ADVANCES IN MATHEMATICS **88**, 200–257 (1991)

Recursion Formulas for the Lie Integral

R. V. CHACON

Department of Mathematics, University of British Columbia, Vancouver, British Columbia, Canada V6T 1W5

AND

A. T. FOMENKO

Department of Topology and Geometry, Moscow State University, 119899 Moscow, USSR

An explicit definition is given of a non-linear integral which is termed the *Lie integral.* This integral is defined for Lie algebra valued functions in terms of the representations of the algebra. Various properties of the integral are investigated, and applications are given to differential equations. Applications to differential geometry will be given elsewhere.

Contents.

1. INTRODUCTION

For certain finite dimensional algebras $\mathscr{A}$ it is clear that there exists a multiple-valued function $A(X, Y) \in \mathscr{A}$ such that

$$e^X e^Y = e^{A(X, Y)}$$

for all $X, Y \in \mathscr{A}$. The Baker (1905)–Hausdorff (1906)–Campbell (1908) (BACH) formula [15] gives the power series expansion of the branch of

0001-8708/91 $7.50

$A(X, Y)$ about $(0, 0)$ with constant term equal to zero. It can be shown that, with BACH $(X, Y) = B(X, Y)$,

$$B(X, Y) = (X + Y) + \tfrac{1}{2}[X, Y] + \cdots,$$

and it is an interesting problem to find efficient recursion formulas for the coefficients for this expansion [11]. Its terms are homogeneous polynomials of increasing degree in Lie products of X and Y. It can also be shown that $B(X, Y)$ converges in a neighborhood of $(0, 0)$ for all $\mathscr{A}$, and that $B(X, Y)$ fails to converge for all $X, Y \in \mathscr{A}$ in general. For example, in the case that $\mathscr{A} = \mathrm{so}(3)$, or more generally for finite dimensional compact matrix groups, one can define an associative product in the algebra to be given by

$$X \cdot Y = B(X, Y)$$

under the assumption that $B(X, Y)$ converges for all $X, Y \in \mathscr{A}$. This product turns $\mathscr{A}$ into a non-compact Lie group whose Lie algebra is also $\mathscr{A}$, yielding a contradiction, since the universal covering group of the group of $\mathscr{A}$ is compact [2].

In this paper we will be concerned with a generalization of the BACH formula. Consider the differential equation

$$\frac{dX}{ds} = K(s) X(s), \qquad a \leqslant s \leqslant b, \tag{1}$$

where K and X are $\mathscr{A}$-valued. Although the results hold more generally if $\mathscr{A}$ is finite dimensional, we shall assume that $\mathscr{A}$ is a finite dimensional matrix algebra. The solution of (1) can be written in the form

$$X(s) = e^{H(s)} X(a). \tag{2}$$

Several authors have considered the problem of finding formulas for the solution of (1) (see e.g. [1, 3–5, 9, 10, 12, 13, 16, 17, 18, 20]).

R. Feynman in [9] found a formal expression for the solution without proving its convergence. Later, Magnus in [13] obtained a well-known differential equation, called the Magnus equation, for the function $H(s)$. The problem of calculating $H(s)$ was also considered by Bialynicki-Birula, Mielnik, and Plebanski in [1]. These authors note that they do not prove convergence of the expressions which they obtain. In 1976 Karasev and Mosolova in [12] obtained an expression for $H(s)$ together with an estimate for its norm, basing their work on an important method developed by Maslov in [14].

K. T. Chen in [3–5] has obtained interesting results concerning the function H. His work was employed in [16] to obtain a formula, which, when specialized to finite dimensional algebras, converges everywhere. As already noted, the BACH formula cannot converge everywhere in general even in the matrix case. We have concentrated on the question of convergence of the series in Lie products of K of increasing degree for the function $H(s)$.

In order to expand $H(s)$ in Lie products of K of increasing degree, we consider the system of equations

$$\frac{dX}{ds} = tK(s)\, X(s), \qquad a \leqslant s \leqslant b,\ 0 \leqslant t \leqslant 1, \tag{1$'$}$$

where we regard t as a parameter. The system then has a solution of the form

$$X(t, s) = e^{H(t,s)} X(a).$$

It is clear that $H(1, s) = H(s)$ and it is shown that

$$H(t, s) = \sum_{0}^{\infty} H_n(s)\, t^n,$$

converges, where the H_n are the homogeneous terms. It is also shown that $H(t, s)$ satisfies the new differential equation

$$\frac{\partial H(t, s)}{\partial t} = \frac{\operatorname{ad}_{H(t,s)}}{1 - e^{-\operatorname{ad}_{H(t,s)}}} \int_a^s e^{-H(t,u)} K(u)\, e^{H(t,u)}\, du.$$

The discrete version of this equation is used to obtain norm estimates and to prove convergence of the series for $H(s)$. It may be contrasted with Magnus' equation

$$\frac{dH(s)}{ds} = \frac{\operatorname{ad}_{H(s)}}{e^{-\operatorname{ad}_{H(s)}} - 1} K(s)$$

from which norm estimates and convergence do not appear to follow easily.

An explicit formula is found for $H(s)$ and it is used to define a new integral, the Lie integral

$$H(s) = L \int_a^s K(u)\, du. \tag{3}$$

The definition of the Lie integral (3) is based on a formula for N

arguments. This formula gives an expression for the matrix $B_N(X_1, X_2, ..., X_N)$ such that

$$e^{X_1} \cdot e^{X_2} \cdots e^{X_N} = e^{B_N(X_1, ..., X_N)} \tag{4}$$

for matrices $X_1, X_2, ..., X_N$. The Lie integral of an arbitrary function F defined on an interval $[a, b]$ is defined in two steps. The integral is first defined for step functions, and then it is shown that the integrals of approximating step functions converge in the case that $F(t)$ is a Riemann integrable function. The Lie integral (3) is only defined for $b-a$ sufficiently small, and the allowed size of the interval of definition depends on the norm of F—the *smaller* F, the *larger* the interval.

Suppose first that $F(s) = \sum_{k=0}^{n-1} F_k \psi_k(s)$, where ψ_k is the characteristic function of the interval $[a_k, a_{k+1})$, $a = a_0 < a_1 < \cdots < a_n = b$. Then the Lie integral (3) is defined to be

$$B_n(F_{n-1} \Delta_{n-1}, ..., F_0 \Delta_0), \tag{5}$$

where $\Delta_k = (a_{k+1} - a_k)$, $k = 0, ..., n-1$. If F is Riemann integrable, then the Lie integral (3) is defined as the limit of expressions of the form (6). The main part of the paper is devoted to calculating explicit formulas for the evaluation of this integral.

We note that two of Lie's theorems, that

$$\lim_{n \to \infty} (e^{X_2/2n} e^{X_1/2n})^n = e^{(X_1 + X_2)/2} \tag{6}$$

and that

$$\lim_{n \to \infty} (e^{X_1/n} e^{X_2/n} e^{-X_1/n} e^{-X_2/n})^{n^2} = e^{[X_1, X_2]}, \tag{7}$$

may be regarded as weak convergence results of the Lie integral. For (7) define $F_n(s)$, $s \in [0, 1]$, by dividing $[0, 1]$ into $2n$ consecutive subintervals of equal length, and then setting $F_n(s)$ equal to X_1 and X_2 alternately on each subinterval, starting with the first one. Real valued functions f_n corresponding to the functions F_n would converge weakly to the constant equal to the average of the two values and thus (7) may be viewed as a special case of a more general result, that if F_n converges to F in a suitable sense, then $L \int_0^1 F_n(s)\, ds \to L \int_0^1 F(s)\, ds$.

For (7), define $G_n(s)$, $s \in [0, 1]$, by first dividing the interval $[0, 1]$ into n^2 consecutive subintervals of equal length, I_k, $k = 1, ..., n^2$, then dividing each of these intervals I_k into four consecutive intervals of equal length, $I_k = I_{k,1} + I_{k,2} + I_{k,3} + I_{k,4}$. Finally, set

$$G_n(s) = \begin{cases} 4nX_1, & s \in I_{k,1}, \\ 4nX_2, & s \in I_{k,2}, \\ -4nX_2, & s \in I_{k,3}, \\ -4nX_1, & s \in I_{k,4}. \end{cases} \tag{8}$$

Real valued functions $g_n(s)$ corresponding to the functions $G_n(s)$ have the property that $\int_0^1 g_n(s)\, ds \to 0$. The problem of convergence is more delicate for Lie integrals, since

$$L \int_0^1 G_n(s)\, ds \to [X_1, X_2]. \tag{9}$$

The Lie integral may be regarded as the logarithm of the product integral discussed by Browder, Dollard, Friedman, and Masani in [8]. The method of this paper may be used to obtain similar results in the case that the parameter t is replaced by a function $\varphi(t, s)$ in the system of equations (1)′.

2. Notation

1. The formal series $B_N(X_1, ..., X_N) = \sum_{n=1}^{\infty} c_n^N(X_1, ..., X_N)$, where $c_n^N(X_1, ..., X_N)$, is the homogeneous polynomial of degree n in $X_1, ..., X_N$ such that the formal identity holds:

$$e^{X_1} \cdots e^{X_N} = e^{B_N(X_1, ..., X_N)}.$$

2. The function $\psi_N(t)$, which is analytic in a neighborhood of 0, is defined as

$$\psi_N(t) = \sum_{n=1}^{\infty} c_n^N t^n \text{ and } e^{tX_1} \cdots e^{tX_N} = e^{\psi_N(t)}.$$

3. The function $T_N(t)$, which is analytic everywhere, is defined as

$$T_N(t) = e^{-tX_N} \cdots e^{-tX_2} X_1 e^{tX_2} \cdots e^{tX_N} + \cdots + e^{-tX_N} X_{N-1} e^{tX_N} + X_N,$$

and for $2 \leqslant i \leqslant N$,

$$t_{N,i}(t, a) = e^{-tX_N} \cdots e^{-tX_i} a e^{tX_i} \cdots e^{tX_N},$$

so that

$$T_N(t) = \sum_{i=2}^{N} t_{N,i}(t, X_{i-1}) + X_N.$$

4. The Lie integral or curve functional is defined by

$$L\int_a^b K(s)\,ds = \sum_{n=1}^{\infty} H_n(K(s), a \leqslant s \leqslant b), \qquad \text{and}$$

$$H[t] = L\int_a^t K(s)\,ds,$$

$$H(t) = L\int_{s=a}^{b} tK(s)\,ds.$$

5. THEOREM 1. *The main recursion formulas*

$$L\int_a^b K(s)\,ds = H_1(a, b) + H_2(a, b) + \cdots,$$

where

$$H_1(a, b) = \int_a^b K(s)\,ds,$$

and for $n \geqslant 1$, *with* $T_0 = H_1$,

$$(n+1)\,H_{n+1} = T_n + \sum_{r=1}^{n} \left\{\frac{1}{2}[H_r, T_{n-r}] + \sum_{\substack{p \geqslant 1 \\ 2p \leqslant r}} k_{2p} \sum_{\substack{m_i > 0 \\ m_1 + \cdots m_{2p} = r}} [H_{m_1}, [\ldots, [H_{m_{2p}}, T_{n-r}]\cdots]\right\},$$

where $k_{2p} \cdot (2p)!$ *are Bernoulli's numbers and where for* $k \geqslant 1$,

$$k!\,T_k = T_{\infty}^{(k)}(0) = \frac{d^k}{dt^k} T_{\infty}(0) = \lim_{N \to \infty} \frac{d^k}{dt^k} T_N(t)\bigg|_{t=0}.$$

6. Formulas for calculating $\{T_k\}$ are

$$T_k[t] = T_k = \int_{u_1=a}^{t} \left[\cdots\left[K(u_1), \int_{u_2=a}^{u_1} K(u_2)\right], \ldots, \int_{u_k=a}^{u_{k-1}} K(u_k)\right], \int_{u_{k+1}=a}^{u_k} K(u_{k+1})\right] du_{k+1}\,du_k \cdots du_1,$$

or equivalently

$$7.\quad T_k[t]=T_k=\int\cdots\int_{t\geqslant u_1\geqslant\cdots\geqslant u_{k+1}\geqslant a}[\cdots[K(u_1),K(u_2)],\ldots$$
$$\ldots,K(u_{k-1})],K(u_k)],K(u_{k+1})]\,du_{k+1}\cdots du_1.$$

3. A Formula for the Logarithm Functional of $e^{X_1}\cdots e^{X_N}$.

The Differential Equation of ψ_N

Let G be a Lie group, L its algebra, and let L have a norm, denoted by $|\cdot|$. If $X_1,\ldots,X_N\in L$ and are sufficiently close to $0\in L$, then there is a unique analytic function $b_N(X_1,\ldots,X_N)\in L$ with $b_N(0,\ldots,0)=0$ such that

$$e^{X_1}\cdots e^{X_N}=e^{b_N(X_1,\ldots,X_N)}.\tag{10}$$

Definition 1. The function $b_N\colon \underbrace{L\times\cdots\times L}_{N}\to L$ is called the logarithm functional.

