Found. Comput. Math. (2001) 1:297–324

DOI: 10.1007/s102080010012

FOUNDATIONS OF COMPUTATIONAL MATHEMATICS
The Journal of the Society for the Foundations of Computational Mathematics

Generalized Polar Decompositions on Lie Groups with Involutive Automorphisms

H. Z. Munthe-Kaas,[1] G. R. W. Quispel,[2] and A. Zanna[3]

[1]Institutt for Informatikk
University of Bergen
Høyteknologisenteret
Thormøhlensgate 55
N-5020 Bergen, Norway
hans@ii.uib.no

[2]Department of Mathematics
La Trobe University
Bundoora
Melbourne 3083, Australia
r.quispel@latrobe.edu.au

[3]Institutt for Informatikk
University of Bergen
Høyteknologisenteret
Thormøhlensgate 55
N-5020 Bergen, Norway
anto@ii.uib.no

Abstract. The polar decomposition, a well-known algorithm for decomposing real matrices as the product of a positive semidefinite matrix and an orthogonal matrix, is intimately related to involutive automorphisms of Lie groups and the subspace decomposition they induce. Such generalized polar decompositions, depending on the choice of the involutive automorphism σ, always exist near the identity although frequently they can be extended to larger portions of the underlying group.

Date received: September 12, 2000. Final version received: April 16, 2001. Communicated by Arieh Iserles. Online publication: June 11, 2001.
Key words and phrases: Lie group, Polar decomposition, Generalized polar decompositions, Involutive automorphism, Cartan decomposition, Iwasawa decomposition.
AMS classification: 22E15, 65F25, 65Q05.

In this paper, first of all we provide an alternative proof to the local existence and uniqueness result of the generalized polar decomposition. What is new in our approach is that we derive differential equations obeyed by the two factors and solve them analytically, thereby providing explicit Lie-algebra recurrence relations for the coefficients of the series expansion.

Second, we discuss additional properties of the two factors. In particular, when σ is a Cartan involution, we prove that the subgroup factor obeys similar optimality properties to the orthogonal polar factor in the classical matrix setting both locally and globally, under suitable assumptions on the Lie group G.

1. Introduction

It is well known in linear algebra that any $N \times N$ matrix A can be decomposed into the product

$$A = HU, \tag{1.1}$$

where U and H are two $N \times N$ matrices, the first unitary and the second Hermitian positive semidefinite [**10**]. Furthermore, if A is invertible, then H is positive definite. The decomposition (1.1) is called *polar decomposition* and was introduced in 1902 by Autonne [**3**] as a matrix analog of the polar form of a complex number

$$z = re^{i\theta}, \qquad r \geq 0, \quad 0 \leq \theta < \pi.$$

The popularity of the polar decomposition is mainly due to the best approximation properties of its factors. It is proved in [**6**] that

$$\min\{\|A - Q\|\colon\ Q^*Q = I\} = \|A - U\|,$$

where $\|\cdot\|$ is any unitary invariant norm, a property saying that U is the best unitary (orthogonal in the real case) approximant to A in any unitary invariant norm. Optimality results for the factor H are discussed in [**8**].

It is well known that when A is real, the matrix U is orthogonal and H is symmetric. In the remaining part of this section, we shall restrict ourselves to the case when A is real and invertible, $A \in G \subset \mathrm{GL}(\mathbb{R}, N)$, hence H is positive definite. We recall that

$$AA^T = HUU^TH^T = H^2,$$

from which it follows that H is the (unique) positive definite square root of the matrix AA^T and, consequently,

$$U = H^{-1}A = (AA^T)^{-1/2}A.$$

In a recent investigation of symmetric spaces and their connection with numerical analysis [**15**], the authors observed that a number of techniques in numerical analysis can be related to involutive automorphisms (defining subgroups of a given group) and symmetric spaces, the polar decomposition being one of such techniques, as we shall see in this paper. In other words, the polar decomposition is equivalent to decomposing a group element in the product of a term in a symmetric subspace and a term in a subgroup of the given Lie group.

There exist a number of papers on the polar decomposition in Lie groups and its generalization to semigroups (*Ol'shanskii decomposition*), many of them rather recent. A proof of the existence and uniqueness of the polar decomposition in a large portion of a Lie group G can be found, for instance, in [**12**].

In this paper, we present an alternative proof to the local existence and uniqueness of the polar decomposition in a Lie group G. We derive differential equations obeyed by the two factors and solve them analytically, thereby obtaining Lie-algebra recurrence relations for the coefficients of the series expansion of each factor. We show that the subgroup factor is expanded in odd powers of time only, a result of interest in the context of numerical analysis applications, such as the development of numerical integrators for ODEs, and finally prove optimality results for semisimple Lie groups with right-invariant metrics inherited from the Killing–Cartan form.

The paper is organized as follows. In Section 2 we introduce background theory on symmetric spaces and Lie triple systems. The main results of this paper are presented in Sections 3–5. The differential equations for the generalized polar factors and the resulting recurrence relations are introduced in Section 3, where we also show that the subgroup factor expands in odd powers of time only. In Section 4 we prove local optimality properties of the subgroup factor as an approximant to the original Lie-group element. Section 5 is devoted to the extension of the optimality result to a global result, under suitable conditions on the Lie group G. This section is based on notes kindly provided to us by E. van den Ban, although any errors are of course ours. In Section 6 we discuss various applications to computations and, finally, Section 7 is devoted to conclusions.

2. Background Theory

Let G be a connected Lie group and $\sigma\colon\ G \to G$ an involutive automorphism, i.e., $\sigma \neq \mathrm{id}$ and $\sigma^2 = \mathrm{id}$. Let G^σ denote the set of fixed points of σ:

$$G^\sigma = \{x \in G\colon\ \sigma(x) = x\},$$

and let G_σ denote the set of anti-fixed points of σ:

$$G_\sigma = \{x \in G\colon\ \sigma(x) = x^{-1}\}.$$

The set G^σ is a subgroup of G and may be disconnected, so that G^σ_e denotes its connected component including the identity. The set G_σ does not have a group structure, but is a *symmetric space* when endowed with the nonassociative multiplication $x \cdot y = xy^{-1}x$. We recall that a symmetric space is a manifold endowed with a differentiable multiplication $\cdot$ obeying the following conditions:

(i) $x \cdot x = x$;
(ii) $x \cdot (x \cdot y) = y$;
(iii) $x \cdot (y \cdot z) = (x \cdot y) \cdot (x \cdot z)$;

and, moreover,

(iv) every x has a neighborhood U such that $x \cdot y = y$ implies $y = x$ for all y in U.

The involutive automorphism σ can be lifted to the Lie algebra $\mathfrak{g}$ of G and this lift will be denoted by $d\sigma$. Let $X \in \mathfrak{g}$ and consider $x = \exp(tX)$. Then

$$d\sigma(X) := \left.\frac{d}{dt}\right|_{t=0} \sigma(\exp(tX)), \qquad \forall\, X \in \mathfrak{g}. \tag{2.1}$$

Consider the spaces

$$\mathfrak{k} = \{X \in \mathfrak{g}\colon\ d\sigma(X) = X\}$$

of fixed points of $d\sigma$, and

$$\mathfrak{p} = \{X \in \mathfrak{g}\colon\ d\sigma(X) = -X\}$$

of anti-fixed points of $d\sigma$. The space $\mathfrak{k}$ is a subalgebra of $\mathfrak{g}$, while $\mathfrak{p}$ is a Lie triple system, namely, a vector space that is closed under the double commutator ad^2. One has

$$\mathfrak{g} = \mathfrak{p} \oplus \mathfrak{k}, \tag{2.2}$$

(direct sum), thus every Lie-algebra element X can be uniquely written as two components, one being fixed under σ and the other being anti-fixed. This is called a generalized Cartan decomposition. The projection in each subspace is given by the formula

$$X = \tfrac{1}{2}(X + d\sigma(X)) + \tfrac{1}{2}(X - d\sigma(X)), \tag{2.3}$$

where $X + d\sigma(X) \in \mathfrak{k}$ and $X - d\sigma(X) \in \mathfrak{p}$. Note also that if $K \in \mathfrak{k}$, then $\exp(tK) \in G^\sigma$. By a similar token, $P \in \mathfrak{p}$ implies that $\exp(tP) \in G_\sigma$.

As an example let $\mathrm{GL}(N)$ be the group of $N \times N$ invertible real matrices. Consider the map

$$\sigma(x) = x^{-T}, \qquad x \in \mathrm{GL}(N). \tag{2.4}$$

It is clear that σ is an involutive automorphism of $\mathrm{GL}(N)$. Then, from above, the set $G_\sigma = \{x \in \mathrm{GL}(N)\colon\ \sigma(x) = x^{-1}\}$ is a symmetric space. The set G_σ is the set of invertible symmetric matrices. The symmetric space G_σ is disconnected. Its connected component containing the identity matrix I is the set of symmetric positive definite matrices. Similarly, G^σ is the set of orthogonal matrices and is a subgroup of $\mathrm{GL}(N)$ and G^σ_e corresponds to orthogonal matrices with unit determinant (the Lie group $\mathrm{SO}(N)$).

