IMA Journal of Numerical Analysis (2006) **26**, 228−251
doi:10.1093/imanum/dri039
Advance Access publication on October 13, 2005

Newton's method on Riemannian manifolds: Smale's point estimate theory under the γ-condition

CHONG LI† AND JINHUA WANG‡
Department of Mathematics, Zhejiang University, Hangzhou 310027, People's Republic of China

[Received on 29 November 2004; revised on 11 August 2005]

The γ-conditions for vector fields on Riemannian manifolds are introduced. The γ-theory and the α-theory for Newton's method on Riemannian manifolds are established under the γ-conditions. Applications to analytic vector fields are provided and the results due to Dedieu *et al.* (2003, *IMA J. Numer. Anal.*, **23**, 395–419) are improved.

Keywords: Newton's method; Riemannian manifold; vector field; Smale's point estimate theory; the γ-condition.

1. Introduction

Newton's method and its variants are among the most efficient methods known for solving systems of non-linear equations when the functions involved are continuously differentiable. Besides its practical applications, Newton's method is also a powerful theoretical tool. Therefore, it has been studied and used extensively. One of the famous results on Newton's method is the well-known Kantorovich theorem (cf. Kantorovich & Akilov, 1982) which guarantees convergence of Newton's sequence to a solution under very mild conditions. Another important result concerning Newton's method is Smale's point estimate theory (cf. Blum *et al.*, 1997, Smale, 1981, 1986 and 1997).

Newton's method has been extended to finding numerically zeros of vector fields on Riemannian manifolds, see, e.g. Edelman *et al.*, 1998; Gabay, 1982; Smith, 1993, 1994; Udriste, 1994. Recent research has focused on extensions of the Kantorovich theorem and Smale's point estimate theory, see Ferreira & Svaiter, 2002; Dedieu *et al.*, 2003. Here we are particularly interested in the work due to Dedieu *et al.* (2003). Let X be an analytic vector field on an analytic Riemannian manifold M. Let $p \in M$ be such that $\mathrm{D}X(p)^{-1}$ exists, and define

$$\gamma(X, p) = \sup_{k\geqslant 2} \left\| \mathrm{D}X(p)^{-1} \frac{\mathrm{D}^k X(p)}{k!} \right\|_p^{\frac{1}{k-1}}.$$

Write

$$\beta(X, p) = \|\mathrm{D}X(p)^{-1} X(p)\| \quad \text{and} \quad \alpha(X, p) = \beta(X, p)\gamma(X, p).$$

†Email: cli@zju.edu.cn
‡Email: wjh@zjut.edu.cn

Let $\alpha_0 = 0.130716944\ldots$ denote the unique root of the equation $2u = \psi(u)^2$ in the interval $(0, 1-\frac{\sqrt{2}}{2})$, where $\psi(u) = 1 - 4u + 2u^2$ for each $u \in (0, 1 - \frac{\sqrt{2}}{2})$. Set

$$\sigma = \sum_{k\geqslant 0} \left(\frac{1}{2}\right)^{2^k-1} = 1.632843018\ldots$$

and

$$s_0 = \frac{1}{\sigma + \frac{(1-\sigma\alpha_0)^2}{\psi(\sigma\alpha_0)}\left(1 + \frac{\sigma}{1-\sigma\alpha_0}\right)} = 0.103621842\ldots.$$

Then the main results in Dedieu *et al.* (2003) are as follows. For the definition of the geometrical constant K_{p^*} and of other undefined notations in the sequel, we refer to Dedieu *et al.* (2003), see also Sections 3 and 6.

THEOREM 1.1 Suppose that $X(p^*) = 0$ and let $p_0 \in M$. If

$$d(p_0, p^*) \leqslant \min\left\{\mathbf{r}_{p^*}, \frac{2 + K_{p^*} - \sqrt{K_{p^*}^2 + 4K_{p^*} + 2}}{2\gamma(X, p^*)}\right\},$$

then Newton's method with the initial point p_0 is well-defined for all $n \geqslant 0$, and

$$d(p_n, p^*) \leqslant \left(\frac{1}{2}\right)^{2^n-1} d(p_0, p^*), \quad n = 0, 1, 2, \ldots.$$

THEOREM 1.2 Let $p_0 \in M$ be such that

$$\beta(X, p_0) \leqslant s_0\mathbf{r}_{p_0} \quad \text{and} \quad \alpha(X, p_0) < \alpha_0. \tag{1.1}$$

Then Newton's method with the initial point p_0 is well-defined for all $n \geqslant 0$ and converges to a zero p^* of X. Moreover,

$$d(p_{n+1}, p_n) \leqslant \left(\frac{1}{2}\right)^{2^n-1} \beta(X, p_0).$$

THEOREM 1.3 Suppose that $X(p^*) = 0$, and let $p_0 \in M$. If

$$d(p_0, p^*) < \min\left\{\hat{s}_0\mathbf{r}_{p^*}, \frac{\hat{u}_0}{\gamma(X, p^*)}\right\}, \tag{1.2}$$

where $\hat{u}_0 = 0.069778332\ldots$ is the smallest positive root of the equation $\frac{\hat{u}_0}{\psi(\hat{u}_0)^2} = \alpha_0$ and $\hat{s}_0 = \frac{s_0}{s_0 + \frac{1-\hat{u}_0}{\psi(\hat{u}_0)}} = 0.075262346\ldots$, then Newton's method with the initial point p_0 is well-defined for all $n \geqslant 0$, and

$$d(p_n, p^*) \leqslant \sigma\left(\frac{1}{2}\right)^{2^n-1} d(p_1, p_0).$$

The γ-conditions for non-linear operators in Banach spaces were first introduced and explored by Wang & Han (1997) for the study of Smale's point estimate theory. The purpose of the present paper is to extend the notion of γ-conditions to the case of vector fields on Riemannian manifolds and to establish the γ-theory and the α-theory of Newton's method on Riemannian manifolds under the γ-conditions. In particular, when the results obtained in the present paper are applied to the special case when the vector field X is analytic, Theorem 1.1 becomes a direct consequence, while Theorems 1.2 and 1.3 are improved in such a way that the criteria (1.1) and (1.2) in Theorems 1.2 and 1.3 are, respectively, replaced by the weaker conditions (1.3) and (1.4) below:

$$\beta \leqslant (2-\sqrt{2})\mathbf{r}_{p_0} \quad \text{and} \quad \alpha = \beta\gamma \leqslant \frac{13-3\sqrt{17}}{4} \approx 0.157671 \tag{1.3}$$

and

$$d(p_0, p^*) < \min\left\{t_0\mathbf{r}_{p^*}, \frac{\bar{u}_0}{\gamma(X, p^*)}\right\}, \tag{1.4}$$

where $\bar{u}_0 = 0.0776121\ldots$ is the smallest positive root of the equation $\frac{\bar{u}_0}{\psi(\bar{u}_0)^2} = \frac{13-3\sqrt{17}}{4}$ while $t_0 = 0.305332\ldots$.

2. Notions and preliminaries

We begin with some basic notions and notations. Most of them are standard, see, e.g. Boothby, 1986; DoCarmo, 1992; Lang, 1995. Let M be a real complete m-dimensional Riemannian manifold. Let $p \in M$ and let T_pM denote the tangent space at p to M. Let $\langle\cdot,\cdot\rangle$ be the scalar product on T_pM with the associated norm $\|\cdot\|_p$, where the subscript p is sometimes omitted. For any two distinct elements $p, q \in M$, let c: $[0, 1] \to M$ be a piecewise smooth curve connecting p and q. Then the arc-length of c is defined by $l(c) := \int_0^1 \|c'(t)\|\,\mathrm{d}t$, and the Riemannian distance from p to q by $d(p, q) := \inf_c l(c)$, where the infimum is taken over all piecewise smooth curves c: $[0, 1] \to M$ connecting p and q. Thus, (M, d) is a complete metric space by the Hopf–Rinow theorem (cf. Boothby, 1986; DoCarmo, 1992; Lang, 1995).

For a finite-dimensional 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, i.e.

$$\mathbf{B}_Z(p, r) = \{q \in Z\colon d(p, q) < r\},$$

$$\overline{\mathbf{B}_Z(p, r)} = \{q \in Z\colon d(p, q) \leqslant r\}.$$

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.

Noting that M is complete, the exponential map at p, i.e. $\exp_p\colon T_pM \to M$, is well-defined on T_pM. Furthermore, the radius of injectivity of the exponential map at p is denoted by $\mathbf{r}_p$. Thus, $\exp_p$ is a one-to-one mapping from $\mathbf{B}_{T_pM}(0, \mathbf{r}_p)$ to $\mathbf{B}(p, \mathbf{r}_p)$. The following proposition gives the relationship of the radii $\mathbf{r}_p$ and $\mathbf{r}_q$, see Dedieu *et al.*, 2003, Lemma 4.4.

PROPOSITION 2.1 Let $p, q \in M$. Then

$$\mathbf{r}_p - d(p, q) \leqslant \mathbf{r}_q.$$

Recall that a geodesic in M connecting p and q is called a minimizing geodesic if its arc-length equals its Riemannian distance between p and q. Note that there is at least one minimizing geodesic connecting p and q. In particular, the curve c: $[0, 1] \to M$ is a minimizing geodesic connecting p and q if, and only if, there exists a vector $v \in T_pM$ such that $\|v\| = d(p, q)$, $q = \exp_p(v)$ and $c(t) = \exp_p(tv)$ for each $t \in [0, 1]$.

Let ∇ denote 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

$\mathrm{D}X(p)\colon T_pM \to T_pM$ by

$$\mathrm{D}X(p)(w) = \nabla_Y X(p) \quad \forall\, w \in T_pM,$$

where Y is a vector field satisfying $Y(p) = w$. Then the value $\mathrm{D}X(p)(w)$ of $\mathrm{D}X(p)$ at w depends only on the tangent vector $w = Y(p) \in T_pM$ since ∇ is tensorial in Y. Let $c\colon \mathbb{R} \to M$ be a C^∞ curve and let $P_{c,\cdot,\cdot}$ denote the parallel transport along c, which is defined by

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

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

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

In particular, we write $P_{q,p}$ for $P_{c,q,p}$ in the case when c is a minimizing geodesic connecting p and q.

