Derivative of the exponential map

In the theory of Lie groups, the exponential map is a map from the Lie algebra $\mathbf{g}$ of a Lie group G into G. In case G is a matrix Lie group, the exponential map reduces to the matrix exponential. The exponential map, denoted $\exp:\mathbf{g} \to G$, is analytic and has as such a derivative $\frac{d}{dt}\exp(X(t)):\mathrm{T}\mathbf{g} \to \mathrm{T}G$, where $X(t)$ is a C^1 path in the Lie algebra, and a closely related differential $d\exp:\mathrm{T}\mathbf{g} \to \mathrm{T}G$.[2]

The formula for $d\exp$ was first proved by Friedrich Schur (1891).[3] It was later elaborated by Henri Poincaré (1899) in the context of the problem of expressing Lie group multiplication using Lie algebraic terms.[4] It is also sometimes known as **Duhamel's formula**.

The formula is important both in pure and applied mathematics. It enters into proofs of theorems such as the Baker–Campbell–Hausdorff formula, and it is used frequently in physics[5] for example in quantum field theory, as in the Magnus expansion in perturbation theory, and in lattice gauge theory.

Throughout, the notations $\exp(X)$ and e^X will be used interchangeably to denote the exponential given an argument, *except* when, where as noted, the notations have dedicated *distinct* meanings. The calculus-style notation is preferred here for better readability in equations. On the other hand, the $\exp$-style is sometimes more convenient for inline equations, and is necessary on the rare occasions when there is a real distinction to be made.

In 1899, Henri Poincaré's investigations into group multiplication in Lie algebraic terms led him to the formulation of the universal enveloping algebra.[1]

Statement

The derivative of the exponential map is given by[6]

$$\frac{d}{dt}e^{X(t)} = e^{X(t)}\frac{1-e^{-\mathrm{ad}_X}}{\mathrm{ad}_X}\frac{dX(t)}{dt}. \qquad (1)$$

Explanation

- $X = X(t)$ is a C^1 (continuously differentiable) path in the Lie algebra with derivative $X'(t) = \frac{dX(t)}{dt}$. The argument t is omitted where not needed.
- ad_X is the linear transformation of the Lie algebra given by $\mathrm{ad}_X(Y) = [X, Y]$. It is the adjoint action of a Lie algebra on itself.
- The fraction $\frac{1-\exp(-\mathrm{ad}_X)}{\mathrm{ad}_X}$ is given by the power series

  $$\frac{1-e^{-\mathrm{ad}_X}}{\mathrm{ad}_X} = \sum_{k=0}^{\infty}\frac{(-1)^k}{(k+1)!}(\mathrm{ad}_X)^k. \qquad (2)$$

  derived from the power series of the exponential map of a linear endomorphism, as in matrix exponentiation.[6]
- When G is a matrix Lie group, all occurrences of the exponential are given by their power series expansion.
- When G is *not* a matrix Lie group, $\frac{1-\exp(-\mathrm{ad}_X)}{\mathrm{ad}_X}$ is still given by its power series (**2**), while the other two occurrences of $\exp$ in the formula, which now are the exponential map in Lie theory, refer to the time-one flow of the left invariant vector field X, i.e. element of the Lie algebra as defined in the general case, on the Lie group G viewed as an analytic manifold. This still amounts to exactly the same formula as in the matrix case. Left multiplication of an element of the algebra $\mathbf{g}$ by an element $\exp(X(t))$ of the Lie group is interpreted as applying the differential of the left translation $\mathrm{dL}_{\exp(X(t))}$.
- The formula applies to the case where $\exp$ is considered as a map on matrix space over $\mathbb{R}$ or $\mathbf{C}$, see matrix exponential. When $G = \mathrm{GL}(n, \mathbf{C})$ or $\mathrm{GL}(n, \mathbf{R})$, the notions coincide precisely.

