Baker - Campbell - Hausdorff relation for special unitary groups $SU(N)$



1997 J. Phys. A: Math. Gen. 30 8739

(http://iopscience.iop.org/0305-4470/30/24/032)

J. Phys. A: Math. Gen. **30** (1997) 8739–8749. Printed in the UK PII: S0305-4470(97)86896-5

# Baker–Campbell–Hausdorff relation for special unitary groups $SU(N)$

Stefan Weigert

Departement für Physik und Astronomie der Universität Basel, Klingelbergstrasse 82, CH-4056 Basel, Switzerland

Received 16 August 1997

**Abstract.** Multiplication of two elements of the special unitary group $SU(N)$ determines uniquely a third group element. A Baker–Campbell–Hausdorff relation is derived which expresses the group parameters of the product (written as an exponential) in terms of the parameters of the exponential factors. This requires the eigenvalues of three $(N \times N)$ matrices. Consequently, the relation can be stated analytically up to $N = 4$, in principle. Similarity transformations encoding the time evolution of quantum-mechanical observables, for example, can be obtained by the same means.

## 1. Introduction

Various questions in physics reduce to the following problem: write the product of exponential functions depending on noncommuting operators $\widehat{A}$ and $\widehat{B}$, respectively, as the exponential of a third operator, $\widehat{C}$,

$$\exp[\widehat{A}]\exp[\widehat{B}] = \exp[\widehat{C}]. \tag{1}$$

The names of Baker, Campbell, and Hausdorff (BCH) are associated [21] with a formula for the operator $\widehat{C}$ expressed in multiple commutators of $\widehat{A}$ and $\widehat{B}$:

$$\widehat{C} = \widehat{A} + \widehat{B} + \tfrac{1}{2}[\widehat{A}, \widehat{B}] + \tfrac{1}{12}([\widehat{A}, [\widehat{A}, \widehat{B}]] + [[\widehat{A}, \widehat{B}], \widehat{B}]) + \ldots . \tag{2}$$

Remarkably, the operator $\widehat{C}$ depends on commutators of $\widehat{A}$ and $\widehat{B}$ *only* implying that it is contained in the same algebra as $\widehat{A}$ and $\widehat{B}$. For this result to hold it is crucial to consider products of *exponential* functions.

Although the expansion (2) for the operator $\widehat{C}$ is explicit, usually the infinite series of repeated commutators cannot be summed in closed form. It may be used, however, to generate an approximate expression for $\widehat{C}$ by directly calculating a finite number of terms [8]. When read from left to right, equation (1) shows how to *entangle* the two factors into a single exponential. An application important in quantum mechanics results for the Heisenberg group of position and momentum operators $\hat{q}$ and $\hat{p}$, where

$$\exp[-\mathrm{i}\hat{p}]\exp[-\mathrm{i}\hat{q}] = \exp[-\mathrm{i}(\hat{p} + \hat{q}) + \mathrm{i}\hbar/2]. \tag{3}$$

The right-hand side is particularly simple because the commutator

$$[\hat{p}, \hat{q}] = \frac{\hbar}{\mathrm{i}} \tag{4}$$

is a constant such that only the first commutator in (2) contributes to the operator $\widehat{C}$. Another situation with the need for entangling two operators is encountered in periodically driven

0305-4470/97/248739+11$19.50  

systems. In specific cases, the propagator over one full period reduces to a product of the propagators for shorter intervals [4, 7, 19]. The Lie algebras involved in these 'quantum maps' may either have a finite or an infinite number of elements.

When read in the opposite sense, equation (1) represents a *disentangling* relation, that is, the decomposition of a single exponential into factors with simple properties. Such a relation is useful to calculate expectation values of basic operators in the group $SU(2)$, for example, since they are easily derived from a generating function in disentangled form [1]. Similarly, changes of the group parametrization [9] are conveniently performed by using BCH relations. In general, the discussion of coherent states for particle and spin systems as well as for arbitrary Lie groups [16] benefits from the knowledge of (de-)composition rules (1).

A closely related question arises from the need to perform similarity transformations according to

$$\exp[-\widehat{A}]\widehat{B}\exp[\widehat{A}] = \widehat{B}'. \tag{5}$$

If the operator $\widehat{A}$ is proportional to i times the Hamiltonian of a quantum system, equation (5) describes the time evolution of the Heisenberg observable $\widehat{B}$ into $\widehat{B}'$.

A number of techniques has been established in order to efficiently treat entangling and disentangling problems, in particular, if the operators involved in the BCH relation are elements of a *finite*-dimensional Lie algebra. Two-dimensional unitary faithful irreducible representations are used to derive explicit results for the group $SU(2)$ [9], and for the group of the harmonic oscillator [15, 10], for example. Applications to more complicated cases involving symplectic groups also have been worked out in detail [20, 11]. However, it is *not* necessary to exclusively work with unitary representations: any faithful representation can be used [10]. This is helpful if one knows a representation consisting of upper and lower triangular matrices since they are easily exponentiated. Disentanglement of Lie group elements is also achieved by using recursion relations for expanded exponentials and Laplace-transform techniques [18]. This approach generalizes a method first applied to particular group elements of $SU(3)$ [17]. The powerful approach in [21] maps the problem of both (dis-)entangling (1) and similarity transformations (5) to the solution of a set of coupled first-order differential equations. This paper also contains theoretical background on BCH relations, applications in physics as well as a large number of references.

In this paper a different method to evaluate BCH relations is developped for the groups $SU(N)$. It is based on the spectral theorem for Hermitian operators in finite-dimensional vector spaces. A 'linearized' version of this theorem is derived by exploiting a specific feature of the algebra $su(N)$ the fundamental of representation going beyond its Lie algebraic properties. In this way, a one-to-one correspondence between an exponential of linearly combined generators and a linear combination of them is established—thus 'removing' the exponential function. It is then straightforward to entangle elements of the group $SU(N)$. Conceptually, this method is related to work performed in the early 1970s where the study of chiral algebras required the evaluation of *finite* transformations for special unitary groups [2, 3]. In that context, however, BCH relations have not been considered.

