MATHEMATICS OF COMPUTATION
Volume 73, Number 246, Pages 761–776
S 0025-5718(03)01602-8
Article electronically published on November 5, 2003

# RECURRENCE RELATIONS AND CONVERGENCE THEORY OF THE GENERALIZED POLAR DECOMPOSITION ON LIE GROUPS

ANTONELLA ZANNA

Abstract. The subject matter of this paper is the analysis of some issues related to *generalized polar decompositions* on Lie groups. This decomposition, depending on an involutive automorphism $\sigma$, is equivalent to a factorization of $z \in G$, $G$ being a Lie group, as $z = xy$ with $\sigma(x) = x^{-1}$ and $\sigma(y) = y$, and was recently discussed by Munthe-Kaas, Quispel and Zanna together with its many applications to numerical analysis.

It turns out that, contrary to $X(t) = \log(x)$, an analysis of $Y(t) = \log(y)$ is a very complicated task. In this paper we derive the series expansion for $Y(t) = \log(y)$, obtaining an explicit recurrence relation that completely defines the function $Y(t)$ in terms of projections on a Lie triple system $\mathfrak{p}_\sigma$ and a subalgebra $\mathfrak{k}_\sigma$ of the Lie algebra $\mathfrak{g}$, and obtain bounds on its region of analyticity.

The results presented in this paper have direct application, among others, to linear algebra, integration of differential equations and approximation of the exponential.

## 1. Introduction

The polar decomposition is well known in linear algebra as a useful canonical form of matrices: it consists in factoring a real matrix as the product of a positive semidefinite matrix and an orthogonal matrix [5]. However, the matrix case is just a particular setting. In fact, structures like Lie groups, endowed with involutive automorphisms, possess all the attributes needed to define the analogue of the matrix polar decomposition, the difference being that the decomposition depends on the choice of the automorphism $\sigma$ — hence the name *generalized polar decomposition.*

The generalized polar decomposition on Lie groups was recently considered in [15], and some application in numerical analysis were studied in [14]: for instance, the generalized Scovel projection [13] to generate self-adjoint numerical schemes from an arbitrary integrator for differential equations, as well as the Thue–Morse sequence [7] for improving the retention of symmetries by a numerical ODE scheme, can be associated to choosing a factor in the generalized polar decomposition of the underlying numerical method in tandem with an appropriate automorphism.

Other applications of the generalized polar decomposition occur in linear algebra. Recently, it has been shown that iterated generalized polar decompositions can be employed to approximate the matrix exponential from a Lie algebra $\mathfrak{g}$ to a Lie group

Received by the editor September 5, 2001 and, in revised form, October 1, 2002.
2000 *Mathematics Subject Classification.* Primary 51A50; Secondary 65L99, 58A99.
*Key words and phrases.* Lie group, Lie algebra, generalized polar decomposition, generalized Cartan decomposition.

$G$ [16], and the resulting algorithms are very competitive with standard methods for approximating the matrix exponential.

To be more specific, let $G$ be a Lie group (i.e., a manifold which is endowed with a multiplication compatible with its topology) and $\sigma$ an involutive automorphism (a one-to-one map $G \to G$ such that $\sigma(xy) = \sigma(x)\sigma(y)$, and $\sigma \neq \mathrm{id}$, $\sigma^2 = \mathrm{id}$). Concrete examples of Lie groups are matrix groups like the general linear group $\mathrm{GL}(n, \mathbb{C})$ of complex invertible matrices, the group of matrices with unit determinant $\mathrm{SL}(n, \mathbb{C})$, and, in general, continous subgroups of $\mathrm{GL}(n, \mathbb{C})$. Concrete example of involutive automorphisms on $G \subseteq \mathrm{GL}(n, \mathbb{C})$ are the maps

$$\sigma_1(z) = \bar{z},$$

where $\bar{z}$ is the complex conjugate of the matrix $z$,

$$\sigma_2(z) = z^{-T}$$

and, finally,

$$\sigma_3(z) = z^{-*},$$

where $z^* = \bar{z^T}$ is the Hermitian adjoint of $z$.

Returning to the more abstract setting, denote by $G^\sigma$ the subgroup of $G$ consisting of fixed points of $\sigma$, by $G^\sigma = \{y | \sigma(y) = y\}$ and $G_\sigma$ the set of 'anti-fixed' points of $G$, i.e., of the type $\sigma(x) = x^{-1}$. The set $G_\sigma$ has the structure not of a group but of a symmetric space [4], namely it is closed under a symmetric type of composition,

$$x_1 \cdot x_2 = x_1 x_2^{-1} x_1 \in G_\sigma, \qquad x_1, x_2 \in G_\sigma,$$

where $x_1 x_2$ denotes the usual product in the Lie group $G$. Now, the generalized polar decomposition of $z \in G$ is equivalent to writing

$$z = xy, \qquad x \in G_\sigma, y \in G^\sigma.$$

As mentioned above, such a decomposition always exists when $z$ is sufficiently close to the identity element $e \in G$; thus, if $z = \exp(tZ)$, then the factors $x$ and $y$ are differentiable functions and $x = \exp(X(t))$, $y = \exp(Y(t))$, for some functions $X(t)$ and $Y(t)$ in $\mathfrak{g}$. Moreover, since $\sigma$ induces an involutive automorphism $\mathrm{d}\sigma$ on the Lie algebra $\mathfrak{g}$, then $\mathfrak{g}$ is split into the direct sum of two vector spaces,

$$\mathfrak{g} = \mathfrak{p}_\sigma \oplus \mathfrak{k}_\sigma,$$

$\mathfrak{p}_\sigma$ corresponding to the eigenvalue $\lambda = -1$ of $\mathrm{d}\sigma$ (i.e., $P \in \mathfrak{p}_\sigma$ implies that $\mathrm{d}\sigma(P) = -P$), $\mathfrak{k}_\sigma$ corresponding to the eigenvalue $\lambda = 1$ of $\mathrm{d}\sigma$ (i.e., $K \in \mathfrak{k}_\sigma$ implies that $\mathrm{d}\sigma(K) = K$). It is easily verified that $\mathfrak{k}_\sigma$ is a subalgebra of $\mathfrak{g}$, while $\mathfrak{p}_\sigma$ has the structure of a *Lie triple system*, meaning that $\mathfrak{p}_\sigma$ is closed not under the commutator but under the double commutator: $[X_1, [X_2, X_3]] \in \mathfrak{p}_\sigma$ whenever $X_1, X_2, X_3 \in \mathfrak{p}_\sigma$. The functions $X(t)$ and $Y(t)$ have the property that

$$X(t) \in \mathfrak{p}_\sigma, \qquad Y(t) \in \mathfrak{k}_\sigma,$$

and are analytic at $t = 0$, provided that the commutator is a bounded operator and that $\mathfrak{g}$ is a Banach space. Thus, $X(t) = \sum_{n=0}^{\infty} t^n X_n$ and $Y(t) = \sum_{n=0}^{\infty} t^n Y_n$. The recurrence relation for the $X_n$s was derived in [15]. However, although in the same paper the authors showed that $Y(t)$ expands in odd powers of $t$ only and obtained a differential equation for $Y(t)$, such an equation was not independent of the knowledge of $X(t)$, and this dependence rendered the derivation of a recurrence relation for $Y(t)$ a rather convoluted task.

