A Baker-Campbell-Hausdorff disentanglement relation for Lie groups

1992 J. Phys. A: Math. Gen. 25 1527

(http://iopscience.iop.org/0305-4470/25/6/013)

J. Phys. A: Math. Gen. 25 (1992) 1527-1534. Printed in the UK

A Baker–Campbell–Hausdorff disentanglement relation for Lie groups

K Raghunathan†, M Seetharaman†, S S Vasan† and J Mary Agnes‡

† Department of Theoretical Physics, University of Madras, Guindy Campus, Madras 600025, India

‡ Department of Physics, Meenakshi College for Women, Madras 600024, India

Received 9 July 1991

Abstract. A BCH formula for groups having faithful matrix representations is derived, by a decomposition of the general group element $\exp(M)$ into factors, each factor being an exponential involving the infinitesimal generators of the group. This is a generalization of our earlier work for SU(3).

1. Introduction

In a recent letter (Raghunathan *et al* 1989) we presented a Baker–Campbell–Hausdorff (BCH)-type disentanglement relation for SU(3). This relation consists in expressing the group element $\exp(M)$, M being some linear combination of the SU(3) generators, as a product of exponential factors in the form

$$\exp(M)=\exp(\alpha_{32}E_{32})\exp(\alpha_{31}E_{31})\exp(\alpha_{21}E_{21})\exp(d_1H_1)\exp(d_2H_2) \times\exp(\alpha_{12}E_{12})\exp(\alpha_{13}E_{13})\exp(\alpha_{23}E_{23}). \tag{1.1}$$

Here E_{ij} are a set of three creation and three annihilation operators for SU(3) and H_i are its two diagonal generators. This factorization of the group element is particularly useful in the study of SU(3) coherent states.

In this paper we generalize the above result and present the complete factorization of $\exp(M)$ for any square matrix M. Our result can be regarded as a result in matrix theory. It acquires added importance when viewed from the point of group theory. If we consider M as a given linear combination of the generators of a group in some faithful $n\times n$ matrix representation, then the result we have obtained is a BCH formula (Gilmore 1974a, b) for the group.

Consider $\exp(\lambda M)$, where M is a square matrix of dimension n and λ is a parameter. M can always be written as

$$M=\sum m_{ij}E_{ij} \tag{1.2}$$

where the basis vectors E_{ij} are $n\times n$ matrices whose elements are given by

$$(E_{ij})_{kl}=\delta_{ik}\delta_{jl}. \tag{1.3}$$

We seek a generalization of the relation (1.1)) in the form

$$\exp(\lambda M)=[\exp(\alpha_{n\,n-1}E_{n\,n-1})\exp(\alpha_{n\,n-2}E_{n\,n-2})\dots\exp(\alpha_{32}E_{32}) \times\exp(\alpha_{31}E_{31})\exp(\alpha_{21}E_{21})][\exp(\alpha_{nn}E_{nn})\dots\exp(\alpha_{11}E_{11})] \times[\exp(\alpha_{12}E_{12})\exp(\alpha_{13}E_{13})\exp(\alpha_{23}E_{23})\dots\exp(\alpha_{n-1\,n}E_{n-1\,n})] \tag{1.4}$$

0305-4470/92/061527+08$04.50

where α_{ij} are coefficients. On the RHS of (1.4) there are $n(n-1)/2$ operators E_{ij} with $i>j$ in the first square bracket, n diagonal operators in the middle square bracket and $n(n-1)/2$ operators E_{ij} with $i<j$ in the last square bracket. The main objective of the present work is to establish the result (1.4) by determining the α_{ij}s as functions of the m_{ij}s.

We adopt a three-step procedure. In section 2 we express $\exp(\lambda M)$ as a polynomial in M whose coefficients are functions of the m_{ij}s. Let $G=G(M)$ be the matrix polynomial so obtained. Turning to the RHS of (1.4), we note that it has the form of a product LDU where L is a lower triangular matrix, D is a diagonal matrix and U is an upper triangular matrix corresponding to the first, middle and last square brackets, respectively. In section 3 the elements of L, D and U are determined in terms of the elements of G. This LDU decomposition is unique. In the last stage, the α_{ij}s, which are to be determined, are found in terms of the elements of L, D and U, and thus in terms of the elements of G. This is carried out in section 4.

The scope of the paper is limited to the derivation of the main result. Its application to SU(4) and other groups in the context of coherent states will be dealt with separately.

