Baker-Campbell-Hausdorff-Dynkin formula

From Wikipedia, the free encyclopedia

In mathematics, the Baker-Campbell-Hausdorff-Dynkin formula is the solution to

\[ Z = \log(e^X e^Y) \]

for non-commuting \( X \) and \( Y \). It links Lie Groups to Lie Algebras, by expressing the logarithm of the product of two Lie group elements as a Lie algebra element in canonical coordinates, a significant guiding connection appreciated before the full development of the theory.

It is named for Henry Frederick Baker, John Edward Campbell, Felix Hausdorff and Eugene Dynkin. It was first noted in print by Campbell (1897); elaborated by Henri Poincaré (1899) and Baker (1902); and systematized geometrically, and linked to the Jacoby identity by Hausdorff (1906). The explicit combinatoric formula furnished below was introduced by Eugene Dynkin (1947).

John Edward Campbell was a British mathematician, best known for his contribution to the Baker-Campbell-Hausdorff formula....

Felix Hausdorff was a German mathematician who is considered to be one of the founders of modern topology

Jules Henri Poincare , generally known as Henri Poincare, was one of France's greatest mathematicians and theoretical physicists...

The formula below was introduced by Eugene Dynkin

Eugene Borisovich Dynkin is a Russian mathematician.
The Lie Tools Package for the symbolic manipulation of Lie algebraic expressions. Please send your questions or comments to miguellt@cem.mcgill.ca

- 10 June 2002: The Campbell-Baker-Hausdorff formula in Dynkin's form (CBHD) has been successfully tested. It will be included within the LTP module files in the coming days. The cbhd procedure is found in the file cbhd1.mws within the directory "Ltp/dev" (the construction and development area of the package).

- 16 July 2002: The series expansion up to Lie brackets of order 10 for the exponent \( Z(X,Y) \) corresponding to the composition of exponentials \( \exp(X)\exp(Y)=\exp(Z) \) has been obtained using the Campbell-Baker-Hausdorff-Dynkin formula and is available as a Maple worksheet cbhd_z3_10.mws.

Dynkin's method of computing the terms of the Baker–Campbell–Hausdorff series


Asok Bose
Département de Physique, Université de Montréal, C.P. 6128, Succ "A", Montréal, Québec H3C 3J7, Canada

The infinite series for \( \log(\exp X \exp Y) \) for noncommuting \( X \) and \( Y \) is expressible in terms of iterated commutators of \( X \) and \( Y \) except for the linear term \( X+Y \). Dynkin derived an explicit expression for the terms as a sum of iterated commutators over a certain set of sequences. This paper presents a practical algorithm for applying Dynkin's formula and gives several illustrative examples. Journal of Mathematical Physics is copyrighted by The American Institute of Physics.

History: Received 8 November 1988; accepted 26 April 1989
Permalink http://link.aip.org/link/?JMAPAQ/30/2035/

On use of Campbell-Baker-Hausdorff-Dynkin formulas in nonholonomic motion planning
Duleba, I.
Inst. of Eng. Cybern., Wroclaw Univ.;

***********************************************************


***********************************************************
Calculation of the coefficients in the
Campbell-Hausdorff formula

E.B. Dynkin
Institute of Mathematics
Moscow State University

(Presented by Acad. A.N. Kolmogorov on 2/2/47)

Let us introduce the exponential and logarithmic functions by the formal series
\[ e^x = \sum_{k=0}^{\infty} \frac{1}{k!} x^k, \quad \log z = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} (z-1)^k \]
and, without assuming the commutativity of \( x \) and \( y \), calculate the series
\[ (1) \quad \Phi(x, y) = \log(e^x e^y) = \sum \frac{(-1)^{k-1}}{p_1! q_1! \cdots p_k! q_k!} \frac{1}{p_1^{p_1} q_1^{q_1} \cdots p_k^{p_k} q_k^{q_k}} \]
(the summation over all systems of non-negative integers \((p_1, q_1; \ldots; p_k, q_k)\), connected by the relations \( p_i + q_i > 0 \) \((i = 1, 2, \ldots, k)\)). Gathering together the terms of this series for which \( p_1 + q_1 + p_2 + q_2 + \cdots + p_k + q_k = m \), we represent it in the form
\[ (2) \quad \Phi(x, y) = \sum_{m=1}^{\infty} P_m(x, y), \]
where \( P_m(x, y) \) is a homogeneous polynomial of degree \( m \) in \( x \) and \( y \).

