# 指数映射的导数

wikipedia

16 March 2023

在李群理论中，指数映射是从一个李群 $G$ 的李代数 $\mathfrak{g}$ 到 $G$ 的映射。如果 $G$ 是矩阵李群，则指数映射可约化为矩阵指数。指数映射，标志为 $\exp : \mathfrak{g} \to G$，是解析的，并且具有这样一个导数 $\frac{d}{dt}\exp(X(t)) : T\mathfrak{g} \to TG$，其中 $X(t)$ 是在李代数中的一个 $C^1$ 路径，以及一个密切相关的微分 $d\exp : T\mathfrak{g} \to TG$。[2]

Friedrich Schur (1891) 首先证明了 $d\exp$ 的公式。[3] 后来，Henri Poincaré (1899) 在用李代数项表示李群乘法的问题中对此进行了阐述。[4] 该公式有时也被称为**杜哈默尔公式** (**Duhamel's formula**)。

这个公式在纯数学和应用数学中都很重要。它用于证明 Baker-Campbell-Hausdorff (BCH) 公式等定理，在物理学中 [5] 经常使用，例如在量子场论中，在微扰论中的 Magnus 展开式中，以及在格点规范理论中。

贯穿始终，符号 $\exp(X)$ 和 $e^X$ 将互换使用，以标志给定参数的指数，除非如前所述，符号具有专用的不同含义。为了在方程中具有更好的可读性，这里首选微积分形式的符号。另一方面，exp 样式有时对于内联公式更方便，并且在需要进行真正区分的少数情况下是必需的。

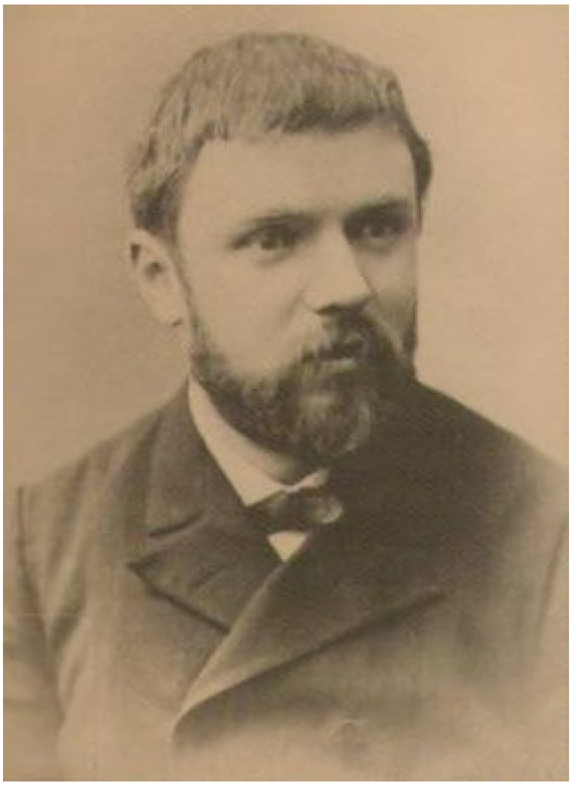

图 1: 在 1899 年，Henri Poincaré 对李代数项中群乘法的研究导致了泛包络代数的形成。[1]

## 1 声明

指数映射的导数给出为：[6]

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

**解释**

- $X = X(t)$ 是在李代数中的一条 $C^1$ (连续可微) 路径，其导数为 $X'(t) = \frac{dX(t)}{dt}$。在不需要时可省略参数 $t$。
- $\mathrm{ad}_X$ 是李代数的线性变换，给出为 $\mathrm{ad}_X(Y) = [X, Y]$。它是李代数对其自身的伴随作用。
- 分数 $\frac{1-\exp(-\mathrm{ad}_X)}{\mathrm{ad}_X}$ 由幂级数给出为

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

  它由线性自同态的指数映射的幂级数推导出，如在矩阵幂运算中。[6]

- 当 $G$ 是矩阵李群时，所有指数的出现都由它们的幂级数展开式给出。
- 当 $G$ 不是矩阵李群时，$\frac{1-\exp(-\mathrm{ad}_X)}{\mathrm{ad}_X}$ 仍然由其幂级数方程 (2) 给出，而公式中 exp 的另外两个出现，现在是在李理论中的指数映射，指左不变向量场 $X$ 的“时间-1”的流，即在一般情况下定义的李代数元素，关于李群 $G$ 的分析流形。这仍然与矩阵情况下的公式完全相同。代数 $\mathfrak{g}$ 的元素与李群的元素 $\exp(X(t))$ 的左乘法被解释为应用左平移 $\mathrm{dL}_{\exp(X(t))}$ 的微分。
- 该公式适用于 exp 被认为是在 $\mathbb{R}$ 或 $\mathbf{C}$ 上的矩阵空间的映射的情况，请参见矩阵指数。当 $G = \mathrm{GL}(n, \mathbf{C})$ 或 $\mathrm{GL}(n, \mathbf{R})$ 时，这些概念正好重合。

