REPORTS
IN
INFORMATICS

ISSN 0333-3590

The polar decomposition on Lie groups with involutive automorphisms

H. Munthe-Kaas, R.G.W. Quispel, A. Zanna

REPORT NO 191 May 2000

Department of Informatics

UNIVERSITY OF BERGEN

Bergen, Norway

This report has URL

`http://www.ii.uib.no/publikasjoner/texrap/ps/2000-191.ps`

Reports in Informatics from Department of Informatics, University of Bergen, Norway, is available at `http://www.ii.uib.no/publikasjoner/texrap/`.

Requests for paper copies of this report can be sent to:

Department of Informatics, University of Bergen, Høyteknologisenteret,
P.O. Box 7800, N-5020 Bergen, Norway

The polar decomposition on Lie groups with involutive automorphisms

H. Munthe-Kaas[*], G. R. W. Quispel[†] and A. Zanna[‡]

11th May 2000

Abstract

The polar decomposition, a well known algorithm for decomposing real matrices in 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. Lawson (1994) proved the existence of the polar decomposition to arbitrary Lie groups endowed with an involutive automorphism σ. Such decomposition, depending on σ, always exists nearby the identity, and its existence in such a general setting implies that one may decompose other mathematical objects other than matrices, like, for instance, flows of differential equations.

In this paper, we provide an alternative proof to the existence and uniqueness result of the polar decomposition of Lawson. In particular, our result is constructive: we derive differential equations obeyed by the two factors and solve them analytically, thereby providing a recurrence relation for the coefficients of the series expansion.

Further results on properties of the two factors are also introduced. We prove that the subgroup factor obeys similar optimality properties to the orhogonal polar factor in the classical matrix setting.

1 Introduction

It is well known in linear algebra that any $N \times N$ matrix A can be decomposed in 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 (Horn & Johnson 1991). 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 (1902) as a matrix analog of the polar form of a complex number

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

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

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

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

The popularity of the polar decomposition is mainly due to the best approximation properties of its factors. Fan & Hoffman (1955) proved that

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

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 is discussed in Higham (1986)

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 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

$$AA^{\mathrm{T}} = HUU^{\mathrm{T}}H^{T} = H^2,$$

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

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

However, in a recent investigation of symmetric spaces and their connection with numerical analysis (Munthe-Kaas, Quispel, Nikolayevsky & Zanna 2000), 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 space and a term in a subgroup of the given Lie group.

There exist a number of papers on the polar decomposition on Lie groups and its generalization to semigroups (*Ol'shanskii decomposition*), many of them rather recent. To our knowledge, the proof of existence and uniqueness of the polar decomposition in Lie groups is due to (Lawson 1994).

In this paper, we present an alternative proof to the existence and uniqueness of the polar decomposition in Lie groups. We derive differential equations obeyed by the two factors and solve one of them analytically, thereby obtaining a recurrence relation of the coefficients of the series expansion of the symmetric-space factor. We also present the recurrence relation that generates the subgroup factor, referring to (Zanna 2000) for its derivation. The latter recurrence relation is rather complicated and its derivation requires a number of series expansions involving various Lie-algebra operatots. Furthermore, we show that the subgroup factor is expanded in odd powers of time only, a result of interest in the setting of numerical integrators, and finally prove some optimality results for semisimple Lie groups with right-invariant metrics inherited from the Killing–Cartan form.

The article is organized as follows. In Section 2 we introduce background theory on symmetric spaces of Lie groups and Lie triple systems. The main results of this paper are presented in Section 3 and Section 4. We derive differential equations for the polar factors and derive a recurrence relation for the coefficients of the series expansion of one of the factor. Additional properties of the factors are also discussed. In Section 4 we prove optimality properties of the subgroup factor as an approximant to the original Lie group element. This is an analogous to the matrix case. Finally, Section 5 is devoted to conclusions.

