Monografías de la Real Academia de Ciencias de Zaragoza **33**, 113–124, (2010).

On the convergence of generalized polar decompositions in Lie groups

Jordi P. García-Seguí and Fernando Casas

Institut de Matemàtiques i Aplicacions de Castelló (IMAC) and

Dept. de Matemàtiques, Universitat Jaume I, 12071 Castellón, Spain.

This paper is dedicated to Prof. Manuel Calvo on the occasion of his 65th anniversary

Abstract

We analyze the so-called generalized polar decomposition determined by an involutive automorphism in a Lie group. This concept generalizes the well known factorization of a matrix as the product of a positive semidefinite matrix and an orthogonal matrix in linear algebra. We provide a different constructive proof of the existence of such a decomposition in a neighborhood of the identity and obtain several explicit bounds on the convergence domain of the series defined each factor.

1 Introduction

The polar decomposition can be seen as the matrix analog of the polar form of a complex number $z = r\,\mathrm{e}^{i\theta}$, $r > 0$. If A is any $n \times n$ matrix, then there exists a unitary matrix U and a unique Hermitian positive semidefinite matrix H such that

$$A = H\,U.$$

Furthermore, if A is invertible, then H is positive definite and U is uniquely determined. If A is real, the matrix U is orthogonal and H is symmetric. It is well known that the factors H and U possess best approximation properties. Specifically, the polar factor U is the best unitary (orthogonal in the real case) approximant to A in any unitarily invariant norm, whereas H is a good Hermitian positive definite approximation to A when it is nonsingular and Hermitian, and $\frac{1}{2}(A + H)$ is a best Hermitian positive semidefinite approximation to A [8].

The polar decomposition has been generalized to abstract Lie groups and even to semigroups [11]. Such generalized polar decomposition has been found to be closely related with the concept of involutive automorphism and the subspace decomposition it induces. In this setting, the polar decomposition is equivalent to expressing a group element as the product of a term in a symmetric subspace and a term in a subgroup of the given Lie group.

More specifically, let G be a Lie group and $\sigma : G \longrightarrow G$ an involutive automorphism. By this we mean a one-to-one map such that $\sigma(xy) = \sigma(x)\sigma(y)$, $\sigma \neq \mathrm{id}$ and $\sigma^2 = \mathrm{id}$. Let G^σ denote the subgroup of G consisting of fixed points of σ, i.e., $G^\sigma = \{x \in G : \sigma(x) = x\}$ 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 not a group, but a symmetric space with the non-associative multiplication $x \cdot y \equiv xy^{-1}x$ [7]. Then, the generalized polar decomposition of $z \in G$ consists in writing

$$z = xy, \qquad x \in G_\sigma, \quad y \in G^\sigma. \tag{1}$$

Since σ induces an involutive automorphism $d\sigma$ on the Lie algebra $\mathfrak{g}$ corresponding to G as

$$d\sigma(X) \equiv \frac{d}{dt}\Big|_{t=0} \sigma(\exp(tX))$$

for all $X \in \mathfrak{g}$, then $\mathfrak{g}$ can be expressed as the direct sum

$$\mathfrak{g} = \mathfrak{p} \oplus \mathfrak{k}, \tag{2}$$

where $\mathfrak{k}$ corresponds to the set of fixed points of $d\sigma$ ($d\sigma(X) = X$) and $\mathfrak{p}$ to the set of anti-fixed points, $d\sigma(X) = -X$. The space $\mathfrak{k}$ is a subalgebra of $\mathfrak{g}$, whereas $\mathfrak{p}$ is a Lie triple system: $\mathfrak{p}$ is a vector space that is not closed under the commutator but under the double commutator, that is, $[X_1, [X_2, X_3]] \in \mathfrak{p}$ for $X_i \in \mathfrak{p}$, whereas $[X_1, X_2] \in \mathfrak{k}$ [7]. In general, the sets $\mathfrak{p}$ and $\mathfrak{k}$ verify the following commutation relations:

$$[\mathfrak{k}, \mathfrak{k}] \subseteq \mathfrak{k}, \qquad [\mathfrak{k}, \mathfrak{p}] \subseteq \mathfrak{p}, \qquad [\mathfrak{p}, \mathfrak{p}] \subseteq \mathfrak{k}.$$

As a result, every element $Z \in \mathfrak{g}$ can be uniquely written as

$$Z = P + K, \qquad P \in \mathfrak{p}, \quad K \in \mathfrak{k} \tag{3}$$

with

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

Moreover, if $K \in \mathfrak{k}$, then $\exp(tK) \in G^\sigma$, whereas $P \in \mathfrak{p}$ implies $\exp(tP) \in G_\sigma$.