An important role in theory of Lie groups is played by the theorem of Campbell [1] and Hausdorff [2], which claims that every polynomial \( P_m(x, y) \) can be expressed in terms of \( x \) and \( y \) by means of a formula involving only operations of addition, multiplication by rational numbers, and taking commutators \((*)\). However the explicit formulas have not been known up to

---


\((*)\) The commutator of two polynomials \( P(x_1, x_2, \ldots, x_n) \) and \( Q(x_1, x_2, \ldots, x_n) \) is the expression \( P \circ Q = PQ - QP \).
now, which made it difficult to apply the theorem of Campbell and Hausdorff. In this note we give a simple expression for the series (1) in terms of commutators (formula (12)).

Let us pose a more general problem. Let \( K \) be an arbitrary field of characteristic zero and \( P(x_1, x_2, \ldots, x_n) \) an arbitrary polynomial over \( K \) in non-commuting indeterminates \( x_1, x_2, \ldots, x_n \). Our goal is to find answers to the following two questions:

1. Can \( P(x_1, x_2, \ldots, x_n) \) be expressed in terms of \( x_1, x_2, \ldots, x_n \) by means of a formula involving only operations of addition, multiplication by elements of \( K \), and taking commutators?

2. If such an expression exists then how to find it?

The set \( \mathcal{R} \) of all non-commuting polynomials in \( x_1, x_2, \ldots \) is the free associative algebra over \( K \) with generators \( x_1, x_2, \ldots \). Denote by \( \mathcal{R}^0 \) the minimal subset of \( \mathcal{R} \) satisfying the conditions: a) \( x_1, x_2, \ldots \in \mathcal{R}^0 \); b) if \( P \in \mathcal{R}^0 \) and \( Q \in \mathcal{R}^0 \) then \( \lambda P + \mu Q \in \mathcal{R}^0 \) (\( \lambda, \mu \in K \)) and \( P \circ Q \in \mathcal{R}^0 \).

Define a linear map \( P \rightarrow P^0 \) from \( \mathcal{R} \) to \( \mathcal{R}^0 \) by setting

\[
(x_{i_1} x_{i_2} \cdots x_{i_k})^0 = \frac{1}{k} x_{i_1} \circ x_{i_2} \circ \cdots \circ x_{i_k}
\]

here by \( x_{i_1} \circ x_{i_2} \circ \cdots \circ x_{i_k} \) we understand

\[
(\cdots (x_{i_1} \circ x_{i_2}) \circ x_{i_3}) \circ \cdots \circ x_{i_k}).
\]

**Theorem.** If \( P(x_1, x_2, \ldots, x_n) \in \mathcal{R}^0 \), then \( P^0 = P \).

This theorem gives answers to the both of our questions. It is sufficient to write polynomial \( P \) in the form

\[
P = \sum a_{i_1 i_2 \ldots i_k} x_{i_1} x_{i_2} \cdots x_{i_k}
\]

(indices \( i_1, i_2, \ldots, i_k \) take values \( 1, 2, \ldots, n \); the number \( k \) of indices varies arbitrarily; the sum contains only finitely many summands) and to calculate

\[
P^0(x_1, x_2, \ldots, x_n) = \sum \frac{1}{k} a_{i_1 i_2 \ldots i_k} x_{i_1} \circ x_{i_2} \circ \cdots \circ x_{i_k}.
\]

If the expressions (4) and (5) are not equal to each other then the presentation of polynomial \( P \) we are interested in is impossible whatsoever. If, however, the equality takes place, it also provides us with an explicit solution of the second question.

We sketch the proof of our theorem.

