

Convergence of the Newton method and uniqueness of zeros of vector fields on Riemannian manifolds

LI Chong & WANG Jinhua

Department of Mathematics, Zhejiang University, Hangzhou 310027, China

Correspondence should be addressed to Li Chong (email: cli@zju.edu.cn)

Received June 16, 2004

Abstract The estimates of the radii of convergence balls of the Newton method and uniqueness balls of zeroes of vector fields on the Riemannian manifolds are given under the assumption that the covariant derivatives of the vector fields satisfy some kind of general Lipschitz conditions. Some classical results such as the Kantorovich's type theorem and the Smale's γ -theory are extended.

Keywords: Riemannian manifold, Newton method, convergence ball, uniqueness ball.

DOI: 10.1360/04ys0147

1 Introduction

The Newton method and its variations are the most efficient methods known for solving systems of nonlinear equations when they are continuously differentiable. Besides its practical applications, the Newton method is also a powerful theoretical tool. One of the famous results on the Newton method is the well-known Kantorovich's theorem^[1], which has the advantage that the Newton sequence converges to a solution under very mild conditions. Another important result on the Newton method is the Smale's point estimate theory which was presented by Smale in his report written for the 20th International Conference of Mathematician^[2]. In this theory, the notion to be an approximation zero was introduced and the rule to judge an initial point of an approximation zero was provided, depending only on the information of the nonlinear operator at the initial point. Other results on the Newton method such as the estimates of the radii of convergence balls were given by Traub and Wozniakowski^[3], and Wang^[4] independently. A big step in this direction was made recently by Wang^[5,6], where the Kantorovich's theorem and the Smale's theory were unified and extended.

While the Newton method in the Riemannian manifolds was studied by many authors^[7-11], extensions of the Kantorovich's theorem and the Smale's α -theory and γ -theory were just made recently by Ferreira and Svaiter^[12] and Dedieu, Priouret and

Malajovich^[13]. Other extensions about the local behavior of the Newton method in the Riemannian manifolds have not been found to be studied in our best knowledge. The purpose of the present paper is to estimate the radii of convergence balls of the Newton method and uniqueness balls of zeros of vector fields on the Riemannian manifolds under general Lipschitz conditions introduced by Wang^[5]. In particular, the Smale's γ -theory is extended.

We conclude this introduction with a short remark that the issue on mappings from manifolds to the n -dimensional spaces can be addressed along almost the same path and does not need further elaboration here.

2 Notions and preliminaries

We begin with some basic notions and notations. Most of them are standard, see for example refs. [14,15]. Let M be a real complete n -dimensional Riemannian manifold. Let $p', p \in M$ and let $c : [0, 1] \rightarrow M$ be a piecewise smooth curve connecting p' and p . Then the arc length of c is defined by $l(c) := \int_0^1 \|c'(t)\| dt$, and the Riemannian distance from p' to p is defined by $d(p', p) := \inf_c l(c)$, where the infimum is taken over all piecewise smooth curves $c : [0, 1] \rightarrow M$ connecting p' and p . Thus (M, d) is a complete metric space by the Hopf-Rinow Theorem^[14,15].

For a Banach space or a Riemannian manifold Z , let $\mathbf{B}_Z(p, r)$ and $\overline{\mathbf{B}}_Z(p, r)$ denote respectively the open metric ball and the closed metric ball at p with radius r , that is,

$$\begin{aligned}\mathbf{B}_Z(p, r) &= \{q \in Z : d(p, q) < r\}, \\ \overline{\mathbf{B}}_Z(p, r) &= \{q \in Z : d(p, q) \leq r\}.\end{aligned}$$

In particular, we write respectively $\mathbf{B}(p, r)$ and $\overline{\mathbf{B}}(p, r)$ for $\mathbf{B}_M(p, r)$ and $\overline{\mathbf{B}}_M(p, r)$ in the case when M is a Riemannian manifold.

Let $T_p M$ denote the tangent space at p to M and let $\langle \cdot, \cdot \rangle$ be the scalar product on $T_p M$ with the associated norm $\|\cdot\|_p$, where the subscript p is sometimes omitted. The tangent bundle of M is denoted by TM and defined by

$$TM = \bigcup_{p \in M} T_p M.$$

Noting that M is complete, the exponential map at p $\exp_p : T_p M \rightarrow M$ is well defined on $T_p M$. Furthermore, the radius of injectivity of the exponential map at p is denoted by $\mathbf{r}_p > 0$. Thus, \exp_p is a one to one mapping from $\mathbf{B}_{T_p M}(0, \mathbf{r}_p)$ to $\mathbf{B}(p, \mathbf{r}_p)$. Recall that a geodesic α in M connecting p' and p is called a minimizing geodesic if its arc length equals its Riemannian distance between p' and p . Note that there is at least one minimizing geodesic connecting p' and p . In particular, the curve $\alpha : [0, 1] \rightarrow M$ is a minimizing geodesic connecting p' and p if and only if there exists a vector $v \in T_{p'} M$ such that $\|v\| = d(p', p)$ and $\alpha(t) = \exp_{p'}(tv)$ for each $t \in [0, 1]$.