As a well know example, let us consider the general linear group $\mathrm{GL}(n)$ of real $n \times n$ invertible matrices and the map $\sigma(x) = (x^{-1})^T$, which is an involutive automorphism. Now the set G_σ is the set of invertible symmetric matrices (a symmetric space), whereas

G^σ is the set of orthogonal matrices (which is a subgroup of $\mathrm{GL}(n)$). It can be checked that $d\sigma(X) = -X^T$, whence $\mathfrak{k}$ is the classical algebra of skew-symmetric matrices and $\mathfrak{p}$ is the set of symmetric matrices. In consequence, the decomposition (3) is nothing but the canonical decomposition of a matrix into its skew-symmetric and symmetric part: $P = (Z + Z^T)/2$, $K = (Z - Z^T)/2$.

In a series of papers [13, 15, 16], Zanna and his collaborators have analyzed the polar decomposition in a generic Lie group G. In particular, they provide a proof of its existence and uniqueness in a neighborhood of the identity $e \in G$, which can be established as the following theorem [13].

Theorem 1.1 *Let $z = \exp(tZ) \in G$, where $Z = P + K$ is the decomposition of Z in $\mathfrak{p} \oplus \mathfrak{k}$, i.e., $d\sigma(P) = -P$ and $d\sigma(K) = K$. Then, for sufficiently small values of t, the element z admits a unique generalized polar decomposition $z = xy$, where $x = \exp(X(t))$, $X(t) \in \mathfrak{p}$, and $y = \exp(Y(t))$ with $Y(t) \in \mathfrak{k}$.*

Moreover, they derive differential equations obeyed by $X(t)$ and $Y(t)$ and solve them perturbatively, thus constructing X and Y as a power series whose terms can be obtained by a recursive procedure. These recurrences are in turn used to prove the convergence of the series when $\mathfrak{g}$ is a Banach algebra. In this way, the function $X(t)$ is shown to be analytic in a sphere of radius

$$\rho = \frac{\delta}{2\alpha} \qquad \text{for some constant} \quad 0 < \delta < \pi \tag{4}$$

and $\alpha = \max\{\|P\|, \|K\|\}$ [13], although no specific value of δ is provided. On the other hand, the radius of convergence of the series $Y(t)$ is given implicitly as $\rho = \frac{r}{2\beta}$ [15], where $\beta = \max\{t\|Z\|, \|X(t)\|\}$ and r is related to the radius of convergence of the Baker–Campbell–Hausdorff (BCH) series. Notice that these estimates are all of a qualitative nature, whereas (at least up to our knowledge) no actual bounds for the convergence domain are found in the literature.

In this paper we try to fill this gap by first proposing new computationally well adapted recurrences for generating the series $X(t)$ and $Y(t)$. These recurrences are used to get numerical estimates on the convergence of the series $X(t)$ and also a bound on $\|X(t)\|$ itself, which is then used to establish the convergence of the series $Y(t)$. These results are supplemented with sharper numerical estimates obtained from the BCH series.

Although of theoretical nature, generalized polar decompositions in Lie groups have found interesting applications in numerical analysis, namely in connection with self-adjoint numerical integrators for differential equations [10] and the numerical approximation of the exponential of a matrix from a Lie algebra to a Lie group [16], especially in $\mathrm{SL}(n)$. From a more abstract point of view, they constitute a particular instance of the Atkinson factorization theorem for Rota–Baxter algebras [3, 5]. We believe that the convergence results provided here will be of interest in these different settings.