We compute $d\sigma$ making use of (2.1). Given $X \in \mathfrak{gl}(N)$:

$$d\sigma(X) = \left.\frac{d}{dt}\right|_{t=0} \sigma(\exp(tX)) = \left.\frac{d}{dt}\right|_{t=0} (I + tX + \mathcal{O}(t^2))^{-T} \tag{2.5}$$

$$= \left.\frac{d}{dt}\right|_{t=0} (I + tX^T + \mathcal{O}(t^2))^{-1} = \left.\frac{d}{dt}\right|_{t=0} (I - tX^T + \mathcal{O}(t^2))$$

$$= -X^T, \tag{2.6}$$

hence we deduce that

$$\mathfrak{k} = \{X \in \mathfrak{gl}(N)\colon\ d\sigma(X) = X\} = \mathfrak{so}(N),$$

the classical algebra of skew-symmetric matrices, while

$$\mathfrak{p} = \{X \in \mathfrak{gl}(N)\colon\ d\sigma(X) = -X\}$$

is the classical set of symmetric matrices. This set is not a subalgebra of $\mathfrak{gl}(N)$ (the commutator of two symmetric matrices is not a symmetric matrix) but is closed under ad^2 (the double commutator of symmetric matrices is a symmetric matrix) and is a Lie triple system.

The decomposition (2.3) is nothing other than the canonical decomposition of a matrix into its skew-symmetric and symmetric part

$$X = \tfrac{1}{2}(X + d\sigma(X)) + \tfrac{1}{2}(X - d\sigma(X)) = \tfrac{1}{2}(X - X^{\mathrm{T}}) + \tfrac{1}{2}(X + X^{\mathrm{T}}).$$

However, as mentioned in the Introduction, the above procedure is very general. If G is a subgroup of the matrix group GL, we may choose a matrix S such that $S^2 = I$ and define an automorphism on G as $\sigma(g) = SgS$, which leads to $d\sigma(X) = SXS$. This type of “inner” automorphisms on GL is discussed in Section 6.

As yet another example one can choose $G = \mathrm{Diff}(M)$,[1] the group of diffeomorphisms of a manifold M, and set $\sigma(\varphi) = \mathcal{R}\varphi\mathcal{R}^{-1}$, with $\mathcal{R}$ an involutive diffeomorphism of M onto M. If F denotes a vector field, so that $\varphi = \exp(tF)$, then the sets G_σ (resp., G^σ) correspond precisely to the vector fields that possess $\mathcal{R}$ as a reversing symmetry (irony of a sort, this is a symmetric space!) (resp. $\mathcal{R}$ as a symmetry). The existence of the polar decomposition in a general Lie-group context implies that, given $\mathcal{R}$, every diffeomorphism[2] can be written as the composition of two flows, one possessing $\mathcal{R}$ as a symmetry and the other possessing

[1] The group of diffeomorphisms of a manifold does not have the structure of a Lie group, and the exponential map is not onto, even in very small neighborhoods of the identity map [**2**],[**5**]. However, our procedure here is intended to be formal and we assume that it is possible to verify the convergence of the formulas presented here by other means. Under these premises, we assume that $\mathrm{Diff}(M)$ is essentially a Lie group.

[2] See footnote 1.

$\mathcal{R}$ as a reversing symmetry. Such a decomposition has fundamental implications in the numerical analysis of differential equations and numerical integration of systems with symmetries and reversing symmetries, an issue that has long been under the spotlight of researchers in the field of numerical analysis and dynamical systems (see, for instance, [**14**] and references therein). An application to vector fields with polynomial coefficients is discussed in Section 6.

For those who are interested in further reading on symmetric spaces and Lie triple systems, we refer to [**7**] and [**13**].

3. Generalized Polar Decomposition in Lie Groups

Given a generic Lie group G, we wish to write $z \in G$ as $z = xy$, where $x \in G_\sigma$ and $y \in G^\sigma$. We call the decomposition $z = xy$ a *generalized polar decomposition* of z, in analogy to the terminology of linear algebra.

Theorem 3.1. *Let $z = \exp(tZ) \in G$, where $Z = K + P$, $d\sigma(K) = K$, and $d\sigma(P) = -P$, is the decomposition of Z in $\mathfrak{k} \oplus \mathfrak{p}$. Then, for sufficiently small t, z admits a differentiable generalized polar decomposition $z = xy$ where $x = \exp(X(t)) \in G_\sigma$, with $X(t) \in \mathfrak{p}$ and $y = \exp(Y(t)) \in G^\sigma$, where $Y(t) \in \mathfrak{k}$. Moreover, such a decomposition is locally unique.*

Proof. Set

$$P = \tfrac{1}{2}(Z - d\sigma(Z)), \qquad K = \tfrac{1}{2}(Z + d\sigma(Z)),$$

so that $Z = K + P$ and $d\sigma(P) = -P$ and $d\sigma(K) = K$. Let

$$\begin{aligned} X(t) &= tX_1 + t^2X_2 + t^3X_3 + \cdots, \\ Y(t) &= tY_1 + t^2Y_2 + t^3Y_3 + \cdots, \end{aligned}$$

where the X_i's are in $\mathfrak{p}$ and the Y_i's are in $\mathfrak{k}$. Imposing

$$\exp(tZ) = \exp(X(t)) \exp(Y(t))$$

and making use of the BCH formula, one could derive the following formal conditions:

$$\begin{aligned} X_1 &= P, & Y_1 &= K, \\ X_2 &= -\tfrac{1}{2}[P, K], & Y_2 &= 0, \\ X_3 &= -\tfrac{1}{6}[K, [P, K]], & Y_3 &= -\tfrac{1}{12}[P, [P, K]], \end{aligned}$$

etc. As we shall see later, all the X_i's and Y_i's can be algorithmically calculated and are uniquely determined in terms of P and K.

In what follows, we derive the Cauchy problem obeyed by X, find its solution as a series expansion, and prove that, in case G is finite dimensional (with more generality, when the adjoint operator ad is bounded), such a series converges for t sufficiently close to zero. The convergence of $Y(t)$ will follow from that of X and by the BCH formula.

Differentiating $\exp(tZ) = \exp(X)\exp(Y)$ and multiplying by $\exp(-tZ)$ on the right, we derive

$$d\exp_X X' = Z - \mathrm{Ad}_{\exp(X)}\, d\exp_Y Y'.$$

We apply to both sides the operator $\mathrm{Ad}_{\exp(-X)}$, resulting in

$$\mathrm{Ad}_{\exp(-X)}\; d\exp_X X' = \mathrm{Ad}_{\exp(-X)}\; Z - d\exp_Y Y'.$$

Recall that $\mathrm{Ad}_{\exp(V)} = \exp(\mathrm{ad}_V)$, hence the equality

$$\left.\frac{e^{-u}-1}{-u}\right|_{u=\mathrm{ad}_X} X' = \exp(\mathrm{ad}_{-X})(K+P) - d\exp_Y Y'.$$

Our goal is to decompose the above expression in $\mathfrak{k}\oplus\mathfrak{p}$. To this end, observe that $d\exp_Y Y' \in \mathfrak{k}$ since $Y, Y' \in \mathfrak{k}$ and $\mathfrak{k}$ is a subalgebra of $\mathfrak{g}$. Next, we analyze the term $\exp(\mathrm{ad}_{-X})(K+P)$. Recall that

$$\exp(\mathrm{ad}_{-X})(K+P) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!}\,\mathrm{ad}_X^k(K+P),$$

hence the term $\mathrm{ad}_X^k(K)$ is in $\mathfrak{k}$ for even k while it is in $\mathfrak{p}$ for odd values of k. Conversely, we have $\mathrm{ad}_X^{2m+1}(P) \in \mathfrak{k}$, and $\mathrm{ad}_X^{2m}(P) \in \mathfrak{p}$ for $m = 0, 1, 2, \ldots$. In summary,

$$\begin{aligned}\exp(\mathrm{ad}_{-X})(K+P) &= \underbrace{(-\sinh u(K) + \cosh u(P))}_{\in\mathfrak{p}} + \underbrace{(\cosh u(K) - \sinh u(P))}_{\in\mathfrak{k}},\\ u &= \mathrm{ad}_X\,.\end{aligned}$$

A similar procedure applies to the term $e^{-u}-1/-u|_{u=\mathrm{ad}_X} X'$: since $X, X' \in \mathfrak{p}$:,

$$\begin{aligned}\left.\frac{e^{-u}-1}{-u}\right|_{u=\mathrm{ad}_X} X' &= \underbrace{\frac{1}{2}\left.\left(\frac{e^{-u}-1}{-u} + \frac{e^{u}-1}{u}\right)\right|_{u=\mathrm{ad}_X} X'}_{\in\mathfrak{p}}\\ &\quad + \underbrace{\frac{1}{2}\left.\left(\frac{e^{-u}-1}{-u} - \frac{e^{u}-1}{u}\right)\right|_{u=\mathrm{ad}_X} X'}_{\in\mathfrak{k}}\\ &= \left.\frac{1}{u}\sinh u\right|_{u=\mathrm{ad}_X} X' - \left.\frac{1}{u}(\cosh u - 1)\right|_{u=\mathrm{ad}_X} X'.\end{aligned}$$

Now, since X evolves in $\mathfrak{p}$, it must depend only on terms that are in $\mathfrak{p}$. As a consequence,

$$\left.\frac{1}{u}\sinh u\right|_{u=\mathrm{ad}_X} X' = -\sinh u\,\Big|_{u=\mathrm{ad}_X}(K) + \cosh u\,\Big|_{u=\mathrm{ad}_X}(P).$$

Inverting the operator on the left-hand side, we deduce that X obeys the differential equation

$$\begin{aligned} X' &= -[X, K] + u \left. \frac{\cosh u}{\sinh u} \right|_{u=\mathrm{ad}_X} (P), \\ X(0) &= 0. \end{aligned} \tag{3.1}$$

Note that

$$u \frac{\cosh u}{\sinh u} = 1 + \sum_{k=1}^{\infty} c_{2k} u^{2k}, \qquad |u| < \pi,$$

is the series expansion of the function $u \coth(u)$, with coefficients

$$c_{2k} = \frac{2^{2k} B_{2k}}{(2k)!}, \qquad k = 1, 2, \ldots,$$

B_k being the kth Bernoulli number [**1**].

