Research Article

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

Jinsu He, Jinhua Wang, and Jen-Chih Yao

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 \( x_0 \in D \),

\[
x_{n+1} = x_n - f'(x_n)^{-1} f(x_n), \quad n = 0, 1, \ldots
\]

(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., \( \alpha \)-theory and \( \gamma \)-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 \( \alpha \)-theory and \( \gamma \)-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 \( \gamma \)-condition for the vector fields and mappings on Riemannian manifolds. The radii of uniqueness balls of singular points of vector fields satisfying the \( \gamma \)-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 solve 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 a unified estimation of uniqueness ball under the Kantorovich condition and the $\gamma$-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. More precisely, under the assumption that $\mathcal{G}$ is Abelian, $\exp$ is also a homomorphism from $\mathfrak{g}$ to $\mathcal{G}$; that is,

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

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 $\gamma$-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 $(\mathcal{G}, \cdot)$ 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 $\mathcal{G}$. Let $q$ be the Lie algebra of the Lie group $\mathcal{G}$ which is the tangent space $T_e \mathcal{G}$ of $\mathcal{G}$ 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 $\mathcal{G}$. We define for each $y \in \mathcal{G}$ the left translation $L_y : \mathcal{G} \to \mathcal{G}$ by

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

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

$$
\sigma_u(0) = e,
$$

$$
\sigma_u(t) = X_u(\sigma_u(t)) = (L'_{\sigma_u(t)})_e(u) \quad \text{for each } t \in \mathbb{R}.
$$

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

$$
\exp(u) = \sigma_u(1).
$$

Moreover, we have that

$$
\exp(tu) = \sigma_{tu}(1) = \sigma_u(t) \quad \text{for each } t \in \mathbb{R}, \ u \in \mathfrak{g},
$$

$$
\exp(t + s)u = \exp(tu) \cdot \exp(su) \quad \text{for any } t, \ s \in \mathbb{R}, \ u \in \mathfrak{g}.
$$

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 $\mathcal{G}$ is Abelian, $\exp$ is also a homomorphism from $\mathfrak{g}$ to $\mathcal{G}$; that is,

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

Note also that in general, there does not exist $v \in \mathfrak{g}$ satisfying $x_0 = x^* \exp v$ because the exponential map is not surjective global, even if $\phi(x^*, x_0) < (3 - 2\sqrt{2})/2\gamma$ in view of this, our results somewhat improve the corresponding results in [19, Corollary 4.1].
In the non-Abelian case, \(\exp\) is not a homomorphism and, by the Baker-Campbell-Hausdorff (BCH) formula (cf. \cite[page 114]{21}), (9) must be replaced by

\[
\exp(w) = \exp(u) \cdot \exp(v),
\]

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

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

In particular,

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

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 \(\Delta_x \in T_x G\) and any nontrivial smooth curve \(c : (-\epsilon, \epsilon) \to G\) with \(c(0) = x\) and \(c'(0) = \Delta_x\), one has

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

Then, by (22), we get

\[
\lambda < 1 \int_{0}^{\hat{r}} L(\tilde{r}) (\tilde{r} - s) ds \leq 1.
\]

Let \(L(g)\) denote the set of all linear operators on \(g\). 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 \(L(g)\). Then \(T\) is said to satisfy the \(L\)-average Lipschitz condition on \(C_r(x_0)\) if

\[
\left\| T(x \cdot \exp u) - T(x) \right\| \leq \int_{g(x, x_0)} L(s) ds
\]

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

### 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 \(B(0, r)\) to denote the open ball at 0 with radius \(r\) on \(g\); that is, 

\[
B(0, r) := \{ v \in g \mid \| v \| < r \}.
\]

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

\[
\frac{1}{\bar{r}} \int_{0}^{\bar{r}} L(u) (\bar{r} - u) du = 1.
\]

**Theorem 2.** Let \(0 < r \leq \bar{r}\). 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 \(v \in g\) such that \(y^* = x^* \exp \nu \) and \(\|\nu\| < r\). As \(L(\cdot)\) is a positive function, it follows from [6] that the function \(\psi\) defined by

\[
\psi(t) = \frac{1}{t} \int_{0}^{t} L(s)(t - s) ds, \quad \forall t \in (0, \bar{r}]
\]

is strictly monotonically increasing. set

\[
\lambda := \frac{1}{\|\nu\|} \int_{0}^{\|\nu\|} L(s)(\|\nu\| - s) ds.
\]