2 Recursion for the factor X

Our starting point is the factorization provided by Theorem 1.1

$$\mathrm{e}^{tZ} = \mathrm{e}^{X(t)}\,\mathrm{e}^{Y(t)}, \tag{5}$$

with $Z = P + K$. Differentiating (5) one arrives at the expression

$$\mathrm{e}^{-X}(Z - d\exp_X(X'))\,\mathrm{e}^{X} = d\exp_Y(Y'), \tag{6}$$

where

$$d\exp_Y(Y') \equiv \sum_{j=0}^{\infty} \frac{1}{(j+1)!}\mathrm{ad}_Y^j(Y') \in \mathfrak{k} \tag{7}$$

since $Y, Y' \in \mathfrak{k}$ and $\mathfrak{k}$ is a subalgebra of the Lie algebra $\mathfrak{g}$. Here ad_A stands for the adjoint operator of $A \in \mathfrak{g}$, which acts according to

$$\mathrm{ad}_A B = [A, B], \qquad \mathrm{ad}_A^j B = [A, \mathrm{ad}_A^{j-1} B], \qquad \mathrm{ad}_A^0 B = B, \qquad j \in \mathbb{N}, B \in \mathfrak{g}. \tag{8}$$

Notice that the left hand side of eq. (6) also belongs to $\mathfrak{k}$. We therefore analyze this term and separate the contribution in $\mathfrak{p}$, which has to be canceled.

First we note that

$$\begin{aligned} \mathrm{e}^{-X} Z\,\mathrm{e}^{X} &= -\sinh(u)(K) + \cosh(u)(P) && (\in \mathfrak{p}) \\ &\quad + \cosh(u)(K) - \sinh(u)(P) && (\in \mathfrak{k}) \end{aligned}$$

where $u \equiv \mathrm{ad}_X$ and the functions involving u have to be understood as power series. On the other hand,

$$\begin{aligned} \mathrm{e}^{-X} d\exp_X(X')\,\mathrm{e}^{X} &= \frac{1}{u}\sinh(u)(X') && (\in \mathfrak{p}) \\ &\quad + \frac{1}{u}(1 - \cosh(u))(X') && (\in \mathfrak{k}) \end{aligned}$$

In consequence,

$$-\sinh(u)(K) + \cosh(u)(P) - \frac{1}{u}\sinh(u)(X') = 0$$

whence, after some algebra, we arrive at the differential equation satisfied by X:

$$X' = -\mathrm{ad}_X K + \sum_{k=0}^{\infty} \frac{2^{2k} B_{2k}}{(2k)!}\mathrm{ad}_X^{2k} P, \qquad X(0) = 0, \tag{9}$$

with B_j denoting the Bernoulli numbers [1]. To solve equation (9), let us introduce a parameter $\epsilon > 0$ in Z and consider instead $\epsilon Z = \epsilon(K + P)$, i.e., the decomposition

$$\mathrm{e}^{t\,\epsilon\,Z} = \mathrm{e}^{X(\epsilon,t)}\,\mathrm{e}^{Y(\epsilon,t)}.$$

The corresponding equation satisfied by $X(\epsilon,t)$ is then

$$\frac{\partial X}{\partial t} = -\epsilon\,\mathrm{ad}_X K + \sum_{k=0}^{\infty} c_{2k}\mathrm{ad}_X^{2k}(\epsilon P), \qquad X(\epsilon,0)=0, \tag{10}$$

where, for simplicity, $c_{2k} = \frac{2^{2k}B_{2k}}{(2k)!}$. Now we try to determine the solution $X(\epsilon,t)$ perturbatively as an infinite series in ϵ,

$$X(\epsilon,t) = \sum_{n=1}^{\infty} \epsilon^n X_n(t). \tag{11}$$

To do that, first we substitute expression (11) into (10), thus obtaining for each terms up to order ϵ^n the expressions

$$\begin{aligned}
\frac{\partial}{\partial t}X(\epsilon,t) &= \sum_{j=1}^{n} \epsilon^j X_j'(t) + \mathcal{O}(\epsilon^n)\\
\mathrm{ad}_X(\epsilon K) &= \sum_{j=1}^{n-1} \epsilon^{j+1}\mathrm{ad}_{X_j}K + \mathcal{O}(\epsilon^{n+1})\\
\sum_{j=1}^{n-1} c_j\,\mathrm{ad}_X^j(\epsilon P) &= \sum_{l=2}^{n}\epsilon^l \sum_{j=1}^{l-1} c_j \sum_{\substack{k_1+\cdots+k_j=l-1\\ k_1\ge 1,\dots,k_j\ge 1}} \mathrm{ad}_{X_{k_1}}\cdots\mathrm{ad}_{X_{k_j}}P + \mathcal{O}(\epsilon^{n+1}).
\end{aligned}$$