Let ∇ be the Levi-Civita connection on M . For any two vector fields X and Y on M , the covariant derivative of X with respect to Y is denoted by $\nabla_Y X$. Define the linear

map $DX(p) : T_pM \rightarrow T_pM$ by

$$DX(p)(u) = \nabla_Y X(p), \quad \forall u \in T_pM,$$

where Y is a vector field satisfying $Y(p) = u$. Note that the value $DX(p)(u)$ of $DX(p)$ at u depends only on the tangent vector $u = Y(p) \in T_pM$ since ∇ is tensorial in Y . We still need the notion of the parallel transport. Let $c : \mathbb{R} \rightarrow M$ be a C^∞ curve. Then the parallel transport along c is denoted by $P_{c, \cdot, \cdot}$ and defined by

$$P_{c,a,b}(v) = V(b), \quad \forall a, b \in \mathbb{R} \text{ and } v \in T_{c(a)}M,$$

where V is the unique vector field on c satisfying $\nabla_{c'(t)}V = 0$ and $V(a) = v$. Then, for any $a, b \in \mathbb{R}$, $P_{c,a,b}$ is an isometry from $T_{c(a)}M$ to $T_{c(b)}M$. Note that, for any $a, b, b_1, b_2 \in \mathbb{R}$,

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

We also need the number K_p ^[13], which is related to the sectional curvature at $p \in M$ and defined by

$$K_p = \sup \frac{d(\exp_q(u), \exp_q(v))}{\|u - v\|_q}, \tag{2.1}$$

where the supremum is taken over all $q \in \mathbf{B}(p, r_p)$, and $v, u - v \in \overline{\mathbf{B}_{T_qM}(0, r_p)}$ with $u \neq v$.

Remark 2.1. (1) K_p measures how fast the geodesics spread apart in M . In particular, if $u = 0$ or more generally if u and v are on the same line through 0, then

$$d(\exp_q(u), \exp_q(v)) = \|u - v\|_q.$$

This means that $K_p \geq 1$.

(2) In the case when M has a non-negative sectional curvature, the geodesics spread apart less than the rays, see ref. [15], so that

$$d(\exp_q(u), \exp_q(v)) \leq \|u - v\|_q$$

and consequently $K_p = 1$. Examples of manifolds with a non-negative sectional curvature are given in ref. [15], see also ref. [13].

Throughout the whole paper, we assume that $X : M \rightarrow TM$ is a C^1 vector field. Then, following ref. [12], for any given initial point p_0 in M , the Newton method for X is defined by

$$p_{n+1} = \exp_{p_n}(-DX(p_n)^{-1}X(p_n)), \quad n = 0, 1, 2, \dots \tag{2.2}$$

The following definition extends the notions of the center Lipschitz condition and the radius Lipschitz condition with the L average in Banach spaces, which were first introduced by Wang^[5] for the study of the existence of the solution of the nonlinear operator equation and the convergence of the Newton method. Let L be a positive integrable function on $[0, R]$, where R is a positive number such that $\int_0^R L(u)du > 1$.

Definition 2.1. Let $0 < r \leq R$. Let p^* be a point of M such that $DX(p^*)^{-1}$ exists. Then $DX(p^*)^{-1}DX$ is said to satisfy

(i) the center Lipschitz condition with the L average in the ball $\mathbf{B}(p^*, r)$ if, for each point p in $\mathbf{B}(p^*, r)$ and any minimizing geodesic $c : [0, 1] \rightarrow M$ connecting p^* and p ,

$$\|DX(p^*)^{-1}(P_{c,1,0}DX(p)P_{c,0,1} - DX(p^*))\| \leq \int_0^{d(p^*,p)} L(u)du; \tag{2.3}$$

(ii) the radius Lipschitz condition with the L average in the ball $\mathbf{B}(p^*, r)$ if, for each point p in $\mathbf{B}(p^*, r)$ and any minimizing geodesic $c : [0, 1] \rightarrow M$ connecting p^* and p ,

$$\begin{aligned} & \|DX(p^*)^{-1}P_{c,1,0}(DX(p) - P_{c,\tau,1}DX(c(\tau))P_{c,1,\tau})\| \\ & \leq \int_{\tau d(p^*,p)}^{d(p^*,p)} L(u)du, \quad \forall 0 \leq \tau \leq 1. \end{aligned} \tag{2.4}$$

In particular, in the special case when the function L is a constant, $DX(p^*)^{-1}DX$ is said to satisfy the center Lipschitz condition (*resp.* the radius Lipschitz condition) with the Lipschitz constant L in the ball $\mathbf{B}(p^*, r)$.

Clearly, the radius Lipschitz condition with the L average implies the center Lipschitz condition with the L average. The following lemma, which is taken from ref. [12], will play a key role in the present paper.

Lemma 2.1. Let $c : \mathbb{R} \rightarrow M$ be a C^∞ curve. Then

$$P_{c,t,0}X(c(t)) = X(c(0)) + \int_0^t P_{c,s,0}(DX(c(s))c'(s))ds.$$

3 Convergence balls of the Newton method

Throughout this section and the next one, we always assume that X is a C^1 vector field on M and that p^* is a point in M such that $DX(p^*)^{-1}$ exists. We begin with the following lemma. Recall that L is a positive integrable function on $[0, R]$ and let $r_0 > 0$ such that

$$\int_0^{r_0} L(u)du = 1. \tag{3.1}$$

Lemma 3.1. Let $r \leq r_0$. Suppose that $DX(p^*)^{-1}DX$ satisfies the center Lipschitz condition with the L average in the ball $\mathbf{B}(p^*, r)$. Then, for each point p in the ball $\mathbf{B}(p^*, r)$, $DX(p)^{-1}$ exists and

$$\|DX(p)^{-1}P_{c,0,1}DX(p^*)\| \leq \frac{1}{1 - \int_0^{d(p,p^*)} L(u)du}, \tag{3.2}$$

where $c : [0, 1] \rightarrow M$ is a minimizing geodesic connecting p^* and p .

