(p) = \int_0^t P_{\gamma,\tau,0}(\mathscr{D}X(\gamma(\tau))\,\gamma'(\tau))\,d\tau.$$

As γ is a geodesic, $\gamma'(t)$ is parallel and $\gamma'(\tau) = P_{\gamma,0,\tau}\gamma'(0)$. Using also that $\gamma'(0) = v$ we get

$$\begin{aligned} P_{\gamma,t,0}X(\gamma(t)) - X(p) &= \int_0^t P_{\gamma,\tau,0}(\mathscr{D}X(\gamma(\tau))\,P_{\gamma,0,\tau}v)\,d\tau \\ &= \int_0^t P_{\gamma,\tau,0}\,\mathscr{D}X(\gamma(\tau))\,P_{\gamma,0,\tau}v\,d\tau. \end{aligned}$$

Therefore,

$$\begin{aligned} &P_{\gamma,t,0}X(\gamma(t)) - X(p) - t\,\mathscr{D}X(p)\,v \\ &\quad= \int_0^t P_{\gamma,\tau,0}\,\mathscr{D}X(\gamma(\tau))\,P_{\gamma,0,\tau}v\,d\tau - t\,\mathscr{D}X(p)\,v \\ &\quad= \int_0^t (P_{\gamma,\tau,0}\,\mathscr{D}X(\gamma(\tau))\,P_{\gamma,0,\tau} - \mathscr{D}X(p))\,v\,d\tau. \end{aligned}$$

Set $R(t)=\int_0^t (P_{\gamma,\tau,0}\,\mathscr{D}X(\gamma(\tau))\,P_{\gamma,0,\tau}-\mathscr{D}X(p))\,v\,d\tau$. By hypothesis $\mathscr{D}X\in \mathrm{Lip}_L(\Omega)$, therefore

$$
\begin{aligned}
\|R(t)\| &\leqslant \int_0^t \|P(\gamma)_\tau^0\,\nabla X_{\gamma(\tau)}\,P(\gamma)_0^\tau-\mathscr{D}X(p)\|_{\mathrm{op}}\,\|v\|\,d\tau\\
&\leqslant \int_0^t L\left(\int_0^\tau \|\gamma'(s)\|\,ds\right)\|v\|\,d\tau.
\end{aligned}
$$

As γ is a geodesic, $\|\gamma'(s)\|$ is constant; in particular, $\|\gamma'(s)\|=\|\gamma'(0)\|=\|v\|$, so $\int_0^\tau \|\gamma'(s)\|\,ds=\tau\,\|v\|$ and

$$
\begin{aligned}
\|R(t)\| &\leqslant \int_0^t L\tau\,\|v\|^2\,d\tau\\
&=\frac{L}{2}\,t^2\,\|v\|^2. \quad ∎
\end{aligned}
$$

COROLLARY 2.4. *Let Ω be an open subset of M. Let X be a continuous vector field defined on $\bar{\Omega}$ and C^1 in Ω with $\mathscr{D}X\in \mathrm{Lip}_L(\Omega)$. Take $p\in\Omega$, $v\in T_pM$, and define*

$$\gamma(t)=\exp_p(tv).$$

If $\gamma([0,1))\subseteq\Omega$ and $\mathscr{D}X(p)\,v=-X(p)$ then

$$\|X(\gamma(1))\|\leqslant\frac{L}{2}\,\|v\|^2.$$

Proof. Since the parallel transport is an isometry, from Lemma 2.3 we obtain the result. ∎

A most useful result, applying $\|\cdot\|_{\mathrm{op}}$, is Banach's Lemma [13, Section 5, Theorems 3 and 4; 17, Lemma 2.3.2]. We quote it here, in a particular form suitable for our use:

LEMMA 2.5 (Banach's Lemma). *Let A, B be linear operators in T_pM. If A is nonsingular and*

$$\|A^{-1}\|_{\mathrm{op}}\,\|B-A\|_{\mathrm{op}}<1,$$

then B is nonsingular and

$$\|B^{-1}\|_{\rm op} \leqslant \frac{\|A^{-1}\|_{\rm op}}{1-\|A^{-1}(B-A)\|_{\rm op}} \leqslant \frac{\|A^{-1}\|_{\rm op}}{1-\|A^{-1}\|_{\rm op}\,\|B-A\|_{\rm op}}.$$

Furthermore,

$$\|B^{-1}A\|_{\rm op} \leqslant \frac{1}{1-\|A^{-1}(B-A)\|_{\rm op}} \leqslant \frac{1}{1-\|A^{-1}\|_{\rm op}\,\|B-A\|_{\rm op}}.$$

Very often, this lemma is stated with $A = I$ (identity), and the general case is a corollary. The last inequality is not included in the "usual form of" Banach's Lemma, but follows trivially for $\widetilde{B} = A^{-1}B$.

3. KANTOROVICH'S THEOREM IN RIEMANNIAN MANIFOLDS

First let us recall Kantorovich's Theorem on Newton's Method (see [13, 16, 17]).