Then, by equating successive powers of ϵ, we get

$$\begin{aligned}
X_1' &= P\\
X_l' &= -\mathrm{ad}_{X_{l-1}}K + \sum_{j=2}^{l-1} c_j \sum_{\substack{k_1+\cdots+k_j=l-1\\ k_1\ge 1,\dots,k_j\ge 1}} \mathrm{ad}_{X_{k_1}}\cdots\mathrm{ad}_{X_{k_j}}P, \qquad l\ge 2.
\end{aligned}$$

From the initial condition, it is clear that $X_l(0)=0$ for all $l\ge 1$, so that finally we arrive at the recursion

$$\begin{aligned}
X_1(t) &= tP \qquad\qquad (12)\\
X_l(t) &= -\int_0^t \mathrm{ad}_{X_{l-1}}K\,ds + \sum_{j=2}^{l-1} c_j \sum_{\substack{k_1+\cdots+k_j=l-1\\ k_1\ge 1,\dots,k_j\ge 1}} \int_0^t \mathrm{ad}_{X_{k_1}}\cdots\mathrm{ad}_{X_{k_j}}P\,ds, \qquad l\ge 2.
\end{aligned}$$

If this recurrence is worked out explicitly, one gets for the first terms

$$\begin{aligned}
&X_2(t) = -\tfrac{t^2}{2}[P,K], \qquad X_3(t) = -\tfrac{t^3}{6}[K,[P,K]],\\
&X_4(t) = \tfrac{t^4}{24}([P,[P,[P,K]]]-[K,[K,[P,K]]])
\end{aligned}$$

3 Recursion for the factor Y

By considering the projection of equation (6) into $\mathfrak{k}$ we have

$$\cosh(u)(K) - \sinh(u)(P) + \frac{\cosh(u)-1}{u}(X') = d\exp_Y(Y'), \tag{13}$$

where, as before, $u \equiv \mathrm{ad}_X$. Inserting equation (9) into (13) results in

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

Taking into account that

$$d\exp_Y^{-1}(Y') = \sum_{j=0}^{\infty} \frac{B_j}{j!} \mathrm{ad}_Y^j(Y')$$

and the power series of the function $(1-\cosh(u))/\sinh(u)$, we get finally

$$Y' = \sum_{j=0}^{\infty} \frac{B_j}{j!} \mathrm{ad}_Y^j \left(K - 2\sum_{k=2}^{\infty} \frac{(2^k-1)B_k}{k!} \mathrm{ad}_X^{k-1}(P) \right), \qquad Y(0) = 0. \tag{14}$$

Notice that solving for $Y(t)$ requires to previously compute $X(t)$. In spite of that, in the sequel we show that it is indeed possible to construct a power series for $Y(t)$ by recurrence. We proceed in a similar way as for the X factor: introduce the parameter $\epsilon > 0$ in Z and determine the successive terms in the expansion

$$Y(\epsilon, t) = \sum_{n=1}^{\infty} \epsilon^n Y_n(t) \tag{15}$$

by inserting it into the corresponding differential equation

$$\frac{\partial Y}{\partial t} = \epsilon \, d\exp_Y^{-1} D, \qquad Y(\epsilon, 0) = 0, \tag{16}$$

where

$$D \equiv K - 2\sum_{k=2}^{\infty} d_k \, \mathrm{ad}_X^{k-1}(P) \qquad \text{and} \qquad d_k = \frac{(2^k-1)B_k}{k!}.$$

It can be shown after some elementary algebra that the r.h.s. of equation (16) can be written as

$$\epsilon \, d\exp_Y^{-1} D = \epsilon K + A + B + C + \mathcal{O}(\epsilon^{n+1})$$