To compute the **differential** $d\exp$ of $\exp$ at X, $d\exp_X\colon \mathrm{T}\mathbf{g}_X \to \mathrm{T}G_{\exp(X)}$, the standard recipe[2]

$$d\exp_X Y = \frac{d}{dt}e^{Z(t)}\Big|_{t=0}, Z(0) = X, Z'(0) = Y$$

is employed. With $Z(t) = X + tY$ the result[6]

$$d\exp_X Y = e^X\frac{1-e^{-\mathrm{ad}_X}}{\mathrm{ad}_X}Y \qquad (3)$$

follows immediately from (**1**). In particular, $d\exp_0:\mathrm{T}\mathbf{g}_0 \to \mathrm{T}G_{\exp(0)} = \mathrm{T}G_e$ is the identity because $\mathrm{T}\mathbf{g}_X \simeq \mathbf{g}$ (since $\mathbf{g}$ is a vector space) and $\mathrm{T}G_e \simeq \mathbf{g}$.

Proof

The proof given below assumes a matrix Lie group. This means that the exponential mapping from the Lie algebra to the matrix Lie group is given by the usual power series, i.e. matrix exponentiation. The conclusion of the proof still holds in the general case, provided each occurrence of $\exp$ is correctly interpreted. See comments on the general case below.

The outline of proof makes use of the technique of differentiation with respect to s of the parametrized expression

$$\Gamma(s,t) = e^{-sX(t)} \frac{\partial}{\partial t} e^{sX(t)}$$

to obtain a first order differential equation for Γ which can then be solved by direct integration in s. The solution is then $e^X\,\Gamma(1,\,t)$.

Lemma

Let Ad denote the adjoint action of the group on its Lie algebra. The action is given by $\mathrm{Ad}_A X = AXA^{-1}$ for $A \in G$, $X \in \mathbf{g}$. A frequently useful relationship between Ad and ad is given by[7][nb 1]

$$\boxed{\mathrm{Ad}_{e^X} = e^{\mathrm{ad}_X}, \quad X \in \mathfrak{g}\,. \qquad \textbf{(4)}}$$

Proof

Using the product rule twice one finds,

$$\frac{\partial\Gamma}{\partial s} = e^{-sX}(-X)\frac{\partial}{\partial t}e^{sX(t)} + e^{-sX}\frac{\partial}{\partial t}\left[X(t)e^{sX(t)}\right] = e^{-sX}\frac{dX}{dt}e^{sX}.$$

Then one observes that

$$\frac{\partial\Gamma}{\partial s} = \mathrm{Ad}_{e^{-sX}}X' = e^{-\mathrm{ad}_{sX}}X',$$

by **(4)** above. Integration yields

$$\Gamma(1,t) = e^{-X(t)}\frac{\partial}{\partial t}e^{X(t)} = \int_0^1 \frac{\partial\Gamma}{\partial s}ds = \int_0^1 e^{-\mathrm{ad}_{sX}}X'\,ds.$$

Using the formal power series to expand the exponential, integrating term by term, and finally recognizing (**2**),

$$\Gamma(1,t) = \int_0^1 \sum_{k=0}^{\infty}\frac{(-1)^k s^k}{k!}(\mathrm{ad}_X)^k\frac{dX}{dt}ds = \sum_{k=0}^{\infty}\frac{(-1)^k}{(k+1)!}(\mathrm{ad}_X)^k\frac{dX}{dt} = \frac{1-e^{-\mathrm{ad}_X}}{\mathrm{ad}_X}\frac{dX}{dt},$$

and the result follows. The proof, as presented here, is essentially the one given in Rossmann (2002). A proof with a more algebraic touch can be found in Hall (2015).[8]

Comments on the general case

The formula in the general case is given by[9]

$$\frac{d}{dt}\exp(C(t)) = \exp(C)\phi(-\mathrm{ad}(C))C',$$

where[nb 2]

$$\phi(z) = \frac{e^z-1}{z} = 1+\frac{1}{2!}z+\frac{1}{3!}z^2+\cdots,$$

which formally reduces to

$$\frac{d}{dt}\exp(C(t)) = \exp(C)\frac{1-e^{-\mathrm{ad}_C}}{\mathrm{ad}_C}\frac{dC(t)}{dt}.$$

Here the $\exp$-notation is used for the exponential mapping of the Lie algebra and the calculus-style notation in the fraction indicates the usual formal series expansion. For more information and two full proofs in the general case, see the freely available Sternberg (2004) reference.