## 2. Some fundamentals of $SU(N)$

An irreducible faithful representation of the group $SU(N)$ [14] is given by the set of all unitary $(N \times N)$ matrices $\mathbf{U}$ with unit determinant,

$$\det\mathbf{U} = 1 \qquad \mathbf{U}_{nn'} \in \mathbb{C} \qquad n, n' = 1, \ldots, N \tag{6}$$

also known as the fundamental or defining representation. Each matrix $\mathsf{U}$ can be written in the form

$$\mathsf{U} = \exp[-\mathrm{i}\mathsf{L}] \qquad \mathsf{L}^\dagger = \mathsf{L} \tag{7}$$

with a traceless Hermitian matrix $\mathsf{L}$. It is conveniently expressed as a linear combination

$$\mathsf{L} = \boldsymbol{L} \cdot \boldsymbol{\Lambda} \equiv \sum_{j=1}^{N^2-1} L_j \boldsymbol{\Lambda}_j \qquad L_j \in \mathbb{R} \tag{8}$$

with the set $\boldsymbol{\Lambda}$ forming a basis for traceless Hermitian matrices. The Hermitian generators $\boldsymbol{\Lambda}_j^\dagger = \boldsymbol{\Lambda}_j$ are a basis of the Lie *algebra* $su(N)$ of $SU(N)$, satisfying the commutation relations:

$$[\boldsymbol{\Lambda}_j, \boldsymbol{\Lambda}_k]_- = 2\mathrm{i} f_{jkl} \boldsymbol{\Lambda}_l \tag{9}$$

where the indices $j, k, l$, take values from 1 to $N^2 - 1$, the summation convention for repeated indices applies, and the $(N \times N)$ unit matrix is denoted by $\mathsf{I}_N$. The structure constants $f_{jkl}$ are elements of a completely antisymmetric tensor (spelled out explicitly in [12] for example) with Jacobi identity

$$f_{klm} f_{mpq} + f_{plm} f_{mkq} + f_{kpm} f_{mlq} = 0. \tag{10}$$

The group $SU(N)$ has rank $(N - 1)$. In other words, any maximal Abelian subalgebra of $su(N)$ consists of $(N - 1)$ elements corresponding to all linearly independent traceless $N$-dimensional diagonal matrices. A 'complete set of commuting variables' for a quantum system described by $SU(N)$ would contain in addition the same number of Casimir operators according to Racah's theorem [12]. The properties given so far are valid for all faithful representations.

A particular feature of the generators in the defining representation of the algebra $su(N)$ is closure under *anti*commutation:

$$[\boldsymbol{\Lambda}_j, \boldsymbol{\Lambda}_k]_+ = \frac{4}{N} \delta_{jk} \mathsf{I}_N + 2 d_{jkl} \boldsymbol{\Lambda}_l \tag{11}$$

where the $d_{jkl}$ form a totally symmetric tensor [12]. For $N = 2$, *all* numbers $d_{jkl}$ are equal to zero, and the generators $\boldsymbol{\Lambda}$ coincide with the Pauli matrices $\boldsymbol{\sigma}$: the anticommutator of two of them is either equal to zero or a multiple of the unit matrix, $\mathsf{I}_2$.

The anticommutation relation is crucial for the following, however, it is neither valid for representations other than the fundamental one nor for other Lie algebras. As a consequence of (11), two generators $\boldsymbol{\Lambda}_j$ and $\boldsymbol{\Lambda}_k$ of $su(N)$ are 'orthogonal' to each other with respect to the trace:

$$\mathrm{Tr}(\boldsymbol{\Lambda}_j \boldsymbol{\Lambda}_k) = 2\delta_{jk}. \tag{12}$$

In addition, a second Jacobi-type identity exists involving both the antisymmetric and the symmetric structure coefficients in (9) and (11):

$$f_{klm} d_{mpq} + f_{kqm} d_{mpl} + f_{kpm} d_{mlq} = 0. \tag{13}$$

For the following, a vector-type notation is useful, based on the structure constants and the Kronecker symbol. Define the scalar product as already employed in equation (8),

$$\boldsymbol{A} \cdot \boldsymbol{B} = A_n \delta_{nm} B_m = A_n B_n \tag{14}$$

where the components of $\boldsymbol{A}$ and $\boldsymbol{B}$ are allowed to be either numbers or generators $\boldsymbol{\Lambda}_n$. Similarly, define an antisymmetric 'cross product' $\otimes$ by

$$(\boldsymbol{A} \otimes \boldsymbol{B})_j = f_{jkl} A_k B_l = -(\boldsymbol{B} \otimes \boldsymbol{A})_j \tag{15}$$

and a symmetric 'dot product' $\odot$:

$$(\boldsymbol{A} \odot \boldsymbol{B})_j = d_{jkl} A_k B_l = +(\boldsymbol{B} \odot \boldsymbol{A})_j . \tag{16}$$

Then, the relations (9) and (11) can be written

$$[\boldsymbol{A} \cdot \boldsymbol{\Lambda}, \boldsymbol{B} \cdot \boldsymbol{\Lambda}]_- = 2\mathrm{i}(\boldsymbol{A} \otimes \boldsymbol{B}) \cdot \boldsymbol{\Lambda} \tag{17}$$

$$[\boldsymbol{A} \cdot \boldsymbol{\Lambda}, \boldsymbol{B} \cdot \boldsymbol{\Lambda}]_+ = \frac{4}{N} \boldsymbol{A} \cdot \boldsymbol{B} \mathsf{I}_N + 2(\boldsymbol{A} \odot \boldsymbol{B}) \cdot \boldsymbol{\Lambda} \tag{18}$$

where $\boldsymbol{A}$ and $\boldsymbol{B}$ are *arbitrary* vectors of dimension $(N^2 - 1)$ with numeric entries. Adding these equations leads to a compact form of the (anti-) commutation relations:

$$(\boldsymbol{A} \cdot \boldsymbol{\Lambda})(\boldsymbol{B} \cdot \boldsymbol{\Lambda}) = \frac{2}{N} \boldsymbol{A} \cdot \boldsymbol{B} \ \mathsf{I}_N + (\boldsymbol{A} \odot \boldsymbol{B} + \mathrm{i} \boldsymbol{A} \otimes \boldsymbol{B}) \cdot \boldsymbol{\Lambda}. \tag{19}$$