with

$$\begin{aligned}
A &= -2\sum_{l=2}^{n} \epsilon^l \sum_{j=1}^{l-1} d_{j+1} \sum_{\substack{k_1+\cdots+k_j=l-1 \\ k_1\geq 1,\ldots,k_j\geq 1}} \mathrm{ad}_{X_{k_1}} \cdots \mathrm{ad}_{X_{k_j}} P \\
B &= \sum_{l=2}^{n} \epsilon^l \sum_{j=1}^{l-1} \frac{B_j}{j!} \sum_{\substack{k_1+\cdots+k_j=l-1 \\ k_1\geq 1,\ldots,k_j\geq 1}} \mathrm{ad}_{Y_{k_1}} \cdots \mathrm{ad}_{Y_{k_j}} K \\
C &= -2\sum_{l=3}^{n} \epsilon^l \sum_{j=2}^{l-1} \left(\sum_{m=1}^{j-1} \frac{B_m}{m!} \sum_{\substack{k_1+\cdots+k_m=j-1 \\ k_1\geq 1,\ldots,k_m\geq 1}} \mathrm{ad}_{Y_{k_1}} \cdots \mathrm{ad}_{Y_{k_m}} \right) \\
&\quad \left(\sum_{p=1}^{l-j} d_{p+1} \sum_{\substack{r_1+\cdots+r_p=l-j \\ r_1\geq 1,\ldots,r_p\geq 1}} \mathrm{ad}_{X_{r_1}} \cdots \mathrm{ad}_{X_{r_p}} P \right)
\end{aligned} \tag{17}$$

Equating powers of ϵ leads one to the recursion

$$\begin{aligned}
Y_1(t) &= t\,K \\
Y_n(t) &= \sum_{j=1}^{n-1} \frac{B_j}{j!} \sum_{\substack{k_1+\cdots+k_j=n-1\\ k_1\geq 1,\ldots,k_j\geq 1}} \int_0^t \mathrm{ad}_{Y_{k_1}} \cdots \mathrm{ad}_{Y_{k_j}} K\, ds \\
&\quad -2\sum_{j=1}^{l-1} d_{j+1} \sum_{\substack{k_1+\cdots+k_j=n-1\\ k_1\geq 1,\ldots,k_j\geq 1}} \int_0^t \mathrm{ad}_{X_{k_1}} \cdots \mathrm{ad}_{X_{k_j}} P\, ds \qquad (18)\\
&\quad -2\sum_{j=2}^{n-1} \int_0^t d\tau \left(\sum_{s=1}^{j-1} \frac{B_s}{s!} \sum_{\substack{k_1+\cdots+k_s=j-1\\ k_1\geq 1,\ldots,k_s\geq 1}} \mathrm{ad}_{Y_{k_1}} \cdots \mathrm{ad}_{Y_{k_s}} \right) \\
&\quad \left(\sum_{p=1}^{n-j} d_{p+1} \sum_{\substack{r_1+\cdots+r_p=n-j\\ r_1\geq 1,\ldots,r_p\geq 1}} \mathrm{ad}_{X_{r_1}} \cdots \mathrm{ad}_{X_{r_p}} P \right) \qquad n\geq 2
\end{aligned}$$

which allows us to get the explicit expression of the first terms as

$$Y_2(t) = 0, \qquad Y_3(t) = -\frac{t^3}{12}[P,[P,K]], \qquad Y_4(t) = 0.$$

As a matter of fact, it is not difficult to prove that $Y(t)$ is and odd function of t, so that in general $Y_{2n}(t) = 0$ for all n. Notice that it is necessary to previously generate the terms X_i through recurrence (12) to obtain the series $Y(t)$ by (18). Although it is indeed possible to derive another recursion involving only terms Y_i, we have found the recursion (18) more convenient not only from a computational point of view (the implementation in a symbolic package is rather straightforward) but also for establishing explicit convergence domains for the series.

4 Convergence of the expansions

We next analyze the convergence of the previous series. For that purpose we assume that $\mathfrak{g}$ is a complete normed Lie algebra endowed with a norm compatible with associative multiplication, i.e., such that $\|AB\| \leq \|A\|\,\|B\|$ for all A, B in $\mathfrak{g}$. Then it is true that

$$\|[A,B]\| \leq 2\|A\|\,\|B\|.$$

First we consider the series

$$v(\varepsilon,t) = \sum_{j=1}^{\infty} \epsilon^j \|X_j(t)\|. \qquad (19)$$

From (12) it is clear that for $l \geq 2$

$$\|X_l(t)\| \leq 2\|K\| \int_0^t \|X_{l-1}\| ds + \|P\| \sum_{j=2}^{l-1} |c_j| 2^j \sum_{\substack{k_1+\cdots+k_j=l-1\\ k_1\geq 1,\ldots,k_j\geq 1}} \int_0^t \|X_{k_1}(s)\| \cdots \|X_{k_j}(s)\| ds$$

and thus

$$\sum_{l=2}^{N} \epsilon^l \|X_l(t)\| \le \sum_{l=2}^{N} 2\epsilon^l \|K\| \int_0^t \|X_{l-1}\| ds$$
$$+ \epsilon \|P\| \sum_{j=1}^{N-1} |c_j| 2^j \sum_{l=p}^{N-1} \varepsilon^l \sum_{\substack{k_1+\cdots+k_p=l \\ k_1\ge 1,\ldots,k_p\ge 1}} \int_0^t \|X_{k_1}(s)\| \cdots \|X_{k_j}\| ds,$$

where we have interchanged the order of summation in the second term.