1. Every element of \( \mathcal{R}^0 \) is represented as a linear combination of expressions \( x_{i_1} \circ x_{i_2} \circ \cdots \circ x_{i_k} \). So it suffices to prove the theorem only for \( P = x_{i_1} \circ x_{i_2} \circ \cdots \circ x_{i_k} \).

2. Let \( P \) and \( Q \) be two polynomials in \( x_1, x_2, \ldots, x_n \), with

\[
P^0(x_1, x_2, \ldots, x_n) = Q(x_1, x_2, \ldots, x_n).
\]

Then for any \( n \)-tuple \( i_1, i_2, \ldots, i_n \) of natural numbers we have

\[
P^0(x_{i_1}, x_{i_2}, \ldots, x_{i_n}) = Q(x_{i_1}, x_{i_2}, \ldots, x_{i_n}).
\]

Hence the theorem will be proved if we check the equality \( P^0 = P \) for \( P = x_1 \circ x_2 \circ \cdots \circ x_n \).
3. By definition of the commutator,

$$x_1 \circ x_2 \circ \cdots \circ x_n = \sum a_{i_1 i_2 \cdots i_n} x_{i_1} x_{i_2} \cdots x_{i_n},$$

where \((i_1, i_2, \ldots, i_n)\) runs over all permutations of the numbers \(1, 2, \ldots, n\) and \(a_{i_1 i_2 \cdots i_k} \in K\). On the other hand, using the identities

$$\begin{align*}
&\{ u \circ v = -v \circ u, \ (\lambda u + \mu v) \circ w = \lambda (u \circ w) + \mu (v \circ w), \\
&\text{for all } u, v, w \in R, \lambda, \mu \in K, \\
&u \circ v \circ w + v \circ w \circ u + w \circ u \circ v = 0,
\end{align*}$$

one can prove, for any fixed \(k \leq n\), the formula

$$x_1 \circ x_2 \circ \cdots \circ x_n = \sum_{(j_2, \ldots, j_n)} c_{k j_2 \cdots j_n} x_k \circ x_{j_2} \circ \cdots \circ x_{j_n}, \quad c_{k j_2 \cdots j_n} \in K$$

(the summation is taken over all permutations \((j_2 j_3 \cdots j_n)\) of the numbers \(1, 2, \ldots, k-1, k+1, \ldots, n\)). Moreover, we have

$$x_k \circ x_{j_2} \circ \cdots \circ x_{j_n} = x_k x_{j_2} \cdots x_{j_n} + \cdots,$$

where dots denote monomials starting not from \(x_k\). Combining (6), (8) and (9) forces \(c_{k j_2 \cdots j_n} = ak_{j_2 \cdots j_n}\), and thus

$$x_1 \circ x_2 \circ \cdots \circ x_n = \sum_{(j_2, \ldots, j_n)} a_{k j_2 \cdots j_n} x_k \circ x_{j_2} \circ \cdots \circ x_{j_n} \quad (k = 1, 2, \ldots, n).$$

Adding together the equalities (10) over \(k\), we get

$$n \cdot x_1 \circ x_2 \circ \cdots \circ x_n = \sum_{(i_1 i_2 \cdots i_n)} a_{i_1 i_2 \cdots i_n} x_{i_1} \circ x_{i_2} \circ \cdots \circ x_{i_n},$$

where \((i_1, i_2, \ldots, i_n)\) runs over all permutations of the numbers \(1, 2, \ldots, n\).

Comparing (3), (6), and (11), we see that

$$(x_1 \circ x_2 \circ \cdots \circ x_n)^0 = x_1 \circ x_2 \circ \cdots \circ x_n.$$  

Remark. From our proof above we can derive more than the theorem claims. Let \(\Pi(x_1, x_2, \ldots, x_n)\) be some expression obtained from the indeterminates \(x_1, x_2, \ldots, x_n\) by means of addition, multiplication by scalars and taking commutators. Let \(P(x_1, x_2, \ldots, x_n)\) be the expression obtained if we exclude all commutators from \(\Pi\) by changing \(u \circ v\) to \(uv - vu\) everywhere. We have proved that \(\Pi\) is equivalent to \(P^0\), i.e. one of these expressions can be transformed into another using only identities (7). Thus, \(\Pi\) is equivalent to zero if and only if \(P\) is equal to zero in the algebra \(R\).