Equation (3.1), in tandem with the series expansion $X(t) = \sum_{k=1}^{\infty} t^k X_k$, implies that the terms X_k obey the recurrence relation

$$\begin{aligned} (k+1) X_{k+1} &= -[X_k, K] + \sum_{\substack{\ell \geq 1 \\ 2\ell \leq k}} c_{2\ell} \sum_{\substack{\ell_1, \ldots, \ell_{2\ell} > 0 \\ \ell_1 + \cdots + \ell_{2\ell} = k}} [X_{\ell_1}, [X_{\ell_2}, \ldots, [X_{\ell_{2\ell}}, P]]], \\ k &= 1, 2, \ldots, \\ X_1 &= P, \end{aligned} \tag{3.2}$$

as can be easily verified by comparison of powers of t.

We do not report here the proof of convergence of $X(t)$, since the existence of such a decomposition is a well-established result. However, for completeness, the convergence of $X(t)$ is proved in the Appendix. □

For completeness' sake, we derive a differential equation obeyed by $Y(t)$. Matching terms in $\mathfrak{k}$, we obtain

$$-\frac{1}{u} (\cosh u - 1) X' = \cosh u (K) - \sinh u (P) - d \exp_Y Y',$$

where $u = \mathrm{ad}_X$, which, after some simple algebra, reduces to

$$d \exp_Y Y' = K + (\operatorname{csch} u - \coth u)(P), \qquad u = \mathrm{ad}_X .$$

Using the series expansion of $\operatorname{csch}(u)$ and $\coth(u)$, and inverting the $d \exp_Y$ operator, we find

$$Y' = d \exp_Y^{-1} \left(K - 2 \sum_{k=1}^{\infty} \left. \frac{(2^{2k} - 1) B_{2k}}{(2k)!} u^{2k-1} \right|_{u=\mathrm{ad}_X} (P) \right),$$

in tandem with the initial condition $Y(0) = 0$. Note that, in this formulation, solving for $Y(t)$ requires the knowledge of the function $X(t)$. In [**19**] the recurrence relation for $Y(t)$ is derived following a different approach, solving an implicit differential equation for Y that does not require the direct knowledge of X but only that of Z and $-d\sigma(Z) = P - K$. More specifically,

$$\begin{aligned}
Y_1 &= K, \\
Y_{2n} &= O, \qquad n = 0, 1, 2, \ldots, \qquad (3.3)\\
2(2n+1)Y_{2n+1} &= -2\sum_{q=1}^{n}\sum_{\substack{k\geq 1\\ k\leq q}}\frac{1}{(2k+1)!}\sum_{\substack{k_1,\ldots,k_{2k}>0\\ k_1+\cdots+k_{2k}=2q}} \\
&\quad \times [Y_{k_1}, \ldots, [Y_{k_{2k}}, Y_{2(n-q)+1}], \ldots] \\
&\quad - \sum_{m=1}^{n}\frac{2(n-m)+1}{(2m)!}\,\mathrm{ad}_Z^{2m}\,Y_{2(n-m)+1} \\
&\quad - \sum_{q=0}^{2(n-1)}\sum_{j=0}^{2(n-1)-q}\frac{(-1)^{2n-q-j-1}(j+1)}{(2n-q-j-1)!}\,\mathrm{ad}_Z^{2n-j-q-1} \\
&\quad \times \sum_{\substack{k\geq 1\\ k\leq q+1}}\frac{1}{(k+1)!}\sum_{\substack{j_1,\ldots,j_k>0\\ j_1+\cdots+j_k=q+1}}[Y_{j_1}, \ldots, [Y_{j_k}, Y_{j+1}]\ldots] \\
&\quad - \sum_{\substack{\ell\geq 1\\ \ell\leq n}}\frac{1}{(2\ell)!}\sum_{\substack{\ell_1,\ldots,\ell_{2\ell}>0\\ \ell_1+\cdots+\ell_{2l}=2n}}[Y_{\ell_1}, \ldots, [Y_{\ell_{2\ell}}, P-K], \ldots]. \qquad (3.4)
\end{aligned}$$

Let $z = xy$ be the differentiable generalized polar decomposition of z defined by σ. Note that

$$\sigma(x) = \sigma\exp(X(t)) = \exp(d\sigma(X(t))) = \exp(-X(t)) = x^{-1},$$

and, by a similar token,

$$\sigma(y) = y.$$

It follows that

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

hence

$$x = (z\sigma(z)^{-1})^{1/2}.$$

In particular, setting $Z = P + K$, one has

$$x = \exp(X(t)), \qquad X(t) = \tfrac{1}{2}bch(t(P+K), t(P-K)), \qquad (3.5)$$

where $bch(\cdot,\cdot)$ is the operator of the BCH formula, so that

$$\exp(V)\exp(W) = \exp(bch(V, W)),$$

for all $V, W \in \mathfrak{g}$ of sufficiently small norm (see [**18**]). Furthermore,

$$y = x^{-1}z = (z\sigma(z)^{-1})^{-1/2}z. \tag{3.6}$$

Equations (3.5) and (3.6) are derived in [**12**] as the polar factors in the polar decomposition of z. Lawson's method of proof is based on generalizing an application of the Bony–Brezis theorem [**9**].

The first terms (up to t^6) in the expansion of X and Y are displayed below:

$$\begin{aligned}
X \;=\; & Pt - \tfrac{1}{2}[P,K]t^2 - \tfrac{1}{6}[K,[P,K]]t^3 \\
& + (\tfrac{1}{24}[P,[P,[P,K]]] - \tfrac{1}{24}[K,[K,[P,K]]])t^4 \\
& + (\tfrac{7}{360}[K,[P,[P,[P,K]]]] - \tfrac{1}{120}[K,[K,[K,[P,K]]]] \\
& \quad - \tfrac{1}{180}[[P,K],[P,[P,K]]])t^5 \\
& + (-\tfrac{1}{240}[P,[P,[P,[P,[P,K]]]]] + \tfrac{1}{180}[K,[K,[P,[P,[P,K]]]]] \\
& \quad - \tfrac{1}{720}[K,[K,[K,[K,[P,K]]]]] + \tfrac{1}{720}[[P,K],[K,[P,[P,K]]]] \\
& \quad + \tfrac{1}{180}[[P,[P,K]],[K,[P,K]]])t^6 + \mathcal{O}(t^7), \\
Y \;=\; & Kt - \tfrac{1}{12}[P,[P,K]]t^3 + (\tfrac{1}{120}[P,[P,[P,[P,K]]]] \\
& + \tfrac{1}{720}[K,[K,[P,[P,K]]]] - \tfrac{1}{240}[[P,K],[K,[P,K]]])t^5 + \mathcal{O}(t^7).
\end{aligned}$$

Proposition 3.2. *Let xy be the differentiable generalized polar decomposition of $z = \exp(tZ)$, where $Z = P + K$ as in Theorem* 3.1. *The function $Y(t)$, such that $y = \exp(Y(t))$, is an odd function of t.*

Proof. Let $\exp(tZ) = \exp(X(t))\exp(Y(t))$ be the generalized polar decomposition of z according to Theorem 3.1. Taking the inverse on both sides, we find $\exp(-tZ) = \exp(-Y(t))\exp(-X(t))$, a term that we write as

$$\exp(-tZ) = \exp(-Y(t))\exp(-X(t))\exp(Y(t))\exp(-Y(t)).$$

Clearly, $\exp(-Y(t)) \in G^\sigma$. Set

$$\tilde{x} = \exp(-Y(t))\exp(-X(t))(\exp(Y(t)) = y^{-1}x^{-1}y.$$

Since $\sigma(\tilde{x}) = \sigma(y^{-1})\sigma(x^{-1})\sigma(y) = y^{-1}xy = \tilde{x}^{-1}$, we deduce that $\tilde{x} \in G_\sigma$, hence $z^{-1} = \tilde{x}y^{-1}$ is the generalized polar decomposition of z^{-1}. On the other hand,

$$\exp(-tZ) = \exp(X(-t))\exp(Y(-t)),$$

and because of the uniqueness of the generalized polar decomposition, we conclude that $\exp(Y(-t)) = \exp(-Y(t))$, from which the assertion follows by taking the logarithm of both sides. □

Corollary 3.2.1. *Given* $z = xy$ *as in Theorem* 3.1, *then we also have*

$$z = \tilde{y}\tilde{x},$$

with $\tilde{y} \in G^\sigma$ *and* $\tilde{x} \in G_\sigma$. *Moreover,* $\tilde{y} = y$ *and* $\tilde{x} = \exp(-X(-t))$.

Proof. As above, one has $\exp(-tZ) = \exp(-Y(t))\exp(-X(t))$. Next, we make use of Proposition 3.2, hence $\exp(-Y(t)) = \exp(Y(-t))$. Replacing t with $-t$, we find $z = \exp(tZ) = \exp(Y(t))\exp(-X(-t)) = y\tilde{x}$, where $\tilde{x} = \exp(-X(-t))$, which concludes the proof. □

4. Local Optimality Results

We have mentioned that the orthogonal factor in the polar decomposition of matrices has certain optimality properties, namely it is the best orthogonal approximant to a given matrix in any unitary invariant norm. In this section we shall see that, as well as for the polar decomposition in a semisimple Lie group a similar optimality result holds, provided σ defines a *Cartan decomposition*. To do so, we need to introduce a distance on the Lie group.