This equation emphasizes the important point that any expression *quadratic* in the generators can be expressed as a *linear* combination of them, including the identity. As a matter of fact, it generalizes the known identity in $SU(2)$ for the Pauli matrices:

$$(\boldsymbol{A} \cdot \boldsymbol{\sigma})(\boldsymbol{B} \cdot \boldsymbol{\sigma}) = \boldsymbol{A} \cdot \boldsymbol{B} \ \mathsf{I}_N + \mathrm{i} \boldsymbol{A} \otimes \boldsymbol{B} \cdot \boldsymbol{\sigma}. \tag{20}$$

In the new notation, the identities (10), (13) read

$$(\boldsymbol{A} \otimes \boldsymbol{B}) \cdot (\boldsymbol{C} \otimes \boldsymbol{D}) + (\boldsymbol{C} \otimes \boldsymbol{B}) \cdot (\boldsymbol{A} \otimes \boldsymbol{D}) + (\boldsymbol{A} \otimes \boldsymbol{C}) \cdot (\boldsymbol{B} \otimes \boldsymbol{D}) = 0 \tag{21}$$

$$(\boldsymbol{A} \otimes \boldsymbol{B}) \cdot (\boldsymbol{C} \odot \boldsymbol{D}) + (\boldsymbol{A} \otimes \boldsymbol{D}) \cdot (\boldsymbol{C} \odot \boldsymbol{B}) + (\boldsymbol{A} \otimes \boldsymbol{C}) \cdot (\boldsymbol{B} \odot \boldsymbol{D}) = 0. \tag{22}$$

Another useful form of equation (13) is given by

$$\boldsymbol{A} \otimes (\boldsymbol{B} \odot \boldsymbol{C}) = (\boldsymbol{A} \otimes \boldsymbol{B}) \odot \boldsymbol{C} + \boldsymbol{B} \odot (\boldsymbol{A} \otimes \boldsymbol{C}) \tag{23}$$

showing that applying $\boldsymbol{A}\otimes$ to a $\odot$ product acts as does a derivative. The 'orthogonality' of the generators (12) becomes

$$\mathrm{Tr}((\boldsymbol{A} \cdot \boldsymbol{\Lambda})(\boldsymbol{B} \cdot \boldsymbol{\Lambda})) = 2 \boldsymbol{A} \cdot \boldsymbol{B} \tag{24}$$

for arbitrary $\boldsymbol{A}$ and B.

## 3. Spectral theorem

Every matrix $\mathsf{M} \in \mathbb{C}^N$ satisfies its own characteristic equation,

$$\sum_{n=0}^{N} a_n \mathsf{M}^n = 0 \qquad a_N = 1, a_0 = \det \mathsf{M} \tag{25}$$

according to the theorem of Cayley–Hamilton. The coefficients $a_n$ define the characteristic polynomial of $\mathsf{M}$. For traceless matrices such as $\mathsf{M} \in su(N)$, the coefficient $a_{N-1}$ in equation (25) is equal to zero since it equals the trace of $\mathsf{M}$. According to equation (25), any power $N' \geqslant N$ of the matrix $\mathsf{M}$ is identical to a linear combination of its powers $\mathsf{M}^n$ with $0 \leqslant n \leqslant N-1$. The expansion of a matrix exponential can thus be written

$$\exp[-\mathrm{i}\mathsf{M}] = \sum_{m=0}^{\infty} \frac{(-\mathrm{i}\mathsf{M})^m}{m!} = \sum_{n=0}^{N-1} e_n(\mathsf{M}) \mathsf{M}^n \tag{26}$$

with uniquely defined coefficients $e_n(\mathsf{M})$. They are determined directly by referring to the *spectral theorem* [13] valid for smooth functions $f$ of a Hermitian matrix $\mathsf{M}$ with (nondegenerate) eigenvalues $m_k$, $k = 1, \ldots, N$:

$$f(\mathsf{M}) = \sum_{k=1}^{N} f(m_k) \mathsf{P}_k \tag{27}$$

and the operator $\mathsf{P}_k = |m_k\rangle\langle m_k|$ projects down to the one-dimensional eigenspace spanned by the eigenvector $|m_k\rangle$ associated with the eigenvalue $m_k$. In terms of powers $\mathsf{M}^k$ and the eigenvalues $m_k$, the matrices $\mathsf{P}_k$ read

$$\mathsf{P}_k = \prod_{n\neq k} \frac{\mathsf{M} - m_n}{m_k - m_n} = \sum_{n=0}^{N-1} P_{kn}\mathsf{M}^n \tag{28}$$

the sum contains powers $\mathsf{M}^{N-1}$ at most since the product runs over $(N-1)$ factors. Combining equations (27) and (28), one obtains

$$f(\mathsf{M}) = \sum_{n=0}^{N-1} \left( \sum_{k=1}^{N} P_{kn} f(m_k) \right) \mathsf{M}^n \equiv \sum_{n=0}^{N-1} f_n \mathsf{M}^n \tag{29}$$

and, upon choosing $f(x) \equiv \exp[-\mathrm{i}x]$, the sum in the round brackets produces the coefficients $e_n$ of the expansion (26) in terms of the eigenvalues $m_k$.

It is possible to express the numbers $f_n$ in (29) differently [18]. Write the coefficient $f_{N-1}(\mathsf{M}, \lambda)$ of $\mathsf{M}^{N-1}$ with a dummy parameter $\lambda$ introduced as follows

$$f_{N-1}(\mathsf{M}, \lambda) = \sum_{n=1}^{N} \Delta_n f(\lambda m_k) \qquad \Delta_n = \prod_{k\neq n} (m_n - m_k)^{-1}. \tag{30}$$

Linear combinations of derivatives with respect to $\lambda$ yield the remaining coefficients $f_n$, $n = 0, 1, \ldots, N-2$, associated with any smooth function $f$:

$$f_n(\mathsf{M}) = \left[ \left( \partial_\lambda^{N-n-1} - \sum_{\nu=1}^{N-n-1} a_{N-\nu} \partial_\lambda^{N-n-1-\nu} \right) f_{N-1}(\mathsf{M}, \lambda) \right]_{\lambda=1} \tag{31}$$

with numbers $a_n$ from the characteristic polynomial (25), and the abbreviaton $\mathrm{d}/\mathrm{d}\lambda \equiv \partial_\lambda$. Since equation (29) requires the eigenvalues of M, analytic expressions will be obtained only for $(4 \times 4)$ matrices at most, i.e. for $SU(4)$.