为了计算在 $X$ 处 exp 的微分 $d\exp$，$d\exp_X : \mathrm{T}\mathfrak{g}_X \to \mathrm{T}G_{\exp(X)}$，标准配方 [2]

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

被使用。对于 $Z(t) = X + tY$，结果 [6]

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

立刻跟随在方程 (1) 之后。特别地，$d\exp_0 : \mathrm{T}\mathfrak{g}_0 \to \mathrm{T}G_{\exp(0)} = \mathrm{T}G_e$ 是恒等式，因为 $\mathrm{T}\mathfrak{g}_X \simeq \mathfrak{g}$ (因为 $\mathfrak{g}$ 是向量空间) 并且 $\mathrm{T}G_e \simeq \mathfrak{g}$。

# 2 证明

下面给出的证明假设一个矩阵李群。这意味着从李代数到矩阵李群的指数映射由通常的幂级数给出，即矩阵指数化。只要 exp 的每一次出现都得到正确的解释，证明结论在一般情况下仍然成立。见下文关于一般情况的评论。

证明大纲使用了参数化表达式相对于 $s$ 微分的技巧

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

以获得 $\Gamma$ 的一阶微分方程，其可在 $s$ 中的直接积分法求解。则该解为 $e^X \Gamma(1, \mathrm{t})$ 。

**引理.** *设 Ad 标志群在其李代数上的伴随作用。对于 $A \in G, X \in \mathfrak{g}$，其作用给出为 $\mathrm{Ad}_A X = AXA^{-1}$。一个 Ad 和 ad 之间的常用关系给出为*[1] *[7]*

$$\boxed{\mathrm{Ad}_{e^X} = e^{\mathrm{ad}_X}, X \in \mathfrak{g}.} \tag{4}$$

[1]可在这里找到该恒等式的一个证明。这个关系很简单，根据李对应 (Lie correspondence) 关系，这就是一个李群的表示和它的李代数的表示之间的关系，因为 Ad 和 ad 都是 $\mathrm{ad} = d\mathrm{Ad}$ 的表示。

证明. 使用乘积规则两次寻找，

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

则通过上述方程 (4)，我们观察到

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

集成产出

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

使用形式幂级数展开指数，逐项积分，并最后识别方程 (2)，

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

接着就得出结果。这里给出的证明基本上是 Rossmann (2002) 给出的证明。Hall 2015 提供了一个更具代数意义的证明。[8] □

## 2.1 对一般情况的评论

一般情况下的公式给出为 [9]

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

其中[2]

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

将其形式上约化为

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

在这里，exp 符号用于李代数的指数映射，而分数中的微积分式符号表示通常的形式级数展开。有关一般情况下的更多信息和两个完整证明，请参阅免费提供的 Sternberg (2004) 参考资料。

## 2.2 直接形式论证

如果答案存在的话，一种立即看到答案的方法如下。在每种情况下都需要分别证明存在性。通过直接微分指数的标准极限定义，并交换微分与极限的阶数，

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

[2]它认为

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

对于 $|z-1|<1$ 成立，其中

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

这里，$\tau$ 是下式的指数生成函数

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

其中 $b_k$ 是 Bernoulli 数字。

其中，每个因子的位置取决于 $X(t)$ 和 $X'(t)$ 的非交换性。

将单位区间分成 $N$ 段 $\Delta s=\frac{\Delta k}{N}$ (由于总和指数为整数，因此 $\Delta k=1$)，并设 $N\to\infty,\Delta k\to dk,\frac{k}{N}\to s,\Sigma\to\int$，产出

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

# 3 应用

## 3.1 指数映射的局部性态

逆函数定理与与指数映射的导数一起提供关于 exp 的局部性态的信息。任意 $C^k,0\le k\le\infty$，在向量空间 (这里首先考虑矩阵李群) 之间，$\omega$ 映射 $f$ 有一个 $C^k$ 逆，使得 $f$ 在定义域中在围绕一个点 $X$ 的开集中是一个 $C^k$ 双射，只要 $df_X$ 是可逆的。从方程 (3) 可以看出，这将精确地发生在当