Proof. Since $c : [0, 1] \rightarrow M$ is a minimizing geodesic connecting p^* and p , by (2.3) and (3.1), we have that

$$\|DX(p^*)^{-1}P_{c,1,0}DX(p)P_{c,0,1} - \mathbf{I}_{T_{p^*}M}\| \leq \int_0^{d(p,p^*)} L(u)du < \int_0^{r_0} L(u)du = 1,$$

where $\mathbf{I}_{T_{p^*}M}$ is the identity on $T_{p^*}M$. It follows from the Banach lemma that the inverse

of $DX(p^*)^{-1}P_{c,1,0}DX(p)P_{c,0,1}$ exists and

$$\| P_{c,1,0}DX(p)^{-1}P_{c,0,1}DX(p^*) \| \leq \frac{1}{1 - \int_0^{d(p,p^*)} L(u)du}.$$

Hence $DX(p)^{-1}$ and (3.2) holds exists because $P_{c,1,0}$ and $P_{c,0,1}$ are isometries. The proof is complete.

In the remainder of this section, we assume in addition that L is nondecreasing on $[0, R]$. Then the main theorem of this section is as follows:

Theorem 3.1. Let $r_c > 0$ satisfy

$$\frac{K_{p^*}}{r_c} \int_0^{r_c} L(u) \left(u + \frac{r_c}{K_{p^*}} \right) du \leq 1. \tag{3.3}$$

Let $r = \min(r_{p^*}, r_c)$. Suppose that $X(p^*) = 0$ and $DX(p^*)^{-1}DX$ satisfies the radius Lipschitz condition with the L average in the ball $\mathbf{B}(p^*, r)$. Then the Newton method (2.2) is well defined and convergent for each $p_0 \in \mathbf{B}(p^*, r)$, and

$$d(p_n, p^*) \leq q^{2^n - 1} d(p_0, p^*), \quad n = 1, 2, \dots, \tag{3.4}$$

where

$$q = \frac{K_{p^*} \int_0^{d(p_0, p^*)} L(u)u du}{d(p_0, p^*) \left(1 - \int_0^{d(p_0, p^*)} L(u)du \right)} < 1. \tag{3.5}$$

Proof. Let $p_0 \in \mathbf{B}(p^*, r)$ and let q be defined by (3.5). Then, by the monotonicity of L , it follows from ref. [5] that the function ϕ defined by

$$\phi(t) = \frac{1}{t^2} \int_0^t L(u)u du, \quad \forall t \in (0, r)$$

is nondecreasing. Since $r \leq r_c \leq r_0$, it follows from (3.3) that

$$\begin{aligned} q &= \frac{K_{p^*} \int_0^{d(p_0, p^*)} L(u)u du}{d^2(p_0, p^*) \left(1 - \int_0^{d(p_0, p^*)} L(u)du \right)} d(p_0, p^*) \\ &\leq \frac{K_{p^*} \int_0^{r_c} L(u)u du}{r_c^2 \left(1 - \int_0^{r_c} L(u)du \right)} d(p_0, p^*) \\ &\leq \frac{d(p_0, p^*)}{r_c} < 1. \end{aligned}$$

Below we will show that $p_n \in \mathbf{B}(p^*, r)$ and (3.4) holds for each $n = 0, 1 \dots$ by mathematical induction. Clearly, the case when $n = 0$ is trivial. Now assume that $p_n \in \mathbf{B}(p^*, r)$ and (3.4) holds for n . We have to prove that p_{n+1} is well defined, $p_{n+1} \in \mathbf{B}(p^*, r)$ and (3.4) holds for $n + 1$. For this purpose, note that there exists $v \in T_{p^*}M$ such that $p_n = \exp_{p^*}(v)$ and $\|v\| = d(p_n, p^*)$. Define $\alpha(t) = \exp_{p^*}(tv)$ for each $t \in [0, 1]$. Then, α is a minimizing geodesic. Hence, by Lemma 3.1, $DX(p_n)^{-1}$ exists and

$$\| DX(p_n)^{-1}P_{\alpha,0,1}DX(p^*) \| \leq \frac{1}{1 - \int_0^{d(p_n, p^*)} L(u)du}. \tag{3.6}$$

This shows that p_{n+1} is well defined. Thus, to complete the proof, it remains to verify that (3.4) holds for $n + 1$. We claim that

$$\| -DX(p_n)^{-1}X(p_n) - (-P_{\alpha,0,1}v) \| \leq \frac{\int_0^{d(p_n,p^*)} L(u)u du}{1 - \int_0^{d(p_n,p^*)} L(u)du}. \tag{3.7}$$

Since $\alpha'(\tau) = P_{\alpha,0,\tau}v$, it follows from Lemma 2.1 that

$$\begin{aligned} P_{\alpha,1,0}X(p_n) - X(p^*) &= \int_0^1 P_{\alpha,\tau,0}DX(\alpha(\tau))\alpha'(\tau)d\tau \\ &= \int_0^1 P_{\alpha,\tau,0}DX(\alpha(\tau))P_{\alpha,0,\tau}v d\tau. \end{aligned} \tag{3.8}$$

Then, using (3.6) and (2.4), we get that

$$\begin{aligned} &\| -DX(p_n)^{-1}X(p_n) - (-P_{\alpha,0,1}v) \| \\ &\leq \| -DX(p_n)^{-1}P_{\alpha,0,1}(P_{\alpha,1,0}X(p_n) - X(p^*)) + P_{\alpha,0,1}v \| \\ &\leq \| DX(p_n)^{-1} \int_0^1 (DX(p_n) - P_{\alpha,\tau,1}DX(\alpha(\tau))P_{\alpha,1,\tau})P_{\alpha,0,1}v d\tau \| \\ &\leq \| DX(p_n)^{-1}P_{\alpha,0,1}DX(p^*) \| \\ &\quad \cdot \int_0^1 \| DX(p^*)^{-1}P_{\alpha,1,0}(DX(p_n) - P_{\alpha,\tau,1}DX(\alpha(\tau))P_{\alpha,1,\tau}) \| \| v \| d\tau \\ &\leq \frac{1}{1 - \int_0^{d(p_n,p^*)} L(u)du} \int_0^1 \int_{\tau d(p_n,p^*)}^{d(p_n,p^*)} L(u)dud(p_n,p^*)d\tau \\ &= \frac{\int_0^{d(p_n,p^*)} L(u)u du}{1 - \int_0^{d(p_n,p^*)} L(u)du}. \end{aligned}$$