Let X be a C^1 vector field on M and let $p_0 \in M$. Following Ferreira & Svaiter (2002), Newton's method with the initial point p_0 for X is defined as follows.

$$p_{n+1} = \exp_{p_n}(-\mathrm{D}X(p_n)^{-1}X(p_n)), \quad n = 0, 1, 2, \ldots. \tag{2.1}$$

The γ-conditions for operators in Banach spaces were first presented by Wang & Han (1997) for the study of Smale's point estimate theory. In the following, we extend this notion to the case of vector fields on a Riemannian manifold M. Let k be a positive integer. We first define the notion of kth covariant derivatives.

DEFINITION 2.1 Let $\{Y_1, \ldots, Y_k\}$ be a finite sequence of vector fields on M. Then, the kth covariant derivative of X with respect to $\{Y_1, \ldots, Y_k\}$ is denoted by $\nabla^k_{\{Y_i\}_{i=1}^k} X$ and is defined inductively by

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

DEFINITION 2.2 Let $p \in M$ and $(v_1, \ldots, v_k) \in (T_pM)^k$. Let $\{Y_1, \ldots, Y_k\}$ be a finite sequence of vector fields on M such that $Y_i(p) = v_i$ for each $i = 1, \ldots, k$. Then, the value of the kth covariant derivative of X with respect to $\{Y_1, \ldots, Y_k\}$ at p is denoted by

$$\mathrm{D}^k X(p) v_1 v_2 \cdots v_k = \nabla^k_{\{Y_i\}_{i=1}^k} X(p).$$

Note that $\mathrm{D}^k X(p) v_1 v_2 \cdots v_k$ only depends on the k-tuple of vectors $(v_1, \ldots, v_k)$ since the covariant derivative is tensorial in each Y_i.

Clearly, by Definition 2.2, the kth covariant derivative $\mathrm{D}^k X(p)$ at a point p is a k-multilinear map from $(T_pM)^k$ to T_pM. We define the norm of $\mathrm{D}^k X(p)$ by

$$\|\mathrm{D}^k X(p)\|_p = \sup \|\mathrm{D}^k X(p) v_1 v_2 \cdots v_k\|_p, \tag{2.2}$$

where the supremum is taken over all k-tuples of vectors $(v_1, \ldots, v_k) \in (T_pM)^k$ each with $\|v_j\| = 1$.

Let $r > 0$ and $\gamma > 0$ be such that $\gamma r \leqslant 1$. Also let k be a positive integer. Throughout the paper, we always assume that X is a C^2 vector field on M.

DEFINITION 2.3 Let $q_0 \in M$ be such that $\mathrm{D}X(q_0)^{-1}$ exists. X is said to satisfy the k-piece γ-condition at q_0 in $\mathbf{B}(q_0, r)$, if

$$\left\|\mathrm{D}X(q_0)^{-1} P_{q_0,q_1} \circ P_{q_1,q_2} \circ \cdots \circ P_{q_{k-1},q_k} \mathrm{D}^2 X(q_k)\right\| \leqslant \frac{2\gamma}{\left(1 - \gamma \sum_{i=1}^{k} d(q_{i-1}, q_i)\right)^3} \tag{2.3}$$

holds for any k points $q_1, q_2, \ldots, q_k \in \mathbf{B}(q_0, r)$ satisfying $\sum_{i=1}^{k} d(q_{i-1}, q_i) < r$.

REMARK 2.1

(i) The $(k+1)$-piece γ-condition at q_0 implies the k-piece γ-condition at q_0 in $\mathbf{B}(q_0, r)$.

(ii) Let b denote the bound of $\left\|\mathrm{D}X(q_0)^{-1} P_{q_0,q_1} \circ P_{q_1,q_2} \circ \cdots \circ P_{q_{k-1},q} \mathrm{D}^2 X(q)\right\|$ on $\mathbf{B}(q_0, r)$. Then it is easy to see that X satisfies the k-piece $\frac{b}{2}$-condition at q_0 in $\mathbf{B}(q_0, r)$. It follows that $\gamma_0 \leqslant \frac{b}{2}$ if γ_0 is the minimum of $\gamma > 0$ such that X satisfies the k-piece γ-condition at q_0 in $\mathbf{B}(q_0, r)$.

The following two lemmas will be used later. The first one is stated in Ferreira & Svaiter (2002, p. 308) while the second one is its consequence.

LEMMA 2.1 Let c: $[0, 1] \to M$ be a geodesic and Z a C^1 vector field on M. Then

$$P_{c,c(0),c(t)} Z(c(t)) = Z(c(0)) + \int_0^t P_{c,c(0),c(s)} (\mathrm{D}Z(c(s))c'(s))\,\mathrm{d}s.$$

LEMMA 2.2 Let c: $[0, 1] \to M$ be a geodesic and Y a C^∞ vector field on M. Then

$$P_{c,c(0),c(t)} \mathrm{D}X(c(t))Y(c(t)) = \mathrm{D}X(c(0))Y(c(0)) + \int_0^t P_{c,c(0),c(s)} (\mathrm{D}^2 X(c(s))Y(c(s))c'(s))\,\mathrm{d}s. \tag{2.4}$$

In particular,

$$P_{c,c(0),c(t)} \mathrm{D}X(c(t))c'(t) = \mathrm{D}X(c(0))c'(0) + \int_0^t P_{c,c(0),c(s)} (\mathrm{D}^2 X(c(s))(c'(s))^2)\,\mathrm{d}s. \tag{2.5}$$

Proof. Clearly, (2.5) is a direct consequence of (2.4). Thus, we only need to show (2.4). Let $Z = \nabla_Y X$. Then Z is a C^1 vector field on M and Lemma 2.1 is applicable. Hence,

$$P_{c,c(0),c(t)} Z(c(t)) = Z(c(0)) + \int_0^t P_{c,c(0),c(s)} (\mathrm{D}Z(c(s))c'(s))\,\mathrm{d}s.$$

Since

$$\mathrm{D}Z(c(s))c'(s) = \mathrm{D}(\mathrm{D}X(c(s))Y(c(s)))c'(s) = \mathrm{D}^2 X(c(s))Y(c(s))c'(s),$$

(2.4) follows and the proof is complete. □

Finally, we state a lemma, which will play a key role in this paper. This lemma is true for the general case although it is stated and proved for the special case when $k \leqslant 2$. For simplicity, we use the function ψ defined by

$$\psi(u) := 1 - 4u + 2u^2, \quad u \in \left[0, 1 - \frac{\sqrt{2}}{2}\right). \tag{2.6}$$

Note that ψ is strictly monotonic decreasing on $[0, 1 - \frac{\sqrt{2}}{2})$.

LEMMA 2.3 Let $r < \frac{1-\frac{\sqrt{2}}{2}}{\gamma}$ and let $k \leqslant 2$. Let $q_0 \in M$ be such that $\mathrm{D}X(q_0)^{-1}$ exists. Suppose that X satisfies the k-piece γ-condition at q_0 in $\mathbf{B}(q_0, r)$. Then, for each point $q \in \mathbf{B}(q_0, r)$, $\mathrm{D}X(q)^{-1}$ exists, and for any $k-1$ points $q_1, q_2, \ldots, q_{k-1}$ in $\mathbf{B}(q_0, r)$ satisfying $\sum_{i=0}^{k-1} d(q_i, q_{i+1}) < r$,

$$\left\| \mathrm{D}X(q)^{-1} P_{q,q_{m-1}} \circ \cdots \circ P_{q_2,q_1} \circ P_{q_1,q_0} \mathrm{D}X(q_0) \right\| \leqslant \frac{\left(1-\gamma \sum_{i=0}^{k-1} d(q_i, q_{i+1})\right)^2}{\psi\left(\gamma \sum_{i=0}^{k-1} d(q_i, q_{i+1})\right)}, \tag{2.7}$$

where $q_k = q$.

Proof. We only prove the lemma for the case when $k = 2$ because the proof for the case when $k = 1$ (and for the general case) is similar. By the Banach lemma, to complete the proof, it is sufficient to show that

$$\begin{aligned} &\left\| \mathrm{D}X(q_0)^{-1} P_{q_0,q_1} \circ P_{q_1,q} \mathrm{D}X(q) P_{q,q_1} \circ P_{q_1,q_0} - \mathbf{I}_{T_{q_0}M} \right\| \\ &\quad \leqslant -1 + \frac{1}{(1-\gamma(d(q_0,q_1)+d(q_1,q)))^2} < 1 \end{aligned} \tag{2.8}$$

because P_{q_0,q_1} and $P_{q_1,q}$ are isometries, where $\mathbf{I}_{T_{q_0}M}$ is the identity on $T_{q_0}M$. To verify (2.8), let $v \in T_{q_0}M$. Let $v_1 \in T_{q_0}M$ and $v_2 \in T_{q_1}M$ be such that the curve $c_1(t) := \exp_{q_0}(tv_1)$, $t \in [0,1]$, is a minimizing geodesic connecting q_0 and q_1 and that the curve $c_2(t) := \exp_{q_1}(tv_2)$, $t \in [0,1]$, is a minimizing geodesic connecting q_1 and q. Note that there exist vector fields Y_1 and Y_2 such that $Y_1(c_1(0)) = v$, $\mathrm{D}_{c_1'(t)} Y_1(c_1(t)) = 0$, $Y_2(c_2(0)) = P_{c_1,q_1,q_0} v$ and $\mathrm{D}_{c_2'(t)} Y_2(c_2(t)) = 0$. Then we apply Lemma 2.2 to conclude that

$$\begin{aligned} &\mathrm{D}X(q_0)^{-1} \left(P_{c_1,q_0,q_1} \mathrm{D}X(q_1) Y_1(c_1(1)) - \mathrm{D}X(q_0) Y_1(c_1(0))\right) \\ &\quad = \mathrm{D}X(q_0)^{-1} \int_0^1 P_{c_1,q_0,c_1(s)} \mathrm{D}^2 X(c_1(s)) Y_1(c_1(s)) c_1'(s)\,\mathrm{d}s \end{aligned} \tag{2.9}$$

and

$$\begin{aligned} &\mathrm{D}X(q_0)^{-1} P_{c_1,q_0,q_1} \left(P_{c_2,q_1,q} \mathrm{D}X(q) Y_2(c_2(1)) - \mathrm{D}X(p) Y_2(c_2(0))\right) \\ &\quad = \mathrm{D}X(q_0)^{-1} P_{c_1,q_0,q_1} \int_0^1 P_{c_2,q_1,c_2(s)} \mathrm{D}^2 X(c_2(s)) Y_2(c_2(s)) c_2'(s)\,\mathrm{d}s. \end{aligned} \tag{2.10}$$