Corollary. By the theorem of Campbell-Hausdorff, homogeneous polynomials \(P_m(x, y)\) in series (2) can be expressed in terms of commutators. So, by virtue of our theorem, \(P^0(x, y) = P_m(x, y)\) and

E. B. DYNKIN 33
\[ \Phi(x, y) = \log(e^x e^y) = \Phi^0(x, y) = \]
\[ = \sum \frac{(-1)^{k-1}}{k} \frac{1}{p_1! q_1! p_2! q_2! \ldots p_k! q_k!} (x^{p_1} y^{q_1} x^{p_2} y^{q_2} \ldots x^{p_k} y^{q_k})^0. \]

Note that equality (12) can be verified by a direct calculation, and this

Note that equality (12) can be verified by a direct calculation, and this

gives a new proof of the theorem of Campbell-Hausdorff. Formula (12) allows

Note that equality (12) can be verified by a direct calculation, and this
gives a new proof of the theorem of Campbell-Hausdorff. Formula (12) allows

one to make the construction of a Lie group from a Lie algebra much more
effective and simple.

1. Classical case. Let \( K \) be the field of complex or real numbers, and \( R \) be a Lie algebra of finite rank over \( K \). We put

\[ x \ast y = \Phi^0(x, y) \]

and show that by this a local group is defined, in a neighbourhood \( U \) of zero of

and show that by this a local group is defined, in a neighbourhood \( U \) of zero of

algebra \( R \). Choose a basis \( e_1, e_2, \ldots, e_n \) in \( R \). Let \( e_i \circ e_j = \sum_{k=1}^{n} c_{ij}^k e_k \) \( (c_{ij}^k \in \text{K}) \) and \( c = \max_j \sum_{i=1}^{n} |c_{ij}^k| \). For \( x = \sum_{k=1}^{n} \lambda_k e_k \), we set \( \|x\| = c \max_j |\lambda_k| \).

One can easily see that:

One can easily see that:

a) if \( x \neq 0 \) then \( \|x\| > 0 \); b) \( \|x + y\| \leq \|x\| + \|y\| \);

c) \( \|\lambda x\| = |\lambda| \|x\| \); d) \( \|x \circ y\| \leq \|x\| \cdot \|y\| \); e) completeness: if \( \|x_n - x_m\| \rightarrow 0 \)

when both \( m \) and \( n \) go to infinity, then there exists such \( x \) that \( \|x - x_n\| \rightarrow 0 \)

when both \( m \) and \( n \) go to infinity, then there exists such \( x \) that \( \|x - x_n\| \rightarrow 0 \)

when \( n \rightarrow \infty \).

It follows from the formula (12) that the series \( \Phi^0(x, y) \) converges for all

It follows from the formula (12) that the series \( \Phi^0(x, y) \) converges for all

\( x \) and \( y \) such that the series \( \log(e^{\|x\|})e^{\|y\|} \) absolutely converges, i.e. \( \Phi^0(x, y) \)

such that the series \( \log(e^{\|x\|})e^{\|y\|} \) absolutely converges, i.e. \( \Phi^0(x, y) \)

converges provided \( \|x\| + \|y\| < \log 2 \). Thus the set of all \( x = \sum_{k=1}^{n} \lambda_k e_k \)

converges provided \( \|x\| + \|y\| < \log 2 \). Thus the set of all \( x = \sum_{k=1}^{n} \lambda_k e_k \)

such that \( |\lambda_k| < \frac{\log 2}{\lambda} \) \( (k = 1, 2, \ldots, n) \) can be taken as the neighbourhood

such that \( |\lambda_k| < \frac{\log 2}{\lambda} \) \( (k = 1, 2, \ldots, n) \) can be taken as the neighbourhood

\( U \). Furthermore, from the remark to the theorem above and formula (12) it

\( U \). Furthermore, from the remark to the theorem above and formula (12) it

follows immediately that \( x \ast (y \ast z) = (x \ast y) \ast z \), \( x \ast 0 = x \), \( x \ast (-x) = 0 \).

follows immediately that \( x \ast (y \ast z) = (x \ast y) \ast z \), \( x \ast 0 = x \), \( x \ast (-x) = 0 \).