Consider the formal expansion

$$B_N(X_1,\ldots,X_N)=\sum_{n=0}^{\infty}c_n^N(X_1,\ldots,X_N).\tag{11}$$

The coefficients $c_n^N(X_1,\ldots,X_N)$ are homogeneous polynomials of degree n in the variables $X_1,\ldots,X_N$, and are formally determined by (11) when the infinite series of (12) is substituted in place of b_N in Eq. (11). It is more convenient to put $\psi_N(t)=b_N(tX_1,\ldots,tX_N)$ and then define $c_n^N(X_1,\ldots,X_N)=(1/n!)(d^n/dt^n)(\psi_N(t))|_{t=0}$ and that is the approach we shall take. We then prove in Theorem 4 that the coefficients $c_n^N(X_1,\ldots,X_N)$ are homogeneous of degree n. We use the symbol B_N for the expansion and b_N for the function as it is not at all clear that B_N converges, or whether if it does converge it converges to b_N. The case $N=2$ is treated in [22]. It is shown that B_2 converges to b_2 sufficiently close to 0×0, $0\in L$. Recursion formulas for c_n^2 are given from which it follows, for example, that

$$\begin{aligned}c_1^2(X_1,X_2)&=X_1+X_2,\\ c_2^2(X_1,X_2)&=\tfrac{1}{2}[X_1,X_2],\\ c_3^2(X_1,X_2)&=\tfrac{1}{12}[(X_1-X_2),[X_1,X_2]].\end{aligned}\tag{12}$$

We outline first a simple argument which shows that $B_N(X_1,\ldots,X_N)$ does in fact converge to $b_N(X_1,\ldots,X_N)$ for $X_1,\ldots,X_N$ sufficiently close to 0, and then turn to the principal task of this section. This principal task is to find the recursion formulas for $c_n^N(X_1, ..., X_N)$. These formulas are useful in computing $B_N(X_1, ..., X_N)$ and in determining its properties.

Recall that the exponential mapping is analytic and its inverse is locally analytic and thus there is a neighborhood $U \subset L$ of 0 such that the mapping $b_N: U \times \cdots \times U$ (N times) $\to L$ is analytic for $X_1, ..., X_N \in U$ and $b_N(0, ..., 0) = 0$. It follows from these remarks that for each $\delta > 0$ there is a $V \subset U$ such that the function $\phi(u_1, ..., u_N) = b_N(u_1 X_1, ..., u_N X_N)$ is an analytic function of R^N if $|u_1 + \cdots + u_N| \leqslant (1+\delta)\sqrt{N}$ and $X_1, ..., X_N \in V$. The function $\psi(t) = \psi_N(t, X_1, ..., X_N) = \phi(t, ..., t)$ is analytic for $|t| \leqslant 1 + \delta$, and

$$c_n^N(X_1, ..., X_N) = \frac{1}{n!}\frac{d^n}{dt}(\psi_N(t))|_{t=0}. \tag{13}$$

Thus

$$\psi(t) = \sum_{n=0}^{\infty} c_n^N t^n \tag{14}$$

and the expansion converges absolutely for $|t| \leqslant 1 + (\delta/2)$. Since $b_N(X_1, ..., X_N) = \phi_N(1, X_1, ..., X_N)$, the problem of the convergence of B_N to b_N is reduced to the problem of showing that the expansion of $\psi(t)$ converges for $t = 1$, and that expansion has been shown to converge. Thus $B_N = b_N$ for $X_1, ..., X_N$ sufficiently small.

DEFINITION 2. The linear operator $\mathrm{ad}_X: L \to L$ is defined by

$$\mathrm{ad}_X(Y) = [X, Y]. \tag{15}$$

The representation $X \to \mathrm{ad}_X$ is called the *adjoint representation.*

If $X \in L$ is an arbitrary element in L, then as exp: $L \to G$, the differential

$$(d\exp)_X: T_X L \to T_g G, \qquad g = e^X.$$

The tangent space $T_X L$ can be naturally identified with the Lie algebra L and we may therefore consider $(d\exp)_X$ as a mapping from L to $T_g G$, i.e. $(d\exp)_X: L \to T_g G$.

It is shown in [15] (and as matrix multiplication corresponds to composition, see also Theorem 5.1 of [8]) that

$$e^{-X} \circ (d\exp)_X = m(\mathrm{ad}_X), \tag{16}$$

where

$$m(\mathrm{ad}_X)=\sum_{n=0}^{\infty}\frac{(-1)^n}{(n+1)!}(\mathrm{ad}_X)^n. \tag{17}$$

It is also shown in [15] that

$$(m(\mathrm{ad}_X))^{-1}=k(\mathrm{ad}_X)+\tfrac{1}{2}\,\mathrm{ad}_X, \tag{18}$$

where

$$k(\mathrm{ad}_X)=1+\sum_{p=1}^{\infty}k_{2p}(\mathrm{ad}_X)^{2p} \tag{19}$$

and where $k_{2p}\cdot(2p)!=b_{2p}$ are the well-known Bernoulli numbers.

We now return to the function $\phi(u_1, ..., u_n)$ and define $\phi_{N,i}=\phi(u_i, u_{i+1}, ..., u_N)$, so that

$$e^{u_iX_i}\cdots e^{u_NX_N}=e^{\phi_{N,i}(u_i,u_{i+1},...,u_N)}.$$

These functions are analytic in a neighborhood if $0\in R^{N-i+1}$, and one may therefore take their partial derivatives. We prove the following lemma.

LEMMA 1. *The following equations are valid in the domain of analyticity of the functions for* $1\leqslant i\leqslant N$:

$$\begin{aligned}\frac{\partial\phi_{N,1}}{\partial u_i}&=k(\mathrm{ad}_{\phi_{N,1}})(e^{-\phi_{N,i+1}}X_ie^{\phi_{N,i+1}})\\&\quad+\frac{1}{2}[\phi_{N,1}, e^{-\phi_{N,i+1}}X_ie^{\phi_{N,i+1}}].\end{aligned} \tag{20}$$

Remark. Since $\phi_{N,1}$, $\phi_{N,i+1}\in L$, all expressions are well defined as $aXa^{-1}\in L$ if $a\in G$ and $X\in L$, where aXa^{-1} is defined in terms of the product of the corresponding matrices.

Proof of Lemma 1. Write $\phi_{N,1}=\phi$. We have

$$e^{u_1X_1}\cdots e^{u_NX_N}=e^{\phi(u_1,...,u_N)}\in G,$$

so that

$$\frac{\partial}{\partial u_i}(e^{u_1X_1}\cdots e^{u_NX_N})=\frac{\partial}{\partial u_i}e^{\phi(u_1,...,u_N)}\in T_{(e^\phi)}G. \tag{21}$$

The right side of (21) is equal to

$$(d\exp_\phi)\frac{\partial}{\partial u_i}\phi(u_1, ..., u_N),$$

where $\phi \in L$. The left side of (21) is equal to

$$e^{u_1X_1}\cdot\ldots\cdot e^{u_{i-1}X_{i-1}}\frac{\partial}{\partial u_i}(e^{u_iX_i})\,e^{u_{i+1}X_{i+1}}\cdots e^{u_NX_N}$$
$$= e^{u_1X_1}\cdots e^{u_{i-1}X_{i-1}}\cdot X_i e^{u_iX_i} e^{u_{i+1}X_{i+1}}\cdots e^{u_NX_N}.$$

Note that $X_ie^{u_iX_i} = e^{u_iX_i}\cdot X_i$, because the matrices X_i and u_iX_i commute as the parameter u_i is a scalar. Substituting these expressions into (21), we obtain

$$e^{u_1X_1}\cdots e^{u_{i-1}X_{i-1}}e^{u_iX_i}\cdot X_i\cdot e^{u_{i+1}X_{i+1}}\cdots e^{u_NX_N}$$
$$= (d\exp_\phi)\frac{\partial}{\partial u_i}\phi(u_1, ..., u_N). \tag{22}$$

Multiplying (22) by $e^{-\phi}\in G$ on the left, where $e^{-\phi} = (e^{u_1X_1}\cdots e^{u_NX_N})^{-1} = e^{-u_NX_N}\cdots e^{-u_1X_1}$, we obtain

$$e^{-u_NX_N}\cdots e^{-u_{i+1}X_{i+1}}\cdot X_i\cdot e^{u_{i+1}X_{i+1}}\cdots e^{u_NX_N}$$
$$= (e^{-\phi}(d\exp_\phi))\frac{\partial}{\partial u_i}\phi(u_1, ..., u_N). \tag{23}$$

From (19) we have $(e^{-\phi}(d\exp_\phi))^{-1} = k(\mathrm{ad}_\phi) + \frac{1}{2}\mathrm{ad}_\phi$, so that

$$\left(k(\mathrm{ad}_\phi) + \frac{1}{2}\mathrm{ad}_\phi\right)(e^{-\phi_{N,i+1}}X_ie^{\phi_{N,i+1}}) = \frac{\partial}{\partial u_i}\phi(u_1, ..., u_N).$$

Since $\phi(u_1, ..., u_N) = \phi_{N,1}(u_1, ..., u_N)$, we obtain (20) and the proof is finished.

LEMMA 2. *Define* $\psi(t, X_1, ..., X_N) = \psi_N(t, X_1, ..., X_N) = \phi_{N,1}(t, ..., t)$ (*i.e.*, $u_1 = \cdots = u_N = t$).

Then

$$\frac{d\psi_N(t)}{dt} = \left\{k(\mathrm{ad}_{\phi_{N,1}})\left(\sum_{i=1}^{N} e^{-\phi_{N,i+1}}\cdot X_i\cdot e^{\phi_{N,i+1}}\right)\right\}$$
$$+\frac{1}{2}\left[\phi_{N,1}, \sum_{i=1}^{N} e^{-\phi_{N,i+1}}\cdot X_i\cdot e^{\phi_{N,i+1}}\right], \tag{24}$$

where $u_1 = \cdots = u_N = t$.

Proof of Lemma 2. It follows from the definition of $\psi_N(t)$ that

$$\frac{d\psi(t)}{dt} = \sum_{i=1}^{N} \frac{\partial \phi}{\partial u_i}\bigg|_{u_1 = \cdots = u_N = t}.$$

Formula (20) of Lemma 1 then yields (24).

Lemma 2 can be formulated in a slightly different way. Consider the new L-valued function

$$\begin{aligned} T_N(t) &= T_N(t, X_1, ..., X_N) \\ &= \sum_{i=1}^{N} e^{-tX_N} \cdots e^{-tX_{i+1}} \cdot X_i \cdot e^{tX_{i+1}} \cdots e^{tX_N} \\ &= e^{-tX_N} \cdots e^{-tX_2} X_1 e^{tX_2} \cdots e^{tX_N} \\ &\quad + \cdots + e^{-tX_N} X_{N-1} e^{tX_N} + X_N, \end{aligned} \tag{25}$$

where we formally set $X_{N+1} = 0$. The function $T_N(t)$ is an analytic function of t and $X_1, ..., X_N$.

LEMMA 2′. *Let $N \geqslant 2$ and $T_N(t, X_1, ..., X_N)$, $\psi_N(t, X_1, ..., X_N)$ be the functions described above. Then the function ψ_N satisfies the following differential equation in a neighborhood of the origin*:

$$\frac{d\psi_N}{dt} = k(\mathrm{ad}_{\psi_N})(T_N) + \frac{1}{2}[\psi_N, T_N]. \tag{26}$$

Proof immediately follows from (24) of Lemma 2, since $\phi_{N,1}(t, ..., t) \equiv \psi_N(t, X_1, ..., X_N)$.

4. A RECURSION FORMULA FOR THE COEFFICIENTS $c_n^N(X_1, ..., X_N)$

The function

$$T_N(t) = e^{-tX_N} \cdots e^{-tX_2} X_1 e^{tX_2} \cdots e^{tX_N} + \cdots + e^{-tX_N} X_{N-1} e^{tX_N} + X_N$$

defined in (25) is very important for all the investigations which follow. Its values in a neighborhood of zero determine the coefficients $c_n^N(X_1, ..., X_N)$ and clearly

$$c_1^N(X_1, ..., X_N) = T_N(0) = X_1 + \cdots + X_n. \tag{27}$$

Consider expansion (10) again: $\psi_N(t, X_1, ..., X_N) = \sum_{n=0}^{\infty} t^n c_n^N(X_1, ..., X_N)$. The coefficients $c_n^N(X_1, ..., X_N)$ do not depend on t. Because $\psi_N(0) = 0$ (see the definition of ψ_N), $c_0^N(X_1, ..., X_N) = 0$ for all N and arbitrary $X_1, ..., X_N$, and therefore the expansion can be written $\psi_N(t) = \sum_{n=1}^{\infty} t^n c_n^N(X_1, ..., X_N)$.

Remark. We write $X = (X_1, ..., X_N)$ for simplicity and omit both X and $(X_1, ..., X_N)$ when the context makes it possible. Recall that the expansion $\sum_{n=1}^{\infty} t^n c_n^N$ converges in a neighborhood of the origin. As noted earlier, the problem is to find expressions for $c_n^N(X)$, and then to show that $\sum_{n=1}^{\infty} c_n^N(X)$ converges to $\psi_N(1, X)$ for $X = (X_1, ..., X_N) \in \underbrace{L \times \cdots \times L}_{N}$ sufficiently small.