## 4. Linearized spectral theorem

A stronger version of relation (27) is now derived. It is valid for Hermitian $(N \times N)$ matrices, and it will be called the *linearized spectral theorem*:

$$f(\boldsymbol{M} \cdot \boldsymbol{\Lambda}) = f_0(\boldsymbol{M})\, \mathsf{I}_N + \boldsymbol{f}(\boldsymbol{M}) \cdot \boldsymbol{\Lambda}. \tag{32}$$

It states that any function $f$ of a linear combination of the generators $\boldsymbol{\Lambda}$ of $SU(N)$ is equal to a linear combination of the identity and the generators with well-defined coefficients $(f_0, \boldsymbol{f})$. In other words, the powers of the generators $\boldsymbol{\Lambda}$ contained in the powers $\mathsf{M}^n \equiv (\boldsymbol{M} \cdot \boldsymbol{\Lambda})^n$ in equation (29) can be reduced to linear combinations of them. In view of the commutation relations of the algebra $su(N)$, equation (19), this is not surprising: the required reduction is carried out in a finite number of steps by repeatedly expressing products of two generators by a linear combination of generators.

A convenient procedure to determine $(f_0, \boldsymbol{f})$ in (32) starts from writing

$$\mathsf{M}^n = \mu_{0,n}\mathsf{I}_N + \boldsymbol{\mu}_n \cdot \boldsymbol{\Lambda} \qquad n = 0, 1, 2, \ldots, N-1 \tag{33}$$

where

$$\mu_{0,0} = 1 \qquad \mu_{0,1} = 0 \tag{34}$$

$$\boldsymbol{\mu}_0 = 0 \qquad \boldsymbol{\mu}_1 = \boldsymbol{M}. \tag{35}$$

A recursion relation for $(\mu_{0,n}, \boldsymbol{\mu}_n)$ follows from writing $\mathsf{M}^{n+1} = \mathsf{M}^n\mathsf{M}$, using (19) and (33),

$$\begin{aligned}\mathsf{M}^{n+1} &= \mu_{0,n}\boldsymbol{M}\cdot\boldsymbol{\Lambda} + (\boldsymbol{\mu}_n\cdot\boldsymbol{\Lambda})(\boldsymbol{M}\cdot\boldsymbol{\Lambda}) \\ &= \frac{2}{N}\boldsymbol{\mu}_n\cdot\boldsymbol{M}\,\mathsf{I}_N + (\mu_{0,n}\boldsymbol{M} + \boldsymbol{\mu}_n\odot\boldsymbol{M} + \mathrm{i}\boldsymbol{\mu}_n\otimes\boldsymbol{M})\cdot\boldsymbol{\Lambda}.\end{aligned} \tag{36}$$

Comparison with (33) for $(n+1)$ instead of $n$ shows that

$$\mu_{0,n+1} = \frac{2}{N}\boldsymbol{\mu}_n\cdot\boldsymbol{M} \tag{37}$$

$$\boldsymbol{\mu}_{n+1} = \mu_{0,n}\boldsymbol{M} + \boldsymbol{\mu}_n\odot\boldsymbol{M} + \mathrm{i}\boldsymbol{\mu}_n\otimes\boldsymbol{M} = \frac{2}{N}(\boldsymbol{\mu}_{n-1}\cdot\boldsymbol{M})\boldsymbol{M} + \boldsymbol{\mu}_n\odot\boldsymbol{M} \tag{38}$$

which recursively defines $(\mu_{0,n}, \boldsymbol{\mu}_n)$ in terms of M, starting with the 'initial values' (34) and (35). The terms $\mathrm{i}\boldsymbol{\mu}_n\otimes\boldsymbol{M}$ do *not* contribute since each $\boldsymbol{\mu}_n$ following from (33) is proportional to $\boldsymbol{M}, \boldsymbol{M}\odot\boldsymbol{M}, (\boldsymbol{M}\odot\boldsymbol{M})\odot\boldsymbol{M}, \ldots$. Using the derivative-like property (23), one always encounters terms $\boldsymbol{M}\otimes\boldsymbol{M}$ being equal to zero. Consequently, the coefficients $(f_0, \boldsymbol{f})$ on the right-hand side of (32) have been expressed explicitly through $\boldsymbol{M}$ and the eigenvalues $m_k$:

$$f_0(\boldsymbol{M}) = \sum_{n=0}^{N-1} f_n\mu_{0,n} \qquad \boldsymbol{f}(\boldsymbol{M}) = \sum_{n=0}^{N-1} f_n\boldsymbol{\mu}_n \tag{39}$$

with $f_n$ from equations (30) and (31). Note that according to (38) the expression for $\boldsymbol{f}(\boldsymbol{M})$ contains only totally symmetric powers $\boldsymbol{M}, \boldsymbol{M}\odot\boldsymbol{M}, (\boldsymbol{M}\odot\boldsymbol{M})\odot\boldsymbol{M}, \ldots$. Given $\boldsymbol{M}$, a simple expression for $f_0$ is provided by taking the trace of equation (32):

$$f_0(\boldsymbol{M}) = \frac{1}{N}\,\mathrm{Tr}(f(\boldsymbol{M}\cdot\boldsymbol{\Lambda})) = \frac{1}{N}\sum_{k=1}^{N} f(m_k). \tag{40}$$

It should be pointed out that $f_0$ is *not* independent of $\boldsymbol{f}$: one can solve the recursion for $\boldsymbol{\mu}_n$, equation (38) without referring to (37). This is reasonable because only then are there the *same* number of parameters in $\boldsymbol{M}$ and on the right-hand side of (32).