In this paper we shall follow a different approach and hence complete the work in [15] by deriving an explicit recurrence relation for the $Y_n$s. This will be accomplished in two phases: first, in Section 2 we derive a differential equation obeyed by $Y$, this time independent of the direct knowledge of $X(t)$. Although the differential equation is implicit in $Y'$, in Section 3 we derive a recurrence relation for the $Y_n$s expanding all the terms and operators in series and comparing powers of $t$, and estimate the radius of convergence for the function $Y(t)$ as a series expansion. Some applications of our results to the polar decomposition of matrices, integration of differential equations with time-reversal symmetries and approximation of the exponential of a matrix from a Lie algebra to a Lie group are described in Section 4, while in Section 5 we present some discussion and concluding remarks on the issues of this paper.

1.1. **Notation and background theory.** We use the notation convention of [4] to denote group and algebra elements. Thus, group elements will be denoted by lowercase Latin letters and algebra elements by uppercase Latin letters. A generic Lie group will be denoted by $G$, while Lie algebras (and their subspaces) will be denoted by Gothic letters. For background theory on symmetric spaces we refer mainly to [4] and [11].

To simplify notation, we will make use time and again of the following Lie-group and Lie-algebra operators,

$$\begin{aligned}
\mathrm{Ad}_a B &= aBa^{-1}, \qquad a \in G, B \in \mathfrak{g}, \\
\mathrm{ad}_A B &= [A, B], \qquad A, B \in \mathfrak{g}, \\
\mathrm{dexp}_A &= \left.\frac{\mathrm{e}^u - 1}{u}\right|_{u=\mathrm{ad}_A} = \sum_{k=0}^{\infty} \frac{1}{(k+1)!} \mathrm{ad}_A^k, \qquad A \in \mathfrak{g}, \\
\mathrm{dexp}_A^{-1} &= \left.\frac{u}{\mathrm{e}^u - 1}\right|_{u=\mathrm{ad}_A} = \sum_{k=0}^{\infty} \frac{\mathrm{B}_k}{k!} \mathrm{ad}_A^k, \qquad A \in \mathfrak{g}
\end{aligned}$$

(the $B_k$s are Bernoulli numbers, see [1]), whose main properties are described at length in [8].

Let $G$ be a Lie group and $\mathfrak{g}$ the corresponding Lie algebra. Denote by $\sigma : G \to G$ an involutive automorphism and let $\mathrm{d}\sigma$ be the corresponding Lie-algebra automorphism, defined as

$$\mathrm{d}\sigma(Z) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \sigma(\exp(tZ)) \tag{1}$$

($\mathrm{d}\sigma = \mathrm{d}\sigma_e$, the tangent map of $\sigma$ at $e$, where $e$ is the identity element of $G$). Set $\mathfrak{p}_\sigma = \{P \in \mathfrak{g} | \, \mathrm{d}\sigma(P) = -P\}$ and $\mathfrak{k}_\sigma = \{K \in \mathfrak{g} | \, \mathrm{d}\sigma(K) = -K\}$. Then, every element $Z \in \mathfrak{g}$ can be uniquely written as

$$Z = P + K, \qquad P \in \mathfrak{p}_\sigma, \quad K \in \mathfrak{k}_\sigma,$$

where

$$P = \frac{1}{2}(Z - \mathrm{d}\sigma(Z)) \qquad \text{and} \qquad K = \frac{1}{2}(Z + \mathrm{d}\sigma(Z)).$$

In other words, $\mathfrak{g}$ decomposes as the direct sum

$$\mathfrak{g} = \mathfrak{p}_\sigma \oplus \mathfrak{k}_\sigma. \tag{2}$$

We refer to the decomposition (2) as a *generalized Cartan decomposition* of $\mathfrak{g}$, to distinguish it from *Cartan decompositions* of semi-simple Lie algebras, which are

again of the form (2) but possess the additional feature that the Cartan-Killing form $\langle A, B\rangle = \mathrm{tr}(\mathrm{ad}_A \mathrm{ad}_B)$ is strictly negative definite on $\mathfrak{k}_\sigma$ and strictly positive definite on $\mathfrak{p}_\sigma$ [4].

The linear subspace $\mathfrak{k}_\sigma$ is a subalgebra of $\mathfrak{g}$ (corresponding to the eigenvalue $\lambda = 1$ of $\mathrm{d}\sigma$), while $\mathfrak{p}_\sigma$ is a Lie triple system (linear subspace of $\mathfrak{g}$ corresponding to the eigenvalue $\lambda = -1$), and is closed under the double commutator $\mathrm{ad}^2$, that is,

$$[A, [B, C]] \in \mathfrak{p}_\sigma, \qquad A, B, C \in \mathfrak{p}_\sigma,$$

while in general it is not true that $[A, B]$ is in $\mathfrak{p}_\sigma$.

The sets $\mathfrak{p}_\sigma$ and $\mathfrak{k}_\sigma$ obey the important inclusion relations

$$\begin{aligned}[\mathfrak{k}_\sigma, \mathfrak{k}_\sigma] &\subseteq \mathfrak{k}_\sigma,\\ [\mathfrak{k}_\sigma, \mathfrak{p}_\sigma] &\subseteq \mathfrak{p}_\sigma,\\ [\mathfrak{p}_\sigma, \mathfrak{p}_\sigma] &\subseteq \mathfrak{k}_\sigma.\end{aligned}$$

Since we are interested in series expansions in a neighbourhood of the identity $e \in G$, we set $z = \exp(tZ)$ and assume that $Z = P + K$ is the canonical projection of $Z$ into $\mathfrak{p}_\sigma \oplus \mathfrak{k}_\sigma$, i.e., $\mathrm{d}\sigma(P) = -P$ and $\mathrm{d}\sigma(K) = K$. We denote $w = (\sigma(z))^{-1} = \sigma(z^{-1})$. Note that, because of (1), it is true that $w = \exp(tW)$, where $W = -\mathrm{d}\sigma(Z) = P - K$.

We also set

$$X(t) = \sum_{n=0}^{\infty} X_n t^n, \qquad Y(t) = \sum_{n=0}^{\infty} Y_n t^n.$$

Before proceeding further, it is useful to recall the following results derived in [15].

**Theorem 1.1.** *The following statements hold.*

(i) *For $t$ sufficiently small, $z = \exp(tZ)$ admits the generalized polar decomposition $z = xy$, where $x = \exp(X(t))$, $X(t) \in \mathfrak{p}_\sigma$, and $y = \exp(Y(t))$, $Y(t) \in \mathfrak{k}_\sigma$.*

(ii) *The coefficients of the series expansion of $X(t)$ obey the recurrence relation*

$$\begin{aligned}(n+1)X_{n+1} &= -[X_n, K]\\ &\quad + \sum_{\substack{\ell \geq 1\\ 2\ell \leq n}} c_{2\ell} \sum_{\substack{\ell_1, \ldots, \ell_{2\ell} > 0\\ \ell_1 + \cdots + \ell_{2\ell} = n}} [X_{\ell_1}, [X_{\ell_2}, \cdots, [X_{\ell_{2\ell}}, P]]],\\ X_1 &= P,\end{aligned} \tag{3}$$

*for $n = 1, 2, \ldots$, where $c_{2\ell} = \frac{2^{2\ell} \mathrm{B}_{2\ell}}{(2\ell)!}$ and $\mathrm{B}_\ell$ is the $\ell$-th Bernoulli number* [1].