2 Background theory

Let G be a connected Lie group and $\sigma : G \to G$ 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,

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

and G_σ the set of anti-fixed points of σ,

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

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 that of a *symmetric space* when endowed with the non-associative multiplication $x \cdot y = xy^{-1}x$. We recall that a symmetric space is 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 will be denoted as $\mathrm{d}\sigma$. Let $X \in \mathfrak{g}$ and consider $x = \exp(tX)$. Then

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

Let λ be an eigenvalue of $\mathrm{d}\sigma$, $\mathrm{d}\sigma X = \lambda X$. Then $\mathrm{d}\sigma^2 X = X$, hence $\lambda = \pm 1$. Consider the spaces

$$\mathfrak{k}_\sigma = \{X \in \mathfrak{g} : \mathrm{d}\sigma(X) = X\}$$

of fixed points of $\mathrm{d}\sigma$, and

$$\mathfrak{p}_\sigma = \{X \in \mathfrak{g} : \mathrm{d}\sigma(X) = -X\}$$

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

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

(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. The projection in each subspace is given by the formula

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

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

In order to keep a parallelism between the general Lie-group theory and the classical matrix theory, let $\mathrm{GL}(N)$ be the group of $N \times N$ invertible real matrices. Consider the map

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

It is clear that σ is an involutive automorphism of $\mathrm{GL}(N)$. Then, from above, the set of $G_\sigma = \{x \in \mathrm{GL}(N) : \sigma(x) = x^{-1}\}$ is a symmetric space. The set G_σ is the set of invertible symmetric matrices. The symmetric space G_σ is disconnected and particular mention deserves its connected component containing the identity matrix I, since it reduces to 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 determinat (the Lie group $\mathrm{SO}(N)$).

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

$$\begin{aligned}
\mathrm{d}\sigma(X) &= \tfrac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0}\sigma\big(\exp(tX)\big) = \tfrac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0}\Big(I + tX + \mathcal{O}(t^2)\Big)^{-\mathrm{T}} && (2.5)\\
&= \tfrac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0}\Big(I + tX^{\mathrm{T}} + \mathcal{O}(t^2)\Big)^{-1} = \tfrac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0}\Big(I - tX^{\mathrm{T}} + \mathcal{O}(t^2)\Big) \\
&= -X^{\mathrm{T}}, && (2.6)
\end{aligned}$$

hence we deduce that

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

the classical algebra of skew-symmetric matrices, while

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

is the classical set of symmetric matrices. Such 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 else than the canonical decomposition of a matrix into its skew-symmetric and symmetric part,

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

However, as mentioned in the introduction, the above procedure is very general: for example one can choose $G = \mathrm{Diff}(M)$, the group of diffeomorphisms of a manifold M, and set $\sigma(\varphi) = \mathcal{R}\varphi\mathcal{R}^{-1}$, with $\mathcal{R}$ involutive diffeomorphism of M onto M. If F denotes a vector field, so that $\varphi = \exp(tF)$, then the sets G_σ and G^σ are precisely the vector fields that possess $\mathcal{R}$ as a reversing symmetry (irony of sort, this is a symmetric space!) and $\mathcal{R}$ as a symmetry. By proving the existence of the polar decomposition in a general Lie-group context we deduce that given $\mathcal{R}$, every diffeomorphism can be written as the composition of two flows, one possessing $\mathcal{R}$ as a symmetry and the one as a reversing symmetry. Such a decomposition has fundamental implication in 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 (McLachlan, Quispel & Turner 1998) and references therein).

For those who are interested in a further reading on symmetric spaces and Lie triple systems, we refer to (Helgason 1978) and (Loos 1969).

3 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 $y \in G^\sigma$, where $G_\sigma = \{x \in G : \sigma(x) = x^{-1}\}$ and $G^\sigma = \{x \in G : \sigma(x) = x\}$. We call the decomposition $z = xy$ a *polar decomposition* of z, in analogy to the terminology of linear algebra.