Suppose now that the *right-hand side* of equation (32) is given, i.e. the parameters $(f_0, \boldsymbol{f})$ are known to define a group element of $SU(N)$. How does one express $\boldsymbol{M}$ in terms of $\boldsymbol{f}$? This is actually the difficult step when deriving a BCH formula: to find the group element in terms of the the original parametrization. Assume the function $f$ to be invertible, then one can write

$$\boldsymbol{M}\cdot\boldsymbol{\Lambda} = f^{-1}(f_0\mathsf{I}_N + \boldsymbol{f}\cdot\boldsymbol{\Lambda}) = F(\boldsymbol{f}\cdot\boldsymbol{\Lambda}) \tag{41}$$

with a new function $F$. The clue to the inversion is to realize that (41) represents an equation of the type (32) again. This follows from reading equation (32) from right to left, replacing $f \to F$, exchanging the role of $\boldsymbol{f}$ and $\boldsymbol{M}$, and setting $f_0$ equal to zero in (32). Now the reasoning leading to equation (39) can be repeated in order to determine $\boldsymbol{M} = \boldsymbol{M}(\boldsymbol{f})$. Therefore, $\boldsymbol{M}$ can be found as a function of $\boldsymbol{f}$ by the means already established.

The orthonormality (12) for the generators $\boldsymbol{\Lambda}$ allows us to formally switch from $\boldsymbol{M}$ to $\boldsymbol{f}$ and *vice versa* in a simple manner: multiply equation (32) with $\boldsymbol{\Lambda}_k$ and take the trace which leads to

$$f_k = \mathrm{Tr}(f_0\boldsymbol{\Lambda}_k + \boldsymbol{f}\cdot\boldsymbol{\Lambda}\boldsymbol{\Lambda}_k) = \mathrm{Tr}(f(\boldsymbol{M}\cdot\boldsymbol{\Lambda})\boldsymbol{\Lambda}_k) \tag{42}$$

while the inverse transformation follows from (41):

$$M_k = \mathrm{Tr}((\boldsymbol{M}\cdot\boldsymbol{\Lambda})\boldsymbol{\Lambda}_k) = \mathrm{Tr}(f^{-1}(f_0\mathsf{I}_N + \boldsymbol{f}\cdot\boldsymbol{\Lambda})\boldsymbol{\Lambda}_k). \tag{43}$$

Before applying the linearized spectral theorem to the derivation of BCH formulae, a comment on the relation between the matrices $\mathsf{M} = \boldsymbol{M} \cdot \boldsymbol{\Lambda}$ and $\mathsf{F}$ in (32),

$$f(\mathsf{M}) = f_0 \mathsf{I}_N + \mathsf{F} \tag{44}$$

should be made. We must have $[\mathsf{M}, \mathsf{F}] = 0$ since equation (44) is an identity. Nevertheless, the matrices involved do not have to be multiples of each other. The vanishing commutator implies that the matrices $\mathsf{M}$ and $\mathsf{F}$ can be diagonalized simultaneously. Having done this $\mathsf{M}$ would be given by a specific linear combination of $(n-1)$ traceless diagonal generators $\mathsf{H}_k$, $k = 1, 2, \ldots, N-1$. The matrix $\mathsf{F}$ commutes with $\mathsf{M}$ and it is therefore only required to be another element of the maximal Abelian subalgebra containing $\mathsf{M}$. For the group $SU(2)$, the dimension of this algebra is equal to one: $\mathsf{M}$ and $\mathsf{F}$ are in this (and only this) case proportional to each other (cf the first example below). For $SU(3)$ this observation is illustrated by a result of [6] where Lie groups are studied from a geometric point of view. In an appropriate local basis, any group element can be written as a function of a linear combination of two commuting operators which span a maximal Abelian subalgebra.

## 5. BCH for $SU(N)$

A BCH relation for composing a group of elements of $SU(N)$ follows from twofold application of the linearized spectral theorem with $f(x) = \exp[-\mathrm{i}x]$. Consider the product of two finite transformations, $\exp[-\mathrm{i}\boldsymbol{M}\cdot\boldsymbol{\Lambda}]$ and $\exp[-\mathrm{i}\boldsymbol{N}\cdot\boldsymbol{\Lambda}]$, which defines a third element of $SU(N)$ characterized by $\boldsymbol{R}$,

$$\exp[-\mathrm{i}\boldsymbol{R} \cdot \boldsymbol{\Lambda}] = \exp[-\mathrm{i}\boldsymbol{M} \cdot \boldsymbol{\Lambda}]\exp[-\mathrm{i}\boldsymbol{N} \cdot \boldsymbol{\Lambda}]. \tag{45}$$

Using equation (32) with the exponential function, we obtain

$$\begin{aligned}\exp[-\mathrm{i}\boldsymbol{R} \cdot \boldsymbol{\Lambda}] &= \mu_0\nu_0 \mathsf{I}_N + (\nu_0\boldsymbol{\mu} + \mu_0\boldsymbol{\nu}) \cdot \boldsymbol{\Lambda} + (\boldsymbol{\mu} \cdot \boldsymbol{\Lambda})(\boldsymbol{\nu} \cdot \boldsymbol{\Lambda}) \\ &= (\mu_0\nu_0 + \frac{2}{N}\boldsymbol{\mu} \cdot \boldsymbol{\nu})\mathsf{I}_N + (\nu_0\boldsymbol{\mu} + \mu_0\boldsymbol{\nu} + \boldsymbol{\mu} \odot \boldsymbol{\nu} + \mathrm{i}\boldsymbol{\mu} \otimes \boldsymbol{\nu}) \cdot \boldsymbol{\Lambda} \\ &= \rho_0 \mathsf{I}_N + \boldsymbol{\rho} \cdot \boldsymbol{\Lambda}\end{aligned} \tag{46}$$

using the commutation relations (19). The quantities $(\rho_0, \boldsymbol{\rho})$ can be read off directly as the coefficients of $\mathsf{I}_N$ and $\Lambda_j$, respectively. The components of $\boldsymbol{R}$ are thus given by equation (43):

