# The Newton Iteration on Lie Groups *

Brynjulf Owren† and Bruno Welfert‡

†Department of Mathematical Sciences, NTNU, N-7034 Trondheim, Norway

‡Department of Mathematics, Arizona State University, Tempe, Arizona 85287-1804, U.S.A.

November, 1996§

submitted to BIT

**Abstract**

We define the Newton iteration for solving the equation $f(y) = 0$, where $f$ is a map from a Lie group to its corresponding Lie algebra. Two versions are presented, which are formulated independently of any metric on the Lie group. Both formulations reduce to the standard method in the Euclidean case, and are related to existing algorithms on certain Riemannian manifolds. In particular, we show that, under classical assumptions on $f$, the proposed method converges quadratically. We illustrate the techniques by solving a fixed-point problem arising from the numerical integration of a Lie-type initial value problem via implicit Euler.

*This work was in part sponsored by The Norwegian Research Council under contract no. 111038/410, through the SYNODE project. WWW: http://www.math.ntnu.no/num/synode

†Email: Brynjulf.Owren@math.ntnu.no, WWW: http://www.math.ntnu.no/~bryn

‡Email: bdw@math.la.asu.edu, WWW: http://math.la.asu.edu/~bdw

§Rev. September 28, 1999

# 1 Motivation

Recently, there has been an increased interest in studying numerical algorithms on manifolds. Traditionally, most numerical algorithms we know from the literature are formulated in vector spaces, and the convergence properties are usually investigated by imposing the additional structure of a Banach space. This approach can be justified through the fact that all manifolds can, by a proper choice of coordinates, be interpreted locally as Euclidian $\mathbb{R}^N$. This may however lead to difficulties if the algorithm operates on different charts. Another frequently used approach is to embed the manifold globally into a higher-dimensional Euclidian space. In this case there is the difficulty that the approximation will typically drift off the manifold if no information about the manifold is encoded in the algorithm. Finally, to accommodate the use of object-oriented implementations of numerical algorithms, it is of importance to phrase the algorithms independently of any particular choice of coordinates. Edelmann et al. [5] introduce conjugate gradient methods for Stiefel and Grassman manifolds by using a Riemannian metric and by following the corresponding geodesics. A similar approach for solving eigenvalue problems as a constrained optimization problem via a Newton iteration has been considered in [12, 13]. The difference between [12] and this paper lies mainly in our focus on solving nonlinear equations on Lie groups arising from the use of implicit integration methods on such manifolds. The idea of using numerical schemes which share qualitative properties with the original problem (structure, reversibility) has also been shown to be critical in determining the properties of perturbed problems [7] [9].

The motivation for this paper comes from recent approaches for solving ordinary differential equations on manifolds, see Munthe-Kaas [14, 15], Crouch and Grossman [4], and Owren and Marthinsen [18]. So far, the efforts have been focusing on explicit methods, but we believe that there is also a need for implicit methods for manifolds, as was recently shown by Zanna et al. [24]. Throughout this paper, we shall use the implicit Euler method as a generic example of an implicit integration method in the Lie group setting. This method is chosen for simplicity rather than for its believed importance among the members of this new class of integration methods. However, the generalization to more advanced implicit schemes is straightforward. Consider first the implicit Euler method applied to the problem $y' = g(y)$, which has the form

$$y^{(k+1)} = y^{(k)} + hg(y^{(k+1)})$$

Now, if instead we consider a differential equation on a Lie matrix group $G$ with corresponding Lie algebra $\mathfrak{g}$, i.e.

$$y' = y \cdot g(y), \quad y \in G \subset \mathrm{GL}(n, \mathbb{R}), \;\; g : G \to \mathfrak{g} \subset \mathfrak{gl}(n, \mathbb{R}),$$

the Crouch-Grossman and Munthe-Kaas versions of the implicit Euler method would be

$$y^{(k+1)} = y^{(k)} \cdot \exp(hg(y^{(k+1)})). \tag{1}$$

For general $g$ it is not possible to solve this equation explicitly. Instead we try to compute a sequence $\left(y_n^{(k+1)}\right)_{n\geq 0}$ that converges to the solution of (1) in some sense. An obvious choice for such a sequence is the one generated by the fixed point iteration

$$y_{n+1}^{(k+1)} = y^{(k)} \cdot \exp(hg(y_n^{(k+1)})), \;\; n = 0, 1, \ldots \tag{2}$$

In the classical version of the implicit Euler method, a fixed point iteration of the above type is frequently not feasible, usually, some version of the Newton iteration is preferred. The generalization of Newton's method to (1) can be obtained by observing that for $y$ "near" $y^{(k)}$, we have the representation $y = y^{(k)} \cdot \exp(v(y))$, with $v(y) \in \mathfrak{g}$, thus we need to solve the equation

$$f(y) = 0, \tag{3}$$

where

$$\begin{array}{rccl} f: & G & \to & \mathfrak{g} \\ & y & \mapsto & v(y) - hg(y). \end{array}$$

In this paper we propose two algorithms for solving problems of the form $f(y) = 0$ where $f$ is a map from a Lie group to its corresponding Lie algebra. We note already here that both versions reduce to the classical Newton method in the case that the Lie group $G$ is the abelian group $\mathbb{R}^n$, i. e. the Euclidean case. However, for other Lie groups, the classical Newton iteration does not generally make sense, because it works by definition on a linear space, whereas Lie groups in general are *not* linear spaces. Thus, here we discuss generalized versions of the classical Newton iteration which make sense for general Lie groups.

We start by reviewing in Section 2 some of the main ideas of interest to us regarding Lie groups and Lie algebras and their relationship. Section 3 introduces two versions of the Newton algorithm, one defined directly on the Lie group and another based on the traditional vector space approach. In Section 4 we study the convergence of the two versions. Section 5 illustrates the numerical behavior of the proposed algorithms with two simple examples arising from an implicit Euler time discretization of a Lie-type ODE. Finally, we conclude with some important remarks on the applicability and possible generalizations to homogeneous manifolds.

# 2 Background and Notation

## 2.1 Lie groups and Lie algebras

A *Lie group* $(G, \cdot)$ is a group which also has the structure of a smooth manifold such that the group product and the inversion are smooth operations in the differentiable structure given on the manifold. The dimension of the Lie group is that of the underlying manifold, and we shall always assume that it is finite. The symbol $e$ designates the identity element of $G$.

We will here refer to the *Lie algebra* $\mathfrak{g}$ of a Lie group as its tangent space at the $e, TG|_e$ equipped with the Lie bracket $[-,-] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$, which is defined for any pair of elements $u, v \in \mathfrak{g}$ as

$$[u, v] = \left. \frac{\partial^2}{\partial t \partial s} \right|_{t=s=0} y(t) \cdot z(s) \cdot y(t)^{-1}$$

where $y(t)$ and $z(s)$ are two smooth curves in $G$ satisfying $g(0) = h(0) = e$ and $g'(0) = u,\ z'(0) = v$.

Any Lie algebra can be shown to be bilinear and skew-symmetric. It also satisfies the Jacobi identity

$$[[u, v], w] + [[v, w], u] + [[w, u], v] = 0 \tag{4}$$

for all $u, v, w \in \mathfrak{g}$.

The Lie bracket defines a linear operator ad on $\mathfrak{g}$ such that

$$\mathrm{ad}_u(v) = [u, v]$$

for all $u, v \in \mathfrak{g}$.

In the sequel we will make use of the natural *left trivialization* of the Lie group $G$. We define for any $y \in G$

$$\begin{array}{rccc} L_y: & G & \to & G \\ & z & \mapsto & y \cdot z, \end{array}$$

the *left multiplication* in the group. Its tangent map $L'_y$ determines an isomorphism of $\mathfrak{g} = TG|_e$ with the tangent space $TG|_y$ via the relation

$$L'_y(\mathfrak{g}) = TG|_y \tag{5}$$

or, equivalently,

$$\mathfrak{g} = \left(L'_y\right)^{-1}(TG|_y) = L'_{y^{-1}}(TG|_y) \tag{6}$$

(see Figure 1).

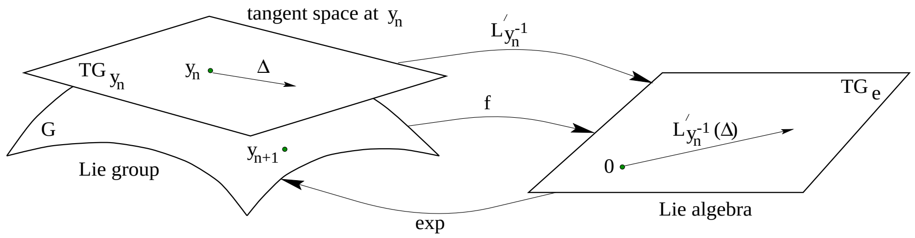


Figure 1: Lie group, tangent space and Lie algebra.

## 2.2 Some examples

Varadarajan [20] and Abraham et al. [1, supplementary chapter 9] give a list of a few of the most important Lie groups. We include two of them here: the first example is the standard Euclidean case, which provides a basis for comparing the proposed Newton algorithm described in Section 3.3 with its classical counterpart. The second example will provide the setting for solving a fixed-point problem arising from a Lie-type differential equation in Section 5.

