Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2013, Article ID 367161, 9 pages
http://dx.doi.org/10.1155/2013/367161

Research Article

Local Convergence of Newton's Method on Lie Groups and Uniqueness Balls

Jinsu He,[1] Jinhua Wang,[2,3] and Jen-Chih Yao[4,5]

[1] *Department of Mathematics, Zhejiang Normal University, Jinhua 321004, China*
[2] *Department of Mathematics, Zhejiang University of Technology, Hangzhou 310032, China*
[3] *Department of Mathematics, National Sun Yat-sen University, Kaohsiung 804, Taiwan*
[4] *Center for Fundamental Science, Kaohsiung Medical University, Kaohsiung 80702, Taiwan*
[5] *Department of Mathematics, King Abdulaziz University, P.O. Box 80203, Jeddah 21589, Saudi Arabia*

Correspondence should be addressed to Jinsu He; hejinsu@zjnu.cn

Received 20 August 2013; Accepted 1 October 2013

Academic Editor: Antonio M. Peralta

An estimation of uniqueness ball of a zero point of a mapping on Lie group is established. Furthermore, we obtain a unified estimation of radius of convergence ball of Newton's method on Lie groups under a generalized L-average Lipschitz condition. As applications, we get estimations of radius of convergence ball under the Kantorovich condition and the γ-condition, respectively. In particular, under the γ-condition, our results improve the corresponding results in (Li et al. 2009, Corollary 4.1) as showed in Remark 17. Finally, applications to analytical mappings are also given.

1. Introduction

In a vector space framework, when f is a differentiable operator from some domain D in a real or complex Banach space X to another Y, Newton's method is one of the most important methods for finding the approximation solution of the equation $f(x) = 0$, which is formulated as follows: for any initial point $x_0 \in D$,

$$x_{n+1} = x_n - f'(x_n)^{-1} f(x_n), \quad n = 0, 1, \ldots. \tag{1}$$

As is well known, one of the most important results on Newton's method is Kantorovich's theorem (cf. [1]), which provides a simple and clear criterion ensuring quadratic convergence of Newton's method under the mild condition that the second Fréchet derivative of f is bounded (or more generally, the first derivative is Lipschitz continuous) and the boundedness of $\|f'(x)^{-1}\|$ on a proper open metric ball of the initial point x_0. Another important result on Newton's method is Smale's point estimate theory (i.e., α-theory and γ-theory) in [2], where the notions of approximate zeros were introduced and the rules to judge an initial point x_0 to be an approximate zero were established, depending on the information of the analytic nonlinear operator at this initial point and at a solution x^*, respectively. There are a lot of works on the weakness and/or the extension of the Lipschitz continuity made on the mappings; see, for example, [3–7] and references therein. In particular, Wang introduced in [6] the notion of Lipschitz conditions with L-average to unify both Kantorovich's and Smale's criteria.

In a Riemannian manifold framework, an analogue of the well-known Kantorovich theorem was given in [8] for Newton's method for vector fields on Riemannian manifolds while the extensions of the famous Smale α-theory and γ-theory in [2] to analytic vector fields and analytic mappings on Riemannian manifolds were done in [9]. In the recent paper [10], the convergence criteria in [9] were improved by using the notion of the γ-condition for the vector fields and mappings on Riemannian manifolds. The radii of uniqueness balls of singular points of vector fields satisfying the γ-conditions were estimated in [11], while the local behavior of Newton's method on Riemannian manifolds was studied in [12, 13]. Furthermore, in [14], Li and Wang extended the generalized L-average Lipschitz condition (introduced in [6]) to Riemannian manifolds and established a unified convergence

criterion of Newton's method on Riemannian manifolds. Similarly, inspired by the previous work of Zabrejko and Nguen in [7] on Kantorovich's majorant method, Alvarez et al. introduced in [15] a Lipschitz-type radial function for the covariant derivative of vector fields and mappings on Riemannian manifolds and gave a unified convergence criterion of Newton's method on Riemannian manifolds.

Note also that Mahony used one-parameter subgroups of a Lie group to develop a version of Newton's method on an arbitrary Lie group in [16], where the algorithm presented is independent of affine connections on the Lie group. This means that Newton's method on Lie groups is different from the one defined on Riemannian manifolds. On the other hand, motivated by looking for approaches to solv ordinary differential equations on Lie groups, Owren and Welfert also studied in [17] Newton's method, independent of affine connections on the Lie group, and showed the local quadratical convergence. Recently, Wang and Li [18] established Kantorovich's theorem (independent of the connection) for Newton's method on Lie group. More precisely, under the assumption that the differential of f satisfies the Lipschitz condition around the initial point (which is in terms of one-parameter semigroups and independent of the metric), the convergence criterion of Newton's method is presented. Extensions of Smale's point estimate theory for Newton's method on Lie groups were given in [19].

The purpose of the present paper is to continue the study of Newton's method on Lie groups. At first, we give an estimation of uniqueness ball of a zero point of a mapping on a Lie group. Second, we establish a unified estimation of radius of convergence ball of Newton's method on Lie groups under a generalized L-average Lipschitz condition. As applications, we obtain estimations of radius of convergence ball under the Kantorovich condition and the γ-condition, respectively. In particular, under the γ-condition, we get that (see Theorem 16) if $f(x^*) = 0$ and

$$\varrho(x^*, x_0) < \frac{3 - 2\sqrt{2}}{2\gamma}, \tag{2}$$

then the sequence $\{x_n\}$ generated by Newton's method (28) with initial point x_0 is well defined and converges quadratically to a zero y^* of f. This improves the corresponding results in [19, Corollary 4.1], where it was proved under the following assumption: there exists $v \in \mathfrak{g}$ such that

$$x_0 = x^* \exp v, \qquad \|v\| \le \frac{a_0}{\gamma}, \tag{3}$$

with $a_0 = 0.081256\ldots$ being the smallest positive root of the equation $a_0/(1 - 4a_0 + 2a_0^2)^2 = 3 - 2\sqrt{2}$. Clearly,

$$\frac{a_0}{\gamma} < \frac{3 - 2\sqrt{2}}{2\gamma}. \tag{4}$$

Note also that in general, there dose not exist $v \in \mathfrak{g}$ satisfying $x_0 = x^* \exp v$ because the exponential map is not surjective global, even if $\varrho(x^*, x_0) < (3 - 2\sqrt{2})/2\gamma$. In view of this, our results somewhat improve the corresponding results in [19, Corollary 4.1].

The remainder of the paper is organized as follows. Some preliminary results and notions are given in Section 2, while the estimation of uniqueness ball is presented in Section 3. In Section 4, the main results about estimations of convergence ball are explored. Theorems under the Kantorovich condition and the γ-condition are provided in Section 5. In the final section, we get the estimations of uniqueness ball and convergence ball under the assumption that f is analytic.

