J. Nonlinear Var. Anal. 6 (2022), No. 3, pp. 199-212
Available online at http://jnva.biemdas.com
https://doi.org/10.23952/jnva.6.2022.3.03

LOCAL CONVERGENCE OF THE NEWTON'S METHOD IN TWO STEP NILPOTENT LIE GROUPS

BÉCHIR DALI[1], MOHAMMED GUEDIRI[2,*]

[1] *Mathematics Department, faculty of Sciences of Bizerte University of Carthage, 7021 Jarzouna Bizerte, Tunisia*
[2] *Department of Mathematics, College of Sciences, King Saud University, Riyadh 11451, Saudi Arabia*

Abstract. In this paper, we consider N, a simply connected two-step nilpotent Lie group with $\mathscr{N}$, its corresponding (two-step nilpotent) Lie algebra, and we study Newton's method for solving the equation $f(x)=0$, where $f: N \to \mathscr{N}$ is a mapping. Under certain generalized Lipschitz condition, we obtain the convergence radius of Newton's method and the estimation of the uniqueness ball of the zero point of f. Some applications to special cases including Kantorovich's condition and γ-condition are provided. The determination of an approximate zero point of an analytic mapping is also presented.

Keywords. Lipschitz condition; Local convergence; Newton's method; Nilpotent Lie group.

1. Introduction

There is an increasing interest in numerical algorithms on Lie groups, such as, the minimization problems with orthogonality constraints, the optimization problems with equality constraints and so on; see, e.g., [1, 2, 3] and the references therein. As is well known, the Lie group is a hausdorff topological group, which has the structure of a smooth manifold such that the group product and the inversion are smooth operations in the differential structure given on the manifold. Mahony [2] used one-parameter subgroups of the Lie group to develop a version of Newton's method on an arbitrary Lie group, where the approach for solving eigenvalue problems as a constrained optimization problem on a Lie group via Newton's method was explored on a Lie group and the local convergence was analyzed. For the approaches to solve ordinary differential equations on manifolds, Owren and Welfert [4] introduced a Newton's method, which is independent of affine connections on the Lie group for solving the equation $f(x)=0$, where f is a map from the Lie group to its Lie algebra, and shows that, under classical assumptions on f, Newton's method converges quadratically. Newton's method is one of the most efficient methods for solving the systems of non-linear equations when the functions involved are continuously differentiable. More precisely, given an initial point x_0, the Newton's method is formulated as follows: $x_{n+1} = x_n - f'(x_n)^{-1} f(x_n)$, $\forall n \geq 0$, where f is a nonlinear operator from some domain D in a real or complex Banach space X to another Y. In addition

*Corresponding author.
E-mail addresses: bechir.dali@yahoo.fr (B. Dali), mguediri@ksu.edu.sa (M. Guediri).
Received December 9, 2021; Accepted December 31, 2021.

to its practical applications, Newton's method is also a powerful theoretical tool. It has been studied and used extensively recently; see, e.g., [5, 6, 7] and the references therein.

Recently, various problems posed on manifolds arise from theoretic and applied areas of mathematics; see, e.g., [3, 8, 9, 10]. Such problems can usually be formulated as computing zeroes of mappings or vector fields on a Riemannian manifold. Hence, Newton's method has been extended to Riemannian manifolds and explored thoroughly; see e.g., [11, 12, 13, 14, 15, 16]. On the other hand, some numerical problems, such as, symmetric eigenvalue problems, the optimization problems with equality constraints, and ordinary differential equations on manifolds can be actually considered as the problems on Lie groups; see, e.g., [1, 2, 3, 17]. Newton's method has also been extended to Lie groups for solving the equation $f(x) = 0$ with f being a mapping from a two-step nilpotent Lie group to its Lie algebra; see, e.g., [2, 17], It only depends on the one-parameter semigroup but not on the affine connections. This means that Newton's methods on Lie groups are completely different from Newton's method on the corresponding Riemannian manifolds, which is defined via the Riemannian connection. In particular, Mahony [2] used the one-parameter subgroups of a Lie group to develop a version of Newton's method on an arbitrary Lie group, where the local convergence of Newton's method was explored. On the other hand, motivated by the approaches to solve ordinary differential equations on Lie groups, Owren and Welfert [4] considered the implicit method for Lie groups where they used the implicit Euler method as a generic example of an implicit integration method in a Lie group setting. In [18], Kantorovich's theorem was established for Newton's method on a Lie group, which is independent of the connection. Furthermore, in [19], the notion of the (one-parameter subgroup) γ-condition for a mapping from a Lie group to its Lie algebra was introduced. The α-theory and γ-theory for Newton's method for a mapping satisfying these conditions were established. As is well known, Lie groups play a central role in the general program of analysis on metric spaces, while simultaneously continuing to figure prominently in applications from other scientific disciplines ranging from robotic control and planning problems to magnetic resonance imaging function to new models of neurobiological visual processing and digital image reconstruction. For example, some fundamental geometric structure that one meets in information transmission, signal analysis, optics, quantization, quantum field theory, and the symplectic group of the plane as well as in its Lie algebra is that of a Heisenberg group and a Heisenberg algebra (see, e.g., [20, 21, 22]). On other hand, the classic technics of Newton's method in $\mathbb{R}^n$ may not be still available for a general Lie groups since, in general, they are non abelian. However, the class of two-step nilpotent Lie groups/algebras is the closest one to the abelian structure. In [23], Dali, Li, and Wang studied the Newton's method for the Heisenberg Lie group.

The purpose of this paper is to study Newton's method on two-step nilpotent Lie groups for solving the equation $f(x) = 0$ with f being a mapping from the two-step nilpotent Lie group to its Lie algebra. Under certain generalized Lipschitz conditions, we obtain the convergence radius of Newton's method and the estimation of the uniqueness ball of the zero point of f. Some applications to special cases including the Kantorovich's condition and the γ-condition are provided. Furthermore, the determination of an approximate zero point of an analytic mapping is also presented. Note that, in a connected nilpotent Lie group, for any two points, there is a one-parameter subgroup to connect them (see [24]). This property will play a key role in our study. Our results improve and extend the corresponding results in [18, Corollary 3.4] and [19, Corollary 4.1].

The paper is organized as follows. In Section 2, we introduce some preliminaries and basic properties about nilpotent Lie groups and essentially two-step nilpotent Lie groups and Lie algebras. In Section 3, the main results on Newton's method on the two-step nilpotent Lie group under certain generalized Lipschitz condition are given. Some applications to special cases are provided in Section 4, the last section.

2. Preliminaries

This section is devoted to recalling basic notations of two-step nilpotent Lie group/algebra. Let us collect a few well-known definitions and the results on Lie algebras (see, e.g., [25, 26]). Let $\mathscr{N}$ be finite dimensional nilpotent Lie algebra over the field of real numbers. For each integer $k \geq 1$, let $\mathscr{N}^k = [\mathscr{N}, \mathscr{N}^{k-1}]$, where $\mathscr{N}^0 = \mathscr{N}$. The Lie algebra $\mathscr{N}$ is said to be a k-step nilpotent Lie algebra if $\mathscr{N}^k = \{0\}$ and $\mathscr{N}^{k-1} \neq \{0\}$ (for some positive integer k). A k-step nilpotent Lie algebra $\mathscr{N}$ has a non trivial center that contains $\mathscr{N}^{k-1}$.