### 2.2.1 Euclidean space $\mathbb{R}^n$

The Euclidean space $\mathbb{R}^n$ is a particular example of an abelian Lie group. In this case we have

$$\begin{aligned} (G, \cdot) &= (\mathbb{R}^n, +) \\ (\mathfrak{g}, [-,-]) &= (\mathbb{R}^n, [u,v] = 0) \\ L_y(z) &= y + z \\ L'_y(u) &= u \end{aligned}$$

### 2.2.2 Matrix groups

The set $\mathrm{GL}(n,\mathbb{R})$ of all invertible $n \times n$ matrices with real coefficients forms a Lie group called *the general linear group* under standard matrix multiplication, and it has several interesting subgroups. One is *the special linear group* $SL(n,\mathbb{R})$. consisting of matrices with unit determinant. Another is *the orthogonal group* $\mathrm{O}(n,\mathbb{R})$ of $\mathrm{GL}(n,\mathbb{R})$ consisting of orthogonal matrices $y$ (such that $y^T y = I_n$, where $I_n$ denotes the $n \times n$ identity matrix). The latter group has two connected components, the rotations $\mathrm{SO}(n,\mathbb{R}) = \{y \in \mathrm{O}(n,\mathbb{R}) : \det(y) = 1\}$ is again subgroup of $\mathrm{O}(n,\mathbb{R})$, (as well as of $SL(n,\mathbb{R})$) and it is frequently denoted the *special orthogonal group*. The Lie algebra of $GL(n,\mathbb{R})$ is usually denoted $\mathfrak{gl}(n,\mathbb{R})$, it is the linear space of all real $n \times n$ matrices with bracket given as the amtrix commutator $[u,v] = uv - vu$. The Lie algebras of the various subgroups of $\mathrm{GL}(n,\mathbb{R})$ are subalgebras of $\mathfrak{gl}(n,\mathbb{R})$. For instance, the Lie algebra of $SL(n,\mathbb{R})$ is the linear space of matrices whose trace is zero. The Lie algebra of both $\mathrm{O}(n,\mathbb{R})$ and $\mathrm{SO}(n,\mathbb{R})$ is the set of all $n \times n$ skew-symmetric matrices

$$\mathfrak{so}(n,\mathbb{R}) = \{v \in \mathfrak{gl}(n,\mathbb{R}), v^T + v = 0\}$$

The special orthogonal group $\mathrm{SO}(3,\mathbb{R})$ is the configuration space of a rigid body, and it plays an important role in computational mechanics (e.g., numerical integration of orthogonal flows, see

[2]). The importance of matrix Lie groups comes also from the fact that any finite dimensional Lie algebra over $\mathbb{R}$ is isomorphic to the Lie algebra associated to a subgroup of $\mathrm{GL}(n,\mathbb{R})$, according to Ado's Theorem [20, p.237].

### 2.3 The exponential map and its differential

The exponential map is a map

$$\begin{array}{rccc} \exp: & \mathfrak{g} & \to & G \\ & u & \mapsto & \exp(u) \end{array}$$

which is certainly the single most important construct associated to $G$ and $\mathfrak{g}$. Given $u \in \mathfrak{g}$, the left invariant vector field $X_u : y \mapsto L'_y(u)$ determines an integral curve $\gamma_u(t)$ solution of the initial value problem

$$\begin{cases} \gamma'_u(t) = X_u(\gamma_u(t)) \\ \gamma_u(0) = e. \end{cases}$$

The exponential map is then defined by the relation

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

In the Euclidean case this map reduces to the identity on $\mathbb{R}^n$, while it corresponds to the classical exponentiation of a matrix

$$\exp(u) = \sum_{k\geq 0} \frac{u^k}{k!} \tag{7}$$

if $u$ belongs to $\mathfrak{gl}(n,\mathbb{R})$. Note that in both cases exp maps the origin 0 of $\mathfrak{g}$ to the identity element $e$ of $G$. This is always the case. In fact, the exponential mapping can be shown [20, p.86] to be an analytic diffeomorphism on an open neighborhood $\mathcal{N}(0)$ of $0 \in \mathfrak{g}$. If

$$N(e) = \exp(\mathcal{N}(0)) \tag{8}$$

denotes the image of $\mathcal{N}(0)$ by exp, this means in particular that any $y \in N(e)$ can be written in the form

$$y = \exp(v)$$

for some $v \in \mathcal{N}(0)$, and that if

$$\exp(u) = \exp(v) \in N(e)$$

for some $u, v \in \mathcal{N}(0)$, then $u = v$.

If $G$ is abelian, exp is also a homomorphism from $\mathfrak{g}$ to $G$, i.e.,

$$\exp(u+v) = \exp(u)\cdot\exp(v) = \exp(v)\cdot\exp(u) \tag{9}$$

for all $u, v \in \mathfrak{g}$. In the non-abelian case (e.g., $G = \mathrm{O}(n,\mathbb{R})$), exp is not a homomorphism and (9) must be replaced by

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

where $w$ is given by the Baker-Campbell-Hausdorff (BCH) formula [20, pp.114] [22]

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

valid[1] for $u, v \in \mathcal{N}(0)$. A nice review of the matrix exponential and its properties can be found in [17].

[1] $\mathcal{N}(0)$ may be taken "small" enough such that the series (11) converges if $\mathfrak{g}$ is also a Banach space, see Section 4.1 and [20, (2.15.21)].

The differential of exp at $u \in \mathfrak{g}$ is a map

$$\begin{array}{rcl} \exp'_u : \quad \mathfrak{g} & \rightarrow & TG|_{\exp(u)} \\ v & \mapsto & \left.\dfrac{d}{dt}\right|_{t=0} \exp(u+tv) \end{array}$$

which is usually expressed by a map $\mathrm{dexp}_u : \mathfrak{g} \rightarrow \mathfrak{g}$ using the identification (6) of $TG|_{\exp(v)}$ with $\mathfrak{g}$, i.e.,

$$\mathrm{dexp}_u = L'_{(\exp(u))^{-1}} \circ \exp'_u = L'_{\exp(-u)} \circ \exp'_u \,. \tag{12}$$

For $v \in \mathfrak{g}$, $\mathrm{dexp}_u(v)$ can be obtained using the sum form [20, p.108] [10]

$$\mathrm{dexp}_u(v) = \sum_{n\geq 0} \frac{(-1)^n}{(n+1)!} \mathrm{ad}_u^n(v) \tag{13}$$

$$= \sum_{n\geq 0} \frac{(-1)^n}{(n+1)!} \underbrace{[u,[u,\ldots[u}_{n \text{ times}},v]\ldots]]$$

$$= v - \frac{1}{2}[u,v] + \frac{1}{6}[u,[u,v]] - \ldots \tag{14}$$

Note that if $u$ and $v$ commute ($[u,v]=0$) then $\mathrm{dexp}_u(v) = v$, so that $\mathrm{dexp}_0$ reduces to the identity map $\mathrm{Id}_{\mathfrak{g}}$ on $\mathfrak{g}$.

# 3 Two versions of Newton iteration

In this section we present a version of Newton's method for solving (3), defined directly on the Lie group $G$, and a variation, in which the standard technique is applied to an equation posed on $g$ obtained by transforming (3) via the exponential map. We first introduce the differential of $f$. The two versions of Newton's method are also compared in the last part of the section, where it is shown that the second version is merely a linearization (in a sense which is made precise below) of the first one.

## 3.1 The differential of the map $f$

The differential of $f$ at a point $y \in G$ is a map $f'_y : TG|_y \rightarrow T\mathfrak{g}|_{f(y)} \cong \mathfrak{g}$ defined as

$$f'_y(\Delta_y) = \left.\frac{d}{dt}\right|_{t=0} f\left(y \cdot \exp\left(tL'_{y^{-1}}(\Delta_y)\right)\right) \tag{15}$$

for any tangent vector $\Delta_y \in TG|_y$ to the manifold $G$ at $y$. The image by $f'_y$ of a tangent vector $\Delta_y$ is obtained by first identifying $\Delta_y$ with an element $v \in \mathfrak{g}$ via left multiplication. The exponential mapping transforms the scaled (by $t$) $v$ to an element $z$ of the Lie group (see Figure 1). Then $f'_y(\Delta_y)$ is obtained as

$$f'_y(\Delta_y) = \lim_{t\rightarrow 0} \frac{f(y\cdot z) - f(y)}{t}. \tag{16}$$

Analogously to (12), the differential $f'_y$ can be expressed via a function $\mathrm{d}f_y : \mathfrak{g} \rightarrow \mathfrak{g}$ given by

$$\mathrm{d}f_y = (f \circ L_y)' = f'_y \circ L'_y. \tag{17}$$

Thus

$$\mathrm{d}f_y(u) = f'_y\left(L'_y u\right) = \left.\frac{d}{dt}\right|_{t=0} f(y \cdot \exp(tu)) \tag{18}$$

(compare (18) with [21, (2) p.107], with $\varphi = f \circ L_y$).

## 3.2 The Newton iteration

Using the formula (15), Newton's method on the Lie group $G$ may proceed as follows: given $y_0 \in G$, we first determine the differential $f'_{y_0}$ according to (15). Then find $\Delta_{y_0} \in TG|_{y_0}$ satisfying the equation

$$f'_{y_0}(\Delta_{y_0}) + f(y_0) = 0,$$

and finally update $y_0$ by $y_1 = y_0 \cdot \exp\left(L'_{y_0^{-1}}(\Delta_{y_0})\right)$. In view of (17), the algorithm can be written as

**Version 1**: NEWTON'S METHOD ON THE LIE GROUP $G$

- Given $y_n \in G$, determine $\mathrm{d}f_{y_n}$ according to (18).
- Find $u_n \in \mathfrak{g}$ such that
$$\mathrm{d}f_{y_n}(u_n) + f(y_n) = 0. \tag{19}$$
- Compute $y_{n+1} = y_n \cdot \exp(u_n)$.

Since $y_{n+1} = y_n \cdot \exp\left(-\mathrm{d}f_{y_n}^{-1} \circ f(y_n)\right) = L_{y_n} \circ \exp \circ\, \mathrm{d}f_{y_n}^{-1} \circ (-f)(y_n)$, the method is a fixed-point iteration on $G$, with iteration function

$$\varphi(y) = L_y \circ \exp \circ\, \mathrm{d}f_y^{-1} \circ (-f)(y).$$

**Remark 3.1** *If the map $f$ is an homomorphism from $G$ to $\mathfrak{g}$ (i.e., if $f(y \cdot z) = f(y) + f(z)$ for all $y, z \in G$), then*

$$\begin{aligned} \mathrm{d}f_y(u) &= \lim_{t\to 0} \frac{f(y\cdot\exp(tu)) - f(y)}{t} = \lim_{t\to 0}\frac{f(\exp(tu)))}{t} \quad (= \mathrm{d}f_e(u)) \\ &= \lim_{n\to\infty} nf(\exp(\frac{1}{n}u)) = \lim_{n\to\infty} f(\exp(u)) = f(\exp(u)), \end{aligned}$$

*so that*

$$f(y_1) = f(y_0 \cdot \exp(u_0)) = f(y_0) + \mathrm{d}f_{y_0}(u_0) = 0$$

*and this version of the algorithm converges in one iteration independently of the initial choice $y_0$.*

Practical implementations of the above algorithm requires the determination of the differential $\mathrm{d}f_y$ for a given $y$, as well as the solution to the equation (19).