2. Notions and Preliminaries

Most of the notions and notations which are used in the present paper are standard; see, for example, [20, 21]. A Lie group $(G, \cdot)$ is a Hausdorff topological group with countable bases which also has the structure of an analytic manifold such that the group product and the inversion are analytic operations in the differentiable structure given on the manifold. The dimension of a Lie group is that of the underlying manifold, and we will always assume that it is m-dimensional. The symbol e designates the identity element of G. Let $\mathfrak{g}$ be the Lie algebra of the Lie group G which is the tangent space T_eG of G at e, equipped with Lie bracket $[\cdot,\cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$.

In the sequel, we will make use of the left translation of the Lie group G. We define for each $y \in G$ the left translation $L_y : G \to G$ by

$$L_y(z) = y \cdot z \quad \text{for each } z \in G. \tag{5}$$

The differential of L_y at z is denoted by $(L'_y)_z$ which clearly determines a linear isomorphism from T_zG to the tangent space $T_{(y\cdot z)}G$. In particular, the differential $(L'_y)_e$ of L_y at e determines a linear isomorphism form $\mathfrak{g}$ to the tangent space T_yG. The exponential map $\exp : \mathfrak{g} \to G$ is certainly the most important construction associated with G and $\mathfrak{g}$ and is defined as follows. Given $u \in \mathfrak{g}$, let $\sigma_u : \mathbb{R} \to G$ be the one-parameter subgroup of G determined by the left invariant vector field $X_u : y \mapsto (L'_y)_e(u)$; that is, σ_u satisfies that

$$\begin{gathered} \sigma_u(0) = e, \\ \sigma'_u(t) = X_u(\sigma_u(t)) = \left(L'_{\sigma_u(t)}\right)_e(u) \quad \text{for each } t \in \mathbb{R}. \end{gathered} \tag{6}$$

The value of the exponential map exp at u is then defined by

$$\exp(u) = \sigma_u(1). \tag{7}$$

Moreover, we have that

$$\begin{gathered} \exp(tu) = \sigma_{tu}(1) = \sigma_u(t) \quad \text{for each } t \in \mathbb{R},\ u \in \mathfrak{g}, \\ \exp(t+s)u = \exp(tu) \cdot \exp(su) \quad \text{for any } t,\ s \in \mathbb{R},\ u \in \mathfrak{g}. \end{gathered} \tag{8}$$

Note that the exponential map is not surjective in general. However, the exponential map is a diffeomorphism on an open neighborhood of $0 \in \mathfrak{g}$. In the case when G is Abelian, exp is also a homomorphism from $\mathfrak{g}$ to G; that is,

$$\exp(u + v) = \exp(u) \cdot \exp(v), \quad \forall u, v \in \mathfrak{g}. \tag{9}$$

In the non-Abelian case, exp is not a homomorphism and, by the Baker-Campbell-Hausdorff (BCH) formula (cf. [21, page 114]), (9) must be replaced by

$$\exp(w) = \exp(u) \cdot \exp(v), \tag{10}$$

for all u, v in a neighborhood of $0 \in \mathfrak{g}$, where w is defined by

$$w := u + v + \frac{1}{2}[u, v] + \frac{1}{12}([u, [u, v]] + [v, [v, u]]) + \cdots. \tag{11}$$

Let $f : G \to \mathfrak{g}$ be a C^1-map and let $x \in G$. We use f'_x to denote the differential of f at x. Then, by [22, Page 9] (the proof given there for a smooth mapping still works for a C^1-map), for each $\Delta_x \in T_xG$ and any nontrivial smooth curve $c : (-\varepsilon, \varepsilon) \to G$ with $c(0) = x$ and $c'(0) = \Delta_x$, one has

$$f'_x \Delta_x = \left(\frac{d}{dt}(f \circ c)(t)\right)_{t=0}. \tag{12}$$

In particular,

$$f'_x \Delta_x = \left(\frac{d}{dt} f\left(x \cdot \exp\left(t\left(L'_{x^{-1}}\right)_x \Delta_x\right)\right)\right)_{t=0} \quad \text{for each } \Delta_x \in T_xG. \tag{13}$$

Define the linear map $df_x : \mathfrak{g} \to \mathfrak{g}$ by

$$df_x u = \left(\frac{d}{dt} f(x \cdot \exp(tu))\right)_{t=0} \quad \text{for each } u \in \mathfrak{g}. \tag{14}$$

Then, by (13),

$$df_x = f'_x \circ \left(L'_x\right)_e. \tag{15}$$

Also, in view of definition, we have that, for all $t \geq 0$,

$$\frac{d}{dt} f(x \cdot \exp(tu)) = df_{x \cdot \exp(tu)} u \quad \text{for each } u \in \mathfrak{g}, \tag{16}$$

$$f(x \cdot \exp(tu)) - f(x) = \int_0^t df_{x \cdot \exp(su)} u \, ds \quad \text{for each } u \in \mathfrak{g}. \tag{17}$$

For the remainder of the present paper, we always assume that $\langle \cdot, \cdot \rangle$ is an inner product on $\mathfrak{g}$ and $\| \cdot \|$ is the associated norm on $\mathfrak{g}$. We now introduce the following distance on G which plays a key role in the study. Let $x, y \in G$ and define

$$\varrho(x, y) := \inf\left\{\sum_{i=1}^{k} \|u_i\| \mid \text{there exist } k \geq 1 \text{ and } u_1, \ldots, u_k \in \mathfrak{g} \text{ such that } y = x \cdot \exp u_1 \cdots \exp u_k\right\}, \tag{18}$$

where we adapt the convention that $\inf \emptyset = +\infty$. It is easy to verify that $\varrho(\cdot, \cdot)$ is a distance on G and that the topology induced by this distance is equivalent to the original one on G.

Let $x \in G$ and $r > 0$. We denoted the corresponding ball of radius r around x of G by $C_r(x)$; that is,

$$C_r(x) := \{y \in G \mid \varrho(x, y) < r\}. \tag{19}$$

Let $\mathscr{L}(\mathfrak{g})$ denote the set of all linear operators on $\mathfrak{g}$. Below, we will modify the notion of the Lipschitz condition with L-average for mappings on Banach spaces to suit sections. Let L be a positive nondecreasing integrable function on $[0, R]$, where R is a positive number large enough such that $\int_0^R (R - s) L(s) ds \geq R$. The notion of Lipschitz condition in the inscribed sphere with the L average for operators on Banach spaces was first introduced in [23] by Wang for the study of Smale's point estimate theory.

Definition 1. Let $r > 0$, $x_0 \in G$ and let T be a mapping from G to $\mathscr{L}(\mathfrak{g})$. Then T is said to satisfy the L-average Lipschitz condition on $C_r(x_0)$ if

$$\|T(x \cdot \exp u) - T(x)\| \leq \int_{\varrho(x_0, x)}^{\varrho(x_0, x) + \|u\|} L(s) \, ds \tag{20}$$

holds for any $x \in C_r(x_0)$ and $u \in \mathfrak{g}$ such that $\|u\| < r - \varrho(x, x_0)$.