THEOREM 3.1 (Kantorovich). *Let Z, W be Banach spaces, $C \subseteq Z$ an open convex set, and $F: \bar{C} \to W$ a continuous function, continuously differentiable on C, with*

$$\|DF(p) - DF(q)\| \leqslant L\,\|p - q\|,$$

for any $p, q \in C$. Suppose that for some $x_0 \in C$, $DF(x_0)$ is invertible and that

$$\|DF(x_0)^{-1}\| \leqslant a, \qquad \|DF(x_0)^{-1}\,F(x_0)\| \leqslant b, \qquad \ell = abL \leqslant 1/2.$$

Suppose also that

$$B(x_0, t_*) \subseteq C,$$

where $t_* = (1/aL)(1 - \sqrt{1 - 2\ell})$. Then the sequence $\{x_k\}$, generated by (1.1), (1.2) *with starting point x_0, is well defined and contained in $B(x_0, t_*)$ and converges to a point x_* which is the unique zero of F in $B[x_0, t_*]$. Furthermore, if $\ell < 1/2$ and $B(x_0, r) \subseteq C$ with*

$$t_* < r \leqslant t_{**} = (1/aL)(1 + \sqrt{1 - 2\ell}),$$

then x_* is also the unique zero of F in $B(x_0, r)$. Regarding the error bound,

$$d(x_k, x_*) \leqslant (2\ell)^{2^k} \frac{b}{\ell}, \qquad k = 1, 2, ...$$

Now we can translate Kantorovich's Theorem to a new context: Newton's method in Riemannian manifolds. The extension of Newton's method to a Riemannian manifold has already been discussed informally in Section 1. The concept of a *singularity* of a vector field X corresponds to a *zero* of X in the classical setting; i.e., a point p where $X(p) = 0$. Our aim is to prove the following result.

Theorem 3.2 (Kantorovich's Theorem in Riemannian Manifolds). *Let M be a complete Riemannian manifold, Ω be an open subset of M, and X be a continuous vector field defined on $\bar{\Omega}$ which is C^1 in Ω with $\mathscr{D}X \in \mathrm{Lip}_L(\Omega)$. Take $p_0 \in \Omega$. Suppose that $\mathscr{D}X(p_0)$ is nonsingular and that, for some $a > 0$ and $b \geqslant 0$,*

$$\|\mathscr{D}X(p_0)^{-1}\|_{\mathrm{op}} \leqslant a, \qquad \|\mathscr{D}X(p_0)^{-1} X(p_0)\| \leqslant b, \qquad \ell = abL \leqslant 1/2,$$

and

$$B(p_0, t_*) \subseteq \Omega,$$

where $t_* = \frac{1}{aL}(1 - \sqrt{1 - 2\ell})$. Then the sequence $\{p_k\}$ generated by (1.3) with starting point p_0 is well defined and contained in $B(p_0, t_*)$ and converges to a point p_* which is the unique singularity of X in $B[p_0, t_*]$. Furthermore, if $\ell < 1/2$ and $B(p_0, r) \subseteq \Omega$ with

$$t_* < r \leqslant t_{**} = (1/aL)(1 + \sqrt{1 - 2\ell}),$$

then p_* is also the unique singularity of X in $B(p_0, r)$. Regarding the error bound,

$$d(p_k, p_*) \leqslant (2\ell)^{2^k} \frac{b}{\ell}, \qquad k = 1, 2, ... \tag{3.1}$$

The proof of this theorem will be broken into two parts. First we will prove the well definedness of Newton's method under the above conditions and convergence to a singularity of the vector field. In the second part, uniqueness in the suitable region will be established. We begin by proving some auxiliary results. From now on, we assume that the hypotheses of Theorem 3.2 hold. In our proof, a most useful auxiliary function is $f\colon \mathbb{R} \to \mathbb{R}$, defined by

$$f(t) = t^2 L/2 - t/a + b/a. \tag{3.2}$$

This is a second degree polynomial with discriminant $(1 - 2Lba)/a^2$. Therefore, due to our assumptions ($\ell = aLb \leqslant 1/2$) the equation $f = 0$ has at least one real root. The smallest root (or the unique one, when $\ell = 1/2$) is t_*. Direct calculation shows that $f'(t) < 0$ for $0 \leqslant t < t_*$, $|f'|$ is strictly decreasing in this range, and if $f'(\tau)(t - \tau) = -f(\tau)$ then $f(t) = (L/2)(t - \tau)^2$. But what about Newton's method applied to f?

PROPOSITION 3.3. *Take $\tau_0 \in [0, t_*)$. For $k = 0, 1, \ldots$ define*

$$\tau_{k+1} = \tau_k - f(\tau_k)/f'(\tau_k).$$

The sequence $\{\tau_k\}$ is well defined for all k and is strictly increasing and converges to t_*. Furthermore, if $\ell < 1/2$ then

$$t_* ^{-\tau_k} \leqslant (2\ell)^{2^k} (b/\ell),$$

for all k.

From now on, we will refer to the sequences in the above proposition as the sequence generated by Newton method (for solving $f = 0$) with starting point τ_0.

Of particular importance will be the sequence generated by the Newton method (for solving $f = 0$) with starting point 0. Let us call it $\{t_k\}$,

$$t_0 = 0,$$
$$t_{k+1} = t_k - f(t_k)/f'(t_k), \qquad k = 0, 1, \ldots .$$

3.1. *Convergence*

In this subsection we will prove the well definedness of Newton's method and the convergence of the generated sequence for some starting points. Instead of obtaining this result only for p_0, for proving uniqueness it is more convenient to enlarge a little bit more the possible choices of starting points in Ω for Newton's method.

Note that $B(p_0, t_*) \subseteq \Omega$. The first thing we want to show is that the Newton iteration is well defined for any $q \in B(p_0, t_*)$.

PROPOSITION 3.4. *If* $q \in B(p_0, t_*)$ *then* $\mathscr{D}X(q)$ *is nonsingular and*

$$\|\mathscr{D}X(q)^{-1}\|_{\mathrm{op}} \leqslant 1/|f'(\lambda)|,$$

where $\lambda = d(p_0, q) < t_*$.