A direct formal argument

An immediate way to see what the answer *must* be, provided it exists is the following. Existence needs to be proved separately in each case. By direct differentiation of the standard limit definition of the exponential, and exchanging the order of differentiation and limit,

$$\begin{aligned}\frac{d}{dt}e^{X(t)} &= \lim_{N\to\infty}\frac{d}{dt}\left(1+\frac{X(t)}{N}\right)^N \\ &= \lim_{N\to\infty}\sum_{k=1}^{N}\left(1+\frac{X(t)}{N}\right)^{N-k}\frac{1}{N}\frac{dX(t)}{dt}\left(1+\frac{X(t)}{N}\right)^{k-1},\end{aligned}$$

where each factor owes its place to the non-commutativity of $X(t)$ and $X'(t)$.

Dividing the unit interval into N sections $\Delta s = \frac{\Delta k}{N}$ ($\Delta k = 1$ since the sum indices are integers) and letting $N \to \infty$, $\Delta k \to dk$, $\frac{k}{N} \to s$, $\Sigma \to \int$ yields

$$\begin{aligned}\frac{d}{dt}e^{X(t)} &= \int_0^1 e^{(1-s)X}X'e^{sX}ds = e^X\int_0^1 \mathrm{Ad}_{e^{-sX}}X'ds\\ &= e^X\int_0^1 e^{-\mathrm{ad}_{sX}}ds X' = e^X\frac{1-e^{-\mathrm{ad}_X}}{\mathrm{ad}_X}\frac{dX}{dt}\ .\end{aligned}$$

Applications

Local behavior of the exponential map

The inverse function theorem together with the derivative of the exponential map provides information about the local behavior of $\exp$. Any C^k, $0 \le k \le \infty, \omega$ map f between vector spaces (here first considering matrix Lie groups) has a C^k inverse such that f is a C^k bijection in an open set around a point x in the domain provided df_x is invertible. From (**3**) it follows that this will happen precisely when

$$\frac{1-e^{\mathrm{ad}_X}}{\mathrm{ad}_X}$$

is invertible. This, in turn, happens when the eigenvalues of this operator are all nonzero. The eigenvalues of $\frac{1-\exp(-\mathrm{ad}_X)}{\mathrm{ad}_X}$ are related to those of ad_X as follows. If g is an analytic function of a complex variable expressed in a power series such that $g(U)$ for a matrix U converges, then the eigenvalues of $g(U)$ will be $g(\lambda_{ij})$, where λ_{ij} are the eigenvalues of U, the double subscript is made clear below.[nb 3] In the present case with $g(U) = \frac{1-\exp(-U)}{U}$ and $U = \mathrm{ad}_X$, the eigenvalues of $\frac{1-\exp(-\mathrm{ad}_X)}{\mathrm{ad}_X}$ are

$$\frac{1-e^{-\lambda_{ij}}}{\lambda_{ij}},$$

where the λ_{ij} are the eigenvalues of ad_X. Putting $\frac{1-\exp(-\lambda_{ij})}{\lambda_{ij}} = 0$ one sees that $d\exp$ is invertible precisely when

$$\lambda_{ij} \ne k2\pi i, k = \pm 1, \pm 2, \ldots.$$

The eigenvalues of ad_X are, in turn, related to those of X. Let the eigenvalues of X be λ_i. Fix an ordered basis e_i of the underlying vector space V such that X is lower triangular. Then

$$Xe_i = \lambda_i e_i + \cdots,$$

with the remaining terms multiples of e_n with $n > i$. Let E_{ij} be the corresponding basis for matrix space, i.e. $(E_{ij})_{kl} = \delta_{ik}\delta_{jl}$. Order this basis such that $E_{ij} < E_{nm}$ if $i - j < n - m$. One checks that the action of ad_X is given by

$$\mathrm{ad}_X E_{ij} = (\lambda_i - \lambda_j)E_{ij} + \cdots \equiv \lambda_{ij}E_{ij} + \cdots,$$

with the remaining terms multiples of $E_{mn} > E_{ij}$. This means that ad_X is lower triangular with its eigenvalues $\lambda_{ij} = \lambda_i - \lambda_j$ on the diagonal. The conclusion is that $d\exp_X$ is invertible, hence $\exp$ is a local bianalytical bijection around X, when the eigenvalues of X satisfy[10][nb 4]

$$\lambda_i - \lambda_j \ne k2\pi i, \quad k = \pm 1, \pm 2, \ldots, \quad 1 \le i, j \le n = \dim V.$$