2. Reduction of exp(λM) to a matrix polynomial

In this section we consider the LHS of (1.4) and reduce it to a matrix polynomial.

Let M be a square matrix of dimension n satisfying the minimal equation

$$M^N=\sum_{k=0}^{N-1} a_k M^k \qquad N\leqslant n. \tag{2.1}$$

By definition,

$$\exp(\lambda M)=\sum_{j=0}^{\infty} \lambda^j M^j/j!. \tag{2.2}$$

By virtue of (2.1), it is clear that M^j for every j could be expressed as

$$M^j=\sum_{k=0}^{N-1} A_{j,k} M^k \tag{2.3}$$

where $A_{j,k}$ are functions of a_k. Therefore, $\exp(\lambda M)$ itself can be written as a polynomial of degree $N-1$ in M. Substituting (2.3) in (2.2), we have

$$\begin{aligned}\exp(\lambda M)&=\sum_{j=0}^{\infty}\frac{\lambda^j}{j!}\sum_{k=0}^{N-1} A_{j,k}M^k\\ &=\sum_k\left(\sum_j \lambda^j A_{j,k}/j!\right)M^k\\ &=\sum_{k=0}^{N-1} B_k M^k.\end{aligned} \tag{2.4}$$

In this section we obtain explicit expressions for B_ks as functions of the a_ks and λ. It turns out, as we shall show, that $B_0, B_1, \ldots, B_{N-2}$ can all be evaluated once B_{N-1} is known. B_{N-1} itself can be elegantly determined by a Laplace transform technique.

First, we note that (2.1) requires that

$$A_{j,k}=\begin{cases}\delta_{jk} & j<N\\ a_k & j=N.\end{cases} \tag{2.5}$$

To determine $A_{j,k}$ for $j>N$, we proceed as follows. Multiplying (2.3) by M we notice that

$$\begin{aligned}\sum_{k=0}^{N-1} A_{j+1,k}M^k \equiv M^{j+1} &= \sum_{k=0}^{N-1} A_{j,k}M^{k+1} \\ &= \sum_{k=0}^{N-2} A_{j,k}M^{k+1}+A_{j,N-1}M^N \\ &= \sum_{k=0}^{N-2} A_{j,k}M^{k+1}+A_{j,N-1}\sum_{k=0}^{N-1} a_kM^k \end{aligned} \tag{2.6}$$

where in the last step we have used (2.1). Comparing coefficients of powers of M in (2.6), we get the following recurrence relation:

$$\begin{aligned} A_{j+1,0} &= a_0A_{j,N-1} \\ A_{j+1,k} &= A_{j,k-1}+a_kA_{j,N-1} \qquad k=1,\ldots,N-1. \end{aligned} \tag{2.7}$$

Successively setting $k=N-1,\ N-2,\ldots,1$ and adding the resulting expressions we get a recurrence relation in which the second index has the fixed value $N-1$. For simplicity, taking

$$A_j \equiv A_{j,N-1} \tag{2.8}$$

this recurrence relation is

$$A_{j+1}=\sum_{k=0}^{N-1} a_{N-k-1}A_{j-k}. \tag{2.9a}$$

This is an $(N+1)$ term recurrence relation with initial conditions

$$A_j=\delta_{j,N-1} \qquad j\leqslant N-1 \tag{2.9b}$$

which is a consequence of (2.5). For our purposes it is not necessary to solve (2.9). It suffices to evaluate the quantity $S(\mu)$ defined by

$$S(\mu)\equiv\mu\sum_{j=0}^{\infty} A_j\mu^j. \tag{2.10}$$

For we see from (2.4) that

$$\begin{aligned} B_{N-1} &= \sum_{j=0}^{\infty}(\lambda^j/j!)A_{j,N-1}=\mathscr{L}^{-1}\mathscr{L}\sum_{j=0}^{\infty}(\lambda^j/j!)A_j \\ &= \mathscr{L}^{-1}\sum_{j=0}^{\infty}A_j/s^{j+1}=\mathscr{L}^{-1}S(1/s) \end{aligned} \tag{2.11}$$

where

$$\mathscr{L}(f(\lambda))=\int_0^\infty \exp(-s\lambda)f(\lambda)\,\mathrm{d}\lambda$$

is the Laplace transform. To compute $S(\mu)$, multiply (2.9a) by μ^{j+1} and sum from $j=0$ to ∞. Because of (2.9b), the LHS becomes $S(\mu)\mu^{-1}$, while the RHS reduces to

$$\mu^{N-1}+(a_{N-1}+\mu a_{N-2}+\cdots+\mu^{N-1}a_0)S(\mu)\mu^{-1}$$

the first term being a consequence of (2.9b). From this the unknown function $S(1/s)$ is readily found to be

$$S(1/s)=(s^N-a_{N-1}s^{N-1}-a_{N-2}s^{N-2}-\cdots-a_0)^{-1}. \tag{2.12}$$

If now the inverse Laplace transform of (2.12) is taken, B_{N-1} can be obtained.