3. Uniqueness Ball of Zero Points of Mappings

This section is devoted to the study of uniqueness ball of zero points of mappings. Let $r > 0$. We use $\mathbf{B}(0, r)$ to denote the open ball at 0 with radius r on $\mathfrak{g}$; that is,

$$\mathbf{B}(0, r) := \{v \in \mathfrak{g} \mid \|v\| < r\}. \tag{21}$$

Write $N(x^*, r) := x^* \exp(\mathbf{B}(0, r))$. Clearly, $N(x^*, r) \subseteq C_r(x^*)$. Let $\hat{r} > 0$ be such that

$$\frac{1}{\hat{r}} \int_0^{\hat{r}} L(u)(\hat{r} - u) \, du = 1. \tag{22}$$

Theorem 2. *Let $0 < r \leq \hat{r}$. Suppose that $f(x^*) = 0$ and $df_{x^*}^{-1} df$ satisfies the L-average Lipschitz condition in $N(x^*, r)$. Then x^* is the unique zero point of f in $N(x^*, r)$.*

Proof. Let $y^* \in N(x^*, r)$ be another zero point of f in $N(x^*, r)$. Then, there exists $v \in \mathfrak{g}$ such that $y^* = x^* \exp v$ and $\|v\| < r$. As $L(\cdot)$ is a positive function, it follows from [6] that the function ψ defined by

$$\psi(t) = \frac{1}{t} \int_0^t L(s)(t - s) \, ds, \quad \forall t \in (0, \hat{r}] \tag{23}$$

is strictly monotonically increasing. set

$$\lambda := \frac{1}{\|v\|} \int_0^{\|v\|} L(s)(\|v\| - s) \, ds. \tag{24}$$

Then, by (22), we get

$$\lambda < \frac{1}{\hat{r}} \int_0^{\hat{r}} L(s)(\hat{r} - s) \, ds = 1. \tag{25}$$

To complete the proof, it suffices to show that

$$\|v\| \le \lambda \|v\|. \tag{26}$$

Granting this, one has that $x^* = y^*$. Now,

$$\begin{aligned}\|v\| &= \left\|-df_{x^*}^{-1}\left(f\left(y^*\right)-f\left(x^*\right)\right)+v\right\| \\ &\le \left\|-df_{x^*}^{-1}\int_0^1 df_{x^*\exp(sv)}v\,ds+v\right\| \\ &\le \int_0^1 \left\|df_{x^*}^{-1}\left(df_{x^*\exp(sv)}-df\left(x^*\right)\right)\right\|\|v\|\,ds \\ &\le \int_0^1\int_0^{s\|v\|} L(t)\,dt\,\|v\|\,ds \\ &= \int_0^{\|v\|} L(t)\left(\|v\|-t\right)dt = \lambda\|v\|,\end{aligned} \tag{27}$$

where the third inequality holds because of (20) by selecting $x = x_0 = x^*$. Therefore, (26) is seen to hold and the proof is completed □

4. Convergence Ball of Newton's Method

Following [17], we define Newton's method with initial point x_0 for f on a Lie group as follows:

$$x_{n+1} = x_n \cdot \exp\left(-df_{x_n}^{-1} f\left(x_n\right)\right) \quad \text{for each } n = 0, 1, \ldots. \tag{28}$$

Let $r_0 > 0$ and $b > 0$ be such that

$$\int_0^{r_0} L(s)\,ds = 1, \qquad b = \int_0^{r_0} L(s)\,s\,ds. \tag{29}$$

Remark 3. (i) Since $L(\cdot)$ is a positive function, we always have $b \le r_0$. Indeed,

$$\begin{aligned} b - r_0 &= \int_0^{r_0} L(s)\,s\,ds - r_0\int_0^{r_0} L(s)\,ds \\ &= \int_0^{r_0} L(s)\left(s-r_0\right)ds \le 0.\end{aligned} \tag{30}$$

(ii) Consider $r_0 \le \hat{r}$. Indeed, recall from [6] that the function ψ defined by

$$\psi(t) = \frac{1}{t}\int_0^t L(s)(t-s)\,ds, \quad \forall t \in (0, \hat{r}] \tag{31}$$

is strictly monotonically increasing. Sine $\psi(r_0) \le 1 = \psi(\hat{r})$, we get $r_0 \le \hat{r}$.

The following proposition plays a key role in this section, which is taken from [24].

Proposition 4. *Suppose that $x_0 \in G$ is such that $df_{x_0}^{-1}$ exists and $df_{x_0}^{-1}df$ satisfies the L-average Lipschitz condition on $C_{r_0}(x_0)$ and that*

$$\beta := \left\|df_{x_0}^{-1} f\left(x_0\right)\right\| \le b. \tag{32}$$

Then the sequence $\{x_n\}$ generated by Newton's method (28) *with initial point x_0 is well defined and converges to a zero point x^* of f and $\varrho(x^*, x_0) < r_0$.*

The remainder of this section is devoted to an estimate of the convergence domain of Newton's method on G around a zero x^* of f. Below we will always assume that $x^* \in G$ is such that $df_{x^*}^{-1}$ exists.

Lemma 5. *Let $0 < r \le r_0$ and let $x_0 \in C_r(x^*)$ be such that there exist $j \ge 1$ and $w_1, \ldots, w_j \in \mathfrak{g}$ satisfying*

$$x_0 = x^* \cdot \exp w_1 \cdots \exp w_j, \tag{33}$$

and $\rho(x^*, x_0) := \sum_{i=1}^{j}\|w_i\| < r$. Suppose that $df_{x^*}^{-1}df$ satisfies the L-average Lipschitz condition on $C_r(x^*)$. Then $df_{x_0}^{-1}$ exists,

$$\left\|df_{x_0}^{-1}df_{x^*}\right\| \le \frac{1}{1-\int_0^{\rho(x^*,x_0)} L(s)\,ds}, \tag{34}$$

$$\begin{aligned}&\left\|df_{x_0}^{-1} f\left(x_0\right)\right\| \\ &\quad\le \frac{\int_0^{\rho(x^*,x_0)} L(s)\left(\rho\left(x^*,x_0\right)-s\right)ds + \rho\left(x^*,x_0\right)}{1-\int_0^{\rho(x^*,x_0)} L(s)\,ds}.\end{aligned} \tag{35}$$