Then, by (22), we get

\[
\lambda < \frac{1}{\bar{r}} \int_{0}^{\bar{r}} L(s)(\bar{r} - s) ds = 1.
\]
To complete the proof, it suffices to show that
\[ \|v\| \leq \lambda \|v\|. \] (26)
Granting this, one has that \( x^* = y^* \). Now,
\[
\|v\| = \left\| -d_{x_{0}}^{-1} f'(y^*) + v \right\|
\leq \left\| -d_{x_{0}}^{-1} \int_{0}^{1} d_{x_{0}}^{-1} f'(x_{0} \exp(tv)) v \, dt + v \right\|
\leq \int_{0}^{1} \left\| d_{x_{0}}^{-1} \left( d_{x_{0}}^{-1} f'(x_{0}) - d_{x_{0}}^{-1} f'(x_{0}) \right) \right\| \|v\| \, dt
\leq \int_{0}^{1} \int_{0}^{1} [L(t) \|v\|] \, dt \|v\| ds
\]
(27)
\[
= \int_{0}^{1} \int_{0}^{1} L(t) (\|v\| - t) \, dt = \lambda \|v\|,
\]
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:
\[
x_{n+1} = x_n \cdot \exp \left( -d_{x_n}^{-1} f(x_n) \right) \quad \text{for each } n = 0, 1, \ldots \tag{28}
\]
Let \( r_0 > 0 \) and \( b > 0 \) be such that
\[
\int_{0}^{r_0} L(s) \, ds = 1, \quad b = \int_{0}^{r_0} L(s) \, ds. \tag{29}
\]

**Remark 3.** (i) Since \( L(\cdot) \) is a positive function, we always have \( b \leq r_0 \). Indeed,
\[
b - r_0 = \int_{0}^{r_0} L(s) \, ds - r_0 \int_{0}^{1} L(s) \, ds
\]
\[
= \int_{0}^{r_0} L(s) (s - r_0) \, ds \leq 0. \tag{30}
\]
(ii) Consider \( r_0 \leq \bar{r} \). Indeed, recall from [6] that the function \( \psi \) defined by
\[
\psi(t) = \frac{1}{t} \int_{0}^{t} L(s) (t - s) \, ds, \quad \forall t \in (0, \bar{r}] \tag{31}
\]
is strictly monotonically increasing. Since \( \psi(r_0) \leq 1 = \psi(\bar{r}) \), we get \( r_0 \leq \bar{r} \).

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 \( d_{x_0}^{-1} f \) exists and \( d_{x_0}^{-1} f \) satisfies the \( L \)-average Lipschitz condition on \( C_{r_0}(x_0) \) and that
\[
\beta := \|d_{x_0}^{-1} f(x_0)\| \leq b. \tag{32}
\]
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 \( \rho(x^*, x_0) < r_0 \).

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 \( d_{x^*}^{-1} f \) exists.

**Lemma 5.** Let \( 0 < r \leq r_0 \) and let \( x_0 \in C_r(x^*) \) be such that there exist \( j \geq 1 \) 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 \( d_{x_0}^{-1} f \) satisfies the \( L \)-average Lipschitz condition on \( C_r(x^*) \). Then \( d_{x_0}^{-1} f \) exists,
\[
\|d_{x_0}^{-1} f(x_0)\| \leq \frac{1}{1 - \int_{0}^{r_0} L(s) \, ds} - 1 \tag{34}
\]
\[
\|d_{x_0}^{-1} f(x_0)\| \leq \frac{\rho(x^*, x_0)}{1 - \int_{0}^{r_0} L(s) \, ds}. \tag{35}
\]