(iii) *Assume that $\mathfrak{g}$ is a Banach space and that there exists $0 < \mu \leq 2$ such that $\|\mathrm{ad}_A B\| \leq \mu \|A\| \|B\|$ for all $A, B \in \mathfrak{g}$. Then, the function $X(t)$ is analytic in a sphere of radius $\rho = \frac{\delta}{\alpha\mu}$, where $0 < \delta < \pi$ and $\alpha = \max\{\|P\|, \|K\|\}$.*

**Lemma 1.2.** *The function $Y(t)$ is an odd function of $t$, that is,*

$$Y_{2n} = 0, \qquad n = 0, 1, 2, \ldots. \tag{4}$$

As a consequence of the above lemma, only the odd coefficients of $Y$ need be determined.

## 2. An implicit differential equation for $Y(t)$

Assume that $z = \exp(tZ)$ admits a generalized polar decomposition $z = xy$ as in Theorem 1.1. Recall that $\sigma(x) = x^{-1}$ and $\sigma(y) = y$, so that

$$z\sigma(z)^{-1} = xy\sigma(xy)^{-1} = xyy^{-1}x = x^2, \tag{5}$$

from which it is easily verified (taking the logarithm on both sides) that

$$X(t) = \frac{1}{2}\mathrm{bch}(tZ, tW) = \frac{1}{2}\mathrm{bch}(t(P+K), t(P-K)),$$

$\mathrm{bch}(\cdot,\cdot)$ being the familiar operator of the Baker–Campbell–Hausdorff (BCH) formula,

$$\exp(Z_1)\exp(Z_2) = \exp(\mathrm{bch}(Z_1, Z_2)),$$

where

$$\mathrm{bch}(Z_1, Z_2) = Z_1 + Z_2 + \frac{1}{2}[Z_1, Z_2] + \frac{1}{12}([Z_1, [Z_1, Z_2]] - [Z_2, [Z_1, Z_2]]) + \cdots.$$

Note that the function

$$\tilde{Z}(t) = \mathrm{bch}(tZ_1, tZ_2) = \sum_{n=0}^{\infty} \tilde{Z}_n t^n \tag{6}$$

obeys a recurrence relation very similar to (3) (see [18]).

Once $x = \exp(X(t))$ is known, one has

$$y = x^{-1}z;$$

hence, taking logarithm on both sides and using the bch formula, it is true that

$$Y(t) = \mathrm{bch}(-X(t), tZ). \tag{7}$$

Thus, the factors $Y_n$ of the expansion of $Y(t)$ can be obtained by carefully tracing powers of $t$.

However, a main complication is that both $X(t)$ and the BCH formula are known by recursion only, which means that (7) is a double infinite recursion. Furthermore, there is another objection to this procedure, on a more philosophical basis: As we have seen in Theorem 1.1, $x = \exp(X(t))$ can be expressed independently of $y = \exp(Y(t))$ and as a function of $z$ and $\sigma(z)$ only. The same must holds for $y$, since—although coupled by $z$ and $\sigma(z)$—after all, the functions $X(t)$ and $Y(t)$ live in different linear subspaces!

Motivated by this thought, our goal is to derive a relation for $y$, defined in terms of $z$ and $\sigma(z)$ only. As a consequence, the differentiation of this relation will allow us to obtain an explicit recurrence relation for $Y$ in terms of $Z = P + K$ and $W = P - K$.

Our point of departure is the observation that $z = xy$ implies $\sigma(z) = x^{-1}y$, and hence

$$x = y\sigma(z)^{-1} = yw.$$

Substitution in $z = xy$ yields

$$z = xy = ywy,$$

a relation invoking only $y, z$ and $w = \sigma(z)^{-1}$, as desired. In order to differentiate, it is however more convenient to multiply the above equality on both sides by $y^{-1}$ on the right, to obtain

$$zy^{-1} = yw. \tag{8}$$

Differentiating (8) in the usual fashion, we obtain

$$\begin{aligned} Zzy^{-1} - z\mathrm{dexp}_{-Y}Y'y^{-1} &= \mathrm{dexp}_Y Y'w + yWw \\ &= \mathrm{dexp}_Y Y'zy^{-1} + zy^{-1}W, \end{aligned}$$

where we have made use of the fact that $w$ and $W$ commute, as well as (8). Recall that also $z$ and $Z$ commute. Hence, multiplication of both sides of the above equation by $z^{-1}$ on the left and by $y$ on the right yields

$$Z - \mathrm{dexp}_{-Y}Y' = z^{-1}\mathrm{dexp}_Y Y'z + y^{-1}Wy. \tag{9}$$

Note that

$$z^{-1}\mathrm{dexp}_Y Y'z = \mathrm{Ad}_{\exp(-tZ)}\mathrm{dexp}_Y Y' = \exp(-t\mathrm{ad}_Z)\mathrm{dexp}_Y Y',$$

and by a similar token,

$$y^{-1}Wy = \exp(-\mathrm{ad}_Y)W.$$

It is convenient to isolate the first term of the two exponential operators above; hence (9) becomes

$$Z - \mathrm{dexp}_{-Y}Y' = \mathrm{dexp}_Y Y' + \{\exp(-t\mathrm{ad}_Z) - 1\}\mathrm{dexp}_Y Y' + W + \{\exp(-\mathrm{ad}_Y) - 1\}W,$$

from which, collecting the free 'dexp' terms on one side,

$$\{\mathrm{dexp}_{-Y} + \mathrm{dexp}_Y\}Y' = 2K - \{\exp(-t\mathrm{ad}_Z) - 1\}\mathrm{dexp}_Y Y' - \{\exp(-\mathrm{ad}_Y) - 1\}W. \tag{10}$$

From

$$\mathrm{dexp}_{-u} = \frac{\mathrm{e}^{-u} - 1}{u}\Big|_{u=\mathrm{ad}_Y}, \qquad \mathrm{dexp}_u = \frac{\mathrm{e}^{u} - 1}{u}\Big|_{u=\mathrm{ad}_Y},$$

we observe that

$$\mathrm{dexp}_{-u} + \mathrm{dexp}_u = \frac{\mathrm{e}^{u} - \mathrm{e}^{-u}}{u} = 2\frac{\sinh u}{u};$$

hence (10) reduces to

$$2\frac{\sinh u}{u}\Big|_{u=\mathrm{ad}_Y} Y' = 2K - \sum_{k=1}^{\infty} \frac{(-t)^k}{k!}\mathrm{ad}_Z^k \mathrm{dexp}_Y Y' - \sum_{k=1}^{\infty} \frac{(-1)^k}{k!}\mathrm{ad}_Y^k W. \tag{11}$$