Let us briefly recall the definition of the exponential mapping, which is the most important construct connecting both of the Lie group to the Lie algebra. Given a Lie group N, one proves that the tangent space $\mathscr{N} = T_eN$, at the identity, has a canonical structure of Lie algebra, which is the Lie algebra corresponding to N. On the other hand, one shows that $\mathscr{N}$ is isomorphic to the Lie algebra of left invariant vector fields on N. Given $U \in \mathscr{N}$, we let $\sigma_U : \mathbb{R} \to N$ be the one-parameter subgroup of N, determined by the left invariant vector field $X_U : y \mapsto (L_y)'_e(U)$, i.e., σ_U satisfies that $\sigma_U(0) = e$ and $\sigma'_U(t) = X_U(\sigma_u(t)) = (L_{\sigma_U(t)})'_e(U)$, $U \in \mathscr{N}$, $t \in \mathbb{R}$, where $L_x : N \to N, \quad y \mapsto xy$. The value of the exponential map at U is defined by $\exp(U) = \sigma_U(1)$. Moreover, we have $\exp(tU) = \sigma_{tU}(1) = \sigma_U(t)$, $t \in \mathbb{R}$, $U \in \mathscr{N}$. Let N be the unique connected and simply connected Lie group corresponding to $\mathscr{N}$. It is well known that, when $\mathscr{N}$ is a nilpotent Lie algebra, the exponential mapping $\exp : \mathscr{N} \to N$ is a global diffeomorphism. Let $\log := \exp^{-1} : N \to \mathscr{N}$ be its inverse. This means that, for any $x \in N$, there exists a unique $U \in \mathscr{N}$ such that $x = \exp U$, in particular, for any pair of points $x, x' \in N$, there exists a unique vector $U \in \mathscr{N}$ such that $x' = x \exp U$. Recall also that, for any $U \in \mathscr{N}$ and $x \in N$, the mapping $t \mapsto x\exp(tU)$ defines a smooth curve in N passing through x. For a nilpotent Lie algebra $\mathscr{N}$, the Campbell-Hausdorff-Baker formula says that, for any $U, V \in \mathscr{N}$

$$\exp U \exp V = \exp\left(U + V + \frac{1}{2}[U,V] + \frac{1}{12}([U,[U,V]] - [V,[U,V]) + \dots\right), \tag{2.1}$$

If $\mathscr{N}$ is a two-step nilpotent Lie algebra, that is, $[[U,V],W] = 0$, $U,V,W \in \mathscr{N}$, then $\exp U \exp V = \exp(U + V + \frac{1}{2}[U,V]$. If $a = \exp U, b = \exp V$, then $\log(ab) = \log(a) + \log(b) + \frac{1}{2}[\log(a), \log(b)]$, $\forall a, b \in N$, In particular, if a^{-1} denotes the inverse of $a \in N$, then $0 = \log(aa^{-1}) = \log(a) + \log(a^{-1}) + \frac{1}{2}[\log(a), \log(a^{-1})]$. Thus

$$\begin{aligned} \log(a^{-1}) &= -\log(a) - \frac{1}{2}[\log(a), \log(a^{-1})] \\ &= -\log(a) - \frac{1}{2}[\log(a), -\log(a) - \frac{1}{2}[\log(a), \log(a^{-1})]], \end{aligned}$$

that is, $\log(a^{-1}) = -\log(a)$.

Lemma 2.1. *Let $\mathscr{N}$ be a 2-step nilpotent Lie algebra, and let N denote the connected and simply connected corresponding two-step nilpotent Lie group. Let* $\exp : \mathscr{N} \to N$ *be the exponential mapping. Then, for any $U, A \in \mathscr{N}$,* $\mathrm{d}\exp_U : T_U\mathscr{N} \equiv \mathscr{N} \to T_{\exp U}N$ *is given by* $\mathrm{d}\exp_U(A) = \mathrm{d}L_{\exp U}\left(A + \frac{1}{2}[A,U]\right)$, *where $L_{\exp U}$ denotes the left translation by* $\exp U$.

Proof. Observe $\mathrm{d}\exp_U(A) = \frac{\mathrm{d}}{\mathrm{d}t}_{|t=0}(\exp(U+tA))$. Since $\mathscr{N}$ is a two-step nilpotent Lie algebra, then $\exp U \exp\left(t(A + \frac{1}{2}[A,U])\right) = \exp(U+tA)$, and thus

$$\mathrm{d}\exp_U(A) = \frac{\mathrm{d}}{\mathrm{d}t}_{|t=0}\left(L_{\exp U}(\exp(t(A + \frac{1}{2}[A,U]))\right).$$

On other hand, from $\mathrm{d}\exp_0 = id_{\mathscr{N}}$, one has $\mathrm{d}\exp_U(A) = \mathrm{d}L_{\exp U}(A + \frac{1}{2}[A,U])$. □

Remark 2.1. It is easy to construct a two-step nilpotent Lie algebras. Let $\mathscr{V}, \mathscr{Z}$ be any finite dimensional real vector spaces with bases $\{V_1, \dots, V_n\}$ for $\mathscr{V}$ and $\{Z_1, \dots, Z_m\}$ for $\mathscr{Z}$. Let $\mathscr{N} = \mathscr{V} \oplus \mathscr{Z}$ and define $[V_i, V_j] = \sum_{k=1}^m C_{ij}^k Z_k$, where the constants (C_{ij}^k) are chosen so that $C_{ij}^k = -C_{ji}^k$ for $1 \le i, j \le n$ and $1 \le k \le m$, but not all of them are zero. Define $[Z,V] = 0$ for all $Z \in \mathscr{Z}$ and $V \in \mathscr{V}$. The Jacoby identity is satisfied due t $[\mathscr{N}, \mathscr{N}] \subset \mathscr{Z}$.

From now on, $\mathscr{N}$ is a two-step nilpotent Lie algebra with nilpotent Lie group N.

3. Newton Method and Its Convergence Criteria

We begin by introducing some essential tools, which will be useful in the rest of this paper. Given a connected simply connected Lie group N, one introduces a distance $\rho(\cdot,\cdot)$ on N, which plays a key role in the study (see [27]). Let $x, y \in N$ and define $\rho(x,y) := \inf\{\sum_{i=1}^k \|U_i\|$, there exist $k \ge 1$ and $U_1, \dots, U_k \in \mathscr{N} : y = x\exp U_1 \dots \exp U_k\}$, where we adapt the convention $\inf \emptyset = +\infty$. We can see that ρ is a distance on N and that the topology induced by this distance is equivalent to the original one. On other hand, since $\mathscr{N}$ is nilpotent Lie algebra (and precisely is a two-step nilpotent Lie algebra), then the exponential mapping $\exp : \mathscr{N} \to N$ is a bijection, thus, for any $x, y \in N$, there exists a unique $U = \exp^{-1}(y^{-1}x) \in \mathscr{N}$ such that $x = y\exp U$. Hence $\rho(x,y) = \|U\|$. For $r > 0$, we use $C_r(x_0)$ to denote the open ball at x_0 with radius r defined by $C_r(x_0) = \{x \in N : \ x = x_0 \exp U, \ U \in \mathscr{N}, \|U\| < r\} = \{x \in N : \ \|\exp^{-1}(x_0^{-1}x)\| < r\}$, Next, let R be a positive constant, and let L be a non-negative nondecreasing function on $[0,R)$ satisfying $\int_0^R L(s)\mathrm{d}s \ge 1$. Then there exists a unique number $r_0 \in (0,R]$ such that

$$r_0 = \sup\left\{r \in (0,R) : \ \int_0^r L(s)\mathrm{d}s \le 1\right\}. \tag{3.1}$$

Define the real-valued function h on $(0,r_0)$ by $h(t) := \frac{1}{t}\int_0^t L(s)(t-s)\mathrm{d}s$ for each $t \in (0,r_0)$.