Hence, in view of (2.3) (with $k = 1, 2$, respectively), we have that

$$\begin{aligned} &\left\| \mathrm{D}X(q_0)^{-1} \left(P_{c_1,q_0,q_1} \mathrm{D}X(q_1) Y_1(c_1(1)) - \mathrm{D}X(q_0) Y_1(c_1(0))\right) \right\| \\ &\quad \leqslant \int_0^1 \left\| \mathrm{D}X(q_0)^{-1} P_{c_1,q_0,c_1(s)} \mathrm{D}^2 X(c_1(s)) \right\| \|Y_1(c_1(s))\| \|c_1'(s)\|\,\mathrm{d}s \\ &\quad = \int_0^1 \frac{2\gamma}{(1-\gamma s\|v_1\|)^3} \|v\| \|v_1\|\,\mathrm{d}s \end{aligned} \tag{2.11}$$

and

$$
\begin{aligned}
&\left\| \mathrm{D}X(q_0)^{-1} P_{c_1,q_0,q_1} \left(P_{c_2,q_1,q} \mathrm{D}X(q) Y_2(c_2(1)) - \mathrm{D}X(p) Y_2(c_2(0)) \right) \right\| \\
&\quad \leqslant \int_0^1 \left\| \mathrm{D}X(q_0)^{-1} P_{c_1,q_0,q_1} \circ P_{c_2,q_1,c_2(s)} \mathrm{D}^2 X(c_2(s)) \right\| \, \|Y_2(c_2(s))\| \|c_2'(s)\| \, \mathrm{d}s \\
&\quad \leqslant \int_0^1 \frac{2\gamma}{(1-\gamma(\|v_1\| + s\|v_2\|))^3} \|v\| \|v_2\| \, \mathrm{d}s.
\end{aligned} \tag{2.12}
$$

Since

$$
\begin{aligned}
&\left(\mathrm{D}X(q_0)^{-1} P_{q_0,q_1} \circ P_{q_1,q} \mathrm{D}X(q) P_{q,q_1} \circ P_{q_1,q_0} - \mathbf{I}_{T_{q_0}M} \right) v \\
&\quad = \mathrm{D}X(q_0)^{-1} P_{c_1,q_0,q_1} \left[P_{c_2,q_1,q} \mathrm{D}X(q) Y_2(c_2(1)) - \mathrm{D}X(q_1) Y_2(c_2(0)) \right] \\
&\qquad + \mathrm{D}X(q_0)^{-1} \left[P_{c_1,q_0,q_1} \mathrm{D}X(q_1) Y_1(c_1(1)) - \mathrm{D}X(q_0) Y_1(c_1(0)) \right],
\end{aligned} \tag{2.13}
$$

it follows from (2.12) and (2.11) that

$$
\begin{aligned}
&\left\| \left(\mathrm{D}X(q_0)^{-1} P_{q_0,q_1} \circ P_{q_1,q} \mathrm{D}X(q) P_{q,q_1} \circ P_{q_1,q_0} - \mathbf{I}_{T_{q_0}M} \right) v \right\| \\
&\quad \leqslant \int_0^1 \frac{2\gamma}{(1-\gamma(\|v_1\| + s\|v_2\|))^3} \|v\| \|v_2\| \, \mathrm{d}s + \int_0^1 \frac{2\gamma}{(1-\gamma s\|v_1\|)^3} \|v\| \|v_1\| \, \mathrm{d}s \\
&\quad = \left(-1 + \frac{1}{(1-\gamma(d(q_0,q_1) + d(q_1,q)))^2} \right) \|v\|,
\end{aligned} \tag{2.14}
$$

where the last equality holds because $\|v_1\| = d(q_0, q_1)$ and $\|v_2\| = d(q_1, q)$. As $v \in T_{q_0}M$ is arbitrary, (2.8) follows. □

3. Generalized γ-theory

Recall that X is a C^2 vector field on M. Let $p^* \in M$ be such that $\mathrm{D}X(p^*)^{-1}$ exists. The approach for the generalized γ-theory in this section depends on the geometrical number K_{p^*} while the approach independent of K_{p^*} will be considered in Section 5. The K_{p^*} is related to the sectional curvature at $p^* \in M$ and is defined by

$$
K_{p^*} = \sup \frac{d(\exp_q(w), \exp_q(v))}{\|w - v\|_q}, \tag{3.1}
$$

where the supremum is taken over all $q \in \mathbf{B}(p^*, \mathbf{r}_{p^*})$, and $v, w - v \in \overline{\mathbf{B}_{T_qM}(0, \mathbf{r}_{p^*})}$ with $w \neq v$, see Dedieu *et al.*, 2003.

REMARK 3.1

(1) K_{p^*} measures how fast the geodesics spread apart in M. In particular, if $w = 0$ or more generally if w and v are on the same line through 0, then

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

This means that $K_{p^*} \geqslant 1$.

(2) In the case when M has non-negative sectional curvature, the geodesics spread apart less than the rays (cf. Dedieu *et al.*, 2003) so that

$$
d(\exp_q(w), \exp_q(v)) \leqslant \|w - v\|_q
$$

and consequently $K_{p^*} = 1$. Examples of manifolds with non-negative sectional curvature are given in Dedieu *et al.* (2003).

The main theorem of this section gives an estimate of the radius of the convergence ball around the zero of X for Newton's method. Recall that ψ is defined by $\psi(u) = 1 - 4u + 2u^2$ for each $u \in [0, 1 - \frac{\sqrt{2}}{2})$.

THEOREM 3.1 Let

$$r_c = \frac{K_{p^*} + 4 - \sqrt{K_{p^*}^2 + 8K_{p^*} + 8}}{4\gamma} \quad \text{and} \quad r = \min\{\mathbf{r}_{p^*}, r_c\}. \tag{3.2}$$

Suppose that $X(p^*) = 0$ and that X satisfies the one-piece γ-condition at p^* in $\mathbf{B}(p^*, r)$. If $d(p_0, p^*) < r$, then Newton's method (2.1) with the initial point p_0 is well-defined , and

$$d(p_n, p^*) \leqslant \lambda^{2^n - 1} d(p_0, p^*), \tag{3.3}$$

for $n = 0, 1, 2, \ldots$, where

$$\lambda = \frac{K_{p^*} \gamma\, d(p_0, p^*)}{\psi(\gamma\, d(p_0, p^*))} < 1. \tag{3.4}$$

Proof. By (3.2),

$$\gamma\, r_c = \frac{K_{p^*} + 4 - \sqrt{K_{p^*}^2 + 8K_{p^*} + 8}}{4} < \frac{K_{p^*} + 4 - (K_{p^*} + 3)}{4} < 1 - \frac{\sqrt{2}}{2}. \tag{3.5}$$

Then $\gamma\, d(p_0, p^*) < \gamma\, r_c < 1 - \frac{\sqrt{2}}{2}$. Since the function ψ is strictly monotonic decreasing on $[0, 1 - \frac{\sqrt{2}}{2})$, it follows that

$$\lambda = \frac{K_{p^*} \gamma\, d(p_0, p^*)}{\psi(\gamma\, d(p_0, p^*))} < \frac{K_{p^*} \gamma\, r_c}{\psi(\gamma\, r_c)} = 1. \tag{3.6}$$

Below we will show that (3.3) holds for each $n = 0, 1, \ldots$ by induction. Clearly, it is trivial in the case when $n = 0$. Now assume that (3.3) holds for n. Note that, for each $n = 0, 1, \ldots$, (3.3) implies $p_n \in \mathbf{B}(p^*, r)$. Then, by (3.5), we have

$$d(p_n, p^*) < r_c < \frac{1 - \frac{\sqrt{2}}{2}}{\gamma}.$$

Hence, Lemma 2.3 is applicable (with $k = 1$). It follows that $\mathrm{D}X(p_n)^{-1}$ exists and

$$\left\| \mathrm{D}X(p_n)^{-1} P_{p_n, p^*} \mathrm{D}X(p^*) \right\| \leqslant \frac{(1 - \gamma\, d(p_n, p^*))^2}{\psi(\gamma\, d(p_n, p^*))}. \tag{3.7}$$

Thus, p_{n+1} is well-defined. Consequently, to complete the proof, it remains to verify that (3.3) holds for $n + 1$. To do this, let $v \in T_{p^*}M$ be such that $p_n = \exp_{p^*}(v)$ and $\|v\| = d(p_n, p^*)$. We claim that

$$\left\| -\mathrm{D}X(p_n)^{-1} X(p_n) - \left(-P_{p_n, p^*} v\right) \right\| \leqslant \frac{\gamma\, d(p_n, p^*)^2}{\psi(\gamma\, d(p_n, p^*))}. \tag{3.8}$$

In fact, since the curve $c(t) := \exp_{p^*}(tv)$, $t \in [0, 1]$, is the minimizing geodesic connecting p^* and p_n, by Lemma 2.1, we have that

$$P_{c,p^*,p_n} X(p_n) - X(p^*) = \int_0^1 P_{c,p^*,c(\tau)} \mathrm{D}X(c(\tau))c'(\tau)\,\mathrm{d}\tau. \tag{3.9}$$

Also, by Lemma 2.2,

$$\mathrm{D}X(p_n)P_{c,p_n,p^*}v - P_{c,p_n,c(\tau)}\mathrm{D}X(c(\tau))c'(\tau) = \int_\tau^1 P_{c,p_n,c(s)}\mathrm{D}^2X(c(s))(c'(s))^2\,\mathrm{d}s. \tag{3.10}$$

Hence, the two equalities above imply that

$$\begin{aligned} &-\mathrm{D}X(p_n)^{-1}X(p_n) - \left(-P_{p_n,p^*}v\right)\\ &= -\mathrm{D}X(p_n)^{-1}P_{c,p_n,p^*}\left(P_{c,p^*,p_n}X(p_n) - X(p^*)\right) + P_{c,p_n,p^*}v\\ &= \mathrm{D}X(p_n)^{-1}\int_0^1 \left(\mathrm{D}X(p_n)P_{c,p_n,p^*}v - P_{c,p_n,c(\tau)}\mathrm{D}X(c(\tau))c'(\tau)\right)\mathrm{d}\tau\\ &= \mathrm{D}X(p_n)^{-1}P_{c,p_n,p^*}\int_0^1\int_\tau^1 P_{c,p^*,c(s)}\mathrm{D}^2X(c(s))(c'(s))^2\,\mathrm{d}s\,\mathrm{d}\tau. \end{aligned} \tag{3.11}$$