Proof. It follows from [24, Lemma 2.1] that $df_{x_0}^{-1}$ exists and (34) holds. Write $y_0 = x^*$, $y_i = y_{i-1}\cdot\exp w_i$ and $\rho_i := \sum_{l=1}^{i}\|w_l\|$ for each $i = 1, \ldots, j$. Thus, by (33), we have $y_j = x_0$ and so $\rho_j = \rho(x^*, x_0)$. Fix i, one has from (17) that

$$f\left(y_i\right) - f\left(y_{i-1}\right) = \int_0^1 df_{y_{i-1}\cdot\exp(\tau w_i)} w_i\,d\tau, \tag{36}$$

which implies that

$$\begin{aligned}&df_{x^*}^{-1}\left(f\left(y_i\right)-f\left(y_{i-1}\right)\right) \\ &\quad= \int_0^1 df_{x^*}^{-1}\left(df_{y_{i-1}\cdot\exp(\tau w_i)} - df_{x^*}\right)w_i\,d\tau + w_i.\end{aligned} \tag{37}$$

Since $df_{x^*}^{-1}df$ satisfies the L-average Lipschitz condition on $C_r(x^*)$, it follows that

$$\begin{aligned}&\left\|df_{x^*}^{-1}\left(df_{y_{i-1}\cdot\exp(\tau w_i)} - df_{x^*}\right)\right\| \\ &\quad\le \left\|df_{x^*}^{-1}\left(df_{y_{i-1}\cdot\exp(\tau w_i)} - df_{y_{i-1}}\right)\right\| \\ &\qquad+ \sum_{l=1}^{i-1}\left\|df_{x^*}^{-1}\left(df_{y_l} - df_{y_{l-1}}\right)\right\| \\ &\quad\le \int_{\rho_{i-1}}^{\rho_{i-1}+\tau\|w_i\|} L(s)\,ds + \sum_{l=1}^{i-1}\int_{\rho_{l-1}}^{\rho_l} L(s)\,ds \\ &\quad\le \int_0^{\rho_{i-1}+\tau\|w_i\|} L(s)\,ds,\end{aligned} \tag{38}$$

where $\rho_0 = 0$. Noting that $f(x_0) = \sum_{i=1}^{j}(f(y_i) - f(y_{i-1}))$, we have from (37) and (38) that

$$\begin{aligned}
&\left\|df_{x^*}^{-1} f(x_0)\right\| \\
&\le \sum_{i=1}^{j}\left(\int_0^1 \left\|df_{x^*}^{-1}\left(df_{y_{i-1}\cdot\exp(\tau w_i)} - df_{x^*}\right)\right\| \right. \\
&\qquad \left. \times \|w_i\|\, d\tau + \|w_i\|\right) \\
&\le \sum_{i=1}^{j}\left(\int_0^1 \int_0^{\rho_{i-1}+\tau\|w_i\|} L(s)\, ds\, \|w_i\|\, d\tau + \|w_i\|\right).
\end{aligned} \tag{39}$$

Write $F(t) := \int_0^t L(s)ds$. Then,

$$\begin{aligned}
&\int_0^1 \int_0^{\rho_{i-1}+\tau\|w_i\|} L(s)\, ds\, \|w_i\|\, d\tau \\
&\quad = \int_0^1 F(\rho_{i-1} + \tau\|w_i\|)\, \|w_i\|\, d\tau \\
&\quad = \int_{\rho_{i-1}}^{\rho_i} F(t)\, dt,
\end{aligned} \tag{40}$$

and so

$$\begin{aligned}
&\sum_{i=1}^{j}\int_0^1 \left\|df_{x^*}^{-1}\left(df_{y_{i-1}\cdot\exp(\tau w_i)} - df_{x^*}\right)\right\| \|w_i\|\, d\tau \\
&\quad \le \sum_{i=1}^{j}\int_{\rho_{i-1}}^{\rho_i} F(t)\, dt \\
&\quad = \int_0^{\rho(x^*,x_0)} F(t)\, dt \\
&\quad = \int_0^{\rho(x^*,x_0)} L(s)\left(\rho(x^*,x_0) - s\right) ds.
\end{aligned} \tag{41}$$

This, together with (39), yields that

$$\begin{aligned}
&\left\|df_{x^*}^{-1} f(x_0)\right\| \\
&\quad \le \int_0^{\rho(x^*,x_0)} L(s)\left(\rho(x^*,x_0) - s\right) ds + \rho(x^*,x_0).
\end{aligned} \tag{42}$$

Combining this with (34) implies that

$$\begin{aligned}
&\left\|df_{x_0}^{-1} f(x_0)\right\| \\
&\quad \le \left\|df_{x_0}^{-1} df_{x^*}\right\| \left\|df_{x^*}^{-1} f(x_0)\right\| \\
&\quad \le \frac{\int_0^{\rho(x^*,x_0)} L(s)\left(\rho(x^*,x_0) - s\right) ds + \rho(x^*,x_0)}{1 - \int_0^{\rho(x^*,x_0)} L(s)\, ds},
\end{aligned} \tag{43}$$

which completes the proof of the lemma. □

We make the following assumption throughout the remainder of the paper:

$$x^* \in G \quad \text{such that } f(x^*) = 0,\ df_{x^*}^{-1} \text{ exists.} \tag{44}$$

Theorem 6 below gives an estimation of convergence ball of Newton's method.

Theorem 6. *Suppose that $df_{x^*}^{-1}df$ satisfies the L-average Lipschitz condition on $C_{r_0}(x^*)$. Suppose that $\varrho(x^*, x_0) < (b/2)$. Then the sequence $\{x_n\}$ generated by Newton's method* (28) *with initial point x_0 is well defined and converges quadratically to a zero point y^* of f and $\varrho(y^*, x^*) < r_0$.*

Proof. Since $\varrho(x^*, x_0) < (b/2)$, there exist $j \ge 1$ and $w_1, \ldots, w_j \in \mathfrak{g}$ satisfying

$$x_0 = x^* \cdot \exp w_1 \cdots \exp w_j, \tag{45}$$

and $\rho(x^*, x_0) := \sum_{i=1}^{j}\|w_i\| < (b/2) \le r_0$, where the last inequality holds because of Remark 3(i). By Lemma 5, $df_{x_0}^{-1}$ exists and

$$\begin{aligned}
\beta &:= \left\|df_{x_0}^{-1} f(x_0)\right\| \\
&\le \frac{\int_0^{\rho(x^*,x_0)} L(s)\left(\rho(x^*,x_0) - s\right) ds + \rho(x^*,x_0)}{1 - \int_0^{\rho(x^*,x_0)} L(s)\, ds}.
\end{aligned} \tag{46}$$

Write

$$\overline{L}(s) := \frac{L(s + \rho(x^*,x_0))}{1 - \int_0^{\rho(x^*,x_0)} L(s)\, ds}. \tag{47}$$

Let $\overline{r}_0$, $\overline{b}$ be such that

$$\int_0^{\overline{r}_0} \overline{L}(s)\, ds = 1, \qquad \overline{b} = \int_0^{\overline{r}_0} \overline{L}(s)\, s\, ds. \tag{48}$$