In particular, in the case of matrix Lie groups, it follows, since $d\exp_0$ is invertible, by the inverse function theorem that $\exp$ is a bi-analytic bijection in a neighborhood of $0 \in \mathbf{g}$ in matrix space. Furthermore, $\exp$, is a bi-analytic bijection from a neighborhood of $0 \in \mathbf{g}$ in $\mathbf{g}$ to a neighborhood of $e \in G$.[11] The same conclusion holds for general Lie groups using the manifold version of the inverse function theorem.

It also follows from the implicit function theorem that $d\exp_\xi$ itself is invertible for ξ sufficiently small.[12]

Derivation of a Baker–Campbell–Hausdorff formula

If $Z(t)$ is defined such that

$$e^{Z(t)} = e^X e^{tY},$$

an expression for $Z(1) = \log(\ \exp X \exp Y\ )$, the Baker–Campbell–Hausdorff formula, can be derived from the above formula,

$$\exp(-Z(t))\frac{d}{dt}\exp(Z(t)) = \frac{1-e^{-\mathrm{ad}_Z}}{\mathrm{ad}_Z}Z'(t).$$

Its left-hand side is easy to see to equal Y. Thus,

$$Y = \frac{1-e^{-\mathrm{ad}_Z}}{\mathrm{ad}_Z}Z'(t),$$

and hence, formally,[13][14]

$$Z'(t) = \frac{\mathrm{ad}_Z}{1-e^{-\mathrm{ad}_Z}}Y \equiv \psi\left(e^{\mathrm{ad}_Z}\right)Y, \quad \psi(w) = \frac{w\log w}{w-1} = 1 + \sum_{m=1}^\infty \frac{(-1)^{m+1}}{m(m+1)}(w-1)^m, \|w\| < 1.$$

However, using the relationship between Ad and ad given by **(4)**, it is straightforward to further see that

$$e^{\mathrm{ad}_Z} = e^{\mathrm{ad}_X} e^{t\mathrm{ad}_Y}$$

and hence

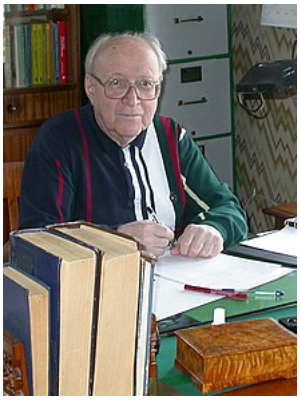
$$Z'(t) = \psi\left(e^{\mathrm{ad}_X} e^{t\mathrm{ad}_Y}\right) Y.$$

Putting this into the form of an integral in t from 0 to 1 yields,

$$Z(1) = \log(\exp X \exp Y) = X + \left(\int_0^1 \psi\left(e^{\mathrm{ad}_X}\ e^{t\,\mathrm{ad}_Y}\right)\, dt\right)\, Y,$$

an integral formula for $Z(1)$ that is more tractable in practice than the explicit Dynkin's series formula due to the simplicity of the series expansion of ψ. Note this expression consists of $X+Y$ and nested commutators thereof with X or Y. A textbook proof along these lines can be found in Hall (2015) and Miller (1972).

Derivation of Dynkin's series formula


Eugene Dynkin at home in 2003. In 1947 Dynkin proved the explicit BCH series formula.[15] Poincaré, Baker, Campbell and Hausdorff were mostly concerned with the *existence* of a bracket series, which suffices in many applications, for instance, in proving central results in the Lie correspondence.[16][17] Photo courtesy of the Dynkin Collection.