**Proof.** It follows from [24, Lemma 2.1] that \( d_{x_0}^{-1} f \) exists and (34) holds. Write \( y_0 = x^* \), \( y_j = y_{j-1} \cdot \exp w_j \) and \( \rho_j := \sum_{i=1}^{j} \|w_i\| \) 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
\[
f(y_i) - f(y_{i-1}) = \int_{0}^{1} d_{y_{j-1} \cdot \exp(tw_j)} w_i \, dt, \tag{36}
\]
which implies that
\[
d_{x_0}^{-1} f(y_i) - d_{x_0}^{-1} f(y_{i-1})
\]
\[
= \int_{0}^{1} d_{x_0}^{-1} \left( d_{y_{j-1} \cdot \exp(tw_j)} - d_{y_{j-1}} \right) w_i \, dt + w_i. \tag{37}
\]
Since \( d_{x_0}^{-1} f \) satisfies the \( L \)-average Lipschitz condition on \( C_r(x^*) \), it follows that
\[
\|d_{x_0}^{-1} (d_{y_{j-1} \cdot \exp(tw_j)} - d_{y_{j-1}})\|
\]
\[
\leq \|d_{x_0}^{-1} (d_{y_{j-1} \cdot \exp(tw_j)} - d_{y_{j-1}})\|
\]
\[
+ \sum_{i=1}^{j} \|d_{x_0}^{-1} (d_{y_j} - d_{y_{j-1}})\|
\]
\[
\leq \int_{0}^{r_0} L(s) \, ds + \sum_{i=1}^{j} \int_{0}^{r_0} L(s) \, ds
\]
\[
\leq \sum_{i=1}^{j} \int_{0}^{r_0} L(s) \, ds,
\]
\[
\leq \rho_j \cdot r \tag{38}
\]
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

$$
\|df^{-1}_{x_0} f(x_0)\|
\leq \sum_{i=1}^{j} \left( \int_0^{\rho_{0} + \tau} \left\| f(y_i) - f(y_{i-1}) \right\| \, d\tau \right) \, \|w_i\|
\leq \sum_{i=1}^{j} \left( \int_0^{\rho_i} F(s) \, ds \right) \|w_i\|
\leq \sum_{i=1}^{j} \left( \int_0^{\rho_i} F(s) \, ds \right) \frac{\|w_i\|}{\rho_i},
$$

(39)

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

$$
\int_0^1 \int_0^{\rho_0 + \tau} L(s) \, ds \, \|w_i\| \, d\tau
= \int_0^1 F(\rho_{0} + \tau) \|w_i\| \, d\tau
= \int_0^{\rho_1} F(t) \, dt,
$$

(40)

and so

$$
\sum_{i=1}^{j} \int_0^{\rho_i} F(t) \, dt
= \int_0^{\rho(x^*, x_0)} L(s) \, ds + \rho(x^*, x_0) \, ds.
$$

(41)

This, together with (39), yields that

$$
\|df^{-1}_{x_0} f(x_0)\|
\leq \int_0^{\rho(x^*, x_0)} L(s) \, ds + \rho(x^*, x_0).
$$

(42)

Combining this with (34) implies that

$$
\|df^{-1}_{x_0} f(x_0)\|
\leq \left\| df^{-1}_{x_0} df \right\| \|df^{-1}_{x_0} f(x_0)\|
\leq \int_0^{\rho(x^*, x_0)} L(s) \, ds + \rho(x^*, x_0)
\leq \frac{\int_0^{\rho(x^*, x_0)} L(s) \, ds + \rho(x^*, x_0)}{1 - \int_0^{\rho(x^*, x_0)} L(s) \, ds},
$$

(43)

which completes the proof of the lemma.

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

$$
x^* \in G \text{ such that } f(x^*) = 0, \, df^{-1}_{x^*} \text{ exists.}
$$

(44)

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

**Theorem 6.** Suppose that $df^{-1}_{x^*} df$ 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 $j \geq 1$ and $w_1, \ldots, w_j \in q$ satisfying

$$
x_0 = x^* \cdot \exp w_1 \cdots \exp w_j,
$$

(45)

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

$$
\beta := \left\| df^{-1}_{x_0} f(x_0) \right\|
\leq \int_0^{\rho(x^*, x_0)} L(s) \left( \rho(x^*, x_0) - s \right) \, ds + \rho(x^*, x_0).
$$

(46)

Write

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

(47)

Let $\tau_0, \bar{b}$ be such that

$$
\int_0^{\tau_0} \bar{L}(s) \, ds = 1, \quad \bar{b} = \int_0^{\tau_0} \bar{L}(s) \, ds.
$$

(48)