Hence (3.7) holds. Consequently,

$$\| -DX(p_n)^{-1}X(p_n) - (-P_{\alpha,0,1}v) \| \leq qd(p_n,p^*) < \mathbf{r}_{p^*}.$$

On the other hand, since

$$\| -P_{\alpha,0,1}v \| = \|v\| = d(p_n,p^*) < \mathbf{r}_{p^*},$$

in view of the definition of K_{p^*} and (3.7), we have

$$d(\exp_{p_n}(-DX(p_n)^{-1}X(p_n)), \exp_{p_n}(-P_{\alpha,0,1}v)) \leq \frac{K_{p^*} \int_0^{d(p_n,p^*)} L(u)u du}{1 - \int_0^{d(p_n,p^*)} L(u)du}. \tag{3.9}$$

As $p_{n+1} = \exp_{p_n}(-DX(p_n)^{-1}X(p_n))$ and $p^* = \exp_{p_n}(-P_{\alpha,0,1}v)$, (3.9) means that

$$\begin{aligned} d(p_{n+1}, p^*) &\leq \frac{K_{p^*} \int_0^{d(p_n, p^*)} L(u)u du}{d^2(p_n, p^*) \left(1 - \int_0^{d(p_n, p^*)} L(u)du\right)} d^2(p_n, p^*) \\ &\leq \frac{K_{p^*} \int_0^{d(p_0, p^*)} L(u)u du}{d^2(p_0, p^*) \left(1 - \int_0^{d(p_0, p^*)} L(u)du\right)} d^2(p_n, p^*) \\ &\leq \frac{q}{d(p_0, p^*)} (q^{2^n} - 1)^2 d(p_0, p^*)^2. \\ &= q^{2^{n+1} - 1} d(p_0, p^*). \end{aligned}$$

Therefore (3.4) holds for $n + 1$, thus completing the proof.

In the case when M has non-negative sectional curvature, $K_{p^*} = 1$ by Remark 2.1. Thus we get immediately the following.

Corollary 3.1. Suppose that M has non-negative sectional curvature. Let $r_c > 0$ satisfy

$$\frac{1}{r_c} \int_0^{r_c} L(u) (u + r_c) du \leq 1. \tag{3.10}$$

Let $r = \min(r_{p^*}, r_c)$. Suppose that $X(p^*) = 0$ and $DX(p^*)^{-1}DX$ satisfies the radius Lipschitz condition with the L average in the ball $\mathbf{B}(p^*, r)$. Then the Newton method (2.2) is well defined and convergent for each $p_0 \in \mathbf{B}(p^*, r)$, and (3.4) holds for $q < 1$ defined by

$$q = \frac{\int_0^{d(p_0, p^*)} L(u)u du}{d(p_0, p^*) \left(1 - \int_0^{d(p_0, p^*)} L(u)du\right)}. \tag{3.11}$$

4 Uniqueness balls of zeros

Throughout this section, we assume that L is positively integrable. Recall that X is a C^1 vector field on M and that p^* is a point of M such that $DX(p^*)^{-1}$ exists. Then the main theorem of this section is as follows:

Theorem 4.1. Let $r_u > 0$ such that

$$\frac{K_{p^*}}{r_u} \int_0^{r_u} L(u)(r_u - u)du \leq 1. \tag{4.1}$$

Let $r = \min(r_{p^*}, r_u)$. Suppose that $X(p^*) = 0$ and $DX(p^*)^{-1}DX$ satisfies the center Lipschitz condition with the L average in the ball $\mathbf{B}(p^*, r)$. Then p^* is the unique solution

of $X(p) = 0$ in $\mathbf{B}(p^*, r)$.

Proof. Let $p_0 \in \mathbf{B}(p^*, r)$. Then there exists $v_0 \in T_{p^*}M$ such that $p_0 = \exp_{p^*}(v_0)$ and $\|v_0\| = d(p_0, p^*)$. Define

$$p_1 = \exp_{p_0}(-P_{\alpha_0,0,1}DX(p^*)^{-1}P_{\alpha_0,1,0}X(p_0)),$$

where the curve $\alpha_0 : [0, 1] \rightarrow M$ is defined by $\alpha_0(t) = \exp_{p^*}(tv_0)$ for each $t \in [0, 1]$. We will define $\{p_n\} \subseteq M$ inductively. Having defined p_n , we take $v_n \in T_{p^*}M$ such that $p_n = \exp_{p^*}(v_n)$ and $\|v_n\| = d(p_n, p^*)$. Then p_{n+1} is defined by

$$p_{n+1} = \exp_{p_n}(-P_{\alpha_n,0,1}DX(p^*)^{-1}P_{\alpha_n,1,0}X(p_n)), \tag{4.2}$$

where the curve $\alpha_n : [0, 1] \rightarrow M$ is defined by

$$\alpha_n(t) = \exp_{p^*}(tv_n). \tag{4.3}$$

Thus, $\{p_n\}$ is well-defined. Set

$$q_0 = \frac{K_{p^*}}{d(p_0, p^*)} \int_0^{d(p_0, p^*)} L(u)(d(p_0, p^*) - u)du.$$