Let $s_1, s_2, \ldots, s_N$ be the zeros of $S^{-1}(1/s)$. Then

$$S(1/s)=\prod (s-s_i)^{-1}. \tag{2.13}$$

Assuming the zeros to be distinct, it is easy to show that

$$S(1/s)=\sum_{i=1}^{N} C_i(s-s_i)^{-1} \qquad C_i=\prod_{j\neq i}(s_i-s_j)^{-1}. \tag{2.14}$$

When S^{-1} has multiple zeros, this simple resolution into partial fractions requires modification, which we discuss in the appendix. For the case of simple zeros, the inverse transform of S is easily found. We therefore have

$$B_{N-1}=\mathscr{L}^{-1}S(1/s)=\sum_{1}^{N} C_i \exp(\lambda s_i). \tag{2.15}$$

The remaining B_js are determined as follows. Differentiating B_k (cf (2.4)) with respect to λ and making use of (2.7), we deduce that

$$\mathrm{d}B_k/\mathrm{d}\lambda = B_{k-1}+a_kB_{N-1}. \tag{2.16}$$

From this it is clear that B_{N-2} can be expressed in terms of B_{N-1} and its derivative. By repeatedly doing this, one is led to the solution

$$B_k=(D^{N-k-1}-a_{N-1}D^{N-k-2}-a_{N-2}D^{N-k-3}-\cdots-a_{k+1})B_{N-1} \tag{2.17}$$

where $D=\mathrm{d}/\mathrm{d}\lambda$, and $k=0,1,\ldots,N-2$. With this we have evaluated all the coefficients B_k required for expressing $\exp(\lambda M)$ as a matrix polynomial.

Relation (2.16) is valid irrespective of whether the minimal polynomial has distinct zeros or not. Therefore, the only change to be made in (2.17) for the case of multiple zeros is to use the expression for B_{N-1} given in the appendix, instead of (2.15). It is remarkable that $\exp(\lambda M)$ can be expressed so simply in terms of the coefficients of the minimal polynomial of M and one scalar function of the eigenvalues, B_{N-1}.

It may be noted that when the s_i are distinct, our result is equivalent to the well known spectral resolution (Jordan 1969, Halmos 1958)

$$\exp(\lambda M)=\sum_i \exp(\lambda s_i)P_i$$

where the P_i are the projection operators given by

$$P_i=\prod_{j\neq i}(s_i-s_j)^{-1}(M-s_j).$$

When the minimal equation admits multiple roots, it is not possible to express $\exp(\lambda M)$ purely in terms of projection operators. There will occur in addition certain nilpotent operators (Gantmacher 1959). We believe that our method for expressing $\exp(\lambda M)$ as a matrix polynomial is of interest for the reason that it is elementary and gives the result directly in powers of M.

3. LDU decomposition of exp(λM)

Let G be the matrix $\exp(\lambda M)$ calculated in section 2. As explained in the introductory section, the intermediate stage in our derivation is that of finding a lower triangular matrix L, a diagonal matrix D and an upper triangular matrix U such that G is their product:

$$G = LDU. \tag{3.1}$$

In our case, the diagonal elements of both L and U are all unity, and D is a non-singular matrix. Therefore, we define

$$\begin{aligned} L_{ij} &= \delta_{ij} + l_{ij} \qquad l_{ij} = 0 \qquad j \geqslant i \\ D_{ij} &= d_i \delta_{ij} \\ U_{ij} &= \delta_{ij} + u_{ij} \qquad u_{ij} = 0 \qquad j \leqslant i. \end{aligned} \tag{3.2}$$

Hence, from (3.1) it follows that

$$G_{ij} = d_i \delta_{ij} + d_i u_{ij} + l_{ij} d_j + \sum_{k=1}^{n} l_{ik} d_k u_{kj}. \tag{3.3}$$