Let G be a Lie group. We say that a distance function on G (obeying the standard metric axioms, i.e., positivity, symmetry, and triangle inequality), $d(\cdot,\cdot)\colon G\times G \to \mathbb{R}^+$ is *right* (resp., *left*) *invariant* if

$$d(xg, yg) = d(x, y) \qquad \forall\, g \in G,$$

(resp., $d(gx, gy) = d(x, y)$). Before proceeding further, let us review some basic facts about invariant norms on Lie groups.

Any inner product $\langle\cdot,\cdot\rangle$ on $\mathfrak{g}$ induces a left (resp., right) invariant Riemannian metric on G. By right trivializing tangents to G in the usual manner, $T_gG = \{Xg \mid X \in \mathfrak{g}\}$, we obtain

$$\langle Xg, Yg\rangle = \langle X, Y\rangle.$$

The Riemannian length of a curve $\gamma(t) \in G$ between $t = 0$ and $t = 1$ is given as

$$\operatorname{length}(\gamma) = \int_0^1 \langle\dot\gamma(t), \dot\gamma(t)\rangle^{1/2} dt.$$

The shortest curve (minimizing geodesic) between two sufficiently close points x and y is given by

$$\gamma(t) = \exp(tZ)x,$$

where $\exp(Z) = yx^{-1}$.

The right invariant metric on G is now defined as

$$d(x, y) = \min_{\gamma(0)=x,\gamma(1)=y} \operatorname{length}(\gamma) = \langle Z, Z\rangle^{1/2}.$$

Let us now introduce a canonical inner product $\langle \cdot, \cdot \rangle$ on $\mathfrak{g}$: Recall that the *Cartan–Killing form* on $\mathfrak{g}$:

$$B(X, Y) = \mathrm{tr}(\mathrm{ad}_X\ \mathrm{ad}_Y) \qquad \forall\ X, Y \in \mathfrak{g},$$

is symmetric and bilinear, moreover, provided that $\mathfrak{g}$ is semisimple, it is also nondegenerate.

Definition 4.1. Let $\mathfrak{g}$ be a semisimple Lie algebra and $d\sigma$ an involutive automorphism on $\mathfrak{g}$ as above, so that $\mathfrak{g} = \mathfrak{p} \oplus \mathfrak{k}$. This decomposition is called a *Cartan decomposition* if the symmetric bilinear form

$$B_{d\sigma}(X, Y) = -B(X, d\sigma(Y)) \qquad \forall\ X, Y \in \mathfrak{g},$$

is positive definite on $\mathfrak{g}$.

The Cartan decomposition is unique up to an inner automorphism [**7**].

Lemma 4.1. *If $\mathfrak{g} = \mathfrak{p} \oplus \mathfrak{k}$ is a Cartan decomposition, then the $\mathfrak{p}$ and $\mathfrak{k}$ are orthogonal with respect to the inner product*

$$\langle X, Y \rangle = B_{d\sigma}(X, Y) \qquad \forall\ X, Y \in \mathfrak{g}. \tag{4.1}$$

Proof. The orthogonality of the two subspaces is immediate: let $X \in \mathfrak{p}$ and $Y \in \mathfrak{k}$. Then

$$\begin{aligned} -\langle X, Y \rangle &= -B_{d\sigma}(X, Y) = B(X, d\sigma(Y)) \\ &= B(X, Y) = B(-d\sigma(X), Y) = -B(d\sigma(X), Y) = B_{d\sigma}(X, Y) \\ &= \langle X, Y \rangle, \end{aligned}$$

from which it follows that $\langle X, Y \rangle = 0$. □

Lemma 4.2. *With respect to the positive definite bilinear form* (4.1), *each* ad_X, $X \in \mathfrak{p}$, *is symmetric and each* ad_Y, $Y \in \mathfrak{k}$, *is skew-symmetric. Moreover, $B_{d\sigma}$ is* $\mathrm{Ad}(k)$ *invariant, $k \in G^\sigma$.*

Proof. See [**7,** Lemma 1.2, p. 253]. □

Let $y \in G$. We say that y *lies* in the *normal neighborhood* of the identity e if there exists $Y \in \mathfrak{g}$ such that the curve $\gamma(t) = \exp(tY)$ is the minimizing geodesic connecting e and y (namely, $\gamma(0) = e$, $\gamma(1) = y$, and γ is the curve of minimal length connecting e and y) [**17**]. Note that if y is in such a domain, then $d(y, e) = \|Y\| = (\langle Y, Y \rangle)^{1/2}$. Moreover, if $x \in G$ and $\gamma(t) = \exp(tV)x$ is the geodesic connecting x and y, then $d(x, y) = \|V\|$.

Lemma 4.3. *Let G be a semisimple Lie group. Let $x \in G_\sigma$ and $y \in G^\sigma$ be sufficiently close to the identity* (*so that the generalized polar decomposition of Theorem* 3.1 *exists*). *Moreover, assume that x and xy^{-1} are in the normal neighborhood of the identity, otherwise arbitrary. Then*

$$d(x, e) \leq d(x, y)$$

in the right invariant metric induced by (4.1).

Proof. Since x is in the normal neighborhood, there exists X in $\mathfrak{g}$ such that $x = \exp(X)$ and, moreover, $d(x, e) = \|X\|$. Set $z = xy^{-1}$ and $Z = \log z$, so that $z = \exp(Z)$. Then, from above it follows that $d(x, y) = \|Z\|$. To prove the lemma is thus sufficient to prove that $\|X\| \leq \|Z\|$.

Let $Z = P + K$ be the Cartan decomposition of Z in $\mathfrak{p} \oplus \mathfrak{k}$ and set $z(t) = \exp(tZ)$. By virtue of Theorem 3.1 we can perform the generalized polar decomposition of $z(t)$ and this decomposition is unique when z is sufficiently close to the identity. Hence there exists $x(t) \in G_\sigma$ and $w(t) \in G^\sigma$ such that $z(t) = x(t)w(t)$. Moreover, $x(t) = \exp(X(t))$, where $X(t) \in \mathfrak{p}$ obeys the differential equation (3.1). However, since $z(1) = z = xy^{-1}$, with $x \in G_\sigma$ and $y^{-1} \in G^\sigma$, it is true that $x(1) = x$ and $w(1) = y^{-1}$. In particular, $\exp X(1) = \exp X$ and $\|X\| = \|X(1)\|$.

Note that

$$\frac{d}{dt}\|X(t)\|^2 = \frac{d}{dt}\langle X, X\rangle = 2\langle X', X\rangle.$$

Making use of (3.1) and recalling that, by virtue of Lemma 4.2, $\langle \mathrm{ad}_X W, Z\rangle = \langle W, \mathrm{ad}_X Z\rangle$ for $X \in \mathfrak{p}$, $W, Z \in \mathfrak{g}$, we deduce that

$$\begin{aligned}\langle X', X\rangle &= \langle P, X\rangle - \langle \mathrm{ad}_X K, X\rangle + \sum_{k=1}^{\infty} c_{2k}\langle \mathrm{ad}_X^{2k} P, X\rangle \\ &= \langle P, X\rangle - \langle K, \mathrm{ad}_X X\rangle + \sum_{k=1}^{\infty} c_{2k}\langle P, \mathrm{ad}_X^{2k} X\rangle \\ &= \langle P, X\rangle \leq \|P\|\|X\|\end{aligned}$$

holds for all t. On the other hand,

$$\frac{d}{dt}\|X(t)\|^2 = 2\|X(t)\|\|X(t)\|',$$

hence

$$\|X(t)\|' \leq \|P\|, \qquad \|X(0)\| = 0,$$

from which we deduce

$$\|X(t)\| \leq \|P\|t,$$

thus

$$\|X\| = \|X(1)\| \leq \|P\|.$$

Furthermore, $\|Z\| = \|P + K\| = (\langle P + K, P + K\rangle)^{1/2} = (\langle P, P\rangle + 2\langle P, K\rangle + \langle K, K\rangle)^{1/2} = (\|P\|^2 + \|K\|^2)^{1/2}$ because of Lemma 4.1, since P and K belong to orthogonal subspaces. Hence

$$d(x, e) = \|X\| \leq \|P\| \leq (\|P\|^2 + \|K\|^2)^{1/2} = \|Z\| = d(x, y),$$

which completes our proof. □

We are ready to present the main result of this section.

In what follows, we assume, without further ado, that all Lie-group elements are in the normal neighborhood of the identity.

Theorem 4.4. *Let $\mathfrak{g}$ be a semisimple Lie algebra and let $d(\cdot, \cdot)$ be the G right-invariant metric induced by the symmetric bilinear form*

$$B_{d\sigma}(X, Y) = -B(X, d\sigma(Y)), \qquad \forall\, X, Y \in \mathfrak{g}.$$

In this norm, $y = \exp(Y(t))$ of Theorem 3.1 *is a differentiable best approximant to $\exp(tZ)$ in the subgroup G^σ in the domain of convergence of the generalized polar decomposition of Theorem* 3.1.

Proof. Let $\tilde{y}$ be any differentiable element in G^σ other than y, such that $\tilde{y}(0) = e$. Now, for the G right-invariant metric it is true that

$$\begin{aligned} d(z, y) &= d(xy, y) = d(x, e), \\ d(z, \tilde{y}) &= d(xy, \tilde{y}) = d(x, \tilde{y}y^{-1}). \end{aligned}$$

Since $\tilde{y}y^{-1} = w \in G^\sigma$, the assertion follows directly from Lemma 4.3. □

5. Global Optimality Results via the Iwasawa (*QR*) Decomposition

In the last section we have shown local optimality results for the subgroup factor in the polar decomposition. There, we have made explicit use of the differential equation for the function $X(t)$ derived in Theorem 3.1 and of the fact that, on Lie groups, the geodesic exponential map Exp equals the usual exponential map exp whenever we consider elements close enough to the identity (with the choice of the "standard connection" derived by the choice of the metric induced by the Cartan–Killing form). In general, even in the case when G is connected, the standard exponential map exp may fail to be surjective on the whole group G (a typical example is the group $G = \mathrm{SL}(2, \mathbb{R})$), hence Exp need not equal exp on the whole group G.