Consequently, by (3.11), (3.7) and (2.3) (with $k = 1$), we obtain that

$$\begin{aligned} &\left\|-\mathrm{D}X(p_n)^{-1}X(p_n) - \left(-P_{p_n,p^*}v\right)\right\|\\ &\leqslant \left\|\mathrm{D}X(p_n)^{-1}P_{c,p_n,p^*}\mathrm{D}X(p^*)\right\|\int_0^1\int_\tau^1 \|\mathrm{D}X(p^*)^{-1}P_{c,p^*,c(s)}\mathrm{D}^2X(c(s))(c'(s))^2\|\,\mathrm{d}s\,\mathrm{d}\tau\\ &\leqslant \frac{(1-\gamma\,d(p_n,p^*))^2}{\psi(\gamma\,d(p_n,p^*))}\int_0^1\int_\tau^1 \frac{2\gamma}{(1-\gamma\,s d(p_n,p^*))^3}d(p_n,p^*)^2\,\mathrm{d}s\,\mathrm{d}\tau\\ &= \frac{\gamma\,d(p_n,p^*)^2}{\psi(\gamma\,d(p_n,p^*))}. \end{aligned} \tag{3.12}$$

This shows that (3.8) holds and hence

$$\left\|-\mathrm{D}X(p_n)^{-1}X(p_n) - \left(-P_{p_n,p^*}v\right)\right\| \leqslant \lambda d(p_n,p^*) \leqslant \mathbf{r}_{p^*}. \tag{3.13}$$

On the other hand, since

$$\left\|-P_{c,p_n,p^*}v\right\| = \|v\| = d(p_n,p^*) < \mathbf{r}_{p^*}, \tag{3.14}$$

in view of the definition of K_{p^*} and (3.8), one gets that

$$d\left(\exp_{p_n}(-\mathrm{D}X(p_n)^{-1}X(p_n)), \exp_{p_n}\left(-P_{c,p_n,p^*}v\right)\right) \leqslant \frac{K_{p^*}\gamma\,d(p_n,p^*)^2}{\psi(\gamma\,d(p_n,p^*))}. \tag{3.15}$$

As $p_{n+1} = \exp_{p_n}(-\mathrm{D}X(p_n)^{-1}X(p_n))$ and $p^* = \exp_{p_n}\left(-P_{c,p_n,p^*}v\right)$, (3.15) means that

$$d(p_{n+1}, p^*) \leqslant \frac{K_{p^*}\gamma\, d(p_n, p^*)^2}{\psi(\gamma\, d(p_n, p^*))} \leqslant \frac{K_{p^*}\gamma\, \left(\lambda^{2^n-1}\right)^2 d(p_0, p^*)^2}{\psi(\gamma\, d(p_0, p^*))} = \lambda^{2^{n+1}-1} d(p_0, p^*).$$

Therefore, (3.3) holds for $n+1$. □

4. Generalized α-theory

The majorizing function h, which is due to Wang (1999) and Wang & Han (1990), will play a key role in this section. Let $\beta > 0$ and $\gamma > 0$. Define

$$h(t) = \beta - t + \frac{\gamma t^2}{1-\gamma t}, \quad \text{for each } 0 \leqslant t < \frac{1}{\gamma}. \tag{4.1}$$

Let $\{t_n\}$ denote the sequence generated by Newton's method with the initial value $t_0 = 0$ for h, i.e.

$$t_{n+1} = t_n - h'(t_n)^{-1}h(t_n), \quad \text{for each } n = 0, 1, \ldots. \tag{4.2}$$

Then we have the following proposition which was proved in Wang (1999) and Wang & Han (1990).

PROPOSITION 4.1 Suppose that $\alpha = \gamma\beta \leqslant 3 - 2\sqrt{2}$. Then the zeros of h are

$$r_1 = \frac{1+\alpha-\sqrt{(1+\alpha)^2-8\alpha}}{4\gamma}, \quad r_2 = \frac{1+\alpha+\sqrt{(1+\alpha)^2-8\alpha}}{4\gamma} \tag{4.3}$$

and they satisfy

$$\beta \leqslant r_1 \leqslant \left(1+\frac{1}{\sqrt{2}}\right)\beta \leqslant \left(1-\frac{1}{\sqrt{2}}\right)\frac{1}{\gamma} \leqslant r_2 \leqslant \frac{1}{2\gamma}. \tag{4.4}$$

Moreover,

$$t_n = \frac{1-\mu^{2^n-1}}{1-\mu^{2^n-1}\eta} r_1 \tag{4.5}$$

and

$$t_{n+1} - t_n = \frac{\left(1-\mu^{2^n}\right)\sqrt{(1+\alpha)^2-8\alpha}}{2\alpha\left(1-\eta\mu^{2^n-1}\right)\left(1-\eta\mu^{2^{n+1}-1}\right)}\eta\mu^{2^n-1}\beta, \quad n = 0, 1, \ldots, \tag{4.6}$$

where

$$\mu = \frac{1-\alpha-\sqrt{(1+\alpha)^2-8\alpha}}{1-\alpha+\sqrt{(1+\alpha)^2-8\alpha}} \tag{4.7}$$

and

$$\eta = \frac{1+\alpha-\sqrt{(1+\alpha)^2-8\alpha}}{1+\alpha+\sqrt{(1+\alpha)^2-8\alpha}}. \tag{4.8}$$

Lemma 4.1 was shown in Wang (1999) and Wang & Han (1990). However, here we give a direct and simpler proof of this lemma.

LEMMA 4.1 Suppose that $\alpha < 3 - 2\sqrt{2}$. Then

$$\frac{\left(1-\mu^{2^n}\right)\sqrt{(1+\alpha)^2-8\alpha}}{2\alpha\left(1-\eta\mu^{2^n-1}\right)\left(1-\eta\mu^{2^{n+1}-1}\right)}\eta \leqslant 1, \quad n = 0, 1, \ldots. \tag{4.9}$$

Proof. Let

$$a_n = \frac{\left(1-\mu^{2^n}\right)\sqrt{(1+\alpha)^2-8\alpha}}{2\alpha\left(1-\eta\mu^{2^n-1}\right)\left(1-\eta\mu^{2^{n+1}-1}\right)}\eta.$$

Since $0 < \eta < 1$ and $\eta\mu^{-1} > 1$, one has that

$$\frac{a_n}{a_{n-1}} = \frac{1-\mu^{2^n}}{1-\eta\mu^{2^{n+1}-1}}\,\frac{1-(\eta\mu^{-1})\mu^{2^{n-1}}}{1-\mu^{2^{n-1}}} \leqslant \frac{1-\mu^{2^n}}{1-\mu^{2^{n+1}-1}}\,\frac{1-\mu^{2^{n-1}}}{1-\mu^{2^{n-1}}} \leqslant 1.$$

Hence,

$$a_n \leqslant a_{n-1} \leqslant \cdots \leqslant a_0 = \frac{t_1 - t_0}{\beta} = 1$$

and (4.9) follows. □

Recall that X is a C^2 vector field. In the remainder of this section, let $p_0 \in M$ be such that $\mathrm{D}X(p_0)^{-1}$ exists, and define

$$\beta = \|\mathrm{D}X(p_0)^{-1}X(p_0)\|, \quad \alpha = \gamma\beta.$$

THEOREM 4.1 Let

$$\beta \leqslant (2-\sqrt{2})\mathbf{r}_{p_0} \quad \text{and} \quad \alpha = \beta\gamma \leqslant 3 - 2\sqrt{2}.$$

Suppose that X satisfies the two-piece γ-condition at p_0 in $\mathbf{B}(p_0, r_1)$. Then Newton's method (2.1) with the initial point p_0 is well-defined and the generated sequence $\{p_n\}$ converges to a zero p^* of X in $\overline{\mathbf{B}(p_0, r_1)}$. Moreover,

$$d(p_{n+1}, p_n) \leqslant \frac{\left(1-\mu^{2^n}\right)\sqrt{(1+\alpha)^2-8\alpha}}{2\alpha\left(1-\eta\mu^{2^n-1}\right)\left(1-\eta\mu^{2^{n+1}-1}\right)}\eta\mu^{2^n-1}d(p_1, p_0), \tag{4.10}$$

for all $n = 0, 1, 2, \ldots$, where μ and η are given by (4.7) and (4.8), respectively.

Proof. Recall from (2.1) that

$$p_{n+1} = \exp_{p_n}(-DX(p_n)^{-1}X(p_n)), \quad n = 0, 1, \ldots. \tag{4.11}$$

Let

$$v_n = -DX(p_n)^{-1}X(p_n). \tag{4.12}$$

We will use induction to prove that

$$d(p_{n+1}, p_n) = \|v_n\| \leqslant t_{n+1} - t_n \tag{4.13}$$

holds for each $n = 0, 1, \ldots$. Since $t_0 = 0$, $t_1 = \beta$, $\|v_0\| = \beta < \mathbf{r}_{p_0}$ and $p_1 = \exp_{p_0}(v_0)$,

$$d(p_1, p_0) = \|v_0\| \leqslant t_1 - t_0.$$

Therefore, the result is clear for the case when $n = 0$. Assume that

$$d(p_{n+1}, p_n) = \|v_n\| \leqslant t_{n+1} - t_n, \quad n = 0, 1, \ldots, k-1. \tag{4.14}$$

Then, we have

$$d(p_k, p_0) \leqslant \sum_{n=0}^{k-1} d(p_{n+1}, p_n) \leqslant \sum_{n=0}^{k-1}(t_{n+1} - t_n) = t_k < r_1 < \frac{1-\frac{\sqrt{2}}{2}}{\gamma}. \tag{4.15}$$

Hence, $DX(p_k)^{-1}$ exists by Lemma 2.3 and p_{k+1} is well-defined. Furthermore, by (2.7) (with $k = 2$), we have that

$$\left\|DX(p_k)^{-1}P_{p_k,p_{k-1}} \circ P_{p_{k-1},p_0}DX(p_0)\right\| \leqslant \frac{(1-\gamma(d(p_{k-1}, p_0) + d(p_k, p_{k-1})))^2}{\psi(\gamma(d(p_{k-1}, p_0) + d(p_k, p_{k-1})))}. \tag{4.16}$$