## 3. The series expansion and recurrence relation

Instead of inverting the operator $\frac{\sinh u}{u}$, $u = \mathrm{ad}_Y$, on the left-hand-side of (11), we observe that (11) can already be used to derive, after some algebra, an explicit recurrence relation for the $Y_n$s. Let us begin by expanding in series the term on the left hand side of (11). Since $Y(t) = \sum_{n=0}^{\infty} t^{2n+1}Y_{2n+1}$, it is easily seen that

$$Y'(t) = \sum_{n=0}^{\infty} (2n+1)t^{2n}Y_{2n+1}.$$

Moreover,

$$\frac{\sinh u}{u}\Big|_{u=\mathrm{ad}_Y} = \sum_{k=0}^{\infty} \frac{1}{(2k+1)!}\mathrm{ad}_Y^{2k}$$

(note that it is an even function of $t$, as it is $Y'$). Hence, changing the order of summation and isolating the first term,

$$\begin{aligned}\frac{\sinh u}{u}\Big|_{u=\mathrm{ad}_Y} Y' &= \sum_{k=0}^{\infty} \frac{1}{(2k+1)!}\mathrm{ad}_Y^{2k} \sum_{n=0}^{\infty}(2n+1)t^{2n}Y_{2n+1} \\ &= \sum_{n=0}^{\infty}(2n+1)t^{2n}Y_{2n+1} + \sum_{n=0}^{\infty}(2n+1)t^{2n}\sum_{k=1}^{\infty}\frac{1}{(2k+1)!}\mathrm{ad}_Y^{2k}(Y_{2n+1}).\end{aligned}$$

We next perform a series expansion of the second term on the above right-hand-side, which, after matching powers of $t$, results in

$$\sum_{n=1}^{\infty} t^{2n}\sum_{q=1}^{n}\sum_{\substack{k\geq 1\\ k\leq q}}\frac{1}{(2k+1)!}\sum_{\substack{k_1,\ldots,k_{2k}>0\\ k_1+\cdots+k_{2k}=2q}}[Y_{k_1},\ldots,[Y_{k_{2k}},Y_{2(n-q)+1}],\ldots].$$

It turns out that the coefficient of the $t^{2n}$ term of the left hand side of (11) is given by

(12)
$$2(2n+1)Y_{2n+1} + 2\sum_{q=1}^{n}\sum_{\substack{k\geq 1\\ k\leq q}}\frac{1}{(2k+1)!}\sum_{\substack{k_1,\ldots,k_{2k}>0\\ k_1+\cdots+k_{2k}=2q}}[Y_{k_1},\ldots,[Y_{k_{2k}},Y_{2(n-q)+1}],\ldots].$$

Next, we turn our attention to the right hand side of (11). Let us start from the term

$$\sum_{k=1}^{\infty}\frac{(-t)^k}{k!}\mathrm{ad}_Z^k \mathrm{dexp}_Y Y'. \tag{13}$$

Since

$$\begin{aligned}\mathrm{dexp}_Y Y' &= \sum_{n=0}^{\infty}(2n+1)t^{2n}\mathrm{dexp}_Y Y_{2n+1} \\ &= \sum_{n=0}^{\infty}(2n+1)t^{2n}Y_{2n+1} + \sum_{n=0}^{\infty}(2n+1)t^{2n}\{\mathrm{dexp}_Y - 1\}Y_{2n+1},\end{aligned}$$

the term (13) can be split into two parts, say $T_1 + T_2$, where

$$T_1 = \sum_{k=1}^{\infty}\frac{(-t)^k}{k!}\mathrm{ad}_Z^k\sum_{n=0}^{\infty}(2n+1)t^{2n}Y_{2n+1},$$

and

$$T_2 = \sum_{k=1}^{\infty}\frac{(-t)^k}{k!}\mathrm{ad}_Z^k\sum_{n=0}^{\infty}(2n+1)t^{2n}\left(\sum_{m=1}^{\infty}\frac{1}{(m+1)!}\mathrm{ad}_Y^m Y_{2n+1}\right).$$

We analyse $T_1$ first. We observe that $T_1$ consists in even and odd powers of $t$; however, given that the left hand side of (11) possesses even powers of $t$ only, we disregard the odd powers in $T_1$ and consider the even part only, $T_{1,\mathrm{even}}$, given by

$$T_{1,\mathrm{even}} = \sum_{n=1}^{\infty} t^{2n}\sum_{m=1}^{n}\frac{2(n-m)+1}{(2m)!}\mathrm{ad}_Z^{2m}Y_{2(n-m)+1}.$$

Hence, the coefficient of $t^{2n}$ is

$$\sum_{m=1}^{n} \frac{2(n-m)+1}{(2m)!} \mathrm{ad}_Z^{2m} Y_{2(n-m)+1}. \tag{14}$$

The analysis of the $T_2$ term is more complicated. First, merging the first two series, we obtain

$$T_2 = -t\mathrm{ad}_Z \sum_{\ell=0}^{\infty} t^\ell \sum_{j=0}^{\ell} \frac{(-1)^{\ell-j}(j+1)}{(\ell-j+1)!} \mathrm{ad}_Z^{\ell-j}.$$

Second, observe that

$$\begin{aligned}
&\sum_{m=1}^{\infty} \frac{1}{(m+1)!} \mathrm{ad}_Y^m Y_{j+1} \\
&\qquad = \sum_{q=0}^{\infty} t^{q+1} \sum_{\substack{k\geq 1 \\ k\leq q+1}} \frac{1}{(k+1)!} \sum_{\substack{j_1,\ldots,j_k>0 \\ j_1+\cdots+j_k=q+1}} [Y_{j_1}, \ldots, [Y_{j_k}, Y_{j+1}] \ldots],
\end{aligned}$$

from which we obtain

$$\begin{aligned}
T_2 \;=\; & \sum_{m=0}^{\infty} t^{m+2} \sum_{q=0}^{m} \sum_{j=0}^{m-q} \frac{(-1)^{m-q-j+1}(j+1)}{(m-q-j+1)!} \mathrm{ad}_Z^{m-j-q+1} \\
& \times \sum_{\substack{k\geq 1 \\ k\leq q+1}} \frac{1}{(k+1)!} \sum_{\substack{j_1,\ldots,j_k>0 \\ j_1+\cdots+j_k=q+1}} [Y_{j_1}, \ldots, [Y_{j_k}, Y_{j+1}] \ldots].
\end{aligned}$$

To obtain the coefficient of $t^{2n}$ in the above expression, it is sufficient to set $m+2 = 2n$, from which we deduce that $m = 2(n-1)$. Substituting above, the required coefficient is

$$\begin{aligned}
&\sum_{q=0}^{2(n-1)} \sum_{j=0}^{2(n-1)-q} \frac{(-1)^{2n-q-j-1}(j+1)}{(2n-q-j-1)!} \mathrm{ad}_Z^{2n-j-q-1} \\
&\quad \times \sum_{\substack{k\geq 1 \\ k\leq q+1}} \frac{1}{(k+1)!} \sum_{\substack{j_1,\ldots,j_k>0 \\ j_1+\cdots+j_k=q+1}} [Y_{j_1}, \ldots, [Y_{j_k}, Y_{j+1}] \ldots].
\end{aligned}$$