**Example 3.1** Given $z \in G$ and $h > 0$, the equation

$$y = z \cdot \exp(hg(y))$$

is equivalent, for $y$ such that $z^{-1}y \in N(e)$, to the equation

$$f(y) \stackrel{\text{def}}{=} v(y) - hg(y) = 0, \tag{20}$$

with

$$L_{z^{-1}}(y) = z^{-1}y = \exp(v(y)). \tag{21}$$

The quantity $f(y_n)$ is thus evaluated as

$$f(y_n) = \exp^{-1}(z^{-1} \cdot y_n) - hg(y_n).$$

The differential $\mathrm{d}f_{y_n}$ is obtained by differentiating $v$ and $g$:

$$\mathrm{d}f_{y_n} = \mathrm{d}v_{y_n} - h\mathrm{d}g_{y_n}.$$

From the relation (21) we obtain

$$L'_{z^{-1}} = \exp'_{v(y)} \circ v',$$

so that

$$\mathrm{dexp}_{v(y_n)}(\mathrm{d}v_{y_n}) = L'_{\exp(-v(y_n))} \circ \exp'_{v(y_n)} \circ v' \circ L'_{y_n} = L'_{\exp(-v(y_n))\cdot z^{-1}\cdot y_n} = L'_e = \mathrm{Id}_{\mathfrak{g}}.$$

Consequently, equation (19) becomes[2]

$$\mathrm{dexp}^{-1}_{v(y_n)}(u_n) - h\mathrm{d}g_{y_n}(u_n) + f(y_n) = 0. \tag{22}$$

Because of the identity

$$\mathrm{dexp}^{-1}_v = \mathrm{Id}_{\mathfrak{g}} + \frac{1}{2}\mathrm{ad}_v + \sum_{p\geq 1} \frac{B_{2p}}{(2p)!}\,\mathrm{ad}_v^{2p} \tag{23}$$

found for instance in [20], where $B_{2p}$ are the Bernoulli numbers, (22) is a problem of Sylvester type. It can be solved for $u_n$ via the solution of a linear system, provided a coordinate system for $\mathfrak{g}$ is given.

From a practical point of view, the quantity $\mathrm{dexp}^{-1}_{v(y_n)}(u)$ can be evaluated by:

- truncating the series (23) to a few terms, since one expects from (20) that $v(y_n)$ be "small" compared to 1 for small $h$, so that $\|\mathrm{ad}_v\| \ll \|\mathrm{Id}_{\mathfrak{g}}\|$. For example,

  $$\mathrm{dexp}^{-1}_{v(y_n)}(u) \simeq u \tag{24}$$

  (as in the commutative case) or

  $$\mathrm{dexp}^{-1}_{v(y_n)}(u) \simeq u + \frac{1}{2}[v(y_n), u]. \tag{25}$$

- using a finite difference approximation

  $$\begin{aligned} \mathrm{dexp}^{-1}_{v(y_n)}(u) &= \mathrm{d}\left(\exp^{-1}\right)_{z^{-1}\cdot y_n}(u) \\ &= \left.\frac{d}{dt}\right|_{t=0} \exp^{-1}(z^{-1}\cdot y_n \cdot \exp(tu)) \\ &\simeq \frac{\exp^{-1}(z^{-1}\cdot y_n\cdot \exp(tu)) - \exp^{-1}(z^{-1}\cdot y_n)}{t} \\ &= \frac{\exp^{-1}(z^{-1}\cdot y_n \cdot \exp(tu)) - v(y_n)}{t} \end{aligned} \tag{26}$$

  for some small value $t > 0$ (e.g., $t = \mathcal{O}(\sqrt{\mathbf{u}})$, where $\mathbf{u}$ is the round-off unit of the machine).

Alternately, the change of variable $u = \mathrm{dexp}_{v(y_n)}(w)$ (resp. $u_n = \mathrm{dexp}_{v(y_n)}(w_n)$) transforms (22) into

$$w_n - h\mathrm{d}g_{y_n} \circ \mathrm{dexp}_{v(y_n)}(w_n) + f(y_n) = 0. \tag{27}$$

A simple calculation shows that

$$\mathrm{d}g_{y_n} \circ \mathrm{dexp}_{v(y_n)} = \mathrm{d}(g \circ L_z \circ \exp)_{v(y_n)}, \tag{28}$$

so that $u_n$ in (22) may also be obtained by solving (28) for $w_n$ using an approximation of $\mathrm{d}g_{y_n} \circ \mathrm{dexp}_{v(y_n)}(w_n)$ similar to (26), i.e.,

$$\mathrm{d}g_{y_n} \circ \mathrm{dexp}_{v(y_n)}(w_n) \simeq \frac{g(z\cdot \exp(v_n + tw_n)) - g(y_n)}{t}, \tag{29}$$

followed by an approximation of $\mathrm{dexp}_{v(y_n)}(w_n)$ similar to (24) or (25),

$$u_n \simeq w_n, \tag{30}$$

or

$$u_n \simeq w_n - \frac{1}{2}[v(y_n), w_n]. \tag{31}$$

[2] The notation $\mathrm{dexp}^{-1}_{v(y_n)}$ is understood as $\left(\mathrm{dexp}_{v(y_n)}\right)^{-1}$. Since $z^{-1}\cdot y_n = \exp(v(n))$, it is also true that

$$\mathrm{dexp}^{-1}_{v(y_n)} = \mathrm{d}\left(\exp^{-1}\right)_{z^{-1}\cdot y_n}.$$

Note however that approximations like (24), (25), (30) or (31) may not be computationally satisfying in cases where the initial guess $y_0$ is "too far" from the solution $y$ (for example when solving (20) with a large value of $h$ and $y_0 = z$), due to the non-linearity of the exponential map (see Remark 4.1). In non-abelian groups, the approximation (26) should be preferred.

For more general functions $f$, the differential $\mathrm{d}f_{y_n}$ can be evaluated at $u$ using (18) in a form similar to (26), namely,

$$\mathrm{d}f_{y_n}(u) \simeq \frac{f(y_n \cdot \exp(tu)) - f(y_n)}{t} \tag{32}$$

for some $t = \mathcal{O}(\sqrt{\mathbf{u}})$.

Without loss of generality, suppose now that there exists a *known* $z \in G$ and some $v_n \in \mathfrak{g}$ such that

$$y_n = z \cdot \exp(v_n) = L_z \circ \exp(v_n).$$

For example, $z$ may be an approximation to the solution $y$ to $f(y) = 0$, such that $z^{-1}y_n \in N(e)$. Then

$$f(y_n) = f \circ L_z \circ \exp(v_n) \stackrel{\text{def}}{=} \tilde{f}(v_n).$$

The function $\tilde{f} = f \circ L_z \circ \exp$ is a map from $\mathfrak{g}$ to $\mathfrak{g}$. In particular, the differential $\mathrm{d}\tilde{f}_{u_n}$ of $\tilde{f}$ at $v_n$ is a linear map defined by

$$\mathrm{d}\tilde{f}_{v_n}(u) = \left.\frac{d}{dt}\right|_{t=0} \tilde{f}(v_n + tu) \tag{33}$$

If $\mathfrak{g}$ is finite dimensional, then the standard Newton procedure can be applied to the problem $\tilde{f}(v) = 0$. This leads to the following algorithm:

**Version 2**: LIE ALGEBRA-BASED NEWTON'S METHOD

- Given $y_n \in G$ such that $y_n = z \cdot \exp(v_n)$, determine $\mathrm{d}\tilde{f}_{v_n}$ according to (33).
- Find $u_n \in \mathfrak{g}$ such that
$$\mathrm{d}\tilde{f}_{v_n}(u_n) + \tilde{f}(v_n) = 0. \tag{34}$$
- Compute $v_{n+1} = v_n + u_n$ and $y_{n+1} = z \cdot \exp(v_{n+1})$.

In practice, one may start Version 2 with an initial guess $v_0$ in $\mathfrak{g}$ rather than a $y_0$ in $G$ and avoid computing the updates $y_n$ until convergence of $v_n$ (in $\mathfrak{g}$). If the solution $y$ to $f(y) = 0$ is expected to be "close" to $z$, then $v_0 = 0$ is appropriate as a starting guess.

**Remark 3.2** *If the map $\tilde{f}$ is linear, then*

$$f(z) = \tilde{f}(0) = \tilde{f}(0 + 0) = f(z) + f(z),$$

*so that $f(z) = 0$, so that $z$ is already a solution of $f(y) = 0$. Note also that $\tilde{f}$ linear is not equivalent to $f$ being an homomorphism from $G$ to $\mathfrak{g}$ (see Remark (3.1)), since*

$$\tilde{f}(u) + \tilde{f}(v) = f(z \cdot \exp(u)) + f(z \cdot \exp(v)) = f(z \cdot \exp(u) \cdot \exp(v)) \neq f(z \cdot \exp(u + v)) = \tilde{f}(u + v)$$

*in general if $f(z) = 0$ and $f$ is an homomorphism.*

**Example 3.2** Using (21) the equation (20) in Example 3.1 can be written

$$\tilde{f}(v) \stackrel{\text{def}}{=} f(z \cdot \exp(v)) = v - hg(z \cdot \exp(v)). \tag{35}$$

The differential of $\tilde{f}$ at $v_n$ is

$$\begin{aligned} \mathrm{d}\tilde{f}_{v_n} &= \mathrm{Id}_{\mathfrak{g}} - h\mathrm{d}(g \circ L_z \circ \exp)_{v_n} \\ &= \mathrm{Id}_{\mathfrak{g}} - h\mathrm{d}g_{y_n} \circ L'_{y_n^{-1}} \circ L'_z \circ \exp'_{v_n} \\ &= \mathrm{Id}_{\mathfrak{g}} - h\mathrm{d}g_{y_n} \circ \mathrm{dexp}_{v_n}, \end{aligned}$$

so that (34) becomes

$$u_n - h\mathrm{d}g_{y_n} \circ \mathrm{dexp}_{v_n}(u_n) + \tilde{f}(v_n) = 0. \tag{36}$$

Since

$$\tilde{f}(v_n) = f(z \cdot \exp(v_n)) = f(y_n) \tag{37}$$

if $v_n = v(y_n)$ with $y_n$ defined in (21) ($y = y_n$), $u_n$ in (36) is equal to $w_n$ in (27). In particular, $\mathrm{d}g_{y_n} \circ \mathrm{dexp}_{v_n}(u_n)$ nay be evaluated as in (29).