Let us denote $v_N(\epsilon, t) = \sum_{l=1}^{N} \epsilon^l \|X_l(t)\|$. Then it is easy to show that

$$(v_N(\epsilon,t))^p = \sum_{l=p}^{pN} \epsilon^l \sum_{\substack{k_1+\cdots+k_p=l \\ k_1\ge 1,\ldots,k_p\ge 1}} \|X_{k_1}\| \cdots \|X_{k_p}\|$$

so that, in the last inequality,

$$\sum_{l=p}^{N-1} \epsilon^l \sum_{\substack{k_1+\cdots+k_p=l \\ k_1\ge 1,\ldots,k_p\ge 1}} \|X_{k_1}\| \cdots \|X_{k_p}\| \le (v_N(\epsilon,t))^p$$

and therefore

$$v_N(\epsilon,t) \le 2\epsilon \|K\| \int_0^t v_{N-1}(\epsilon,s) ds + \epsilon \|P\| \sum_{j=0}^{N-1} |c_j| 2^j \int_0^t v_N(\epsilon,s)^j \, ds.$$

Taking the limit $N \to \infty$ in the last expression we have

$$v(\epsilon,t) \le 2\epsilon \|K\| \int_0^t v(\epsilon,s) ds + \epsilon \|P\| \int_0^t g(2v(\epsilon,s)))\, ds \tag{20}$$

since

$$\sum_{j=0}^{\infty} |c_j| \, (2x)^j = \sum_{j=0}^{\infty} \frac{|B_j|}{j!} (2x)^j = 2 + x(1-\cot x) \equiv g(x). \tag{21}$$

We proceed now as follows. Let us denote $k = \|K\|$ and $p = \|P\|$ and introduce the function $G(x) = \beta x + g(x)$, with $\beta = k/p \ge 0$. Then (20) can be written as

$$v(\epsilon,t) \le \epsilon\, p \int_0^t G(2v(\epsilon,s)) ds \equiv F(\epsilon,t).$$

In this way

$$\frac{\partial F(\epsilon,t)}{\partial t} = \epsilon\, p\, G(2v(\epsilon,t)) \le \epsilon\, p\, G(2F(\epsilon,t))$$

since G is a non-decreasing function on the domain $[0, \pi)$. In fact $G(z)$ is analytic for $|z| < \pi$ with positive coefficients in the power series and has no zeros in the ball $|z| < \pi$. The last inequality can be expressed as

$$\frac{\partial F(\epsilon,t)}{\partial t} \frac{1}{G(2F(\epsilon,t))} \le \epsilon\, p$$

so that, by integrating, we get

$$H(2F(\epsilon,t)) \leq 2\,\epsilon\, p\, t$$

where $H(t) \equiv \int_0^t \frac{1}{G(x)}dx$. Now $H(z)$ is also analytic in $|z| < \pi$ and $H'(z) = \frac{1}{G(z)} \neq 0$. Then $y = H(z)$ has an inverse function $z = H^{-1}(y)$ for y in the ball $|y| < H(\pi)$, which is also analytic there. In consequence,

$$v(\epsilon,t) \leq F(\epsilon,t) \leq \frac{1}{2} H^{-1}(2\,\epsilon\, p\, t)$$

for t such that $2\,\epsilon\, p\, t$ belongs to the domain of H^{-1}, i.e.,

$$2\,\epsilon\, p\, t < H(\pi) = \int_0^\pi \frac{1}{G(x)}dx \equiv \xi(\beta).$$

Therefore the series $X(t)$ is assured to be convergent for $0 \leq t \leq t_c$, with

$$t_c \equiv \frac{1}{2\,\epsilon\, p}\xi(\beta) = \frac{1}{2\,\epsilon\, p}\int_0^\pi \frac{1}{2 + (1+\beta)x - x\cot x}dx. \tag{22}$$

If we take $\epsilon = 1$ in (19), then

$$v(\epsilon = 1) = \|X(t)\| = \sum_{n=1}^{\infty} \|X_n(t)\| < \frac{1}{2}H^{-1}(\xi(\beta)) = \frac{\pi}{2} \tag{23}$$

in the convergence domain defined by (22).