As $L(u) > 0$, it follows from ref. [5] that the function ψ defined by

$$\psi(t) = \frac{1}{t} \int_0^t L(u)(t - u)du, \quad \forall t \in [0, r]$$

is strictly monotonically increasing. Consequently, by (4.1), we get that

$$q_0 < \frac{K_{p^*}}{r_u} \int_0^{r_u} L(u)(r_u - u)du \leq 1.$$

We will show by mathematical induction that

$$d(p_n, p^*) \leq q_0^n d(p_0, p^*) \tag{4.4}$$

holds for $n = 0, 1, 2, \dots$. In fact, it is trivial in the case when $n = 0$. Now assume that (4.4) holds for n . Thus, to complete the proof, it suffices to prove that (4.4) holds for $n + 1$. We claim that

$$\begin{aligned} & \| -P_{\alpha_n,0,1}DX(p^*)^{-1}P_{\alpha_n,1,0}X(p_n) - (-P_{\alpha_n,0,1}v_n) \| \\ & \leq \int_0^{d(p_n, p^*)} L(u)(d(p_n, p^*) - u)du. \end{aligned} \tag{4.5}$$

Indeed, using Lemma 2.1 and (2.3), we obtain that

$$\begin{aligned} & \| -P_{\alpha_n,0,1}DX(p^*)^{-1}P_{\alpha_n,1,0}X(p_n) - (-P_{\alpha_n,0,1}v_n) \| \\ & \leq \| -P_{\alpha_n,0,1}DX(p^*)^{-1}(P_{\alpha_n,1,0}X(p_n) - X(p^*)) + P_{\alpha_n,0,1}v_n \| \\ & \leq \| -P_{\alpha_n,0,1}DX(p^*)^{-1} \int_0^1 (P_{\alpha_n,\tau,0}DX(\alpha(\tau))P_{\alpha_n,0,\tau} - DX(p^*))v_n d\tau \| \\ & \leq \int_0^1 \| DX(p^*)^{-1}(P_{\alpha_n,\tau,0}DX(\alpha(\tau))P_{\alpha_n,0,\tau} - DX(p^*)) \| \cdot \| v_n \| d\tau \\ & \leq \int_0^1 \int_0^{\tau d(p_n, p^*)} L(u)du \cdot d(p_n, p^*)d\tau \\ & = \int_0^{d(p_n, p^*)} L(u)(d(p_n, p^*) - u)du. \end{aligned}$$

Therefore the claim stands. From (4.5), it is seen that

$$\| -P_{\alpha_n,0,1}DX(p^*)^{-1}P_{\alpha_n,1,0}X(p_n) - (-P_{\alpha_n,0,1}v_n) \| \leq q_0d(p_n, p^*) < \mathbf{r}_{p^*}.$$

On the other hand, as

$$\| -P_{\alpha_n,0,1}v_n \| = \| v_n \| = d(p_n, p^*) < \mathbf{r}_{p^*},$$

by (2.1) and (4.5), one gets that

$$\begin{aligned} & d(\exp_{p_n}(-P_{\alpha_n,0,1}DX(p^*)^{-1}P_{\alpha_n,1,0}X(p_n)), \exp_{p_n}(-P_{\alpha_n,0,1}v_n)) \\ & \leq K_{p^*} \int_0^{d(p_n, p^*)} L(u)(d(p_n, p^*) - u)du. \end{aligned} \tag{4.6}$$

Since

$$p_{n+1} = \exp_{p_n}(-P_{\alpha_n,0,1}DX(p^*)^{-1}P_{\alpha_n,1,0}X(p_n))$$

and

$$p^* = \exp_{p_n}(-P_{\alpha_n,0,1}v_n),$$

(4.6)implies that

$$d(p_{n+1}, p^*) \leq K_{p^*} \int_0^{d(p_n, p^*)} L(u)(d(p_n, p^*) - u)du \leq q_0d(p_n, p^*) \leq q_0^{n+1}d(p_0, p^*).$$

Therefore, (4.4) holds for $n + 1$. By (4.4), we have

$$\lim_{n \rightarrow +\infty} p_n = p^*. \tag{4.7}$$

Below we will prove that p^* is the unique solution of $X(p) = 0$ in $\mathbf{B}(p^*, r)$. In fact, if otherwise, there exists a point q^* in $\mathbf{B}(p^*, r) \setminus \{p^*\}$ such that $X(q^*) = 0$. Then let $\{p_n\}$ be the sequence generated by (4.2) with the initial point $p_0 = q^*$. Then $p_n = q^*$ for all $n = 0, 1, \dots$ and hence $q^* = p^*$ by (4.7). This is a contradiction and hence completes the proof.

In particular, in the case when M has non-negative sectional curvature, we have the following.

Corollary 4.1. Suppose that M has non-negative sectional curvature. Let $r_u > 0$ such that

$$\frac{1}{r_u} \int_0^{r_u} L(u)(r_u - u)du \leq 1. \tag{4.8}$$

Let $r = \min(\mathbf{r}_{p^*}, r_u)$. Suppose that $X(p^*) = 0$ and $DX(p^*)^{-1}DX$ satisfies the center Lipschitz condition with the L average in the ball $\mathbf{B}(p^*, r)$. Then p^* is the unique solution of $X(p) = 0$ in $\mathbf{B}(p^*, r)$.

5 Applications

This section is devoted to some applications of the results obtained in the preceding sections. At first, in the case when the Lipschitz conditions with the Lipschitz constant $L > 0$ in the ball $\mathbf{B}(p^*, r)$ are satisfied, it is easy to get, by (3.3) and (4.1),

$$r_c = \frac{2}{2L + K_{p^*}L} \quad \text{and} \quad r_u = \frac{2}{LK_{p^*}}. \tag{5.1}$$

Hence, by Theorems 3.1 and 4.1, the radii of the convergence balls of the Newton method and the radii of the uniqueness balls of zeros of vector fields are equal to $\min\{\mathbf{r}_{p^*}, \frac{2}{2L+K_{p^*}L}\}$ and $\min\{\mathbf{r}_{p^*}, \frac{2}{LK_{p^*}}\}$, respectively. Moreover, (3.4) holds for q defined by

$$q = \frac{K_{p^*}Ld(p_0, p^*)}{2(1 - Ld(p_0, p^*))}. \tag{5.2}$$

The subsections below will be focused on the Smale’s γ -theory as well as the approximation zero, which are the generalizations of the classical Smale’s γ -theory for the analytic operators on the Banach spaces (see for example ref. [2]) and were also studied recently by Dedieu, Priouret and Malajovich^[13].