It remains to analyse the last term of (11),

$$\sum_{k=1}^{\infty} \frac{(-1)^k}{k!} \mathrm{ad}_Y^k W,$$

a task that offers no particular difficulties. Let us recall, however, that $Y$ is an odd function of $t$; therefore terms of the form $\mathrm{ad}_Y^{2k+1}$ only contribute with odd powers of $t$ to the expansion, while we are interested in even powers of $t$. Therefore, we disregard the odd powers altogether, to consider the expansion of

$$\sum_{k=1}^{\infty} \frac{1}{(2k)!} \mathrm{ad}_Y^{2k} W$$

instead. This reduces to

$$\sum_{n=1}^{\infty} t^{2n} \sum_{\substack{\ell \geq 1 \\ \ell \leq n}} \frac{1}{(2\ell)!} \sum_{\substack{\ell_1,\ldots,\ell_{2\ell}>0 \\ \ell_1+\cdots+\ell_{2l}=2n}} [Y_{\ell_1}, \ldots, [Y_{\ell_{2\ell}}, W], \ldots],$$

and the coefficient of $t^{2n}$ is

$$\sum_{\substack{\ell \geq 1 \\ \ell \leq n}} \frac{1}{(2\ell)!} \sum_{\substack{\ell_1,\ldots,\ell_{2\ell}>0 \\ \ell_1+\cdots+\ell_{2l}=2n}} [Y_{\ell_1}, \ldots, [Y_{\ell_{2\ell}}, W], \ldots]. \tag{15}$$

Now we have all the elements to complete the proof of the main result of this paper.

**Theorem 3.1.** *Let $G$ be a Lie group and $\sigma$ an involutive automorphism so that the generalized polar decomposition $z = \exp(tZ) = xy = \exp(X(t))\exp(Y(t))$ exists in the sense of* [15]. *The function $Y(t)$ admits a series expansion $Y(t) = \sum_{n=0}^{\infty} Y_n t^n$, whose coefficient obey the explicit recurrence relation*

$$\begin{aligned}
Y_1 &= K, \qquad\qquad (16)\\
Y_{2n} &= 0, \qquad n = 0, 1, 2, \ldots,\\
2(2n+1)Y_{2n+1} &= -2\sum_{q=1}^{n} \sum_{\substack{k \geq 1 \\ k \leq q}} \frac{1}{(2k+1)!} \sum_{\substack{k_1,\ldots,k_{2k}>0 \\ k_1+\cdots+k_{2k}=2q}} [Y_{k_1}, \ldots, [Y_{k_{2k}}, Y_{2(n-q)+1}], \ldots]\\
&\quad - \sum_{m=1}^{n} \frac{2(n-m)+1}{(2m)!} \mathrm{ad}_Z^{2m} Y_{2(n-m)+1} \qquad (17)\\
&\quad - \sum_{q=0}^{2(n-1)} \sum_{j=0}^{2(n-1)-q} \frac{(-1)^{2n-q-j-1}(j+1)}{(2n-q-j-1)!} \mathrm{ad}_Z^{2n-j-q-1}\\
&\qquad \times \sum_{\substack{k \geq 1 \\ k \leq q+1}} \frac{1}{(k+1)!} \sum_{\substack{j_1,\ldots,j_k>0 \\ j_1+\cdots+j_k=q+1}} [Y_{j_1}, \ldots, [Y_{j_k}, Y_{j+1}] \ldots]\\
&\quad - \sum_{\substack{\ell \geq 1 \\ \ell \leq n}} \frac{1}{(2\ell)!} \sum_{\substack{\ell_1,\ldots,\ell_{2\ell}>0 \\ \ell_1+\cdots+\ell_{2l}=2n}} [Y_{\ell_1}, \ldots, [Y_{\ell_{2\ell}}, W], \ldots].
\end{aligned}$$

*Here, $Z = P + K$ and $W = P - K$, where $\mathrm{d}\sigma(P) = -P$ and $\mathrm{d}\sigma(K) = K$ (i.e., $P \in \mathfrak{p}_\sigma$ and $K \in \mathfrak{k}_\sigma$).*

*Proof.* Having derived the series expansion of the differential equation (11) obeyed by $Y(t)$, we need only to match the coefficients of the $t^{2n}$ term in (11), for $n = 0, 1, \ldots$, since we already know that the function $Y(t)$ is odd in $t$ [15]. Thus, (12)–(15), in tandem with (11), result in the relations (16) and (17), thus proving the first part of the theorem. □

The first terms in the expansion of the function $Y(t)$ are

$$\begin{aligned}
Y = Kt &- \tfrac{1}{12}[P,[P,K]]t^3 + \Big(\tfrac{1}{120}[P,[P,[P,[P,K]]]] \\
&+ \tfrac{1}{720}[K,[K,[P,[P,K]]]] - \tfrac{1}{240}[[P,K],[K,[P,K]]]\Big)t^5 + \mathcal{O}\big(t^7\big)\,.
\end{aligned} \tag{18}$$

If $\mathfrak{g}$ is a Banach space and ad is a bounded operator, i.e., $\|\mathrm{ad}_A B\| \leq \mu \|A\| \|B\|$ for some finite $\mu$, then $Y(t)$ is analytic at the origin with radius of convergence $\rho = \frac{\delta_{\mathrm{bch}}}{\beta\mu}$, where $\delta_{\mathrm{bch}}$ is the radius of convergence of the BCH formula and $\beta = \max\{t\|Z\|, \|X(t)\|\}$. This is a consequence of the convergence of the series $X(t)$, proven in [15] and that of the BCH formula [18]. In particular, the series $Y(t)$ always converges for $t$ sufficiently small in the finite dimensional case.

However, a better estimate of the radius of convergence of $Y$ in terms of $\alpha = \max\{\|P\|, \|K\|\}$ and $\mu$ is given in the result below.

**Theorem 3.2.** *Under the same assumptions as Theorem 1.1, assume that the series*

$$\sum_{n=1}^{\infty} w_n t^n$$

*has radius of convergence $\bar{\delta}$, where $w_n$ is the solution of the recurrence relation*

$$\begin{aligned}
w_1 &= \frac{1}{2}, \\
w_{2n} &= 0, \quad n = 1, 2, \ldots, \\
2(2n+1) w_{2n+1} &= 2 \sum_{q=1}^{n} w_{2(n-q)+1} \sum_{\substack{k \geq 1 \\ k \leq q}} \frac{1}{(2k+1)!} \sum_{\substack{k_1, \ldots, k_{2k} > 0 \\ k_1 + \cdots + k_{2k} = 2q}} w_{k_1} \cdots w_{k_{2k}} \\
&\quad + \sum_{m=1}^{n} \frac{2(n-m)+1}{(2m)!} w_{2(n-m)+1} \\
&\quad + \sum_{q=0}^{2(n-1)} \sum_{j=0}^{2(n-1)-q} \frac{(j+1)}{(2n-q-j-1)!} w_{j+1} \\
&\qquad \times \sum_{\substack{k \geq 1 \\ k \leq q+1}} \frac{1}{(k+1)!} \sum_{\substack{j_1, \ldots, j_k > 0 \\ j_1 + \cdots + j_k = q+1}} w_{j_1} \cdots w_{j_k} \\
&\quad + \sum_{\substack{\ell \geq 1 \\ \ell \leq n}} \frac{1}{(2\ell)!} \sum_{\substack{\ell_1, \ldots, \ell_{2\ell} > 0 \\ \ell_1 + \cdots + \ell_{2l} = 2n}} w_{\ell_1} \cdots w_{\ell_{2\ell}}.
\end{aligned}$$