Dynkin's formula mentioned may also be derived analogously, starting from the parametric extension

$$e^{Z(t)} = e^{tX} e^{tY},$$

whence

$$e^{-Z(t)} \frac{de^{Z(t)}}{dt} = e^{-t\,\mathrm{ad}_Y} X + Y\ ,$$

so that, using the above general formula,

$$Z' = \frac{\mathrm{ad}_Z}{1 - e^{-\mathrm{ad}_Z}}\ \left(e^{-t\,\mathrm{ad}_Y} X + Y\right) = \frac{\mathrm{ad}_Z}{e^{\mathrm{ad}_Z} - 1}\ \left(X + e^{t\,\mathrm{ad}_X} Y\right).$$

Since, however,

$$\begin{aligned}\mathrm{ad}_Z &= \log(\exp(\mathrm{ad}_Z)) = \log(1 + (\exp(\mathrm{ad}_Z) - 1)) \\ &= \sum_{n=1}^{\infty} \frac{(-)^{n+1}}{n} (\exp(\mathrm{ad}_Z) - 1)^n\ , \quad \|\mathrm{ad}_Z\| < \log 2\ ,\end{aligned}$$

the last step by virtue of the Mercator series expansion, it follows that

$$Z' = \sum_{n=1}^{\infty} \frac{(-)^{n-1}}{n} \left(e^{\mathrm{ad}_Z} - 1\right)^{n-1}\ \left(X + e^{t\,\mathrm{ad}_X} Y\right)\ , \qquad \textbf{(5)}$$

and, thus, integrating,

$$Z(1) = \int_0^1 dt\ \frac{dZ(t)}{dt} = \sum_{n=1}^{\infty} \frac{(-)^{n-1}}{n} \int_0^1 dt\ \left(e^{t\,\mathrm{ad}_X} e^{t\mathrm{ad}_Y} - 1\right)^{n-1}\ \left(X + e^{t\,\mathrm{ad}_X} Y\right).$$

It is at this point evident that the qualitative statement of the BCH formula holds, namely Z lies in the Lie algebra generated by X, Y and is expressible as a series in repeated brackets **(A)**. For each k, terms for each partition thereof are organized inside the integral $\int dt\ t^{k-1}$. The resulting Dynkin's formula is then

$$\boxed{Z = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} \sum_{s \in S_k} \frac{1}{i_1 + j_1 + \cdots + i_k + j_k} \frac{\left[X^{(i_1)} Y^{(j_1)} \ldots X^{(i_k)} Y^{(j_k)}\right]}{i_1! j_1! \cdots i_k! j_k!}, \quad i_r, j_r \geq 0, \quad i_r + j_r > 0, \quad 1 \leq r \leq k.}$$

For a similar proof with detailed series expansions, see Rossmann (2002).

Combinatoric details

Change the summation index in (5) to $k = n - 1$ and expand

$$\frac{dZ}{dt} = \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} \left\{\left(e^{\mathrm{ad}_{tX}} e^{\mathrm{ad}_{tY}} - 1\right)^k X + \left(e^{\mathrm{ad}_{tX}} e^{\mathrm{ad}_{tY}} - 1\right)^k e^{\mathrm{ad}_{tX}} Y\right\} \qquad \textbf{(97}$$

in a power series. To handle the series expansions simply, consider first $Z = \log(e^X e^Y)$. The $\log$-series and the $\exp$-series are given by

$$\log(A) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k}(A-I)^k, \quad \text{and} \quad e^X = \sum_{k=0}^{\infty} \frac{X^k}{k!}$$

respectively. Combining these one obtains

$$\log\left(e^X e^Y\right) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k}\left(e^X e^Y - I\right)^k = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k}\left(\sum_{i=0}^{\infty} \frac{X^i}{i!}\sum_{j=0}^{\infty} \frac{Y^j}{j!} - I\right)^k = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k}\left(\sum_{i,j\geq 0, i+j>1}^{\infty} \frac{X^i Y^j}{i!j!}\right)^k.$$

This becomes

$$Z = \log\left(e^X e^Y\right) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} \sum_{s\in S_k} \frac{X^{i_1}Y^{j_1}\cdots X^{i_k}Y^{j_k}}{i_1!j_1!\cdots i_k!j_k!}, \quad i_r, j_r \geq 0, \quad i_r + j_r > 0, \quad 1 \leq r \leq k, \qquad \text{(99)}$$

where S_k is the set of all sequences $s = (i_1, j_1, \ldots, i_k, j_k)$ of length $2k$ subject to the conditions in **(99)**.