For illustration, we collect next the values of $\xi(\beta)$ for several values of β:

β	0	1	10
$\xi(\beta)$	1.08687	0.83751	0.31228

If instead of using the norm ratio β we work with $\alpha = \max\{k,p\}$ then a similar argument shows that the series $\|X(t)\|$ is convergent for $0 \leq t \leq t_c$ with

$$t_c = \frac{\xi(1)}{2\alpha} \simeq \frac{0.83751}{2\alpha}. \tag{24}$$

Notice that we have obtained a numerical value for the constant δ in (4).

A enlarged convergence domain can indeed be established by means of the Baker–Campbell–Hausdorff (BCH) formula. As is well known, the BCH formula deals with the expansion of Z in $\mathrm{e}^{X_1}\,\mathrm{e}^{X_2} = \mathrm{e}^{Z}$ in terms of nested commutators of X_1 and X_2 when they are assumed to be non-commuting operators. Specifically,

$$Z = X_1 + X_2 + \sum_{n=2}^{\infty} G_n(X_1, X_2), \tag{25}$$

where $G_n(X_1,X_2)$ is a homogeneous Lie polynomial in X_1 and X_2 of grade n; in other words, G_n can be expressed in terms of X_1 and X_2 by addition, multiplication by rational numbers and nested commutators. This result proves to be very useful in various fields of mathematics (theory of linear differential equations [12], Lie group theory [4], numerical analysis [6]) and theoretical physics (perturbation theory, transformation theory, Quantum Mechanics and Statistical Mechanics. In particular, in the theory of Lie groups, with this theorem one can explicitly write the operation of multiplication in a Lie group in canonical coordinates in terms of the Lie bracket operation in its algebra and also prove the existence of a local Lie group with a given Lie algebra [4].

The following theorem concerning the convergence of the BCH series has been proved (see [2]).

Theorem 4.1 *Let X_1 and X_2 be two bounded elements in a Hilbert space $\mathcal{H}$ with $\dim \mathcal{H} \geq 2$. Then the BCH formula in the form (25), i.e., expressed as a series of homogeneous Lie polynomials in X_1 and X_2, converges absolutely when $\|X_1\| + \|X_2\| < \pi$.*

Here the norm is taken as the 2-norm induced by the scalar product in $\mathcal{H}$. This result can be generalized, of course, to any number of non commuting operators $X_1, X_2, \ldots, X_q$. Specifically, the series

$$Z = \log(\mathrm{e}^{X_1}\, \mathrm{e}^{X_2} \, \cdots \, \mathrm{e}^{X_q}),$$

converges absolutely if $\|X_1\| + \|X_2\| + \cdots + \|X_q\| < \pi$.

We next show how this result can be used to get a sharper bound on the convergence domain of X. As usual, we set $z = \exp(tZ)$ with $Z = P + K$ the decomposition of Z into $\mathfrak{p} \oplus \mathfrak{k}$, and denote $w \equiv (\sigma(z))^{-1} = \sigma(z^{-1})$. Then it is true that $w = \exp(tW)$, with $W = P - K$. Now, since $\sigma(x) = x^{-1}$ and $\sigma(y) = y$ in the generalized polar decomposition $z = xy$, it is clear that

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

so that

$$\mathrm{e}^{2X(t)} = \mathrm{e}^{tZ}\, \mathrm{e}^{tW}. \tag{26}$$

As a matter of fact, it is possible to apply the algorithm proposed in [2] to generate the series $X(t) = \frac{1}{2}\log(\exp(tZ)\exp(tW))$ in an arbitrary generalized Hall basis of the free Lie algebra generated by P and K. Applying now Theorem 4.1 we conclude that the series $X(t)$ is convergent as long as $t(\|Z\| + \|W\|) < \pi$ or equivalently, when $0 \leq t < t_{bch}$, with

$$t_{bch} = \frac{\pi}{\|P+K\| + \|P-K\|}. \tag{27}$$

This estimate can be compared with (24) by taking into account that

$$\|P+K\| \leq p + k \leq 2\alpha, \qquad \|P-K\| \leq 2\alpha$$

where, as before, $\alpha = \max\{p, k\}$. In consequence, $t_{bch} > \pi/(4\alpha)$ and the convergence of $X(t)$ is guaranteed for

$$t \leq \frac{\pi}{4\alpha}.$$

Unfortunately, no bound for $\|X(t)\|$ in this domain can be obtained from the BCH series, contrarily to the estimate (23), valid when $t < t_c$.