2. Let \( K \) be a field of characteristic zero and \( R \) be a Lie algebra over \( K \).

The construction of item 1 is applicable to the algebra \( R \) provided:

The construction of item 1 is applicable to the algebra \( R \) provided:

\( A \). The field \( K \) can be endowed with a real norm \( |\lambda| \) satisfying the conditions:

\( A \). The field \( K \) can be endowed with a real norm \( |\lambda| \) satisfying the conditions:

a) if \( x \neq 0 \) then \( |\lambda| > 0 \); b) \( |\lambda + \mu| \leq |\lambda| + |\mu| \); c) \( |\lambda \mu| = |\lambda||\mu| \); d)\) the completeness condition.

b) \( |\lambda + \mu| \leq |\lambda| + |\mu| \); c) \( |\lambda \mu| = |\lambda||\mu| \); d)\) the completeness condition.

\( B \). The algebra \( R \) can be endowed with a norm \( \|x\| \) satisfying the conditions a)-e) of item 1.

\( B \). The algebra \( R \) can be endowed with a norm \( \|x\| \) satisfying the conditions a)-e) of item 1.

If algebra \( R \) is of finite rank over \( K \) then condition \( B \) follows from condition \( A \) as in item 1. R. Hooke [3] studied a special case of finite rank algebras over the field of \( p \)-adic numbers, and G. Birkhoff [4] studied algebras of infinite rank over the field of complex numbers.
Theorem 9 (Friedrichs). Let $\mathfrak{H} = \mathfrak{F}(x_1, \ldots, x_r)$ be the free algebra generated by the $x_i$ over a field of characteristic 0. Let $\delta$ be the diagonal mapping of $\mathfrak{H}$, that is, the homomorphism of $\mathfrak{H}$ into $\mathfrak{H} \otimes \mathfrak{H}$ such that $x_i \delta = x_i \otimes 1 + 1 \otimes x_i$. Then $a \in \mathfrak{H}$ is a Lie element if and only if $a \delta = a \otimes 1 + 1 \otimes a$.

Proof: We have $[a \otimes 1 + 1 \otimes a, b \otimes 1 + 1 \otimes b] = [ab] \otimes 1 + 1 \otimes [ab]$ which implies that the set of elements $a$ satisfying $a \delta = a \otimes 1 + 1 \otimes a$ is a subalgebra of $\mathfrak{H}_L$. This includes the $x_i$; hence it contains $\mathfrak{H}$. Let $y_1, y_2, \ldots$ be a basis for $\mathfrak{H}$, and $y_1, y_2, \ldots, y_m$, $m$ arbitrary, $k_i \geq 0$ ($y_1 = 1$) form a basis for $\mathfrak{H}$. Hence the product

\[(y_1^{k_1}y_2^{k_2} \cdots y_m^{k_m}) \otimes (y_1^{l_1}y_2^{l_2} \cdots y_n^{l_n})\]

form a basis for $\mathfrak{H} \otimes \mathfrak{H}$. We have

\[
(y_1^{k_1}y_2^{k_2} \cdots y_m^{k_m}) \delta = (y_1 \otimes 1 + 1 \otimes y_1)^{k_1}(y_2 \otimes 1 + 1 \otimes y_2)^{k_2} \cdots (y_m \otimes 1 + 1 \otimes y_m)^{k_m}
\]

\[
= y_1^{k_1}y_2^{k_2} \cdots y_m^{k_m} \otimes 1 + k_1 k_2 \cdots y_m^{k_m} \otimes y_1 \]

\[
+ k_1 k_2 \cdots k_m y_1^{k_1}y_2^{k_2-1} \cdots y_m^{k_m-1} \otimes y_m + * \tag{30}
\]