In this section, we will extend the local optimality results derived in Section 4 to global results under suitable assumptions on the underlying Lie group G. We shall see that the optimality of the subgroup factor holds globally whenever the polar decomposition is global as, for instance, in the case of connected real semisimple Lie groups with a finite center [**7**]. To this end, we introduce some further background on Riemannian manifolds.

Let M be a Riemannian manifold with distance function d, let H be a Lie group acting smoothly on M by isometries, and let $S \subset M$ be a smooth manifold. Assume that the map

$$\varphi\colon H \times S \to M, \qquad (h, s) \in H \times S \to hs \in M,$$

is a diffeomorphism from $H \times S$ onto M. Let $\mathfrak{h}$ denote the Lie algebra of H. The infinitesimal action of $\mathfrak{h}$ on M is denoted by

$$X \in \mathfrak{h} \quad \mapsto \quad v_X(m) = \left.\frac{d}{dt}\right|_{t=0} \exp(tX)m \in M, \qquad m \in M.$$

Lemma 5.1. *Let $a \in S$ and let T_aS and $T_a(Ha)$ be the tangent spaces to S and Ha, respectively. Assume that for every $a \in S$ it is true that T_aS is orthogonal to $T_a(Ha)$:*

$$T_aS \perp T_a(Ha). \tag{5.1}$$

Then, for all $s_1, s_2 \in S$ and $h_1, h_2 \in H$, it is true that

$$d(s_1, s_2) \le d(h_1s_1, h_2s_2). \tag{5.2}$$

Moreover, M is complete if and only if S, equipped with the restriction of the metric d, is complete.

If M is complete, then (5.2) *hold as an equality if and only if $h_1 = h_2$.*

Finally, S is a completely geodesic submanifold of M.[3]

Proof. Consider $c(t)\colon [0, 1] \to M$ as a C^1 curve joining the two manifold points h_1s_1 ($c(0) = h_1s_1$) and h_2s_2 ($c(1) = h_2s_2$). Denote by $\beta(t) \in H$ and $\gamma(t) \in S$ the unique C^1 curves such that $c(t) = \beta(t)\gamma(t)$. Then

$$\begin{aligned} c'(t) &= \left.\frac{d}{ds}\right|_{s=0} \beta(t+s)\gamma(t) + \left.\frac{d}{ds}\right|_{s=0} \beta(t)\gamma(t+s) \\ &= L'_{\beta(t)}[v_{X(t)}(\gamma(t)) + \gamma'(t))], \end{aligned}$$

where the infinitesimal action is given in terms of $X(t) = d/ds|_{s=0}\ \beta(t)^{-1}\beta(t+s) \in \mathfrak{h}$ and L' denotes the tangent map of the left multiplication by β (lifting φ to the corresponding tangent space, keeping the S argument fixed).

---

[3] A submanifold N of a Riemannian manifold M is said to be *totally geodesic* if for all $x \in N$ the geodesic $\mathrm{Exp}(tX) \in N$ coincides with $\exp(tX)$ for all $X \in T_xN$ for small values of t. N is *completely geodesic* if it is totally geodesic and complete.

Since H acts by isometries, one has

$$\|c'(t)\| = \|v_{X(t)}(\gamma(t)) + \gamma'(t)\|,$$

and, because of the orthogonality condition (5.1), it follows that

$$\|c'(t)\|^2 = \|v_{X(t)}(\gamma(t))\|^2 + \|\gamma'(t)\|^2,$$

from which it is easily deduced that

$$\|\gamma'(t)\| \le \|c'(t)\|, \qquad t \in [0, 1],$$

hence

$$d(s_1, s_2) \le \text{length}(\gamma) \le \text{length}(c)$$

(recall that $\text{length}(\gamma) = \int_0^1 \|\gamma'(t)\| dt$). Because of the arbitrariness of $c(t)$, the above relation holds passing to the infimum over all connecting curves c and, in particular, for the geodesic curve. Hence,

$$d(s_1, s_2) \le d(h_1 s_1, h_2 s_2).$$

We next prove the completeness of M given that S is complete. Let $\{m_j\}_{j\ge1}$ be a Cauchy sequence in M. Since φ is onto, there exist h_j, s_j in H, S, respectively, such that $m_j = h_j s_j$, $j = 1, 2, \dots$. However, from the above it follows that $d(s_i, s_j) \le d(h_i s_i, h_j s_j) = d(m_i, m_j)$, which implies that $\{s_j\}_{j\ge1}$ is a Cauchy sequence in S. Hence

$$\lim_{j\to\infty} s_j = s \in S,$$

since S is complete. Let $\varepsilon > 0$, such that $\bar{B}(s; \varepsilon)$, the closed ball of center s and radius ε, is compact. Then $\bar{B}(hs, \varepsilon)$ is compact for every $h \in H$. Fix $N \in \mathbb{N}$ such that $d(s_j, s) < \varepsilon/2$ and $d(m_i, m_j) < \varepsilon/2$ for all indices $i, j \ge N$. Then $d(m_j, h_N s) < \varepsilon$ for all $j \ge N$ and a fixed element $h_N \in H$. Thus, the sequence $\{m_j\}$ is a Cauchy sequence in $\bar{B}(h_N; \varepsilon)$ and hence it converges to a limit $m \in \bar{B}(h_N; \varepsilon)$, from which the completeness of M follows.

In what follows, let us assume that M is complete. Since φ is a diffeomorphism, S is closed in M and hence is complete as well. Now, if $d(h_1 s_1, h_2 s_2) = d(s_1, s_2)$, there exists a geodesic curve c connecting $h_1 s_1$ and $h_2 s_2$ with length equal to $d(s_1, s_2)$. With the notation above, $\text{length}(c) = \text{length}(\gamma)$, and because $\|c'(t)\| \ge \|\gamma'(t)\|$, it is true that $\|c'(t)\| = \|\gamma'(t)\|$ for all $t \in [0, 1]$. Thus,

$$v_{X(t)}(\gamma(t)) = O,$$

and, in turn, $X(t) = O$ for all t. Hence $\beta(t)$ is a constant curve, from which

$$h_1 = \beta(0) = \beta(1) = h_2,$$

and this part of the lemma follows.

To complete the proof of the lemma, it remains to show that S is totally geodesic at every point $a \in S$. For this purpose, consider the geodesic exponential map $\mathrm{Exp}_a\colon T_aM \to M$ at a, which is well defined because of the completeness of M. Let Ω be a open ball centered at $0 \in T_aM$ with the property that $\mathrm{Exp}_a|_\Omega$ is a diffeomorphism, and for every $X \in \Omega$ the geodesic $c\colon t \mapsto \mathrm{Exp}_a(tX)$ has length exactly equal to $d(a, \mathrm{Exp}_a X)$. Thus, if $\mathrm{Exp}_a(X) \in S$, then the curve c is contained in S. Hence the smooth manifold S' of Ω, consisting of $X \in \Omega$ such that $\mathrm{Exp}_a(X) \in S$, must equal $T_aS \cup \Omega$. From this it follows that S is totally geodesic at a. $\square$

We apply the above results to the Iwasawa decomposition for a real semisimple Lie group G. Let K be a subgroup of G such that $G_e^\sigma \subset K \subset G^\sigma$, and let $\mathfrak{k}$ be its Lie algebra. Assume, moreover, that G is connected and has a finite center. Then the map

$$(k, X) \in K \times \mathfrak{p} \to G \mapsto k \exp(X)$$

is a diffeomorphism onto (see [**7**],[**18**]), hence in this case, the polar decomposition

$$G = K \exp \mathfrak{p}$$

is global. Note that here, for simplicity, we consider the decomposition $G = K \exp \mathfrak{p}$ instead of $G = \exp \mathfrak{p} K$ as in the previous sections. However, as already observed, these two formalisms are equivalent, hence the same results apply to the other case as well, since the subgroup factor is essentially the same.

From the Cartan decomposition, we construct the Iwasawa decomposition as follows. Let $\mathfrak{a}$ be the maximal abelian subspace of $\mathfrak{p}$. One has the following properties:

- Every element of $\mathfrak{p}$ is K-conjugate to an element of $\mathfrak{a}$, i.e.,

$$\mathfrak{p} = \mathrm{Ad}(K)\mathfrak{a}.$$

- If λ is a linear functional on $\mathfrak{a}$, then define

$$\mathfrak{g}_\lambda = \{X \in \mathfrak{g}\colon [H, X] = \lambda(H)X, H \in \mathfrak{a}\}.$$

  Since $d\sigma(X) = -X$ for $X \in \mathfrak{a}$, one has $d\sigma(\mathfrak{g}_\lambda) = \mathfrak{g}_{-\lambda}$, and $\mathfrak{g}_0$ is invariant under $d\sigma$. Hence $\mathfrak{g}_0 \cup \mathfrak{p}_\sigma = \mathfrak{a}$. Set $\mathfrak{m} = \mathfrak{g}_0 \cup \mathfrak{k}_\sigma$, the centralizer of $\mathfrak{a}$ in $\mathfrak{k}_\sigma$. Then, $\mathfrak{g}_0 = \mathfrak{m} \oplus \mathfrak{a}$.
- Denote by Σ the set of roots of $\mathfrak{a}$ in $\mathfrak{g}$. Then Σ is a root system and

$$\mathfrak{g} = \mathfrak{m} \oplus \mathfrak{a} \oplus \bigoplus_{\alpha \in \Sigma} \mathfrak{g}_\alpha.$$

- The set Σ is the disjoint union of Σ^+ and $-\Sigma^+$, where Σ^+ is the set of roots that are positive on a fixed Weyl chamber C. Set

$$\mathfrak{n} = \bigoplus_{\alpha \in \Sigma^+} \mathfrak{g}_\alpha.$$

  Then $\mathfrak{n}$ is a nilpotent subalgebra of $\mathfrak{g}$.