Proof. Let $\lambda = d(p_0, q)$ and let $\alpha : [0, 1] \to M$ be a geodesic with $\alpha(0) = p_0$, $\alpha(1) = q$, and $\|\alpha'(0)\| = \lambda$. Define $\Phi : T_q M \to T_q M$,

$$\Phi = P_{\alpha, 0, 1} \mathscr{D}X(p_0)\, P_{\alpha, 1, 0}.$$

Since $P_{\alpha, 1, 0}$ is a linear isometry between $T_q M$ and $T_{p_0} M$, Φ is linear. Furthermore, as by hypothesis $\mathscr{D}X(p_0)$ is nonsingular, we conclude that Φ is also nonsingular and

$$\|\Phi^{-1}\|_{\mathrm{op}} = \|\mathscr{D}X(p_0)^{-1}\|_{\mathrm{op}} \leqslant 1/|f'(0)|. \tag{3.3}$$

Since $\alpha([0, 1]) \subset B(p_0, t_*) \subseteq \Omega$ and $\mathscr{D}X \in Lip_L(\Omega)$,

$$\|\mathscr{D}X(q) - \Phi\|_{\mathrm{op}} \leqslant L\lambda. \tag{3.4}$$

From (3.3) and (3.4), we have

$$\|\Phi^{-1}\|_{\mathrm{op}}\, \|\mathscr{D}X(q) - \Phi\|_{\mathrm{op}} \leqslant L\lambda/|f'(0)| < Lt_* a \leqslant 1.$$

Now we can apply Banach's Lemma to conclude that $\mathscr{D}X(q)$ is nonsingular and

$$\begin{aligned}\|\mathscr{D}X(q)^{-1}\|_{\mathrm{op}} &\leqslant \frac{\|\Phi^{-1}\|_{\mathrm{op}}}{1 - \|\Phi^{-1}\|_{\mathrm{op}}\, \|\mathscr{D}X(q) - \Phi\|_{\mathrm{op}}} \\ &\leqslant \frac{1/|f'(0)|}{1 - L\lambda/|f'(0)|} \\ &= 1/|f'(\lambda)|,\end{aligned}$$

where the last equality follows from (3.2) also. ∎

Therefore, for any $q \in B(p_0, t_*)$ one can apply Newton iteration to obtain $q_+ = \exp_q(-\mathscr{D}X(q)^{-1} X(q))$. This is enough to guarantee the well definedness only of the next iterated. To ensure that Newton iterations may be repeated indefinitely from some starting point, additional conditions must be imposed. Now we define a family of subsets of Ω which are very well behaved with respect to Newton iteration. Define for $t \in [0, t_*)$

$$K(t) = \{q \in B[p_0, t];\ \|\mathscr{D}X(q)^{-1} X(q)\| \leqslant f(t)/|f'(t)|\}. \tag{3.5}$$

Note that since in the above definition we assume that $0 \leqslant t < t_*$, if $q \in B[p_0, t]$ then from Proposition 3.4 it follows that $\mathscr{D}X(q)$ is nonsingular. So the above definition is consistent.

Lemma 3.5. *Take $0 \leqslant t < t_*$ and $q \in K(t)$. Define*

$$t_+ = t - f(t)/f'(t),$$
$$q_+ = \exp_q(-\mathscr{D}X(q)^{-1} X(q)).$$

Then $t < t_+ < t_*$ and $q_+ \in K(t_+)$.

Proof. Take the geodesic γ defined by

$$\gamma(\theta) = \exp_q(-\theta \mathscr{D}X(q)^{-1} X(q)).$$

Using triangular inequality and the definition of $K(t)$, we conclude that, for any $\theta \in [0, 1]$,

$$\begin{aligned} d(p_0, \gamma(\theta)) &\leqslant d(p_0, q) + d(q, \gamma(\theta)) \\ &\leqslant t + \theta\ \|\mathscr{D}X(q)^{-1} X(q)\| \\ &\leqslant t + f(t)/|f'(t)| \\ &= t - f(t)/f'(t) = t_+. \end{aligned}$$

In particular, $q_+ = \gamma(1) \in B[p_0, t_+] \subseteq B(p_0, t_*)$. Furthermore, by Proposition 3.4, $\mathscr{D}X(q_+)$ is nonsingular. The above inequality tells us that $\gamma([0, 1]) \subseteq \Omega$. So, using the hypothesis on the Lipschitzian quality of $\mathscr{D}X$ in Ω, the definition of $K(t)$, and Corollary 2.4, we get

$$\begin{aligned}\|\mathscr{D}X(q_+)^{-1}\,X(q_+)\| &\leqslant \|\mathscr{D}X(q_+)^{-1}\|_{\mathrm{op}}\,\|X(q_+)\|\\ &\leqslant \|\mathscr{D}X(q_+)^{-1}\|_{\mathrm{op}}\,\frac{L}{2}\,\|\mathscr{D}X(q)^{-1}\,X(q)\|^2\\ &\leqslant (1/|f'(t_+)|)\,\frac{L}{2}\,(f(t)/|f'(t)|)^2\\ &= (1/|f'(t_+)|)\,\frac{L}{2}\,(t_+-t)^2\\ &= f(t_+)/|f'(t_+)|,\end{aligned}$$

which is the statement of the lemma. ▌

Now we are ready to prove that any point in $K(t)$ can be used as a starting point of Newton's method to generate a sequence converging to a singularity of X.

COROLLARY 3.6. *Take $0 \leqslant t < t_*$ and $q \in K(t)$. Define*

$$\tau_0 = t,$$
$$\tau_{k+1} = \tau_k - f(\tau_k)/f'(\tau_k), \qquad k = 0, 1, \ldots .$$

The sequence $\{q_k\}$ generated by Newton's method with starting point $q_0 = q$ is well defined and satisfies, for all k,

$$q_k \in K(\tau_k). \tag{3.6}$$

Furthermore, $\{q_k\}$ converges to some $q_* \in B[p_0, t_*]$ singular point of X and

$$d(q_k, q_*) \leqslant t_* - \tau_k,$$

for all k. If additionally $\ell < 1/2$ then

$$d(q_k, q_*) \leqslant (2\ell)^{2^k}\, b/\ell \tag{3.7}$$

for all k.