THEOREM 2. *A recursion formula for the coefficients $c_n^N(X)$ is as follows. Let $X_1, ..., X_N$ be arbitrary elements in L (i.e., not necessarily small). Define*

$$c_n^N(X) = \frac{1}{n!} \frac{d^n}{dt^n} (\psi_N(t)) \bigg|_{t=0},$$

where $\psi_N(t) = B_N(tX_1, ..., tX_N)$. Then the coefficients c_n^N uniquely satisfy the recursion relations in terms of $T_N^{(k)}(0) \in L$, valid for $n \geq 0$,

$$(n+1)\, c_{n+1}^N = \frac{1}{n!} T_N^{(n)}(0) + \sum_{r=1}^{n} \frac{1}{(n-r)!} \left\{ \frac{1}{2} [c_r^N, T_N^{(n-r)}(0)] + \sum_{\substack{p \geq 1 \\ 2p \leq r}} k_{2p} \sum_{\substack{m_i > 0 \\ m_1 + \cdots + m_{2p} = r}} [c_{m_1}^N, [..., [c_{m_{2p}}^N, T_N^{(n-r)}(0)]...]] \right\}, \tag{28}$$

so that if $n = 0$,

$$c_1^N = T_N(0) = X_1 + \cdots + X_N. \tag{29}$$

Remark. After this theorem is proved, the main task which remains is to find simple formulas for $T_N^k(0)$ so that Theorem 2 can be used to obtain useful information about the coefficients c_n^N. If Eq. (28) is compared with the usual recursion formulas for the case $N = 2$ (see e.g. [15]) one finds an additional complication in our case: the presence of the terms $T_2^{(k)}(0)$.

Proof of Theorem 2. Introduce the expansion

$$\psi_N(t, X) = \sum_{n=1}^{\infty} t^n c_n^N(X)$$

into the basic differential equation (26) of Theorem 1. It follows that

$$\frac{d\psi_N(t)}{dt} = c_1^N + 2tc_2^N + \cdots + (n+1)\, t^n c_n^N + o(t^n). \tag{28'}$$

From the analyticity of the ad representation we have

$$\mathrm{ad}_{\psi_N}(t) = t\,\mathrm{ad}_{c_1^N} + t^2\,\mathrm{ad}_{c_2^N} + \cdots + t^n\,\mathrm{ad}_{c_n^N} + o(t^n). \tag{29$'$}$$

From (29)′ we obtain (for $2p \leqslant n$):

$$(\mathrm{ad}_{\psi_{N(t)}})^{2p} = \sum_{2p \leqslant q \leqslant n} t^q \sum_{\substack{m_i > 0 \\ 1 \leqslant i \leqslant 2p \\ m_1 + \cdots + m_{2p} = q}} \mathrm{ad}_{c_{m_2}^N}\,\mathrm{ad}_{c_{m_2}^N} \cdots \mathrm{ad}_{c_{m_{2p}}^N} + o(t^n). \tag{30}$$

Using (19) for $k(\mathrm{ad}_*)$ we obtain

$$k(\mathrm{ad}_{\psi_{N(t)}}) = 1 + \sum_{\substack{p \geqslant 1 \\ 2p \leqslant n}} k_{2p} (\mathrm{ad}_{\psi_{N(t)}})^{2p} + o(t^n)$$

$$= 1 + \sum_{1 \leqslant q \leqslant n} t^q \sum_{\substack{p \geqslant 1 \\ 2p \leqslant q}} k_{2p}$$

$$\times \sum_{\substack{m_i > 0 \\ m_1 + \cdots + m_{2p} = q}} \mathrm{ad}_{c_{m_1}^N}\,\mathrm{ad}_{c_{m_2}^N} \cdots \mathrm{ad}_{c_{m_2}^N} \cdots \mathrm{ad}_{c_{m_{2p}}^N} + o(t^n). \tag{31}$$

Now we can substitute (28)′, (29)′, and (31) in the differential equation (26).

A simple direct calculation yields

$$\sum_{n=0}^{\infty} t^n (n+1)\, c_{n+1}^N(X)$$

$$= \sum_{n=0}^{\infty} t^n \left\{ \frac{1}{2} [c_n^N(X), T_N(t, X)] + \sum_{\substack{p \geqslant 1 \\ 2p \leqslant n}} k_{2p} \right.$$

$$\left. \times \sum_{\substack{m_i > 0 \\ m_1 + \cdots + m_{2p} = n}} [c_{m_1}^N(X), [\ldots, [c_{m_{2p}}^N(X), T_N(t, X)]\ldots]] \right\} + T_N(t, X). \tag{32}$$

As the coefficients of powers of t in the right side of Eq. (32) involve the function $T_N(t)$, they need not equal the constant coefficients of the corresponding powers of t on the left side. We may, however, consider derivatives $(d^n/dt^n)|_{t=0}$ of both sides of (32). Introduce the notation

$$Q_r(t) = \frac{1}{2} [c_r^N, T_N]$$

$$+ \sum_{\substack{p \geqslant 1 \\ 2p \leqslant n}} k_{2p} \sum_{\substack{m_i > 0 \\ m_1 + \cdots + m_{2p} = r}} [c_{m_1}^N, [\ldots, [c_{m_{2p}}^N, T_N]\ldots]] \tag{33}$$

for $r \geqslant 1$ and

$$Q_0(t) = T_N(t) \qquad \text{for} \quad r = 0. \tag{34}$$

Then (32) can be rewritten as

$$\sum_{r=1}^{\infty} t^r (n+1)\, c_{r+1}^N = \sum_{r=1}^{\infty} t^r Q_r(t). \tag{32'}$$

We also have

$$\begin{aligned} \frac{d^n}{dt^n}(t^r Q_r) &= \sum_{\alpha=0}^{n} C_\alpha^n (t^r)^{(\alpha)} Q_r^{(n-\alpha)} \Big|_{t=0} \\ &= r!\, C_r^n Q_r^{(n-r)}, \qquad n \geqslant r \geqslant 0, \end{aligned}$$

where $C_\alpha^n = n!/(n-\alpha)!\,\alpha!$. Applying $(d^n/dt^n)|_{t=0}$ to both sides of (32)′ then yields

$$(n+1)!\, c_{n+1}^N = \sum_{r=0}^{n} C_r^n r!\, (Q_r(t)) \Big|_{t=0}^{(n-r)},$$

and since

$$\frac{C_r^n r!}{n!} = \frac{1}{(n-r)!}$$

this last equation can be rewritten as

$$(n+1)\, c_{n+1}^N = \sum_{r=0}^{n} \frac{1}{(n-r)!} Q_r^{(n-r)} \Big|_{t=0}. \tag{35}$$

Next we calculate $Q_r^{(n-r)}|_{t=0}$. Using (33), it easily follows that

$$\begin{aligned} Q_r^{(n-r)}|_{t=0} &= \frac{1}{2}[c_r^N, T_N^{(n-r)}(0)] \\ &\quad + \sum_{\substack{p \geqslant 1 \\ 2p \leqslant n}} k_{2p} \sum_{\substack{m_i > 0 \\ m_1 + \cdots + m_{2p} = r}} [c_{m_1}^N, [\ldots[c_{m_{2p}}^N, T_N^{(n-r)}(0)]\ldots] \end{aligned} \tag{36}$$

for all $r \geqslant 1$ and

$$Q_0^{(n-0)}|_{t=0} = T_N^{(n-0)}(0) \qquad \text{for} \quad r = 0.$$

Substituting (36) into (35) yields (28), and the theorem is proved.

In order to make Theorem 2 useful it is necessary to give explicit formulas for $T_N^{(k)}(0)$. These formulas will be proved in the next section, and have the following form.

For $k \geqslant 0$,

$$T_N^{(k)}(0) = ad^k_{-X_N}(X_{N-1}) + \sum_{i=2}^{N-1} \sum_{\alpha_1=0}^{\alpha_0=k} \sum_{\alpha_2=0}^{\alpha_1} \cdots \sum_{\alpha_{N-i}=0}^{\alpha_{N-i-1}} C^k_{\alpha_1} C^{\alpha_1}_{\alpha_2} \cdots C^{\alpha_{N-i-1}}_{\alpha_{N-i}} \times \mathrm{ad}^{k-\alpha_1}_{-X_N} \mathrm{ad}^{\alpha_1-\alpha_2}_{-X_{N-1}} \cdots \mathrm{ad}^{\alpha_{N-i-1}-\alpha_{N-i}}_{-X_{i+1}} \mathrm{ad}^{\alpha_{N-i}}_{-X_i}(X_{i-1}) + X_N^{(k)},$$

where, as X_N does not depend on t, $X_N^{(k)} = 0$ if $k > 0$ and $X_N^{(0)} = X_N$. Note that

$$T_N^{(0)}(0) = T_N(0) = X_1 + X_2 + \cdots + X_N.$$

Convergence of B_N

It has been indicated that $B_N(X_1, ..., X_N) = \sum_{n=1}^{\infty} c_n^N(X_1, ..., X_N)$ converges to $b_N(X_1, ..., X_N)$ if $X_1, ..., X_N$ are in a certain neighborhood V of $0 \in L$. We actually need a stronger result, that the convergence is *uniform* with respect to N provided that the norms of $X_1, ..., X_N$ are required to decrease in proportion to $(1/N)$. A more precise statement is given below. Recall that the Lie algebra L has a norm $|\cdot|$. Let $M \geqslant 1$ be such that $|[X, Y]| \leqslant M|X|\,|Y|$. It is well known that such a number M always exists [19], and in the case of a matrix algebra this number M can be calculated directly, but we do not need the precise value of M. For an arbitary $X = (X_1, ..., X_N)$, $X_i \in L$, denote by $|X|$ the sum $|X| = \sum_{i=1}^N |X_i|$.

LEMMA 3 (Estimation of $T_N^{(k)}(0)$). *Let $T_N(t, X)$, $X = (X_1, ..., X_N)$ be the function defined by* (25) *where X is arbitrary. Then for $k \geqslant 1$*

$$|T_N^{(k)}(0)| \leqslant \frac{1}{M}(M|X|)^{k+1}$$

and

$$|T_N^{(0)}(0)| = |T_N(0)| \leqslant \frac{1}{M}(M|X|),$$

so that for all $k \geqslant 0$,

$$|T_N^{(k)}(0)| \leqslant \frac{1}{M}(M|X|)^{k+1}.$$

Remark. The inequality which we use is the last inequality, valid for $k \geqslant 0$.

As the proof of Lemma 3 is quite complicated it is broken up into three parts, Lemmas 4, 5, and 6. We note, however, that the case $k=0$ is quite simple. What is needed is that $|T_N(0)| \leqslant N|X|$. From (25) it follows that $T_N(0) = X_1 + \cdots + X_N$ and thus $|T_N(0)| \leqslant |X_1| + \cdots + |X_N| \leqslant N|X|$. Before stating Lemmas 4 and 5, write the function $T_N(t, X)$ in the form

$$\begin{aligned} T_N(t, X) = t_{N,2}(t, X_1) + t_{N,3}(t, X_2) \\ + \cdots + t_{N,N}(t, X_{N-1}) + t_{N,N+1}(t, X_N), \end{aligned} \tag{37}$$

where the functions $t_{N,i}$, $2 \leqslant i \leqslant N$ are defined as by

$$\begin{aligned} t_{N,2}(t, a) &= e^{-tX_N} \cdots e^{-tX_2} a e^{tX_2} \cdots e^{tX_N} \\ t_{N,3}(t, a) &= e^{-tX_N} \cdots e^{-tX_3} a e^{tX_3} \cdots e^{tX_N} \\ &\cdots\cdots\cdots\cdots \\ t_{N,i}(t, a) &= e^{-tX_N} \cdots e^{-tX_i} a e^{tX_i} \cdots e^{tX_N} \\ &\cdots\cdots\cdots\cdots \\ t_{N,N}(t, a) &= e^{-tX_N} a e^{tX_N}, \\ t_{N,N+1}(t, a) &= a, \end{aligned} \tag{38}$$

and for all N, so that $t_{N,N+1}(t, X_N) = X_N$. We have thus

$$T_N(t, X) = \sum_{i=2}^{N} t_{N,i}(t, X_{i-1}) + X_N$$

and

$$T_N^{(k)}(t, X) = \sum_{i=2}^{N} t_{N,i}^{(k)}(t, X_{i-1}). \tag{39}$$

LEMMA 4. *The following equality is valid for $2 \leqslant i \leqslant N$ and $k \geqslant 0$:*

$$t_{N,i}^{(k)}(t, a) = e^{-tX_N} \left(\sum_{\alpha=0}^{k} C_\alpha^k \operatorname{ad}_{-X_N}^{k-\alpha} (t_{N-1,i}^{(\alpha)}(t, a)) \right) e^{tX_N}. \tag{40}$$

To prove Lemma 4 we need the following result.

LEMMA 5. *Consider the L-valued smooth function $f(t) = e^{-tX_N} g(t) e^{tX_N}$,* *where $g(t)$ is an arbitrary smooth L-valued function of the real variable t. Then*

$$\frac{d^k}{dt^k} f(t) = e^{-tX_N} \left(\frac{d}{dt} + \mathrm{ad}_{-X_N} \right)^k g(t) \Big) e^{tX_N}, \tag{41}$$

where

$$\left(\frac{d}{dt} + \mathrm{ad}_X\right)^k = \sum_{\alpha=0}^{k} C_\alpha^k \, \mathrm{ad}_X^{k-\alpha} \circ \frac{d^\alpha}{dt^\alpha} = \sum_{\alpha=0}^{k} C_\alpha^k \frac{d^\alpha}{dt^\alpha} \circ \mathrm{ad}_X^{k-\alpha}$$

for any constant $X \in L$.

