ENLARGING THE CONVERGENCE BALL
OF NEWTON'S METHOD ON LIE GROUPS

IOANNIS K. ARGYROS* and SAÏD HILOUT†

Dedicated to prof. I. Păvăloiu on the occasion of his 75th anniversary

Abstract. We present a local convergence analysis of Newton's method for approximating a zero of a mapping from a Lie group into its Lie algebra. Using more precise estimates than before [55, 56] and under the same computational cost, we obtain a larger convergence ball and more precise error bounds on the distances involved. Some examples are presented to further validate the theoretical results.

MSC 2010. 65G99, 65J15, 47H17, 49M15, 90C90.
Keywords. Newton's method, Lie group, Lie algebra, Riemannian manifold, convergence ball, Kantorovich hypothesis.

1. INTRODUCTION

In this paper, we are concerned with the problem of approximating a zero $x^\star$ of $\mathcal{C}^1$-mapping $F : G \longrightarrow Q$, where G is a Lie group and Q the Lie algebra of G that is the tangent space $T_e\, G$ of G at e, equipped with the Lie bracket $[.,.] : Q \times Q \longrightarrow Q$ [7, 21, 28, 29, 53].

The study of numerical algorithms on manifolds for solving eigenvalue or optimization problems on Lie groups [1]–[12], [33]–[36], [54]–[57] is very important in Computational Mathematics. Newton-type methods are the most popular iterative procedures used to solve equations, when these equations contain differentiable operators. The study about convergence matter of Newton's method is usually centered on two types: semilocal and local convergence analysis. The semilocal convergence matter is, based on the information around an initial point, to give criteria ensuring the convergence of Newton's method; while the local one is, based on the information around a solution, to find estimates of the radii of convergence balls. There is a plethora of studies on the weakness and/or extension of the hypothesis made on the underlying operators; see for example (cf. [7, 11, 16] and references theiren). A local as

---

*Department of Mathematical Sciences, Cameron University, Lawton, OK 73505, USA, e-mail: ioannisa@cameron.edu.

†Poitiers University, Laboratoire de Mathématiques et Applications, Bd. Pierre et Marie Curie, Téléport 2, B.P. 30179, 86962 Futuroscope Chasseneuil Cedex, France, e-mail: said.hilout@math.univ-poitiers.fr.

well as a semilocal convergence of Newton-type methods has been given by several authors under various conditions [2]–[58]. Recently in [15], we presented a finer semilocal convergence analysis for Newton's method than in earlier studies [7, 32, 34, 36, 54, 55, 56].

Newton's method with initial point $x_0 \in G$ was first introduced by Owren and Welfert [43] in the form

$$(1.1) \qquad x_{n+1} = x_n \cdot \exp\left(-dF_{x_n}^{-1} F(x_n)\right) \quad \text{for each} \quad n = 0, 1, \dots.$$

Newton's method is undoubtedly the most popular method for generating a sequence $\{x_n\}$ approximating $x^\star$ [7, 11, 15, 16, 32, 34, 36], [54]–[56]. In the present paper we establish a finer local convergence analysis with the advantages $(\mathcal{A}_l)$:

(i) Larger convergence ball;

and

(ii) Tighter error bounds on the distances involved.

The necessary background on Lie groups can be found in [7, 11, 15, 16] and the references therein. The paper is organized as follows. In Section 2, we present the local convergence analysis of Newton's method. Finally, numerical examples are given in the concluding Section 3.

2. LOCAL CONVERGENCE ANALYSIS FOR NEWTON'S METHOD

We shall study the semilocal/local convergence of Newton's method. In the rest of the paper we assume $\langle \cdot, \cdot \rangle$ the inner product and $\| \cdot \|$ on Q. As in [56] we define a distance on G for $x, y \in G$ as follows:

$$(2.1) \qquad m(x, y) = \inf\left\{ \sum_{i=1}^{k} \| z_i \| : \text{ there exist } k \geq 1 \text{ and } z_1, \cdots, z_k \in Q \right.$$

$$\left. \text{such that } y = x \cdot \exp z_1 \cdots \exp z_k \right\}.$$

By convention $\inf \emptyset = +\infty$. It is easy to see that $m(\cdot, \cdot)$ is a distance on G and the topology induced is equivalent to the original one on G. Let $w \in G$ and $r > 0$, we denote by

$$U(w, r) = \{y \in G : m(w, y) < r\}$$

the open ball centered at w and of radius r. Moreover, we denote the closure of $U(w, r)$ by $\overline{U}(w, r)$. Let also $L(Q)$ denotes the set of all linear operators on Q.

We need the definition of Lipschitz continuity for a mapping.

DEFINITION 2.1. *Let* $M : G \longrightarrow L(Q)$, $x^\star \in G$ *and* $r > 0$. *We say that* M *satisfies the L-Lipschitz condition on* $U(x^\star, r)$ *if*

$$(2.2) \qquad \| M(x \cdot \exp u) - M(x) \| \leq L \| u \|$$

holds for any $u \in Q$ *and* $x \in U(x^\star, r)$ *such that* $\| u \| + m(x, x^\star) < r$.

It follows from (2.2) that there exists $L_\star$ such that

$$(2.3) \qquad \| M(x \cdot \exp u) - M(x^\star) \| \leq L_\star (\| u \| + m(x, x^\star))$$

holds for any $u \in Q$ and $x \in U(x^\star, r)$ such that $\| u \| + m(x, x^\star) < r$.

We then say M satisfies the $L_\star$-center Lipschitz condition at $x^\star \in G$ on $U(x^\star, r)$. Note that if M satisfies the L-Lipschitz condition on $U(x^\star, r)$, then it also satisfies the $L_\star$-center Lipschitz condition at $x^\star \in G$ on $U(x^\star, r)$. Clearly,

$$(2.4) \qquad\qquad\qquad L_\star \leq L$$

holds in general and $L/L_\star$ can be arbitrarily large [4]–[16].