Once $x = \exp(X(t))$ is known, one has $y = x^{-1}z$, or

$$\mathrm{e}^{Y} = \mathrm{e}^{-X(t)}\,\mathrm{e}^{tZ}.$$

Therefore, the series $Y(t)$ converges if $\|X(t)\| + t\|Z\| < \pi$. For $t < t_c$, we have shown that $\|X(t)\| < \pi/2$, so that one has convergence if $t\|Z\| < \pi/2$ or

$$t < \frac{\pi}{2\|P + K\|},$$

which is also greater than $\pi/(4\alpha)$. We then conclude that the generalized polar decomposition (5) exists with analytic functions $X(t)$ and $Y(t)$ at least for $t < t_c = \xi(1)/(2\alpha)$, whereas the series $X(t)$ is absolutely convergent for $0 \leq t < t_{bch}$.

Acknowledgements

This work has been partially supported by Ministerio de Ciencia e Innovación (Spain) under project MTM2007-61572 (co-financed by the ERDF of the European Union) and Fundació Bancaixa through project P1.1B2009-55. FC would like especially to thank the Instituto de Matemática Aplicada de la Universidad de Zaragoza (IUMA) for inviting him to participate in the workshop celebrating Manuel Calvo's birthday in September 2009.

References

[1] M. Abramowitz and I.A. Stegun. *Handbook of Mathematical Functions*. Dover, 1965.

[2] F. Casas and A. Murua. An efficient algorithm for computing the Baker–Campbell–Hausdorff series and some of its applications. *J. Math. Phys.* **50** (2009), 033513 (23 pages).

[3] K. Ebrahimi-Fard, J.M. Gracia-Bondía, and F. Patras. Rota–Baxter algebras and new combinatorial identities. *Lett. Math. Phys.* **81** (2007), 61–75.

[4] V.V. Gorbatsevich, A.L. Onishchik, and E.B. Vinberg. *Foundations of Lie Theory and Lie Transformation Groups*. Springer-Verlag, 1997.

[5] L. Guo. What is a Rota–Baxter algebra? *Notices of the AMS* **56** (2009), 1436–1437.

[6] E. Hairer, Ch. Lubich, and G. Wanner. *Geometric Numerical Integration. Structure-Preserving Algorithms for Ordinary Differential Equations*. Springer-Verlag, Second edition, 2006.

[7] S. Helgason. *Differential Geometry, Lie Groups, and Symmetric Spaces.* American Mathematical Society, 2001.

[8] N.J. Higham. Computing the polar decomposition. *SIAM J. Sci. Stat. Comput.* **7** (1986), 1160–1174.

[9] A. Iserles, H.Z. Munthe-Kaas, S.P. Nørsett, and A. Zanna. Lie-group methods. *Acta Numerica* **9** (2000), 215–365.

[10] S. Krogstad, H.Z. Munthe-Kaas, and A. Zanna. Generalized polar coordinates on Lie groups and numerical integrators. *Tech. Rep. 244, Department of Informatics, University of Bergen*, 2003.

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

[12] W. Magnus. On the exponential solution of differential equations for a linear operator. *Comm. Pure and Appl. Math.* **7** (1954), 649–673.

[13] H.Z. Munthe-Kaas, G.R.W. Quispel, and A. Zanna. Generalized polar decompositions on Lie groups with involutive automorphisms. *Found. Comput. Math.* **1** (2001), 297–324.

[14] V. S. Varadarajan. *Lie Groups, Lie Algebras, and Their Representations.* Springer-Verlag, 1984.

[15] A. Zanna. Recurrence relations and convergence theory of the generalized polar decomposition on Lie groups. *Math. Comp.* **73** (2004), 761–776.

[16] A. Zanna and H.Z. Munthe-Kaas. Generalized polar decompositions for the approximation of the matrix exponential. *SIAM J. Matrix Anal. Appl.* **23** (2001), 840–862.