**Remark 3.3** *More generally, it is easy to see that*

$$\mathrm{d}\tilde{f}_{v_n}(u_n) = \mathrm{d}f_{y_n} \circ \mathrm{dexp}_{v_n}(u_n) \tag{38}$$

*holds.*

## 3.3 Comparison of the two versions

In view of (27) and (36), the two algorithms are related. Indeed, we show here that the second version is a "linearized" version of the first algorithm. A similar result was included also for the problem discussed in [13]. To understand why this is true, we use the following lemma:

**Lemma 3.1** *Let $u, v \in \mathfrak{g}$, and consider $w = w(v)$ given by the BCH formula (11). Then*

$$w = u + \mathrm{dexp}_u^{-1}(v) + \mathcal{O}(v^2),$$

*where $\mathcal{O}(v^2)$ is a smooth $\mathfrak{g}$-valued function of $v$ such that $\mathcal{O}(v^2)/\|v^2\|$ is bounded at $v = 0$.*

*Proof:* Differentiating the relation

$$\exp(u) \cdot \exp(v) = \exp(w)$$

defining $w$ in terms of $u$ and $v$, with respect to $v$ yields

$$L'_{\exp(u)} \circ \exp'_v = \exp'_{w(v)} \circ w'_v$$

so that

$$w'_v = \mathrm{dexp}^{-1}_{w(v)} \circ L'_{\exp(-w(v))} \circ L'_{\exp(u)} \circ L'_{\exp(v)} \circ \mathrm{dexp}_v = \mathrm{dexp}^{-1}_{w(v)} \circ \mathrm{dexp}_v.$$

Taking $v = 0$ and noting that $w(0) = u$ (recall that $\mathrm{dexp}_0 = \mathrm{Id}_{\mathfrak{g}}$) leads to the result. **Q.E.D.**

Using Lemma 3.1 one obtains

**Theorem 3.1** *Version 2 of the Newton algorithm is a linearization of Version 1, in the sense that if $y_n^{(1)} = z \cdot \exp(v_n^{(2)})$ then*

$$y_{n+1}^{(1)} = z \cdot \exp\left(v_{n+1}^{(1)}\right)$$

*with*

$$v_{n+1}^{(1)} = v_{n+1}^{(2)} + \mathcal{O}\left(\left(u_n^{(1)}\right)^2\right),$$

*where $\mathcal{O}(\cdot)$ is as in Lemma 3.1. Here the superscripts (1)and (2) refer to the version number in which the quantity is computed.*

*Proof:* Suppose that $y_n^{(1)} = z \cdot \exp(v_n^{(2)})$. From (37) and (38) we then have $u_n^{(1)} = \mathrm{dexp}_{v_n^{(2)}}\left(u_n^{(2)}\right)$, where $u_n^{(1)}$ is defined by (19) and $u_n^{(2)}$ is determined by from (34). Thus

$$v_{n+1}^{(2)} = v_n^{(2)} + \mathrm{dexp}_{v_n^{(2)}}^{-1}\left(u_n^{(1)}\right),$$

while

$$y_{n+1}^{(1)} = y_n^{(1)} \cdot \exp\left(u_n^{(1)}\right) = z \cdot \exp(v_n^{(2)}) \cdot \exp\left(u_n^{(1)}\right) = z \cdot \exp\left(v_{n+1}^{(1)}\right)$$

where $v_{n+1}^{(1)}$ is given by the BCH formula. The result follows from Lemma 3.1 with $w = v_{n+1}^{(1)} = v_{n+1}^{(1)}\left(u_n^{(1)}\right)$. **Q.E.D.**

A direct consequence of Theorem 3.1 is that both versions of the algorithm coïncide whenever the bracket $[u_n^{(1)}, v_n^{(2)}]$ vanishes for all $n$.

**Corollary 3.1** *If $G$ is abelian, then the versions 1 and 2 of Newton's algorithm are identical.*

It is easy to see that both algorithms reduce to the standard Newton's method in the Euclidean case, since in that case the exponential map reduces to the identity map on $\mathfrak{g} = G$. The quantity $f'(y_n)$ is then the usual Jacobian matrix of partial derivatives $\dfrac{\partial f}{\partial y_{i,j}}$ evaluated at $y_n$, and $\mathrm{d}f_{y_n}(u_n)$ is simply the directional derivative in the (tangent) direction $\Delta_{y_n} = L'_{y_n}(u_n) = u_n$.

We also note that the update $y_{n+1} = y_n \cdot \exp(u_n)$ in Version 1 corresponds to the evaluation at $t = 1$ of the integral curve $\gamma_{u_n}(t) = y_n \cdot \exp(tu_n)$ solution of

$$\begin{cases} \gamma'_{u_n}(t) = X_{u_n}(\gamma_{u_n}(t)) \\ \gamma_{u_n}(0) = y_n, \end{cases}$$

i.e., is determined by moving from $y_n$ in the direction of the tangent vector $X_{u_n}$ on $\gamma_{u_n}$. In some Riemannian cases, this curve can be shown to be a geodesic (curve of minimal length) on $G$ [5, eq.(34)].

# 4 Convergence analysis

## 4.1 Convergence in the Lie group

How "large" the neighborhood $\mathcal{N}(0)$ (resp. $N(e)$, see Section 2.3) is depends on the existence of a metric defined on the Lie algebra (identified with the vector space $TG|_e$). In $\mathbb{R}^n$, we have $\mathcal{N}(0) = \mathbb{R}^n = \mathfrak{g} = G = N(e)$ $(e = 0)$. However, Varadarajan [20, p.90] gives the example of the Lie group $G = \mathrm{GL}(2, \mathbb{R})$ and a $y \in G$ such that there is no $u \in \mathfrak{g} = \mathfrak{gl}(2, \mathbb{R})$ satisfying $y = \exp(u)$, i.e., $\exp(\mathfrak{g})$ is only a subset on $G$ in general.

Since a Lie group is a manifold, it is also a topological space, and hence, the notion of convergence is defined with respect to this topology. However, we shall find it useful to impose a Banach space structure on the Lie algebra $\mathfrak{g}$, and use this to analyse convergence. Thus, we assume that we have to our disposal a norm $\| - \|$ on $\mathfrak{g}$ which induces a complete metric. This requirement is more general than in [5], where a Newton iteration based on Hilbert spaces[3] is proposed.

[3]a Hilbert space is an inner product space with a complete metric [1]. In particular, a Hilbert space is also a Banach space.

Since we only deal with finite dimensional Lie algebras, every linear mapping $\varphi : \mathfrak{g} \to \mathfrak{g}$ is bounded and we can define its norm subordinate to that on $\mathfrak{g}$

$$\|\varphi\| = \sup_{u \neq 0} \frac{\|\varphi(u)\|}{\|u\|} = \sup_{\|u\|=1} \|\varphi(u)\| < \infty.$$

**Definition 4.1** *Let $N(e)$ be as in (8). A sequence $\{y_n\}_{n\geq 0}$ of points of $N(e)$ is said to converge to e if and only if the associated sequence $\{u_n\}_{n\geq 0}$ defined by*

$$y_n = \exp(u_n)$$

*for all $n \geq 0$ converges to 0, in the sense that*

$$\lim_{n\to\infty} \|u_n\| = 0. \tag{39}$$

**Definition 4.2** *If there exists $y \in G$ such that*

$$y^{-1} y_n = \exp(u_n) \in N(e) \tag{40}$$

*and (39) holds, then the sequence $\{y_n\}_{n\geq 0}$ is said to converge to $y$. Furthermore, the convergence is of order $p$ if $\|u_{n+1}\| \leq \mathcal{O}(\|u_n\|^p)$.*

As usual, linear convergence corresponds to $p = 1$, and quadratic convergence to $p = 2$.

**Remark 4.1** *In the non-abelian case, the map* $\mathrm{dexp} : u \mapsto \mathrm{dexp}_u$ *is only "almost" linear in the vicinity of $u = 0$, the origin of $\mathfrak{g}$, see (14). The non-linearity of the exponential map may then result in a deterioration or the absence of convergence if $y^{-1}y_0$ lies outside a neighborhood of $e$ contained in $N(e)$.*

The continuity of the Lie bracket $[-,-]$ guarantees the existence of a constant $\mu \geq 0$ such that

$$\|[u, v]\| \leq \mu \|u\| \, \|v\| \tag{41}$$

for all $u, v \in \mathfrak{g}$. This is equivalent to the boundedness of the operator $\mathrm{ad}_u$ on compact sets of $\mathfrak{g}$:

$$\|\mathrm{ad}_u\| \leq \mu \|u\|. \tag{42}$$

Because $\mathrm{dexp}_0 = \mathrm{Id}_{\mathfrak{g}}$, remark that for small enough $\|u\| \leq \beta$ the map $\mathrm{dexp}_u$ is well-defined and we have

$$\|\mathrm{dexp}_u\| \simeq \|\mathrm{dexp}_u^{-1}\| \simeq \Theta(\mu\beta) \tag{43}$$

for some generic function $\Theta(t)$ analytic around $t = 0$ satisfying $\Theta(0) = 1$. A more precise statement for (43) can be made using the functions

$$\Phi(t) = \frac{e^t - 1}{t} \tag{44}$$

and $h(t)$ (defined below), but this would only obscure the exposition.