Let us show that (2.2) indeed implies (2.3). If (2.2) holds, then for $x = x^\star$ there exists $K \in (0, L]$ such that

$$\| M(x^\star \cdot \exp u) - M(x^\star) \| \leq K \| u \| \quad \text{for any} \quad u \in Q \quad \text{such that} \quad \| u \| < r.$$

There exist $k \geq 1$ and $z_0, z_1, \cdots, z_k \in Q$ such that $x = x^\star \cdot \exp z_0 \cdots \exp z_k$. Let $y_0 = x^\star$, $y_{i+1} = y_i \cdot \exp u_i$, $i = 0, 1, \cdots, k$. Then, we have that $x = y_{k+1}$. We can write the identity

$$\begin{aligned}
M(x \cdot \exp u) &- M(x^\star) = \\
&= (M(x \cdot \exp u) - M(x)) + (M(x) - M(x^\star)) \\
&= (M(x \cdot \exp u) - M(x)) + (M(y_k \cdot \exp u_k) - M(y_k)) + (M(y_k) - M(x^\star)) \\
&= (M(x \cdot \exp u) - M(x)) + (M(y_k \cdot \exp u_k) - M(y_k)) \\
&\quad + (M(y_{k-1} \cdot \exp u_{k-1}) - M(x^\star)) \\
&= (M(x \cdot \exp u) - M(x)) + (M(y_k \cdot \exp u_k) - M(y_k)) + \cdots + \\
&\quad + (M(y_1 \cdot \exp u_1) - M(y_1)) + (M(x^\star \cdot \exp u_0) - M(x^\star)).
\end{aligned}$$

Using (2.2) and the triangle inequality, we obtain in turn that

$$\begin{aligned}
\| M(x \cdot \exp u) &- M(x^\star) \| \leq \\
&\leq \| M(x \cdot \exp u) - M(x) \| + \| M(y_k \cdot \exp u_k) - M(y_k) \| + \cdots \\
&\quad + \| M(y_1 \cdot \exp u_1) - M(y_1) \| + \| M(x^\star \cdot \exp u_0) - M(x^\star) \| \\
&\leq L (\| u \| + \| u_k \| + \cdots + \| u_1 \|) + K \| u_0 \| \\
&= L (\| u \| + \| u_k \| + \cdots + \| u_0 \|) + (K - L) \| u_0 \| \\
&\leq L (\| u \| + \| u_k \| + \cdots + \| u_0 \|),
\end{aligned}$$

which implies (2.3).

We need a Banach type lemma on invertible mappings.

LEMMA 2.2. *Suppose that $dF_{x^\star}^{-1}$ exists and let $0 < r \leq \frac{1}{L_\star}$. Suppose $dF_{x^\star}^{-1} dF$ satisfies the $L_\star$-center Lipschitz condition at $x^\star \in G$ on $U(x^\star, r)$; for $x \in U(x^\star, r)$, there exist $k \geq 1$ and $z_0, z_1, \cdots, z_k \in Q$, such that $x = x^\star \cdot \exp z_0 \cdots \exp z_k$ and $\sum_{i=0}^{k} \| z_i \| < r$. Then, linear mapping dF_x^{-1} exists and*

$$(2.5) \qquad\qquad \| dF_x^{-1} dF_{x^\star} \| \leq \left(1 - L_\star \sum_{i=0}^{k} \| z_i \| \right)^{-1}.$$

Proof. Let $y_k = x^\star \cdot \exp z_0 \cdot \exp z_1 \cdots \exp z_{k-1}$. Then, we deduce $y_k \in U(x^\star, r)$, since $\sum_{i=0}^{k-1} \parallel z_i \parallel < r$. Note that $x = y_k \cdot \exp z_k$. Using (2.3) for $M = dF_{x^\star}^{-1} dF$, we get in turn

(2.6)
$$\parallel dF_{x^\star}^{-1} (dF_x - dF_{x^\star}) \parallel \leq L_\star (\parallel z_k \parallel + m(y_k, x^\star)) \leq L_\star \sum_{i=0}^{k} \parallel z_i \parallel < L_\star r \leq 1.$$

It follows from (2.6) and the Banach lemma on invertible operators [7], [32] that dF_x^{-1} exists and (2.5) holds. That completes the proof of Lemma 2.2. $\square$

Next, we study the convergence domain of Newton's method around a zero $x^\star$ of mapping F. First from Lemma 2.3 until Corollary 2.9, we present the local result for Newton's method when G is an Abelian group. Then, the corresponding local results follow when G is not necessarily an Abelian group.

LEMMA 2.3. *Let G be an Abelian group and $0 < r \leq \frac{1}{L_\star}$. Let $F : G \longrightarrow Q$ be a $\mathcal{C}^1$-mapping. Let $x^\star \in G$ be a zero of mapping F. Suppose $dF_{x^\star}^{-1}$ exists and let $L_\star > 0$; there exist $j \geq 1$ and $z_1, z_2, \cdots, z_j \in Q$ such that*

(2.7)
$$x_0 = x^\star \cdot \exp z_1 \cdots \exp z_j \quad for \quad \sum_{i=1}^{j} \parallel z_i \parallel < r;$$

$dF_{x^\star}^{-1} dF$ satisfies the $L_\star$-center Lipschitz condition at $x^\star$ on $U(x^\star, r)$. Then, linear mapping $dF_{x_0}^{-1}$ exists and

(2.8)
$$\parallel dF_{x_0}^{-1} F(x_0) \parallel \leq \frac{\left(2 + L_\star \sum_{i=1}^{j} \parallel z_i \parallel \right) \sum_{i=1}^{j} \parallel z_i \parallel}{2 \left(1 - L_\star \sum_{i=1}^{j} \parallel z_i \parallel \right)}.$$

Proof. Using hypothesis $dF_{x^\star}^{-1} dF$ satisfies the $L_\star$-center Lipschitz condition at $x^\star$ on $U(x^\star, \frac{1}{L_\star})$, we have that

(2.9)
$$\parallel dF_{x^\star}^{-1} (dF_{x \cdot \exp u} - dF_{x^\star}) \parallel \leq L_\star (\parallel u \parallel + m(x, x^\star))$$

for all $u \in Q$, $x \in U(x^\star, r)$ such that $\parallel u \parallel + m(x, x^\star) < r$. Let $z = z_1 + z_2 + \cdots + z_j$. Then, since G is an Abelian group, $\exp z_1 \cdot \exp z_2 \cdots \exp z_j = \exp(z_1 + z_2 + \cdots + z_j) = \exp z$, so, we can write $x_0 = x^\star \cdot \exp z$. It then follows from Lemma 2.2 that $dF_{x_0}^{-1}$ exists and