Proof. First note that $\{\tau_k\}$ is the sequence generated by Newton's method for solving $f = 0$. Since $0 \leqslant \tau_0 < t_*$, from Proposition 3.3 the sequence $\{\tau_k\}$ is well defined and strictly increasing and converges to t_*. To prove the well definedness of the sequence $\{q_k\}$ and (3.6) we proceed by induction. For $k = 0$, q_0 is well defined (it's the starting point q) and satisfies (3.6) by hypothesis. Suppose now that, for some k, the points

$q_0, ..., q_k$ are well defined and that (3.6) holds for such a k. Then, using Lemma 3.5 we conclude that q_{k+1} is well defined and (3.6) still holds for $k+1$.

Now, to prove convergence observe that $d(q_{k+1}, q_k) \leqslant \|\mathscr{D}X(q_k)^{-1} X(q_k)\|_{\mathrm{op}}$. Therefore, using (3.6),

$$d(q_{k+1}, q_k) \leqslant f(\tau_k)/|f'(\tau_k)| = \tau_{k+1} - \tau_k. \tag{3.8}$$

Hence, for $k \geqslant l$ ($k, l \in \mathbb{N}$),

$$d(q_k, q_l) \leqslant \tau_k - \tau_l. \tag{3.9}$$

It follows that $\{q_k\}$ is a Cauchy sequence, hence it converges to some $q_* \in M$. But $q_k \in K(\tau_k) \subseteq B[p_0, t_*]$, therefore $q_* \in B[p_0, t_*]$.

It remains to prove that q_* is a singular point of X. Observe that $d(p_0, q_k) \leqslant \tau_k$. Therefore

$$\|\mathscr{D}X(q_k)\|_{\mathrm{op}} \leqslant \|\mathscr{D}X(p_0)\|_{\mathrm{op}} + L\tau_k. \tag{3.10}$$

Using (3.10), (3.6), and (3.8) we get,

$$\begin{aligned} \|X(q_k)\| &\leqslant \|\mathscr{D}X(q_k)\|_{\mathrm{op}} \, \|\mathscr{D}X(q_k)^{-1} X(q_k)\| \\ &\leqslant (\|\mathscr{D}X(p_0)\|_{\mathrm{op}} + L\tau_k) \, f(\tau_k)/|f'(\tau_k)| \\ &\leqslant (\|\mathscr{D}X(p_0)\|_{\mathrm{op}} + Lt_*)(\tau_{k+1} - \tau_k). \end{aligned} \tag{3.11}$$

Taking the limit in (3.11), we conclude that $X(q_*) = 0$. Taking the limit in (3.9), we get $d(q_k, q_*) \leqslant t_* - \tau_k$. Finally, the inequality (3.7) results from the last inequality and Proposition 3.3. ∎

Observe, by hypothesis, that $p_0 \in K(0)$. Therefore, by using Corollary 3.6 it follows that the sequence $\{p_k\}$ generated by (1.3) with starting point p_0 is well defined and contained in $B(p_0, t_*)$ and converges to some p_*, which is a singular point of X in $B(p_0, t_*]$; and if $\ell < 1/2$ then

$$d(p_k, p_*) \leqslant (2l)^{2^k} b/\ell,$$

for all k. This concludes the first part.

3.2. *Uniqueness*

We have already proved that the sequence $\{p_k\}$ converges to some p_* singular point of X in the region $B[p_0, t_*]$ and that

$$d(p_k, p_*) + t_k \leqslant t_*, \qquad \text{for all } k.$$

To prove uniqueness, we will need a "stronger" version of Lemma 3.5.

LEMMA 3.7. *Take $0 \leqslant t < t_*$ and $q \in K(t)$. Define for $\theta \in \mathbb{R}$,*

$$\tau(\theta) = t - \theta f(t)/f'(t),$$
$$\gamma(\theta) = \exp_q(-\theta \mathcal{D}X(q)^{-1} X(q)).$$

Then for $\theta \in [0, 1]$, $t \leqslant \tau(\theta) < t_*$ and $\gamma(\theta) \in K(\tau(\theta))$.

Proof. As γ is a geodesic, for $\theta \geqslant 0$,

$$\begin{aligned} d(p_0, \gamma(\theta)) &\leqslant d(p_0, q) + \theta \, \|\mathcal{D}X(q)^{-1} X(q)\| \\ &\leqslant t + \theta f(t)/|f'(t)| = \tau(\theta), \end{aligned} \tag{3.12}$$

where the triangular inequality, the definition of γ, and the hypothesis $q \in K(t)$ were used. Note that $\tau(\theta)$ is increasing in θ. Take now some fixed $\theta \in [0, 1]$. Trivially

$$t \leqslant \tau(\theta) \leqslant \tau(1) < t_*. \tag{3.13}$$

Therefore

$$\gamma([0, \theta]) \subseteq B[p_0, \tau(\theta)] \subseteq B(p_0, t_*). \tag{3.14}$$

Using (3.12), (3.13), and Proposition 3.4 we conclude that

$$\|\mathcal{D}X(\gamma(\theta))^{-1}\|_{\text{op}} \leqslant 1/|f'(\tau(\theta))|. \tag{3.15}$$

From Lemma 2.3 and (3.14) it follows that

$$X(\gamma(\theta)) = P_{\gamma, 0, \theta}(X(q) + \mathcal{D}X(q)(-\theta \mathcal{D}X(q)^{-1} X(q)) + R(\theta))$$

with

$$\|R(\theta)\| \leqslant \frac{L}{2} \, \|\theta \mathcal{D}X(q)^{-1} X(q)\|^2. \tag{3.16}$$

Therefore, after some simple algebraic manipulations we get

$$\|\mathscr{D}X(\gamma(\theta))^{-1}\, X(\gamma(\theta))\| \leqslant (1-\theta)\, \|\mathscr{D}X(\gamma(\theta))^{-1}\, P_{\gamma,0,\theta}X(q)\| + \cdots$$
$$+\, \|\mathscr{D}X(\gamma(\theta))^{-1}\, P_{\gamma,0,\theta}P(\theta)\|. \qquad (3.17)$$

