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

$$e^x = \sum_{k=0}^{\infty} \frac{1}{k!} x^k, \qquad \log z = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} (z-1)^k$$

and, without assuming the commutativity of x and y, calculate the series

$$(1) \qquad \Phi(x,y) = \log(e^x e^y) = \sum \frac{(-1)^{k-1}}{k} \frac{1}{p_1! q_1! \dots p_k! q_k!} x^{p_1} y^{q_1} \dots x^{p_k} y^{q_k}$$

(the summation over all systems of non-negative integers $(p_1, q_1; \dots; p_k, q_k)$, connected by the relations $p_i + q_i > 0$ $(i = 1, 2, \dots, k)$). Gathering together the terms of this series for which $p_1 + q_1 + p_2 + q_2 + \dots + p_k + q_k = m$, we represent it in the form

$$(2) \qquad \Phi(x,y) = \sum_{m=1}^{\infty} P_m(x,y),$$

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

Originally published in Doklady Akad. Nauk SSSR (N.S.), 57(1947), 323-326. English translation by A. Kleshchev (1997).

(*) The commutator of two polynomials $P(x_1, x_2, \dots, x_n)$ and $Q(x_1, x_2, \dots, x_n)$ 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 $P(x_1, x_2, \dots, x_n)$ an arbitrary polynomial over K in non-commuting indeterminates $x_1, x_2, \dots, x_n$. Our goal is to find answers to the folowing two questions:

1. Can $P(x_1, x_2, \dots, x_n)$ be expressed in terms of $x_1, x_2, \dots, x_n$ 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 $x_1, x_2, \dots$ is the free associative algebra over K with generators $x_1, x_2, \dots$. Denote by $\mathcal{R}^0$ the minimal subset of $\mathcal{R}$ satisfying the conditions: a) $x_1, x_2, \dots \in \mathcal{R}^0$; 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 $P \to P^0$ from $\mathcal{R}$ to $\mathcal{R}^0$ by setting

$$(3) \qquad (x_{i_1} x_{i_2} \dots x_{i_k})^0 = \frac{1}{k} x_{i_1} \circ x_{i_2} \circ \dots \circ x_{i_k};$$

here by $x_{i_1} \circ x_{i_2} \circ \dots \circ x_{i_k}$ we understand

$$(\dots((x_{i_1} \circ x_{i_2}) \circ x_{i_3}) \circ \dots \circ x_{i_k}).$$

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

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

$$(4) \qquad P = \sum a_{i_1 i_2 \dots i_k} x_{i_1} x_{i_2} \dots x_{i_k}$$

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

$$(5) \qquad P^0(x_1, x_2, \dots, x_n) = \sum \frac{1}{k} a_{i_1 i_2 \dots i_k} x_{i_1} \circ x_{i_2} \circ \dots \circ x_{i_k}.$$

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 $x_{i_1} \circ x_{i_2} \circ \dots \circ x_{i_k}$. So it suffices to prove the theorem only for $P = x_{i_1} \circ x_{i_2} \circ \dots \circ x_{i_k}$.

2. Let P and Q be two polynomials in $x_1, x_2, \dots, x_n$, with

$$P^0(x_1, x_2, \dots, x_n) = Q(x_1, x_2, \dots, x_n).$$

Then for any n-tuple $i_1, i_2, \dots, i_n$ of natural numbers we have

$$P^0(x_{i_1}, x_{i_2}, \dots, x_{i_n}) = Q(x_{i_1}, x_{i_2}, \dots, x_{i_n}).$$

Hence the theorem will be proved if we check the equality $P^0 = P$ for $P = x_1 \circ x_2 \circ \dots \circ x_n$.