Proof of Lemma 5. Because $f = e^{-tX_N} g e^{tX_N}$, then $f' = (-X_N) e^{-tX_N} g e^{tX_N} + e^{-tX_N} g' e^{tX_N} + e^{-tX_N} g X_N e^{tX_N} = e^{-tX_N}(-X_N g + g' + g X_N) e^{tX_N} = e^{-tX_N}((d/dt + \mathrm{ad}_{-X_N}) g) e^{tX_N}$. Here $-X_N g + g X_N = -[X_N, g] = \mathrm{ad}_{-X_N}(g)$, since $X_N \cdot e^{tX_N} = e^{tX_N} \cdot X_N$. We see that the first derivative f' of the initial function f again has the form $e^{-tX_N} g_1 e^{tX_N}$, where $g_1 = (d/dt + \mathrm{ad}_{-X_N}) g$. This means that we can put $f' = f_1 = e^{-tX_N} g_1 e^{tX_N}$ and repeat the preceding calculation to obtain $f_1' = e^{-tX_N}((d/dt + \mathrm{ad}_{-X_N}) g_1) e^{tX_N}$, i.e., $f'' = e^{-tX_N}((d/dt + \mathrm{ad}_{-X_N})^2 g) e^{tX_N}$. Iterating the process clearly yields Lemma 5.

Remark. Because X_N is a fixed and constant element in the Lie algebra, the two operators d/dt and ad_{-X_N} commute. Indeed, $((d/dt) \circ \mathrm{ad}_{-X_N}) q(t) = [-X_N, q(t)]' = [-X_N, q'(t)] = (\mathrm{ad}_{-X_N} \circ d/dt) q(t)$ for an arbitrary smooth function $q(t)$. This justifies the use of the usual binomial formula for the polynomial $(d/dt + \mathrm{ad}_{-X_N})^k = \sum_{\alpha=0}^{k} C_\alpha^k \, \mathrm{ad}_{-X_N}^{k-\alpha} (d^\alpha/dt^\alpha)$.

Proof of Lemma 4. Letting $f(t) = t_{N,i}(t, a) = e^{-tX_N} \cdots e^{-tX_i} \cdot a \cdot e^{tX_i} \cdots e^{tX_N} = e^{-tX_N}(e^{-tX_{N-1}} \cdots e^{-tX_i} \cdot a \cdot e^{tX_i} \cdots e^{tX_{N-1}}) e^{tX_N} = e^{-tX_N} \cdot t_{N-1,i}(t, a) \cdot e^{tX_N}$, we can put $g(t) = t_{N-1,i}(t, a)$ to obtain that $f(t) = e^{-tX_N} g(t) e^{tX_N}$. Lemma 5 then implies

$$f^{(k)}(t) = t_{N,i}^{(k)}(t, a) = e^{-tX_N} \left(\left(\frac{d}{dt} + \mathrm{ad}_{-X_N} \right)^k t_{N-1,i}(t, a) \right) e^{tX_N},$$

and the binomial formula (see preceding remark) then yields (40) and Lemma 4 is proved.

LEMMA 6. *For all $k \geqslant 0$ and $1 \leqslant i \leqslant N$ we have*

$$|t_{N,i}^{(k)}(0, a)| \leqslant M^k |a| (|X_i| + \cdots + |X_N|)^k, \tag{41$'$}$$

where the functions $t_{N,i}(t, a)$ are defined by (38).

Proof of Lemma 6. We use induction with respect to N, with i and k arbitrary but within the restrictions listed above.

(a) We first prove the statement for $i = N \geqslant 1$, and all $k \geqslant 0$ (only needed for $N = 1$ in the initial induction step). We have $t_{N,i}(t, a) = t_{i,i}(t, a) = e^{-tX_i} \cdot a \cdot e^{tX_i}$, and so $|t_{i,i}(0, a)| = |a| \leqslant |X| = (1/M) N^O(M|X|)^{O+1}$, i.e., the estimate is valid for $k = 0$. Consider an arbitrary $k \geqslant 0$. Lemma 4 implies that $t_{i,i}^{(k)}(t, a) = e^{-tX_i}((d/dt + \mathrm{ad}_{-X_i})^k t_{i-1,i}(t, a))\, e^{tX_i} = e^{-tX_i}((d/dt + \mathrm{ad}_{-X_i})^k a)\, e^{tX_i}$ because $t_{i-1,i}(a) = a$ as we can see from (38), and as $a \in L$ is a constant, $t_{i,i}^{(k)}(t, a) = e^{-tX_i}(\mathrm{ad}^k_{-X_i}(a))\, e^{tX_i}$, and thus $|t_{i,i}^{(k)}(O, a)| = |\mathrm{ad}^k_{-X_i}(a)| \leqslant M^k|X_i|^k\, |a|$. Thus the statement of the lemma is proved for all $k \geqslant 0$ and $i = N \geqslant 1$.

Remark. If the norm $|\cdot|$ of the Lie algebra L is invariant under the adjoint action of Lie group G, i.e., $|X| = |gXg^{-1}|$ for all $X \subset L$ and $g \in G$, then it follows from the preceding calculations that a stronger estimate is valid: $|t_{i,i}^{(k)}(t, a)| \leqslant M^k|X_i|^k\, |a|$, $k \geqslant 0$, $i = N \geqslant 1$, for all t (not only for $t = 0$). If $L \subset GL(m, \mathbb{R})$ one may assume that the norm $|\;|$ is invariant under the adjoint action of G.

(b) Using induction, let us suppose that (41)′ is valid for all $k \geqslant 0$ and $1 \leqslant i \leqslant N-1$. We need to prove (41)′ for all $k \geqslant 0$ and $1 \leqslant i \leqslant N$.

We have from Lemma 4 that $|t_{N,i}^{(k)}(0, a)| \leqslant \sum_{\alpha=0}^k C_\alpha^k |\mathrm{ad}^{k-\alpha}_{-X_N}(t_{N-1,i}^{(\alpha)}(0, a))| \leqslant \sum_{\alpha=0}^k C_\alpha^k \cdot M^{k-\alpha} \cdot |X_N|^{k-\alpha} \cdot |t_{N-1,i}^{(\alpha)}(o, a)| \leqslant$ (using the induction hypothesis for $i \leqslant N-1$) $\leqslant \sum_{\alpha=0}^k C_\alpha^k \cdot M^{k-\alpha}|X_N|^{k-\alpha}\, |a|\, M^\alpha(|X_i| + \cdots + |X_N|)^k = |a|\, M^k(|X_i| + \cdots + |X_N|)^k$, and thus we have proved (41)′ for all $k \geqslant 0$ and $i \leqslant N-1$. The case $i = N$ has already been established in part (a). This proves Lemma 6.

Remark. Again in the case when the norm $|\cdot|$ is invariant under the adjoint action of G, we have actually shown that $|t_{N,i}^{(k)}(t, a)| \leqslant |a|\, M^k(|X_i| + \cdots + |X_N|)^k$ for arbitrary t in the domain of analyticity.

Proof of Lemma 3. We use induction with respect to k with N an arbitrary fixed number.

(a) $k = 0$. Then $|T_N(0)| = |\sum_{i=2}^N t_{N,i}(0, X_{i-1}) + X_N| = |\sum_{i=1}^N X_i|$.

(b) Consider the case of arbitrary $k \geqslant 1$. Then $|T_N^{(k)}(0)| = |\sum_{i=2}^N t_{N,i}^{(k)}(0, X_{i-1})| \leqslant \sum_{i=2}^N |t_{N,i}^{(k)}(0, X_{i-1})| \leqslant$ (using Lemma 6 and (41)′) $\leqslant \sum_{i=2}^N M^k|X_{i-1}|(|X_i| + \cdots + |X_N|)^k \leqslant M^k \sum_{i=2}^N |X_{i-1}|(|X_i| + \cdots + |X_N|)^k \leqslant M^k(|X_2| + \cdots + |X_N)^k(|X_1| + \cdots + |X_{N-1}|) \leqslant M^k(|X_1| + \cdots + |X_N|)^{k+1}$ and thus Lemma 3 is proved.

We have noted that in the case of an invariant norm, we can obtain the estimate not only for $T_N^{(k)}(0)$, but even for $T_N^{(k)}(t)$. In fact, the following statement is valid.

LEMMA 3* (Estimate of $T_N^{(k)}(t)$). *Let* $T_N(t, X_1, ..., X_N)$ *be defined by*

(25) *where* $X_1, ..., X_N \in L$, *and* L *has an adjoint invariant norm. Then for each* t *in the interval of analyticity of* T_N *we have*

$$|T_N^{(k)}(t, X)| \leqslant \frac{1}{M}(M|X|)^{k+1}, \qquad k \geqslant 1$$

and

$$|T_N(t, X)| \leqslant |X|.$$

The proof of Lemma 3* is similar to that of Lemma 3, with modifications suggested by the remarks following Lemmas 4, 5, and 6.

We are now prepared to prove the main theorem of this section, the uniform convergence of B_N for certain $X_1, ..., X_N$.

THEOREM 3. *There exists* $\delta > 0$ *such that if* $|X_1| + \cdots + |X_N| < \delta$, *then* $B_N(X_1, ..., X_N)$ *converges absolutely uniformly in* N, *and converges to* $b_N(X_1, ..., X_N)$. *Furthermore the function*

$$B_N(X_1, ..., X_N) = \sum_{n=1}^{\infty} c_n^N(X_1, ..., X_N)$$

defines an analytic mapping from $U \subset L \times \cdots \times L$ (N *times*), *where*

$$U = \{(X_1, ..., X_N): |X_1| + \cdots + |X_N| < \delta\},$$

into L, *and the identity*

$$e^{X_1} \cdots e^{X_N} = e^{B_N(X_1, ..., X_N)}$$

is valid for all $X_1, ..., X_N \in U$.

The proof of Theorem 3 will be divided into two lemmas, Lemmas 7 and 8. We first introduce a certain scalar differential equation whose solution is very closely connected with the L-valued function $\psi_N(t, X)$.

The Main Differential Equation

Let $q(z)$ be a complex function of a complex variable $z \in \mathbb{C}$, defined by

$$q(z) = 1 + \sum_{p=1}^{\infty} |k_{2p}| \, z^{2p}, \tag{42}$$

where the k_{2p} are those of (16). This equation is a scalar version of (16), which was used in the calculation of the differential of the operator inverse to $e^{-X} \circ (d \exp_X)$. Consider also the associated function $k(z) = 1 + \sum_{p=1}^{\infty} k_{2p} z^{2p}$. Then by the Cauchy–Hadamard formula we have that the radius of convergence of the expansion $1+\sum_{p=1}^{\infty} k_{2p}z^{2p}$ for the function $k(z)$ is the same as the radius of convergence of the expansion for $q(z)$. Because $k(z)=1/(1-e^{-z})-(z/2)$ (see [15]), the singularities of the function $k(z)$ closest to zero are the points $\pm 2\pi i$. Therefore, the radius of convergence of the expansions $1+\sum_{p=1}^{\infty} k_{2p}z^{2p}$ and $1+\sum_{p=1}^{\infty} |k_{2p}|\, z^{2p}$ is equal to 2π.

Consider the differential equation

$$\frac{dh(z)}{dz}=e^{z}\left(q(h)+\frac{h}{2}\right). \tag{43}$$

According to the general theory of differential equations, there is a positive number $\hat{b}$, where $\hat{b}<2\pi$, such that the equation (43) has an analytic solution $h(z)$ in the disk $(|z|<\hat{b})$ for which $h(0)=0$. Let us fix this number $\hat{b}$.

We next investigate the properties of the solution $h(z)$ of (43). Let $h=h(z)$ be the analytic function on the disk $|z|<\hat{b}$ that solves Cauchy's problem with $h(0)=0$ for (43). Consider the expansion $(|z|<\hat{b})$

$$h(z)=\sum_{n=1}^{\infty} \rho_n z^n, \tag{44}$$

where $\rho_0=0$ because $h(0)=0$.

LEMMA 7. *The coefficients ρ_n of the function $h(z)$ satisfy the recursion formulas*

$$(n+1)\,\rho_{n+1}=\frac{1}{n!}+\sum_{r=1}^{n}\frac{1}{(n-r)!}\times\left(\frac{1}{2}\rho_r+\sum_{\substack{p\geqslant 1\\ 2p\leqslant r}}|k_{2p}|\sum_{\substack{m_i>0\\ m_1+\cdots+m_{2p}=r}}\rho_{m_1}\cdots\rho_{m_{2p}}\right) \tag{45}$$

for $n\geqslant 1$, and $\rho_1=1$. It then follows that $\rho_n>0$ for all $n\geqslant 1$.

Remark. The recursion formulas can be written formally as

$$(n+1)\,\rho_{n+1}=\sum_{r=0}^{n}\frac{1}{(n-r)!}\times\left(\frac{1}{2}\rho_r+\sum_{\substack{p\geqslant 1\\ 2p\leqslant r}}|k_{2p}|\sum_{\substack{m_i>0\\ m_1+\cdots+m_{2p}=r}}\rho_{m_1}\cdots\rho_{m_{2p}}\right),$$

where $1/n!$ is the factor corresponding to $r=0$.

Remark. Compare the formulas of (45) with the formulas of (28). It is remarkable and surprising that these formulas are very similar. This is the reason that the coefficients ρ_n have a pivotal rôle in the definition of the logarithm functional.