where $*$ represents a linear combination of base elements of the form $y_1^{l_1}y_2^{l_2} \cdots y_i^{l_i} \otimes y_1^{l_1}y_2^{l_2} \cdots y_i^{l_i}$ with $\sum l_i > 1$. The second through the $(m + 1)$-st term do not occur in the expressions of this type for any other base element $y_1^{l_1}y_2^{l_2} \cdots y_i^{l_i}$. It follows that in order that $a \delta$ shall be a linear combination of the base elements of the form $y_1^{k_1} \cdots y_m^{k_m} \otimes 1$ and $1 \otimes y_1^{l_1} \cdots y_i^{l_i}$ it is necessary that in the expression for $a$ in terms of the chosen basis only base elements $y_1^{k_1} \cdots y_m^{k_m}$ with one $k_i = 1$ and all the other $k_i = 0$ occur with non-zero coefficients. This means that $a$ is a linear combination of the $y_1$; hence $a \in \mathfrak{H}$. Hence $a \delta = a \otimes 1 + 1 \otimes a$ if and only if $a \in \mathfrak{H}$. 

5.2.1 Exponential Map Revisited

5.2.1.1 Local Diffeomorphism Let $G$ be a Lie subgroup of $GL(n, \mathbb{C})$. We already know from Theorem 4.6 that $\exp : \mathfrak{g} \rightarrow G$ is a local diffeomorphism near 0. In fact, more is true. Before beginning, use power series to define

$$\frac{I - e^{-\operatorname{ad}X}}{\operatorname{ad} X} = \sum_{n=0}^{\infty} \frac{(-1)^n}{(n+1)!} (\operatorname{ad} X)^n$$

for $X \in \mathfrak{g}$.

Theorem 5.14. (a) Let $G$ be a Lie subgroup of $GL(n, \mathbb{C})$ and $\gamma : \mathbb{R} \rightarrow \mathfrak{g}$ a smooth curve. Then

$$\frac{d}{dt} e^{\gamma(t)} = e^{\gamma(t)} \left( \frac{I - e^{-\operatorname{ad} \gamma(t)}}{\operatorname{ad} \gamma^\prime(t)} \right) (\gamma^\prime(t))$$

$$= \left[ \left( \frac{e^{\operatorname{ad} \gamma(t)} - I}{\operatorname{ad} \gamma(t)} \right) (\gamma^\prime(t)) \right] e^{\gamma(t)}.$$

(b) For $X \in \mathfrak{g}$, the map $\exp : \mathfrak{g} \rightarrow G$ is a local diffeomorphism near $X$ if and only if the eigenvalues of $\operatorname{ad} X$ on $\mathfrak{g}$ are disjoint from $2\pi i \mathbb{Z} \setminus \{0\}$.

Proof. In part (a), consider the special case of, say, $\gamma(t) = X + tY$ for $Y \in \mathfrak{g}$. Using the usual tangent space identifications at $t = 0$, the first part of (a) calculates the differential at $X$ of the map $\exp : \mathfrak{g} \rightarrow G$ evaluated on $Y$. If $(\operatorname{ad} X)Y = \lambda Y$ for $\lambda \in \mathbb{C}$, then

$$\left( \frac{I - e^{-\operatorname{ad}X}}{\operatorname{ad} X} \right) (Y) = \begin{cases} \frac{1-e^{-1}}{\lambda} Y & \text{if } \lambda \neq 0 \\ \frac{1-e^{-1}}{\lambda} Y & \text{if } \lambda = 0 \end{cases}$$
which is zero if and only if $\lambda \in 2\pi i \mathbb{Z} \setminus \{0\}$. Since Lemma 5.6 shows that $\text{ad } X$ is normal and therefore diagonalizable, part (b) follows from the Inverse Mapping Theorem and part (a).