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

是可逆时。当这个算子的特征值都非零时，就会发生这种情况。$\frac{1-\exp(-\mathrm{ad}_X)}{\mathrm{ad}_X}$ 的特征值与 $\mathrm{ad}_X$ 的特征值相关，具体如下。如果 $g$ 是一个用幂级数表达的复变量的解析函数，使得矩阵 $U$ 的 $g(U)$ 收敛，则 $g(U)$ 的特征值为 $g(\lambda_{ij})$，其中 $\lambda_{ij}$ 是 $U$ 的特征值，双下标使得下面的表述清晰。[3]在当前情况下，当 $g(U)=\frac{1-\exp(-U)}{U}$ 和 $U=\mathrm{ad}_X$ 时，则 $\frac{1-\exp(-\mathrm{ad}_X)}{\mathrm{ad}_X}$ 的特征值为

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

其中 $\lambda_{ij}$ 是 $\mathrm{ad}_X$ 的特征值。设置 $\frac{1-\exp(-\lambda_{ij})}{\lambda_{ij}}=0$，我们可以看到 $d\exp$ 是精确可逆的，当

$$\lambda_{ij}\neq k2\pi i,k=\pm1,\pm2,\ldots.$$

$\mathrm{ad}_X$ 的特征值依次与 $X$ 的特征值相关。设 $X$ 的特征值为 $\lambda_i$。固定底层向量空间 $V$ 的有序基 $e_i$，以使得 $X$ 是下三角矩阵。则

$$Xe_i=\lambda_ie_i+\cdots,$$

带着剩余项乘以 $e_n$，其中 $n>i$。设 $E_{ij}$ 为矩阵空间的相应基，即 $(E_{ij})_{kl}=\delta_{ik}\delta_{jl}$。以此基为序，如果 $i-j<n-m$，以便 $E_{ij}<E_{nm}$。我们检查 $\mathrm{ad}_X$ 的作用，给出为

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

带着剩余项为 $E_{mn}>E_{ij}$ 的倍数。这意味着 $\mathrm{ad}_X$ 是下三角矩阵，其特征值 $\lambda_{ij}=\lambda_i-\lambda_j$ 在对角线上。结论是 $d\exp_X$ 是可逆的，因此 exp 是围绕 $X$ 的局部双解析双射，当 $X$ 的特征值满足条件时[4] [10]

$$\lambda_i-\lambda_j\neq k2\pi i,\quad k=\pm1,\pm2,\ldots,\quad 1\le i,j\le n=\mathrm{dim}V.$$

[3]这可以通过为底层向量空间选择一个基来看出，使得 $U$ 是三角形矩阵，特征值是对角元素。则 $U^k$ 是三角形矩阵，具有对角元素 $\lambda_j^k$。因此，$U$ 的特征值为 $f(\lambda_j)$。参见 Rossmann (2002) 第 1.2 节的引理 6。

[4]特征值 $\lambda$ 满足 $|\mathrm{Im}\lambda|<\pi$ 的矩阵是在指数下与特征值 $\mu$ 不在负实数线上或为零的矩阵双射。$\lambda$ 和 $\mu$ 由复指数关联。参见 Rossmann (2002) 第 1.2 节的备注 2c。

特别地，在矩阵李群的情况下，由于 $d\exp_0$ 是可逆的，因此根据反函数定理，在矩阵空间中在 $0 \in \mathfrak{g}$ 邻域中，exp 是一个双解析双射。此外，exp 是在 $\mathfrak{g}$ 中从 $0 \in \mathfrak{g}$ 的邻域到 $e \in G$ 的邻域的双解析双射。[11] 使用流形版本的反函数定理，对于一般的李群，同样的结论成立。

也由此得出，根据隐函数定理，当 $\xi$ 足够小时，$\exp_\xi$ 本身是可逆的。[12]

### 3.2 Baker-Campbell-Hausdorff 公式的推导

如果 $Z(t)$ 被定义，使得

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

对于 $Z(1) = \log(\exp X \exp Y)$ 的表达式，Baker-Campbell-Hausdorff 公式，可以从上述公式中推导出来，

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

它的左侧很容易看出等于 $Y$，因此，

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

并因此，在形式上，[13][14]

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

然而，通过方程 (4) 给出的 Ad 和 ad 之间的关系，可以直接进一步看出

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

并因此

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

将其设置为在 $t$ 中从 0 到 1 的积分形式，产出，

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

对于 $Z(1)$，由于 $\psi$ 级数展开式的简单性，在实际中比显式 Dynkin 级数公式更易于处理。注意，该表达式由 $X + Y$ 及其与 $X$ 或 $Y$ 的嵌套交换子组成。沿这条线的教科书证明可在 Hall (2015) 和 Miller (1972) 中找到。