**Lemma 4.1** *There exists a constant $\delta > 0$ and a real function $h = h(t)$, analytic in the disk $|t| \leq \delta$, non-decreasing for $t \geq 0$, with $h(0) = 1$, such that for all $u, v \in \mathfrak{g}$*

$$\|u\| + \|v\| \leq \beta \quad \Rightarrow \quad \|w\| \leq \|u + v\| \; h(\mu\beta) \tag{45}$$

*and*

$$\|u\| \leq \frac{\beta}{2}, \quad \|u\| + \|v\| \leq \beta \quad \Rightarrow \quad \|w - u\| \leq \|v\| \; h(\mu\beta) \tag{46}$$

*for any $\beta \leq \delta/\mu$, where $w$ is given by (11).*

Lemma 4.1 shows that (9) "almost" holds in terms of the norm $\|-\|$ for small $\|u\|$ and $\|v\|$. Note that in the abelian case, $\mu = 0$ and $h(\mu\beta) = 1$ for any $\beta \geq 0$, so that (45) and (46) are tight.

The proof of the lemma is based on a variation of the proof of [20, Theorem 2.15.4]. It is rather technical and is omitted here. The interested reader will find it as well as the proofs of the lemmas below in the Appendix of [19].

**Lemma 4.2**

$$\max(\|u\|, \|v\|) \leq \beta \quad \Rightarrow \quad \|\mathrm{dexp}_u - \mathrm{dexp}_v\| \leq \mu\|u - v\|\Phi'(\mu\beta). \tag{47}$$

*where $\Phi(t)$ is defined by (44).*

**Lemma 4.3** *For $y \in G$, $u \in \mathfrak{g}$ and $t \in \mathbb{R}$ we have*

$$\frac{d}{dt} f(y \cdot \exp(-tu)) = -\mathrm{d}f_{y \cdot \exp(-tu)}(u).$$

In the following we consider the radius

$$\rho = \max_{\|u\|=1} \{t \in \mathbb{R}, su \in \mathcal{N}(0), \forall 0 \leq s \leq t\} > 0 \tag{48}$$

of the maximal ball in $\mathfrak{g}$ centered at 0 and contained in $\mathcal{N}(0)$ (defined in Section 2.3). For $\sigma \leq \rho$ we then introduce the corresponding ball of radius $\sigma$ around $y \in G$ as

$$B_\sigma(y) = \{z \in G, y^{-1} \cdot z = \exp(u) \in N(e), \|u\| \leq \sigma\}.$$

For completeness we first restate the standard convergence result of Newton's method in Banach spaces which applies to the second version of the algorithm. We suppose that $z$ (see Version 2) is in $B_{\tilde{\sigma}}(y)$. Then $v$ denotes the solution of $\tilde{f}(v) = 0$, related to $y$ by $y = z \cdot \exp(v)$.

**Theorem 4.1** *Suppose that*

- *there exists a constant $\tilde{\tau}$ such that $0 < \tilde{\tau} \leq \rho$ and a constant $\tilde{\gamma} \geq 0$ such that*

$$\|\mathrm{d}\tilde{f}_u - \mathrm{d}\tilde{f}_w\| \leq \tilde{\gamma}\|u - w\| \tag{49}$$

  *for all $u, w \in \mathfrak{g}$ such that $\|u - v\|, \|w - v\| \leq \tilde{\tau}$ (local Lipschitz condition),*

- *the map $\mathrm{d}\tilde{f}_v$ is one-to-one and $\mathrm{d}\tilde{f}_v^{-1}$ is bounded with constant $\tilde{\eta}$, i.e., $\|\mathrm{d}\tilde{f}_v^{-1}\| \leq \tilde{\eta}$.*

*Then there exists a constant $\tilde{\sigma} > 0$ such that, if $\|v - v_0\| \leq \tilde{\sigma}$*

- *$\|v_n - v_0\| \leq \tilde{\sigma}$ for all $n \geq 0$ (the sequence $\{v_n\}_{n \geq 0}$ is well-defined),*
- *$\lim_{n \to \infty} v_n = v$ and $\|v - v_{n+1}\| \leq \tilde{\eta}\tilde{\gamma}\|v - v_n\|^2$ for all $n \geq 0$, i.e., the sequence $\{v_n\}_{n \geq 0}$ converges quadratically to $v$.*

Theorem 4.1 does not prove, as stated, that the sequence $(y_n)_{n \geq 0}$ obtained from Version 2 converges quadratically to $y = z \cdot \exp(v)$ in the sense of Definition 4.2, since

$$y^{-1} \cdot y_n = \exp(-v) \cdot z^{-1} \cdot z \cdot \exp(v_n) = \exp(-v) \cdot \exp(v_n) \neq \exp(-v + v_n)$$

in general. Nonetheless, it can serve as a basis to show the following result.

**Theorem 4.2 (convergence of Version 2)** *Suppose that*

- *there exists a constant* $\tau$ *such that* $0 < \tau \le \rho$ *and a constant* $\gamma \ge 0$ *such that*
$$\|\mathrm{d}f_{z\cdot\exp(u)} - \mathrm{d}f_z\| \le \gamma\|u\| \tag{50}$$
*for all* $z \in B_\tau(y)$ *and* $u \in \mathfrak{g}$ *such that* $\|u\| \le \tau$ *(local Lipschitz condition),*
- *the map* $\mathrm{d}f_y$ *is one-to-one and* $\mathrm{d}f_y^{-1}$ *is bounded with constant* $\eta$*, i.e.,* $\|\mathrm{d}f_y^{-1}\| \le \eta$.

*Then there exists a constant* $\sigma > 0$ *and two functions* $h_1(t)$ *and* $h_2(t)$*, analytic for* $|t| \le \delta/\mu$ *(where* $\delta$ *and* $\mu$ *are as in Lemma 4.1), with* $h_1(0) = 2h_2(0) = 1$*, such that, if* $y_0 \in B_\sigma(y)$

- $y_n \in B_\sigma(y)$ *for all* $n \ge 0$ *(the sequence* $\{y_n\}_{n\ge0}$ *is well-defined),*
- $y^{-1}y_n = \exp(v_n)$*,* $\lim_{n\to\infty} v_n = 0$ *and* $\|v_{n+1}\| \le \eta\left(\gamma h_1(\mu\sigma) + \|\mathrm{d}f_y\|\mu h_2(\mu\sigma)\right)\|v_n\|^2$ *for all* $n \ge 0$*, i.e., the sequence* $\{y_n\}_{n\ge0}$ *converges quadratically to* $y$.

*Proof:* We only give a sketch of the proof, and avoid details with the rigorous selection of an appropriate neighborhood of $y$ (e.g., $\sigma$) in which the result holds. The result is shown by relating the Lipschitz conditions (49) and (50) as well as $\tilde{\eta}$ and $\eta$ to one another, then applying Theorem 4.1, finally connecting $v_n$ in Theorem 4.2 (hereby denoted $w_n$ to avoid duplication of symbol) to $v - v_n$ in Theorem 4.1. For our purpose we shall suppose that all manipulations are valid by supposing that $\sigma$ is "small enough". This implies in particular that $z = y \cdot \exp(v)$ with $\|v\| \le \sigma$ small ($v$ is the solution to $\tilde{f}(v) = 0$).

• First we have, for $u$ and $w$ such that $\|u\|, \|w\| \le \sigma$
$$\begin{aligned}
\|\mathrm{d}\tilde{f}_u - \mathrm{d}\tilde{f}_w\| &\le \|\mathrm{d}f_{z\cdot\exp(u)} \circ \mathrm{dexp}_u - \mathrm{d}f_{z\cdot\exp(w)} \circ \mathrm{dexp}_w\| \\
&\le \|(\mathrm{d}f_{z\cdot\exp(u)} - \mathrm{d}f_{z\cdot\exp(w)}) \circ \mathrm{dexp}_u + \mathrm{d}f_{z\cdot\exp(w)} \circ (\mathrm{dexp}_u - \mathrm{dexp}_w)\| \\
&\le \gamma\|\hat{v}\|\,\|\mathrm{dexp}_u\| + \|\mathrm{d}f_{z\cdot\exp(w)}\|\,\|\mathrm{dexp}_u - \mathrm{dexp}_w\|
\end{aligned}$$
where $\hat{v}$ is such that $\exp(\hat{v}) = \exp(-u) \cdot \exp(w)$. Since $u$ and $w$ are small, $\hat{v}$ is also small, and Lemma 4.1, equation (45) yields $\|\hat{v}\| \le \|-u + w\|h(\mu\beta)$ for some small $\beta$. From (43), Lemma 4.2, and the fact that
$$\|\mathrm{d}f_{z\cdot\exp(w)}\| \le \|\mathrm{d}f_{z\cdot\exp(w)} - \mathrm{d}f_z\| + \|\mathrm{d}f_z - \mathrm{d}f_y\| + \|\mathrm{d}f_y\| \le \gamma(\|w\| + \|v\|) + \|\mathrm{d}f_y\|,$$
we obtain
$$\|\mathrm{d}\tilde{f}_u - \mathrm{d}\tilde{f}_w\| \le \left(\gamma\varphi_1(\mu\sigma) + \|\mathrm{d}f_y\|\mu\Phi'(\mu\sigma)\right)\|u - w\| \stackrel{\text{def}}{=} \tilde{\gamma}\|u - w\|$$
with $\varphi_1(t) = h(t)\Theta(t) + 2t\Phi'(t)$ and $\varphi_1(0) = 2\Phi'(0) = 1$.