5.1 Smale’s γ -theory

In this subsection and the next one, we always assume that M is an analytic Riemannian Manifold. We begin with the following definitions, see for example refs. [14,15].

Definition 5.1. Let $p \in M$. The vector field $X : M \rightarrow TM$ is said to be analytic at p if there exist a local coordinate system $(U, \{x^i\})$ of p and n analytic functions $X^i : U \rightarrow \mathbb{R}, i = 1, 2, \dots, n$ such that

$$X|_U = \sum_{i=1}^n X^i \frac{\partial}{\partial x^i}.$$

The vector field X is said to be analytic on M if it is analytic at each point of M .

Definition 5.2. Let $\{Y_1, \dots, Y_k\}$ be a finite sequence of vector fields on M . Then the k -th covariant derivative of X with respect to $\{Y_1, \dots, Y_k\}$ is denoted by $\nabla_{\{Y_i\}_{i=1}^k}^k X$ and defined inductively by

$$\nabla_{\{Y_i\}_{i=1}^k}^k X = \nabla_{Y_k}(\nabla_{\{Y_i\}_{i=1}^{k-1}}^{k-1} X).$$

Definition 5.3. Let $p \in M$ and $(u_1, \dots, u_k) \in (T_p M)^k$. Let $\{Y_1, \dots, Y_k\}$ be a finite sequence of vector fields on M such that $Y_i(p) = u_i$ for each $i = 1, \dots, k$. Then, the value of the k -th covariant derivative of X with respect to $\{Y_1, \dots, Y_k\}$ at p is denoted by

$$D^k X(p)(u_1, \dots, u_k) = \nabla_{\{Y_i\}_{i=1}^k}^k X(p).$$

(Note that $D^k X(p)(u_1, \dots, u_k)$ only depends on the vector (u_1, \dots, u_k) since the covariant derivative is tensorial in each Y_i .) In particular, in the case when $u_1 = u_2 = \dots = u_k = u$, we write $D^k X(p)u^k$ for $D^k X(p)(u, \dots, u)$.

Clearly, the k -th covariant derivative $D^k X(p)$ at a point p is a k -multilinear map from $(T_p M)^k$ to $T_p M$. We define the norm of $D^k X(p)$ by

$$\|D^k X(p)\|_p = \sup \|D^k X(p)(u_1, \dots, u_k)\|_p, \tag{5.3}$$

where the supremum is taken over all vectors $(u_1, \dots, u_n) \in (T_p M)^k$ with $\|u_j\| = 1$ for each $j = 1, \dots, n$.

In the remainder of this section, we assume that X is analytic on M and p^* is a point of M such that $DX(p^*)^{-1}$ exists. Following ref. [13], we define

$$\gamma(X, p^*) = \sup_{k \geq 2} \left\| DX(p^*)^{-1} \frac{D^k X(p^*)}{k!} \right\|_{p^*}^{\frac{1}{k-1}}. \tag{5.4}$$

Note that, by analyticity, $\gamma(X, p^*)$ is finite. The following Taylor formula for the vector fields can be found in refs. [13,15].

Lemma 5.1. Let $r = \min(r_{p^*}, \frac{1}{\gamma})$, where $\gamma = \gamma(X, p^*)$. Then, for any $p \in \mathbf{B}(p^*, r)$ and any $u \in T_{p^*}M$ with $p = \exp_{p^*}(u)$,

$$X(p) = P_{\alpha,0,1} \left(\sum_{k=0}^{\infty} \frac{1}{k!} D^k X(p^*) u^k \right),$$

where the curve α on $[0, 1]$ is defined by $\alpha(t) := \exp_{p^*}(tu)$ for $t \in [0, 1]$.

Taking the l -th covariant derivative in Lemma 5.1 gives the following

Corollary 5.1. With the same hypothesis, for any $l \geq 0$, we have

$$D^l X(p) = P_{\alpha,0,1} \left(\sum_{k=0}^{\infty} \frac{1}{k!} D^{k+l} X(p^*) u^k \right) P_{\alpha,1,0}^l,$$

where $P_{\alpha,1,0}^l$ stands for the map from $(T_p M)^l$ to $(T_{p^*} M)^l$ defined by

$$P_{\alpha,1,0}^l(v_1, \dots, v_l) = (P_{\alpha,1,0} v_1, \dots, P_{\alpha,1,0} v_l), \quad \forall (v_1, \dots, v_l) \in (T_p M)^l.$$

The following lemma shows that an analytic vector field satisfies the radius Lipschitz condition with the L average in the ball $\mathbf{B}(p^*, r)$ for the L defined by

$$L(u) = \frac{2\gamma}{(1 - \gamma u)^3}, \quad \forall u \in \mathbf{B}(p^*, r), \tag{5.5}$$

where $r = \min(r_{p^*}, \frac{1}{\gamma})$ and $\gamma = \gamma(X, p^*)$ is defined by (5.4).