(2.10)
$$\parallel dF_{x_0}^{-1} dF_{x^\star}) \parallel \leq \left(1 - L_\star \sum_{i=1}^{j} \parallel z_i \parallel \right)^{-1}.$$

We also have the following identity

(2.11)
$$\begin{aligned} F(x_0) &= F(x^\star \cdot \exp z) - F(x^\star) \\ &= \int_0^1 dF_{x^\star \cdot \exp(tz)} z \, dt = \int_0^1 (dF_{x^\star \cdot \exp(tz)} - dF_{x^\star}) z \, dt + dF_{x^\star} z. \end{aligned}$$

In view of (2.9) and (2.11), we get that
(2.12)

$$
\begin{aligned}
\| dF_{x^\star}^{-1} F(x_0) \| &\leq \int_0^1 \| dF_{x^\star}^{-1} (dF_{x^\star \cdot \exp(t z)} - dF_{x^\star}) \| \| z \| dt + \| z \| \\
&\leq \int_0^1 \| L_\star (t \| z \|) \| z \| dt + \| z \| \\
&\leq \left(\frac{L_\star}{2} \sum_{i=1}^j \| z_i \| + 1 \right) \sum_{i=1}^j \| z_i \| .
\end{aligned}
$$

Moreover, by (2.10) and (2.12), we obtain that

$$
\| dF_{x_0}^{-1} F(x_0) \| \leq \| dF_{x_0}^{-1} dF_{x^\star} \| \| dF_{x^\star}^{-1} F(x_0) \| \leq \frac{\left(2 + L_\star \sum_{i=1}^j \| z_i \| \right) \sum_{i=1}^j \| z_i \|}{2 \left(1 - L_\star \sum_{i=1}^j \| z_i \| \right)}.
$$

That completes the proof of Lemma 2.3.   □

REMARK 2.4. The proof of Lemma 2.3 reduces to [56, Lemma 3.2] if $L_\star = L$. Otherwise (*i.e.*, if $L_\star < L$) it constitutes an improvement. We have also included the proof although similar to the corresponding one in [56] because it is not straightforward to see that $L_\star$ can replace L in the derivation of the crucial estimate (2.8). Note also that (2.9) can holds for some $L_\star^1 \in (0, L_\star]$. If $L_\star^1 < L_\star$, then according to the proof of Lemma 2.3, $L_\star^1$ can replace $L_\star$ in (2.8).

Let us define parameter α for $\beta = \frac{L_\star}{L}$ by

(2.13)
$$
\alpha = \begin{cases} 4 \frac{\sqrt{1 + 3\beta} - (1+\beta)}{\beta (1-\beta)}, & \text{if } L_\star \neq L \\ 1, & \text{if } L_\star = L. \end{cases}
$$

Then, it is can easily be seen that

(2.14)
$$
\alpha \geq 1.
$$

We have the local convergence result for Newton's method.

THEOREM 2.5. *Let G be an Abelian group. Let $0 < r \leq \frac{\alpha}{4L}$, where α in given in (2.13). Let $F : G \longrightarrow Q$ be a $\mathcal{C}^1$-mapping. Suppose there exists $x^\star \in G$ such that $F(x^\star) = 0$, $dF_{x^\star}^{-1}$ exists and $dF_{x^\star}^{-1} dF$ satisfies the L-Lipschitz condition on $U(x^\star, \frac{3r}{1 - L_\star r})$. Then, sequence $\{x_n\}$ ($n \geq 0$) generated by Newton's method starting at $x_0 \in U(x^\star, r)$ is well defined, remains in $U(x^\star, \frac{3r}{1 - L_\star r})$ for all $n \geq 0$ and converges to a zero $y^\star$ of mapping F such that*

$$
m(x^\star, y^\star) < \frac{3r}{1 - L_\star r}.
$$

Proof. It follows from hypothesis $x_0 \in U(x^\star, r)$ that there exist $j \geq 1$ and $z_1, z_2, \cdots, z_j \in Q$ such that (2.7) holds. If $dF_{x^\star}^{-1} dF$ is L-Lipschitz on $U(x^\star, \frac{3r}{1 - L_\star r})$, then it is $L_\star$-center Lipschitz at $x^\star$ on $U(x^\star, \frac{3r}{1 - L_\star r})$. Note also

that $\frac{\alpha}{4L} \leq \frac{1}{L_\star}$ by the choice of α. It follows from Lemma 2.3 that linear mapping $dF_{x_0}^{-1}$ exists and

$$(2.15) \qquad \eta = \parallel dF_{x_0}^{-1} F(x_0) \parallel \leq \frac{\left(2+L_\star \sum\limits_{i=1}^{j} \parallel z_i \parallel \right) \sum\limits_{i=1}^{j} \parallel z_i \parallel}{2\left(1-L_\star \sum\limits_{i=1}^{j} \parallel z_i \parallel \right)}.$$

Set

$$\overline{L} = \frac{L}{1-L_\star \sum\limits_{i=1}^{j} \parallel z_i \parallel} \quad \text{and} \quad \overline{r} = \frac{(2+L_\star r)\,r}{1-L_\star r}.$$

We shall show mapping $dF_{x_0}^{-1} dF$ satisfies the $\overline{L}$-Lipschitz condition on $U(x_0, \overline{r})$. Indeed, let $x \in U(x_0, \overline{r})$ and $z \in Q$ be such that $\parallel z \parallel + m(x_0, x) < \overline{r}$. Then, we get that

$$\parallel z \parallel + m(x, x^\star) < \parallel z \parallel + m(x, x_0) + m(x_0, x^\star) < \overline{r} + r \leq \frac{3\,r}{1-L_\star r}.$$

Using Lemma 2.2, we have that

$$\begin{aligned} \parallel dF_{x_0}^{-1}(dF_{x\cdot\exp z} - dF_x) \parallel &\leq\ \parallel dF_{x_0}^{-1} dF_{x^\star} \parallel \parallel dF_{x^\star}^{-1}(dF_{x\cdot\exp z} - dF_x) \parallel \\ &\leq\ \frac{L\,\|z\|}{1-L_\star \sum\limits_{i=1}^{j} \parallel z_i \parallel} = \overline{L} \parallel z \parallel. \end{aligned}$$