• Next we have
$$\|\mathrm{d}\tilde{f}_v^{-1}\| = \|\mathrm{dexp}_v^{-1} \circ \mathrm{d}f_y^{-1}\| \le \|\mathrm{dexp}_v^{-1}\|\,\|\mathrm{d}f_y^{-1}\| \le \Theta(\mu\sigma)\eta \stackrel{\text{def}}{=} \tilde{\eta}.$$

• Now $y^{-1}y_n = \exp(-v) \cdot \exp(v_n) = \exp(w_n)$ with $w_n = -v + v_n + \frac{1}{2}[-v, v_n] + \ldots$. Using successively Lemma 4.1 (45), Theorem 4.1 and Lemma 4.1 (46) (with $w = v_n$ after rewriting $\exp(v_n) = \exp(v) \cdot \exp(w_n)$), it becomes
$$\begin{aligned}
\|w_{n+1}\| &\le \|v - v_{n+1}\|h(2\mu\sigma) \\
&\le \tilde{\eta}\tilde{\gamma}\|v - v_n\|^2 h(2\mu\sigma) \\
&\le \tilde{\eta}\tilde{\gamma}\left(\|w_n\|h(2\mu\sigma)\right)^2 h(2\mu\sigma) \;=\; \eta\left(\gamma h_1(\mu\sigma) + \|\mathrm{d}f_y\|\mu h_2(\mu\sigma)\right)\|v_n\|^2
\end{aligned}$$
with $h_1(t) = (h(2t))^3\Theta(t)\varphi_1(t)$ and $h_2(t) = (h(2t))^3\Theta(t)\Phi'(t)$, for all $n \ge 0$. The verification that $h_1(0) = 2h_2(0) = 1$ is trivial. **Q.E.D.**

**Theorem 4.3 (convergence of Version 1)** *Suppose that*

- *there exists a constant* $\tau$ *such that* $0 < \tau \le \rho$ *and a constant* $\gamma \ge 0$ *such that*
$$\|\mathrm{d}f_{z\cdot\exp(u)} - \mathrm{d}f_z\| \le \gamma\|u\| \tag{51}$$
*for all* $z \in B_\tau(y)$ *and* $u \in \mathfrak{g}$ *such that* $\|u\| \le \tau$ *(local Lipschitz condition),*
- *the map* $\mathrm{d}f_y$ *is one-to-one and* $\mathrm{d}f_y^{-1}$ *is bounded with constant* $\eta$*, i.e.,* $\|\mathrm{d}f_y^{-1}\| \le \eta$*.*

*Then there exists a constant* $\sigma > 0$ *such that, if* $y_0 \in B_\sigma(y)$

- $y_n \in B_\sigma(y)$ *for all* $n \ge 0$ *(the sequence* $\{y_n\}_{n\ge0}$ *is well-defined),*
- $y^{-1}y_n = \exp(v_n)$*,* $\lim_{n\to\infty} v_n = 0$ *and* $\|v_{n+1}\| \le \beta\gamma h(3\mu\sigma)\|v_n\|^2$ *for all* $n \ge 0$*, where* $h$ *is as in Lemma 4.1, i.e., the sequence* $\{y_n\}_{n\ge0}$ *converges quadratically to* $y$*.*

*Proof of Theorem 4.3:* Let $\delta > 0$ and $h(-)$ be as in Lemma 4.1, $M = \sup_{t\in[0,\delta]} h(t) = h(\delta) \ge 1$ and pick $\sigma = \min\left(\frac{1}{2\eta\gamma M}, \frac{\tau}{2}, \frac{\delta}{3\mu}\right)$, where $\mu$ is defined in (41).

Consider $y_0 \in B_\sigma(y)$. Then $y^{-1}\cdot y_0 = \exp(v_0)$ for some $v_0 \in \mathfrak{g}$ such that $\|v_0\| \le \sigma$. First note that
$$\|\mathrm{Id}_{\mathfrak{g}} - \mathrm{d}f_y^{-1}\circ \mathrm{d}f_{y_0}\| = \|\mathrm{d}f_y^{-1}\circ\left(\mathrm{d}f_y - \mathrm{d}f_{y_0}\right)\| \le \|\mathrm{d}f_y^{-1}\|\,\|\mathrm{d}f_y - \mathrm{d}f_{y_0}\| \le \eta\gamma\|v_0\| \le \frac{1}{2M} \le \frac{1}{2}.$$
Thus $\mathrm{d}f_{y_0}$ is bijective by virtue of the Banach lemma [6, p.59]. Furthermore,
$$\|\mathrm{d}f_{y_0}^{-1}\| \le \frac{\|\mathrm{d}f_y^{-1}\|}{1 - \|\mathrm{Id}_{\mathfrak{g}} - \mathrm{d}f_y^{-1}\circ\mathrm{d}f_{y_0}\|} \le 2\|\mathrm{d}f_y^{-1}\| \le 2\eta.$$
Then, if $u_0$ is defined by (19), we obtain, using Lemma 4.3,
$$\begin{aligned} u_0 + v_0 &= \mathrm{d}f_{y_0}^{-1}\left(f(y) - f(y_0) + \mathrm{d}f_{y_0}(v_0)\right) \\ &= \mathrm{d}f_{y_0}^{-1}\int_0^1 \left(\frac{d}{dt}f(y_0\cdot\exp(-tv_0)) + \mathrm{d}f_{y_0}(v_0)\right)dt \\ &= \mathrm{d}f_{y_0}^{-1}\int_0^1 \left(-\mathrm{d}f_{y_0\cdot\exp(-tv_0)}(v_0) + \mathrm{d}f_{y_0}(v_0)\right)dt \end{aligned}$$
so that (note that $\|tv_0\| \le \sigma$ for $t \in [0,1]$)
$$\begin{aligned} \|u_0 + v_0\| &\le \|\mathrm{d}f_{y_0}^{-1}\|\int_0^1 \|\mathrm{d}f_{y_0}(v_0) - \mathrm{d}f_{y_0\cdot\exp(-tv_0)}(v_0)\|\,dt \\ &\le 2\eta\int_0^1 \gamma\|tv_0\|\,\|v_0\|\,dt \\ &= \eta\gamma\|v_0\|^2. \end{aligned}$$
In particular,
$$\|u_0\| \le \|v_0\| + \|u_0 + v_0\| \le \left(1 + \frac{1}{2M}\right)\|v_0\| \le \frac{3}{2}\|v_0\| \le \frac{3}{2}\sigma \le \frac{3}{4}\tau < \rho,$$
so that $u_0 \in \mathcal{N}(0)$. Thus
$$y^{-1}y_1 = \exp(v_0)\cdot\exp(u_0) = \exp(v_1)$$

with (see (11))

$$v_1 = v_0 + u_0 + \frac{1}{2}[v_0, u_0] + \ldots \in \mathfrak{g}.$$

The inequality (45) in Lemma (4.1) then yields, with $\beta = 3\sigma \leq \delta/\mu$,

$$\|v_1\| \leq \|v_0 + u_0\| \; h(3\mu\sigma) \leq h(3\mu\sigma)\eta\gamma\|v_0\|^2.$$

In particular, $\|v_1\| \leq \eta\gamma M\|v_0\|^2 \leq \frac{1}{2}\|v_0\| \leq \sigma$, so that $y_1 \in B_\sigma(y)$ as well. The result follows by induction (the fact that $v_n \to 0$ as $n \to \infty$ is a consequence of the fact that $\|v_{n+1}\| \leq \frac{1}{2}\|v_n\|$ for all $n$). **Q.E.D.**

**Remark 4.2** *The Lipschitz condition (50)-(51) arises naturally in the proof of Theorem 4.3, and reduces to the standard Lipschitz condition of the type (49) in the Euclidean case (with $u \leftarrow z + u$ and $v \leftarrow z$).*

# 5 Numerical examples

In this section we consider the initial value problem

$$\begin{cases} y' = y \cdot g(y) \\ y(0) = y^{(0)} \end{cases} \tag{52}$$

for $y \in G = \mathrm{O}(N, \mathbb{R})$, where $g(y) \in \mathfrak{g} = \mathfrak{so}(N, \mathbb{R})$ and $y^{(0)}$ is a random starting point.

The application of one step of the backward Euler method on (52) leads to the fixed-point problem

$$y = y^{(0)} \cdot \exp(hg(y))$$

which was used in Examples 3.1 and 3.2 (with $z = y^{(0)}$), where $h$ represents the size of the discretization step.

The quantity $u_n$ determined by (19) was computed by solving a linear system of size $q = \dim(G) = \dim(\mathfrak{g}) = \frac{N(N-1)}{2}$ resulting from the expansion of $u_n = \sum_{i=1}^{q} \alpha_i \mathbf{b}_i$ onto a basis of $\mathfrak{g}$. This system reads

$$\mathbb{D}\alpha + \mathbb{F} = \mathbf{0},$$

where

$$\mathbb{D} = (\mathbb{D}_{i,j})_{1 \leq i,j \leq q}, \quad \alpha = (\alpha_1, \ldots, \alpha_q)^T, \quad \mathbb{F} = (f_1, \ldots, f_q)^T$$

with

$$\mathrm{d}f_{y_n}(\mathbf{b}_i) = \sum_{j=1}^{q} \mathbb{D}_{i,j} \mathbf{b}_j \quad \text{and} \quad f(y_n) = \exp^{-1}(y^{(0)-1} y_n) - hg(y_n) = \sum_{i=1}^{q} f_i \mathbf{b}_i.$$

A similar strategy was used to solve (34). The matrix exponential is carried out via the MATLAB function `expm.m`.

The two versions of the Newton iteration are compared to the fixed-point iteration (2) for a range of step-sizes $10^{-3} \simeq 2^{-10} \leq h \leq 2^{10} = 1024$ and dimensions $2 \leq N \leq 15$. In all cases, the algorithm is started with $y_0 = y^{(0)}$. In all cases we shall say that the iteration has converged when

$$\|f(y_n)\|_2 \leq tol = 10^{-13}$$

(note that convergence for Version 2 of the Newton algorithm would normally be measured by $\|\tilde{f}(v_n)\|_2 \leq tol$). An upper limit of 100 iterations is set for achieving the above convergence criterion.

The convergence rate for each method is determined by considering the ratio

$$r = \log\left(\frac{\|f(y_{p-1})\|_2}{\|f(y_{p-2})\|_2}\right) \Big/ \log\left(\frac{\|f(y_{p-2})\|_2}{\|f(y_{p-3})\|_2}\right)$$

where $p$ is the number of iterations necessary to reach convergence. This avoids wrong estimations of the rate of convergence due to the fact that we may have $\|f(y_p)\|_2 = \mathcal{O}(\mathbf{u})$, but is only possible if $p \geq 3$. The rate of convergence is not computed if $p < 3$.