Proof of Lemma 7. We will use in part arguments from the proof of Theorem 2. Consider the expansion $q(h)+h/2=\sum_{r=0}^{\infty} d_r z^r$.

Let us calculate the coefficients d_r. Because $q(h)=1+\sum_{p=1}^{\infty} |k_{2p}|\, h^{2p}$ (where $k_{2p}(2p)!$ are Bernoulli's numbers) and $h=\sum_{n=1}^{\infty} \rho_n z^n$, then, using the method of calculation which was used in the proof of Theorem 2, we obtain the system of equations

$$d_r=\frac{1}{2}\rho_r+\sum_{\substack{p\geqslant 1\\ 2p\leqslant r}} |k_{2p}| \sum_{\substack{m_i>0\\ m_1+\cdots+m_{2p}=r}} \rho_{m_1}\cdots\rho_{m_{2p}} \tag{46}$$

for all $r\geqslant 1$ and $d_0=1$. Now consider Eq. (43) and substitute in this equation relations (44) and (46). We obtain

$$\sum_{n=0}^{\infty}(n+1)\,\rho_{n+1}z^n=\left(1+\frac{z}{1!}+\frac{z^2}{2!}+\cdots+\frac{z^r}{r!}+\cdots\right)\left(\sum_{r=0}^{\infty} d_r z^r\right).$$

It is clear that the coefficient of z^n in the right side of this equation is equal to $\sum_{r=0}^{n}(1/(n-r)!)\,d_r$, i.e., $z^n((n+1)\,\rho_{n+1})=z^n\sum_{r=0}^{n}(1/(n-r)!)\,d_r$. From this equation we obtain (45) and Lemma 7 is proved.

Let us note an important property of the coefficients ρ_n: *the expansion* $\sum_{n=1}^{\infty}\rho_n|z|^n$ *converges for all z such that* $|z|<\hat{b}$ (see above).

LEMMA 8 (Estimate for the coefficients c_n^N). *Consider the coefficients* c_n^N *of* B_N *(which satisfy formulas* (28) *and* (29) *for* $n\geqslant 1$). *The following inequalities hold for all* $n\geqslant 1$:

$$|c_n^N(X_1,\ldots,X_N)|\leqslant\frac{1}{M}(M|X|)^n\cdot\rho_n \tag{47}$$

for arbitrary $X_1,\ldots,X_N\in L$.

Proof of Lemma 8. We use induction with respect to n.

(a) $n=1$. Then we have

$$c_n^N(X_1,\ldots,X_N)=\sum_{i=1}^{N} X_i \qquad\text{and}\qquad |c_1^N|\leqslant\sum_{i=1}^{N}|X_i|=|X|.$$

(b) Let us assume that (47) is proved for all c_s^N, where $s=1,2,\ldots,n$. Consider c_{n+1}^N. Then

$$|c_{n+1}^N(X)| = (\text{from } (35)) = \left| \frac{1}{n+1} \sum_{r=0}^{n} \frac{1}{(n-r)!} Q_r^{(n-r)}(0) \right|$$

$$\leqslant \frac{1}{n+1} \sum_{r=0}^{n} \frac{1}{(n-r)!} \left\{ \frac{1}{2} \right| [c_r^N, T_N^{(n-r)}(0)] + \sum_{\substack{p \geqslant 1 \\ 2p \leqslant r}} |k_{2p}|$$

$$\times \sum_{\substack{m_i > 0 \\ m_1 + \cdots + m_{2p} = r}} [c_{m_1}^N, [\ldots, [c_{m_{2p}}^N, T_N^{(n-r)}(0)]\ldots]] \left| \right\}$$

$$\leqslant \frac{1}{n+1} \sum_{r=0}^{n} \frac{1}{(n-r)!} \left\{ \frac{1}{2} M |c_r^N| \, |T_N^{(n-r)}(0)| + \sum_{\substack{p \geqslant 1 \\ 2p \leqslant r}} |k_{2p}| \right.$$

$$\left. \times \sum_{\substack{m_i > 0 \\ m_1 + \cdots + m_{2p} = r}} M^{2p} |T_N^{(n-r)}(0)| \, |c_{m_1}^N| \cdots |c_{m_{2p}}^N| \right\}$$

and using induction, Lemma 3, Lemma 7, and (47)

$$\leqslant \frac{1}{n+1} \sum_{r=0}^{n} \frac{1}{(n-r)!} \frac{1}{M} (M|X|)^{n-r+1}$$

$$\times \left(\frac{1}{2} M \frac{1}{M} (M|X|)^r \rho_r + \sum_{\substack{p \geqslant 1 \\ 2p \leqslant r}} |k_{2p}| \right.$$

$$\left. \times \sum_{\substack{m_i > 0 \\ m_1 + \cdots + m_{2p} = r}} M^{2p} M^{-2p} (M|X|)^{m_1 + \cdots + m_{2p}} \rho_{m_1} \cdots \rho_{m_{2p}} \right)$$

and since $m_1 + \cdots + m_{2p} = r$,

$$= \frac{(M|X|)^{n+1}}{M(n+1)} \sum_{r=0}^{n} \frac{1}{(n-r)!}$$

$$\times \left\{ \frac{1}{2} \rho_r + \sum_{\substack{p \geqslant 1 \\ 2p \leqslant r}} |k_{2p}| \sum_{\substack{m_i > 0 \\ m_1 + \cdots + m_{2p} = r}} \rho_{m_1} \cdots \rho_{m_{2p}} \right\}$$

and using 45,

$$= \frac{(M|X|)^{n+1}}{M(n+1)} (n+1) \rho_{n+1}$$

$$= \frac{(M|X|)^{n+1}}{M} \rho_{n+1}$$

and thus Lemma 8 is proved.

Proof of Theorem 3. Consider the expansion

$$\sum_{n=1}^{\infty} c_n^N(X_1, ..., X_N) = \sum_{n=1}^{\infty} c_n^N(X).$$

Then

$$\begin{aligned}\left|\sum_{n=1}^{\infty} c_n^N(X)\right| &\leqslant \sum_{n=1}^{\infty} |c_n^N(X)| \\ &\leqslant \frac{1}{M} \sum_{n=1}^{\infty} \rho_n(M|X|)^n \qquad \text{(using Lemma 8)} \\ &= \frac{1}{M} \sum_{n=1}^{\infty} \rho_n |z_0|^n \qquad \text{for} \quad |z_0| = M|X|.\end{aligned}$$

We know that $\sum_{n=1}^{\infty} \rho_n |z|^n$ converges if $|z| < \hat{b}$. Since $\sum_{n=1}^{\infty} \rho_n |z_0|^n$ converges for $M|X| < \hat{b}$, this implies absolute uniform convergence if

$$\sum_{i=1}^{N} |X_i| < \hat{b}/M$$

and the first part of the theorem is proved with $\delta = (\hat{b}/M)$.

Note that e^{B_N} is analytic where B_N is analytic, and $e^{X_1} \cdots e^{X_N}$ is analytic everywhere. Since as already noted, $e^{B_N} = e^{X_1} \cdots e^{X_N}$ in a neighborhood of $(0 \times \cdots \times 0)$, this implies that

$$e^{B_N} = e^{X_1} \cdots e^{X_N}$$

in the domain of analyticity of B_N. This concludes the proof of Theorem 3.

Remark. Theorem 3 is used in proving the continuous form of the logarithm functional.

THEOREM 4. *For $X_1, ..., X_N \in L$ arbitrary and real μ,*

$$c_n^N(\mu X_1, ..., \mu X_N) = \mu^n c_n^N(X_1, ..., X_N).$$

Proof of Theorem 4. The proof is obtained using induction on n.

(a) For $n = 1$, $c_1^N(\mu X_1, ..., \mu X_N) = \mu X_1 + \cdots + \mu X_N = \mu c_1^N(X_1, ..., X_N)$.

(b) By induction and by the recursion formulas of (28) it follows that

$$(n+1)\,c^N_{n+1}(\mu X)$$
$$=\frac{1}{n!}T^{(n)}_N(0,\mu X)+\sum_{r=1}^{n}\frac{1}{(n-r)!}\left\{\frac{1}{2}\mu^r[c^N_r(X),T^{(n-r)}_N(0,\mu X)]\right.$$
$$+\sum_{\substack{p\geqslant 1\\ 2p\leqslant r}}k_{2p}\sum_{\substack{m_i>0\\ m_1+\cdots+m_{2p}=r}}\mu^{m_1+\cdots+m_{2p}=r}$$
$$\left.\times[c^N_{m_1}(X),[...,[c^N_{m_{2p}}(X),T^{(n-r)}_N(0,\mu X)]...]\right\}.$$

Theorem 4 will thus be proved if we show that $T^{(n-r)}_N(0,\mu X)=\mu^{n-r}\cdot T^{(n-r)}_N(0,X)$, or what is the same thing, that $T^{(k)}_N(0,\mu X)=\mu^k\cdot T^{(k)}_N(0,X)$ for all $k\geqslant 0$. For $k=0$ we have $T^{(0)}_N(0,\mu X)=T_N(0,\mu X)=\mu X_1+\cdots+\mu X_N=\mu T^{(0)}_N(0,X)$. Consider the case $k\geqslant 1$. From (39) we obtain $T^{(k)}_N(0,\mu X)=\sum_{i=2}^N t^{(k)}_{N,i}(0,\mu X_{i-1})$. The problem is thus reduced to the following statement:

$$t^{(k)}_{N,i}(0,\mu a)=\mu^k t^{(k)}_{N,i}(0,a)\qquad\text{for any}\quad a\in L. \tag{48}$$

From (40), (see Lemma 4) we have

$$t^{(k)}_{N,i}(0,\mu a)=\sum_{\alpha=0}^{k}C^k_\alpha\,\mathrm{ad}^{k-\alpha}_{-\mu X_N}(t^{(\alpha)}_{N-1,i}(0,\mu a))$$
$$=\sum_{\alpha=0}^{k}C^k_\alpha\mu^{k-\alpha}\,\mathrm{ad}^{k-\alpha}_{-X_N}(t^{(\alpha)}_{N-1,i}(0,\mu a)). \tag{49}$$

From this formula it follows that it is sufficient to prove that

$$t^{(\alpha)}_{N-1,i}(0,\mu a)=\mu^\alpha t^{(\alpha)}_{N-1,i}(0,a).$$

We prove this using induction on N. Recall that for each N, $t_{N,i}$ satisfies (49) for $2\leqslant i\leqslant N$, and thus by the induction hypothesis we see that it is sufficient to prove

$$t^{(\alpha)}_{i,i}(0,\mu a)=\mu^\alpha t^{(\alpha)}_{i,i}(0,a)$$

for $i=2$ and $i=N$. From the proof of Lemma 6 we have

$$t^{(\alpha)}_{i,i}(t,a)=e^{-tX_i}(\mathrm{ad}^\alpha_{-X_i}(a))\,e^{tX_i},\qquad\text{and so}$$
$$t^{(\alpha)}_{i,i}(0,\mu a)=\mathrm{ad}^\alpha_{-\mu X_i}(a)=\mu^\alpha\,\mathrm{ad}^\alpha_{-X_i}(a)=\mu^\alpha t^{(\alpha)}_{i,i}(a).$$

This calculation finishes the proof of Theorem 4.

5. Algebraic Properties of c_n^N.

First we calculate the first few coefficients $c_n^N(X_1, ..., X_N)$. Their exact expressions are useful in applications and also in understanding their algebraic structure.

(1) The first term $c_1^N(X_1, ..., X_N)$. From (29) we see that

$$c_1^N(X_1, ..., X_N) = T_N(0) = X_1 + \cdots + X_N.$$

(2) The second term $c_2^N(X_1, ..., X_N)$. From the recursion formulas (28) we see that

$$2c_2^N = T_N^{(1)}(0) + \tfrac{1}{2}[c_1^N, T_N(0)] = T_N^{(1)}(0),$$

and therefore $c_2^N = \frac{1}{2}T_N^{(1)}(0)$. Let us calculate $T_N^{(1)}(0)$. Consider $T_N(t) = e^{-tX_N} \cdots e^{-tX_2} X_1 e^{tX_2} \cdots e^{tX_N} + \cdots + e^{-tX_N} X_{N-1} e^{tX_N} + X_N = e^{-tX_N} T_{N-1}(t) e^{tX_N} + X_N$. Then $T_N^{(1)}$ = (using (41)) = $e^{-tX_N}[T_{N-1}(t), X_N]\, e^{tX_N} + e^{-tX_N} T_{N-1}^{(1)}(t)\, e^{tX_N}$. For $t = 0$ we have

$$\begin{aligned} T_N^{(1)}(0) &= [T_{N-1}(0), X_N] + T_{N-1}^{(1)}(0) \\ &= \sum_{i=1}^{N-1} [X_i, X_N] + T_{N-1}^{(1)}(0). \end{aligned} \tag{50}$$

Lemma 9. *The following formula is valid for all* N:

$$2c_2^N = T_N^{(1)}(0) = \sum_{1 \leqslant i < j \leqslant N} [X_i, X_j]. \tag{51}$$

Proof of Lemma 9. We can use the recursion formula (50) and induction.

(a) For $N = 2$ we have $T_2(t) = e^{-tX_2} X_1 e^{tX_2} + X_2$ and it is clear that $T_2^{(1)}(0) = [X_1, X_2]$.