Now substitute $(e^X e^Y - 1)$ for $(e^{\mathrm{ad}_{tX}} e^{\mathrm{ad}_{tY}} - 1)$ in the LHS of (**98**). Equation **(99)** then gives

$$\frac{dZ}{dt} = \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} \sum_{s\in S_k, i_{k+1}\geq 0} t^{i_1+j_1+\cdots+i_k+j_k} \frac{{\mathrm{ad}_X}^{i_1}{\mathrm{ad}_Y}^{j_1}\cdots{\mathrm{ad}_X}^{i_k}{\mathrm{ad}_Y}^{j_k}}{i_1!j_1!\cdots i_k!j_k!}X$$
$$+ t^{i_1+j_1+\cdots+i_k+j_k+i_{k+1}} \frac{{\mathrm{ad}_X}^{i_1}{\mathrm{ad}_Y}^{j_1}\cdots{\mathrm{ad}_X}^{i_k}{\mathrm{ad}_Y}^{j_k}X^{i_{k+1}}}{i_1!j_1!\cdots i_k!j_k!i_{k+1}!}Y, \quad i_r, j_r \geq 0, \quad i_r + j_r > 0, \quad 1 \leq r \leq k,$$

or, with a switch of notation, see An explicit Baker–Campbell–Hausdorff formula,

$$\frac{dZ}{dt} = \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} \sum_{s\in S_k, i_{k+1}\geq 0} t^{i_1+j_1+\cdots+i_k+j_k} \frac{\left[X^{(i_1)}Y^{(j_1)}\cdots X^{(i_k)}Y^{(j_k)}X\right]}{i_1!j_1!\cdots i_k!j_k!}$$
$$+ t^{i_1+j_1+\cdots+i_k+j_k+i_{k+1}} \frac{\left[X^{(i_1)}Y^{(j_1)}\cdots X^{(i_k)}Y^{(j_k)}X^{(i_{k+1})}Y\right]}{i_1!j_1!\cdots i_k!j_k!i_{k+1}!}, \quad i_r, j_r \geq 0, \quad i_r + j_r > 0, \quad 1 \leq r \leq k.$$

Note that the summation index for the rightmost $e^{\mathrm{ad}_{tX}}$ in the second term in (**97**) is denoted i_{k+1}, but is *not* an element of a sequence $s \in S_k$. Now integrate $Z = Z(1) = \int \frac{dZ}{dt} dt$, using $Z(0) = 0$,

$$Z = \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} \sum_{s\in S_k, i_{k+1}\geq 0} \frac{1}{i_1+j_1+\cdots+i_k+j_k+1} \frac{\left[X^{(i_1)}Y^{(j_1)}\cdots X^{(i_k)}Y^{(j_k)}X\right]}{i_1!j_1!\cdots i_k!j_k!}$$
$$+ \frac{1}{i_1+j_1+\cdots+i_k+j_k+i_{k+1}+1} \frac{\left[X^{(i_1)}Y^{(j_1)}\cdots X^{(i_k)}Y^{(j_k)}X^{(i_{k+1})}Y\right]}{i_1!j_1!\cdots i_k!j_k!i_{k+1}!}, \quad i_r, j_r \geq 0, \quad i_r + j_r > 0, \quad 1 \leq r \leq k.$$