The following two lemmas provide some useful properties about the functions L and h. The first one was studied in [6, p. 170]. For the whole paper, let r_0 be given by (3.1).

Lemma 3.1. *The function h is strictly increasing on* $(0,r_0)$.

For the second lemma, note that the real-valued function χ on $(0,r_0)$ defined by

$$\chi(t) := 1 - \int_0^t L(s)\mathrm{d}s - \frac{1}{2}\int_0^t L(s)(t-s)\mathrm{d}s - \frac{1}{2}t, \quad \forall t \in (0,r_0),$$

is strictly decreasing on $(0,r_0)$, and satisfies that $\chi(0)>0$ and $\chi(r_0)<0$. Thus we use $\bar{r}_0$ to denote the unique number $\bar{r}_0\in(0,r_0)$ satisfying

$$1-\int_0^{\bar{r}_0}L(s)\mathrm{d}s-\frac{1}{2}\int_0^{\bar{r}_0}L(s)(\bar{r}_0-s)\mathrm{d}s-\frac{1}{2}\bar{r}_0=0.$$

Let q be the real-valued function on $(0,\bar{r}_0)$, defined by

$$q(t):=\frac{\frac{1}{t}\int_0^t L(s)s\mathrm{d}s}{1-\int_0^t L(s)\mathrm{d}s-\frac{1}{2}\int_0^t L(s)(t-s)\mathrm{d}s-\frac{1}{2}t}\quad \text{for each } t\in(0,\bar{r}_0). \tag{3.2}$$

Lemma 3.2. *([23] p.220) The functions $t\mapsto q(t)$ and $t\mapsto\frac{q(t)}{t}$ are strictly increasing on $(0,\bar{r}_0)$, and there is a unique number $r_1\in(0,\bar{r}_0)$ such that $q(r_1)=1$.*

The notion of Lipschitz conditions with L average in the following definition is an extension to the nilpotent Lie group setting of the corresponding one in Banach spaces, introduced by Wang [5, 6]. For similar extensions to general Lie groups, we refer to [18].

Definition 3.1. Let $r>0$ and $x^*\in N$ such that $\mathrm{d}f_{x^*}^{-1}$ exists. Then $\mathrm{d}f_{x^*}^{-1}\mathrm{d}f$ is said to satisfy:

(1) the center Lipschitz condition with L average on $C_r(x^*)$ if $\|\mathrm{d}f_{x^*}^{-1}(\mathrm{d}f_{x^*\exp U}-\mathrm{d}f_{x^*})\|\le\int_0^{\|U\|}L(s)\mathrm{d}s$ holds for each $U\in\mathcal{N}$ with $\|U\|<r$.

(2) the radius Lipschitz condition with L average on $C_r(x^*)$ if $\|\mathrm{d}f_{x^*}^{-1}(\mathrm{d}f_{x^*\exp U}-\mathrm{d}f_{x^*\exp\tau U})\|\le\int_{\tau\|U\|}^{\|U\|}L(s)\mathrm{d}s$ holds for any $\tau\in[0,1]$ and $U\in\mathcal{N}$ with $\|U\|<r$.

Obviously, the radius Lipschitz condition implies the center Lipschitz condition.

Lemma 3.3. *Let $x^*\in N$ be such that $\mathrm{d}f_{x^*}^{-1}$ exists. Suppose that $\mathrm{d}f_{x^*}^{-1}\mathrm{d}f$ satisfies the center Lipschitz condition with L average on $C_{r_0}(x^*)$. Let $x\in C_{r_0}(x^*)$ such that $x=x^*\exp U$ with $\|U\|<r_0$. Then $\mathrm{d}f_x^{-1}$ exists and satisfies $\|\mathrm{d}f_x^{-1}\mathrm{d}f_{x^*}\|=\|\mathrm{d}f_{x^*\exp U}^{-1}\mathrm{d}f_{x^*}\|\le\left(1-\int_0^{\|U\|}L(s)\mathrm{d}s\right)^{-1}$.*

Proof. Since $\mathrm{d}f_{x^*}^{-1}\mathrm{d}f$ satisfies the center Lipschitz condition with L average on $C_{r_0}(x^*)$. If $x=x^*\exp U$ with $\|U\|<r_0$, we can write

$$\|\mathrm{d}f_{x^*\exp U}^{-1}\mathrm{d}f_{x^*}-\mathbf{I}\|=\|\mathrm{d}f_{x^*\exp U}^{-1}(\mathrm{d}f_{x^*}-\mathrm{d}f_{x^*\exp U})\|\le\int_0^{\|U\|}L(s)\mathrm{d}s<\int_0^{r_0}L(s)\mathrm{d}s<1.$$

Then, we have that $\mathrm{d}f_{x^*\exp U}^{-1}$ exists and satisfies

$$\|\mathrm{d}f_{x^*\exp U}^{-1}\mathrm{d}f_{x^*}\|\le\left(1-\int_0^{\|U\|}L(s)\mathrm{d}s\right)^{-1},\quad U\in\mathcal{N},\|U\|<r_0.$$

□

Newton's method with initial point $x_0\in N$ for f is defined as follows:

$$x_{k+1}=x_k\exp(-\mathrm{d}f_{x_k}^{-1}(f(x_k)))\quad k=0,1,\cdots. \tag{3.3}$$

Our first main theorem is as follows.

Theorem 3.1. *Let $f:N\mapsto\mathcal{N}$ be a C^1-mapping, and and let $x^*\in N$ with $f(x^*)=0$. Let $r\in(0,r_1]$. Suppose that $\mathrm{d}f_{x^*}^{-1}$ exists and that $\mathrm{d}f_{x^*}^{-1}\mathrm{d}f$ satisfies the radius Lipschitz condition with L average on $C_r(x^*)$. Then, for any $x_0\in C_r(x^*)$, the sequence generated by Newton's method*

(3.3) *with initial point* x_0 *is well defined, and is quadratically convergent to* x^*. *Moreover, the following estimates hold, for any* $k = 0, 1, \ldots,$

$$\|\exp^{-1}(x^{*-1}x_{k+1})\| \leq \frac{q_0}{\|u_0\|}\|\exp^{-1}(x^{*-1}x_k)\|^2 \tag{3.4}$$

and

$$\|\exp^{-1}(x^{*-1}x_k)\| \leq q_0^{2^k-1}\|U_0\|, \tag{3.5}$$

where $U_0 := \exp^{-1}(x^{*-1}x_0)$ *and*

$$q_0 := q(\|U_0\|) = \frac{\frac{1}{\|u_0\|}\int_0^{\|U_0\|} L(s)s\mathrm{d}s}{1 - \int_0^{\|U_0\|} L(s)\mathrm{d}s - \frac{1}{2}\int_0^{\|U_0\|} L(s)(\|U_0\| - s)\mathrm{d}s - \frac{1}{2}\|U_0\|} < 1.$$