Using the isometric property of parallel transport, (3.15), (3.16), and the hypothesis $q \in K(t)$, we conclude that

$$\begin{aligned}\|\mathscr{D}X(\gamma(\theta))^{-1}\, P_{\gamma,0,\theta}R(\theta)\| &\leqslant \|\mathscr{D}X(\gamma(\theta))^{-1}\|_{\mathrm{op}}\, \|P_{\gamma,0,\theta}R(\theta)\| \\ &\leqslant (1/|f'(\tau(\theta))|)\, \|R(\theta)\| \\ &\leqslant (1/|f'(\tau(\theta))|)\, \frac{L}{2}\,(\theta f(t)/|f'(t)|)^2 \\ &= (1/|f'(\tau(\theta))|)\, \frac{L}{2}\,(\tau(\theta)-t)^2. \qquad (3.18)\end{aligned}$$

The first term on the right side of the inequality (3.17) must also be estimated. Write

$$\begin{aligned}&\mathscr{D}X(\gamma(\theta))^{-1}\, P_{\gamma,0,\theta}X(q) \\ &\quad = (\mathscr{D}X(\gamma(\theta))^{-1}\, P_{\gamma,0,\theta}\mathscr{D}X(q)\, P_{\gamma,\theta,0})\, P_{\gamma,0,\theta}\mathscr{D}X(q)^{-1}\, X(q).\end{aligned}$$

Recall that $q \in K(t)$. So, using also the isometric property of the parallel transport,

$$\begin{aligned}&\|\mathscr{D}X(\gamma(\theta))^{-1}\, P_{\gamma,0,\theta}X(q)\| \\ &\quad \leqslant \|\mathscr{D}X(\gamma(\theta))^{-1}\, P_{\gamma,0,\theta}\mathscr{D}X(q)\, P_{\gamma,\theta,0}\|_{\mathrm{op}}\;\, f(t)/|f'(t)|. \qquad (3.19)\end{aligned}$$

To simplify the manipulations, set

$$\Phi = P_{\gamma,0,\theta}\mathscr{D}X(q)\, P_{\gamma,\theta,0}\,.$$

Since $d(p_0, q) \leqslant t < t_*$, by using Proposition 3.4 and the isometric properties of parallel transport we get

$$\|\Phi^{-1}\|_{\mathrm{op}} = \|\mathscr{D}X(q)^{-1}\|_{\mathrm{op}} \leqslant 1/|f'(t)|.$$

using the Lipschitzian quality of $\mathscr{D}X$ in Ω, the hypothesis $q \in K(t)$, and (3.14), we conclude that

$$\begin{aligned}\|\mathscr{D}X(\gamma(\theta)) - \Phi\|_{\mathrm{op}} &\leqslant \theta L \, \|\mathscr{D}X(q)^{-1} \, X(q)\| \\ &\leqslant \theta L f(t)/|f'(t)|.\end{aligned}$$

Therefore

$$\|\Phi^{-1}\|_{\mathrm{op}} \, \|\mathscr{D}X(\gamma(\theta)) - \Phi\|_{\mathrm{op}} \leqslant \theta L f(t)/|f'(t)|^2 < 1/2.$$

Applying Banach's Lemma and taking into account that $f'(t) < 0$ and $f'(\tau(\theta)) < 0$, we get

$$\begin{aligned}\|\mathscr{D}X(\gamma(\theta))^{-1} \, P_{\gamma, 0, \theta} \mathscr{D}X_q P_{\gamma, \theta, 0}\|_{\mathrm{op}} &\leqslant \frac{1}{1 - \theta L f(t)/|f'(t)|^2} \\ &= \frac{|f'(t)|}{|f'(t)| - L\theta f(t)/|f'(t)|} \\ &= |f'(t)|/|f'(\tau(\theta))|.\end{aligned} \tag{3.20}$$

By substituting (3.20) into (3.19) we obtain

$$\|\mathscr{D}X(\gamma(\theta))^{-1} \, P_{\gamma, 0, \theta} X(q)\| \leqslant f(t)/|f'(\tau(\theta))|, \tag{3.21}$$

and now by substituting (3.18) and (3.21) into (3.17) we obtain

$$\begin{aligned}\|\mathscr{D}X(\gamma(\theta))^{-1} \, X(\gamma(\theta))\| &\leqslant (1 - \theta) \, f(t)/|f'(\tau(\theta))| + \cdots \\ &\quad + (1/|f'(\tau(\theta))|) \, L(\tau(\theta) - t)^2/2 \\ &= (f(t) - \theta f(t) + (L/2)(\tau(\theta) - t)^2)/|f'(\tau(\theta))|.\end{aligned}$$

Trivial algebraic manipulation gives

$$\|\mathscr{D}X(\gamma(\theta))^{-1} \, X(\gamma(\theta))\| \leqslant f(\tau(\theta))/|f'(\tau(\theta))|,$$

which, together with (3.13), (3.12), is the statement of the lemma. ∎

LEMMA 3.8. *Take $0 \leqslant t < t_*$ and $q \in K(t)$. Suppose that $q_* \in B[p_0, t_*]$ is a singular point of X and*

$$t + d(q, q_*) = t_*.$$

Then

$$d(p_0, q) = t.$$

Furthermore, define

$$t_+ = t - f(t)/f'(t),$$
$$q_+ = \exp_q(-\mathscr{D}X(q)^{-1}\, X(q)).$$

Then $t < t_+ < t_*$, $q_+ \in K(t_+)$ *and*

$$t_+ + d(q_+, q_*) = t_*.$$

Proof. From Lemma 3.5, $t < t_+ < t_*$ and $q_+ \in K(t_+)$. Consider the minimizing geodesic $\alpha : [0, 1] \to M$ joining q to q_*. Using the triangular inequality and that $q \in K(t)$, we conclude that for any $\theta \in [0, 1)$,

$$\begin{aligned} d(p_0, \alpha(\theta)) &\leqslant d(p_0, q) + d(q, \alpha(\theta)) \\ &\leqslant d(p_0, q) + \theta d(q_*, q) \\ &< t + d(q_*, q) \\ &= t_*. \end{aligned}$$