Set

$$(2.16) \qquad \overline{h} = \overline{L}\,\eta \quad \text{and} \quad \overline{r_1} = \frac{2\,\eta}{1+\sqrt{1-2\,\overline{h}}} \leq 2\,\eta.$$

Then, by (2.15), we obtain that

$$(2.17) \qquad \overline{r_1} \leq \frac{\left(2+L_\star \sum\limits_{i=1}^{j} \parallel z_i \parallel \right) \sum\limits_{i=1}^{j} \parallel z_i \parallel}{1-L_\star \sum\limits_{i=1}^{j} \parallel z_i \parallel} \leq \frac{(2+L_\star r)\,r}{1-L_\star r} = \overline{r}$$

and

$$(2.18) \qquad \overline{h} \leq \frac{(2+L_\star r)\,L\,r}{2\,(1-L_\star r)^2} \leq \tfrac{1}{2},$$

by the choice of r and α. By standard majorization techniques [7], $\{x_k\}$ converges to some zero $y^\star$ of mapping F and $m(x_0, y^\star) \leq \overline{r_1}$. Furthermore, we obtain that

$$m(x^\star, y^\star) \leq m(x^\star, x_0) + m(x_0, y^\star) \leq r + \overline{r_1} \leq r + \overline{r} \leq \frac{3\,r}{1-L_\star r}.$$

That completes the proof of Theorem 2.5.   □

     The proofs of remaining results in this Section are omitted, since they can be obtained from the corresponding ones in [56].

COROLLARY 2.6. *Let G be an Abelian group. Let $F : G \longrightarrow Q$ be a $\mathcal{C}^1$-mapping. Let $x^\star \in G$ be a zero of mapping F. Suppose $dF_{x^\star}^{-1}$ exists and $dF_{x^\star}^{-1} dF$ satisfies the L-Lipschitz condition on $U(x^\star, \frac{1}{L_\star})$. Let also $x_0 \in U(x^\star, \frac{\alpha}{4L})$. Then, sequence $\{x_n\}$ ($n \geq 0$) generated by Newton's method starting at x_0 is well defined, remains in $U(x^\star, \frac{\alpha}{4L})$ for all $n \geq 0$ and converges to a zero $y^\star$ of mapping F such that $m(x^\star, y^\star) < \frac{1}{L_\star}$.*

We denote in the following Corollary

$$B(0, r) = \{z \in Q : \| z \| < r\}.$$

COROLLARY 2.7. *Let G be an Abelian group. Let $F : G \longrightarrow Q$ be a $\mathcal{C}^1$-mapping. Let $x^\star \in G$ be a zero of mapping F. Suppose $dF_{x^\star}^{-1}$ exists and $dF_{x^\star}^{-1} dF$ satisfies the L-Lipschitz condition on $U(x^\star, \frac{1}{L_\star})$. Let also $x_0 \in U(x^\star, \frac{1}{L_\star})$. Let σ be the maximal number such that $U(e, \sigma) \subseteq \exp(B(0, \frac{1}{L_\star}))$. Set $r = \min\left\{\frac{\sigma}{3+L_\star \sigma}, \frac{\alpha}{4L}\right\}$ and $N(x^\star, r) = x^\star \exp(B(0, r))$. Then, sequence $\{x_n\}$ ($n \geq 0$) generated by Newton's method starting at $x_0 \in N(x^\star, r)$ converges to $x^\star$.*

COROLLARY 2.8. *Let G be an Abelian group. Let $F : G \longrightarrow Q$ be a $\mathcal{C}^1$-mapping, where G is a compact connected Lie group equipped with a bi-invariant Riemannian metric. Let $x^\star \in G$ be a zero of mapping F and $0 < r < \frac{\alpha}{4L}$. Suppose $dF_{x^\star}^{-1}$ exists and $dF_{x^\star}^{-1} dF$ satisfies the L-Lipschitz condition on $U(x^\star, \frac{3r}{1-L_\star r})$. Let also $x_0 \in U(x^\star, r)$. Then, sequence $\{x_n\}$ ($n \geq 0$) generated by Newton's method starting at x_0, remains in $\overline{U}(x^\star, \frac{3r}{1-L_\star r})$ for all $n \geq 0$ and converges to $x^\star$.*

COROLLARY 2.9. *Let G be an Abelian group. Let $F : G \longrightarrow Q$ be a $\mathcal{C}^1$-mapping, where G is a compact connected Lie group equipped with a bi-invariant Riemannian metric. Let $x^\star \in G$ be a zero of mapping F. Suppose $dF_{x^\star}^{-1}$ exists and $dF_{x^\star}^{-1} dF$ satisfies the L-Lipschitz condition on $U(x^\star, \frac{1}{L_\star})$. Let also $x_0 \in U(x^\star, \frac{\alpha}{4L})$. Then, sequence $\{x_n\}$ ($n \geq 0$) generated by Newton's method starting at x_0, remains in $\overline{U}(x^\star, \frac{\alpha}{4L})$ for all $n \geq 0$ and converges to $x^\star$.*

LEMMA 2.10. *Let $0 < r \leq \frac{1}{L}$. Let $F : G \longrightarrow Q$ be a $\mathcal{C}^1$-mapping. Let $x^\star \in G$ be a zero of mapping F. Suppose $dF_{x^\star}^{-1}$ exists; there exist $j \geq 1$ and $z_1, z_2, \cdots, z_j \in Q$ such that $x_0 = x^\star \cdot \exp z_1 \cdots \exp z_j$ for $\sum_{i=1}^{j} \| z_i \| < r$ and $dF_{x^\star}^{-1} dF$ satisfies the L-Lipschitz condition at $x^\star$ on $U(x^\star, r)$. Then, the following assertions hold*

(i)

$$\| dF_{x^\star}^{-1} (dF_{x^\star \cdot \exp z} - dF_{x^\star}) \| \leq K_\star \| z \|$$

for each $z \in Q$ with $\| z \| < r$ and some $K_\star \in (0, L]$.