Proof. Write

$$U_k := \exp^{-1}(x^{*-1}x_k), \quad \forall k = 0, 1, \cdots. \tag{3.6}$$

Fix k. We first verify the following implication:

$$[x_k \text{ exists and } \|U_k\| < r] \Rightarrow [x_{k+1} \text{ exists and } \|U_{k+1}\| \leq q(\|U_k\|)\,\|U_k\| \leq \|U_k\| < r] \tag{3.7}$$

To do this, we assume that x_k exists and $\|U_k\| < r$. Then, by (3.6), we have $x_k = x^* \exp U_k$. From the assumption that $\mathrm{d}f_{x^*}^{-1} \circ \mathrm{d}f$ satisfies the radius Lipschitz condition in $C_r(x^*)$, it follows from Lemma 3.3 that $\mathrm{d}f_{x_k}^{-1}$ exists, and

$$\|\mathrm{d}f_{x_k}^{-1}\mathrm{d}f_{x^*}\| \leq \left(1 - \int_0^{\|U_k\|} L(s)\mathrm{d}s\right)^{-1}. \tag{3.8}$$

Hence, $x_{k+1} = x_k \exp(-\mathrm{d}f_{x_k}^{-1}f(x_k))$ exists. Write $V_k = -\mathrm{d}f_{x_k}^{-1}f(x_k)$. Noting that $x_k = x^* \exp U_k$, we find from (2.1) that $x_{k+1} = x^* \exp U_k \exp V_k = x^* \exp(U_k + V_k + \frac{1}{2}[U_k, V_k])$. On the other hand, it follows from (3.6) that $x_{k+1} = x^* \exp U_{k+1}$. Since exp is a global diffeomorphism, we have

$$U_{k+1} = U_k + V_k + \frac{1}{2}[U_k, V_k]. \tag{3.9}$$

Furthermore, since $\mathscr{N}$ is two-step nilpotent Lie algebra, we obtain that $[U_{k+1}, V_k] = [U_k, V_k]$. This, together with (3.9), gives that

$$U_{k+1} = U_k + V_k + \frac{1}{2}[U_{k+1}, V_k]. \tag{3.10}$$

Note that

$$\begin{aligned} \|U_k + V_k\| &= \|U_k - \mathrm{d}f_{x_k}^{-1}(f(x_k) - f(x^*))\| \\ &= \|(\mathrm{d}f_{x_k}^{-1} \circ \mathrm{d}f_{x^*})\int_0^1 \mathrm{d}f_{x^*}^{-1}(\mathrm{d}f_{x_k} - \mathrm{d}f_{x^* \exp sU_k})(U_k)\mathrm{d}s \\ &\leq \|\mathrm{d}f_{x_k}^{-1} \circ \mathrm{d}f_{x^*}\|\|\int_0^1 \mathrm{d}f_{x^*}^{-1}(\mathrm{d}f_{x_k} - \mathrm{d}f_{x^* \exp sU_k})(U_k)\mathrm{d}s\|. \end{aligned} \tag{3.11}$$

Then, we conclude that

$$\|\int_0^1 \mathrm{d}f_{x^*}^{-1}(\mathrm{d}f_{x_k} - \mathrm{d}f_{x^* \exp sU_k})(U_k)\mathrm{d}s\| \leq \int_0^1 \int_{s\|U_k\|}^{\|U_k\|} L(\tau)\|U_k\|\mathrm{d}\tau\mathrm{d}s = \int_0^{\|U_k\|} L(s)s\mathrm{d}s.$$

Combining this, (3.8), and (3.11) gives that

$$\|U_k + V_k\| \leq \frac{\int_0^{\|U_k\|} L(s)s\mathrm{d}s}{1 - \int_0^{\|U_k\|} L(s)\mathrm{d}s}. \tag{3.12}$$

Furthermore, we observe that

$$\begin{aligned}
\|\mathrm{d}f_{x^*}^{-1}(f(x_k))\| &= \|\mathrm{d}f_{x^*}^{-1}(f(x_k)-f(x^*))\| \\
&= \|\int_0^1 \mathrm{d}f_{x^*}^{-1}\circ(\mathrm{d}f_{x^*\exp sU_k}-\mathrm{d}f_{x^*})(U_k)\mathrm{d}s+U_k\| \\
&\leq \int_0^1 \|\mathrm{d}f_{x^*}^{-1}\circ(\mathrm{d}f_{x^*\exp sU_k}-\mathrm{d}f_{x^*})\|\mathrm{d}s\|U_k\|+\|U_k\| \\
&\leq \int_0^1\int_0^{s\|U_k\|} L(\tau)\mathrm{d}\tau\mathrm{d}s\|U_k\|+\|U_k\| \\
&= \int_0^{\|U_k\|} L(s)(\|U_k\|-s)\mathrm{d}s+\|U_k\|.
\end{aligned}$$

This together with (3.8) implies that

$$\|V_k\| \leq \|\mathrm{d}f_{x_k}^{-1}\circ\mathrm{d}f_{x^*}\|\|\mathrm{d}f_{x^*}^{-1}(f(x_k))\| \leq \frac{\int_0^{\|U_k\|} L(s)(\|U_k\|-s)\mathrm{d}s+\|U_k\|}{1-\int_0^{\|U_k\|} L(s)\mathrm{d}s}. \tag{3.13}$$

Therefore, we have that

$$\|\frac{1}{2}[U_{k+1},V_k]\| \leq \frac{1}{2}\|U_{k+1}\|\|V_k\| \leq \|U_{k+1}\|\frac{\int_0^{\|U_k\|} L(s)(\|U_k\|-s)\mathrm{d}s+\|U_k\|}{2(1-\int_0^{\|U_k\|} L(s)\mathrm{d}s)}.$$

Thus it follows from (3.10) and (3.12) that

$$\|U_{k+1}\| \leq \frac{\int_0^{\|U_k\|} L(s)s\mathrm{d}s}{1-\int_0^{\|U_k\|} L(s)\mathrm{d}s}+\|U_{k+1}\|\frac{\int_0^{\|U_k\|} L(s)(\|U_k\|-s)\mathrm{d}s+\|U_k\|}{2(1-\int_0^{\|U_k\|} L(s)\mathrm{d}s)},$$

Observe that q is defined by (3.2). One sees that $\|U_{k+1}\| \leq q(\|U_k\|)\|U_k\| \leq \|U_k\| < r$, because $q(\|U_k\|) \leq q(r) \leq q(r_1) = 1$ by assertions (ii) and (iii) of Lemma 3.2. Hence, implication (3.7) holds. Below, we show that (3.4) holds for each $k = 0,1,\cdots$. Let $x_0 \in C_r(x^*)$. Then $x_0 = x^*\exp(\exp^{-1}(x^{*-1}x_0)) = x^*\exp U_0$ and $\|U_0\| < r$. It follows from (3.7) that x_1 exists and $\|U_1\| \leq q(\|U_0\|)\|U_0\| \leq \|U_0\| < r$. Moreover, $\|U_1\| \leq q(\|U_0\|)\|U_0\| = \frac{q(\|U_0\|)}{\|U_0\|}\|U_0\|^2$. Hence, (3.4) holds for $k = 0$. Inductively, assume that x_j exists and $\|U_j\| < r$. Then, by (3.7), x_{j+1} exists and $\|U_{j+1}\| \leq q(\|U_j\|)\|U_j\| \leq \|U_j\| < r$. Thus, it follows that

$$\|U_{j+1}\| \leq q(\|U_j\|)\|U_j\| = \frac{q(\|U_j\|)}{\|U_j\|}\|U_j\|^2 \leq \frac{q(\|U_0\|)}{\|U_0\|}\|U_j\|^2,$$

because $t \to \frac{q(t)}{t}$ is strictly increasing due to Lemma 3.2. Therefore, (3.4) holds for $k = j$.