$$R_k = \mathrm{i}\,\mathrm{Tr}\left\{\ln\left[\left(\mu_0\nu_0 + \frac{2}{N}\boldsymbol{\mu} \cdot \boldsymbol{\nu}\right)\mathsf{I}_N + (\nu_0\boldsymbol{\mu} + \mu_0\boldsymbol{\nu} + \boldsymbol{\mu} \odot \boldsymbol{\nu} + \mathrm{i}\boldsymbol{\mu} \otimes \boldsymbol{\nu}) \cdot \boldsymbol{\Lambda}\right]\Lambda_k\right\} \tag{47}$$

providing the relation $\boldsymbol{R} = \boldsymbol{R}(\boldsymbol{M}, \boldsymbol{N})$. The explicit evaluation requires diagonalization of the matrices $\mathsf{M}$ an $\mathsf{N}$ in order to determine $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$; finally, $\boldsymbol{\rho} \cdot \boldsymbol{\Lambda}$ has to be diagonalized in order to evaluate the logarithm in equation (47). In total, three $(N \times N)$ matrices have to be diagonalized to achieve the entangling. Derivation of (47) makes use of the anticommutation relations which are peculiar to the defining representation. Nevertheless the result is valid for *all* representations since the operator $\hat{C}$ in (2) is a uniquely defined linear combination of the generators. If it has been determined in one faithful representation it is known in all others.

As an illustration, the familiar example of $SU(2)$ will be looked at from the point of view developed here. However, the $\odot$ product being identical to zero, this case does not exhibit the full complexity. Therefore, $SU(4)$ will also be discussed briefly. Before giving the examples, the use of the linearized spectral theorem for the determination of similarity transformations in the group $SU(N)$ will be indicated.

## 6. Similarity transformations

The transformation of the operator $\mathsf{N} = \boldsymbol{N} \cdot \boldsymbol{\Lambda} \in su(N)$ under $\mathsf{M} = \boldsymbol{M} \cdot \boldsymbol{\Lambda} \in su(N)$ according to

$$\exp[-\mathrm{i}\mathsf{M}]\mathsf{N}\exp[\mathrm{i}\mathsf{M}] = \mathsf{N}' \tag{48}$$

could be determined from the linearized spectral theorem in the following way. Write the group element as

$$\exp[\mathrm{i}\mathsf{M}] = \mu_0 \mathsf{I}_N + \boldsymbol{\mu} \cdot \boldsymbol{\Lambda} \tag{49}$$

and its inverse follows from the adjoint of this equation as

$$\exp[-\mathrm{i}\mathsf{M}] = {\mu_0}^* \mathsf{I}_N + \boldsymbol{\mu}^* \cdot \boldsymbol{\Lambda} \tag{50}$$

where the star denotes complex conjugation. Plugging these expressions into (48), one encounters triple products of generators $\boldsymbol{\Lambda}$ which when reduced to a linear combination lead to a somewhat involved expression. It is more convenient to first multiply equation (48) with $\exp[\mathrm{i}\mathsf{M}]$, and to work out the terms *quadratic* in the generators. Comparison of the coefficients of $\mathsf{I}_N$ and $\boldsymbol{\Lambda}$ leads to

$$\boldsymbol{\mu} \cdot \boldsymbol{\nu} = \boldsymbol{\mu} \cdot \boldsymbol{N}' \tag{51}$$

$$\mu_0 \boldsymbol{N} + \boldsymbol{N} \odot \boldsymbol{\mu} + \mathrm{i}\boldsymbol{N} \otimes \boldsymbol{\mu} = \mu_0 \boldsymbol{N}' + \boldsymbol{\mu} \odot \boldsymbol{N}' + \mathrm{i}\boldsymbol{\mu} \otimes \boldsymbol{N}'. \tag{52}$$

It is the vector $\boldsymbol{N}'$ which must be determined from these equations. It is useful to rewrite equation (52) with matrices

$$\mathsf{K}_\pm \equiv \mu_0 \mathsf{I}_N + \boldsymbol{\mu} \odot \pm \mathrm{i}\boldsymbol{\mu} \otimes \tag{53}$$

acting on the vectors $\boldsymbol{N}$ and $\boldsymbol{N}'$, respectively,

$$\mathsf{K}_- \boldsymbol{N} = \mathsf{K}_+ \boldsymbol{N}'. \tag{54}$$

The matrix $\mathsf{K}_+$ *does* have an inverse, $\mathsf{K}_+^{-1}$, since it describes the action of $\exp[\mathrm{i}\mathsf{M}]$ on $\mathsf{N}'$ which *is* invertible. Consequently, the vector N' is determined by the relation

$$\begin{aligned} \boldsymbol{N}' &= \mathsf{K}_+^{-1}\mathsf{K}_- \boldsymbol{\nu} \\ &= (\mu_0 \mathsf{I}_N + \boldsymbol{\mu} \odot + \mathrm{i}\boldsymbol{\mu}\otimes)^{-1}(\mu_0 \mathsf{I}_N + \boldsymbol{\mu} \odot - \mathrm{i}\boldsymbol{\mu}\otimes)\boldsymbol{N} \end{aligned} \tag{55}$$

as a function of $\boldsymbol{\mu}$ and $\boldsymbol{N}$ as required.