(ii)

$$\| dF_{x^\star}^{-1} (F'(x_0) - F'(x^\star)) \| \le L_\star \sum_{i=1}^{j} \| z_i \| \quad for\ some \quad L_\star \in (0, L].$$

(iii) *Linear mapping* $dF_{x_0}^{-1}$ *exists and*

$$\| dF_{x_0}^{-1} F(x_0) \| \le \frac{\left(2 + L \sum_{i=1}^{j} \| z_i \|\right) \sum_{i=1}^{j} \| z_i \|}{2\left(1 - L_\star \sum_{i=1}^{j} \| z_i \|\right)}.$$

Proof.   (i) This assertion follows from the hypothesis that $dF_{x^\star}^{-1} dF$ satisfies the L-Lipschitz condition at $x^\star$ on $U(x^\star, r)$.

(ii) Let $w_0 = x^\star$, $w_{i+1} = w_i \cdot \exp z_{i+1}$, $i = 1, 2, \cdots, j - 1$. Then, we have $w_j = w_{j-1} \cdot \exp z_j = x_0$. Using the L-Lipschitz condition, we get in turn that

$\| dF_{x^\star}^{-1} (dF_{w_j} - dF_{x^\star}) \|$
$\le \| dF_{x^\star}^{-1} (dF_{w_{j-1} \cdot \exp z_j} - dF_{w_{j-1}}) \| + \cdots + \| dF_{x^\star}^{-1} (dF_{w_1 \cdot \exp z_2} - dF_{w_1}) \| +$
$\| dF_{x^\star}^{-1} (dF_{x^\star \cdot \exp z_1} - dF_{x^\star}) \|$
$\le L(\| z_j \| + \cdots + \| z_2 \|) + K_\star \| z_1 \|$
$= L (\| z_j \| + \cdots + \| z_1 \|) + (K_\star - L) \| z_1 \| \le L (\| z_j \| + \cdots + \| z_1 \|)$

which shows (ii).

(iii) We have by (ii) that

$$\| dF_{x^\star}^{-1} (dF_{x_0} - dF_{x^\star}) \| \le L_\star (\| z_j \| + \cdots + \| z_1 \|) \le L\, r < 1.$$

Hence, $dF_{x_0}^{-1}$ exists and

(2.19) $$\| dF_{x_0}^{-1} dF_{x^\star} \| \le \left(1 - L_\star \sum_{i=1}^{j} \| z_i \|\right)^{-1}.$$

We have that

$$dF_{x^\star}^{-1} (F(w_i) - F(w_{i-1})) = dF_{x^\star}^{-1} \int_0^1 dF_{w_{i-1} \cdot \exp(t\, z_i)} z_i\, dt$$
$$= \int_0^1 dF_{x^\star}^{-1} (dF_{w_{i-1} \cdot \exp(t\, z_i)} - dF_{x^\star}) z_i\, dt + z_i.$$

Hence, we get that

$$\| dF_{x^\star}^{-1} (dF_{w_{k-1} \cdot \exp(z_k)} - dF_{w_{k-1}}) \| \le L \| z_k \| \quad \text{for each} \quad k = 1, 2, \cdots, j.$$

Therefore, we obtain that

$\| dF_{x^\star}^{-1} (dF_{w_{i-1} \cdot \exp(t\, z_i)} - dF_{x^\star}) \| \le$
$\le \sum_{k=1}^{i-1} \| dF_{x^\star}^{-1} (dF_{w_{k-1} \cdot \exp(z_k)} - dF_{w_{k-1}}) \| + \| dF_{x^\star}^{-1} (dF_{w_{i-1} \cdot \exp(tw_i)} - dF_{w_{i-1}}) \|$
$\le L \sum_{k=1}^{i-1} \| z_k \| + Lt\| z_i \|.$

We have that $F(x_0) = \sum\limits_{i=1}^{j} (F(w_i) - F(w_{i-1}))$. That is, we can get

$$(2.20) \quad
\begin{aligned}
& \parallel dF_{x^\star}^{-1} F(x_0) \parallel \\
& \leq \sum_{i=1}^{j} \left(\int_0^1 \parallel dF_{x^\star}^{-1} (dF_{w_{i-1}\cdot\exp(t z_i)} - dF_{x^\star}) \parallel \parallel z_i \parallel dt + \parallel z_i \parallel \right) \\
& \leq \sum_{i=1}^{j} \left(\int_0^1 L \left(\sum_{k=1}^{i-1} \parallel z_k \parallel + t \parallel z_i \parallel \right) \parallel z_i \parallel dt + \parallel z_i \parallel \right) \\
& \leq \left(\frac{L}{2} \sum_{i=1}^{j} \parallel z_i \parallel + 1 \right) \sum_{i=1}^{j} \parallel z_i \parallel.
\end{aligned}$$

The result now follows from (2.19), (2.20) and

$$\parallel dF_{x_0}^{-1} F(x_0) \parallel \leq \parallel dF_{x_0}^{-1} dF_{x^\star} \parallel \parallel dF_{x^\star}^{-1} F(x_0) \parallel \leq \frac{\left(2 + L \sum\limits_{i=1}^{j} \parallel z_i \parallel \right) \sum\limits_{i=1}^{j} \parallel z_i \parallel}{2 \left(1 - L_\star \sum\limits_{i=1}^{j} \parallel z_i \parallel \right)}.$$

The proof of Lemma 2.10 is complete. $\qquad\square$

Let us define parameter δ by

$$(2.21) \qquad \delta = \begin{cases} 4 \frac{\sqrt{2(1+\beta)} - (1+\beta)}{1-\beta^2}, & \text{if} \quad L_\star \neq L \\ 1, & \text{if} \quad L_\star = L. \end{cases}$$

We have that

$$\delta \geq 1.$$

THEOREM 2.11. *Let* $0 < r \leq \frac{\delta}{4L}$, *where* δ *in given in (2.21). Let* $F : G \longrightarrow Q$ *be a* $\mathcal{C}^1$*-mapping. Suppose there exists* $x^\star \in G$ *such that* $F(x^\star) = 0$, $dF_{x^\star}^{-1}$ *exists and* $dF_{x^\star}^{-1} dF$ *satisfies the* L*-Lipschitz condition on* $U(x^\star, R = \frac{3+(L-L_\star)r}{1-L_\star r} r)$. *Then, sequence* $\{x_n\}$ *(n $\geq$ 0) generated by Newton's method starting at* $x_0 \in U(x^\star, r)$ *is well defined, remains in* $U(x^\star, R)$ *for all* $n \geq 0$ *and converges to a zero* $y^\star$ *of mapping* F *such that* $m(x^\star, y^\star) < R$.