(b) Consider the arbitrary case. Here we have $T_N^{(1)}(0) = T_{N-1}^{(1)}(0) + \sum_{i=1}^{N-1} [X_i, X_N] = \sum_{1 \leqslant i < j \leqslant N-1} [X_i, X_j] + \sum_{i=1}^{N-1} [X_i, X_N]$ (from the induction hypothesis) $= \sum_{1 \leqslant i < j \leqslant N} [X_i, X_j]$. Lemma 9 is proved.

Thus we have

$$c_2^N(X_1, ..., X_N) = \frac{1}{2} \sum_{1 \leqslant i < j \leqslant N} [X_i, X_j]. \tag{52}$$

(3) The third term $c_3^N(X_1, ..., X_N)$. From the recursion formulas (28) we have

$$\begin{aligned}3c_3^N &= \tfrac{1}{2}T_N^{(2)}(0) + \tfrac{1}{2}[c_1^N, T_N^{(1)}(0)] + \tfrac{1}{2}[c_2^N, T_N(0)] + k_2[c_1^N, [c_1^N, T_N(0)]] \\ &= \tfrac{1}{2}T_N^{(2)}(0) + \tfrac{1}{2}[T_N(0), T_N^{(1)}(0)] + \tfrac{1}{2}[\tfrac{1}{2}T_N^{(1)}(0), T_N(0)] \\ &\qquad \text{(using (52) and (29))} \\ &= \tfrac{1}{2}T_N^{(2)}(0) + \tfrac{1}{4}[T_N(0), T_N^{(1)}(0)].\end{aligned}$$

Thus

$$c_3^N = \frac{1}{6}T_N^{(2)}(0) + \frac{1}{12}\left[X_1 + \cdots + X_N, \sum_{1 \leqslant i < j \leqslant N} [X_i, X_j]\right]. \tag{53}$$

A direct calculation of the other c_n^N is more difficult. Now it is interesting to compare our general formulas in the case $N = 2$ with the classical ones. We consider only the first three terms for simplicity.

(1) In the classical case $c_1(X, Y) = X + Y$. Clearly our function $c_1^{N=2} = X_1 + X_2$ and thus agrees with the classical c_1.

(2) In the classical case $c_2(X, Y) = \frac{1}{2}[X, Y]$. It is clear from (51) that our function $c_2^{N=2}$ is equal to $\frac{1}{2}[X_1, X_2]$, i.e., agrees with the classical one.

(3) In the classical case $c_3(X, Y) = \frac{1}{12}[X - Y, [X, Y]]$. In our case we have

$$c_3^2(X_1, X_2) = \tfrac{1}{6}T_2^{(2)}(0) + \tfrac{1}{12}[X_1 + X_2, [X_1, X_2]].$$

A direct calculation shows that $T_2^{(2)}(0) = [[X_1, X_2], X_3]$, so that

$$\begin{aligned}c_3^2(X_1, X_2) &= \tfrac{1}{6}[[X_1, X_2], X_2] + \tfrac{1}{12}[X_1 + X_2, [X_1, X_2]] \\ &= \tfrac{1}{12}[X_1 - X_2, [X_1, X_2]],\end{aligned}$$

which agrees with the classical expression.

Using induction it is possible and not too difficult to show that our expansion $B_2(X_1, X_2)$ coincides with the classical one. We omit details. For example, $c_4^2(X_1, X_2) = -\frac{1}{48}[X_2, [X_1, [X_1, X_2]]] - \frac{1}{48}[X_1, [X_2, [X_1, X_2]]]$.

We next calculate the derivatives $T_N^{(k)}(0)$, as they appear in nearly all the basic recursion formulas.

Lemma 10 (A formula for $T_N^{(k)}(0)$). *The following formula is valid for* $k \geqslant 0$:

$$\begin{aligned}T_N^{(k)}(0) = X_N^{(k)} + \sum_{i=2}^{N-1} \sum_{\alpha_1=0}^{\alpha_0=k} \sum_{\alpha_2=0}^{\alpha_1} \cdots \sum_{\alpha_{N-i}=0}^{\alpha_{N-i-1}} C_{\alpha_1}^{k} C_{\alpha_2}^{\alpha_1} \\ \cdots C_{\alpha_{N-i}}^{\alpha_{N-i-1}} \operatorname{ad}_{-X_N}^{k-\alpha_1} \operatorname{ad}_{-X_{N-1}}^{\alpha_1-\alpha_2} \\ \cdots \operatorname{ad}_{-X_{i+1}}^{\alpha_{N-i-1}-\alpha_{N-i}} \operatorname{ad}_{-X_i}^{\alpha_{N-i}}(X_{i-1}) + \operatorname{ad}_{-X_N}^{k}(X_{N-1}).\end{aligned}$$

Here $\alpha_{N-i+1}=0$ *if* $i>N$ *and*

$$X_N^{(k)}=\frac{d^k}{dt^k}X_N=\begin{cases}0, & k>1\\ X_N, & k=0\end{cases}.$$

If $k=0$,

$$T_N^{(0)}(0)=T_N(0)=X_1+X_2+\cdots+X_N. \tag{54}$$

Proof of Lemma 10. It follows from (37) that

$$T_N(t, X)=t_{N,2}(t, X_1)+\cdots+t_{N,N}(t, X_{N-1})+X_N$$

and for $k\geqslant 1$

$$T_N^{(k)}(0, X)=t_{N,2}^{(k)}(0, X_1)+\cdots+t_{N,N}^{(k)}(0, X_{N-1}).$$

Note that according to this notation $T_N^{(0)}(0, X)\neq T_N(0, X)$, which is inconvenient, and so we write

$$T_N^{(k)}(0, X)=t_{N,2}^{(k)}(0, X_1)+\cdots+t_{N,N}^{(k)}(0, X_{N-1})+X_N^{(k)},$$

which is valid for all $k\geqslant 0$.

From (40) we have $(2\leqslant i\leqslant N)$

$$t_{N,i}^{(k)}(0, X_{i-1})=\sum_{\alpha=0}^{k} C_\alpha^k \operatorname{ad}_{-X_N}^{k-\alpha}(t_{N-1,i}^{(\alpha)}(0, X_{i-1})),$$

for all $k\geqslant 0$. For $k=0$ we have

$$t_{N,i}^{(0)}(0, X_{i-1})=C_0^0 \operatorname{ad}_{-X_N}^{0-0}(t_{N-1,i}^{(0)}(0, X_{i-1}))=X_{i-1}.$$

We can iterate the process and obtain

$$t_{N,i}^{(k)}(0, X_{i-1})=\sum_{\alpha_1=0}^{k=\alpha_0} C_{\alpha_1}^k \operatorname{ad}_{-X_N}^{k-\alpha_1}\left(\sum_{\alpha_2=0}^{\alpha_1} C_{\alpha_2}^{\alpha_1} \operatorname{ad}_{-X_{N-1}}^{\alpha_1-\alpha_2}(t_{N-2,i}^{(\alpha_2)}(0, X_{i-1}))\right)$$

and so on. Finally, we obtain

$$t_{N,i}^{(k)}(0, X_{i-1})=\sum_{\alpha_1=0}^{k=\alpha_0}\sum_{\alpha_2=0}^{\alpha_1}\cdots\sum_{\alpha_{N-i-1}=0}^{\alpha_{N-i-2}} C_{\alpha_1}^k C_{\alpha_2}^{\alpha_1}\cdots C_{\alpha_{N-i-1}}^{\alpha_{N-i-2}} \operatorname{ad}_{-X_N}^{k-\alpha_1}\operatorname{ad}_{-X_{N-1}}^{\alpha_1-\alpha_2}\cdots\operatorname{ad}_{-X_{i+1}}^{\alpha_{N-i-1}-\alpha_{N-i}} t_{i,i}^{(\alpha_{N-1})}(0, X_{i-1}).$$

But $t_{i,i}(t, X_{i-1})=e^{-tX_i}X_{i-1}e^{tX_i}$ and for all $\alpha\geqslant 0$

$$t_{i,i}^{(\alpha)}(t, X_{i-1})=e^{-tX_i}(\operatorname{ad}_{-X_i}^{\alpha}(X_{i-1}))\,e^{tX_i}.$$

Thus $t_{i,i}^{(\alpha)}(0, X_{i-1}) = \mathrm{ad}^{\alpha}_{-X_i}(X_{i-1})$ for all $\alpha \geqslant 0$. This implies that for $k \geqslant 0$, $N > i$,

$$t_{N,i}^{(k)}(0, X_{i-1}) = \sum_{\alpha_1=0}^{k=\alpha_0} \sum_{\alpha_2=0}^{\alpha_1} \cdots \sum_{\alpha_{N-i}=0}^{\alpha_{N-i-1}} C^{k}_{\alpha_1} C^{\alpha_1}_{\alpha_2} \cdots C^{\alpha_{N-i-1}}_{\alpha_{N-i}} \mathrm{ad}^{k-\alpha_1}_{-X_N} \mathrm{ad}^{\alpha_1-\alpha_2}_{-X_{N-1}} \cdots \mathrm{ad}^{\alpha_{N-i-1}-\alpha_{N-i}}_{-X_{i+1}} \mathrm{ad}^{\alpha_{N-i}}_{-X_i}(X_{i-1}).$$

Since $T_N^{(k)}(0) = \sum_{i=2}^{N} t_{N,i}^{(k)} + X_N^{(k)}$ we obtain (54) and Lemma 10 is proved.

Remark. It is also possible to prove the estimate

$$|T_N^{(k)}(0)| \leqslant (1/M)((N-1)\, M|X|)^{k+1}, \qquad k \geqslant 1$$

given in Lemma 3 using (54) instead of the estimates (41)′ of Lemma 6. Those estimates are useful for other purposes as well, however.

We rewrite (54) in a more compact form:

$$T_N^{(k)}(0) = X_N^{(k)} + \sum_{i=2}^{N-1} \sum_{\alpha_1=0}^{\alpha_0=k} \sum_{\alpha_2=0}^{\alpha_1} \cdots \sum_{\alpha_{N-i}=0}^{\alpha_{N-i-1}} \left(\prod_{p=0}^{N-i-1} C^{\alpha_p}_{\alpha_{p+1}} \right) \left(\prod_{p=0}^{N-i} \mathrm{ad}^{\alpha_p - \alpha_{p+1}}_{-X_{N-p}} (X_{i-1}) \right) + \mathrm{ad}^{k}_{-X_N}(X_{N-1}),$$

$k \geqslant 0$, where $\alpha_{N-i+1} = 0$ if $i > N$, and for $k = 0$

$$T_N^{(0)}(0) = T_N(0) = \sum_{i=1}^{N} X_i.$$

Formulas (53) and (54) then give an exact expression for the coefficients c_n^N. The formula is quite complicated, but fortunately in the main application to the continuous case it can be written as the sum of two parts, one which has a simple geometric interpretation, and one which tends to zero. It is curious that if one returns to the discrete case from the continuous case, it is possible to give a geometric interpretation to both parts. We indicate how this can be done later.

Formula (54) can be used in principle to compute $T_N^{(2)}(0)$ and thus c_3^N. The calculations turn out to be complicated, and as an exact expression for c_3^N is needed later, we calculate it directly here.

First note that

$$t_{N,i}^{(2)}(0)=2\sum_{N\geqslant p>q>i-1}\operatorname{ad}_{-X_p}\operatorname{ad}_{-X_q}(X_{i-1})+\sum_{N\geqslant M>i-1}\operatorname{ad}_{-X_m}\operatorname{ad}_{-X_m}(X_{i-1}).$$

This formula may be obtained in the following way. $t_{N,i}(0)=X_{i-1}$ for $N\geqslant i$ because $t_{N,i}(t)=e^{-tX_N}\cdots e^{-tX_i}X_{i-1}e^{tX_i}\cdots e^{tX_N}$. Also $t_{ii}(t)=e^{-tX_i}X_{i-1}e^{tX_i}$ and $t_{ii}^{(k)}(0)=\operatorname{ad}^k_{-X_i}(X_{i-1})$ (see Lemma 4), for $k\geqslant 0$.