## 7. Example 1: $SU(2)$

The group $SU(2)$ is used to describe rotations in quantum mechanics and it is isomorphic [5, 9] to the group of unimodular quaternions, $Sl(1, q)$. The multiplication rules of quaternions being known, explicit expressions for the product of two elements of the group $SU(2)$ are obtained easily. In quantum mechanics, as a first step one usually establishes the relation

$$\exp[-\mathrm{i}\boldsymbol{\alpha} \cdot \boldsymbol{\sigma}/2] = \cos(\alpha/2)\mathsf{I}_2 - \mathrm{i}\sin(\alpha/2)\boldsymbol{e}_\alpha \cdot \boldsymbol{\sigma} \qquad \boldsymbol{\alpha} = \alpha \boldsymbol{e}_\alpha \qquad \boldsymbol{e}_\alpha \cdot \boldsymbol{e}_\alpha = 1 \tag{56}$$

by expansion (26) of the exponential exploiting the simple properties of the $(2 \times 2)$ Pauli matrices. The three-vector $\boldsymbol{\alpha}$ determines both the axis of rotation, $\boldsymbol{e}_\alpha$, and the turning angle, $0 \leqslant \alpha \leqslant 4\pi$. Equation (56) is special since the matrix in the exponent and the second term on the right are proportional to each other. As was mentioned before this is due to the fact that the group $SU(2)$ has rank one, implying that all traceless $(2 \times 2)$ matrices are multiples of each other. Calculating the product of two rotations characterized by $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, respectively, one obtains

$$\begin{aligned}\exp[-\mathrm{i}\boldsymbol{\gamma}\cdot\boldsymbol{\sigma}/2] = {} & (\cos(\alpha/2)\cos(\beta/2)+\boldsymbol{\alpha}\cdot\boldsymbol{\beta})\mathbf{I}_2 - \mathrm{i}(\sin(\alpha/2)\cos(\beta/2)\boldsymbol{e}_\alpha \\ & +\cos(\alpha/2)\sin(\beta/2)\boldsymbol{e}_\beta+\sin(\alpha/2)\sin(\beta/2)\boldsymbol{e}_\alpha\wedge\boldsymbol{e}_\beta)\cdot\boldsymbol{\sigma}. \qquad (57)\end{aligned}$$

The vector $\boldsymbol{\gamma}$ which points along the axis of the composed rotation can be read off directly.

Equations (56) and (57) are derived easily from the spectral method. First, write down the quantities introduced in the derivation of equation (39). The spectral theorem (32) involves the projection operators $\mathbf{P}_\pm$ (with $(\pm)\equiv(1,2)$) which for $SU(2)$ are found from (28) to be

$$\mathbf{P}_\pm = \frac{\boldsymbol{\alpha}\cdot\boldsymbol{\sigma}-\alpha_\mp}{\alpha_\pm-\alpha_\mp} = \frac{1}{2}(\mathbf{I}_2\pm\boldsymbol{e}_\alpha\cdot\boldsymbol{\sigma}) \qquad (58)$$

using the fact that the operator $\boldsymbol{\alpha}\cdot\boldsymbol{\sigma}$ has eigenvalues $\alpha_\pm=\pm\alpha$. This immediately reproduces equation (56) via

$$\mathrm{e}^{-\mathrm{i}\alpha_+}\mathbf{P}_+ + \mathrm{e}^{-\mathrm{i}\alpha_-}\mathbf{P}_- = \exp[-\mathrm{i}\boldsymbol{\alpha}\cdot\boldsymbol{\sigma}/2]. \qquad (59)$$

Writing down the right-hand side of equation (46) for the parameters $(\mu_0=\cos(\alpha/2), \boldsymbol{\mu} = -\sin(\alpha/2)\boldsymbol{e}_\alpha)$ and similarly for $(\nu_0,\boldsymbol{\nu})$, one finds that (keep $\odot\equiv 0$ in mind)

$$\gamma_0 = \cos(\alpha/2)\cos(\beta/2)+\sin(\alpha/2)\sin(\beta/2)\boldsymbol{e}_\alpha\cdot\boldsymbol{e}_\beta \qquad (60)$$

$$\boldsymbol{\gamma} = (\sin(\alpha/2)\cos(\beta/2)\boldsymbol{e}_\alpha+\cos(\alpha/2)\sin(\beta/2)\boldsymbol{e}_\beta+\sin(\alpha/2)\sin(\beta/2)\boldsymbol{e}_\alpha\otimes\boldsymbol{e}_\beta)\cdot\boldsymbol{\sigma}. \qquad (61)$$

This reproduces equation (57) because $\otimes$ coincides with the familiar cross product in three dimensions. Note that the results have been derived here without explicitly expanding the exponentials involved.

## 8. Example 2: $SU(4)$

The example of $SU(2)$ is exceptional in the sense that (i) the product $\odot$ is identically zero, (ii) the spectral theorem and its linearized version coincide, and (iii) the matrices $\mathbf{M}$ and $\mathbf{F}$ in equation (44) are multiples of each other. None of these properties holds for $SU(N)$, $N\geqslant 3$, all of which do provide *generic* examples to illustrate the BCH-composition rule. Analytic solvability of the third- and fourth-order characteristic polynomials is a pleasant accident but it does not have any structural consequences in this context. To give a nontrivial example, $SU(4)$ will be studied below.

The interesting point is the reduction of the spectral theorem for an element of $SU(4)$ to linear form. Let us assume that the coefficients $e_n(\boldsymbol{M})$ of the powers of $\mathbf{M}$ in equation (26) have been determined (use $f(x)\equiv\exp[-ix]$) by solving the characteristic polynomial of $\mathbf{M}$ and by employing equations (30) and (31):

$$\begin{aligned}\exp[-\mathrm{i}\mathbf{M}] &= e_0\mathbf{I}_4+e_1\boldsymbol{M}\cdot\boldsymbol{\Lambda}+e_2(\boldsymbol{M}\cdot\boldsymbol{\Lambda})^2+e_3(\boldsymbol{M}\cdot\boldsymbol{\Lambda})^3 \\ &= (e_1+e_2\tfrac{1}{2}\boldsymbol{M}^2+e_3\tfrac{1}{2}(\boldsymbol{M}\odot\boldsymbol{M})\cdot\boldsymbol{M})\mathbf{I}_4+((e_1+e_3\tfrac{1}{2}\boldsymbol{M}^2)\boldsymbol{M} \\ &\quad +e_2\boldsymbol{M}\odot\boldsymbol{M}+e_3(\boldsymbol{M}\odot\boldsymbol{M})\odot\boldsymbol{M})\cdot\boldsymbol{\Lambda} \qquad (62)\end{aligned}$$