Lemma 5.2. Let $r = \min(r_{p^*}, \frac{1}{\gamma})$ and let L be defined by (5.5). Then, $DX(p^*)^{-1}DX$ satisfies the radius Lipschitz condition with the L average in the ball $\mathbf{B}(p^*, r)$, namely, for each point $p \in \mathbf{B}(p^*, r)$ and each minimizing geodesic $\alpha : [0, 1] \rightarrow M$ such that $\alpha(0) = p^*, \alpha(1) = p$, we have that, for each $0 \leq \tau \leq 1$,

$$\| DX(p^*)^{-1}P_{\alpha,1,0}(DX(p) - P_{\alpha,\tau,1}DX(\alpha(\tau))P_{\alpha,1,\tau}) \| \leq \int_{\tau d(p^*,p)}^{d(p^*,p)} \frac{2\gamma}{(1 - \gamma u)^3} du. \tag{5.6}$$

Proof. Let $p \in \mathbf{B}(p^*, r)$ and let $\alpha : [0, 1] \rightarrow M$ be a minimizing geodesic such that $\alpha(0) = p^*, \alpha(1) = p$. Then there exists $u \in T_{p^*}M$ such that $\|u\| = d(p^*, p)$ and $\alpha(t) = \exp_{p^*}(tu)$ for each $t \in [0, 1]$. Let $p_\tau = \exp_{p^*}(\tau u)$. Applying Corollary 5.1 (with $l = 1$), we get

$$DX(p) = P_{\alpha,0,1}(DX(p^*) + \sum_{k \geq 2} \frac{k}{k!} D^k X(p^*) u^{k-1}) P_{\alpha,1,0}$$

and

$$DX(\alpha(\tau)) = DX(p_\tau) = P_{\alpha,0,\tau}(DX(p^*) + \sum_{k \geq 2} \frac{k}{k!} D^k X(p^*) (\tau u)^{k-1}) P_{\alpha,\tau,0}.$$

Hence,

$$\begin{aligned} & \|DX(p^*)^{-1}P_{\alpha,1,0}(DX(p) - P_{\alpha,\tau,1}DX(\alpha(\tau))P_{\alpha,1,\tau})\| \\ & \leq \sum_{k \geq 2} \frac{k}{k!} \|DX(p^*)^{-1}D^kX(p^*)\| (1 - \tau^{k-1}) \|u\|^{k-1} \\ & \leq \sum_{k \geq 2} k(1 - \tau^{k-1})\gamma(X, p^*)^{k-1}(d(p^*, p))^{k-1} \\ & = \frac{1}{(1 - d(p^*, p)\gamma(X, p^*))^2} - \frac{1}{(1 - \tau d(p^*, p)\gamma(X, p^*))^2} \\ & = \int_{\tau d(p^*, p)}^{d(p^*, p)} \frac{2\gamma}{(1 - \gamma u)^3} du. \end{aligned}$$

This completes the proof.

Suppose that L is defined by (5.5). Then, by (3.3), (3.5) and (4.1), we compute r_c, r_u and q to get that

$$r_c = \frac{K_{p^*} + 4 - \sqrt{K_{p^*}^2 + 8K_{p^*} + 8}}{4\gamma}, \quad r_u = \frac{1}{(K_{p^*} + 1)\gamma} \tag{5.7}$$

and

$$q = \frac{K_{p^*}\gamma d(p_0, p^*)}{1 - 4\gamma d(p_0, p^*) + 2(\gamma d(p_0, p^*))^2}. \tag{5.8}$$

Clearly, $r_c < r_u < \frac{1}{\gamma}$. Thus, by Lemma 5.2, $DX(p^*)^{-1}DX$ satisfies the radius Lipschitz condition with the L average in the ball $\mathbf{B}(p^*, r)$ with $r = \min(r_{p^*}, r_u)$. Therefore, applying Theorems 3.1 and 4.1, we get the following theorems and corollaries. Recall that X is analytic on M, p^* is a point of M such that $DX(p^*)^{-1}$ exists and $\gamma = \gamma(X, p^*)$.

Theorem 5.1. Suppose $X(p^*) = 0$. Let

$$r = \min(r_{p^*}, r_c),$$

where r_c is defined by (5.7). Then the Newton method (2.2) is well defined and convergent for each $p_0 \in \mathbf{B}(p^*, r)$, and (3.4) holds for $q < 1$ defined by (5.8).

Theorem 5.2. Suppose $X(p^*) = 0$. Let $r = \min(r_{p^*}, r_u)$, where r_u is given by (5.7). Then p^* is the unique solution of $X(p) = 0$ in $\mathbf{B}(p^*, r)$.

Corollary 5.2. Assume that M has non-negative sectional curvature. Suppose $X(p^*) = 0$. Let

$$r = \min\left(r_{p^*}, \frac{5 - \sqrt{17}}{4\gamma}\right).$$

Then the Newton method (2.2) is well defined and convergent for each $p_0 \in \mathbf{B}(p^*, r)$, and (3.4) holds for $q < 1$ defined by

$$q = \frac{\gamma d(p_0, p^*)}{1 - 4\gamma d(p_0, p^*) + 2(\gamma d(p_0, p^*))^2}. \tag{5.9}$$

Corollary 5.3. Assume that M has non-negative sectional curvature. Suppose $X(p^*) = 0$. Let $r = \min(r_{p^*}, \frac{1}{2\gamma})$. Then p^* is the unique solution of $X(p) = 0$ in $\mathbf{B}(p^*, r)$.

5.2 Determination of an approximation zero

To study the computational complexity of the Newton method for the nonlinear operators in the Banach spaces, Smale^[16] proposed the notion of an approximation zero for the Newton method in 1981. But it was found that it did not describe the property of quadratic convergence of the Newton method and was inconvenient for the application in the study of the computational complexity. Hence, he^[2] proposed two kinds of the modifications of the notion again in 1986. The second one, which was also presented and studied by Wang^[17], see also ref. [18], is regarded as the definition of the approximation zero in the recent works^[19,20]. The following definition is a generalization of this notion to vector fields on the Riemannian manifolds.

Definition 5.4. Suppose $X(p^*) = 0$. Then $p_0 \in M$ is called an approximation zero of the associated zero p^* if the Newton method (2.2) with the initial point p_0 is well defined and (3.4) holds for $q = \frac{1}{2}$.