- The exp function is a diffeomorphism from $\mathfrak{n}$ to N, a closed subgroup of G. Denote $A = \exp(\mathfrak{a})$. Then the map

$$(k, a, n) \in K \times A \times N \quad \rightarrow \quad kan \in G$$

  is a diffeomorphism and the decomposition $G = KAN$ is called the *Iwasawa decomposition*.

In the case of matrix Lie groups with the choice $\sigma(z) = z^{-T}$, one can observe that K corresponds to the group of orthogonal matrices, A corresponds to matrices with positive diagonal entries, and N corresponds to the group of upper triangular matrices, with one on the main diagonal. Then, the decomposition $z = kan$ is equivalent to the well-known QR decomposition of matrices, by identifying Q with the k factor and R with the an factor.

The inner product $\langle X, Y\rangle = -B(X, d\sigma(Y))$, introduced in Section 4, induced a right G-invariant Riemannian metric on G, with distance function d. Then, as a consequence of Lemma 5.1, G is a complete Riemannian manifold (choose $M = G$, $S = \{e\}$, and $H = G$).

Lemma 5.2. *For all $a \in A$, the tangent spaces $T_a A$ and $T_a(KaN)$ are orthogonal.*

Proof. Let $a \in A$. We show first that $aNa^{-1} = N$. For this purpose, let $v \in N$. Then $v = \exp V$, $V \in \mathfrak{n}$. One has

$$ava^{-1} = a \exp Va^{-1} = \exp \mathrm{Ad}(a)V \in N,$$

since $\mathrm{Ad}(A) = \mathrm{e}^{\mathrm{ad}_\mathfrak{a}}$ leaves $\mathfrak{n}$ invariant.

Thus, right multiplication by a^{-1} maps the set A onto itself and the set KaN onto $KaNa^{-1} = KN$ as a consequence of the argument above. Therefore, it is sufficient to establish the result for $a = e$. Recall that

$$T_e A = \mathfrak{a} \qquad \text{and} \qquad T_e(KN) = \mathfrak{k} + \mathfrak{n},$$

hence it suffices to show that $\mathfrak{a} \perp \mathfrak{k} + \mathfrak{n}$. We already know that $\mathfrak{a} \perp \mathfrak{n}$ because of the properties of the root-space decomposition described above. Therefore it remains to show that $\mathfrak{a} \perp \mathfrak{k}$. To this end, let $X \in \mathfrak{k}$. Then X decomposes as

$$X = X_\mathfrak{m} + X_\mathfrak{a} + \sum_{\alpha \in \Sigma^+} (X_\alpha + X_{-\alpha}),$$

where $X_\mathfrak{m} \in \mathfrak{m}$, $X_\mathfrak{a} \in \mathfrak{a}$, and $X_\alpha \in \mathfrak{g}_\alpha$ for $\alpha \in \Sigma$. Since $d\sigma(X) = X$ and $d\sigma(X_\alpha) = X_{-\alpha}$, we deduce that $X_\mathfrak{a} = O$, hence $X \perp \mathfrak{a}$. ☐

Corollary 5.2.1. *Let $z = kan \in G$, the Iwasawa decomposition of z. Then*

$$d(a, e) \leq d(kan, e).$$

whereby the above relation holds as an equality if and only if $k = n = e$. *Furthermore, A is a totally geodesic submanifold of G and* $\mathrm{Exp}_e|_{\mathfrak{a}} = \exp|_{\mathfrak{a}}$.

Proof. We apply Lemma 5.1 with $M = G$, $S = A$, and $H = K \times N$. The action of H on G is given by

$$(k, n) \cdot g = kgn^{-1}.$$

Note that, because of Lemma 4.2, the scalar product $\langle\cdot,\cdot\rangle$ is $\mathrm{Ad}(K)$ invariant, hence the Riemannian metric on G is left K invariant as well. Thus, the group H acts by isometries. Furthermore, as a consequence of the above lemma, all hypotheses of Lemma 5.1 are satisfied hence all assertions but the last follow.

To prove the last assertion, note that, since $\mathfrak{a}$ is totally geodesic, $\mathrm{Exp}_e|_{\mathfrak{a}}$ is the geodesic exponential map of A equipped with the restriction metric. Since the set $\mathfrak{a}$ is abelian, the exponential map is an isometry from the Euclidean space $\mathfrak{a}$ onto A. Thus, if $X \in \mathfrak{a}$, then the geodesic curve c in $\mathfrak{a}$, emanating from 0 with velocity $c'(0) = X$, is given by $c(t) = tX$. Hence, $\mathrm{Exp}_e(X) = \exp(c(t))|_{t=1} = \exp(X)$, which completes the proof of the result. □

Corollary 5.2.2. *Let* $x \in G_\sigma = \exp \mathfrak{p}$. *Then, for all* $k \in K$, *it is true that*

$$d(x, e) \le d(kx, e),$$

with equality if and only if $k = e$. *Moreover,* $\mathrm{Exp}_e|_{\mathfrak{p}} = \exp|_{\mathfrak{p}}$.

Proof. Let $x \in G_\sigma = \exp \mathfrak{p}$. Since $\mathfrak{p}$ is $\mathrm{Ad}(K)$ conjugate to $\mathfrak{a}$, there exists $k_1 \in K$ and $a \in A$ such that $x = k_1 a k_1^{-1}$. Then,

$$d(kx, e) = d(kk_1ak_1^{-1}, e) = d(k_1^{-1}kk_1a, e),$$

where the last passage follows by left invariance of the metric under the subgroup K. Because of the previous corollary,

$$d(k_1^{-1}kk_1a, e) \ge d(a, e) = d(x, e),$$

with equality if and only if $k_1^{-1}kk_1 = e$, in other words, $k = e$.

For the last assertion, let $X \in \mathfrak{p}$. Then there exists $k \in K$ and $V \in A$ such that $X = \mathrm{Ad}(k)V$. Since $\mathrm{Ad}_k\colon x \mapsto kxk^{-1}$ is an isometry with tangent map $\mathrm{Ad}(k)$, it follows that

$$\mathrm{Exp}_e(X) = k\,\mathrm{Exp}_e(V)k^{-1} = k\exp(V)k^{-1} = \exp \mathrm{Ad}(k)V = \exp X$$

(the second equality follows from Corollary 5.2.1), which proves our statement. □

Now we have all the tools to extend the optimality of the polar decomposition to a global result.

Theorem 5.3. *Let $z \in G$ and denote $z = kp$ the polar decomposition of z, with $k \in G^\sigma$ and $p \in G_\sigma = \exp \mathfrak{p}$. Then, for all $k_1 \in G^\sigma$, it is true that*

$$d(z, k) \le d(z, k_1),$$

with equality if and only if $k_1 = k$. Hence, the polar factor k is the best approximant of z in G^σ.

Proof. Let $z = kp$ be the polar decomposition of z. Now, by left invariance of the metric under G^σ, one has

$$d(z, k) = d(kp, k) = d(p, e).$$

On the other hand, if k_1 is an arbitrary element in G^σ, one has

$$d(z, k_1) = d(kp, k_1) = d(k_1^{-1}kp, e),$$

where again we have made use of the left invariance of the metric under G^σ. Thus, the result follows by Corollary 5.2.2. □

6. Applications to Computations

6.1. *Approximation of the Matrix Exponential*

The first example consists in the numerical approximation of the exponential of a matrix from a Lie algebra to a Lie group. Such computations are ubiquitous in the numerical solution of ODEs on Lie groups [**11**]. A particularly hard case is when $\mathfrak{g} = \mathfrak{sl}(N)$ and $G = \mathrm{SL}(N)$, since any analytical approximation of the matrix exponential is bound to fail unless the exponential is computed exactly [**4**]. The procedure described below is a very general nonanalytical approximation that can be applied, among others, to the approximation of the exponential in $\mathrm{GL}(N)$, $\mathrm{SO}(N)$, and $\mathrm{SO}(p, q)$. With small modifications, a very similar approach can be applied to the case of the symplectic group $\mathrm{Sp}(N)$. We refer to [**16**] for further details.

Assume that we wish to approximate the exponential $\exp(tZ) \in \mathrm{SL}(N)$ of a matrix $Z \in \mathfrak{sl}(N)$, so that the approximation $F(t, Z) \in \mathrm{SL}(N)$. 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 $d\sigma(Z) = SZS$ and

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

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

We truncate the series $X(t)$ and $Y(t)$ to a given order of accuracy p, and approximate

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

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 $X(t)$ only possess one row and one column and its exponential is very easy to compute exactly. Commutators of matrices in $\mathfrak{p}$ and $\mathfrak{k}$ can be computed employing only matrix–vector products, amounting to $\mathcal{O}(N^2)$ operations. The procedure is repeated for the new matrix $Z^{[1]} = \sum_{n=0}^{p} Y_n t^n$, with the new automorphism $d\sigma^{[1]}V = S^{[1]}VS^{[1]}$, where

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

Proceeding in a similar manner, after n steps one obtains the approximation

$$\exp(tZ) \approx \exp(X^{[1]}(t)) \exp(X^{[2]}(t)) \cdots \exp(X^{[N]}(t)),$$

where all the matrices $X^{[i]}$, for $i = 1, 2, \ldots, N-1$, possess only one row and one column, while $X^{[N]}$ is a diagonal matrix.