- In the first example, $g$ and $y^{(0)}$ are given by (in MATLAB notations)

$$\mathtt{g(y)} = \mathtt{diag(diag(y,1),1)} - \mathtt{diag(diag(y,1),-1)}$$

and

$$\mathtt{y}^{(0)} = \mathtt{qr(rand(N,N))},$$

respectively. This test problem was used in [14] and [23] to investigate the numerical order of explicit integrators on Lie groups. Note that $g$ is linear in $y$, so that, using (7),

$$g(y \cdot \exp(tu)) = g\left(y \sum_{n \geq 0} \frac{t^n}{n!} u^n\right) = \sum_{n \geq 0} \frac{t^n}{n!} g(yu^n)$$

and

$$\mathrm{d}g_y(u) = \left.\frac{d}{dt}\right|_{t=0} g((y \cdot \exp(tu)) = g(yu),$$

yielding the explicit form

$$\mathrm{d}f_y(u) = \mathrm{dexp}^{-1}_{v(y)}(u) - h\mathrm{d}g_y(u) = u + \frac{1}{2}[\exp^{-1}(y^{(0)-1}y), u] + \ldots - hg(yu).$$

The numerical experiments were however conducted by using the finite difference approximation (32).

It is easy to check that the conditions of Theorems 4.2-4.3 are satisfied with $\gamma = 1 + \mathcal{O}(h)$ and $\eta = 1 + \mathcal{O}(h)$ (since $\mathrm{d}f_y(u) \simeq u - hg(yu)$).

Figure 2 shows that the Newton iteration (Version 1) converges significantly better for larger step-sizes than the Version 2 for smaller values of $N$ ($N = 3$ is relevant in practical problems in mechanics), while the fixed-point iteration does not converge. Note that in the case $N = 3$ the cost of both versions of the Newton algorithm is similar (one can use Rodrigues formula and its inverse to compute $\exp(u)$ and $\exp^{-1}(y)$). In practical use of the Newton iteration for solving the equations arising from implicit integration methods, fairly good initial values are available due to the continuity of the solution to the underlying ODE, thus, we have chosen not to consider the behaviour when the initial error is “large”.

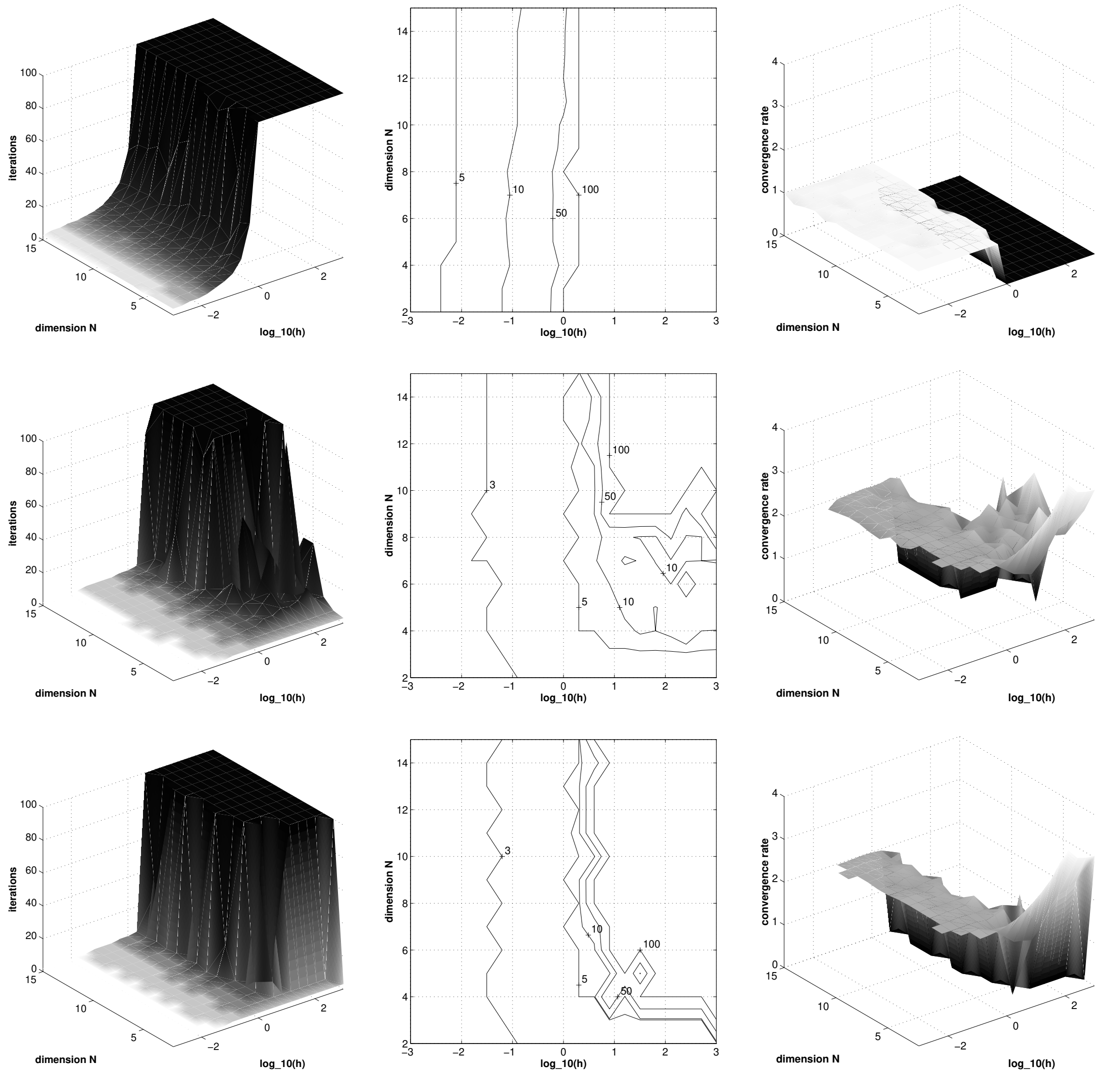


Figure 2: Number of iterations (left and center) and rate of convergence (right) to reach convergence to $tol = 10^{-13}$ for the fixed-point iteration (top), Version 1 (middle) and Version 2 (bottom) of the Newton algorithm applied to $g(y) = u - u^T$, with $\mathtt{u} = \mathrm{diag}(\mathrm{diag}(y, 1), 1)$, as a function of the step-size $h$ and the dimension $N$ of the Lie group $G$.

As expected, the Newton iterations converge quadratically in the range of $h$- and $N-$ values where the methods converge, compared to the linear convergence of the fixed-point iteration.

- The second example is similar to the first one, except that $g$ is now defined by

$$g(y) = u - u^T, \quad u = \sin(y)(2y - 5y^2).$$

The numerical results in Figure 3 show a behavior of the three algorithms similar to what occurred the first example. Again, the first version of the algorithm seems to converge better for smaller size problems.

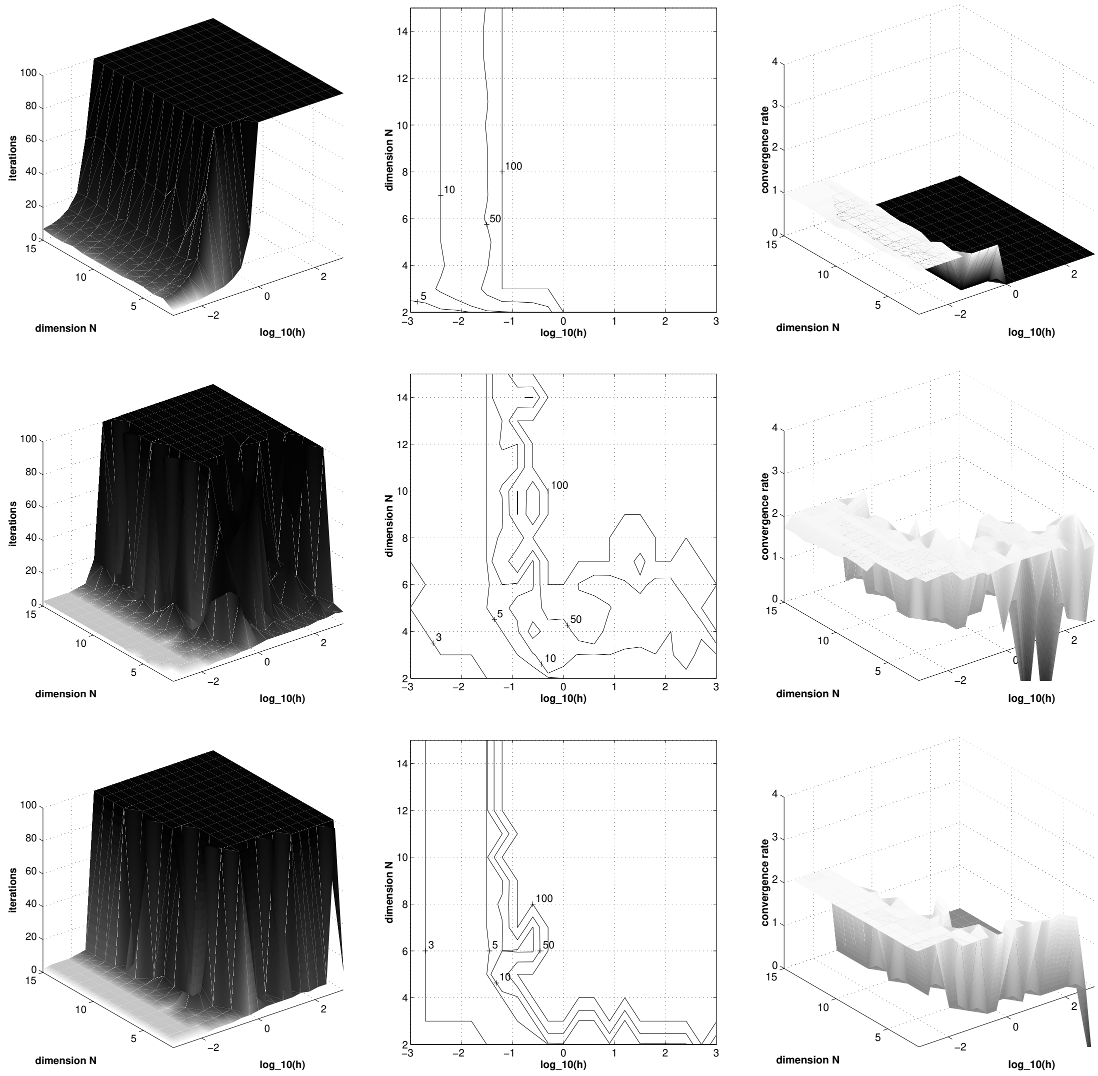


Figure 3: Number of iterations (left and center) and rate of convergence (right) to reach convergence to $tol = 10^{-13}$ for the fixed-point iteration (top), Version 1 (middle) and Version 2 (bottom) of the Newton algorithm applied to $g(y) = u - u^T$ with $u = \sin(y)(2y - 5y^2)$, as a function of the step-size $h$ and the dimension $N$ of the Lie group $G$.