(a)
$$\begin{aligned}t_{i+1,i}^{(2)}(0)&=\operatorname{ad}^2_{-X_{i+1}}t_{ii}(0)+2\operatorname{ad}_{-X_{i+1}}t_{ii}^{(1)}(0)+t_{ii}^{(2)}(0)\\&=\operatorname{ad}^2_{-X_{i+1}}(X_{i-1})\\&\quad+2\operatorname{ad}_{-X_{i+1}}\operatorname{ad}_{-X_i}(X_{i-1})+\operatorname{ad}^2_{-X_i}(X_{i-1})\\&=(\operatorname{ad}^2_{-X_{i+1}}+2\operatorname{ad}_{-X_{i+1}}\operatorname{ad}_{-X_i}+\operatorname{ad}^2_{-X_i})(X_{i-1}).\end{aligned}$$

(b)
$$\begin{aligned}t_{i+2,i}^{(2)}(0)&=\operatorname{ad}^2_{-X_{i+2}}t_{i+1,i}(0)+2\operatorname{ad}_{-X_{i+2}}t_{i+1,i}^{(1)}(0)+t_{i+1,i}^{(2)}(0)\\&=\operatorname{ad}^2_{-X_{i+2}}(X_{i-1})+2\operatorname{ad}_{-X_{i+2}}(\operatorname{ad}_{-X_{i+1}}t_{ii}(0)+t_{ii}^{(1)}(0))\\&\quad+(\operatorname{ad}^2_{-X_{i+1}}+2\operatorname{ad}_{-X_{i+1}}\operatorname{ad}_{-X_i}+\operatorname{ad}^2_{-X_i})(X_{i-1})\\&=(\operatorname{ad}^2_{-X_{i+2}}+\operatorname{ad}^2_{-X_{i+1}}+2\operatorname{ad}_{-X_{i+1}}\operatorname{ad}_{-X_i}+\operatorname{ad}^2_{-X_i})(X_{i-1})\\&\quad+2\operatorname{ad}_{-X_{i+2}}(\operatorname{ad}_{-X_{i+1}}X_{i-1}+\operatorname{ad}_{-X_i}X_{i-1})\\&=(\operatorname{ad}^2_{-X_{i+2}}+\operatorname{ad}^2_{-X_{i+1}}+\operatorname{ad}^2_{-X_i}+2\operatorname{ad}_{-X_{i+2}}\operatorname{ad}_{-X_{i+1}}\\&\quad+2\operatorname{ad}_{-X_{i+2}}\operatorname{ad}_{-X_i}+2\operatorname{ad}_{-X_{i+1}}\operatorname{ad}_{-X_i})(X_{i-1}).\end{aligned}$$

Evidently this process can be iterated and finally we obtain

$$t_{N,i}^{(2)}(0)=2\sum_{N\geqslant p>q>i-1}\operatorname{ad}_{-X_p}\operatorname{ad}_{-X_q}(X_{i-1})+\sum_{N\geqslant m>i-1}\operatorname{ad}_{-X_m}\operatorname{ad}_{-X_m}(X_{i-1}).\tag{54$'$}$$

Hence

$$\begin{aligned}T_N^{(2)}(0)&=\sum_{i=2}^{N}t_{N,i}^{(2)}(0)\\&=2\sum_{N\geqslant p>q>j\geqslant 1}\operatorname{ad}_{-X_p}\operatorname{ad}_{-X_q}(X_j)\\&\quad+\sum_{N\geqslant m>j\geqslant 1}\operatorname{ad}_{-X_m}\operatorname{ad}_{-X_m}(X_j),\end{aligned}\tag{54$''$}$$

or

$$T_N^{(2)}(0) = 2 \sum_{N \geqslant p > q > j \geqslant 1} [X_p, [X_q, X_j]] + \sum_{N \geqslant m > j \geqslant 1} [X_m, [X_m, X_j]] \tag{54'''}$$

and from (53) it follows that

$$c_3^N = \frac{1}{6} T_N^{(2)}(0) + \frac{1}{12}\left[X_1 + \cdots + X_N, \sum_{1 \leqslant i < j \leqslant N} [X_i, X_j]\right].$$

We summarize the results thus far. Let $B_N(X_1, X_2, ..., X_N) = \sum_{n=1}^{\infty} c_n^N(X_1, ..., X_N)$, where the coefficients c_n^N are the homogeneous polynomials of degree n in $X_1, ..., X_N$ such that

$$e^{X_1} \cdots e^{X_N} = e^{B_N(X_1, ..., X_N)}$$

formally, as the expansion need not converge for arbitrary $X_1, ..., X_N \in L$. Then consider $\psi_N(t) = \sum_{n=1}^{\infty} c_n^N t^n$, which converges in a neighborhood of the origin. It is shown to satisfy the differential equation

$$\frac{d\psi_N}{dt} = k(\mathrm{ad}_{\psi_N})(T_N) + \tfrac{1}{2}[\psi_N, T_N],$$

where $T_N(t) = \sum_{i=2}^{N} e^{-tX_N} \cdots e^{-tX_i} X_{i-1} e^{tX_i} \cdots e^{tX_N} + X_N$ and where $k(z) = 1 + \sum_{p=1}^{\infty} k_{2p} z^{2p}$, $k_{2p}(2p)! = b_{2p}$ are Bernoulli's numbers.

Step 1. Consider the differential operator $U_{a_1, ..., a_n}$ of degree n with coefficients in L and which depends on the parameters $a_i \in L$ defined by

$$U_{a_1, ..., a_n} = \sum_{r=0}^{m} \frac{1}{(n-r)!\,(n+1)} \times \left(\frac{1}{2}\mathrm{ad}_{a_r} + \sum_{\substack{p \geqslant 1 \\ 2p \leqslant r}} k_{2p} \sum_{\substack{m_i > 0 \\ m_1 + \cdots + m_{2p} = r}} \mathrm{ad}_{a_{m_1}} \cdots \mathrm{ad}_{a_{m_{2p}}}\right) \circ \frac{d^{n-r}}{dt^{n-r}}. \tag{55}$$

Note that this operator does not depend on N and includes all powers of d/dt less than or equal to n. This operator acts on L-valued functions of t, and produces L-valued functions of t.

Step 2. Show that the coefficients satisfy the recursion formulas

$$c_1^N = T_N(0)$$

$$c_{n+1}^N = U_{c_1^N, ..., c_n^N}(T_N)_{t=0}, \qquad n \geqslant 1.$$

Step 3. Use the recursion formulas of Step 2 and the estimate $|T_N^{(k)}(0)| \leqslant (1/M)(M|X|)^{k+1}$, $k \geqslant 0$, to obtain that

$$|c_{n+1}^N| \leqslant \frac{(M|X|)^{n+1}}{M} \rho_{n+1},$$

where $h(z) = \sum_{n=1}^{\infty} \rho_n z^n$ satisfies the differential equation

$$\frac{dh(z)}{dz} = e^z \left(q(h) + \frac{h}{2} \right),$$

and thus that the expansion converges uniformly in N if $|X_1| + \cdots + |X_N| < \hat{b}/M$, $\hat{b}$ the radius of convergence of $h(z)$.

6. The Logarithm Functional and the Definition of the Integral

Recall that for simplicity we are considering the elements of L as being represented by finite dimensional matrices. Assume that $\gamma = \{K(s), a \leqslant s \leqslant b\}$ is a Riemann integrable curve in L in the sense that each matrix element has that property. This implies that

$$\int_a^b K(s)\, ds = \lim_{N \to \infty} \sum_{i=1}^{N} K(s_i^*)\, \Delta s_i \tag{56}$$

for any partition $P = \{a = s_0 < \cdots < s_N = b\}$ of (a, b) whose mesh size tends to zero as $N \to \infty$, where $\Delta s_i = (s_i - s_{i-1})$ and $s_{i-1} \leqslant s_i^* \leqslant s_i$. Assume as part of the Riemann integrability condition for the curve γ that

$$\int_a^b |K(s)|\, ds < \infty.$$

The elements $K(s_1^*)\, \Delta s_1, ..., K(s_N^*)\, \Delta s_N \in L$. Let $X_1 = K(s_N^*)\, \Delta s_N, ..., X_N = K(s_1^*)\, \Delta s_1$, noting that $X_1, ..., X_N$ is indexed in reverse order from $K(s_1^*)\, \Delta s_1^*, ..., K(s_N)\, \Delta s_N$. The reason for the inverse order is that it is the order in which we want to consider the operators as acting, i.e., $e^{K(s_1^*)\, \Delta s_1}$ first, $e^{K(s_2^*)\, \Delta s_2}$ second, etc., so that the product of interest is

$$e^{K(s_N^*) \Delta s_N} \cdots e^{K(s_1^*) \Delta s_1} = e^{B_N(X_1, ..., X_N)}$$

with $X_1 = K(s_N^*)\, \Delta s_N, ..., X_N = K(s_1^*)\, \Delta s_1$. According to Theorem 3, the infinite expansion denoted by $B_N(X_1, ..., X_N)$ converges absolutely uniformly in N if $|X_1| + \cdots + |X_N| < \delta$, or what is the same thing, if $\sum_{i=1}^N |K(s_i^*)|\, \Delta s_i < \delta$. This last condition holds for a sufficiently small partition if $\int_a^b |K(s)|\, ds < \delta$.

THEOREM 5. *Assume that* $\gamma = \{K(s),\ a \leqslant s \leqslant b\}$ *is Riemann integrable, and that*

$$\int_a^b |K(s)|\, ds < \delta = \frac{\hat{b}}{M}. \tag{57}$$

Then

$$H(\gamma) = B_\infty(\gamma) = \lim_{N \to \infty} B_N(X_1, ..., X_N)$$

exists, where $X_i = K(s^*_{N-i+1})\, \Delta s_{N-i+1}$. *The Lie integral, denoted by* $L \int_a^b K(s)\, ds$, *is defined as* $H(\gamma)$.

Proof of Theorem 5. It has already been indicated that Theorem 3 yields absolute uniform convergence of the series $B_N(X_1, ..., X_N)$. Although the series converges absolutely uniformly, in order to prove that $\lim_{N \to \infty} B_N(X_1, ..., X_N)$ exists, it is necessary to prove more, that the coefficients $c_n^N(X_1, ..., X_N)$ converge as $N \to \infty$.

Formula (28) yields

$$\begin{aligned}
&\lim_{N \to \infty} (n+1)\, c_{n+1}^N(X_1, ..., X_N) \\
&\quad = \frac{1}{n!} \lim_{N \to \infty} T_N^{(n)}(0, X_1, ..., X_N) + \sum_{r=1}^{n} \frac{1}{(n-r)!} \\
&\quad \times \Bigg\{ \frac{1}{2} \Big[\lim_{N \to \infty} c_r^N(X_1, ..., X_N),\ \lim_{N \to \infty} T_N^{(n-r)}(0, X_1, ..., X_N) \Big] \\
&\quad + \sum_{\substack{p \geqslant 1 \\ 2p \leqslant r}} k_{2p} \sum_{\substack{m_i > 0 \\ m_1 + \cdots + m_{2p} = r}} [\lim_{N \to \infty} c_{m_1}^N(X_1, ..., X_N)], \\
&\quad [\cdots [\lim_{N \to \infty} c_{m_{2p}}^N(X_1, ..., X_N),\ \lim_{N \to \infty} T_N^{(n-r)}(0, X_1, ..., X_N)] \cdots] \Bigg\}.
\end{aligned}$$

In the next section, in Theorem 7, it is shown that

$$\lim_{N \to \infty} T_N^{(k)}(0) = T_\infty^{(k)}(0)$$

so that we can prove that $\lim_{N \to \infty} c_n^N$ exists using induction on n. It only remains to show that the limit exists for $n = 1$. In that case, however, (29) implies that $\lim_{N \to \infty} c_1^N = \lim_{N \to \infty} T_N(0, K_N\, \Delta s_N, ..., K_1\, \Delta s_1) = \lim_{N \to \infty} \sum_{i=1}^N K(s_i^*)\, \Delta s_i = \int_a^b K(s)\, ds$, and Theorem 5 is proved.

DEFINITION. The curve $\gamma_s = \{K(u),\ a \leqslant u \leqslant s\}$ is the curve whose initial point is $K(a)$ and whose final point is $K(s)$. The curve $t\gamma_s = \{tK(u),$

$a \leqslant u \leqslant s\}$ is the curve whose initial point is $tK(a)$ and whose final point is $tK(b)$.

Note that since

$$\int_a^b |tK(s)|\, ds \leqslant |t| \int_a^b |K(s)|\, ds$$

all integral are such that Theorem 5 applies if $|t| \leqslant 1$.

THEOREM 6. *Let $\gamma = \{K(s), a \leqslant s \leqslant b\}$ be any Riemann integrable curve such that $\int_a^b |K(s)|\, ds < \delta$. Then $\lim_{N \to \infty} T_N(t)$ exists, and*

$$\lim_{N \to \infty} T_N(t) = T_\infty(t, \gamma) = \int_a^b e^{-H(t\gamma_s)} K(s)\, e^{H(t\gamma_s)}\, ds. \tag{58}$$

Proof of Theorem 6. Consider the function

$$T_N(t, X_1, ..., X_N) = \sum_{i=1}^{N} e^{-tX_N} \cdots e^{-tX_{i+1}} X_i e^{tX_{i+1}} \cdots e^{tX_N},$$

where we define $X_{N+1} = 0$, and $X_i = K(s^*_{N-i+1})\, \Delta s_{N-i+1}$. We can write

$$T_N = \sum_{i=1}^{N} e^{B_{N-i}(-tX_N, ..., -tX_{i+1})} X_i e^{B_{N-i}(tX_{i+1}, ..., tX_N)}$$

and as

$$e^{B_{N-i}(-tX_N, ..., -tX_{i+1})} e^{B_{N-i}(tX_{i+1}, ..., tX_N)} = \text{identity},$$

we have

$$B_{N-i}(-tX_N, ..., -tX_{i+1}) = -B_{N-i}(tX_{i+1}, ..., tX_N)$$

and thus

$$T_N = \sum_{i=1}^{N} e^{-B_{N-i}(tX_N, ..., tX_{i+1})} X_i e^{B_{N-i}(tX_{i+1}, ..., tX_N)}.$$

This implies that

$$\begin{aligned} T_N(t, K_1\, \Delta s_1, ..., K_N\, \Delta s_N) &= \sum_{i=1}^{N} e^{-B_{N-i}(tK_1\, \Delta s_1, ..., tK_{N-i}\, \Delta s_{N-i})} \\ &\quad \cdot K(s_i)\, e^{B_{N-i}(tK_1\, \Delta s_1, ..., tK_{N-1}\, \Delta s_{N-i})} \cdot \Delta s_i. \end{aligned}$$

We have from Theorem 5 that

$$B_{N-i}(tK_1\,\Delta s_1, \ldots, tK_{N-i}\,\Delta s_{N-i}) \to H(t\gamma_s) \tag{58$'$}$$

for a given s. It follows from Theorem 3 that the expansion for B_{N-i} converges uniformly in t and s. The proof of Theorem 7 yields that the coefficients of B_{N-i} converge uniformly in s and t, and thus the convergence in (58)$'$ is uniform in s and t. We therefore have that the following limit exists:

$$\lim_{N\to\infty} T_N(t) = \int_a^b e^{-H(t\gamma_s)} K(s)\, e^{H(t\gamma_s)} ds.$$

The limit $\lim_{N\to\infty} T_N(t)$ is analytic because the functions $T_N(t)$ are and the convergence is uniform. This also implies that $\lim_{N\to\infty} T_N^{(k)}(t) = T_\infty^{(k)}(t)$.