Practical algorithms that follow this approach are described at greater length in [**16**]. For large N, these algorithms approximate the matrix exponential from $\mathfrak{sl}(N)$ to $\mathrm{SL}(N)$ in $\mathcal{O}(3\frac{1}{3}N^3)$, $\mathcal{O}(7N^3)$, $\mathcal{O}(9N^3)$ floating point operations and order 2, 3 and 4, respectively—very competitive with standard methods for the computation of the matrix exponential.

6.2. *Centro-Symmetric Matrix Approximants*

We will now discuss a matrix decomposition which is much less familiar than the classical polar decomposition. Let F be the $N \times N$ "flip" matrix, defining the following action on an arbitrary n vector:

$$(F)(j) = (N + 1 - j) \qquad \text{for} \quad j = 1, \ldots, N.$$

That is, F is a matrix with ones on the main lower-left to upper-right diagonal, and zeros elsewhere. We say that a matrix A is *centro-symmetric* if $FAF = A$, and that it is *centro-skew* if $FAF = -A$. Since $A \mapsto FAF$ is a 180° rotation about the matrix center, this means that elements situated symmetrically with respect to the center are either equal or equal under a sign change. As a motivation, consider, e.g., discretizations of spatial differential operators. A central difference approximation of d^2/dx^2 is centro-symmetric, while d/dx is centro-skew. On the Lie group of all real nonsingular $N \times N$ matrices we define the involutive authomorphism $\sigma(a) = FaF$, which induces the same authomorphism on the Lie algebra $\mathfrak{g} = \mathfrak{gl}(N)$:

$$d\sigma(A) = FAF.$$

The projection $\pi_{\mathfrak{k}} = \frac{1}{2}(I + d\sigma)$ projects $\mathfrak{g}$ onto the subalgebra $\mathfrak{k} \subset \mathfrak{g}$ of centro-symmetric matrices, while $\pi_{\mathfrak{p}} = \frac{1}{2}(I - d\sigma)$ projects onto the Lie triple system consisting of centro-skew matrices. Thus, in the algebra, we have introduced a decomposition of a general matrix into the sum of a centro-symmetric and a centro-skew matrix. At the group level this automorphism defines the two spaces $G^\sigma = \{k \in G \mid FkF = k\}$ (centro-symmetric) and $G_\sigma = \{p \in G \mid FpF = p^{-1}\}$, which we may call *centro-orthogonal* matrices. It is easily verified that the matrix exponential maps $\mathfrak{k}$ into G^σ and $\mathfrak{p}$ into G_σ. The centro-symmetric matrices G^σ form a Lie group while the centro-orthogonal matrices form a symmetric space closed under the symmetric product $x \cdot y = xy^{-1}x$. Compare this to the classical polar decomposition, where the symmetric matrices form a symmetric space and the orthogonal matrices form a Lie group.

Theorem 3.1 thus states that any matrix $z = \exp(Z)$ can be written as $z = xy$, where $x = \exp(X)$ is centro-orthogonal and $y = \exp(Y)$ is centro-symmetric. Letting $Z = P + K$ where P is centro-skew and K is centro-symmetric, we find that $X = \sum_{k=1}^{\infty} X_k$ and $Y = \sum_{k=1}^{\infty} Y_k$, where X_k and Y_k are given in (3.2) and (3.4), and we find that z is decomposed as

$$z = \exp(Z) = \exp(X)\exp(Y). \tag{6.1}$$

From Section 4 we know that $y = \exp(Y)$ is the best centro-symmetric approximation to z, and that the distance between y and z is given as

$$d(z, y) = \|X\| = \|P - \tfrac{1}{2}[P, K] + \cdots\|, \tag{6.2}$$

where $\|\cdot\|$ is the Frobenius norm.

6.3. *Decomposition of Vector Fields with Polynomial Coefficients*

In this last example we apply the generalized polar decomposition to vector fields with polynomial coefficients. Consider the differential equation

$$\begin{aligned} \frac{du}{dt} &= p(u, v), \\ \frac{dv}{dt} &= q(u, v), \end{aligned} \tag{6.3}$$

where $p(u, v)$ and $q(u, v)$ are two polynomials in u and v. We wish to decompose the vector field $F(u, v) = p(u, v)\partial_u + q(u, v)\partial_v$ into two vector fields, one that has the transformation R: $(u, v) \mapsto (u, -v)$ as a symmetry, and the other one that has R as a reversal symmetry. To this goal, we consider the Lie group of flows φ_t of vector fields with polynomial coefficients and consider the automorphism

$$\sigma\varphi_t = R\varphi_t R,$$

where R is the matrix

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

Since the map R is a linear transformation, the corresponding algebra automorphism on the vector field Z is

$$d\sigma Z = RZR,$$

hence

$$F(u, v) = P(u, v) + K(u, v),$$

where

$$P(u, v) = \tfrac{1}{2}((p(u, v) - p(u, -v))\partial_u + (q(u, v) + q(u, -v))\partial_v)$$

and

$$K(u, v) = \tfrac{1}{2}((p(u, v) + p(u, -v))\partial_u + (q(u, v) - q(u, -v))\partial_v).$$

We can construct the series $X(t)$ and $Y(t)$ in terms of the vector fields P and K above as in Section 3—the commutator being the usual Jacobi bracket of vector fields. In this manner, one obtains the decomposition

$$\varphi_t = \exp(tF) = \exp(X(t)) \exp(Y(t)),$$

where now $\exp(Y(t))$ has R as a symmetry and $\exp(X(t))$ has R as a reversal symmetry, in other words,

$$R \exp(X(t))R = \exp(-X(t)), \qquad R \exp(Y(t))R = \exp(Y(t)).$$

For instance, assume $p(u, v) = (1-u)v$, $q(u, v) = 2u - 3v$. One has $P(u, v) = (1-u)v\partial_u + 2v\partial v$, $K(u, v) = -3v\partial_v$, $[P, K] = 3v(1-u)\partial_u - 6u\partial_v$, etc. Note that vector fields in $\mathfrak{k}$ and $\mathfrak{p}$ have reflection with respect to the u axis as a symmetry and antisymmetry, respectively.

Let us restrict our attention to polynomial vector fields of degree at most $r < \infty$. To obtain a Lie algebra, we modify the Jacobi bracket truncating all the polynomial terms of degree higher than r. For instance, if $P(u, v) = (1-u)v\partial_u + 2u\partial_v$, $K(u, v) = -3v\partial_v$, and $r = 2$, then

$$[P, [P, K]]_r = 6(u - u^2 + v^2)\partial_u,$$

whereas the classical Jacobi bracket would give $[P, [P, K]] = 6(u - u^2 + v^2 - uv^2)\partial_u$.

The corresponding Lie algebra $\mathfrak{g}$ is now finite dimensional and we can consider the distance d induced by the canonical Cartan–Killing form, $\mathrm{tr}(\mathrm{ad}_A\ \mathrm{ad}_B)$, which now reduces to computing traces of appropriate matrices describing the adjoint operator.[4] Thus one has the following optimality result: in the metric defined by d, the time-1 map of the r-degree polynomial vector field $Y(t)$ is the best approximant to $\exp(tF)$ among all time-1 flows of polynomial vector fields of degree r that are symmetric with respect to the u axis. The distance in the approximation is given by $\|X(t)\|$.

7. Conclusions

In this paper we have discussed some issues related to the polar decomposition in Lie groups as analogous to the polar decomposition of matrices. Such a factorization always exists and its factors are differentiable, provided that the group element we wish to factorize is sufficiently close to the group identity. Moreover, the decomposition is well known to be global under certain assumptions on the group G. Our contribution to the understanding of the subject is twofold: first of all, we have derived explicitly the Lie-algebra series expansions that determine the factors uniquely near the identity. Second, we have proved that the subgroup factor possesses certain optimality properties, namely, given a Lie-group element z that has polar decomposition $z = xy$, where $x \in G_\sigma$ is the symmetric space factor and $y \in G^\sigma$ is the subgroup factor, then y is the best approximant to z, both locally and globally, under the assumption that σ defines a Cartan decomposition. This result is a Lie-group analog of the optimality of the orthogonal factor in the polar decomposition of matrices.

The existence of this decomposition in such a general setting is very relevant, among others, in the context of integration of dynamical systems possessing geometric attributes such as symmetries and reversing symmetries. Numerical inte-

[4] We assume that this Lie algebra is semisimple, otherwise one can always perform the Levi decomposition of $\mathfrak{g}$ into a semisimple and nilpotent subalgebra.