Write this as

$$Z = \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} \sum_{s\in S_k, i_{k+1}\geq 0} \frac{1}{i_1+j_1+\cdots+i_k+j_k+(i_{k+1}=1)+(j_{k+1}=0)} \frac{\left[X^{(i_1)}Y^{(j_1)}\cdots X^{(i_k)}Y^{(j_k)}X^{(i_{k+1}=1)}Y^{(j_{k+1}=0)}\right]}{i_1!j_1!\cdots i_k!j_k!(i_{k+1}=1)!(j_{k+1}=0)!}$$
$$+ \frac{1}{i_1+j_1+\cdots+i_k+j_k+i_{k+1}+(j_{k+1}=1)} \frac{\left[X^{(i_1)}Y^{(j_1)}\cdots X^{(i_k)}Y^{(j_k)}X^{(i_{k+1})}Y^{(j_{k+1}=1)}\right]}{i_1!j_1!\cdots i_k!j_k!i_{k+1}!(j_{k+1}=1)!},$$
$$(i_r, j_r \geq 0, \quad i_r + j_r > 0, \quad 1 \leq r \leq k).$$

This amounts to

$$Z = \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} \sum_{s\in S_{k+1}} \frac{1}{i_1+j_1+\cdots+i_k+j_k+i_{k+1}+j_{k+1}} \frac{\left[X^{(i_1)}Y^{(j_1)}\cdots X^{(i_k)}Y^{(j_k)}X^{(i_{k+1})}Y^{(j_{k+1})}\right]}{i_1!j_1!\cdots i_k!j_k!i_{k+1}!j_{k+1}!},$$

where $i_r, j_r \geq 0, \quad i_r + j_r > 0, \quad 1 \leq r \leq k+1$, using the simple observation that $[T, T] = 0$ for all T. That is, in (**100**), the leading term vanishes unless j_{k+1} equals 0 or 1, corresponding to the first and second terms in the equation before it. In case $j_{k+1} = 0$, i_{k+1} must equal 1, else the term vanishes for the same reason ($i_{k+1} = 0$ is not allowed). Finally, shift the index, $k \to k-1$,

$$Z = \log e^X e^Y = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} \sum_{s\in S_k} \frac{1}{i_1+j_1+\cdots+i_k+j_k} \frac{\left[X^{(i_1)}Y^{(j_1)}\cdots X^{(i_k)}Y^{(j_k)}\right]}{i_1!j_1!\cdots i_k!j_k!}, \; i_r, j_r \geq 0, \; i_r + j_r > 0, \; 1 \leq r \leq k.$$

This is Dynkin's formula. The striking similarity with (99) is not accidental: It reflects the **Dynkin–Specht–Wever map**, underpinning the original, different, derivation of the formula.[15] Namely, *if*

$$X^{i_1}Y^{j_1}\cdots X^{i_k}Y^{j_k}$$

is expressible as a bracket series, then necessarily[18]

$$X^{i_1}Y^{j_1}\cdots X^{i_k}Y^{j_k} = \frac{\left[X^{(i_1)}Y^{(j_1)}\cdots X^{(i_k)}Y^{(j_k)}\right]}{i_1+j_1+\cdots+i_k+j_k}. \qquad \textbf{(B)}$$

Putting observation **(A)** and theorem (**B**) together yields a concise proof of the explicit BCH formula.

See also

- Adjoint representation (ad)
- Baker-Campbell-Hausdorff formula
- Exponential map
- Matrix exponential
- Matrix logarithm
- Magnus expansion

Remarks

1. A proof of the identity can be found in here. The relationship is simply that between a representation of a Lie group and that of its Lie algebra according to the Lie correspondence, since both Ad and ad are representations with $\mathrm{ad} = d\mathrm{Ad}$.
2. It holds that

   $$\tau(\log z)\phi(-\log z) = 1$$

   for |z − 1| < 1 where

   $$\tau(w) = \frac{w}{1-e^{-w}}.$$

   Here, τ is the exponential generating function of

   $$(-1)^k b_k,$$

   where b_k are the Bernoulli numbers.
3. This is seen by choosing a basis for the underlying vector space such that U is triangular, the eigenvalues being the diagonal elements. Then U^k is triangular with diagonal elements λ_i^k. It follows that the eigenvalues of U are $f(\lambda_i)$. See Rossmann 2002, Lemma 6 in section 1.2.
4. Matrices whose eigenvalues λ satisfy $|\mathrm{Im}\ \lambda| < \pi$ are, under the exponential, in bijection with matrices whose eigenvalues μ are not on the negative real line or zero. The λ and μ are related by the complex exponential. See Rossmann (2002) Remark 2c section 1.2.