This gives that

$$
\int_0^{\bar{b}} \bar{L}(s) \, ds = 1
\quad \int_0^{\tau_0} \bar{L}(s) \, ds
= 1 - \int_0^{\rho(x^*, x_0)} L(s) \, ds
= \int_0^{\rho(x^*, x_0)} L(s) \, ds.
$$

(49)

Hence,

$$
\int_{\tau_0}^{\tau_0 + \rho(x^*, x_0)} L(s) \, ds = \int_{\bar{b}}^{\tau_0} \bar{L}(s) \, ds.
$$

(50)

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

$$
\bar{b} + \rho(x^*, x_0) = r_0.
$$

(51)
Therefore,

\[ b = \int_0^r L(s) s \, ds = \int_0^r \frac{\rho(x^*, x_0)}{L(s)} L(s) s \, ds \]

\[ = \int_0^r \frac{\rho(x^*, x_0)}{1 - \int_0^{\rho(x^*, x_0)} L(s) \, ds} L(s) s \, ds. \tag{52} \]

Below, we will show that

\[ \beta = \|df_{x_0}^{-1} f(x_0)\| \leq b. \tag{53} \]

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

\[
\int_0^{\rho(x^*, x_0)} L(s) (s - \rho(x^*, x_0)) \, ds + \rho(x^*, x_0) \leq \int_0^{\rho(x^*, x_0)} L(s) (s - \rho(x^*, x_0)) \, ds; \tag{54}
\]

that is,

\[
\rho(x^*, x_0) \leq \int_0^{\rho(x^*, x_0)} L(s) (s - \rho(x^*, x_0)) \, ds \]

\[ = \int_0^{\rho(x^*, x_0)} L(s) \, ds \rho(x^*, x_0) \int_0^{\rho(x^*, x_0)} L(s) \, ds \]

\[ = b - \rho(x^*, x_0), \]

which always holds because \( \rho(x^*, x_0) \leq (b/2) \) 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 \( L \)-average Lipschitz condition in \( C_r(x_0) \). To do this, let \( x \in C_r(x_0) \) be such that there exist \( v, v_1, \ldots, v_l \) satisfying \( x = x_0 \exp v \cdot \exp v_1 \cdots \exp v_l \), \( \rho(x_0, x) := \sum_{i=1}^{l} \|v_i\| + \|v\| + \rho(x_0, x) < \tau_0 \). Since \( df_{x_0}^{-1} df \) satisfies the \( L \)-average Lipschitz condition in \( C_r(x_0) \) and

\[
\rho(x^*, x_0) + \|v\| + \rho(x_0, x) < \rho(x^*, x_0) + \tau_0 = r_0 \tag{56}
\]

thanks to (51), we obtain that

\[
\|df_{x_0}^{-1}(df_{x \exp} \cdot df(x))\| \leq \int_0^{\rho(x^*, x_0) + \rho(x_0, x) + \|v\|} L(s) \, ds. \tag{57}
\]

Combining this with (34) yields that

\[
\|df_{x_0}^{-1}(df_{x \exp} \cdot df(x))\| \leq \int_0^{\rho(x^*, x_0) + \rho(x_0, x) + \|v\|} L(s) \, ds.
\]

\[ \leq 1 - \int_0^{\rho(x^*, x_0) + \rho(x_0, x) + \|v\|} L(s) \, ds \]

\[ \int_0^{\rho(x^*, x_0) + \rho(x_0, x) + \|v\|} L(s) \, ds \]

\[ = \int_0^{\rho(x^*, x_0) + \rho(x_0, x) + \|v\|} L(s) \, ds. \tag{58} \]

Hence, \( df_{x_0}^{-1} df \) satisfies the \( L \)-average Lipschitz condition in \( C_r(x_0) \). 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 \( \rho(y^*, x_0) < \tau_0 \). And

\[
\rho(y^*, x^*) \leq \rho(x^*, x_0) + \rho(x^*, x_0)
\]

\[ \leq \rho(x^*, x_0) + \tau_0 = r_0. \tag{59} \]

The proof of the theorem is completed. \( \square \)