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

Proof. Set

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

so that $Z = K + P$ and $\mathrm{d}\sigma(P) = -P$ and $\mathrm{d}\sigma(K) = K$. 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 $\mathfrak{p}_\sigma$ and the Y_i's are in $\mathfrak{k}_\sigma$. Imposing

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

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

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

et cetera. As we shall see later, all the X_i's and Y_i's are 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 function and prove that 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

$$\mathrm{dexp}_X X' = Z - \mathrm{Ad}_{\exp(X)}\mathrm{dexp}_Y Y'.$$

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

$$\mathrm{Ad}_{\exp(-X)}\mathrm{dexp}_X X' = \mathrm{Ad}_{\exp(-X)}Z - \mathrm{dexp}_Y Y'.$$

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

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

Our goal is to decompose the above expression in $\mathfrak{k}_\sigma \oplus \mathfrak{p}_\sigma$. To this aim, observe that $\mathrm{dexp}_Y Y' \in \mathfrak{k}_\sigma$ since $Y, Y' \in \mathfrak{k}_\sigma$ and $\mathfrak{k}_\sigma$ 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 $\mathfrak{k}_\sigma$ for even k while it is in $\mathfrak{p}_\sigma$ for odd values of k. Conversely, we have $\mathrm{ad}_X^{2m+1}(P) \in \mathfrak{k}_\sigma$, and $\mathrm{ad}_X^{2m}(P) \in \mathfrak{p}_\sigma$ for $m = 0, 1, 2, \ldots$. In summary,

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

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

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

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

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

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

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

Note that

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

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

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

B_k being the k-th Bernoulli number (Abramowitz & Stegun 1965).

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

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

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

We astain to 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 the differential equation obeyed by $Y(t)$. Matching terms in $\mathfrak{k}_\sigma$, we obtain

$$-\frac{1}{u}(\cosh u - 1)X' = \cosh u(K) - \sinh u(P) - \mathrm{dexp}_Y Y',$$

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

$$\mathrm{dexp}_Y Y' = K + (\mathrm{csch} u - \coth u)(P), \quad u = \mathrm{ad}_X.$$

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

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

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 (Zanna 2000) 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 $-\mathrm{d}\sigma(Z) = P - K$. More specifically,

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

(see (Zanna 2000)).

Let $z = xy$ the differentiable polar decomposition of z which exists at least for sufficiently small t. Note that

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

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

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

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

$$\exp(V)\exp(W) = \exp(\mathrm{bch}(V,W)),$$

for all $V, W \in \mathfrak{g}$ of sufficiently small norm (see (Varadarajan 1984)). Furthermore,

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

Equations (3.5) and (3.6) are derived in (Lawson 1994) 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 (Hilgert, Hoffmann & Lawson 1989).

The first terms (up to t^6) in the expansion of X and Y are displayed in Table 1 below.

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

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

Table 1: Taylor expansion at $t = 0$ of $X(t)$ and $Y(t)$ in terms of P and K.

Proposition 3.2 *Let xy be the differentiable 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 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 polar decomposition of z^{-1}. On the other hand,

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

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

Corollary 3.2.1 *In the same assumptions of Proposition 3.2, it is true that z can be uniquely decomposed as*

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

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. $\square$

4 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 also for the polar decomposition in a semisimple Lie group similar optimality result hold. 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) $d(\cdot,\cdot) : G \times G \to \mathbb{R}^+$ is *right* (resp. *left*) *invariant* if

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

(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, $T_g G = \{Xg | X \in \mathfrak{g}\}$, we obtain the right invariant riemaniann as

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

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

$$\mathrm{length}(\gamma) = \int_0^1 \langle\dot\gamma(t), \dot\gamma(t)\rangle^{\frac{1}{2}}\,\mathrm{d}t.$$

The shortest curve (minimising geodesic) between sufficiently close points x and y is given as

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

where $\exp(Z) = yx^{-1}$. The right invariant metric on G is now defined as

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

Let us now introduce a canonical innerproduct $\langle\cdot,\cdot\rangle$ 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 non degenerate.