This gives that

$$\begin{aligned}
&\int_0^{\overline{r}_0} L(s + \rho(x^*,x_0))\, ds \\
&\quad = 1 - \int_0^{\rho(x^*,x_0)} L(s)\, ds \\
&\quad = \int_0^{r_0} L(s)\, ds - \int_0^{\rho(x^*,x_0)} L(s)\, ds \\
&\quad = \int_{\rho(x^*,x_0)}^{r_0} L(s)\, ds.
\end{aligned} \tag{49}$$

Hence,

$$\int_{\rho(x^*,x_0)}^{\overline{r}_0+\rho(x^*,x_0)} L(s)\, ds = \int_{\rho(x^*,x_0)}^{r_0} L(s)\, ds. \tag{50}$$

As $L(\cdot)$ is a nondecreasing and positive integrable function, one has

$$\overline{r}_0 + \rho(x^*,x_0) = r_0. \tag{51}$$

Therefore,

$$\begin{aligned}\bar{b} &= \int_0^{\bar{r}_0} \bar{L}(s)\, s\, ds = \int_0^{r_0-\rho(x^*,x_0)} \bar{L}(s)\, s\, ds \\ &= \frac{\int_{\rho(x^*,x_0)}^{r_0} L(s)\left(s-\rho\left(x^*,x_0\right)\right) ds}{1-\int_0^{\rho(x^*,x_0)} L(s)\, ds}.\end{aligned} \tag{52}$$

Below, we will show that

$$\beta = \left\| df_{x_0}^{-1} f\left(x_0\right)\right\| \le \bar{b}. \tag{53}$$

To do this, by (46), it remains to show that

$$\begin{aligned}&\int_0^{\rho(x^*,x_0)} L(s)\left(\rho\left(x^*,x_0\right)-s\right) ds + \rho\left(x^*,x_0\right) \\ &\quad \le \int_{\rho(x^*,x_0)}^{r_0} L(s)\left(s-\rho\left(x^*,x_0\right)\right) ds;\end{aligned} \tag{54}$$

that is,

$$\begin{aligned}\rho\left(x^*,x_0\right) &\le \int_0^{r_0} L(s)\left(s-\rho\left(x^*,x_0\right)\right) ds \\ &= \int_0^{r_0} L(s)\, s\, ds - \rho\left(x^*,x_0\right)\int_0^{r_0} L(s)\, ds \\ &= b-\rho\left(x^*,x_0\right),\end{aligned} \tag{55}$$

which always holds because $\rho(x^*,x_0) \le (b/2)$ by assumption. Hence, (53) is seen to hold.

Then in order to ensure that Proposition 4 is applicable, we have to show the following assertion: $df_{x_0}^{-1}df$ satisfies the $\bar{L}$-average Lipschitz condition in $C_{\bar{r}_0}(x_0)$. To do this, let $x \in C_{\bar{r}_0}(x_0)$ be such that there exist $v, v_1, \ldots, v_l \in \mathfrak{g}$ satisfying $x = x_0 \exp v_1 \cdots \exp v_l$, $\rho(x_0,x) := \sum_{i=1}^{l} \|v_i\|$ and $\|v\| + \rho(x_0,x) < \bar{r}_0$. Since $df_{x^*}^{-1}df$ satisfies the L-average Lipschitz condition in $C_{r_0}(x_0)$ and

$$\rho\left(x^*,x_0\right) + \|v\| + \rho\left(x_0,x\right) < \rho\left(x^*,x_0\right) + \bar{r}_0 = r_0 \tag{56}$$

thanks to (51), we obtain that

$$\left\| df_{x^*}^{-1}\left(df_{x\exp v} - df(x)\right)\right\| \le \int_{\rho(x^*,x_0)+\rho(x_0,x)}^{\rho(x^*,x_0)+\rho(x_0,x)+\|v\|} L(s)\, ds. \tag{57}$$

Combining this with (34) yields that

$$\begin{aligned}&\left\| df_{x_0}^{-1}\left(df_{x\exp v} - df(x)\right)\right\| \\ &\quad \le \left\| df_{x_0}^{-1} df_{x^*}\right\| \left\| df_{x^*}^{-1}\left(df_{x\exp u} - df(x)\right)\right\| \\ &\quad \le \frac{1}{1-\int_0^{\rho(x^*,x_0)} L(s)\, ds}\int_{\rho(x^*,x_0)+\rho(x_0,x)}^{\rho(x^*,x_0)+\rho(x_0,x)+\|v\|} L(s)\, ds \\ &\quad = \int_{\rho(x_0,x)}^{\rho(x_0,x)+\|v\|} \frac{L\left(s+\rho\left(x^*,x_0\right)\right)}{1-\int_0^{\rho(x^*,x_0)} L(s)\, ds} ds \\ &\quad = \int_{\rho(x_0,x)}^{\rho(x_0,x)+\|v\|} \bar{L}(s)\, ds.\end{aligned} \tag{58}$$

Hence, $df_{x_0}^{-1}df$ satisfies the $\bar{L}$-average Lipschitz condition in $C_{\bar{r}_0}(x_0)$. Thus, we apply Proposition 4 to conclude that the sequence $\{x_n\}$ generated by Newton's method (28) with initial point x_0 is well defined and converges to a zero y^* of f and $\varrho(y^*,x_0) < \bar{r}_0$. And

$$\begin{aligned}\varrho\left(y^*,x^*\right) &\le \varrho\left(y^*,x_0\right) + \varrho\left(x^*,x_0\right) \\ &\le \rho\left(x^*,x_0\right) + \bar{r}_0 = r_0.\end{aligned} \tag{59}$$

The proof of the theorem is completed. □

Theorem 6 gives an estimate of the convergence domain for Newton's method. However, we do not know whether the limit y^* of the sequence generated by Newton's method with initial point x_0 from this domain is equal to the zero x^*. The following corollary provides the convergence domain from which the sequence generated by Newton's method with initial point x_0 converges to the zero x^*. Recall that e designates the identity element of G.

Corollary 7. *Suppose that $df_{x^*}^{-1}df$ satisfies the L-average Lipschitz condition on $C_{r_0}(x^*)$. Suppose that $C_{r_0}(e) \subseteq \exp(\mathbf{B}(0,r_0))$ and $\varrho(x^*,x_0) < (b/2)$. Then, the sequence $\{x_n\}$ generated by Newton's method* (28) *with initial point x_0 is well defined and converges quadratically to x^*.*

Proof. Since $\varrho(x^*,x_0) < (b/2)$, we apply Theorem 6 to conclude that the sequence $\{x_n\}$ generated by Newton's method (28) with initial point x_0 is well defined and converges quadratically to a zero point y^* of f and $\varrho(y^*,x^*) < r_0$; that is, $\varrho((x^*)^{-1}y^*,e) < r_0$. Since $C_{r_0}(e) \subseteq \exp(\mathbf{B}(0,r_0))$, there exists $u \in \mathfrak{g}$ such that $\|u\| \le r_0$ and $(x^*)^{-1}y^* = \exp u$; that is, $y^* = x^* \exp u$. Hence, $y^* \in N(x^*,r_0) := x^* \exp(\mathbf{B}(0,r_0))$. As $r_0 \le r_u$ by Remark 3(ii), Theorem 2 is applicable, and so $y^* = x^*$. □

Recall that in the special case when G is a compact connected Lie group, G has a bi-invariant Riemannian metric (cf. [22, page 46]). Below, we assume that G is a compact connected Lie group and endowed with a bi-invariant Riemannian metric. Therefore, an estimate of the convergence domain with the same property as in Corollary 7 is described in the following corollary.