3. By definition of the commutator,

$$x_1 \circ x_2 \circ \cdots \circ x_n = \sum a_{i_1 i_2 \ldots i_n} x_{i_1} x_{i_2} \ldots x_{i_n}, \tag{6}$$

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

$$\begin{cases} u \circ v = -v \circ u, \ (\lambda u + \mu v) \circ w = \lambda(u \circ w) + \mu(v \circ w), \\ u \circ v \circ w + v \circ w \circ u + w \circ u \circ v = 0, \\ u, v, w \in \mathcal{R}, \ \lambda, \mu \in K, \end{cases} \tag{7}$$

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

$$x_1 \circ x_2 \circ \cdots \circ x_n = \sum_{(j_2, \ldots, j_n)} c_{kj_2 \ldots j_n} x_k \circ x_{j_2} \circ \cdots \circ x_{j_n}, \quad c_{kj_2 \ldots k_n} \in K \tag{8}$$

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

$$x_k \circ x_{j_2} \circ \cdots \circ x_{j_n} = x_k x_{j_2} \ldots x_{j_n} + \ldots, \tag{9}$$

where dots denote monomials starting not from x_k. Combining (6), (8) and (9) forces $c_{kj_2 \ldots j_n} = a_{kj_2 \ldots j_n}$, and thus

(10)
$$x_1 \circ x_2 \circ \cdots \circ x_n = \sum_{(j_2 \ldots j_n)} a_{kj_2 \ldots j_n} x_k \circ x_{j_2} \circ \cdots \circ x_{j_n} \quad (k = 1, 2, \ldots, n).$$

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

$$n \cdot x_1 \circ x_2 \circ \cdots \circ x_n = \sum_{(i_1 i_2 \ldots i_n)} a_{i_1 i_2 \ldots i_n} x_{i_1} \circ x_{i_2} \circ \cdots \circ x_{i_n}, \tag{11}$$

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 Π by changing $u \circ v$ to $uv - vu$ everywhere. We have proved that Π is equivalent to P^0, i.e. one of these expressions can be transformed into another using only identities (7). Thus, Π is equivalent to zero if and only if P is equal to zero in the algebra $\mathcal{R}$.

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

$$\Phi(x,y) = \log(e^x e^y) = \Phi^0(x,y) =$$
$$= \sum \frac{(-1)^{k-1}}{k} \frac{1}{p_1!q_1!p_2!q_2!\dots p_k!q_k!} (x^{p_1}y^{q_1}x^{p_2}y^{q_2}\dots x^{p_k}y^{q_k})^0. \tag{12}$$

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

$$x * y = \Phi^0(x,y) \tag{13}$$

and show that by this a local group is defined, in a neighbourhood U of zero of algebra R. Choose a basis $e_1, e_2, \dots, e_n$ in R. Let $e_i \circ e_j = \sum_{k=1}^n c_{ij}^k e_k$ ($c_{ij}^k \in K$) and $c = \max_k \sum_{i,j=1}^n |c_{ij}^k|$. For $x = \sum_{k=1}^n \lambda_k e_k$, we set $\|x\| = c \max_k |\lambda_k|$.

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 $\|x_n - x_m\| \to 0$ when both m and n go to infinity, then there exists such x that $\|x - x_n\| \to 0$ when $n \to \infty$.

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. $\Phi^0(x,y)$ converges provided $\|x\| + \|y\| < \log 2$. Thus the set of all $x = \sum_{k=1}^n \lambda_k e_k$ such that $|\lambda_k| < \frac{\log 2}{2c}$ $(k = 1, 2, \dots, n)$ can be taken as the neighbourhood U. Furthermore, from the remark to the theorem above and formula (12) it follows immediately that $x * (y * z) = (x * y) * z$, $x * 0 = x$, $x * (-x) = 0$.

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:

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. 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.

Submitted 2/2/47

References

[1] J.E. Campbell. *Introductory Treatise on Lie Theory of Finite Continuous Transformation Groups*, Oxford, 1903.

[2] F. Hausdorff, Ber. Sachs. Ges., 58, 19 (1906).

[3] R. Hooke, Ann. Math., 43, 641 (1942).

[4] G. Birkhoff, Trans. Am. Math. Soc., 43, 61 (1938).