Since $\frac{(1-u)^2}{\psi(u)} = \frac{-1}{h'\left(\frac{u}{\gamma}\right)}$, it follows that

$$\begin{aligned}\left\|DX(p_k)^{-1}P_{p_k,p_{k-1}} \circ P_{p_{k-1},p_0}DX(p_0)\right\| &\leqslant -h'(d(p_{k-1}, p_0) + d(p_k, p_{k-1}))^{-1}\\ &\leqslant -h'(t_k)^{-1}\end{aligned} \tag{4.17}$$

because $h'(t)$ is monotonic increasing on $[0, \frac{1-\frac{\sqrt{2}}{2}}{\gamma})$. Define the curve

$$c(t) := \exp_{p_{k-1}}(tv_{k-1}), \quad t \in [0, 1].$$

By (4.14), $d(p_k, p_{k-1}) = \|v_{k-1}\|$; hence, c is the minimizing geodesic connecting p_{k-1} and p_k. Using Lemma 2.1, we obtain that

$$\begin{aligned}P_{c,p_{k-1},p_k}X(p_k) &= P_{c,p_{k-1},p_k}X(p_k) - X(p_{k-1}) - DX(p_{k-1})v_{k-1}\\ &= \int_0^1 \left(P_{c,p_{k-1},c(\tau)}DX(c(\tau))c'(\tau) - DX(p_{k-1})v_{k-1}\right) d\tau.\end{aligned} \tag{4.18}$$

By Lemma 2.2, it follows that

$$P_{c,p_{k-1},c(\tau)}DX(c(\tau))c'(\tau) - DX(p_{k-1})v_{k-1} = \int_0^\tau P_{c,p_{k-1},c(s)}D^2X(c(s))(c'(s))^2\, ds. \tag{4.19}$$

Note that $h''(u) = \frac{2\gamma}{(1-\gamma u)^3}$; hence by (4.18), (4.19) and (2.3) (with $k = 2$), we get that

$$\begin{aligned}&\left\|DX(p_0)^{-1}P_{p_0,p_{k-1}} \circ P_{c,p_{k-1},p_k}X(p_k)\right\|\\ &\leqslant \int_0^1\int_0^\tau \left\|DX(p_0)^{-1}P_{p_0,p_{k-1}} \circ P_{c,p_{k-1},c(s)}D^2X(c(s))\right\| \|c'(s)\|^2\, ds\, d\tau\\ &\leqslant \int_0^1\int_0^\tau h''(t_{k-1} + s\|v_{k-1}\|)\|v_{k-1}\|^2\, ds\, d\tau\\ &\leqslant \int_0^1\int_0^\tau h''(t_{k-1} + s(t_k - t_{k-1}))(t_k - t_{k-1})^2\, ds\, d\tau\\ &= h(t_k) - h(t_{k-1}) - h'(t_{k-1})(t_k - t_{k-1})\\ &= h(t_k),\end{aligned} \tag{4.20}$$

where the last equality holds because $-h(t_{k-1}) - h'(t_{k-1})(t_k - t_{k-1}) = 0$ by (4.2) (with $n = k$). Therefore, (4.17) and (4.20) imply that

$$\begin{aligned}\|DX(p_k)^{-1}X(p_k)\| &\leqslant \left\|DX(p_k)^{-1}P_{c,p_k,p_{k-1}} \circ P_{p_{k-1},p_0}DX(p_0)\right\| \\ &\quad \times \left\|DX(p_0)^{-1}P_{p_0,p_{k-1}} \circ P_{c,p_{k-1},p_k}X(p_k)\right\| \\ &\leqslant -h'(t_k)^{-1}h(t_k) \\ &= t_{k+1} - t_k.\end{aligned}$$

Hence, in view of (4.12),

$$\|v_k\| = \|DX(p_k)^{-1}X(p_k)\| \leqslant t_{k+1} - t_k. \tag{4.21}$$

As $\beta < (2-\sqrt{2})\mathbf{r}_{p_0}$, it follows from (4.4) that $r_1 \leqslant \mathbf{r}_{p_0}$. Thus, (4.15) and (4.21) yield that

$$\|v_k\| + d(p_k, p_0) \leqslant t_{k+1} < r_1 \leqslant \mathbf{r}_{p_0}.$$

This, together with Proposition 2.1, implies that

$$\|v_k\| \leqslant \mathbf{r}_{p_0} - d(p_k, p_0) \leqslant \mathbf{r}_{p_k}.$$

Since $p_{k+1} = \exp_{p_k}(v_k)$, it follows from the definition of $\mathbf{r}_{p_k}$ that $d(p_{k+1}, p_k) = \|v_k\|$. Hence, it is seen that (4.13) holds for $n = k$ thanks to (4.21). Combining (4.13) and (4.6), we get (4.10) and complete the proof. □

By (4.9), we arrive at the following corollary.

COROLLARY 4.1 Let

$$\beta \leqslant (2-\sqrt{2})\mathbf{r}_{p_0} \quad \text{and} \quad \alpha < 3 - 2\sqrt{2}.$$

Suppose that X satisfies the two-piece γ-condition at p_0 in $\mathbf{B}(p_0, r_1)$. Then Newton's method (2.1) with the initial point p_0 is well-defined and the generated sequence $\{p_n\}$ converges to a zero p^* of X in $\overline{\mathbf{B}(p_0, r_1)}$. Moreover,

$$d(p_{n+1}, p_n) \leqslant \mu^{2^n-1} d(p_1, p_0), \quad n = 0, 1, \ldots,$$

where μ is defined by (4.7).

5. Alternative formulation of the generalized γ-theorem

This section will provide an alternative formulation of the generalized γ-theorem, which is independent of the geometric number K_{p^*}. Recall that X is a C^2 vector field on M and that $p^* \in M$ is such that $DX(p^*)^{-1}$ exists. Recall from (2.6) that the function ψ is defined by

$$\psi(u) = 1 - 4u + 2u^2, \quad u \in \left[0, 1 - \frac{\sqrt{2}}{2}\right).$$

The following lemma estimates the value of the quantity $\|DX(p_0)^{-1}X(p_0)\|$, which will be used in the proof of the main theorem of this section.

LEMMA 5.1 Let $r < \frac{1-\frac{\sqrt{2}}{2}}{\gamma}$ and let $p_0 \in \mathbf{B}(p^*, r)$. Suppose that X satisfies the one-piece γ-condition at p^* in $\mathbf{B}(p^*, r)$. Then $\mathrm{D}X(p_0)^{-1}$ exists and

$$\|\mathrm{D}X(p_0)^{-1}X(p_0)\| \leqslant \frac{1-u}{\psi(u)} d(p_0, p^*), \tag{5.1}$$

where $u = \gamma\, d(p_0, p^*)$.

Proof. By Lemma 2.3, $\mathrm{DX}(p_0)^{-1}$ exists and

$$\left\|\mathrm{D}X(p_0)^{-1} P_{p_0,p^*}\mathrm{D}X(p^*)\right\| \leqslant \frac{(1-u)^2}{\psi(u)}. \tag{5.2}$$

Below, we will show that

$$\left\|\mathrm{D}X(p^*)^{-1} P_{p^*,p_0}X(p_0)\right\| \leqslant \frac{d(p_0, p^*)}{(1-u)}. \tag{5.3}$$

Granting this, by (5.2), we have that

$$\|\mathrm{DX}(p_0)^{-1}X(p_0)\| \leqslant \left\|\mathrm{D}X(p_0)^{-1} P_{p_0,p^*}\mathrm{D}X(p^*)\right\| \left\|\mathrm{D}X(p^*)^{-1} P_{p^*,p_0}X(p_0)\right\| \leqslant \frac{1-u}{\psi(u)} d(p_0, p^*)$$

and so (5.1) is seen to hold. To verify (5.3), let c: $[0, 1] \to M$ be a minimizing geodesic connecting p^* and p_0. Then there exists $v \in T_{p^*}M$ such that $\|v\| = d(p_0, p^*)$ and $c(t) = \exp_{p^*}(tv)$ for each $t \in [0, 1]$. Observe that

$$\begin{aligned} P_{c,p^*,p_0}X(p_0) &= P_{c,p^*,p_0}X(p_0) - X(p^*) - \mathrm{D}X(p^*)v + \mathrm{D}X(p^*)v \\ &= \int_0^1 P_{c,p^*,c(\tau)}\mathrm{D}X(c(\tau))c'(\tau)\,\mathrm{d}\tau - \mathrm{D}X(p^*)v + \mathrm{D}X(p^*)v \\ &= \int_0^1 \int_0^\tau P_{c,p^*,c(s)}\mathrm{D}^2X(c(s))c'(s)^2\,\mathrm{d}s\,\mathrm{d}\tau + \mathrm{D}X(p^*)v, \end{aligned} \tag{5.4}$$

where the second equality holds because of Lemma 2.1 while the third equality is valid because of Lemma 2.2. Thus, by (2.3) (with $k = 1$),

$$\begin{aligned} \left\|\mathrm{D}X(p^*)^{-1} P_{c,p^*,p_0}X(p_0)\right\| &\leqslant \int_0^1 \int_0^\tau \|\mathrm{D}X(p^*)^{-1} P_{c,p^*,c(s)}\mathrm{D}^2X(c(s))c'(s)^2\|\,\mathrm{d}s\,\mathrm{d}\tau + \|v\| \\ &\leqslant \int_0^1 \int_0^\tau \frac{2\gamma}{(1-s\gamma\,\|v\|)^3}\|v\|^2\,\mathrm{d}s\,\mathrm{d}\tau + \|v\| \\ &= \frac{d(p_0, p^*)}{(1-u)} \end{aligned}$$

and hence (5.3) holds. □

Let $u_0 = 0.080851\ldots$ be the smallest positive root of the equation

$$\frac{u_0}{\psi(u_0)^2} = 3 - 2\sqrt{2}. \tag{5.5}$$

Also, let

$$t_0 = \frac{2-\sqrt{2}}{2-\sqrt{2}+\frac{1-u_0}{\psi(u_0)}} = 0.305332\ldots. \tag{5.6}$$

Then

$$\frac{1-u_0}{\psi(u_0)}\frac{t_0}{1-t_0} = 2-\sqrt{2}. \tag{5.7}$$

Recall that $u = \gamma\, d(p_0, p^*)$ and $\beta = \|\mathrm{D}X(p_0)^{-1}X(p_0)\|$. Furthermore, set

$$\bar{\gamma} = \frac{\gamma}{\psi(u)(1-u)} \quad \text{and} \quad \bar{\alpha} = \beta\bar{\gamma}.$$

THEOREM 5.1 Let

$$r = \min\left\{t_0\mathbf{r}_{p^*}, \frac{u_0}{\gamma}\right\} \quad \text{and} \quad \bar{r} = \min\left\{\mathbf{r}_{p^*}, \frac{2-\sqrt{2}}{2\gamma}\right\}.$$