Corollary 8. *Let G be a compact connected Lie group and endowed with a bi-invariant Riemannian metric. Suppose that $df_{x^*}^{-1}df$ satisfies the L-average Lipschitz condition on $C_{r_0}(x^*)$. Suppose that $\varrho(x^*,x_0) < (b/2)$. Then, the sequence $\{x_n\}$ generated by Newton's method* (28) *with initial point x_0 is well defined and converges quadratically to x^*.*

Proof. By Theorem 6, the sequence $\{x_n\}$ generated by Newton's method (28) with initial point x_0 is well defined and converges to a zero, say y^*, of f with $\varrho(x^*,y^*) < r_0$. Clearly, there is a minimizing geodesic c connecting $x^{*^{-1}} \cdot y^*$ and e. Since G is a compact connected Lie group and endowed with a bi-invariant Riemannian metric, it follows from [20, page 224] that c is also a one-parameter subgroup of G.

Consequently, there exists $u \in \mathfrak{g}$ such that $y^* = x^* \cdot \exp u$ and $\|u\| = \varrho(x^*, y^*) < r_0$. Hence, $y^* \in N(x^*, r_0) := x^* \exp(\mathbf{B}(0, r_0))$. As $r_0 \leq \hat{r}$ by Remark 3(ii), Theorem 2 is applicable, and so $y^* = x^*$. □

5. Theorems under the Kantorovich Condition and the γ-Condition

This section is devoted to the study of some applications of the results obtained in the preceding sections. At first, if $L(\cdot)$ is a constant, then the L-average Lipschitz condition is reduced to the classical Lipschitz condition.

Let $r > 0$, $x_0 \in G$ and let T be a mapping from G to $\mathscr{L}(\mathfrak{g})$. Then T is said to satisfy the L Lipschitz condition on $C_r(x_0)$ if

$$\|T(x \cdot \exp u) - T(x)\| \leq L\|u\| \tag{60}$$

holds for any $u, u_0, \ldots, u_k \in \mathfrak{g}$ and $x \in C_r(x_0)$ such that $x = x_0 \exp u_0 \exp u_1 \cdots \exp u_k$ and $\|u\| + \rho(x, x_0) < r$, where $\rho(x, x_0) = \sum_{i=0}^{k} \|u_i\|$.

Hence, in the case when $L(\cdot) \equiv L$, we obtain from (22) and (29) that

$$r_0 = \frac{1}{L}, \qquad b = \frac{1}{2L}, \qquad \hat{r} = \frac{2}{L}. \tag{61}$$

Thus, by Theorems 2 and 6, we have the following results, where Theorem 10 has been given in [18].

Theorem 9. *Let $0 < r \leq (2/L)$, Suppose that $df_{x^*}^{-1} df$ satisfies the L Lipschitz condition in $N(x^*, r)$. Then x^* is the unique zero point of f in $N(x^*, r)$.*

Theorem 10. *Suppose that $df_{x^*}^{-1} df$ satisfies the L Lipschitz condition on $C_{1/L}(x^*)$. Suppose that $\varrho(x^*, x_0) < (1/4L)$. Then the sequence $\{x_n\}$ generated by Newton's method (28) with initial point x_0 is well defined and converges quadratically to a zero y^* of f and $\varrho(y^*, x^*) < (1/L)$.*

Furthermore, by Corollaries 7 and 8, one has the following results.

Corollary 11. *Suppose that $df_{x^*}^{-1} df$ satisfies the L Lipschitz condition on $C_{1/L}(x^*)$. Suppose that $C_{1/L}(e) \subseteq \exp(\mathbf{B}(0, (1/L)))$ and $\varrho(x^*, x_0) < (1/4L)$. Then, the sequence $\{x_n\}$ generated by Newton's method (28) with initial point x_0 is well defined and converges quadratically to x^*.*

Corollary 12. *Let G be a compact connected Lie group and endowed with a bi-invariant Riemannian metric. Suppose that $df_{x^*}^{-1} df$ satisfies the L Lipschitz condition on $C_{1/L}(x^*)$. Suppose that $\varrho(x^*, x_0) < (1/4L)$. Then, the sequence $\{x_n\}$ generated by Newton's method (28) with initial point x_0 is well defined and converges quadratically to x^*.*

Let k be a positive integer, and assume further that $f : G \to \mathfrak{g}$ is a C^k-map. Define the map $d^k f_x : \mathfrak{g}^k \to \mathfrak{g}$ by

$$d^k f_x u_1 \cdots u_k = \left(\frac{\partial^k}{\partial t_k \cdots \partial t_1} f(x \cdot \exp t_k u_k \cdots \exp t_1 u_1) \right)_{t_k = \cdots = t_1 = 0}, \tag{62}$$

for each $(u_1, \ldots, u_k) \in \mathfrak{g}^k$. In particular,

$$d^k f_x u^k = \left(\frac{d^k}{dt^k} f(x \cdot \exp tu) \right)_{t=0} \quad \text{for each } u \in \mathfrak{g}. \tag{63}$$

Let $1 \leq i \leq k$. Then, in view of the definition, one has

$$d^k f_x u_1 \cdots u_k = d^{k-i}\left(d^i f \cdot (u_1 \cdots u_i)\right)_x u_{i+1} \cdots u_k \quad \text{for each } (u_1, \ldots, u_k) \in \mathfrak{g}^k. \tag{64}$$

In particular, for fixed $u_1, \ldots, u_{i-1}, u_{i+1}, \ldots, u_k \in \mathfrak{g}$,

$$d^i f_x u_1 \cdots u_{i-1}(u) = d\left(d^{i-1} f(u_1 \cdots u_{i-1})\right)(u) \quad \text{for each } u \in \mathfrak{g}. \tag{65}$$

This implies that $d^i f_x u_1 \cdots u_{i-1} u$ is linear with respect to $u \in \mathfrak{g}$ and so is $d^k f_x u_1 \cdots u_{i-1} u u_{i+1} \cdots u_k$. Consequently, $d^k f_x$ is a multilinear map from $\mathfrak{g}^k$ to $\mathfrak{g}$ because $1 \leq i \leq k$ is arbitrary. Thus, we can define the norm of $d^k f_x$ by

$$\|d^k f_x\| := \sup\left\{ \|d^k f_x u_1 u_2 \cdots u_k\| \mid (u_1, \ldots, u_k) \in \mathfrak{g}^k \text{ with each } \|u_j\| = 1 \right\}. \tag{66}$$