It follows that $\alpha([0, 1)) \subset B(p_0, t_*) \subseteq \Omega$. Setting $v = \alpha'(0)$ and applying Lemma 2.3, we conclude that

$$P_{\alpha, 1, 0} X(q_*) = X(q) + \mathscr{D}X(q)\, v + R(1),$$

with $\|R(1)\| \leqslant L\, \|v\|^2/2$. Since $X(q_*) = 0$ and $\|v\| = d(q, q_*)$, from the last equality we get

$$Ld(q_*, q)^2/2 \geqslant \|X(q) + \mathscr{D}X(q)\, v\|. \tag{3.22}$$

Set

$$\lambda = d(p_0, q). \tag{3.23}$$

As $q \in K(t)$, $0 \leqslant \lambda \leqslant t < t_*$, and using Proposition 3.4 we conclude that $\mathscr{D}X(q)$ is nonsingular,

$$\frac{1}{\|\mathscr{D}X(q)^{-1}\|} \geqslant |f'(\lambda)| \geqslant |f'(t)| > 0, \tag{3.24}$$

where we used also the fact that $|f'|$ is strictly decreasing in $[0, t_*]$. Applying the definition of $\|\cdot\|_{\mathrm{op}}$, the triangular inequality, (3.24), and again the hypothesis $q \in K(t)$, we get

$$\begin{aligned}
\|X(q)+\mathscr{D}X(q)\,v\| &\geqslant \frac{1}{\|\mathscr{D}X(q)^{-1}\|_{\mathrm{op}}}\,\|\mathscr{D}X(q)^{-1}\,X(q)+v\| \\
&\geqslant \frac{1}{\|\mathscr{D}X(q)^{-1}\|_{\mathrm{op}}}\,(\|v\|-\|\mathscr{D}X(q)^{-1}\,X(q)\|) \\
&\geqslant |f'(\lambda)|(\|v\|-\|\mathscr{D}X(q)^{-1}\,X(q)\|) \\
&\geqslant |f'(t)|(\|v\|-\|\mathscr{D}X(q)^{-1}\,X(q)\|) \\
&\geqslant |f'(t)|(\|v\|-f(t)/|f'(t)|). \qquad (3.25)
\end{aligned}$$

Recall that f is a degree 2 polynomial with $f(t_*)=0$. Since $\|v\| = d(q, q_*) = t_* - t > 0$,

$$\begin{aligned}
|f'(t)|(\|v\|-f(t)/|f'(t)|) &= -((t_*-t)\,f'(t)+f(t)) \\
&= Ld(q, q_*)^2/2 + \cdots \\
&\quad -((t_*-t)^2\,f''(t)/2+(t_*-t)\,f'(t)+f(t)) \\
&= Ld(q, q_*)^2/2 \\
&> 0.
\end{aligned}$$

Therefore, the first term in inequality (3.22) is equal to the last term in the chain of inequalities (3.25). Hence, all inequalities in (3.25) hold as equalities between nonzero terms, which implies that

$$1/\|\mathscr{D}X(q)^{-1}\|_{\mathrm{op}} = |f'(\lambda)| = |f'(t)|, \qquad (3.26)$$

$$\|\mathscr{D}X(q)^{-1}\,X(q)+v\| = \|v\| - \|\mathscr{D}X(q)^{-1}\,X(q)\| > 0, \qquad (3.27)$$

$$\|\mathscr{D}X(q)^{-1}\,X(q)\| = f(t)/|f'(t)|. \qquad (3.28)$$

As $|f'|$ is strictly decreasing in $[0, t_*]$ and $0 \leqslant \lambda \leqslant t < t_*$, by (3.26), (3.23) we get

$$d(p_0, q) = t.$$

Recall that the norm we are using in T_qM cames from the Riemannian structure of M. So, there is an inner product $\langle \cdot, \cdot \rangle_q$ defined in T_qM and, for $w \in T_qM$, $\|w\| = \sqrt{\langle w, w \rangle_q}$. Hence

$$\|\mathscr{D}X(q)^{-1} X(q) + v\|^2 = \|\mathscr{D}X(q)^{-1} X(q)\|^2 + 2\langle \mathscr{D}X(q)^{-1} X(q), v \rangle_q + \|v\|^2.$$

From (3.27) we also have

$$\|\mathscr{D}X(q)^{-1} X(q) + v\|^2 = \|\mathscr{D}X(q)^{-1} X(q)\|^2 - 2 \|\mathscr{D}X(q)^{-1} X(q)\| \, \|v\| + \|v\|^2.$$

Therefore,

$$\|\mathscr{D}X(q)^{-1} X(q)\| \, \|v\| = \langle \mathscr{D}X(q)^{-1} X(q), -v \rangle_q$$

and $\mathscr{D}X(q)^{-1} X(q)$, v are linearly dependent. As $\|v\| > 0$,

$$\mathscr{D}X(q)^{-1} X(q) = -rv$$

for some $r \geqslant 0$. By (3.28), (3.27), $0 < r < 1$. Hence

$$\begin{aligned} q_+ &= \exp_q(rv) \\ &= \alpha(r). \end{aligned}$$

But, α is a minimizing geodesic joining q to q_* and $0 < r < 1$. Therefore,

$$d(q, q_*) = d(q; q_+) + d(q_+, q_*)$$

and

$$\begin{aligned} d(q, q_+) &= \|rv\| \\ &= \|\mathscr{D}X(q)^{-1} X(q)\| \\ &= f(t)/|f'(t)| = t_+ - t, \end{aligned}$$

where (3.28) was also used. So, using also the hypotheses of the lemma we get

$$\begin{aligned} d(q, q_+) &= d(q, q_*) - d(q_+, q) \\ &= (t_* - t) - (t_+ - t) \\ &= t_* - t_+. \end{aligned}$$