Suppose that $X(p^*) = 0$ and that X satisfies the three-piece γ-condition at p^* in $\mathbf{B}(p^*, \bar{r})$. If $d(p_0, p^*) < r$, then Newton's method (2.1) with the initial point p_0 is well-defined and

$$d(p_n, p^*) \leqslant \sigma(\bar{\mu})^{2^n-1} d(p_1, p_0),$$

where

$$\bar{\mu} = \frac{1-\bar{\alpha}-\sqrt{(1+\bar{\alpha})^2-8\bar{\alpha}}}{1-\bar{\alpha}+\sqrt{(1+\bar{\alpha})^2-8\bar{\alpha}}} \tag{5.8}$$

and

$$\sigma = \sum_{n\geqslant 0}(\bar{\mu})^{2^n-1}. \tag{5.9}$$

Proof. By Lemma 5.1, $\mathrm{D}X(p_0)^{-1}$ exists and

$$\beta = \|\mathrm{D}X(p_0)^{-1}X(p_0)\| \leqslant \frac{1-u}{\psi(u)} d(p_0, p^*) = \frac{1-u}{\psi(u)}\frac{u}{\gamma}. \tag{5.10}$$

Then

$$\bar{\alpha} = \beta\bar{\gamma} \leqslant \frac{u}{\psi(u)^2} < \frac{u_0}{\psi(u_0)^2} = 3-2\sqrt{2} \tag{5.11}$$

because the function $u \mapsto \frac{u}{\psi(u)^2}$ is strictly monotonic increasing on $\left[0, 1-\frac{\sqrt{2}}{2}\right)$. Let

$$\bar{r}_1 = \frac{1+\bar{\alpha}-\sqrt{(1+\bar{\alpha})^2-8\bar{\alpha}}}{4\bar{\gamma}}. \tag{5.12}$$

Then, by (4.4),

$$\beta \leqslant \bar{r}_1 \leqslant \left(1+\frac{1}{\sqrt{2}}\right)\beta \leqslant \left(1-\frac{1}{\sqrt{2}}\right)\frac{1}{\bar{\gamma}} \tag{5.13}$$

thanks to (5.11). Since $d(p_0, p^*) \leqslant t_0 \mathbf{r}_{p^*}$ and the function $u \mapsto \frac{1-u}{\psi(u)}$ is strictly monotonic increasing on $[0, 1-\frac{\sqrt{2}}{2})$, by (5.10), we have that

$$\beta \leqslant \frac{1-u_0}{\psi(u_0)} \frac{t_0}{1-t_0}(1-t_0)\mathbf{r}_{p^*} = (2-\sqrt{2})(1-t_0)\mathbf{r}_{p^*} \tag{5.14}$$

thanks to (5.7). Hence, by (5.13),

$$\bar{r}_1 \leqslant (1-t_0)\mathbf{r}_{p^*}. \tag{5.15}$$

By Proposition 2.1,

$$\mathbf{r}_{p^*} \leqslant \mathbf{r}_{p_0} + d(p_0, p^*) \leqslant \mathbf{r}_{p_0} + t_0 \mathbf{r}_{p^*}. \tag{5.16}$$

Therefore,

$$\mathbf{r}_{p^*} \leqslant \frac{\mathbf{r}_{p_0}}{1-t_0}. \tag{5.17}$$

Thus, (5.14) implies that

$$\beta \leqslant (2-\sqrt{2})\mathbf{r}_{p_0}. \tag{5.18}$$

Then, in order to ensure that Corollary 4.1 is applicable, we have to show the following assertion: there exists $\hat{r} \geqslant \bar{r}_1$ such that X satisfies the two-piece $\bar{\gamma}$-condition at p_0 in $\mathbf{B}(p_0, \hat{r})$.

For this purpose, let

$$\hat{r} = \bar{r} - d(p_0, p^*) = \min\left\{\mathbf{r}_{p^*}, \frac{2-\sqrt{2}}{2\gamma}\right\} - d(p_0, p^*). \tag{5.19}$$

We claim that $\hat{r}$ is the number desired. First, we have that

$$\hat{r} \geqslant \bar{r}_1. \tag{5.20}$$

In fact, if $\bar{r} = \mathbf{r}_{p^*}$, then

$$\hat{r} = \mathbf{r}_{p^*} - d(p_0, p^*) \geqslant (1-t_0)\mathbf{r}_{p^*} \geqslant \bar{r}_1;$$

if $\bar{r} = \frac{2-\sqrt{2}}{2\gamma}$, then

$$\hat{r} = \frac{1-\frac{\sqrt{2}}{2}}{\gamma} - \frac{u}{\gamma} \geqslant \left(1-\frac{\sqrt{2}}{2}\right)\frac{\psi(u)(1-u)}{\gamma} = \left(1-\frac{\sqrt{2}}{2}\right)\frac{1}{\bar{\gamma}} \geqslant \bar{r}_1.$$

Therefore, (5.20) is proved. Next, we have that X satisfies the two-piece $\bar{\gamma}$-condition at p_0 in $\mathbf{B}(p_0, \hat{r})$. Indeed, for any two points $p, q \in \mathbf{B}(p_0, \hat{r})$ with

$$d(p_0, p) + d(p, q) < \hat{r} \tag{5.21}$$

since X satisfies the three-piece γ-condition at p^* in $\mathbf{B}(p^*, \bar{r})$ and $d(p^*, p_0) + d(p_0, p) + d(p, q) < \bar{r}$, we obtain that

$$\left\| \mathrm{D}X(p^*)^{-1} P_{p^*, p_0} \circ P_{p_0, p} \circ P_{p,q} \mathrm{D}^2 X(q) \right\| \leqslant \frac{2\gamma}{(1-\gamma(d(p^*, p_0) + d(p_0, p) + d(p, q)))^3}. \tag{5.22}$$

Consequently, using Lemma 2.3 (with $k = 1$) and (5.22), we conclude that

$$
\begin{aligned}
&\left\| DX(p_0)^{-1} P_{p_0,p} \circ P_{p,q} D^2 X(p) \right\| \\
&= \left\| DX(p_0)^{-1} P_{p_0,p^*} DX(p^*) \right\| \left\| DX(p^*)^{-1} P_{p^*,p_0} \circ P_{p_0,p} \circ P_{p,q} D^2 X(q) \right\| \\
&\leqslant \frac{(1-u)^2}{\psi(u)} \frac{2\gamma}{(1-\gamma(d(p^*,p_0)+d(p_0,p)+d(p,q)))^3} \\
&= \frac{2\gamma}{\psi(u)(1-u)} \frac{(1-u)^3}{(1-u-\gamma(d(p_0,p)+d(p,q)))^3} \\
&= \frac{2\bar{\gamma}}{\left(1-\frac{\gamma}{1-u}(d(p_0,p)+d(p,q))\right)^3} \\
&\leqslant \frac{2\bar{\gamma}}{\left(1-\frac{\gamma}{\psi(u)(1-u)}(d(p_0,p)+d(p,q))\right)^3} \\
&= \frac{2\bar{\gamma}}{(1-\bar{\gamma}(d(p_0,p)+d(p,q)))^3}
\end{aligned}
$$

because $0 < \psi(u) < 1$ for all $u \in (0, 1 - \frac{\sqrt{2}}{2})$. Therefore, X satisfies the two-piece $\bar{\gamma}$-condition at p_0 in $\mathbf{B}(p_0, \hat{r})$ and the assertion holds. Thus, we apply Corollary 4.1 to conclude that the sequence $\{p_n\}$ generated by Newton's method (2.1) with the initial point p_0 converges to a zero q^* of X in $\overline{\mathbf{B}(p_0, \bar{r}_1)}$ and

$$d(p_{n+1}, p_n) \leqslant (\bar{\mu})^{2^n - 1} d(p_1, p_0), \quad n = 0, 1, 2, \ldots. \tag{5.23}$$

To complete the proof, it remains to verify that $p^* = q^*$. To this end, let $v \in T_{p^*}M$ be such that $q^* = \exp_{p^*}(v)$ and $\|v\| = d(p^*, q^*)$. Then the curve c defined by $c(t) = \exp_{p^*}(tv)$, $t \in [0, 1]$, is a minimizing geodesic connecting p^* and q^*. As $c'(t) = P_{c,c(t),p^*} v$, it follows from Lemma 2.1 that

$$
\begin{aligned}
\left(\int_0^1 DX(p^*)^{-1} P_{c,p^*,c(t)} DX(c(t)) P_{c,c(t),p^*} \, dt \right) v &= DX(p^*)^{-1} [P_{c,p^*,q^*} X(q^*) - X(p^*)] \\
&= 0.
\end{aligned} \tag{5.24}
$$

We claim that $\int_0^1 DX(p^*)^{-1} P_{c,p^*,c(t)} DX(c(t)) P_{c,c(t),p^*} \, dt$ is invertible. Granting this, (5.24) implies that $v = 0$ and so $p^* = q^*$. Let $\bar{v} \in T_{p^*}M$ and let Y be the unique vector field such that $Y(c(0)) = \bar{v}$ and $D_{c'(t)} Y(c(t)) = 0$. By Lemma 2.2, one has that

$$
\begin{aligned}
&\int_0^1 DX(p^*)^{-1} [P_{c,p^*,c(t)} DX(c(t)) P_{c,c(t),p^*} - DX(p^*)] \bar{v} \, dt \\
&\quad = \int_0^1 \int_0^t DX(p^*)^{-1} P_{c,p^*,c(s)} D^2 X(c(s)) Y(c(s)) c'(s) \, ds \, dt.
\end{aligned} \tag{5.25}
$$

Note that, by (5.19) and (5.20),

$$d(q^*, p^*) \leqslant d(q^*, p_0) + d(p_0, p^*) \leqslant \bar{r}_1 + d(p_0, p^*) \leqslant \hat{r} + d(p_0, p^*) = \bar{r}. \tag{5.26}$$