For the remainder of the paper, we always assume that f is a C^2-map from G to $\mathfrak{g}$. Then taking $i = 2$, we have

$$d^2 f_z v u = d(df \cdot v)_z u \quad \text{for any } u, v \in \mathfrak{g} \text{ and each } z \in G. \tag{67}$$

Thus, (17) is applied (with $df \cdot v$ in place of $f(\cdot)$ for each $v \in \mathfrak{g}$) to conclude the following formula:

$$df_{x \cdot \exp(tu)} - df_x = \int_0^t d^2 f_{x \cdot \exp(su)} u \, ds \quad \text{for each } u \in \mathfrak{g},\ t \in \mathbb{R}. \tag{68}$$

The γ-condition for nonlinear operators in Banach spaces was first introduced by Wang and Han [25], and explored further by Wang [26] to study Smale's point estimate theory, which has been extended in [19] for a map f from a Lie group to its Lie algebra in view of the map $d^2 f$ as given in Definition 13 below. Let $r > 0$ and $\gamma > 0$ be such that $\gamma r \leq 1$.

Definition 13. Let $x_0 \in G$ be such that $df_{x_0}^{-1}$ exists. f is said to satisfy the γ-condition at x_0 on $C_r(x_0)$ if, for any $x \in C_r(x_0)$ with $x = x_0 \exp u_0 \exp u_1 \cdots \exp u_k$ such that $\rho(x, x_0) := \sum_{i=0}^{k} \|u_i\| < r$,

$$\|df_{x_0}^{-1} d^2 f_x\| \leq \frac{2\gamma}{(1 - \gamma\rho(x, x_0))^3}. \tag{69}$$

As shown in Proposition 20, if f is analytic at x_0, then f satisfies the γ-condition at x_0.

Let $\gamma > 0$ and let L be the function defined by

$$L(s) = \frac{2\gamma}{(1-\gamma s)^3} \quad \text{for each } 0 < s < \frac{1}{\gamma}. \tag{70}$$

The following proposition shows that the γ-condition implies the L-average Lipschitz condition, which is taken from [24].

Proposition 14. *Suppose that f satisfies the γ-condition at x_0 on $C_r(x_0)$. Then $df_{x_0}^{-1}df$ satisfies the L-average Lipschitz condition on $C_r(x_0)$ with L being defined by* (70).

In the case when L is given by (70), we have from (22) and (29) that

$$r_0 = \frac{2-\sqrt{2}}{2\gamma}, \qquad b = \frac{3-2\sqrt{2}}{\gamma}, \qquad \hat{r} = \frac{1}{2\gamma}. \tag{71}$$

Thus, by Theorems 2 and 6, we have the following results.

Theorem 15. *Let $0 < r \le (1/2\gamma)$. Suppose that f satisfies the γ-condition in $N(x^*, r) := x^* \exp(\mathbf{B}(0, r))$. Then x^* is the unique zero point of f in $N(x^*, r)$.*

Theorem 16. *Suppose that f satisfies the γ-condition on $C_{(2-\sqrt{2})/2\gamma}(x^*)$. Suppose that $\varrho(x^*, x_0) < (3-2\sqrt{2})/2\gamma$. Then the sequence $\{x_n\}$ generated by Newton's method* (28) *with initial point x_0 is well defined and converges quadratically to a zero point y^* of f and $\varrho(y^*, x^*) < (2-\sqrt{2})/2\gamma$.*

Remark 17. Theorem 16 improves the corresponding results in [19, Corollary 4.1], where it was proved under the following assumption: there exists $v \in \mathfrak{g}$ such that $x_0 = x^* \exp v$ and $\|v\| \le (a_0/\gamma)$ with $a_0 = 0.081256\ldots$ being the smallest positive root of the equation $a_0/(1-4a_0+2a_0^2)^2 = 3-2\sqrt{2}$. Clearly, $(a_0/\gamma) < (3-2\sqrt{2})/2\gamma$. Note also that in general, there dose not exist $v \in \mathfrak{g}$ satisfying $x_0 = x^* \exp v$ because the exponential map is not surjective global, even if $\varrho(x^*, x_0) < (3-2\sqrt{2})/2\gamma$. In view of this, our results somewhat improves the corresponding results in [19, Corollary 4.1].

Moreover, we get the following two corollaries from Corollaries 7 and 8.

Corollary 18. *Suppose that f satisfies the γ-condition on $C_{(2-\sqrt{2})/2\gamma}(x^*)$. Suppose that $C_{(2-\sqrt{2})/2\gamma}(e) \subseteq \exp(\mathbf{B}(0, (2-\sqrt{2})/2\gamma))$ and $\varrho(x^*, x_0) < (3-2\sqrt{2})/2\gamma$. Then, the sequence $\{x_n\}$ generated by Newton's method* (28) *with initial point x_0 is well defined and converges quadratically to x^*.*

Corollary 19. *Let G be a compact connected Lie group and endowed with a bi-invariant Riemannian metric. Suppose that f satisfies the γ-condition on $C_{(2-\sqrt{2})/2\gamma}(x^*)$. Suppose that $\varrho(x^*, x_0) < (3-2\sqrt{2})/2\gamma$. Then, the sequence $\{x_n\}$ generated by Newton's method* (28) *with initial point x_0 is well defined and converges quadratically to x^*.*

6. Applications to Analytic Maps

Throughout this section, we always assume that f is analytic on G. For $x \in G$ such that df_x^{-1} exists, we define

$$\gamma_x := \gamma(f, x) = \sup_{i \ge 2} \left\| \frac{df_x^{-1} d^i f_x}{i!} \right\|^{1/(i-1)}. \tag{72}$$

Also we adopt the convention that $\gamma(f, x) = \infty$ if df_x is not invertible. Note that this definition is justified, and, in the case when df_x is invertible, $\gamma(f, x)$ is finite by analyticity.

The following proposition is taken from [19].

Proposition 20. *Let $\gamma_{x^*} := \gamma(f, x^*)$ and let $r = (2-\sqrt{2})/2\gamma_{x^*}$. Then f satisfies the γ_{x^*}-condition at x^* on $C_r(x^*)$.*

Thus, by Theorems 15 and 16 and Proposition 20, we have the following results.

Theorem 21. *Let $0 < r \le (1/2\gamma_{x^*})$. Then x^* is the unique zero point of f in $N(x^*, r)$.*

Theorem 22. *Suppose that $\varrho(x^*, x_0) < (3-2\sqrt{2})/2\gamma_{x^*}$. Then the sequence $\{x_n\}$ generated by Newton's method* (28) *with initial point x_0 is well defined and converges quadratically to a zero point y^* of f and $\varrho(y^*, x^*) < (2-\sqrt{2})/2\gamma_{x^*}$.*

Moreover, we get the following two corollaries from Corollaries 7 and 8 and Proposition 20.