Lemma 4.1 *Let $\mathfrak{g}$ be a semisimple Lie group and* $\mathrm{d}\sigma$ *and involutive automorphism on $\mathfrak{g}$ as above so that $\mathfrak{g} = \mathfrak{p}_\sigma \oplus \mathfrak{k}_\sigma$ is a Cartan decomposition. Then the symmetric bilinear form*

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

is positive definite on $\mathfrak{g}$. Moreover, the $\mathfrak{p}_\sigma$ and $\mathfrak{k}_\sigma$ are orthogonal with respect to the inner product

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

Proof. The positivity of the symmetric bilinear form $B_{\mathrm{d}\sigma}$ is proven in Helgason (1978), Prop. 7.4 page 184. The orthogonality of the two subspaces is also immediate: let $X \in \mathfrak{p}_\sigma$ and $Y \in \mathfrak{k}_\sigma$. Then

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

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

Lemma 4.2 *With respect to the positive definite bilinear form* (4.1), *each* ad_X, $X \in \mathfrak{p}_\sigma$ *is symmetric and each* ad_Y, $Y \in \mathfrak{k}_\sigma$, *is skew-symmetric.*

Proof. See Helgason (1978), Lemma 1.2, p. 253. □

Let $y \in G$. We say that y *lies* in the *normal domain* 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). Note that if y is in such a domain, then $d(y, e) = \|Y\| = (\langle Y, Y\rangle)^{\frac{1}{2}}$. 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 polar decomposition of Theorem 3.1 exists). Moreover, assume that x and xy^{-1} are in the normal domain of the identity, otherwise arbitrary. Then*

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

in the right invariant norm induced by (4.1).

Proof. Since x is in the fundamental domain of the exponential mapping, 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 $\|X\| \le \|Z\|$.

Let $Z = P + K$ be the Cartan decomposition of Z in $\mathfrak{p}_\sigma \oplus \mathfrak{k}_\sigma$ and set $z(t) = \exp(tZ)$. By virtue of Theorem 3.1 we can perform the 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 $X(t) \in \mathfrak{p}_\sigma$ 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

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

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 $X \in \mathfrak{p}_\sigma$, $W, Z \in \mathfrak{g}$, 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,

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

hence

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

from which we deduce

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

thus

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

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

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

which completes our proof. $\square$

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 neighbourhood of the identity.

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

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

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 polar decomposition of Theorem 3.1.

Proof. Let $\tilde{y}$ any differetiable element in G^σ other than y, such that $\tilde{y}(0) = e$. Now, for the G right-invariant norm 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 assert follows directly from Lemma 4.3 $\square$

5 Conclusions

In this paper we have presented a constructive proof of the polar decomposition in Lie groups as analogous of the polar decomposition of matrices. Such factorization always exists and its factors are differentiable, provided that the group element we wish to factorize is sufficiently close to the group identity. At the moment it is not clear whether the such decomposition exists globally: Lie groups as manifolds can possess very complex topological features, which might for instance undermine the uniqueness of the factorization for group elements further away from the identity. Lawson (1994) shows that in its most general form, this decomposition does not exist globally, but for a "large" semigroup of the group G, claiming also that the semigroup setting is the most appropriate to study the polar decomposition.

Although the result of the existence and uniqueness of the polar decomposition is not new and originally due to Lawson, the existence of such decomposition in such a general setting is very relevant, among others, in the context of integration of dynamical systems possessing geometric attributes as symmetries and reversing symmetries. Numerical integrators for differential equations usually introduce perturbation of the underlying vector field, so that the numerical flow (numerical approximation thought as the flow of a modified vector field) seldom preserves the same symmetries/reversing symmetries and ad hoc numerical schemes need be introduced instead or projection technique 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 (McLachlan et al. 1998) corresponds to an approximation of the $\mathfrak{p}_\sigma$-factor of the polar decomposition introduced in this paper, under the choice of an opportune automorphism σ (Munthe-Kaas et al. 2000).

It would also be of great interest to investigate whether linear algebra techniques for computing the polar factorization of matrices can be extended to some respect to the general Lie-group case.

Acknowledgements. H. 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. Among them, special thanks to Stig Faltinsen and Erik van den Ban.