Proof. The proof is similar to the proof of Theorem 2.5. Note that $\bar{r}$ in the proof of Theorem 2.5 is replaced by

$$\bar{\bar{r}} = \frac{(2+Lr) r}{1-L_\star r}$$

and η satisfies the new condition

$$\eta \leq \frac{\left(2 + L \sum\limits_{i=1}^{j} \parallel z_i \parallel \right) \sum\limits_{i=1}^{j} \parallel z_i \parallel}{2 \left(1 - L_\star \sum\limits_{i=1}^{j} \parallel z_i \parallel \right)}.$$

The proof of Theorem 2.11 is complete. $\qquad\square$

COROLLARY 2.12. *Let $F : G \longrightarrow Q$ be a $\mathcal{C}^1$-mapping. Let $x^\star \in G$ be a zero of mapping F. Suppose $dF_{x^\star}^{-1}$ exists and $dF_{x^\star}^{-1} dF$ satisfies the L-Lipschitz condition on $U(x^\star, \frac{1}{L_\star})$. Let also $x_0 \in U(x^\star, \frac{\delta}{4L})$. Then, the conclusions of Corollary 2.6 hold.*

COROLLARY 2.13. *Let $F : G \longrightarrow Q$ be a $\mathcal{C}^1$-mapping. Let $x^\star \in G$ be a zero of mapping F. Suppose $dF_{x^\star}^{-1}$ exists and $dF_{x^\star}^{-1} dF$ satisfies the L-Lipschitz condition on $U(x^\star, \frac{1}{L_\star})$. Let also $x_0 \in U(x^\star, \frac{1}{L_\star})$. Let σ and $N(x^\star, r)$ as in Corollary 2.7. Set $r = \min\left\{ r_0, r_1, \frac{\delta}{4L} \right\}$, where*

$$r_0 = \begin{cases} \frac{-(L_\star + 3L) + \sqrt{(L_\star + 3L)^2 + 4L(L - L_\star)}}{2L(L - L_\star)}, & \text{if} \quad L_\star \neq L \\ \frac{1}{4L}, & \text{if} \quad L_\star = L \end{cases}$$

and

$$r_1 = \begin{cases} \frac{-(3 + \sigma L_\star) + \sqrt{(3 + \sigma L_\star)^2 + 4\sigma(L - L_\star)}}{2(L - L_\star)}, & \text{if} \quad L_\star \neq L \\ \frac{\sigma}{3 + \sigma L}, & \text{if} \quad L_\star = L. \end{cases}$$

Then, sequence $\{x_n\}$ ($n \geq 0$) generated by Newton's method starting at $x_0 \in N(x^\star, r)$ converges to $x^\star$.

COROLLARY 2.14. *Let $F : G \longrightarrow Q$ be a $\mathcal{C}^1$-mapping, where G is a compact connected Lie group equipped with a bi-invariant Riemannian metric. Let $x^\star \in G$ be a zero of mapping F and $0 < r < \frac{\delta}{4L}$. Suppose $dF_{x^\star}^{-1}$ exists and $dF_{x^\star}^{-1} dF$ satisfies the L-Lipschitz condition on $U(x^\star, R)$, where R is defined in Theorem 2.11. Let also $x_0 \in U(x^\star, r)$. Then, sequence $\{x_n\}$ ($n \geq 0$) generated by Newton's method starting at x_0, remains in $\overline{U}(x^\star, R)$ for all $n \geq 0$ and converges to $x^\star$.*

COROLLARY 2.15. *Let $F : G \longrightarrow Q$ be a $\mathcal{C}^1$-mapping, where G is a compact connected Lie group equipped with a bi-invariant Riemannian metric. Let $x^\star \in G$ be a zero of mapping F. Suppose $dF_{x^\star}^{-1}$ exists and $dF_{x^\star}^{-1} dF$ satisfies the L-Lipschitz condition on $U(x^\star, \frac{1}{L_\star})$. Let also $x_0 \in U(x^\star, \frac{\delta}{4L})$. Then, sequence $\{x_n\}$ ($n \geq 0$) generated by Newton's method starting at x_0, remains in $\overline{U}(x^\star, \frac{\delta}{4L})$ for all $n \geq 0$ and converges to $x^\star$.*

REMARK 2.16. (a) The local results reduce to the corresponding ones in [56] if $L = L_\star$. Otherwise (*i.e.*, if $L_\star < L$), they constitute an improvement under the same computational cost with advantages as already stated in the Introduction of this study. Note also that $\alpha > 1$, $\delta > 1$ if $L_\star < L$ and $\alpha \longrightarrow \infty$, $\delta \longrightarrow \infty$ if $\frac{L}{L_\star} \longrightarrow \infty$.

(b) The local results if G is an Abelian group are weaker and tighter than the ones when G is not necessarily an Abelian group. We have for example that $\overline{r} < \overline{\overline{r}}$, $\delta < \alpha$, $\frac{3r}{1 - L_\star r} < R$ and the upper bound on η is smaller if $L_\star < L$.

(c) It is obvious that finer results can be immediately obtained if similar conditions as the semilocal case (see [15, Section 1, $L_\star$ instead L_0]) are used

instead of Kantorovich condition for $L_\star < L$. However, we decided to leave this part of analysis to the motivated reader. We refer the reader to [14] for such results involving nonlinear equations in a Banach space setting. We also refer the reader to [7, 11, 15, 16] for examples.

3. NUMERICAL EXAMPLES

In this Section we present two numerical examples in the more general setting of a nonlinear equation on a Hilbert space where $L_\star < L$.

EXAMPLE 3.1. Let $\mathcal{X} = \mathcal{Y} = \mathbb{R}$, $x^\star = 0$. Define F by $F(x) = -d_2 \sin 1 + d_1 x + d_2 \sin e^{d_3 x}$, where d_1, d_2, d_3 are given real numbers. We have that $F(x^\star) = 0$. Moreover, if d_3 is sufficiently large and d_2 sufficiently small, $L/L_\star$ can be arbitrarily large.