By Theorem 5.1, we have the following result on the approximation zero of the associated zero p^* , which was obtained by Dedieu, Priouret and Malajovichsee^[13] with a different but more complex technique. Recall that X is analytic on M and p^* is a point of M such that $DX(p^*)^{-1}$ exists and $\gamma = \gamma(X, p^*)$.

Corollary 5.4. Suppose $X(p^*) = 0$. Let p_0 be a point of M such that

$$d(p_0, p^*) < r = \min \left(r_{p^*}, \frac{2 + K_{p^*} - \sqrt{K_{p^*}^2 + 4K_{p^*} + 2}}{2\gamma} \right). \tag{5.10}$$

Then p_0 is an approximation zero of the associated zero p^* .

Proof. Write

$$\delta = \frac{2 + K_{p^*} - \sqrt{K_{p^*}^2 + 4K_{p^*} + 2}}{2}.$$

Then p_0 lies in the ball $\mathbf{B}(p^*, \min\{r_{p^*}, r_c\})$ because $\frac{\delta}{\gamma} < r_c$, where r_c is defined by (5.7). Hence, by Theorem 5.1, the Newton method (2.2) is well defined for p_0 and (3.4) holds for q defined by (5.8). As

$$q = \frac{K_{p^*}\gamma d(p_0, p^*)}{1 - 4\gamma d(p_0, p^*) + 2(\gamma d(p_0, p^*))^2} \leq \frac{K_{p^*}\delta}{1 - 4\delta + 2\delta^2} = \frac{1}{2},$$

(3.4) holds for $q = \frac{1}{2}$ and hence p_0 is an approximation zero of the associated zero p^* and the proof is complete.

6 Conclusions

We have established the unified estimates of radii of convergence balls of the Newton method and the uniqueness balls of the zeroes of vector fields on the Riemannian manifolds. As remarked in the introduction, all the results obtained in the preceding sections are also true for mappings on the Riemannian manifolds. Applications to some classical cases including the Kantorovich’s type and Smale’s type are provided. Most part of the results obtained in the present paper are new. Some of them are the extensions of the

known results. As is known, the famous Kantorovich's theorem and the Smale's α -theory have been extended to the Newton method on the Riemannian manifolds respectively by Ferreira and Svaiter^[12] and Dedieu et al.^[13]. In view of Wang's idea in ref. [6], one natural extension is to provide a unified approach to the semi-local convergence by virtue of the similar Lipschitz condition with the L average, which indeed will be done in our next paper.

Acknowledgements This work was supported in part by the National Natural Science Foundation of China (Grant No. 10271025) and Program for New Century Excellent Talents in University.

References

1. Kantorovich, L. V., Akilov, G. P., *Functional Analysis*, Oxford: Pergamon, 1982.
2. Smale, S., Newton's method estimates from data at one point, in *The Merging of Disciplines: New Directions in Pure, Applied and Computational Mathematics* (eds. Ewing, R. Gross, K. Martin, C.), New York: Springer, 1986, 185–196.
3. Traub, J. F., Wozniakowski, H., Convergence and complexity of Newton iteration, *J. Assoc. Comput. Math.*, 1979, 29: 250–258.
4. Wang, X. H., The convergence on Newton's method, *Kexue Tongbao (A Special Issue of Mathematics, Physics and Chemistry, in Chinese)*, 1980, 25: 36–37.
5. Wang, X. H., Convergence of Newton's method and uniqueness of the solution of equations in Banach space, *IMA J. Numer. Anal.*, 2000, 20(1): 123–134.
6. Wang, X. H., Convergence of Newton's method and inverse function theorem in Banach space, *Mathematics of Computation.*, 1999, 225(68): 169–186.
7. Edelman, A., Arias, T. A., Smith, T., The geometry of algorithms with orthogonality constraints, *SIAM J. Matrix Anal. Appl.*, 1998, 20(2): 303–353.
8. Gabay, D., Minimizing a differentiable function over a differential manifold, *J. Optim. Theory Appl.*, 1982, 37(2): 177–219.
9. Smith, S. T., Optimization techniques on Riemannian manifolds, *Fields institute communications 3*, AMS., 1994, 113–146.
10. Smith, S. T., Geometric optimization method for adaptive filtering, Ph. D. thesis, Harvard University, Cambridge, MA, 1993.
11. Udriste, C., Convex functions and optimization methods on Riemannian manifolds, in *Mathematics and Its Applications*, Vol. 297, Dordrecht: Kluwer Academic, 1994.
12. Ferreira, O. P., Svaiter, B. F., Kantorovich's theorem on Newton's method in Riemannian manifolds, *Journal of Complexity.*, 2002, 18: 304–329.
13. Dedieu, J. P., Priouret, P., Malajovich, G., Newton's method on Riemannian manifolds: Covariant alpha theory, *IMA Journal of Numerical Analysis.*, 2003, 23: 395–419.
14. Chen, W. H., Li, X. X., *An Introduction to Riemannian Geometry*(in Chinese), Beijing: Beijing University Press, 2002.
15. DoCarmo, M. P., *Riemannian Geometry*, Boston: Birkhauser, 1992.
16. Smale, S., The fundamental theorem of algebra and complexity theory, *Bull. AMS*, 1981, 4: 1–36.
17. Wang, X. H., Xuan, X. H., Random polynomial space and computational complexity theory, *Science in China, Ser. A*, 1987, 30: 673–684.
18. Wang, X. H., Li, C., Local and global behaviors for algorithms of solving equations, *Chinese Science Bulletin*, 2001, 46(6): 441–448.
19. Smale, S., Complexity theory and numerical analysis, *Acta Numer.*, 1997, 6: 523–551.
20. Blum, L., Cucker, F., Shub, M., Smale, S., *Complexity and Real Computation*, New York: Springer-Verlag, 1997.