*Then, the radius of convergence of*

$$Y(t) = \sum_{n=0}^{\infty} Y_n t^n \quad \textit{is } \bar{\rho} \geq \frac{\bar{\delta}}{2\alpha\mu}.$$

*Proof.* First observe that the $w_n$ are all positive. Next, we show that $\|Y_j\| \leq \frac{(2\alpha\mu)^j}{\mu} w_j$ for $j = 1, 2, \ldots$. Clearly, the assertion is true for $j = 1$, since $\|Y_1\| = \|K\| \leq \alpha = 2\alpha w_1$, and for even values of $j$. Assume now that the statement is true for all $j = 2, 3, \ldots, 2n$. Passing to the norm in (17), using the triangle inequality and the fact that $\|Z\|, \|W\| \leq \|P\| + \|K\| \leq 2\alpha$, together with the induction hypothesis,

we have

$$
\begin{aligned}
&2(2n+1)\|Y_{2n+1}\| \\
&\quad\le 2\sum_{q=1}^{n}\frac{(2\alpha\mu)^{2(n-q)+1}}{\mu}w_{2(n-q)+1}\sum_{\substack{k\ge 1\\ k\le q}}\frac{\mu^{2k}}{(2k+1)!}\sum_{\substack{k_1,\dots,k_{2k}>0\\ k_1+\cdots+k_{2k}=2q}}\frac{(2\alpha\mu)^{2q}}{(\mu)^{2k}}w_{k_1}\cdots w_{k_{2k}} \\
&\qquad+\sum_{m=1}^{n}\frac{2(n-m)+1}{(2m)!}(2\alpha\mu)^{2m}\frac{(2\alpha\mu)^{2(n-m)+1}}{\mu}w_{2(n-m)+1} \\
&\qquad+\sum_{q=0}^{2(n-1)}\sum_{j=0}^{2(n-1)-q}\frac{(j+1)}{(2n-q-j-1)!}(2\alpha\mu)^{2n-j-q-1}\frac{(2\alpha\mu)^{j+1}}{\mu}w_{j+1} \\
&\qquad\qquad\times\sum_{\substack{k\ge 1\\ k\le q+1}}\frac{\mu^k}{(k+1)!}\sum_{\substack{j_1,\dots,j_k>0\\ j_1+\cdots+j_k=q+1}}\frac{(2\alpha\mu)^{q+1}}{\mu^k}w_{j_1}\cdots w_{j_k} \\
&\qquad+\sum_{\substack{\ell\ge 1\\ \ell\le n}}\frac{2\alpha}{(2\ell)!}\mu^{2\ell}\frac{(2\alpha\mu)^{2n}}{\mu^{2\ell}}\sum_{\substack{\ell_1,\dots,\ell_{2\ell}>0\\ \ell_1+\cdots+\ell_{2l}=2n}}w_{\ell_1}\cdots w_{\ell_{2\ell}} \\
&\quad=2(2n+1)\frac{(2\alpha\mu)^{2n+1}}{\mu}w_{2n+1},
\end{aligned}
$$

from which we deduce that the assertion is true also for $j=2n+1$. Thus, the series $\sum_{n=1}^{\infty}\|Y_n\|t^n$ is bounded by the converging series $\frac{1}{\mu}\sum_{n=1}^{\infty}w_n(2\alpha\mu)^n t^n$, from which we deduce that the radius of convergence of $Y(t)$ is at least $\bar{\delta}/(2\alpha\mu)$. □

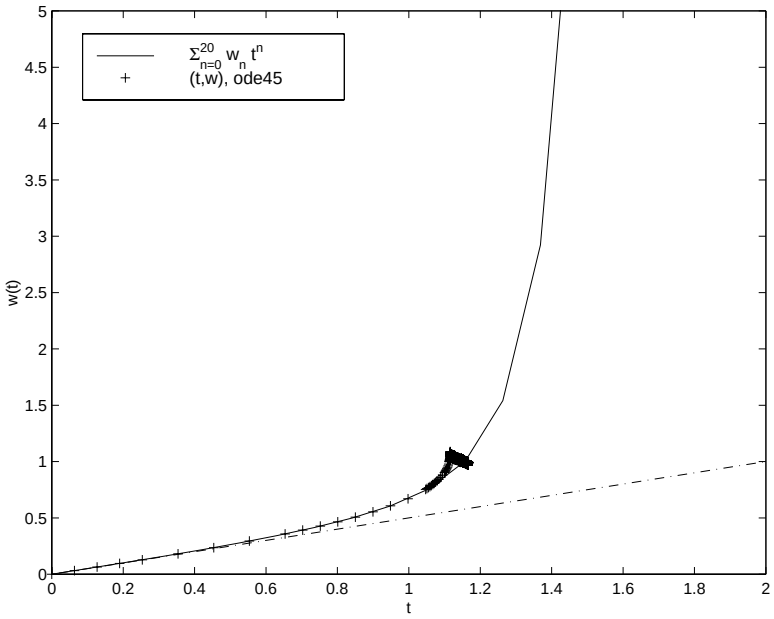


FIGURE 1. Graph of the function $w(t)$ obtained by the Matlab routine `ode45` and by the recurrence relation in Theorem 3.2.

The function $w(t)$ is the solution of the differential equation

$$\left[2\Big(2-\frac{\sinh w}{w}\Big)-\cosh t+1-\frac{1}{2}\Big((\mathrm{e}^{t}-1)\frac{\mathrm{e}^{w}-1-w}{w} - (\mathrm{e}^{-t}-1)\frac{\mathrm{e}^{-w}-1+w}{w}\Big)\right]w' \quad = \quad \cosh w$$

with initial condition $w(0)=0$. In Figure 1 we plot the approximation of the above ordinary differential equation obtained with the Matlab routine `ode45` and initial condition $w(0)=\sqrt{\epsilon}$, where $\epsilon$ is the machine epsilon (about 2.22e-16). The tolerance on relative and absolute error is set to 1.0e-4. The solid line corresponds to the approximation of $w(t)\approx\sum_{n=1}^{20}w_n t^n$ where the coefficients $w_n$, $n=1,\dots,20$, are obtained by means of the recurrence relation in Theorem 3.2. The dash-dotted line corresponds to the tangent at $t=0$ to $w(t)$, whose slope is $w_1=\frac{1}{2}$. The radius of convergence $\bar\delta$ is the distance from the first singularity, which is detected by the ODE routine between 1.1 and 1.2.