Finally, we prove by induction that (3.5) holds. It is clear that (3.5) holds for $k = 0$. Suppose that (3.5) holds for $k = i$, that is, $\|U_i\| \leq q_0^{2^i-1}\|U_0\|$. It follows from (3.4) that

$$\|U_{i+1}\| \leq \frac{q_0}{\|U_0\|}\|U_i\|^2 \leq \frac{q_0}{\|U_0\|}(q_0^{2^i-1}\|U_0\|)^2 = q_0^{2^{i+1}-1}\|U_0\|,$$

that is, (3.5) holds for $k = i+1$. The proof is complete. □

Remark 3.1. There is another approach to estimate the radius of convergence balls for the Newton method. In fact, in the proof of Theorem 3.1, we could use (3.9), instead of (3.10), to estimate the norm $\|U_{k+1}\|$ to obtain

$$\|U_{k+1}\| = \|U_k+V_k\|+\frac{1}{2}\|[U_k,V_k]\| \leq \|U_k+V_k\|+\frac{1}{2}\|U_k\|\|V_k\|.$$

Thus, by (3.12) and (3.13), we have that

$$\|U_{k+1}\| \leq \frac{\int_0^{\|U_k\|} L(s)s\mathrm{d}s}{1-\int_0^{\|U_k\|} L(s)\mathrm{d}s} + \frac{1}{2}\|U_k\| \frac{\int_0^{\|U_k\|} L(s)(\|U_k\|-s)\mathrm{d}s + \|U_k\|}{1-\int_0^{\|U_k\|} L(s)\mathrm{d}s} = p(\|U_k\|)\|U_k\|, \quad (3.14)$$

where the function p on $(0,r_0)$ is defined by

$$p(t) := \frac{\frac{1}{t}\int_0^t L(s)s\mathrm{d}s + \frac{1}{2}\int_0^t L(s)(t-s)\mathrm{d}s + \frac{t}{2}}{1-\int_0^t L(s)\mathrm{d}s} \quad \text{for each } t \in (0,r_0).$$

With a similar arguments to the proof of Lemma 3.2, we can verify that p is strictly increasing on $(0,r_0)$, and there is a unique solution $r_2 \in (0,r_0)$ such that $p(r_2)=1$, and the function $t \mapsto \frac{p(t)}{t}$ is strictly increasing on $(0,r_0)$. Moreover, we have that $r_1 = r_2$. Then, as in the proof, we can use (3.14) to show by mathematical induction that

$$\|\exp^{-1}(x^{*-1}x_{k+1})\| = \|U_{k+1}\| \leq \frac{p(\|U_0\|)}{\|U_0\|}\|U_k\|^2 = \frac{p(\|U_0\|)}{\|U_0\|}\|\exp^{-1}(x^{*-1}x_k)\|^2,$$

and $\|\exp^{-1}(x^{*-1}x_k)\| \leq p(\|U_0\|)^{2^k-1}\|U_0\|$. However, estimate (3.5) is better than the above one because $q(t) \leq p(t)$, $\forall t \in (0,r_1)$.

Our next theorem is concerned with the estimate of the uniqueness ball for zeros of f.

Theorem 3.2. *Let $f : N \mapsto \mathcal{N}$ be a C^1-mapping and and let $x^* \in N$ with $f(x^*) = 0$. Let $R_0 > 0$ be such that*

$$\frac{1}{R_0}\int_0^{R_0} L(s)(R_0-s)\mathrm{d}s \leq 1. \quad (3.15)$$

Suppose that $\mathrm{d}f_{x^*}^{-1}$ exists and that $\mathrm{d}f_{x^*}^{-1}\mathrm{d}f$ satisfies the center Lipschitz condition with L average on $C_{R_0}(x^*)$. Then x^* is the unique zero of f on $C_{R_0}(x^*)$.

Proof. Assume that on the contrary that y^* is another zero of f in $C_{R_0}(x^*)$. Then, there exists $U \in \mathcal{N}$ such that $y^* = x^* \exp U$ with $0 < \|U\| < R_0$. Since $\mathrm{d}f_{x^*}^{-1}\mathrm{d}f$ satisfies the center Lipschitz condition with L average on $C_{R_0}(x^*)$, we have

$$\begin{aligned}
\|U\| &= \|-\mathrm{d}f_{x^*}^{-1}\textstyle\int_0^1 \mathrm{d}f_{x^*\exp(\tau U)}\mathrm{d}\tau + U\| \\
&= \|-\textstyle\int_0^1 \mathrm{d}f_{x^*}^{-1}(\mathrm{d}f_{x^*\exp(\tau U)} - \mathrm{d}f_{x^*}U\mathrm{d}\tau\| \\
&\leq \textstyle\int_0^1\int_0^{\tau\|U\|} L(s)\mathrm{d}s\|U\|\mathrm{d}\tau \\
&= \textstyle\int_0^{\|U\|} L(s)(\|U\|-s)\mathrm{d}s.
\end{aligned}$$

This gives that $1 \leq \frac{1}{\|U\|}\int_0^{\|U\|} L(s)(\|U\|-s)\mathrm{d}s$. It follows from Lemma 3.1 that

$$1 \leq \frac{1}{\|U\|}\int_0^{\|U\|} L(s)(\|U\|-s)\mathrm{d}s < \frac{1}{R_0}\int_0^{R_0} L(s)(r-s)\mathrm{d}s \leq 1,$$

which is a contradiction. Hence, we have $y^* = x^*$. □

4. APPLICATIONS

This section is devoted to the applications of our main results for some special cases, such as, the classical Lipschitz condition and the γ-condition.

4.1. **The classical Lipschitz condition.** Let $r>0$ and $L>0$. Let $x^*\in N$ be such that $\mathrm{d}f_{x^*}^{-1}$ exists. Recall that $\mathrm{d}f_{x^*}^{-1}\mathrm{d}f$ is said to satisfy ([5]):

(i) the center Lipschitz condition with L on $C_r(x^*)$ if $\|\mathrm{d}f_{x^*}^{-1}(\mathrm{d}f_{x^*\exp u}-\mathrm{d}f_{x^*})\|\le L\|u\|$ holds for each $u\in\mathcal{N}$ with $\|u\|\le r$.

(ii) the radius Lipschitz condition with L on $C_r(x^*)$ if $\|\mathrm{d}f_{x^*}^{-1}(\mathrm{d}f_{x^*\exp u}-\mathrm{d}f_{x^*\exp \tau u})\|\le L(1-\tau)\|u\|$ holds for any $\tau\in[0,1]$ and $u\in\mathcal{N}$ with $\|u\|\le r$.

Since function L is a constant, we have from (3.2) that

$$q(t):=\frac{\frac{L}{2}t}{1-(L+\frac{1}{2})t-L(\frac{t}{2})^2}.$$

As $q(r_1)=1$, it follows that

$$r_1=\frac{-(3L+1)+\sqrt{(3L+1)^2+4L}}{L}. \tag{4.1}$$

The following corollary directly follows from Theorem 3.1.