With sufficient patience, the proof of part (a) can be accomplished by explicit power series calculations. As is common in mathematics, we instead resort to a trick. Define $\varphi : \mathbb{R}^2 \to G$ by

$$\varphi(s, t) = e^{-x\gamma(t)} \frac{\partial}{\partial t} e^{x\gamma(t)}.$$ 

To prove the first part of (a), it is necessary to show $\varphi(1, t) = \left(1 - e^{-\text{ad}\gamma(t)} \right) \gamma'(t)$. Begin by observing that $\varphi(0, t) = 0$, and so $\varphi(1, t) = \int_0^1 \frac{\partial}{\partial s} \varphi(s, t) \, ds$. However,

$$\frac{\partial}{\partial s} \varphi(s, t) = -\gamma(t) e^{-x\gamma(t)} \frac{\partial}{\partial t} e^{x\gamma(t)} + e^{-x\gamma(t)} \frac{\partial}{\partial t} \left[ -\gamma(t) e^{x\gamma(t)} \right]$$

$$= -e^{-x\gamma(t)} \gamma(t) e^{x\gamma(t)} + e^{-x\gamma(t)} \gamma'(t) e^{x\gamma(t)} + e^{-x\gamma(t)} \gamma(t) \frac{\partial}{\partial t} e^{x\gamma(t)}$$

$$= e^{-x\gamma(t)} \gamma'(t) e^{x\gamma(t)},$$

so that $\frac{\partial}{\partial s} \varphi(s, t) = \text{Ad}(e^{-x\gamma(t)} \gamma'(t)) = e^{-\text{ad}x\gamma(t)} \gamma'(t)$ by Equation 4.11. Thus

$$\varphi(1, t) = \int_0^1 e^{-\text{ad}x\gamma(t)} \gamma'(t) \, ds = \int_0^1 \sum_{n=0}^{\infty} \frac{(-s)^n}{n!} (\text{ad}\gamma(t))^n \gamma'(t) \, ds$$

$$= \left( \sum_{n=0}^{\infty} \frac{(-1)^n s^{n+1}}{(n+1)!} (\text{ad}\gamma(t))^n \gamma'(t) \right) \bigg|_{s=0}^{s=1} = \left( 1 - e^{-\text{ad}x\gamma(t)} \right) \gamma'(t),$$

as desired. To show the second part of (a), use the relation $I_{e^\alpha} = r_{e^\alpha} \circ \text{Ad}(e^\alpha) = r_{e^\alpha} \circ e^{\text{ad}x\gamma(t)}$, where $I_{e^\alpha}$ and $r_{e^\alpha}$ stand for left and right multiplication by $e^\alpha$. 

**5.2.1.2 Dynkin's Formula** Let $G$ be a Lie subgroup of $GL(n, \mathbb{C})$. For $X_j \in g$, write $[X_n, \ldots, X_3, X_2, X_1]$ for the *iterated Lie bracket*

$$[X_n, \ldots, X_3, X_2, X_1]$$

and write $[X_n^{(i_n)}, \ldots, X_1^{(i_1)}]$ for the *iterated Lie bracket*

$$[X_n^{i_n}, \ldots, X_n, \ldots, X_1^{i_1}, \ldots, X_1].$$

Although now known as the *Campbell–Baker–Hausdorff Series* ([21], [5], and [49]), the following explicit formula is actually due to Dynkin ([35]). In the proof we use the well-known fact that $\ln(X)$ inverts $e^X$ on a neighborhood of $I$, where $\ln(I + X) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} X^n$ converges absolutely on a neighborhood of $0$ (Exercise 5.15).
Theorem 5.15 (Dynkin's Formula). Let $G$ be a Lie subgroup of $GL(n, \mathbb{C})$. For $X, Y \in \mathfrak{g}$ in a sufficiently small neighborhood of 0,

$$e^X e^Y = e^Z,$$

where $Z$ is given by the formula

$$Z = \sum \frac{(-1)^{n+1}}{n} \frac{1}{(i_1 + j_1) + \cdots + (i_n + j_n)} \frac{[X^{(i_1)}, Y^{(j_1)}, \ldots, X^{(i_n)}, Y^{(j_n)}]}{i_1! j_1! \cdots i_n! j_n!},$$

where the sum is taken over all 2n-tuples $(i_1, \ldots, i_n, j_1, \ldots, j_n) \in \mathbb{N}^{2n}$ satisfying $i_k + j_k \geq 1$ for positive $n \in \mathbb{N}$.