By suitable back substitutions one can determine from this equation all the elements of L, D and U successively. For instance, if we set $i = 1$,

$$G_{1j} = d_1 \delta_{1j} + d_1 u_{1j}$$

which gives

$$d_1 = G_{11} \qquad u_{1j} = G_{1j}/G_{11} \qquad j > 1.$$

Similarly, setting $j = 1$, one gets

$$l_{i1} = G_{i1}/G_{11} \qquad i > 1.$$

Having determined the elements in the first row of U and the first column of L, one can make use of these to determine the second row of U and the second column of L. It is seen easily that at each stage the elements of L, D and U are fixed uniquely by this procedure. It turns out that the elements u_{ij}, l_{ij} and d_i can all be expressed as ratios of determinants. Let us define a $k \times k$ determinant ($r, s > k-1$)

$$\Delta(11, 22, \ldots, k-1\; k-1, rs) = \begin{vmatrix} G_{11} & G_{12} & \cdots & G_{1\,k-1} & G_{1s} \\ G_{21} & G_{22} & \cdots & G_{2\,k-1} & G_{2s} \\ \vdots & & & \vdots & \vdots \\ G_{k-1\,1} & & \cdots & G_{k-1\,k-1} & G_{k-1\,s} \\ G_{r1} & & \cdots & G_{r\,k-1} & G_{rs} \end{vmatrix}. \tag{3.4}$$

This is nothing but the determinant of a $k \times k$ submatrix of G obtained by adjoining to the $(k-1) \times (k-1)$ principal submatrix of G, the first $k-1$ elements from the row and column containing the element G_{rs} along with the G_{rs} as shown. In terms of these, the elements of L, D and U are given by (Gantmacher 1959)

$$\begin{aligned} d_k &= \frac{\Delta(11, 22, \ldots, kk)}{\Delta(11, 22, \ldots, k-1\; k-1)} \qquad k = 2, \ldots, n \\ u_{kj} &= \frac{\Delta(11, 22, \ldots, k-1\; k-1, kj)}{\Delta(11, 22, \ldots, kk)} \qquad j > k \\ l_{jk} &= \frac{\Delta(11, 22, \ldots, k-1\; k-1, jk)}{\Delta(11, 22, \ldots, kk)} \qquad j > k. \end{aligned} \tag{3.5}$$

4. Determination of the α_{ij}

The final task is to express the α_{ij}s in terms of elements of L, D and U. Recalling that

$$L=\exp(\alpha_{n\,n-1}E_{n\,n-1})\exp(\alpha_{n\,n-2}E_{n\,n-2})\ldots\exp(\alpha_{32}E_{32})\exp(\alpha_{31}E_{31})\exp(\alpha_{21}E_{21}) \quad (4.1)$$

where E_{ij} are the matrices defined by (1.3), satisfying the relation

$$E_{ij}E_{kl}=\delta_{jk}E_{il}. \quad (4.2)$$

Hence, every exponential on the RHS of (4.1) reduces to a lower triangular matrix of the form $(\mathbb{1}+\alpha_{ij}E_{ij})$ (no summation over the indices). Since the product of any number of lower triangular matrices is a lower triangular matrix, L is a lower triangular matrix, as supposed, whose diagonal elements are all unity.

Let us now define a lower triangular matrix α_L whose elements are the α_{ij} of (4.1) and whose diagonal elements are all zero:

$$\alpha_L=\begin{bmatrix} 0 & 0 & 0 & \cdots & 0\\ \alpha_{21} & 0 & 0 & \cdots & 0\\ \alpha_{31} & \alpha_{32} & 0 & \cdots & 0\\ \vdots & \vdots & \vdots & & \vdots\\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{n\,n-1} & 0\end{bmatrix}. \quad (4.3)$$

Clearly, $\alpha_L^n=0$. Multiplying out the RHS *of (4.1) and collecting terms, we find that* (4.1) can be written in a compact form as

$$L=\mathbb{1}+\alpha_L+\alpha_L^2+\cdots+\alpha_L^{n-1}. \quad (4.4)$$

Using the fact that $\alpha_L^n=0$, this equation can be inverted to obtain α_L in terms of L explicitly:

$$\alpha_L=-\sum_{r=1}^{n-1}(\mathbb{1}-L)^r. \quad (4.5)$$

This relation determines all the α_{ij}s with $i>j$ in terms of the elements of L.