This implies that $d(c(s), p^*) < \bar{r}$ for each $s \in (0, 1)$. It follows that, for each $s \in (0, 1)$,

$$\|DX(p^*)^{-1}P_{c,p^*,c(s)}D^2X(c(s))Y(c(s))c'(s)\| \leqslant \frac{2\gamma}{(1-\gamma s d(p^*, q^*))^3}\|\bar{v}\|d(p^*, q^*) \tag{5.27}$$

since X satisfies the three-piece γ-condition (and therefore the one-piece γ-condition) at p^* in $\mathbf{B}(p^*, \bar{r})$. Consequently, by (5.25) and (5.27),

$$\begin{aligned}
&\left\|\int_0^1 DX(p^*)^{-1}[P_{c,p^*,c(t)}DX(c(t))P_{c,c(t),p^*} - DX(p^*)]\bar{v}\,dt\right\| \\
&\quad \leqslant \int_0^1\int_0^t \frac{2\gamma}{(1-\gamma s d(p^*, q^*))^3}\|\bar{v}\|d(p^*, q^*)\,ds\,dt \\
&\quad < \int_0^1\int_0^1 \frac{2\gamma}{(1-\gamma s d(p^*, q^*))^3}\|\bar{v}\|d(p^*, q^*)\,ds\,dt \\
&\quad = \left(\frac{1}{(1-\gamma d(p^*, q^*))^2} - 1\right)\|\bar{v}\| \\
&\quad \leqslant \|\bar{v}\|,
\end{aligned} \tag{5.28}$$

where the last inequality follows from (5.26) and the fact that $\bar{r} \leqslant \frac{2-\sqrt{2}}{2\gamma}$. Hence,

$$\left\|\int_0^1 DX(p^*)^{-1}[P_{c,p^*,c(t)}DX(c(t))P_{c,c(t),p^*} - X(p^*)]\,dt\right\| < 1.$$

By the Banach lemma, the claim holds and the proof is complete. □

6. Application to analytic vector fields

Throughout this section, we shall always assume that M is an analytic complete m-dimensional Riemannian manifold. Let $p \in M$. Recall from Boothby (1986) and DoCarmo (1992) that a vector field X is said to be analytic at p if there exists a local coordinate system $(U, \{x^i\})$ of p and m analytic functions $X^i: U \to \mathbb{R}$, $i = 1, 2, \ldots, m$, such that

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

Then the vector field X is analytic on M if it is analytic at each point of M.

In the remainder of this section, we assume that X is analytic on M. Let $p \in M$ be such that $DX(p)^{-1}$ exists. Following Dedieu *et al.* (2003), we define

$$\gamma(X, p) = \sup_{k\geqslant 2}\left\|DX(p)^{-1}\frac{D^kX(p)}{k!}\right\|_p^{\frac{1}{k-1}}. \tag{6.1}$$

Also we adopt the convention that $\gamma(X, p) = \infty$ if $DX(p)$ is not invertible. Note that this definition is justified and in the case when $DX(p)$ is invertible, by analyticity, $\gamma(X, p)$ is finite. The following Taylor formula for vector fields can be found in Dedieu *et al.* (2003).

LEMMA 6.1 Let $r = \min\left\{\mathbf{r}_p, \frac{1}{\gamma(X,p)}\right\}$. Let $q \in \mathbf{B}(p, r)$ and $v \in T_pM$ be such that $q = \exp_p(v)$. Then

$$X(q) = P_{q,p}\left(\sum_{k=0}^{\infty} \frac{1}{k!} \mathrm{D}^k X(p) v^k\right).$$

Taking the lth covariant derivative in Lemma 6.1 gives the following corollary.

COROLLARY 6.1 Under the same hypotheses as in Lemma 6.1, for any $l \geqslant 0$, we have

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

where $P_{p,q}^l$ stands for the map from $(T_qM)^l$ to $(T_pM)^l$ defined by

$$P_{p,q}^l(v_1, \ldots, v_l) = (P_{p,q}v_1, \ldots, P_{p,q}v_l) \quad \forall (v_1, \ldots, v_l) \in (T_qM)^l.$$

The following two lemmas were given in Dedieu *et al.* (2003). Let $q_0 \in M$ be such that $\mathrm{D}X(q_0)^{-1}$ exists.

LEMMA 6.2 Let $|r| < 1$ and let k be a positive integer. Then

$$\sum_{l=0}^{\infty} \frac{(k+l)!}{k!l!} r^l = \frac{1}{(1-r)^{k+1}}.$$

LEMMA 6.3 Let $u = \gamma(X, q_0) d(p, q_0)$. If $d(p, q_0) < \min\left\{\mathbf{r}_{q_0}, \frac{2-\sqrt{2}}{2\gamma(X,q_0)}\right\}$, then

$$\gamma(X, p) \leqslant \frac{\gamma(X, q_0)}{(1-u)\psi(u)}. \tag{6.2}$$

The following lemma shows that an analytic vector field satisfies the three-piece γ-condition at q_0 in $\mathbf{B}(q_0, r)$, where $\gamma = \gamma(X, q_0)$ and $r = \min\left\{\mathbf{r}_{q_0}, \frac{2-\sqrt{2}}{2\gamma(X,q_0)}\right\}$.

LEMMA 6.4 Let $0 < r \leqslant \min\left\{\mathbf{r}_{q_0}, \frac{2-\sqrt{2}}{2\gamma(X,q_0)}\right\}$. Then X satisfies the three-piece γ-condition at q_0 in $\mathbf{B}(q_0, r)$.

Proof. Let $p_0, p, q \in \mathbf{B}(q_0, r)$ be such that

$$d(q_0, p_0) + d(p_0, p) + d(p, q) < r. \tag{6.3}$$

Set

$$\hat{r}_1 = \min\left\{\mathbf{r}_p, \frac{1}{\gamma(X, p)}\right\}, \quad \hat{r}_2 = \min\left\{\mathbf{r}_{p_0}, \frac{1}{\gamma(X, p_0)}\right\}, \quad \hat{r}_3 = \min\left\{\mathbf{r}_{q_0}, \frac{1}{\gamma(X, q_0)}\right\}. \tag{6.4}$$

Below, we claim that

$$q \in \mathbf{B}(p, \hat{r}_1), \quad p \in \mathbf{B}(p_0, \hat{r}_2), \quad p_0 \in \mathbf{B}(q_0, \hat{r}_3). \tag{6.5}$$

We only show that $q \in \mathbf{B}(p, \hat{r}_1)$ since the proofs for $p \in \mathbf{B}(p_0, \hat{r}_2)$ and $p_0 \in \mathbf{B}(q_0, \hat{r}_3)$ are similar. As $d(q_0, p_0) + d(p_0, p) + d(p, q) < r \leqslant \mathbf{r}_{q_0}$, it follows from Proposition 2.1 that

$$d(p, q) \leqslant \mathbf{r}_{q_0} - d(q_0, p) \leqslant \mathbf{r}_p. \tag{6.6}$$

Write $u = \gamma(X, q_0)d(p, q_0)$. Since $p \in \mathbf{B}(q_0, r)$, Lemma 6.3 is applicable. It follows that

$$\gamma(X, p) \leqslant \frac{\gamma(X, q_0)}{(1-u)\psi(u)}. \tag{6.7}$$

By a simple calculation, we see that

$$\frac{(1-u)\psi(u)}{\gamma(X, q_0)} \geqslant \frac{1-\frac{\sqrt{2}}{2}}{\gamma(X, q_0)} - d(p, q_0).$$

Then, by (6.7) and (6.3),

$$d(p, q) \leqslant \frac{1}{\gamma(X, p)}. \tag{6.8}$$

This, together with (6.6), implies that $q \in \mathbf{B}(p, \hat{r}_1)$; hence, our claim holds. Thus, by (6.5), Corollary 6.1 is applicable to conclude that

$$\begin{aligned}
&\mathrm{D}X(q_0)^{-1}P_{q_0,p_0} \circ P_{p_0,p} \circ P_{p,q}\mathrm{D}^2X(q)\\
&\quad = \mathrm{D}X(q_0)^{-1}P_{q_0,p_0} \circ P_{p_0,p}\sum_{l=0}^{\infty}\frac{1}{l!}\mathrm{D}^{l+2}X(p)v_3^lP_{p,q}^2\\
&\quad = \mathrm{D}X(q_0)^{-1}P_{q_0,p_0}\sum_{l=0}^{\infty}\frac{1}{l!}\sum_{j=0}^{\infty}\frac{1}{j!}\mathrm{D}^{l+j+2}X(p_0)v_2^jP_{p_0,p}^{l+2}v_3^lP_{p,q}^2\\
&\quad = \sum_{l=0}^{\infty}\frac{1}{l!}\sum_{j=0}^{\infty}\frac{1}{j!}\sum_{k=0}^{\infty}\frac{1}{k!}\mathrm{D}X(q_0)^{-1}\mathrm{D}^{l+j+k+2}X(q_0)v_1^kP_{q_0,p_0}^{l+j+2}v_2^jP_{p_0,p}^{l+2}v_3^lP_{p,q}^2,
\end{aligned} \tag{6.9}$$

where $v_1 \in T_{q_0}M$, $v_2 \in T_{p_0}M$ and $v_3 \in T_pM$ satisfy that $p_0 = \exp_{q_0}(v_1)$, $p = \exp_{p_0}(v_2)$ and $q = \exp_p(v_3)$, respectively. Since

$$\frac{\|\mathrm{D}X(q_0)^{-1}\mathrm{D}^{l+j+k+2}X(q_0)\|}{(l+j+k+2)!} \leqslant \gamma(X, q_0)^{l+j+k+1},$$

one has from (6.9) that

$$\begin{aligned}
&\left\|\mathrm{D}X(q_0)^{-1}P_{q_0,p_0} \circ P_{p_0,p} \circ P_{p,q}\mathrm{D}^2X(q)\right\|\\
&\quad \leqslant \sum_{l=0}^{\infty}\frac{1}{l!}\sum_{j=0}^{\infty}\frac{(l+j+2)!}{j!}\sum_{k=0}^{\infty}\frac{(l+j+k+2)!}{k!(l+j+2)!}\gamma(X, q_0)^{l+j+k+1}\|v_1\|^k\|v_2\|^j\|v_3\|^l.
\end{aligned} \tag{6.10}$$

Using Lemma 6.2 to calculate the quantity on the right-hand side of the inequality (6.10), we get that

$$\left\|\mathrm{D}X(q_0)^{-1}P_{q_0,p_0} \circ P_{p_0,p} \circ P_{p,q}\mathrm{D}^2X(q)\right\| \leqslant \frac{2\gamma(X, q_0)}{(1-\gamma(X, q_0)(\|v_1\|+\|v_2\|+\|v_3\|))^3}. \tag{6.11}$$

Since $\|v_1\| = d(q_0, p_0)$, $\|v_2\| = d(p_0, p)$ and $\|v_3\| = d(p, q)$, it follows from (6.11) that

$$\left\|\mathrm{D}X(q_0)^{-1}P_{q_0,p_0} \circ P_{p_0,p} \circ P_{p,q}\mathrm{D}^2X(q)\right\| \leqslant \frac{2\gamma(X, q_0)}{(1-\gamma(X, q_0)(d(q_0, p_0)+d(p_0, p)+d(p, q)))^3}.$$

Hence, X satisfies the three-piece γ-condition at q_0 in $\mathbf{B}(q_0, r)$ and the proof is complete. □

Then, by Theorem 3.1, we have the following corollary which was obtained in Dedieu *et al.* (2003) with a different technique.