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 \( C_r(x^*) \). Suppose that \( C_{\tau_0}(e) \subseteq \exp(B(0, r_0)) \) and \( \rho(x^*, x_0) < (b/2) \). 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 \( \rho(x^*, x_0) < (b/2) \), 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 \( \rho(y^*, x^*) < r_0 \); that is, \( \rho((x^*)^{-1} y^*, e) < r_0 \). Since \( C_{\tau_0}(e) \subseteq \exp(B(0, r_0)) \), there exists \( u \in G \) such that \( \|u\| \leq r_0 \) and \( (x^*)^{-1} y^* = \exp u \); that is, \( y^* = x^* \exp u \). Hence, \( y^* \in N(x^*, r_0) = x^* \exp(B(0, r_0)) \). As \( r_0 \leq r_{\alpha} \) 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 \( df_G^{-1} df \) satisfies the \( L \)-average Lipschitz condition on \( C_r(x^*) \). Suppose that \( \rho(x^*, x_0) < (b/2) \). 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 \( \rho(x^*, y^*) < r_0 \). Clearly, there is a minimizing geodesic \( c \) connecting \( x^* \) \( \cdot y^* \) 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 \( u \in g \) such that \( y^* = x^* \cdot \exp u \) and \( \|u\| = \rho(x^*, y^*) < r_0 \). Hence, \( y^* \in N(x^*, r_0) := x^* \exp(B(0, r_0)) \). As \( r_0 \leq \bar{r} \) by Remark 3(ii), Theorem 2 is applicable, and so \( y^* = x^* \).

5. Theorems under the Kantorovich Condition and the \( \gamma \)-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 \( \mathcal{L}(g) \). 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 \( u, u_0, \ldots, u_k \in g \) 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

\[
r_0 = \frac{1}{L}, \quad b = \frac{1}{2L}, \quad \bar{r} = \frac{2}{L}. \tag{61}
\]

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 \( df_{x_0} \frac{\partial f}{\partial x} \) 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 \( df_{x_0} \frac{\partial f}{\partial x} \) satisfies the \( L \) Lipschitz condition on \( C_{1/L}(x^*) \). Suppose that \( \rho(x^*, x_0) < (1/4L) \). 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 \( g(y^*, x^*) < (1/L) \).

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

**Corollary 11.** Suppose that \( df_{x_0} \frac{\partial f}{\partial x} \) satisfies the \( L \) Lipschitz condition on \( C_{1/L}(x^*) \). Suppose that \( C_{1/L}(x_0) \subseteq \exp(B(0, (1/L))) \) and \( g(x^*, x_0) < (1/4L) \). 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 right-invariant Riemannian metric. Suppose that \( df_{x_0} \frac{\partial f}{\partial x} \) satisfies the \( L \) Lipschitz condition on \( C_{1/L}(x^*) \). Suppose that \( \rho(x^*, x_0) < (1/4L) \). 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 \( f : G \to g \) is a \( C^k \)-map. Define the map \( d^k f_x : g^k \to g \) by

\[
d^k f_x u_1 \cdots u_k = \left( \frac{\partial^{k+1} f}{\partial t_1 \cdots \partial t_k} (x \cdot \exp t_1 u_1 \cdots \exp t_k u_k) \right)_{t_1 = \cdots = t_k = 0}, \tag{62}
\]

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

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

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

\[
d^k f_x u_1 \cdots u_{i-1} u_i = d^k f_x (d^{i-1} f(u_1 \cdots u_{i-1}) u_i) = d^k f_x (d^{i-1} f(u_1 \cdots u_{i-1}) (u)) \tag{64}
\]

for each \( (u_1, \ldots, u_k) \in g^k \). In particular, for fixed \( u_1, \ldots, u_{i-1}, u_{i+1}, \ldots, u_k \in g \),

\[
d^k f_x u_1 \cdots u_{i-1} u_i = d (d^{i-1} f(u_1 \cdots u_{i-1})) (u) \tag{65}
\]

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

\[
\|d^k f_x\| := \sup \left\{ \|d^k f_x u_1 u_2 \cdots u_k\| : (u_1, \ldots, u_k) \in g^k \right\} \tag{66}
\]

with each \( \|u\| = 1 \).

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

\[
da^2 f_x u v = (d^2 f_x \cdot v)_u \quad \text{for any } u, v \in g \text{ and each } z \in G. \tag{67}
\]

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

\[
da^2 f_x (\exp(ts)u) - df_x = \int_0^t d^2 f_x (\exp(\lambda s)u) ds \tag{68}
\]

for each \( u \in g, t \in \mathbb{R} \).