Corollary 23. *Suppose that $C_{(2-\sqrt{2})/2\gamma_{x^*}}(e) \subseteq \exp(\mathbf{B}(0, (2-\sqrt{2})/2\gamma_{x^*}))$ and $\varrho(x^*, x_0) < ((3-2\sqrt{2})/2\gamma_{x^*})$. Then, the sequence $\{x_n\}$ generated by Newton's method* (28) *with initial point x_0 is well defined and converges quadratically to x^*.*

Corollary 24. *Let G be a compact connected Lie group and endowed with a bi-invariant Riemannian metric. Suppose that $\varrho(x^*, x_0) < (3-2\sqrt{2})/2\gamma_{x^*}$. Then, the sequence $\{x_n\}$ generated by Newton's method* (28) *with initial point x_0 is well defined and converges quadratically to x^*.*

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The research of the second author was partially supported by the National Natural Science Foundation of China (Grant nos. 11001241 and 11371325) and by Zhejiang Provincial Natural Science Foundation of China (Grant no. LY13A010011). The research of the third author was partially supported by a Grant from NSC of Taiwan (NSC 102-2221-E-037-004-MY3).

References

[1] L. V. Kantorovich and G. P. Akilov, *Functional Analysis*, Pergamon Press, Oxford, UK, 2nd edition, 1982.

[2] S. Smale, "Newton's method estimates from data at one point," in *The Merging of Disciplines: New Directions in Pure, Applied, and Computational Mathematics (Laramie, Wyo., 1985)*, R. Ewing, K. Gross, and C. Martin, Eds., pp. 185–196, Springer, New York, NY, USA, 1986.

[3] J. A. Ezquerro and M. A. Hernández, "Generalized differentiability conditions for Newton's method," *IMA Journal of Numerical Analysis*, vol. 22, no. 2, pp. 187–205, 2002.

[4] J. A. Ezquerro and M. A. Hernández, "On an application of Newton's method to nonlinear operators with w-conditioned second derivative," *BIT Numerical Mathematics*, vol. 42, no. 3, pp. 519–530, 2002.

[5] J. M. Gutiérrez and M. A. Hernández, "Newton's method under weak Kantorovich conditions," *IMA Journal of Numerical Analysis*, vol. 20, no. 4, pp. 521–532, 2000.

[6] X. Wang, "Convergence of Newton's method and uniqueness of the solution of equations in Banach space," *IMA Journal of Numerical Analysis*, vol. 20, no. 1, pp. 123–134, 2000.

[7] P. P. Zabrejko and D. F. Nguen, "The majorant method in the theory of Newton-Kantorovich approximations and the Pták error estimates," *Numerical Functional Analysis and Optimization*, vol. 9, no. 5-6, pp. 671–684, 1987.

[8] O. P. Ferreira and B. F. Svaiter, "Kantorovich's theorem on Newton's method in Riemannian manifolds," *Journal of Complexity*, vol. 18, no. 1, pp. 304–329, 2002.

[9] J.-P. Dedieu, P. Priouret, and G. Malajovich, "Newton's method on Riemannian manifolds: convariant alpha theory," *IMA Journal of Numerical Analysis*, vol. 23, no. 3, pp. 395–419, 2003.

[10] C. Li and J. Wang, "Newton's method on Riemannian manifolds: Smale's point estimate theory under the γ-condition," *IMA Journal of Numerical Analysis*, vol. 26, no. 2, pp. 228–251, 2006.

[11] J. Wang and C. Li, "Uniqueness of the singular points of vector fields on Riemannian manifolds under the γ-condition," *Journal of Complexity*, vol. 22, no. 4, pp. 533–548, 2006.

[12] C. Li and J. Wang, "Convergence of the Newton method and uniqueness of zeros of vector fields on Riemannian manifolds," *Science in China. Series A*, vol. 48, no. 11, pp. 1465–1478, 2005.

[13] J. H. Wang, "Convergence of Newton's method for sections on Riemannian manifolds," *Journal of Optimization Theory and Applications*, vol. 148, no. 1, pp. 125–145, 2011.

[14] C. Li and J. Wang, "Newton's method for sections on Riemannian manifolds: generalized covariant α-theory," *Journal of Complexity*, vol. 24, no. 3, pp. 423–451, 2008.

[15] F. Alvarez, J. Bolte, and J. Munier, "A unifying local convergence result for Newton's method in Riemannian manifolds," *Foundations of Computational Mathematics*, vol. 8, no. 2, pp. 197–226, 2008.

[16] R. E. Mahony, "The constrained Newton method on a Lie group and the symmetric eigenvalue problem," *Linear Algebra and Its Applications*, vol. 248, pp. 67–89, 1996.

[17] B. Owren and B. Welfert, "The Newton iteration on Lie groups," *BIT Numerical Mathematics*, vol. 40, no. 1, pp. 121–145, 2000.

[18] J.-H. Wang and C. Li, "Kantorovich's theorems for Newton's method for mappings and optimization problems on Lie groups," *IMA Journal of Numerical Analysis*, vol. 31, no. 1, pp. 322–347, 2011.

[19] C. Li, J.-H. Wang, and J.-P. Dedieu, "Smale's point estimate theory for Newton's method on Lie groups," *Journal of Complexity*, vol. 25, no. 2, pp. 128–151, 2009.

[20] S. Helgason, *Differential Geometry, Lie Groups, and Symmetric Spaces*, vol. 80 of *Pure and Applied Mathematics*, Academic Press, New York, NY, USA, 1978.

[21] V. S. Varadarajan, *Lie Groups, Lie Algebras, and Their Representations*, vol. 102 of *Graduate Texts in Mathematics*, Springer, New York, NY, USA, 1984.

[22] M. P. do Carmo, *Riemannian Geometry*, Mathematics: Theory & Applications, Birkhäuser, Boston, Mass, USA, 1992.

[23] X. Wang, "Convergence of Newton's method and inverse function theorem in Banach space," *Mathematics of Computation*, vol. 68, no. 225, pp. 169–186, 1999.

[24] J. He, J. H. Wang, and J. C. Yao, "Convergence criteria of Newton's method on Lie groups," *Fixed Point Theory and Applications*, to appear.

[25] X. H. Wang and D. F. Han, "Criterion α and Newton's method under weak conditions," *Chinese Journal of Numerical Mathematics and Applications*, vol. 19, no. 2, pp. 96–105, 1997.

[26] X. Wang, "Convergence on the iteration of Halley family in weak conditions," *Chinese Science Bulletin*, vol. 42, no. 7, pp. 552–555, 1997.