EXAMPLE 3.2. Let $\mathcal{X} = \mathcal{Y} = \mathbb{R}^3$, $\mathcal{D} = \overline{U}(0,1)$ and $x^\star = (0,0,0)$. Define function F on $\mathcal{D}$ for $w = (x,y,z)$ by

$$(3.1) \qquad F(w) = (e^x - 1, \tfrac{e-1}{2} y^2 + y, z).$$

Then, the Fréchet derivative of F is given by

$$F'(w) = \begin{pmatrix} e^x & 0 & 0 \\ 0 & (e-1)\,y + 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Notice that we have $F(x^\star) = 0$, $F'(x^\star) = F'(x^\star)^{-1} = \operatorname{diag}\{1,1,1\}$ and $L_\star = e - 1 < L = e$.

REFERENCES

[1] ABSIL, P.–A., BAKER, C.G. and GALLIVAN, K.A., *Trust-region methods on Riemannian manifolds*, Found. Comput. Math., **7** (2007), pp. 303–330. ⬀

[2] ADLER, R.L., DEDIEU, J.P., MARGULIES, J.Y., MARTENS, M. and SHUB, M., *Newton's method on Riemannian manifolds and a geometric model for the human spine*, IMA J. Numer. Anal., **22** (2002), pp. 359–390. ⬀

[3] ALVAREZ, F., BOLTE, J. and MUNIER, J., *A unifying local convergence result for Newton's method in Riemannian manifolds*, Found. Comput. Math., **8** (2008), pp. 197–226. ⬀

[4] ARGYROS, I.K., *An improved unifying convergence analysis of Newton's method in Riemannian manifolds*, J. Appl. Math. Comput., **25** (2007), pp. 345–351. ⬀

[5] ARGYROS, I.K., *On the local convergence of Newton's method on Lie groups*, Pan Amer. Math. J., **17** (2007), pp. 101–109.

[6] ARGYROS, I.K., *A Kantorovich analysis of Newton's method on Lie groups*, J. Concrete Appl. Anal., **6** (2008), pp. 21–32.

[7] ARGYROS, I.K., *Convergence and Applications of Newton-type Iterations*, Springer Verlag Publ., New York, 2008. ⬀

[8] ARGYROS, I.K., *Newton's method in Riemannian manifolds*, Rev. Anal. Numér. Théor. Approx., **37** (2008), pp. 119–125. ⬀

[9] ARGYROS, I.K., *Newton's method on Lie groups*, J. Appl. Math. Comput., **31** (2009), pp. 217–228.

[10] ARGYROS, I.K., *A semilocal convergence analysis for directional Newton methods*, Math. Comp. AMS, **80** (2011), pp. 327–343. ⬀

[11] ARGYROS, I.K., CHO, Y.J. and HILOUT, S., *Numerical methods for equations and its applications*, CRC Press/Taylor and Francis Group, New-York, 2012.

[12] ARGYROS, I.K. and HILOUT, S., *Newton's method for approximating zeros of vector fields on Riemannian manifolds*, J. Appl. Math. Comput., **29** (2009), pp. 417–427. ⬀

[13] ARGYROS, I.K. and HILOUT, S., *Extending the Newton-Kantorovich hypothesis for solving equations*, J. Comput. Appl. Math., **234** (2010), pp. 2993–3006. ⬀

[14] ARGYROS, I.K. and HILOUT, S., *Weaker conditions for the convergence of Newton's method*, J. Complexity, **28** (2012), pp. 364–387. ⬀

[15] ARGYROS, I.K. and HILOUT, S., *Extending the applicability of Newton's method on Lie groups*, Appl. Math. Comput., to appear.

[16] ARGYROS, I.K. and HILOUT, S., *Numerical methods in nonlinear analysis*, World Scientific Publ., ISBN-13: 978-9814405829, to appear.

[17] BAYER, D.A. and LAGARIAS, J.C., *The nonlinear geometry of linear programming. I. Affine and projective scaling trajectories*, Trans. Amer. Math. Soc., **314** (1989), pp. 499–526. ⬀

[18] BROCKETT, R.W., *Differential geometry and the design of gradient algorithms. Differential geometry: partial differential equations on manifolds* (Los Angeles, CA, 1990), pp. 69–92, Proc. Sympos. Pure Math., **54**, Part 1, Amer. Math. Soc., Providence, RI, 1993.

[19] CHU, M.T. and DRIESSEL, K.R., *The projected gradient method for least squares matrix approximations with spectral constraints*, SIAM J. Numer. Anal., **27** (1990), pp. 1050–1060. ⬀

[20] DEDIEU, J.P., PRIOURET, P. and MALAJOVICH, G., *Newton's method on Riemannian manifolds: Convariant α theory*, IMA J. Numer. Anal., **23** (2003), pp. 395–419. ⬀

[21] DO CARMO, M.P., *Riemannian Geometry*, Boston, Birkhauser, 1992.

[22] EDELMAN, A., ARIAS, T.A. and SMITH, S.T., *The geometry of algorithms with orthogonality constraints*, SIAM J. Matrix Anal. Appl., **20** (1999), pp. 303–353. ⬀

[23] EZQUERRO, J.A. and HERNÁNDEZ, M.A., *Generalized differentiability conditions for Newton's method*, IMA J. Numer. Anal., **22** (2002), pp. 187–205. ⬀

[24] EZQUERRO, J.A. and HERNÁNDEZ, M.A., *On an application of Newton's method to nonlinear operators with ω-conditioned second derivative*, BIT, **42** (2002), pp. 519–530.

[25] FERREIRA, O.P. and SVAITER, B.F., *Kantorovich's theorem on Newton's method in Reimannian manifolds*, J. Complexity, **18** (2002), pp. 304–329. ⬀

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

[27] GUTIÉRREZ, J.M. and HERNÁNDEZ, M.A., *Newton's method under weak Kantorovich conditions*, IMA J. Numer. Anal., **20** (2000), pp. 521–532. ⬀

[28] HALL, B.C., *Lie groups, Lie algebras and representations. An elementary introduction*, Graduate Texts in Mathematics, **222**, Springer-Verlag, New York, 2003.