Corollary 4.1. *Let $f:N\mapsto\mathcal{N}$ be a C^1-mapping, and and let $x^*\in N$ with $f(x^*)=0$. Let $r\in(0,r_1]$, where r_1 is given by (4.1). Suppose that $\mathrm{d}f_{x^*}^{-1}$ exists, and $\mathrm{d}f_{x^*}^{-1}\mathrm{d}f$ satisfies the radius Lipschitz condition with constant L on $C_r(x^*)$. Then, for any $x_0\in C_r(x^*)$, the sequence generated by the Newton method (3.3) with initial point x_0 is well defined, and is quadratically convergent to x^*. Moreover, the following estimates hold, for any $k=0,1,\ldots$,*

$$\|\exp^{-1}(x^{*-1}x_{k+1})\|\le\frac{q_0}{\|U_0\|}\|\exp^{-1}(x^{*-1}x_k)\|^2$$

and $\|\exp^{-1}(x^{*-1}x_k)\|\le q_0^{2^k-1}\|U_0\|$, where $U_0:=\exp^{-1}(x^{*-1}x_0)$, and

$$q_0:=q(\|U_0\|)=\frac{\frac{L}{2}\|U_0\|}{1-(L+\frac{1}{2})\|U_0\|-\frac{L}{4}\|U_0\|^2}<1.$$

Remark 4.1. It was proved in [18, Corollary 3.4] that, under the L-Lipschitz condition, Newton's method (3.3) with initial point $x_0\in C_{\frac{1}{4L}}(x^*)$ is well-defined and convergent. Note that

$$\frac{1}{4L}<\frac{-(3L+1)+\sqrt{(3L+1)^2+2L}}{L}\Longleftrightarrow L>\frac{9}{40}.$$

This means that Corollary 4.1 improves [18, Corollary 3.4] in the case that $L>\frac{9}{40}$.

The following corollary directly follows from Theorem 3.2.

Corollary 4.2. *Let $L>0$. Let $f:N\mapsto\mathcal{N}$ be a C^1-mapping, and and let $x^*\in N$ with $f(x^*)=0$. Let $0<r\le\frac{2}{L}$. Suppose that $\mathrm{d}f_{x^*}^{-1}$ exists and that $\mathrm{d}f_{x^*}^{-1}\mathrm{d}f$ satisfies the center Lipschitz condition with constant L on $C_r(x^*)$. Then x^* is the unique zero of f on $C_r(x^*)$.*

Proof. Since $L(t)\equiv L$, it follows from (3.15) that $\frac{1}{R_0}\int_0^{R_0}L(R_0-s)\mathrm{d}s\le 1$. So $R_0\le\frac{2}{L}$. In view of Theorem 3.2, one immediately concludes the desired conclusion. $\square$

4.2. **The γ-condition.** Let k be a positive integer, and assume that $f : N \to \mathscr{N}$ is a C^k-map. Following [18], define the map $\mathrm{d}^k f_x : \mathscr{N}^k \to \mathscr{N}$ by

$$\mathrm{d}^k f_x U_1 \cdots U_k = \left(\frac{\partial^k}{\partial t_k \cdots \partial t_1} f(x \cdot \exp t_k U_k \cdots \exp t_1 U_1) \right)_{t_k = \cdots = t_1 = 0}$$

for each $(U_1, \cdots, U_k) \in \mathscr{N}^k$. In particular,

$$\mathrm{d}^k f_x U^k = \left(\frac{\mathrm{d}^k}{\mathrm{d}t^k} f(x \cdot \exp tU) \right)_{t=0}, \quad \forall U \in \mathscr{N}.$$

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

$$\mathrm{d}^k f_x U_1 \cdots U_k = \mathrm{d}^{k-i} \left(\mathrm{d}^i f_\cdot (U_1 \cdots U_i) \right)_x U_{i+1} \cdots U_k \quad \text{for each } (U_1, \cdots, U_k) \in \mathscr{N}^k. \tag{4.2}$$

In particular, for fixed $U_1, \cdots, U_{i-1}, U_{i+1}, \cdots, U_k \in \mathscr{N}$, $\mathrm{d}^i f_x U_1 \cdots U_{i-1} = \mathrm{d} \left(\mathrm{d}^{i-1} f_\cdot (U_1 \cdots U_{i-1}) \right)_x (\cdot)$. This implies that, for each $i = 1, \cdots, k$, $\mathrm{d}^i f_x U_1 \cdots U_{i-1} u$ is linear with respect to $u \in \mathscr{N}$, so is $\mathrm{d}^k f_x U_1 \cdots U_{i-1} u U_{i+1} \cdots U_k$ by (4.2). Consequently, $\mathrm{d}^k f_x$ is a multilinear map from $\mathscr{N}^k$ to $\mathscr{N}$ because $1 \le i \le k$ is arbitrary. Thus we can define the norm of $\mathrm{d}^k f_x$ by $\| \mathrm{d}^k f_x \| = \sup\{ \| \mathrm{d}^k f_x U_1 U_2 \cdots U_k \| : (U_1, \cdots, U_k) \in \mathscr{N}^k \text{ with each } \|U_j\| = 1\}$.

For the rest, we assume that $k = 2$, that is, $f : N \to \mathscr{N}$ is a C^2-map. Let $\gamma, r > 0$ be such that $\gamma r < 1$. Following [18], we extend the notion of the γ-condition for mappings on two-step nilpotent Lie groups.

Definition 4.1. Let $x^* \in N$ be such that $\mathrm{d}f_{x^*}^{-1}$ exists. f is said to satisfy the γ-condition at x^* on $\mathbf{B}(x^*, r)$ (the open ball at x^* of radius r in $\mathscr{N}$) if, for any $x \in C_r(x^*)$ with $x = x^* \exp(U)$, $U \in \mathscr{N}$ and $\|U\| < r$,

$$\|\mathrm{d}f_{x^*}^{-1} \mathrm{d}^2 f_x\| \le \frac{2\gamma}{(1 - \gamma \|U\|)^3}. \tag{4.3}$$

The following proposition shows that the γ-condition implies the radius Lipschitz condition with L average, where the function L is defined by

$$L(s) := \frac{2\gamma}{(1 - \gamma s)^3}, \quad \forall s \in [0, r). \tag{4.4}$$

Proposition 4.1. *Let L be given by (4.4). Let $x^* \in N$ be such that $\mathrm{d}f_{x^*}^{-1}$ exists. Suppose that f satisfies the γ-condition at x^* on $\mathbf{B}(x^*, r)$. Then $\mathrm{d}f_{x^*}^{-1}\mathrm{d}f$ satisfies the radius Lipschitz condition with L average on $C_r(x^*)$.*

Proof. Let $\tau \in [0, 1]$ and $u \in \mathscr{N}$ with $\|U\| < r$. Note that

$$\mathrm{d}f_{x^* \exp U} - \mathrm{d}f_{x^* \exp \tau U} = \int_\tau^1 \mathrm{d}^2 f_{x^* \exp(sU)} U \mathrm{d}s. \tag{4.5}$$

Since f satisfies the γ-condition at x^* on $\mathbf{B}(x^*, r)$, it follows from (4.3) and (4.5) that

$$\begin{aligned} \|\mathrm{d}f_{x^*}^{-1}(\mathrm{d}f_{x^* \exp U} - \mathrm{d}f_{x^* \exp \tau U})\| &= \|\mathrm{d}f_{x^*}^{-1} \int_\tau^1 \mathrm{d}^2 f_{x^* \exp(sU)} U \mathrm{d}s\| \\ &\le \int_\tau^1 \frac{2\gamma\|U\|}{(1 - \gamma s\|U\|)^3} \mathrm{d}s \\ &= \int_{\tau\|U\|}^{\|U\|} \frac{2\gamma}{(1 - \gamma t)^3} \mathrm{d}t. \end{aligned}$$

Thus, the conclusion follows and the proof is complete. $\square$

Since the function L is given by (4.4), we have from (3.2) that

$$q(t) := \frac{\gamma t}{1-(\frac{1}{2}+4\gamma)t+(2\gamma^2+\frac{\gamma}{2})t^2}.$$

As $q(r_1)=1$, it follows that

$$r_1 = \frac{5\gamma+\frac{1}{2}-\sqrt{17\gamma^2+3\gamma+\frac{1}{4}}}{4\gamma^2+\gamma}. \tag{4.6}$$

The following corollary directly follows from Proposition 4.1 and Theorem 3.1.