COROLLARY 6.2 Let $p^* \in M$ be such that $\mathrm{D}X(p^*)^{-1}$ exists. Suppose $X(p^*) = 0$ and let $p_0 \in M$. If

$$d(p_0, p^*) \leqslant \min\left\{\mathbf{r}_{p^*}, \frac{2 + K_{p^*} - \sqrt{K_{p^*}^2 + 4K_{p^*} + 2}}{2\gamma(X, p^*)}\right\},$$

then Newton's method (2.1) with the initial point p_0 is well-defined, and

$$d(p_n, p^*) \leqslant \left(\frac{1}{2}\right)^{2^n-1} d(p_0, p^*), \quad n = 0, 1, 2, \ldots. \tag{6.12}$$

Proof. Write

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

Let $\gamma = \gamma(X, p^*)$. Then $p_0 \in \mathbf{B}(p^*, r_c)$ because $\frac{\delta}{\gamma} < r_c$, where r_c is defined by (3.2). As $r_c < \frac{2-\sqrt{2}}{2\gamma}$, it follows from Lemma 6.4 that X satisfies the one-piece γ-condition at p^* in $\mathbf{B}(p^*, r)$ with $r = \min\{\mathbf{r}_{p^*}, r_c\}$. Hence, Theorem 3.1 is applicable to conclude that Newton's method (2.1) is well-defined for p_0, and (3.3) holds for λ defined by (3.4). As

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

(6.12) holds by (3.3). The proof is complete. □

Similarly, by Corollary 4.1 and Theorem 5.1, we also have the following two corollaries, which improve the corresponding results due to Dedieu *et al.* (2003). Recall that $p_0 \in M$ is such that $\mathrm{D}X(p_0)^{-1}$ exists and also that $\beta = \|\mathrm{D}X(p_0)^{-1}X(p_0)\|$ and $\alpha = \beta\gamma$, where $\gamma = \gamma(X, p_0)$. Let

$$r_1 = \frac{1 + \alpha - \sqrt{(1+\alpha)^2 - 8\alpha}}{4\gamma}.$$

Let $\bar{u}_0 = 0.0776121\ldots$ be the smallest positive root of the equation

$$\frac{\bar{u}_0}{\psi(\bar{u}_0)^2} = \frac{13 - 3\sqrt{17}}{4}. \tag{6.13}$$

COROLLARY 6.3 If

$$\beta \leqslant (2 - \sqrt{2})\mathbf{r}_{p_0} \quad \text{and} \quad \alpha \leqslant \frac{13 - 3\sqrt{17}}{4} \approx 0.157671,$$

then Newton's method (2.1) with the initial point p_0 is well-defined and the generated sequence $\{p_n\}$ converges to a zero p^* of X in $\overline{\mathbf{B}(p_0, r_1)}$. Moreover,

$$d(p_{n+1}, p_n) \leqslant \left(\frac{1}{2}\right)^{2^n-1} d(p_1, p_0).$$

Proof. As $\beta \leqslant (2-\sqrt{2})\mathbf{r}_{p_0}$, it follows from (4.4) that $r_1 \leqslant \min\{\mathbf{r}_{p_0}, \frac{2-\sqrt{2}}{2\gamma(X,p_0)}\}$. Thus, by Lemma 6.4, X satisfies the two-piece γ-condition at p_0 in $\mathbf{B}(p_0, r_1)$ with $\gamma = \gamma(X, p_0)$. Note also that

$$\alpha \leqslant \frac{13-3\sqrt{17}}{4} < 3-2\sqrt{2}.$$

Corollary 4.1 is applicable to conclude that

$$d(p_{n+1}, p_n) \leqslant \mu^{2^n-1} d(p_1, p_0),$$

where

$$\mu = \frac{1-\alpha-\sqrt{(1+\alpha)^2-8\alpha}}{1-\alpha+\sqrt{(1+\alpha)^2-8\alpha}}.$$

Therefore, we have that $\mu \leqslant \frac{1}{2}$ because μ increases as α does on $\left[0, \frac{13-3\sqrt{17}}{4}\right]$ and the value of μ at $\alpha = \frac{13-3\sqrt{17}}{4}$ is $\frac{1}{2}$. □

COROLLARY 6.4 Let $p^* \in M$ be such that $\mathrm{D}X(p^*)^{-1}$ exists. Suppose $X(p^*) = 0$. Let $p_0 \in M$ and t_0 be given by (5.6). If

$$d(p_0, p^*) < \min\left\{t_0\mathbf{r}_{p^*}, \frac{\bar{u}_0}{\gamma(X, p^*)}\right\},$$

then Newton's method (2.1) with the initial point p_0 is well-defined and

$$d(p_n, p^*) \leqslant \sigma \left(\frac{1}{2}\right)^{2^n-1} d(p_1, p_0),$$

where $\sigma = \sum_{n\geqslant 0} \left(\frac{1}{2}\right)^{2^n-1}$.

Proof. Let

$$\gamma = \gamma(X, p^*), \quad \bar{\gamma} = \frac{\gamma}{\psi(u)(1-u)} \quad \text{and} \quad \bar{\alpha} = \beta\bar{\gamma},$$

where $u = \gamma d(p_0, p^*)$. By Lemma 6.4, X satisfies the three-piece γ-condition at p^* in $\mathbf{B}(p^*, \bar{r})$ with $\bar{r} = \min\{\mathbf{r}_{p^*}, \frac{2-\sqrt{2}}{2\gamma}\}$. Since $\bar{u}_0$ determined by (6.13) is less than u_0 given by (5.5), Theorem 5.1 is applicable. Thus, we have that Newton's method (2.1) with the initial point p_0 is well-defined and

$$d(p_{n+1}, p_n) \leqslant \bar{\mu}^{2^n-1} d(p_1, p_0),$$

where

$$\bar{\mu} = \frac{1-\bar{\alpha}-\sqrt{(1+\bar{\alpha})^2-8\bar{\alpha}}}{1-\bar{\alpha}+\sqrt{(1+\bar{\alpha})^2-8\bar{\alpha}}}.$$

By Lemma 5.1, we have that

$$\beta = \|\mathrm{D}X(p_0)^{-1}X(p_0)\| \leqslant \frac{1-u}{\psi(u)} d(p_0, p^*).$$

It follows that

$$\bar{\alpha} \leqslant \frac{u}{\psi(u)^2} \leqslant \frac{\bar{u}_0}{\psi(\bar{u}_0)^2} = \frac{13 - 3\sqrt{17}}{4} \tag{6.14}$$

because the function $u \mapsto \frac{u}{\psi(u)^2}$ is strictly monotonic increasing on $[0, 1 - \frac{\sqrt{2}}{2})$. Hence, we have that $\bar{\mu} \leqslant \frac{1}{2}$ and the proof is complete. □

7. Conclusion

We have established the γ-theory and the α-theory under the γ-conditions. In particular, when these results are applied to analytic vector fields, some results due to Dedieu *et al.* (2003) are improved. In addition, we should remark that the issue on mappings from manifolds to m-dimensional spaces can be addressed in almost the same way, so we do not elaborate further on this here.

Acknowledgements

We wish to thank the referees for their valuable comments and suggestions. This work was supported in part by the National Natural Science Foundation of China (grant 10271025) and Program for New Century Excellent Talents in University.

References

BLUM, L., CUCKER, F., SHUB, M. & SMALE, S. (1997) *Complexity and Real Computation*. New York: Springer.

BOOTHBY, W. M. (1986) *An Introduction to Differentiable Manifolds and Riemannian Geometry*, 2nd edn. New York: Academic Press, Inc.

DEDIEU, J. P., PRIOURET, P. & MALAJOVICH, G. (2003) Newton's method on Riemannian manifolds: covariant alpha theory. *IMA J. Numer. Anal.*, **23**, 395–419.

DOCARMO, M. P. (1992) *Riemannian Geometry*. Boston, MA: Birkhauser.

EDELMAN, A., ARIAS, T. A. & SMITH, T. (1998) The geometry of algorithms with orthogonality constraints. *SIAM J. Matrix Anal. Appl.*, **20**, 303–353.

FERREIRA, O. P. & SVAITER, B. F. (2002) Kantorovich's theorem on Newton's method in Riemannian manifolds. *J. Complex.*, **18**, 304–329.

GABAY, D. (1982) Minimizing a differentiable function over a differential manifold. *J. Optim. Theory Appl.*, **37**, 177–219.

KANTOROVICH, L. V. & AKILOV, G. P. (1982) *Functional Analysis*. Oxford: Pergamon.

LANG, S. (1995) *Differential and Riemannian Manifolds, GTM 160*. New York: Springer.

SMALE, S. (1981) The fundamental theorem of algebra and complexity theory. *Bull. Am. Math. Soc.*, **4**, 1–36.

SMALE, S. (1986) Newton's method estimates from data at one point. *The Merging of Disciplines: New Directions in Pure, Applied and Computational Mathematics* (R. Ewing, K. Gross & C. Martin eds). New York: Springer, pp. 185–196.

SMALE, S. (1997) Complexity theory and numerical analysis. *Acta Numer.*, **6**, 523–551.

SMITH, S. T. (1993) Geometric optimization method for adaptive filtering. *Ph.D. Thesis*, Harvard University, Cambridge, MA.

SMITH, S. T. (1994) Optimization techniques on Riemannian manifolds. *Fields Institute Communications*, vol. 3. (A. Bloch ed.) Providence, RI: American Mathematical Society, pp. 113–146.

UDRISTE, C. (1994) *Convex Functions and Optimization Methods on Riemannian Manifolds.* Dordrecht: Kluwer Academic.

WANG, X. H. (1999) Convergence of Newton's method and inverse function theorem in Banach space. *Math. Comput.*, **225**, 169–186.

WANG, X. H. & HAN, D. F. (1990) On the dominating sequence method in the point estimates and Smale's theorem. *Sci. Sin. Ser. A*, **33**, 135–144.

WANG, X. H. & HAN, D. F. (1997) Criterion α and Newton's method. *Chin. J. Numer. Appl. Math.*, **19**, 96–105.