The above equation readily implies $t_+ + d(q_+, q_*) = t_*$. ∎

COROLLARY 3.9. *Suppose that* $q_* \in B[p_0, t_*]$ *is a singular point of* X. *If for some* $\tilde{t}$, $\tilde{q}$

$$0 \leqslant \tilde{t} < t_*, \qquad \tilde{q} \in K(\tilde{t}),$$

and $\tilde{t} + d(\tilde{q}, q_*) = t_*$, *then*

$$d(p_0, q_*) = t_*.$$

Proof. By substituting $t = \tilde{t}$ and $q = \tilde{q}$ into Corollary 3.6 we obtain, for all k,

$$q_k \in K(\tau_k).$$

Furthermore, $\{q_k\}$ converges to some $\tilde{q}_* \in B[p_0, t_*]$ singular point of X. Now by induction; applying Lemma 3.8, it is easy to show that for all k,

$$d(p_0, q_k) = \tau_k \tag{3.29}$$

and

$$\tau_k + d(q_k, q_*) = t_*. \tag{3.30}$$

By Proposition 3.3, $\{\tau_k\}$ converges to t_*. Taking the limit in (3.30), we conclude that $\tilde{q}_* = q_*$, and taking the limit in (3.29), we obtain $d(p_0, q_*) = t_*$. ∎

LEMMA 3.10. *The limit* p_* *of the sequence* $\{p_k\}$ *is the unique singular point of* X *in* $B[p_0, t_*]$.

Proof. Let q_* be a singular point of X in $B[p_0, t_*]$. We claim that, for all k,

$$d(p_k, q_*) + t_k \leqslant t_*. \tag{3.31}$$

To prove this we will analyze two possibilities:

(a) $d(p_0, q_*) < t_*$

(b) $d(p_0, q_*) = t_*$.

a. In this case, we use induction to prove that for all k,

$$d(p_k, q_*) + t_k < t_*. \tag{3.32}$$

Indeed, for $k=0$ this is true because $t_0=0$. Suppose (3.32) is true for some k. Define $\gamma_k(\theta)=\exp_{p_k}(-\theta\mathscr{D}X(p_k)^{-1}X(p_k))$. Then $p_{k+1}=\gamma_k(1)$ and, from Lemma 3.7, for any $\theta\in[0,1]$

$$\gamma_k(\theta)\in K(t_k+\theta(t_{k+1}-t_k)).$$

Define $\varphi\colon[0,1]\to\mathbb{R}$ by

$$\varphi(\theta)=d(\gamma_k(\theta),q_*)+(t_k+\theta(t_{k+1}-t_k)).$$

This is a continuous function and $\varphi(0)=d(p_k,q_*)+t_k<t_*$. We show indirectly that $\varphi(\theta)\neq t_*$ for any $\theta\in[0,1]$. Assume the contrary: for some $\tilde{\theta}\in[0,1]$,

$$\varphi(\tilde{\theta})=t_*. \tag{3.33}$$

Take $\tilde{q}=\gamma_k(\tilde{\theta})$, $\tilde{t}=t_k+\tilde{\theta}(t_{k+1}-t_k)$. Using Lemma 3.7 and (3.33) it follows that

$$\tilde{q}\in K(\tilde{t}),\qquad d(\tilde{q},q_*)+\tilde{t}=t_*.$$

Applying Corollary 3.9, we conclude that $d(p_0,q_*)=t_*$, which contradicts our assumption. Hence $\varphi(\theta)\neq t_*$ for any $\theta\in[0,1]$. Since $\varphi(0)<t_*$, it follows that $\varphi(1)<t_*$; i.e., $d(p_{k+1},q_*)+t_{k+1}<t_*$ and (3.32) holds also for $k+1$.

b. Use induction again to prove that for all k,

$$d(p_k,q_*)+t_k=t_*. \tag{3.34}$$

Indeed, for $k=0$ this is true because $t_0=0$. Suppose (3.34) is true for some k. Since $p_k\in K(t_k)$ and (3.34) is true, use Lemma 3.8 to conclude that $d(p_{k+1},q_*)+t_{k+1}=t_*$. Therefore, (3.34) also holds for $k+1$.

From (3.32) and (3.34) we obtain (3.31). Using the convergence of $\{t_k\}$ to t_*, it follows that $q_*=\lim p_k=p_*$. ∎

LEMMA 3.11. *If $\ell<1/2$ and $B(p_0,r)\subseteq\Omega$ with $t_*<r\leqslant t_{**}=(1/aL)(1+\sqrt{1-2\ell})$, then the limit p_* of the sequence $\{p_k\}$ is the unique singular point of X in $B(p_0,r)$.*

Proof. Let q_* be a singular point of X in $B(p_0, r)$. Now consider the minimizing geodesic $\alpha : [0, 1] \to M$ joining p_0 to q_*. Observe that $\alpha([0, 1]) \subseteq \Omega$. Setting $v = \alpha'(0)$ and applying Lemma 2.3,

$$P_{\alpha, 0, 1} X(q_*) = X(p_0) + \mathscr{D}X(p_0)\, v + R(1),$$

with $\|R(1)\| \leqslant L\, \|v\|^2/2$. Since $X(q_*) = 0$ and $\|v\| = d(p_0, q_*)$, from the last equality, the definition of $\|\cdot\|_{\text{op}}$, the hypotheses of Theorem 3.2, and the triangular inequality we get

$$\begin{aligned} Ld(p_0, q_*)^2/2 &\geqslant \|X(p_0) + \mathscr{D}X(p_0)\, v\| \\ &\geqslant \frac{1}{\|\mathscr{D}X(p_0)^{-1}\|_{\text{op}}} \|\mathscr{D}X(p_0)^{-1}\, X(p_0) + v\| \\ &\leqslant \frac{1}{a} \|\mathscr{D}X(p_0)^{-1}\, X(p_0) + v\| \\ &\geqslant \frac{1}{a} (\|v\| - \|\mathscr{D}X(p_0)^{-1}\, X(p_0)\|) \\ &\geqslant \frac{1}{a} (d(p_0, q_*) - b) \\ &= d(p_0, q_*)/a - b/a. \end{aligned}$$