[29] HELGASON, S., *Differential geometry, Lie groups and symmetric spaces*, Pure and Applied Mathematics, **80**, Academic Press, Inc. Harcourt Brace Jovanovich, Publishers, New York-London, 1978.

[30] HELMKE, U. and MOORE, J.B., *Singular-value decomposition via gradient and self-equivalent flows*, Linear Algebra Appl., **169** (1992), pp. 223–248. ⬀

[31] HELMKE, U. and MOORE, J.B., *Optimization and dynamical systems. With a foreword by R. Brockett*, Communications and Control Engineering Series. Springer-Verlag London, Ltd., London, 1994.

[32] KANTOROVICH, L.V. and AKILOV, G.P., *Functional analysis in normed spaces*, Pergamon Press, New York, 1982.

[33] LI, C., LÓPEZ, G. and MARTÍN-MÁRQUEZ, V., *Monotone vector fields and the proximal point algorithm on Hadamard manifolds*, J. Lond. Math. Soc., **(2) 79** (2009), pp. 663–683. ⬈

[34] LI, C. and WANG, J.H., *Convergence of the Newton method and uniqueness of zeros of vector fields on Riemannian manifolds*, Sci. China Ser. A, **48** (2005), pp. 1465–1478. ⬈

[35] LI, C. and WANG, J.H., *Newton's method on Riemannian manifolds: Smale's point estimate theory under the γ-condition*, IMA J. Numer. Anal., **26** (2006), pp. 228–251. ⬈

[36] LI, C., WANG, J.H. and DEDIEU, J.P., *Smale's point estimate theory for Newton's method on Lie groups*, J. Complexity, **25** (2009), pp. 128–151. ⬈

[37] PROINOV, P.D., *General local convergence theory for a class of iterative processes and its applications to Newton's method*, J. Complexity, **25** (2009), pp. 38–62. ⬈

[38] LUENBERGER, D.G., *The gradient projection method along geodesics*, Management Sci., **18** (1972), pp. 620–631.

[39] MAHONY, R.E., *Optimization algorithms on homogeneous spaces: with applications in linear systems theory*, Ph.D. Thesis, Department of Systems Engineering, University of Canberra, Australia, 1994.

[40] MAHONY, R.E., *The constrained Newton method on a Lie group and the symmetric eigenvalue problem*, Linear Algebra Appl., **248** (1996), pp. 67–89. ⬈

[41] MAHONY, R.E., HELMKE, U. and MOORE, J.B., *Pole placement algorithms for symmetric realisations*, Proceedings of 32nd IEEE Conference on Decision and Control, San Antonio, TX, 1993, pp. 355–1358.

[42] MOORE, J.B., MAHONY, R.E. and HELMKE, U., *Numerical gradient algorithms for eigenvalue and singular value calculations*, SIAM J. Matrix Anal. Appl., **15** (1994), pp. 881–902. ⬈

[43] OWREN, B. and WELFERT, B., *The Newton iteration on Lie groups*, BIT, **40** (2000), pp. 121–145. ⬈

[44] PROINOV, P.D., *New general convergence theory for iterative processes and its applications to Newton-Kantorovich type theorems*, J. Complexity, **26** (2010), pp. 3–42. ⬈

[45] RHEINBOLDT, W.C., *An adaptive continuation process for solving system of nonlinear equations*, Banach Center Publ., **3** (1977), pp. 129–142.

[46] SHUB, M. and SMALE, S., *Complexity of Bézout's theorem. I. Geometric aspects*, J. Amer. Math. Soc., **6** (1993), pp. 459–501. ⬈

[47] SMALE, S., *Newton's method estimates from data at a point*, The merging of disciplines: New directions in Pure, Applied and Computational Mathematics, R. Ewing, K. Gross and C. Martin eds, New-York, Springer Verlag Publ., (1986), pp. 185–196.

[48] SMITH, S.T., *Dynamical systems that perform the singular value decomposition*, Systems Control Lett., **16** (1991), pp. 319–327. ⬈

[49] SMITH, S.T., *Smith, S.T., Geometric optimization methods for adaptive filtering*, Thesis (Ph.D.)–Harvard University, ProQuest LLC, Ann Arbor, MI, 1993.

[50] SMITH, S.T., *Smith, S.T., Optimization techniques on Riemannian manifolds, Hamiltonian and gradient flows, algorithms and control*, Fields Inst. Commun., **3**, Amer. Math. Soc., Providence, RI, 1994, pp. 113–136.

[51] TRAUB, J.F. and WOZNIAKOWSKI, H., *Convergence and complexity of Newton iteration for operator equations*, J. Assoc. Comput. Mach., **26** (1979), pp. 250–258. ⬈

[52] UDRIŞTE, C., *Convex functions and optimization methods on Riemannian manifolds*, Mathematics and its Applications, **297**, Kluwer Academic Publishers Group, Dordrecht, 1994.

[53] VARADARAJAN, V.S., *Lie groups, Lie algebras and their representations*, Reprint of the 1974 edition. Graduate Texts in Mathematics, **102**, Springer-Verlag, New York, 1984.

[54] WANG, J.H. and LI, C., *Uniqueness of the singular points of vector fields on Riemannian manifolds under the γ-condition*, J. Complexity, **22** (2006), pp. 533–548. ☐

[55] WANG, J.H. and LI, C., *Kantorovich's theorem for Newton's method on Lie groups*, J. Zheijiang Univ. Sci. A, **8** (2007), pp. 978–986. ☐

[56] WANG, J.H. and LI, C., *Kantorovich's theorems for Newton's method for mappings and optimization problems on Lie groups*, IMA J. Numer. Anal., **31** (2011), pp. 322–347. ☐

[57] WANG, X.H., *Convergence of Newton's method and uniqueness of the solution of equations in Banach space*, IMA J. Numer. Anal., **20** (2000), pp. 123–134. ☐

[58] ZABREJKO, P.P. and NGUEN, D.F., *The majorant method in the theory of Newton-Kantorovich approximations and the Pták error estimates*, Numer. Funct. Anal. Optim., **9** (1987), pp. 671–684. ☐