### 3.3 Dynkin 级数公式的推导

所提到的 Dynkin 公式也可以类似地推导出，从参数扩展开始

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

其中

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

因此，使用上述一般公式，

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

然而，由于，

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

最后一步，通过 Mercator 级数展开式，接着是

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

图 2: 2003 年，Eugene Dynkin 在家中。1947 年 Dynkin 证明了显式 BCH 级数公式。[15] Poincaré、Baker、Campbell 和 Hausdorff 主要研究括号级数的存在性，这在许多应用中都是足够的，例如证明了李对应中的中心结果。[16][17] 照片由 Dynkin 收藏提供。

并就此集成方程，

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

显然，BCH 公式的定性陈述成立，即 $Z$ 位于由 $X$，$Y$ 生成的李代数中，并且可表达为重复括号方程 (A) 中的级数。对于每个 $k$，其每个分区的项都组织在积分 $\int dt\, t^{k-1}$ 内。由此得到的 Dynkin 公式是

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

有关详细级数展开式的类似证明，参见 Rossmann (2002)。

### 3.4 组合细节

将方程 (5) 中的求和索引更改为 $k = n - 1$，并在幂级数中展开为

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

为了简单地处理级数扩展，首先考虑 $Z = \log\left(e^Xe^Y\right)$。log 级数和 exp 级数分别给出为

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

组合这些，我们获得

$$\begin{aligned}\log\left(e^X e^Y\right) &= \sum_{k=1}^{\infty}\frac{(-1)^{k+1}}{k}\left(e^X e^Y - I\right)^k \\ &= \sum_{k=1}^{\infty}\frac{(-1)^{k+1}}{k}\left(\sum_{i=0}^{\infty}\frac{X^i}{i!}\sum_{j=0}^{\infty}\frac{Y^j}{j!} - I\right)^k \\ &= \sum_{k=1}^{\infty}\frac{(-1)^{k+1}}{k}\left(\sum_{i,j\geq 0, i+j>1}^{\infty}\frac{X^i Y^j}{i!j!}\right)^k. \end{aligned} \tag{98}$$

这将变为

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

其中 $S_k$ 是长度为 $2k$ 的所有序列 $s = (i_1, j_1, \ldots, i_k, j_k)$ 的集合，受方程 (99) 中的条件约束。

现在将方程 (98) 左侧中的 $\left(e^X e^Y - 1\right)$ 替换为 $\left(e^{\mathrm{ad}_{tX}} e^{\mathrm{ad}_{tY}} - 1\right)$。则方程 (99) 给出

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

或者，使用一个符号切换，参见“An explicit Baker-Campbell-Hausdorff formula”一文，

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

注意，在方程 (97) 中第二项中最右边 $e^{\mathrm{ad}_{tX}}$ 的求和索引标志为 $i_{k+1}$，但这**不是**序列 $s\in S_k$ 的一个元素。现在使用 $Z(0)=0$，积分 $Z = Z(1) = \int\frac{dZ}{dt}dt$,

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

将其写为

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

这相当于

$$Z=\sum_{k=0}^{\infty}\frac{(-1)^k}{k+1}\sum_{s\in S_{k+1}}\frac{1}{i_1+j_1+\cdots+i_k+j_k+i_{k+1}+j_{k+1}}\frac{\left[X^{(i_1)}Y^{(j_1)}\dots X^{(i_k)}Y^{(j_k)}X^{(i_{k+1})}Y^{(j_{k+1})}\right]}{i_1!j_1!\cdots i_k!j_k!i_{k+1}!j_{k+1}!}, \tag{100}$$

其中 $i_r,j_r\geq 0,\quad i_r+j_r>0,\quad 1\leq r\leq k+1$，使用简单的观察，即对于所有的 $T$，$[T,T]=0$ 。也就是，在方程 (100) 中，除非 $j_{k+1}$ 等于 0 或 1，否则前导项消失，对应于它前面的方程中的第一项和第二项。如果 $j_{k+1}=0$，则 $i_{k+1}$ 必须等于 1，否则该项消失的原因相同 (不允许 $i_{k+1}=0$)。最后，移动索引 $k\to k-1$，

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

这就是 Dynkin 的公式。与方程 (99) 惊人的相似不是偶然的：它反映了 Dynkin-Specht-Wever 映射，加强了公式的原始的、不同的推导。[15] 即，如果

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

表达为一个括号级数，则必然地 [18]

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

将观测方程 (A) 和定理方程 (B) 放在一起，就得到显式 BCH 公式的简明证明。

# 4 See also

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

# 5 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 2015 Section 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.

# 6 References


- Dynkin, Eugene Borisovich (1947), " 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=D9ZF50_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, WIlard (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/10.1063%2F1.1705306).


# 7 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)