From these experiments it appears that

- for small $N$, the Version 1 of Newton's method converges for a larger range of step-sizes compared to Version 2 and the fixed-point iteration.
- for larger $N$, the Newton iteration may not be competitive compared to the fixed-point iteration, but is guaranteed to converge for smaller step-sizes, while the fixed-point method is not.

For MATLAB files implementing the experiments of this section, see [19].

## 6 Concluding remarks

We have presented two versions of the Newton iteration for solving equations of the type $f(y) = 0$ where $f$ is a map from a Lie group to its corresponding Lie algebra. The main motivation has been to facilitate the use of implicit methods for solving ordinary differential equations on manifolds, an area that has recently been subject to considerable interest. Clearly, the approach presented here is only applicable to the case where the manifold in question is a Lie group. This is however not an unimportant case, there are many examples from applications of initial value problems whose solution evolve on a Lie group. For instance, the matrix groups $\mathrm{SO}(3,\mathbb{R})$ and $\mathrm{SL}(3,\mathbb{R})$ are frequently used to model rotations and volume preservation respectively in 3D space. It should be the subject of future research to investigate the behavior of the proposed algorithms applied to such problems.

- Generalization to homogeneous manifolds:

  A more general class of manifolds that contains the Lie groups as a special case are the *homogeneous manifolds*. Such a manifold consists of a Lie group $G$ that acts transitively on a manifold $\mathcal{M}$ through a group action $\lambda : G \times \mathcal{M} \to \mathcal{M}$. It is a well known result [3] that $\mathcal{M}$ is naturally diffeomorphic to $G/G_m$ for any $m \in \mathcal{M}$, the left cosets of the *stabilizer subgroup* $G_m = \{y \in G, \lambda(y,m) = m\}$. In the case where $G_m = \{e\}$, the trivial group, one obtains $\mathcal{M} = G$ and the group action $\lambda$ reduces to the group multiplication $\cdot$ in $G$. However, if $G_m$ contains more elements, the induced map $\lambda_m : y \mapsto \lambda(y,m)$ is not injective. The integration methods proposed by Crouch and Grossman [4] can be interpreted in the context of homogeneous manifolds, see [16]. Certain difficulties are encountered if a straightforward generalization of the Newton iteration presented in this paper is attempted. The obvious approach of identifying the iteration sequence on $\mathcal{M}$ with elements of the Lie algebra of $G$ is impeded by the fact that there is no longer a unique local correspondence between points in $\mathcal{M}$ and elements of $G$ relative to a reference point $m \in \mathcal{M}$, as is the case between $G$ and $\mathfrak{g}$ via the exponential mapping around 0. This may suggest that for homogeneous manifolds, it may be more useful to consider representations in $T\mathcal{M}_m$ where $m$ is a fixed reference point on $\mathcal{M}$. This amounts to the introduction of a Riemannian metric on $T\mathcal{M}$. This idea conforms well with the work of Edelman et al. [5].

- Application to high-order implicit integrators:

  To illustrate the use of the new algorithms for solving initial value problems with implicit methods, we have used the implicit Euler method. It should be noted that the technique we use can be extended in a straightforward way to general RK methods of Crouch-Grossman type for Lie groups. For instance, with DIRK type Crouch-Grossman methods one simply needs to solve, in sequence, equations of the type

  $$Y_i = Z_i \cdot \exp(h a_{ii} g(Y_i)),$$

  $$\text{where} \qquad Z_i = y^{(k)} \cdot \exp(h a_{i,1} g(Y_1)) \cdots \exp(h a_{i,i-1} g(Y_{i-1})),\ i = 1,\ldots,s.$$

Finally, it should be noted that there are still many interesting problems to be solved in order to obtain streamlined implementations of the Newton iteration on Lie groups. Although we believe that the abstract formulation of the iteration technique is satisfactory, several challenges remain on the more detailed level. Computing the differential of $f$ and solving the iteration equations are examples of such challenges. Another difficulty is that in the evaluation of the function $f$ in version 1, one makes use of the inverse of the exponential map. If $G$ is a matrix Lie group, the `logm` function in Matlab may be used for this purpose, but the computation is fairly expensive. However, there are special cases in which an explicit formula is available for the inverse of the exponential. An example of this is $\mathrm{SO}(3,\mathbb{R})$ where the logarithm can be obtained in a very stable way and very efficiently (with less arithmetic operations than the Rodrigues formula which is commonly used for computing the matrix exponential in $\mathfrak{so}(3)$).

Although many questions are still to be addressed regarding the solution of nonlinear equations on manifolds, we believe that the numerical results from the algorithms proposed in this paper indicate that this new approach may be a worthwhile alternative to existing methods.


## Acknowledgements

This work was completed while the second author was visiting the Numerical group at Institutt for matematiske fag at NTNU. He thanks everyone for making this collaboration possible and pleasant. Both authors also thank Hans Munthe-Kaas for valuable discussions and inputs.


## References


[1] R. Abraham, J. E. Marsden and T. Ratiu, *Manifolds, Tensor Analysis, and Applications*, Springer-Verlag (1980).

[2] V. I. Arnold, *Mathematical Methods of Classical Mechanics*, 2nd edition, Springer GTM no. 60, New York, (1989).

[3] R. L. Bryant, *An Introduction to Lie Groups and Symplectic Geometry*, in Geometry and Quantum field Theory, Eds. Freed and Uhlenbeck, AMS IAS/Park City math. series vol. 1 (1995).

[4] P. E. Crouch and R. Grossman, *Numerical Integration of Ordinary Differential Equations on Manifolds*, J. Nonlinear Sci. Vol.3 (1993) 1-33.

[5] A. Edelman, T. Arias and S. T. Smith, *Conjugate Gradient and Newton's Method on the Grassmann and Stiefel Manifolds*, Preprint, Department of Mathematics, MIT (1996).

[6] G. Golub and C. Van Loan, *Matrix Computations*, John Hopkins University Press (1992).

[7] O. Gonzalez and A. M. Stuart, *Remarks on the Qualitative Properties of Modified Equations*, Tech. Report SCCM-96-03, Department of Mechanical Engineering, Stanford University.

[8] S. Helgason, *Differential Geometry, Lie Groups and Symmetric Spaces*, Academic Press (1978).

[9] D. Higham, private communication.

[10] A. Iserles, *Solving Linear Ordinary Differential Equations by Exponentials of Iterated Commutators*, Numer. Math. 45 (1984) 183-199.

[11] S. Klarsfeld and J. A. Oteo, *The Baker-Campbell-Hausdorff formula and the convergence of the Magnus expansion*, J. Phys. A 22, #21 (1989) 4565-4572

[12] R. E. Mahony, *The Constrained Newton Method on a Lie-Group and the Symmetric Eigenvalue Problem*, Linear Algebra and Its Applications 248, #1-3 (1996) 67-89.

[13] R. E. Mahony, *Optimization Algorithms on Homogeneous Spaces: With Applications in Linear Systems Theory*, Ph. D. Thesis, Dept. of Systems Engineering, Univ. of Canberra, Australia (1994).

[14] H. Munthe-Kaas, *Runge-Kutta Methods on Lie Groups*, Preprint, Department of Informatics, University of Bergen (1996).

[15] H. Munthe-Kaas, *Lie-Butcher Theory for Runge-Kutta Methods*, BIT 35 (1996) 572-587.

[16] H. Munthe-Kaas and A. Zanna, *Numerical Integration of Differential Equations on Homogeneous Manifolds*, Preprint, Department of Informatics, University of Bergen (1996).

[17] I. Najfeld and T. F. Havel, *Derivatives of the Matrix Exponential and their Computation*, Adv. Appl. Math. 16, #3 (1995) 321-375.

[18] B. Owren and A. Marthinsen, *Runge–Kutta methods adapted to manifolds and based on rigid frames*, BIT, 39, (1999), 116–142.

[19] B. Owren and B. Welfert, *The Newton iteration on Lie groups*, Technical report 3/96, Department of Mathematical Sciences, NTNU (1996).

[20] V. S. Varadarajan, *Lie Groups, Lie Algebras and Their Representations*, Springer GTM no. 102, New York (1984).

[21] F. W. Warner, *Foundations of Differentiable Manifolds and Lie Groups*, Springer GTM no. 94, New York (1983).

[22] Y.-P. Lin and H. L.-P. Hwang, *Efficient Computation of the Matrix Exponential using Padé Approximation*, Computers & Chemistry 16, #4 (1992) 285-293.

[23] A. Zanna, *The Method of Iterated Commutators for Ordinary Differential Equations on Lie Groups*, Tech. Rept, University of Cambridge, DAMTP 1996/NA12 (1996).

[24] A. Zanna and K. Engø and H. Z. Munthe-Kaas, *Adjoint and Selfadjoint Lie-Group Methods*, Technical report 1999/NA02, DAMTP, Cambridge University, 1999.