Corollary 4.3. *Let $\gamma>0$. Let $f: N \mapsto \mathcal{N}$ be a C^1-mapping, and and let $x^* \in N$ with $f(x^*)=0$. Let $r\in(0,r_1]$ with r_1 given by (4.6). Suppose that $\mathrm{d}f_{x^*}^{-1}$ exists and that f satisfies the γ-condition at x^* on $\mathbf{B}(x^*,r)$. Then, for any $x_0\in C_r(x^*)$, the sequence generated by Newton method (3.3) with initial point x_0 is well defined and is quadratically convergent to x^*. Moreover, the following estimates hold, for any $k=0,1,\ldots$, $\|\exp^{-1}(x^{*-1}x_{k+1})\| \le \frac{q_0}{\|U_0\|}\|\exp^{-1}(x^{*-1}x_k)\|^2$, and $\|\exp^{-1}(x^{*-1}x_k)\| \le q_0^{2^k-1}\|U_0\|$, where $U_0:=\exp^{-1}(x^{*-1}x_0)$ and*

$$q_0 := q(\|U_0\|) = \frac{\gamma\|U_0\|}{1-(\frac{1}{2}+4\gamma)\|U_0\|+(2\gamma^2+\frac{\gamma}{2})\|U_0\|^2} < 1.$$

Remark 4.2. Let $a_0 = 0.080851\cdots$ be the smallest positive root of the equation $\frac{u}{1-4u+2u^2} = 3-2\sqrt{2}$. It was improved in [19, Corollary 4.1] that, under the γ-condition, Newton's method (3.3) with any initial point $x_0\in C_{\frac{a_0}{\gamma}}(x^*)$ is well-defined and convergent.

Note that if $\gamma > \frac{\sqrt{a_0}}{2}$, then $\frac{a_0}{\gamma} < r_1 = \frac{5\gamma+\frac{1}{2}-\sqrt{17\gamma^2+3\gamma+\frac{1}{4}}}{4\gamma^2+\gamma}$. This means that Corollary 4.3 improves [19, Corollary 4.1] in the case that $\gamma > \frac{\sqrt{a_0}}{2}$.

Since the radius Lipschitz condition $\Longrightarrow$ the center Lipschitz condition, the following corollary is direct from Theorem 3.2 and Proposition 4.1.

Corollary 4.4. *Let $f: N\mapsto\mathcal{N}$ be a C^1-mapping and and let $x^*\in N$ with $f(x^*)=0$. Let $\gamma,r>0$ be such that $\gamma r<\frac{1}{2}$. Suppose that $\mathrm{d}f_{x^*}^{-1}$ exists and that $\mathrm{d}f_{x^*}^{-1}\mathrm{d}f$ satisfies the γ-condition at x^* on $\mathbf{B}(x^*,r)$. Then x^* is the unique zero of f on $C_r(x^*)$.*

Proof. Since $\mathrm{d}f_{x^*}^{-1}$ exists and that $\mathrm{d}f_{x^*}^{-1}\mathrm{d}f$ satisfies the γ-condition at x^* on $\mathbf{B}(x^*,r)$, it follows from Proposition 4.1 that that $\mathrm{d}f_{x^*}^{-1}\mathrm{d}f$ satisfies the center Lipschitz condition with L average on $C_r(x^*)$, where L is given by (4.4). Thus, it follows from (3.15) that $\frac{1}{R_0}\int_0^{R_0}\frac{2\gamma}{(1-\gamma s)^3}(R_0-s)\mathrm{d}s \le 1$. This gives that $R_0\gamma\le\frac{1}{2}$. Hence, the conclusion follows from Theorem 3.2. $\square$

4.3. **Approximate zeros.** Throughout this section, we always assume that f is analytic on N.

For $x\in N$ such that $\mathrm{d}f_x^{-1}$ exists, we define

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

Also we adopt the convention that $\gamma(f,x)=\infty$ if $\mathrm{d}f_x$ is not invertible. Note that this definition is justified and, in the case that $\mathrm{d}f_x$ is invertible, $\gamma(f,x)$ is finite by analyticity. The Taylor formula for a real-valued function on N can be found in [24, P.95]; and its extension to the map from N to $\mathcal{N}$ is trivial.

Proposition 4.2. *Let $\gamma=\gamma(f,x_0)$. Then f satisfies the γ-condition at x_0 on $\mathbf{B}(x_0,\frac{1}{\gamma})$.*

This section is devoted to the determination of an approximate zero point of an analytic mapping. For the purpose, we first recall the notion of the approximate zero of an analytic mapping F from the domain U in a Banach space to another. The following unified definition is taken from [5, Definition 7.1]. Consider Newton's iteration with initial point x_0:

$$x_{k+1}=x_k-F'(x_k)^{-1}F(x_k), \quad \forall k=0,1,2,\cdots. \tag{4.7}$$

Definition 4.2. Suppose that $x_0\in U$ is such that Newton's iteration (4.7) is well-defined for F and satisfies $\|x_k-x^*\|\le\left(\frac{1}{2}\right)^{2^{k-1}}\|x_0-x^*\|$, $\forall k=1,2,\cdots$, where x^* is a solution of $F(x)=0$. Then x_0 is said to be an approximate zero of F.

The notion of the approximate zero in the sense of $\|x_{n+1}-x_n\|$ was introduced in [7]; while the second kind of approximate zero was defined in the sense of $\|x_n-x^*\|$ in [7]. A more reasonable definition for the second kind was given in [28] and [29]. It was also studied by Wang [30]. The notion of the approximate zero in the sense of $\|F'(x_0)^{-1}F(x_n)\|$ was defined in [31] and, as shown in [5], it is equivalent to that in the sense of $\|x_{n+1}-x_n\|$, or equivalently, in the sense of $\|F'(x_n)^{-1}F(x_n)\|$. Definition 4.3 in the following extends the notion of the approximate zero to the case of the mappings on Heisenberg group N.

Definition 4.3. Suppose $x_0\in N$ is such that Newton's method (3.3) is well-defined for f, and satisfies $\|\exp^{-1}(x^{*-1}x_k)\|\le\left(\frac{1}{2}\right)^{2^{k-1}}\|\exp^{-1}(x^{*-1}x_0)\|$, $\forall k=1,2,\cdots$, where x^* is a solution of $f(x)=0$. Then x_0 is said to be an approximate singular point of f.