Notes

1. Schmid 1982
2. Rossmann 2002 Appendix on analytic functions.
3. Schur 1891
4. Poincaré 1899
5. Suzuki 1985
6. Rossmann 2002 Theorem 5 Section 1.2
7. Hall 2015 Proposition 3.35
8. See also Tuynman 1995 from which Hall's proof is taken.
9. Sternberg 2004 This is equation (1.11).
10. Rossmann 2002 Proposition 7, section 1.2.
11. Hall 2015 Corollary 3.44.
12. Sternberg 2004 Section 1.6.
13. Hall 2015Section 5.5.
14. Sternberg 2004 Section 1.2.
15. Dynkin 1947
16. Rossmann 2002 Chapter 2.
17. Hall 2015 Chapter 5.
18. Sternberg 2004 Chapter 1.12.2.

References

- Dynkin, Eugene Borisovich (1947), "Вычисление коэффициентов в формуле Campbell–Hausdorff" [Calculation of the coefficients in the Campbell–Hausdorff formula], *Doklady Akademii Nauk SSSR* (in Russian), **57**: 323–326 ; translation from Google books (https://books.google.com/books?id=D9ZF5O_JH2gC&dq=Dynkin+Yushkevich++Campbell&pg=PA31).
- Hall, Brian C. (2015), *Lie groups, Lie algebras, and representations: An elementary introduction*, Graduate Texts in Mathematics, vol. 222 (2nd ed.), Springer, ISBN 978-3319134666
- Miller, Wllard (1972), *Symmetry Groups and their Applications*, Academic Press, ISBN 0-12-497460-0
- Poincaré, H. (1899), "Sur les groupes continus", *Cambridge Philos. Trans.*, **18**: 220–55
- Rossmann, Wulf (2002), *Lie Groups – An Introduction Through Linear Groups*, Oxford Graduate Texts in Mathematics, Oxford Science Publications, ISBN 0-19-859683-9
- Schur, F. (1891), "Zur Theorie der endlichen Transformationsgruppen", *Abh. Math. Sem. Univ. Hamburg*, **4**: 15–32
- Suzuki, Masuo (1985). "Decomposition formulas of exponential operators and Lie exponentials with some applications to quantum mechanics and statistical physics". *Journal of Mathematical Physics*. **26** (4): 601–612. Bibcode:1985JMP....26..601S (https://ui.adsabs.harvard.edu/abs/1985JMP....26..601S). doi:10.1063/1.526596 (https://doi.org/10.1063%2F1.526596).
- Tuynman (1995), "The derivation of the exponential map of matrices", *Amer. Math. Monthly*, **102** (9): 818–819, doi:10.2307/2974511 (https://doi.org/10.2307%2F2974511), JSTOR 2974511 (https://www.jstor.org/stable/2974511)
- Veltman, M, 't Hooft, G & de Wit, B (2007). "Lie Groups in Physics", online lectures (http://www.staff.science.uu.nl/~hooft101/lectures/lieg07.pdf).
- Wilcox, R. M. (1967). "Exponential Operators and Parameter Differentiation in Quantum Physics". *Journal of Mathematical Physics*. **8** (4): 962–982. Bibcode:1967JMP.....8..962W (https://ui.adsabs.harvard.edu/abs/1967JMP.....8..962W). doi:10.1063/1.1705306 (https://doi.org/10.10

63%2F1.1705306).

External links

- Sternberg, Shlomo (2004), *Lie Algebras* (http://www.math.harvard.edu/~shlomo/docs/lie_algebras.pdf) (PDF)
- Schmid, Wilfried (1982), "Poincaré and Lie groups" (https://www.ams.org/journals/bull/1982-06-02/S0273-0979-1982-14972-2/S0273-0979-1982-14972-2.pdf) (PDF), *Bull. Amer. Math. Soc.*, **6** (2): 175–186, doi:10.1090/s0273-0979-1982-14972-2 (https://doi.org/10.1090%2Fs0273-0979-1982-14972-2)