To find the α_{ij}s for $i<j$, one proceeds in a similar manner. If we denote by α_U the upper triangular matrix with elements

$$\begin{aligned}(\alpha_U)_{ij}&=\alpha_{ij} \qquad i<j\\ (\alpha_U)_{ij}&=0 \qquad i\geqslant j\end{aligned} \quad (4.6)$$

then the upper triangular matrix U defined by

$$U=\exp(\alpha_{12}E_{12})\ldots\exp(\alpha_{n-1\,n}E_{n-1\,n}) \quad (4.7)$$

is expressible as

$$U=\mathbb{1}+\alpha_U+\alpha_U^2+\cdots+\alpha_U^{n-1} \quad (4.8)$$

which gives

$$\alpha_U=-\sum_{r=1}^{n-1}(\mathbb{1}-U)^r. \quad (4.9)$$

The diagonal elements α_{ii} in (1.5) are easily found in terms of the elements of

$$D=\exp(\alpha_{11}E_{11})\exp(\alpha_{22}E_{22})\ldots\exp(\alpha_{nn}E_{nn}) \quad (4.10)$$

and

$$\alpha_{ii}=(\ln D)_{ii}=\ln d_i. \quad (4.11)$$

This completes the determination of all the α_{ij}s in terms of elements of G defined in (3.1).

5. Discussion

We have shown that the exponential of an arbitrary $n \times n$ matrix can be factorized and written as a product of exponentials involving the natural basis, E_{ij}, for the set of $n \times n$ matrices. We have introduced the Laplace transform technique as a tool to reduce the exponential of a matrix to a matrix polynomial. This technique is general enough to accommodate both diagonalizable and non-diagonalizable matrices. For diagonalizable matrices, our results in section 2 are equivalent to the spectral decomposition of functions of matrices. For the case of non-diagonalizable matrices the spectral decomposition involves both projection operators and associated nilpotent operators. The Laplace transform method determines the coefficient B_l directly without having to deal with the projection and nilpotent operators separately.

The disentanglement relation derived by us should prove useful as a BCH formula for SU(n) because α_{ij}s determined as functions of M in one faithful representation are valid in all faithful representations. Our result is of particular interest in the study of SU(n) coherent states. Its application to physically interesting groups such as SU(4), SU(6), etc., will be considered separately.

Acknowledgments

We thank P M Mathews and M D Srinivas for discussions.

Appendix

Let (s_i, α_i), $i = 1, \ldots, r \leqslant N$, denote the distinct roots and their multiplicities of the polynomial equation

$$s^N - a_{N-1}s^{N-1} - a_{N-2}s^{N-2} - \cdots - a_0 = 0$$

and consider the expression

$$F(s) = \prod_{i=1}^{r} 1/(s - s_i)^{\alpha_i}.$$

The partial fraction resolution of F is given by

$$F(s) = \sum_{i=1}^{r} \sum_{j=1}^{\alpha_i} C_{i,j}/(s - s_i)^j$$

where the coefficients $C_{i,j}$ are given by

$$C_{i,j} = [1/(\alpha_i - j)!][(\mathrm{d}/\mathrm{d}s)^{\alpha_i - j}(F(s)(s - s_i)^{\alpha_i})]_{s = s_i}.$$

Consequently, $\mathscr{L}^{-1}F(s)$ is easily found to be

$$\mathscr{L}^{-1}F(s) = \sum_{i=1}^{r} \sum_{j=1}^{\alpha_i} C_{i,j} \exp(s_i\lambda)\lambda^{j-1}/(j-1)! = \sum_{i=1}^{r} \tilde{C}_i(\lambda) \exp(s_i\lambda)$$

where

$$\tilde{C}_i(\lambda) = \sum_{j=1}^{\alpha_i} C_{i,j}\lambda^{j-1}/(j-1)!.$$

Hence

$$B_{N-1} = \sum_{i=1}^{r} \tilde{C}_i(\lambda) \exp(s_i\lambda).$$

This replaces (2.15) for the case of multiple zeros of $S(1/s)$.

References

Gantmacher F R 1959 *The Theory of Matrices* (New York: Chelsea)
Gilmore R 1974a *J. Math. Phys.* **15** 2090
—— 1974b *Lie Groups, Lie Algebras and Some of Their Applications* (New York: Wiley)
Halmos P R 1958 *Finite Dimensional Vector Spaces* (New York: van Nostrand)
Jordan T F (ed) 1969 *Linear Operators for Quantum Mechanics* (Duluth: Jordan)
Raghunathan K, Seetharaman M and Vasan S S 1989 *J. Phys. A: Math. Gen.* **22** L1089