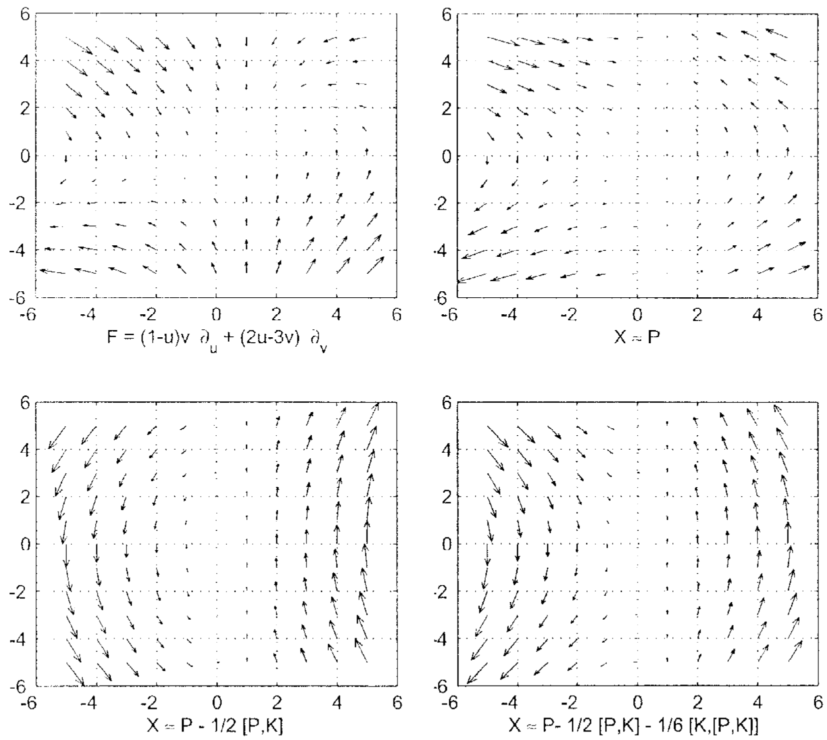

Fig. 1. Vector field F and the first terms in the X series.

grators for differential equations usually introduce a perturbation of the underlying vector field, so that the numerical flow (numerical approximation thought of as the flow of a modified vector field) seldom preserves the same symmetries/reversing

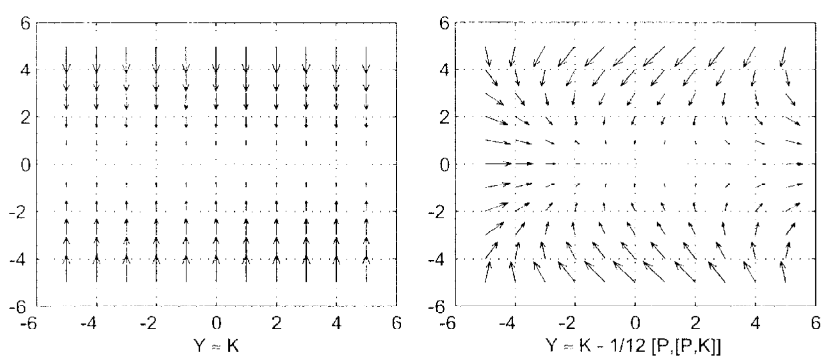

Fig. 2. First terms in the Y series.

symmetries, and ad hoc numerical schemes need to be introduced instead or projection techniques to be employed. Thus, one could perform a “polar factorization” of the numerical integrator and discard the factor that is not relevant for the given problem. For instance, the well-known *generalized Scovel projection* for numerical methods to preserve reversing symmetries [**14**] corresponds to an approximation of the $\mathfrak{p}$ factor of the polar decomposition introduced in this paper, under the choice of a suitable automorphism σ [**15**].

It would also be interesting to investigate whether linear algebra techniques for computing the polar factorization of matrices can be extended in some respect to the general Lie-group case. It would also be of interest as to whether the recurrence relations for the polar factors presented in this paper can be used to devise numerical methods competitive with the standard algorithms for computing the polar decomposition [**8**]. These issues are currently under investigation and we hope to come up with further results in a future paper.

8. Appendix

With the same notation as Theorem 3.1, we prove that the series $\sum_{k=1}^{\infty} X_k t^k$, with the X_k's as in (3.2), is absolutely converging.

Assume that $\mathfrak{g}$ is a Banach space and that the ad operator is bounded (which is always the case when $\mathfrak{g}$ is finite dimensional). Let μ be the smallest constant such that

$$\|[Y_1, Y_2]\| \leq \mu \|Y_1\| \|Y_2\|, \qquad \forall\, Y_1, Y_2 \in \mathfrak{g},$$

and denote $\alpha = \max(\|K\|, \|P\|)$. From (3.2) we deduce that

$$(k+1)\|X_{k+1}\| \leq \alpha \left(\mu \|X_k\| + \sum_{\substack{\ell \geq 1 \\ 2\ell \leq k}} |c_{2\ell}| \mu^{2\ell} \sum_{\substack{\ell_1, \ldots, \ell_{2\ell} > 0 \\ \ell_1 + \cdots + \ell_{2\ell} = k}} \|X_{\ell_1}\| \cdot \|X_{\ell_2}\| \cdots \|X_{\ell_{2\ell}}\| \right). \tag{8.1}$$

Consider next the differential equation

$$\frac{dw}{du} = h(w), \qquad w(0) = 0, \tag{8.2}$$

where

$$h(u) = 1 + u + \sum_{k=1}^{\infty} |c_{2k}| u^{2k},$$

which has an analytic solution for some constant $0 < \delta < \pi$. Set $w(u) = \sum_{k=1}^{\infty} w_k u^k$ for $|u| < \delta$. It is easily verified that the w_k's are positive and obey the recurrence relation

$$(k+1) w_{k+1} = w_k + \sum_{\substack{\ell \geq 1 \\ 2\ell \leq k}} |c_{2\ell}| \sum_{\substack{\ell_1, \ldots, \ell_{2\ell} > 0 \\ \ell_1 + \cdots + \ell_{2\ell} = k}} w_{\ell_1} w_{\ell_2} \cdots w_{\ell_{2\ell}}, \qquad k = 1, 2, \ldots, \tag{8.3}$$

with starting value $w_1 = 1$. We claim that $\|X_k\| \le \mu^{k-1}\alpha^k w_k$ for $k = 1, 2, \ldots$. Clearly, the statement is true for $k = 1$, since $\|X_1\| = \|P\| \le \alpha$ and $w_1 = 1$. Next, assume that the statement is true for $m = 1, 2, \ldots, k$. From the induction hypothesis, together with (8.1) and (8.3), we deduce

$$\begin{aligned}(k+1)\|X_{k+1}\| &\le \mu^k\alpha^{k+1}w_k + \alpha\sum_{\substack{\ell\ge 1\\ 2\ell\le k}}|c_{2\ell}|\mu^{2\ell}\sum_{\substack{\ell_1,\ldots,\ell_{2\ell}>0\\ \ell_1+\cdots+\ell_{2\ell}=k}}\mu^{k-2\ell}\alpha^k w_{\ell_1}w_{\ell_2}\cdots w_{\ell_{2\ell}}\\ &= \alpha^{k+1}\mu^k(k+1)w_{k+1}.\end{aligned}$$

It follows that the series $\sum_{k=1}^{\infty}\|X_k\|t^k$ is converging in the disk of radius $\delta/(\alpha\mu)$, being bounded by the absolutely converging series $(1/\mu)\sum_{k=0}^{\infty}w_k(t\alpha\mu)^k$. This completes the proof of Theorem 3.1. □

Acknowledgments

H. Z. Munthe-Kaas and G. R. W. Quispel wish to thank the Norwegian Research Council, while A. Zanna is grateful to the Australian Research Council for financial support. The authors are also grateful to a number of colleagues and friends with whom they have discussed the material presented in this paper. Special thanks to Erik van der Ban and Stig Faltinsen.

References

[1] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*, Dover, New York, 1965.
[2] V. I. Arnold and B. A. Khesin, *Topological Methods in Hydrodynamics*, Springer-Verlag, New York, 1998.
[3] L. Autonne, Sur les groupes linéaires, réel et orthogonaux, *Bull. Soc. Math. France* **30** (1902), 121–134.
[4] E. Celledoni and A. Iserles, Methods for the approximation of the matrix exponential in a Lie-algebraic setting, Technical Report NA1999/03, DAMTP, University of Cambridge, *IMA J. Numer. Anal.* (to appear).
[5] D. G. Ebin and J. E. Marsden, Group of diffeomorphisms and the motion of an incompressible fluid, *Ann. of Math.* **92** (1970), 102–163.
[6] K. Fan and A. J. Hoffman, Some metric inequalities in the space of matrices, *Proc. Amer. Math. Soc.* **6** (1955), 111–116.
[7] S. Helgason, *Differential Geometry, Lie Groups and Symmetric Spaces*, Academic Press, New York, 1978.
[8] N. J. Higham, Computing the polar decomposition—with applications, *SIAM J. Sci. Statist. Comput.* **7** (1986), 1160–1174.
[9] J. Hilgert, K. H. Hoffmann, and J. D. Lawson, *Lie Groups, Convex Cones and Semigroups*, Oxford University Press, New York, 1989.
[10] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, New York, 1985.
[11] A. Iserles, H. Munthe-Kaas, S. P. Nørsett, and A. Zanna, Lie-group methods, *Acta Numerica* **9** (2000), 215–365.
[12] J. D. Lawson, Polar and Ol'shanskii decompositions, *J. Reine Angew. Math.* **448** (1994), 191–219.
[13] O. Loos, *Symmetric Spaces I: General Theory*, W. A. Benjamin, New York, 1969.

[14] R. I. McLachlan, G. R. W. Quispel, and G. S. Turner, Numerical integrators that preserve symmetries and reversing symmetries, *SIAM J. Numer. Anal*. **35**(2) (1998), 586–599.
[15] H. Munthe-Kaas, G. R. Quispel, and A. Zanna, Application of symmetric spaces and Lie triple systems in numerical analysis (in preparation).
[16] H. Munthe-Kaas and A. Zanna, Generalized polar decompositions for the approximation of the matrix exponential, Technical Report, no. 200, Department of Informatics, University of Bergen, Norway, August 2000.
[17] B. O'Neill, *Semi-Riemaniann Geometry*, Academic Press, New York, 1983.
[18] V. S. Varadarajan, *Lie Groups, Lie Algebras, and Their Representation*, Graduate Text in Mathematics, Vol. 102, Springer-Verlag, New York, 1984.
[19] A. Zanna, Recurrence relation for the factors in the polar decomposition on Lie groups, Technical Report, no. 192, Department of Informatics, University of Bergen, May 2000.