The \( \gamma \)-condition for nonlinear operators in Banach spaces was first introduced by Wang and Han [25], and explored further by Wang et al. [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 \( d^2 f_x \) 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 \( f_{x_0}^{-1} \) exists. \( f \) is said to satisfy the \( \gamma \)-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 \( \rho(x, x_0) := \sum_{i=0}^k \|u_i\| < r \),

\[
\left\|d^2 f_{x_0} d^2 f_x\right\| \leq \frac{2\gamma}{(1 - \gamma \rho(x, x_0))^3}. \tag{69}
\]
Abstract and Applied Analysis

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

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

$$L(s) = \frac{2\gamma}{(1-\gamma)s^3}$$

for each $0 < s < \frac{1}{\gamma}$. (70)

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

**Proposition 14.** Suppose that $f$ satisfies the $\gamma$-condition at $x_0$ on $C_r(x_0)$. Then $df_x^{-1}df$ 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

$$r_0 = \frac{2 - \sqrt{2}}{2\gamma}, \quad b = \frac{3 - 2\sqrt{2}}{\gamma}, \quad r = \frac{1}{2\gamma}. \quad (71)$$

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

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

**Theorem 16.** Suppose that $f$ satisfies the $\gamma$-condition on $C_{(2-\sqrt{2})/2\gamma}(x^*)$. Suppose that $\varphi(x^*, x_0) < (3 - 2\sqrt{2})/2\gamma$. Then the sequence $\{x_i\}$ 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 $\varphi(y^*, x) < (2 - \sqrt{2})/2\gamma$.

**Remark 17.** Theorem 16 improves the corresponding results in [19, Corollary 4.1], where it was proved under the following assumption: there exists $v \in G$ such that $x_0 = x^* \exp v$ and $\|v\| \leq (a_0/\gamma)$ with $a_0 = 0.081256 \ldots$ being the smallest positive root of the equation $a_0(1 - 4a_0 + 2a_0^2) = 3 - 2\sqrt{2}$. Clearly, $(a_0/\gamma) < (3 - 2\sqrt{2})/2\gamma$. Note also that in general, there does not exist $v \in G$ satisfying $x_0 = x^* \exp v$ because the exponential map is not surjective global, even if $\varphi(x^*, x_0) < (3 - 2\sqrt{2})/2\gamma$. 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 $\gamma$-condition on $C_{(2-\sqrt{2})/2\gamma}(x^*)$. Suppose that $C_{(2-\sqrt{2})/2\gamma}(e) \subseteq \exp(B(0, (2 - \sqrt{2})/2\gamma))$ and $\varphi(x^*, x_0) < (3 - 2\sqrt{2})/2\gamma$. Then, the sequence $\{x_i\}$ 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 $\gamma$-condition on $C_{(2-\sqrt{2})/2\gamma}(x^*)$. Suppose that $\varphi(x^*, x_0) < (3 - 2\sqrt{2})/2\gamma$. Then, the sequence $\{x_i\}$ 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

$$\gamma_x := \gamma(f, x) = \sup_{i \geq 1} \frac{\|df_x^{-1}df_x^i\|^{1/(i-1)}}{i!}.$$ (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 $\gamma, \gamma_x := \gamma(f, x^*)$ and let $r = (2 - \sqrt{2})/2\gamma_x$. Then $f$ satisfies the $\gamma_x$-condition at $x^*$ on $C_r(x^*)$.

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

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

**Theorem 22.** Suppose that $\varphi(x^*, x_0) < (3 - 2\sqrt{2})/2\gamma_x$. Then the sequence $\{x_i\}$ 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 $\varphi(y^*, x) < (2 - \sqrt{2})/2\gamma_x$.

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

**Corollary 23.** Suppose that $C_{(2-\sqrt{2})/2\gamma_x}(e) \subseteq \exp(B(0, (2 - \sqrt{2})/2\gamma_x))$ and $\varphi(x^*, x_0) < (3 - 2\sqrt{2})/2\gamma_x$. Then, the sequence $\{x_i\}$ 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


Submit your manuscripts at
http://www.hindawi.com