Proof. The approach of this proof follows [34]. Using Theorem 4.6, choose a neighborhood $U_0$ of 0 in $\mathfrak{g}$ on which exp is a local diffeomorphism and where ln is well defined on exp $U$. Let $U \subseteq U_0$ be an open ball about of 0 in $\mathfrak{g}$, so that $(\exp U)^2 (\exp U)^{-2} \subseteq \exp U_0$ (by continuity of the group structure as in Exercise 1.4). For $X, Y \in U$, define $\gamma(t) = e^{Xt} e^{Yt}$ mapping a neighborhood of $[0, 1]$ to exp $U$. Therefore there is a unique smooth curve $Z(t) \in U_0$, so that $e^{Z(t)} = e^{Xt} e^{Yt}$. Apply $\frac{d}{dt}$ to this equation and use Theorem 5.14 to see that

$$\left[ \left( \frac{\exp Z(t)}{\exp Z(t) - I} \right) (Z'(t)) \right] e^{Z(t)} = X e^{Z(t)} + e^{Z(t)} Y.$$

Since $Z(t) \in U_0$, exp is a local diffeomorphism near $Z(t)$. Thus the proof of Theorem 5.14 shows that $\left( \frac{\exp Z(t)}{\exp Z(t) - I} \right)$ is an invertible map on $\mathfrak{g}$. As $e^{Z(t)} = e^{Xt} e^{Yt}$, $\text{Ad}(e^{Z(t)}) = \text{Ad}(e^{Xt}) \text{Ad}(e^{Yt})$, so that $e^{\text{ad} Z(t)} = e^{\text{ad} X} e^{\text{ad} Y}$ by Equation 4.11. Thus

$$Z'(t) = \left( \frac{\text{ad} Z(t)}{\text{ad} Z(t) - I} \right) \left( X + \text{Ad}(e^{Z(t)}) Y \right) = \left( \frac{\text{ad} Z(t)}{\text{ad} Z(t) - I} \right) \left( X + e^{\text{ad} X} e^{\text{ad} Y} \right).$$

Using the relation $A = \ln(I + (e^A - I)) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} (e^A - I)^n$ for $A = \text{ad} Z(t)$ and $e^A = e^{t \text{ad} X} e^{t \text{ad} Y}$, we get

$$\frac{\text{ad} Z(t)}{\text{ad} Z(t) - I} = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} \left( e^{t \text{ad} X} e^{t \text{ad} Y} - I \right)^{n-1}.$$

Hence

$$Z'(t) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} \left( e^{t \text{ad} X} e^{t \text{ad} Y} - I \right)^{n-1} \left( X + e^{t \text{ad} X} e^{t \text{ad} Y} \right)$$

$$= \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} \left[ \sum_{i,j=0}^{\infty} \frac{t^{i+j}}{i! j!} (\text{ad} X)^i (\text{ad} Y)^j \right] \left( X + \sum_{i=0}^{\infty} \frac{t^i}{i!} (\text{ad} X)^i \right)^n.$$
\[ \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} \left[ \frac{i_1! + j_1! + \ldots + i_{n-1}! + j_{n-1}!}{i_1! j_1! \cdots i_{n-1}! j_{n-1}!} [X^{i_1}, (Y)^{j_1}, \ldots, (X)^{i_{n-1}}, (Y)^{j_{n-1}}, X] \right] \\
+ \sum \frac{i_1! + j_1! + \ldots + i_n! + j_n!}{i_1! j_1! \cdots i_{n-1}! j_{n-1}! i_n! j_n!} [X^{i_1}, (Y)^{j_1}, \ldots, (X)^{i_{n-1}}, (Y)^{j_{n-1}}, (X)^{i_n}, Y] \]

where the second and third sum are taken over all \( i_k, j_k \in \mathbb{N} \) with \( i_k + j_k \geq 1 \). Since \( Z(0) = 0 \), \( Z(1) = \int_0^1 \frac{d}{dt} Z(t) \, dt \). Integrating the above displayed equation finishes the proof. 