Let $\gamma=\gamma(f,x_0)$. Let r_1 be the smaller solution of

$$q(t):=\frac{\gamma t}{1-(\frac{1}{2}+4\gamma)t+(2\gamma^2+\frac{\gamma}{2})t^2}=\frac{1}{2},$$

that is,

$$r_1=\frac{6\gamma+\frac{1}{2}-\sqrt{28\gamma^2+4\gamma+\frac{1}{4}}}{\gamma(4\gamma+1)}. \tag{4.8}$$

Clearly, $r_1<\frac{1}{\gamma}$.

Corollary 4.5. *Let $\gamma=\gamma(f,x_0)$. Let $f:N\mapsto\mathcal{N}$ be analytic, and and let $x^*\in N$ with $f(x^*)=0$. Let $r\in(0,r_1]$ with r_1 given by (4.8). Suppose that $\mathrm{d}f_{x^*}^{-1}$ exists. Let $x_0\in C_r(x^*)$. Then x_0 is an approximate singular point of f.*

Proof. Since f is analytic, it follows from 4.2 that f satisfies the pieces γ-condition at x_0 on $\mathbf{B}(x_0,\frac{1}{\gamma})$. Then, Corollary 4.3 is applicable to conclude that the sequence generated by Newton method (3.3) with initial point x_0 is well defined, and is quadratically convergent to x^*. Moreover, the following estimate holds, for any $k=0,1,\ldots$, $\|\exp^{-1}(x^{*-1}x_k)\|\le q_0^{2^k-1}\|U_0\|$, where

$U_0 := \exp^{-1}(x^{*-1}x_0)$, and

$$q_0 := q(\|U_0\|) = \frac{\gamma\|U_0\|}{1-(\frac{1}{2}+4\gamma)\|U_0\|+(2\gamma^2+\frac{\gamma}{2})\|U_0\|^2}.$$

Since $\|U_0\| < r_1$, $q(r_1) = \frac{1}{2}$, and $q(t)$ is strictly increasing on $(0, r_1)$, we obtain that, for any $k = 0, 1, \ldots$, $\|\exp^{-1}(x^{*-1}x_k)\| \le (\frac{1}{2})^{2^k-1}\|U_0\|$. Hence, x_0 is an approximate singular point of f. This completes the proof. □

Acknowledgments

This work was supported by NSTIP strategic technologic program number (13-MAT874-02) in the Kingdom of Saud Arabia.

REFERENCES

[1] S.T. Smith, Optimization techniques on Riemannian manifolds, in: Fields Institute Communications, vol. 3, pp. 113-146, American Mathematical Society, Providence, RI, 1994.

[2] R.E. Mahony, The constrained Newton method on a Lie group and the symmetric eigenvalue problem, Linear Algebra Appl. 248 (1996), 67-89.

[3] R. Adler, J.P. Dedieu, J. Margulies, M. Martens, M. Shub, Newton's method on Riemannian manifolds and a geometric model for human spine, IMA J. Numer. Anal. 22 (2002), 1-32.

[4] B. Owren, B. Welfert, The Newton iteration on Lie groups, BIT Numer. 40 (2000), 121-145.

[5] X. Wang, Convergence of Newton's method and uniqueness of the solution of equations in Banach space, IMA J. Numer. Anal. 20 (2000), 123-134.

[6] X. Wang, Convergence of Newton's method and inverse function theorem in Banach spaces, Math. Comput. 225 (1999), 169-186.

[7] S. Smale, Newton's method estimates from data at one point, In: R. Ewing, K. Gross, C. Martin (Eds.), The Merging of Disciplines NewDirections in Pure, Applied and Computational Mathematics, pp. 185-196, Springer, New York, 1986.

[8] A. Edelman, T.A. Arias, T. Smith, The geometry of algorithms with orthogonality constraints, SIAM J. Matrix Anal. Appl. 20 (1998), 303-353.

[9] C. Udriste, Convex Functions and Optimization Methods on Riemannian Manifolds, Mathematics and Its Applications, Vol. 297. Kluwer Academic Publishers, Dordrecht, 1994.

[10] P.A. Absil, C.G. Baker, K.A. Gallivan, Trust-region methods on Riemannian manifolds, Found. Comput. Math. 7 (2007), 303-330.

[11] F. Alvarez, J., Bolte, J. Munier, A unifying local convergence result for Newton's method in Riemannian manifolds, Found. Comput. Math. 8 (2008), 197-226.

[12] J.P. Dedieu, P. Priouret, G. Malajovich, Newton's method on Riemannian manifolds: covariant alpha theory, IMA J. Numer. Anal. 23 (2003), 395-419.

[13] O.P. Ferreira, B.F. Svaiter, Kantorovich's theorem on Newton's method in Riemannian manifolds, J. Complex. 18 (2002), 304-329.

[14] C. Li, J.H. Wang, Newton's Method for Sections on Riemannian Manifolds: Generalized Covariant α-Theory, J. Complexity, 24 (2008), 423-451.

[15] C. Li, J.H. Wang, Newton's method on Riemannian manifolds: Smale's point estimate theory under the γ-condition, IMA J. Numer Anal. 26 (2006), 228-251.

[16] J. Wang, C. Li, Uniqueness of the singular point of vector field on Riemannian manifold under the γ-condition, J. Complexity 22 (2006), 533-548.

[17] R.E. Mahony, U. Helmke, J.B. Moore,) Pole placement algorithms for symmetric realisations, Proceedings of 32nd IEEE Conference on Decision and Control, San Antonio, TX, pp. 1355-1358, 1993.

[18] J. Wang, C. Li, Kantorovich's theorems for Newton's method for mappings and optimization problems on Lie groups, IMA J. Numer Anal. 31 (2011), 322-347.

[19] C. Li, J. Wang, J.P. Dedieu, Newton's Method on Lie groups: Smale's point estimate theory under the γ-condition, J. Complexity, 25 (2009), 128-151.

[20] S.T. Ali, J.P. Antoine, J.P. Gazeau, Coherent States, Wavelets, and their Generalizations, Springer, New York, 2014. doi:10.1007/978-1-4614-8535-3

[21] V. Guillemin, S. Sternberg, Symplectic Techniques in Physics, Cambridge University Press, Cambridge, 1991.

[22] K.H. Grochenig, Foundations of Time-Frequency Analysis, Birkhauser, Boston, 2000.

[23] B. Dali, C. Li, J. Wang, Local convergence of Newton's method on the Heisenberg group, J. Comput. Appl. Math. 300 (2016), 217-232.

[24] V.S. Varadarajan, Lie Groups, Lie Algebras and Their Representations, GTM no.102, Springer-Verlag, New York, 1984.

[25] Capogna, L., Danielli, D., Pauls, S. D., Tyson, J. T., (2007) An introduction to the Heisenberg Group and the Sub-Riemannian Isoperimetric Problem, Birkhaser Verlag AG.

[26] V.S. Varadarajin, Lie Groups, Lie Algebras and Their Representations, Graduate Texts in Mathematics, Springer, New York, 1984.

[27] J. He, J. Wang, J.C. Yao, Convergence criteria of Newton's Method on Lie groups, Fixed point Theory Appl. 2013 (2013), 293.

[28] L. Blum, F. Cucker, M. Shub, S. Smale, Complexity and Real Computation, Springer, New York, 1997.

[29] S. Smale, Complexity theory and numerical analysis, Acta Numer. 6 (1997), 523-551.

[30] X. Wang, X. Xuan, Random polynomial space and computational complexity theory, Sci. China A 30 (1987), 673-684.

[31] P. Chen, Approximate zeros of quadratically convergent algorithms, Math. Comput. 63 (1994), 247-270.