## 4. On some applications

In this section we briefly describe some applications of the results derived in this paper. First, let us consider the case when $G=\mathrm{GL}(n,\mathbb{R})$, the group of real-coefficient invertible $n\times n$ matrices, and set $\sigma(z)=z^{-T}$. Assume that $z=\exp(Z)$. Then $P=\frac{1}{2}(Z-Z^T)$ and $K=\frac{1}{2}(Z+Z^T)$, and the decomposition $z=xy$ is the classical polar decomposition of the matrix $z$, and the functions $X(t)$ and $Y(t)$ are the logarithms of the factor $x$ (a symmetric and positive definite matrix) and $y$ (an orthogonal matrix) respectively, which are completely described as analytic functions at the origin by the recurrence relations discussed in this paper. In Figure 2 we plot the norm of the truncation of $X(t)$ and $Y(t)$ to include the $\mathcal{O}\big(t^{14}\big)$-terms in the series expansions for values of $t$ between $10^{-2}$ and 10. The matrix $Z$ is a $10\times 10$ random matrix with entries between $-1$ and 1, while $\mu=2$, since an obvious bound for the commutator operator is $\|AB-BA\|\le 2\|A\|\|B\|$ when $A$ and $B$ are arbitrary matrices. We also plot the bounds on $\alpha$ for convergence as obtained in Theorems 1.1 and 3.2. The convergence of the series ceases when the norms behave like $t^{14}$, and Figure 2 seems to indicate that the bound $\alpha<\pi/\mu$ for $X(t)$ is quite sharp, and that the bound for $Y(t)$, $\alpha<\bar\delta/2\mu$, can also be improved to the same value.

In the second example we consider the application of our results to differential equations with time-reversal symmetries. Assume for instance that $\varphi_t\begin{pmatrix}p_0\\ q_0\end{pmatrix}$ is the evolution map of a Hamiltonian system with a linear time-reversal symmetry defined by the matrix

$$R=\begin{pmatrix}-1 & 0\\ 0 & 1\end{pmatrix}.$$

Then $R\varphi_t R=\varphi_{-t}$. We write $\varphi_t=\exp(t\mathcal{F})$ to indicate that $\varphi$ is the flow of the Hamiltonian vector field $\mathcal{F}$. Clearly, $\varphi_t\in G_\sigma$, where $\sigma$ is the involutive automorphism $\sigma(\varphi)=R\varphi R$, and moreover $\mathcal{F}\in\mathfrak{p}_\sigma$; namely, $R\mathcal{F}R=-\mathcal{F}$. Now, let $\varphi_h=\exp(h\mathcal{F}_h)$ be the flow of a numerical method used to approximate $\varphi_t$ with stepsize $h$. We assume that the shadow vector field $\mathcal{F}_h$ is known (it can be approximated with an exponentially small error by means of backward error analysis, see

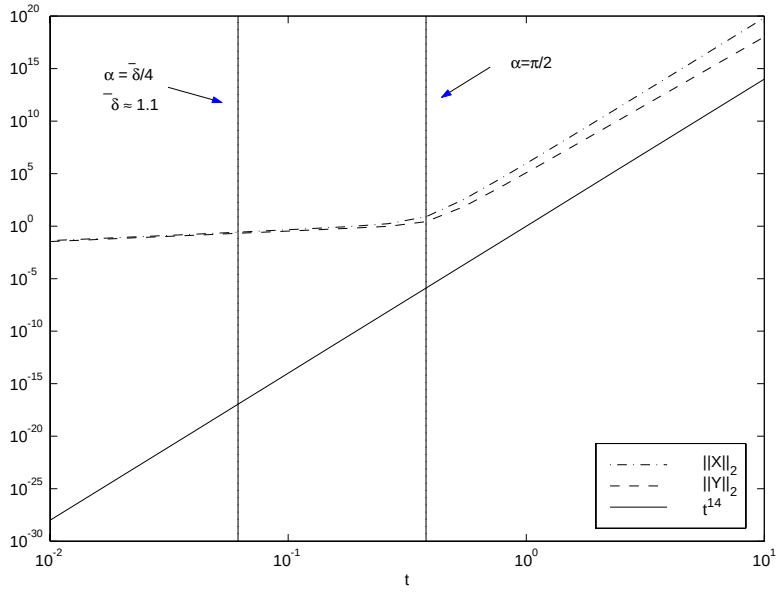


FIGURE 2. Norms of the truncated series of $X(t)$ and $Y(t)$ to order 14 for a $10 \times 10$ random matrix $Z$.

[17]). Usually, $\mathcal{F}_h \notin \mathfrak{p}_\sigma$. However, denote

$$\mathcal{P}_h = \frac{1}{2}(\mathcal{F}_h - R\mathcal{F}_h R),$$
$$\mathcal{K}_h = \frac{1}{2}(\mathcal{K}_h + R\mathcal{F}_h R).$$

Then, we can apply the generalized polar decomposition

$$\exp(h\mathcal{F}_h) = \exp(h\mathcal{X}_h)\exp(h\mathcal{Y}_h),$$

where $\mathcal{X}_h$ and $\mathcal{Y}_h$ are the vector fields corresponding to the series in Theorems 1.1 and 3.2 in terms of $\mathcal{P}_h$ and $\mathcal{K}_h$. The commutator 'ad' corresponds to the Jacobi bracket of vector fields. The vector fields $\mathcal{X}_h$ and $\mathcal{Y}_h$ can be approximated to a given order of accuracy by means of the recurrence relations discussed in this paper. Note that $\mathcal{X}_h$ possesses $R$ as a time-reversal symmetry; hence, if $\varphi_h$ is a numerical method of order $p$, then the approximation

$$\varphi_h = \exp(h\mathcal{F}_h) \approx \exp(h\mathcal{X}_h) = \tilde{\varphi}_h$$

yields a numerical method $\tilde{\varphi}_h$ that has order $p$ and $R$ as a time-reversal symmetry. In this case, though, the series $\mathcal{X}_h$ and $\mathcal{Y}_h$ might not converge, because the Jacobi bracket of vector fields need not be bounded, and one might need to truncate the infinite expansions to an appropriate index so that exponentially small estimates hold. This is not a practical method for devising new numerical methods that have $R$ as a reversing symmetry, but we believe that it might be useful on a more theoretical level.

In the third and last example, perhaps the most practical, let us denote by $\mathrm{SL}(n)$ the group of matrices with unit determinant and by $\mathfrak{sl}(n)$ the corresponding

Lie algebra of traceless matrices. We wish to describe how to approximate the exponential $\exp(tZ) \in \mathrm{SL}(n)$ of a matrix $Z \in \mathfrak{sl}(n)$, so that the approximation $F(t,Z) \in \mathrm{SL}(n)$. To this purpose, consider the involutive automorphism $\sigma(z) = SzS$ on $\mathrm{SL}(n)$, where $S$ is the diagonal matrix

$$S = \begin{pmatrix} -1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix}.$$

Then, it easily verified that $\mathrm{d}\sigma(Z) = SZS$ and

$$P = \frac{1}{2}(Z - SZS) = \left(\begin{array}{c|ccc} 0 & z_{1,2} & \cdots & z_{1,n} \\ \hline z_{2,1} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ z_{n,1} & 0 & \cdots & 0 \end{array}\right),$$

$$K = \frac{1}{2}(Z + SZS) = \left(\begin{array}{c|ccc} z_{1,1} & 0 & \cdots & 0 \\ \hline 0 & z_{2,2} & \cdots & z_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & z_{n,2} & \cdots & z_{n,n} \end{array}\right).$$

Thus, one can truncate the series $X(t)$ and $Y(t)$ discussed in this paper to a given order of accuracy $p$, and approximate

$$\exp(tZ) \approx F(t,Z) = \exp(\sum_{n=0}^{p} X_n t^n) \exp(\sum_{n=0}^{p} Y_n t^n),$$

where the $X_n$s and $Y_n$s are determined in terms of the matrices $P$ and $K$ above. Note that both $\sum_{n=0}^{p} X_n t^n$ and $\sum_{n=0}^{p} Y_n t^n$ are zero-trace matrices; hence their exact exponential is a matrix with determinant equal to one. Hence $F(t,Z)$ is guaranteed to sit in $\mathrm{SL}(n)$. The truncation of $X(t)$ has only one row and one column, and its exponential is very easy to compute exactly. Commutators of matrices in $\mathfrak{p}_\sigma$ and $\mathfrak{k}_\sigma$ can be computed employing only matrix-vector products, using $\mathcal{O}(n^2)$ operations. In [16] an iterative approach of this kind is used to devise methods for the approximation of the exponential from $\mathfrak{sl}(n)$ to $\mathrm{SL}(n)$ for large $n$, costing $\mathcal{O}(3\frac{1}{3}n^3), \mathcal{O}(7n^3), \mathcal{O}(9n^3)$ floating point operations for approximations of order 2, 3 and 4 respectively. Also, this method can be generalized to other matrix groups and is very attractive when the given matrices are in a banded/Hessenberg form [10].