6 Appendix

With the same notation of Theorem 3.1, we prove that the series $\sum_{k=1}^{\infty} X_k$, with the X_ks as in (3.2), is absolutely converging.

Let μ 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

$$(k+1)\|X_{k+1}\| \leq \alpha \left(\mu \|X_k\| + \sum_{\substack{\ell \geq 1 \\ 2\ell \leq k}} |c_{2\ell}| \mu^{2\ell} \sum_{\substack{\ell_1, \ldots, \ell_{2\ell} > 0 \\ \ell_1 + \cdots + \ell_{2\ell} = k}} \|X_{\ell_1}\| \cdot \|X_{\ell_2}\| \cdots \|X_{\ell_{2\ell}}\| \right). \tag{6.1}$$

Consider next the differential equation

$$\frac{\mathrm{d}w}{\mathrm{d}u} = h(w), \quad w(0) = 0, \tag{6.2}$$

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_ks are positive and obey the recurrence relation

$$(k+1)w_{k+1} = w_k + \sum_{\substack{\ell \geq 1 \\ 2\ell \leq k}} |c_{2\ell}| \sum_{\substack{\ell_1,\ldots,\ell_{2\ell}>0 \\ \ell_1+\cdots+\ell_{2\ell}=k}} w_{\ell_1} w_{\ell_2} \cdots w_{\ell_{2\ell}}, \quad k = 1, 2, \ldots, \tag{6.3}$$

with starting value $w_1 = 1$. We claim that $\|X_k\| \leq \mu^{k-1}\alpha^k w_k$, for $k = 1, 2, \ldots$. Clearly, the statement is true for $k = 1$, since $\|X_1\| = \|P\| \leq \alpha$ and $w_1 = 1$. Next, assume that the statement is true for $m = 1, 2, \ldots, k$. From the induction hypothesis, together with (6.1) and (6.3), we deduce

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

It follows that the series $\sum_{k=1}^{\infty} \|X_k\|$ is converging in the disc of radius $\delta/(\alpha\mu)$, being bounded by an absolute converging series. Hence, there exists a sufficiently small t such that $X(t) = \sum_{k=1}^{\infty} t^k X_k$ is absolutely converging. This completes the proof of the Theorem 3.1.

References

Abramowitz, M. & Stegun, I. A. (1965), *Handbook of Mathematical Functions*, Dover, New York.

Autonne, L. (1902), 'Sur les groupes linéaires, réel et orthogonaux', *Bulletin de la Société Mathématique de France* **30**, 121–134.

Fan, K. & Hoffman, A. J. (1955), 'Some metric inequalities in the space of matrices', *Proc. Amer. Math. Soc.* **6**, 111–116.

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

Higham, N. J. (1986), 'Computing the polar decomposition—with applications', *SIAM J. Sci. Stat. Comput.* **7**, 1160–1174.

Hilgert, J., Hoffmann, K. H. & Lawson, J. D. (1989), *Lie Groups, convex Cones and Semigroups*, Oxford University Press.

Horn, R. A. & Johnson, C. R. (1991), *Topics in Matrix Analysis*, Cambridge University Press.

Lawson, J. D. (1994), 'Polar and Ol'shankii decompositions', *J. Reine Angew. Math.* **448**, 191–219.

Loos, O. (1969), *Symmetric Spaces I: General Theory*, W. A. Benjamin, Inc.

McLachlan, R. I., Quispel, G. R. W. & Turner, G. S. (1998), 'Numerical integrators that preserve symmetries and reversing symmetries', *SIAM J. Numer. Anal.* **35**(2), 586–599.

Munthe-Kaas, H., Quispel, G. R., Nikolayevsky, Y. & Zanna, A. (2000), Application of symmetric spaces and Lie triple systems in Numerical Analysis, In preparation.

Varadarajan, V. S. (1984), *Lie Groups, Lie Algebras, and Their Representation*, GTM 102, Springer-Verlag.

Zanna, A. (2000), Recurrence relation for the factors in the polar decomposition on Lie groups, Technical Report Report no. 192, Department of Informatics, University of Bergen.