and that the reduction has been carried out via equation (19), using the antisymmetry of the $\otimes$ product. Alternatively, one employs formula (39) based on the recursion relations. The quadratic and cubic terms lead to vectors with third powers of $\boldsymbol{M}$ at most. As an identity the left- and right-hand side of (62) must commute which is not trivial only for the last two terms multiplying $\boldsymbol{\Lambda}$:

$$[\boldsymbol{M}\cdot\boldsymbol{\Lambda}, (\boldsymbol{M}\odot\boldsymbol{M})\cdot\boldsymbol{\Lambda}] = 2\mathrm{i}\{\boldsymbol{M}\otimes(\boldsymbol{M}\odot\boldsymbol{M})\}\cdot\boldsymbol{\Lambda} = 0 \qquad (63)$$

as follows from (22) applied to the quantity in curly brackets. Similarly, for the fourth term one finds

$$[\boldsymbol{M} \cdot \boldsymbol{\Lambda}, \{(\boldsymbol{M} \odot \boldsymbol{M}) \odot \boldsymbol{M}\} \cdot \boldsymbol{\Lambda}] = 2\mathrm{i}(\boldsymbol{M} \otimes \{(\boldsymbol{M} \odot \boldsymbol{M}) \odot \boldsymbol{M}\}) \cdot \boldsymbol{\Lambda} = 0. \tag{64}$$

Furthermore, one shows along the same line that these two terms commute among themselves,

$$[(\boldsymbol{M} \odot \boldsymbol{M}) \cdot \boldsymbol{\Lambda}, \{(\boldsymbol{M} \odot \boldsymbol{M}) \odot \boldsymbol{M}\} \cdot \boldsymbol{\Lambda}] = 2\mathrm{i}((\boldsymbol{M} \odot \boldsymbol{M}) \otimes \{(\boldsymbol{M} \odot \boldsymbol{M}) \odot \boldsymbol{M}\}) \cdot \boldsymbol{\Lambda} = 0. \tag{65}$$

Hence, in the process of 'linearization', *three* commuting linear combinations of the $(N^2-1)$ matrices $\boldsymbol{\Lambda}$ arise naturally for $SU(4)$. They span the maximal Abelian subalgebra associated with the element $\boldsymbol{M} \cdot \boldsymbol{\Lambda}$. Knowing (62) it is straightforward to (i) multiply two elements $\exp[-\mathrm{i}\mathbf{M}]$ and $\exp[-\mathrm{i}\mathbf{N}]$ of $SU(4)$; (ii) reduce the product to linear form by removing the single term quadratic in $\boldsymbol{\Lambda}$ in analogy to (46) and (iii) to re-exponentiate using the prescription in (47).

## 9. Summary and discussion

It has been shown how to explicitly calculate BCH relations for the group $SU(N)$. The essential ingredients are: (i) the property that products of generators $\boldsymbol{\Lambda}_j \in SU(N)$ are expressible as linear combinations of generators, and (ii) the reduction of the spectral theorem to linearized form. It has been assumed throughout that the operators involved have no degenerate eigenvalues (this case could be included along the lines shown in [18], for example). Applications of these results are expected to deal with coherent states for the group $SU(N)$, useful for the description of lasers with $N$ levels.

Both steps, (i) and (ii), are based on a surplus of structure in the fundamental representation of the algebra $su(N)$, i.e. the specific form of the anticommutator (11). Therefore, the generalization of this approach to other groups is possible whenever there is a representation such that the *product* of two generators defines another element of the original algebra. In general, this is guaranteed only for the *Lie product*, the commutator, but not for the anticommutator. To put it differently, one must have a representation of the Lie algebra which is closed under both commutation *and* anticommutation of its elements. Apart from $SU(N)$, this property also holds for the general linear group in $N$ dimensions, $GL(N)$, for example.

## Acknowledgment

This work was supported by the Swiss National Science Foundation.

## References

[1] Arecchi F T, Courtens E, Gilmore R and Thomas H 1972 *Phys. Rev.* A **6** 2211
[2] Barnes K J, Dondi P H and Sarkar S C 1972 *J. Phys. A: Math. Gen.* **5** 555
[3] Barnes K J and Delbourgo R 1972 *J. Phys. A: Math. Gen.* **5** 1043
[4] Berry M V, Balazs N L, Tabor M and Voros A 1979 Ann Phys. **122** 26
[5] Biedenharn L C and Louck J 1981 *Angular Momentum in Quantum Physics* (London: Addison Wesley)
[6] Bulgac A and Kusnezow D 1990 *Ann. Phys.* **199** 187
[7] Casati G, Chirikov D V, Ford J and Izraelev F M 1979 *Stochastic Behaviour in Classical and Quantum Systems (Lecture Notes in Physics 93)* ed G Casati and J Ford (Berlin: Springer)
[8] Erikson E 1968 *J. Math. Phys.* **8** 790

[9] Gilmore R 1974 *Lie Groups, Lie Algebras and Some of Their Applications* (New York: Wiley)
[10] Gilmore R 1974 *J. Math. Phys.* **15** 2090
[11] Gilmore R and Yuan B 1989 *J. Chem. Phys.* **91** 917
[12] Greiner W and Müller B 1990 *Quantenmechanik Teil 2: Symmetrien* (Thun: Harri Deutsch)
[13] Halmos P R 1958 *Finite Dimensional Vector Spaces* (New York: van Nostrand)
[14] Itzykson C and Nauenberg M 1966 *Rev. Mod. Phys.* **38** 95
[15] Miller W Jr 1972 *Symmetry Groups and their Applications* (New York: Academic)
[16] Perelomov A 1986 *Generalized Coherent States and their Applications* (Berlin: Springer)
[17] Rhagunathan K, Seetharaman M and Vasan S S 1989 *J. Phys. A: Math. Gen.* **22** L1089
[18] Rhagunathan K, Seetharaman M, Vasan S S and Mary Agnes J 1992 *J. Phys. A: Math. Gen.* **25** 1527
[19] Schatzer L and Weigert St 1998 *Phys. Rev.* A to be published
[20] Twamley J 1993 *Phys. Rev.* A **48** 2627
[21] Wilcox R M 1967 *J. Math. Phys.* **8** 962