## 5. Concluding remarks

In this paper we have presented a combinatorial formula that gives the solution of the implicit Lie-algebra differential equation (11) and we have discussed its convergence, in the spirit, say, of Chacon and Fomenko [3]. However, there exist at least two alternative solution techniques, both already present in the geometric-integration literature:

- the Picard iteration approach, and
- the rooted-trees approach.

The Picard iteration approach consists in approximating a nonlinear ordinary differential equation

$$y' = f(y), \qquad y(0) = y_0,$$

by a sequence of linear problems of the type

$$(y^{[m+1]})' = f(y^{[m]}(t)), \qquad m = 0, 1, 2, \ldots, \quad y^{[0]} = y_0,$$

so that, at each iteration, $y^{[m+1]}$ can be computed explicitly,

$$y^{[m+1]}(t) = \int_0^t f(y^{[m]}(\tau))\,\mathrm{d}\tau.$$

In the Lie-group context, Picard iterations were used as long ago as Magnus [12], and in more recent years by Iserles and Nørsett [9] and Blanes, Casas, Oteo and Ros [2].

The rooted-trees approach was first introduced by Iserles and Nørsett in [9] as an alternative expansion to Picard iterations for differential equations on Lie groups. In this approach, binary trees represent multiple integrals of commutator terms. Lately, this approach has been extended to the case of trees with coloured leaves [6].

The main disadvantage of the two approaches above is that they need to be applied to an explicit differential equation, while (11) is implicit. However, an alternative is to use the explicit differential equation for $Y$,

$$Y' = \mathrm{dexp}_Y^{-1}\left(K - 2\sum_{k=1}^{\infty} \frac{(2^{2k}-1)\mathrm{B}_{2k}}{(2k)!} u^{2k-1}\Big|_{u=\mathrm{ad}_X}(P)\right),$$

given in [15], in tandem with the initial condition $Y(0) = 0$. Note that, in this formulation, solving for $Y(t)$ requires the knowledge of the function $X(t)$, which is not necessary in the formulation (11): indeed, the $Y(t)$ term is independent of the function $X(t)$, but depends on the terms $P$ and $K$ that decompose $Z$. Clearly, each of the three approaches has its advantages and disadvantages. In our opinion, the approach presented in this paper is more favourable for symbolic manipulations of the expansions.

In passing, we mention that symbolic software to compute the series expansion for the functions $X(t)$ and $Y(t)$ described in this paper is now included in the latest release of the Matlab package *DiffMan*, written by K. Engø, A. Marthinsen and H. Munthe-Kaas and freely available from the home page

`http://www.ii.uib.no/diffman/`

*DiffMan* is an object-oriented MATLAB toolbox for solving differential equations on manifolds.

## References


[1] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*, Dover, New York, 1965. MR **34:**8606

[2] S. Blanes, F. Casas, J. A. Oteo, and J. Ros, *Magnus and Fer expansion for matrix differential equations: The convergence problem*, J. Phys. A **31** (1998), 259–268. MR **98m:**34019

[3] R. V. Chacon and A. T. Fomenko, *Recursion formulas for the Lie integral*, Adv. Math. **98** (1991), no. 2, 200–257. MR **93e:**22013

[4] S. Helgason, *Differential Geometry, Lie Groups and Symmetric Spaces*, Academic Press, 1978. MR **80k:**53081

[5] R. A. Horn and C. R. Johnson, *Topics in matrix analysis*, Cambridge University Press, 1991. MR **92e:**15003

[6] A. Iserles, *On the discretization of double-bracket flows*, Found. Comp. Math. **2** (2002), 305–329. MR **2003g:**37157
[7] A. Iserles, R. McLachlan, and A. Zanna, *Approximately preserving symmetries in numerical integration*, Euro. J. Appl. Math. **10** (1999), 419–445. MR **2000j:**65070
[8] A. Iserles, H. Munthe-Kaas, S. P. Nørsett, and A. Zanna, *Lie-group methods*, Acta Numerica **9** (2000), 215–365. MR **2003a:**37123
[9] A. Iserles and S. P. Nørsett, *On the solution of differential equations in Lie groups*, Phil. Trans. R. Soc. Lond. A **357** (1999), 983–1019. MR **2000d:**34022
[10] A. Iserles and A. Zanna, *Efficient computation of the matrix exponential by generalized polar decompositions*, D.A.M.T.P. Technical Report NA2003/02, University of Cambridge, U.K. To appear in SIAM J. Numer. Anal.
[11] O. Loos, *Symmetric Spaces I: General Theory*, W. A. Benjamin, Inc., 1969. MR **39:**365a
[12] W. Magnus, *On the Exponential Solution of Differential Equations for a Linear Operator*, Comm. Pure and Appl. Math. **VII** (1954), 649–673. MR **16:**790a
[13] R. I. McLachlan, G. R. W. Quispel, and G. S. Turner, *Numerical integrators that preserve symmetries and reversing symmetries*, SIAM J. Numer. Anal. **35** (1998), no. 2, 586–599. MR **99b:**65083
[14] H. Munthe-Kaas, G. R. Quispel, and A. Zanna, *Application of symmetric spaces and Lie triple systems in Numerical Analysis*, Tech. Report 217, Department of Informatics, University of Bergen, Norway, 2001.
[15] H. Munthe-Kaas, R. G. W. Quispel, and A. Zanna, *Generalized polar decompositions on Lie groups with involutive automorphisms*, Found. Comp. Math. **1** (2001), no. 3, 297–324. MR **2002g:**22014
[16] H. Munthe-Kaas and A. Zanna, *Generalized polar decompositions for the approximation of the matrix exponential*, SIAM J. Matrix Anal. Applic. **23** (2002), no. 3, 840–862. MR **2003e:**65066
[17] S. Reich, *Dynamical systems, numerical integration, and exponentially small estimates*, Habilitationsschrift, 1998.
[18] V. S. Varadarajan, *Lie Groups, Lie Algebras, and Their Representation*, GTM 102, Springer-Verlag, 1984. MR **85e:**22001

Institutt for informatikk, University of Bergen, Høyteknologisenteret, Thormøhlensgate 55, N-5020 Bergen, Norway
*E-mail address*: anto@ii.uib.no
*URL*: www.ii.uib.no/~anto