Therefore,

$$\begin{aligned} f(d(p_0, q_*)) &= \frac{L}{2} d(p_0, q_*)^2 - \frac{1}{a} d(p_0, q_*) + \frac{a}{b} \\ &\geqslant 0. \end{aligned} \tag{3.35}$$

Since $d(p_0, q_*) < t_{**}$, the equation (3.35) implies that

$$d(p_0, q_*) \leqslant t_*.$$

From Lemma 3.10 we have $q_* = p_*$. ∎

This concludes the second part. Therefore Theorem 3.2 is proved.

4. FINAL REMARKS

In the original Kantorovich Theorem, finite dimensionality plays no role. We restricted our study to finite dimensional Riemannian manifolds. It seems that under suitable assumptions the present results may be extended to infinite dimensional Riemannian manifolds.

ACKNOWLEDGMENTS

We thank the anonymous referee for the corrections and constructive suggestions which led to improvement of the paper.

REFERENCES

1. J. Appel, E. De Pascale, J. V. Lysenko, and P. P. Zabrejko, New results on Newton–Kantorovich approximations with applications to nonlinear integral equations, *Numer. Funct. Anal. Optim.* **18**, Nos. 1/2 (1997), 1–17.
2. I. K. Argyros, A unified approach for constructing fast tow-step Newton-like methods, *Monatsh. Math.* **119** (1995), 1–22.
3. A. Cauchy, Sur la détermination des racines d'une équation algébrique ou transcendente, *in* "Œuvres Complétes (II)," Vol. 4, pp. 573–609, Gauthier–Villars, Paris, 1899.
4. J. X. da Cruz Neto, "Métodos Geodésicos na Progamação Matemática," Ph.D. thesis, COPPE/UFRJ, Rio de Janeiro, Brazil, 1995.
5. J. E. Dennis, On the Kantorovich hypothesis for Newton's method, *SIAM J. Numer. Anal.* **6** (1969), 493–507.
6. M. P. do Carmo, "Riemannian Geometry," Birkhäuser, Boston, 1992.
7. A. Edelman, T. A. Arias, and T. Smith, The geometry of algorithms with orthogonality constraints, *SIAM J. Matrix Anal. Appl.* **20**, No. 2 (1998), 303–353.
8. D. Gabay, Minimizing a differentiable function over a differential manifold, *J. Optim. Theory Appl.* **37**, No. 2 (1982), 177–219.
9. K. P. Hadeler, Shadowing orbits and Kantorovich's theorem, *Numer. Math.* **73** (1996), 65–73.
10. F. B. Hildebrand, "Introduction to Numerical Analysis," McGraw–Hill, New York, 1956.
11. L. V. Kantorovich, On Newton's method for functional equations, *Dokl. Akad. Nauk. SSSR* **59** (1984), 1237–1240.
12. L. V. Kantorovich, Further applications of Newton's Method, *Vestnik Leningr. Univ.* **2**, No. 2 (1957), 68–103.
13. L. V. Kantorovich and G. P. Akilov, "Functional Analysis in Normed Spaces," Pergamon, Oxford, 1964.
14. J. Moser, A new techniques for the construction of solutions of nonlinear differential equations, *Proc. Natl. Acad. Sci. USA* **47** (1961), 1824–1831.
15. J. Nash, The embedding problem for Riemannian manifolds, *Ann. Math.* **63** (1956), 20–63.
16. J. M. Ortega, The Newton–Kantorovich theorem, *Amer. Math. Monthly* **75** (1968), 658–660.

17. J. M. Ortega and W. C. Rheimboldt, "Interactive Solution of Nonlinear Equations in Several Variables," Academic Press, New York, 1970.

18. C. Runge, Separation und Approximation der Wurzeln von Gleichungen, *in* "Enzyklopädie der Mathematichen Wissenschaften," Vol. 1, pp. 405–449, Teubner, Leipzig.

19. M. Shub and S. Smale, Complexity of Bezout's theorem. I. Geometric aspects, *J. Amer. Math. Soc.* **6**, No. 2 (1993), 459–499.

20. S. Smale, Newton method estimates form data at one point, *in* "The Merging of Disciplines: New Directions in Pure, Applied and Computational Mathematics" (R. Ewing, K. Gross, and C. Martin, Eds.), pp. 185–196, Springer-Verlag, New York, 1986.

21. S. T. Smith, Optimization Techniques on Riemannian Manifolds, *in* "Fields Institute Communications," Vol. 3, pp. 113–146, American Mathematical Society, Providence, RI, 1994.

22. S. T. Smith, "Geometric Optimization Method for Adaptive Filtering," Ph.D. thesis, Harvard University, Cambridge, MA, 1993.

23. J. L. Synge and A. Schild, "Tensor Calculus," Dover, New York, 1949.

24. C. Udriste, "Convex Functions and Optimization Methods on Riemannian Manifolds," Mathematics and Its Applications, Vol. 297, Kluwer Academic, Dordrecht, 1994.

25. T. Yamamoto, A method for finding sharp error bounds for Newton's method under the Kantorovich assumptions, *Numer. Math.* **49** (1986), 203–220.

26. C. E. Wayne, "An Introduction to KAM Theory. Dynamical Systems and Probabilistic Methods in Partial Differential Equation," Lectures in Applied Mathematics, Vol. 31, American Mathematical Society, Providence, RI, 1996.