COROLLARY 1.

$$\lim_{N\to\infty} T_N^{(k)}(t) = \frac{d^k}{dt^k} T_\infty(t, \gamma)$$

$$= \int_b^a \frac{d^k}{dt^k} (e^{-H(t\gamma_s)} K(s)\, e^{H(t\gamma_s)})\, ds. \tag{59}$$

Proof. The proof of the first equality has already been given at the end of the proof of Theorem 6. The second equality follows from Theorem 6 by differentiation under the integral sign. That this is possible follows from uniform analyticity in t of the integrand. This uniform analyticity follows from the remarks at the end of the proof of Theorem 6.

The limit $\lim_{N\to\infty} c_n^N(K_n\,\Delta s_N, \ldots, K_1\,\Delta s_1)$ will be denoted by $c_n^\infty(K(s))$ or by $H_n(K(s))$. It has been shown to exist provided that $\int_a^b |K(s)|\, ds$ is sufficiently small, and further, that $H(K(s)) = \sum_{n=0}^\infty H_n(K(s))$ converges, under the same assumption.

COROLLARY 2. *Recursion formulas for* H_n: $H_1 = \int_{s=a}^b K(s)\, ds$ *and for all* $n \geq 1$:

$$(n+1)\, H_{n+1} = \frac{1}{n!} T_\infty^{(n)}(0) + \sum_{r=1}^n \frac{1}{(n-r)!} \Bigg[\frac{1}{2} [H_r, T_\infty^{(n-r)}(0)] + \sum_{\substack{p\geq 1\\ 2p\leq r}} k_{2p} \sum_{\substack{m_i>0\\ m_1+\cdots+m_{2p}=r}} [H_{m_1}, [\ldots, [H_{m_{2p}}, T_\infty^{(n-r)}(0)]\cdots \Bigg]. \tag{60}$$

We now calculate $T_\infty^{(k)}(0, \gamma)$ for $k=0$ and $k=1$. For $k=0$ we have from (58) that

$$T_\infty(0, \gamma) = H_1(\gamma) = \int_a^b K(s)\, ds. \tag{61}$$

For $k=1$ we have from (51) that

$$\begin{aligned} T_N^{(1)}(0) &= \sum_{1 \leqslant i < j \leqslant N} [X_i, X_j] \\ &= \sum_{1 \leqslant i \leqslant N} \left[X_i, \sum_{i<j} X_j \right] \\ &= \sum_{1 \leqslant i \leqslant N} \left[K(s_i), \left(\sum_{i>j} K(s_j)\, \Delta s_j \right) \right] \Delta s_i. \end{aligned}$$

As the mesh size of the partition tends to zero, this double sum tends to the integral

$$T_\infty^{(1)}(0) = \int_{s=a}^b \left[K(s), \left(\int_{t=a}^b K(t)\, dt \right) \right] ds. \tag{62}$$

The integral $\int_{t=a}^s K(t)\, dt$ is an element of the Lie algebra L, and therefore its commutator with $K(s)$ is well defined. It is possible to obtain a similar formula for $T_\infty^{(2)}(0)$ using (54), etc.

It is also interesting to obtain formulas for the first few terms H_n in the expansion $H(K(s)) = \sum_{n=1}^\infty H_n(K(s))$. It follows from (60), (61), and (62) that

$$\begin{aligned} H_1(K(s)) &= \int_{s=a}^b K(s)\, ds; \\ H_2(K(s)) &= \frac{1}{2} T_\infty^{(1)}(0) \qquad \text{(from (51))} \\ &= \frac{1}{2} \int_{s=a}^b \left[K(s), \int_{t=a}^s K(t)\, dt \right] ds. \end{aligned} \tag{63}$$

COROLLARY 3. *An estimate of the norm* $H(K(s))$: *we have that* $|(K(s))| \leqslant (1/M)\, h(z_0)$, *where* h *is the solution of the differential equation* $(dh/dz) = e^z(q(h) + \frac{1}{2}h)$ *with* $h(0) = 0$ *and* $z_0 = M \int_a^b |K(s)|\, ds$ (*we assume* $\int_a^b |K(s)|\, ds < \delta$).

Proof. Follows at once from the proof of Theorem 3.

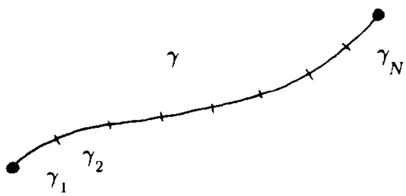

FIGURE 1

COROLLARY 4. *Suppose that* $\gamma = \gamma_1 \cup \cdots \cup \gamma_N$, *i.e., that* γ *is the union of* N *smaller curves (see Fig. 1). Then there exists* $\delta > 0$ *such that if* $\int_a^b |K(s)|\, ds < \delta$ *then*

$$H(\gamma) = B_N(H(\gamma_N), ..., H(\gamma_1))$$

provided the parameter of γ *traces* γ_1 *first,* γ_2 *second, etc.*

COROLLARY 5. $H(\gamma)$ *is independent of parameterizations which preserve order. Consider the curve* $K(s)$ *in* L, $a \leqslant s \leqslant b$, *and the family of curves* $tK(s)$, *where* $0 \leqslant t \leqslant \infty$. *We have defined the function*

$$T(t) = T_{\infty}(t) = \lim_{N \to \infty} T_N(t)$$

and the integral

$$L \int_{s=a}^{b} tK(s)\, ds = H(t).$$

Then the derivative $H'(t) = dH(t)/dt \in L$ *is well defined (subject to the usual restrictions). The value of the derivative* $H'(t)$ *for each* t *is an element in* L.

COROLLARY 6. *For all* t *for which the functions are defined we have*

$$\left(\frac{1 - e^{-\mathrm{ad}_H}}{\mathrm{ad}_H}\right) H' = T, \tag{64a}$$

i.e.,

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{(n+1)!} (\mathrm{ad}_H)^n H' = T. \tag{64b}$$

Proof. It has been shown that for all N,

$$\frac{d\psi_N}{dt} = \left(k\left(\mathrm{ad}_{\psi_N}\right) + \frac{1}{2}\,\mathrm{ad}_{\psi_N}\right) T_N,$$

where $k(\mathrm{ad}_X)+\frac{1}{2}\mathrm{ad}_X=(m(\mathrm{ad}_X))^{-1}$. Then $m(\mathrm{ad}_{\psi_N})\,\psi'_N=T_N$ and as $N\to\infty$, we obtain $m(\mathrm{ad}_H)\,H'=T$, where

$$m(z)=\sum_{n=0}^{\infty}\frac{(-1)^n}{(n+1)!}z^n=\frac{1-e^{-z}}{z}.$$

The corollary is proved.

In an arbitrary finite-dimensional Lie algebra ad_H is a matrix. Consequently, all expressions $(\mathrm{ad}_H)^n$ are matrices. An equivalent form of (64a) is

$$H'=\frac{\mathrm{ad}_H}{1-e^{-\mathrm{ad}_H}}(T).$$

7. The Main Result

In this section we find an explicit formula for the solution of the equation

$$\frac{dX}{dt}=K(t)\,X(t)$$

in the case that K and X are matrices. The solution is of the form

$$X(t)=e^{L\int_0^t K(s)\,ds}X(0),$$

where $L\int_0^t K(s)\,ds$ is the new type of integral defined in Theorem 5. It may be regarded as the logarithm of the curve in L traced by the function X which is the solution of the differential equation.

The Calculation of $T_\infty^{(k)}(0)=\lim_{N\to\infty}T_N^{(k)}(0)$

We shall make use of the formulas derived in this section in the proofs of Theorems 5 and 6 of the previous section, and so no results which depend on those theorems can be used in the calculations of this section. We assume that

$$\int_a^b |K(s)|\,ds<\delta$$

and consider a partition $P=\{a=s_0<\cdots<s_N=b\}$ of (a,b) of mesh size Δ, i.e., $\Delta=\max_{1\leqslant i\leqslant N}\Delta_i$, $\Delta_i=(s_i-s_{i-1})$. Let $X_i=K(s^*_{N-i+1})\,\Delta_{N-i+1}$.

We start from formula (54) and first change it to a more convenient form. Note that $C^k_{\alpha_1} C^{\alpha}_{\alpha_2} \cdots C^{\alpha_{N-i-1}}_{\alpha_{N-i}}$ is equal to

$$\frac{k!}{\alpha_1!\,(k-\alpha_1)!}\cdot\frac{\alpha_1!}{\alpha_2!\,(\alpha_1-\alpha_2)!}\cdots\frac{\alpha_{N-i-1}!}{\alpha_{N-i}!\,(\alpha_{N-i-1}-\alpha_{N-i})!}$$
$$=\frac{k!}{(k-\alpha_1)!\,(\alpha_1-\alpha_2)!\cdots(\alpha_{N-i-1}-\alpha_{N-i})!\,\alpha_{N-i}!}.$$

Consequently, from (54) we have, for all $k \geqslant 0$,

$$T^{(k)}_N(0) = X^{(k)}_N + \sum_{i=2}^{N} \sum_{\alpha_1=0}^{\alpha_0=k} \sum_{\alpha_2=0}^{\alpha_1} \cdots$$
$$\cdots \sum_{\alpha_{N-i}=0}^{\alpha_{N-i-1}} \frac{k!}{(k-\alpha_1)!\,(\alpha_1-\alpha_2)!\cdots(\alpha_{N-i-1}-\alpha_{N-i})!\,\alpha_{N-i}!}$$
$$\times \operatorname{ad}^{k-\alpha_1}_{-X_N} \operatorname{ad}^{\alpha_1-\alpha_2}_{-X_{N-1}} \cdots \operatorname{ad}^{\alpha_{N-i-1}-\alpha_{N-i}}_{-X_{i+1}} \operatorname{ad}^{\alpha_{N-i}}_{-X_i}(X_{i-1}). \tag{65}$$

The limit as $N \to \infty$ exists in (65) and the following theorem is obtained when the limit is evaluated.

THEOREM 7. *Let $K(s)$, $a \leqslant s \leqslant b$, be a Riemann integrable curve with values in L such that $\int_a^b |K(s)|\, ds < \delta$. Then the limit may be written in the following equivalent ways for $k \geqslant 0$:*

(1) $$T^{(k)}_{\infty}(0) = \lim_{N\to\infty} T^{(k)}_N(0)$$
$$= (-1)^k k! \int\cdots\int_{a \leqslant u_1 \leqslant \cdots \leqslant u_{k+1} \leqslant b} [K(u_1), [K(u_2),$$
$$\ldots, [K(u_k), K(u_{k+1})] \cdots]\, du_1\, du_2 \cdots du_{k+1}, \tag{66}$$

(2) $$T^{(k)}_{\infty}(0) = k! \int_{u_1=a}^{b} \left[\cdots \left[K(u_1), \int_{u_2=a}^{u_1} K(u_2) \right],\right.$$
$$\left. \cdots \int_{u_k=a}^{u_{k-1}} K(u_k) \right], \int_{u_{k+1}=a}^{u_k} K(u_{k+1}) \Bigg]\, du_{k+1}\, du_k \cdots du_2\, du_1, \tag{66$'$}$$

(3) $$T^{(k)}_{\infty}(0) = k! \int\cdots\int_{b \geqslant u_1 \geqslant \cdots \geqslant u_{k+1} \geqslant a} [\cdots[K(u_1), K(u_2)],$$
$$\ldots, K(u_k)], K(u_{k+1})] \cdots]\, du_{k+1} \cdots du_2\, du_1. \tag{66$''$}$$

Proof of Theorem 7. First note that $(1) \Rightarrow (3)$ if the commutators are interchanged in pairs starting from the left, and if the names of the variables are interchanged. This follows as there are k interchanges and each interchange in a commutator yields a factor of (-1). Also $(3) \Rightarrow (2)$ as (2) is just a rewriting of (3). This means that the proof is reduced to proving case (1).

In order to make the proof more transparent, we consider three special cases first, and then prove the theorem in full.

(a) If $k = 0$, then

$$T_N^{(0)}(0) = X_N^{(0)} + \sum_{i=2}^{N} X_{i-1} \qquad \text{(from (54))}$$
$$= X_1 + X_2 + \cdots + X_N$$

and it is clear that

$$\lim_{\Delta \to 0} \sum_{i=1}^{N} X_{N-i+1} = \int_a^b K(s)\, ds.$$

(b) If $k = 1$ then, recalling the inverse ordering,

$$T_N^{(1)}(0) = -\sum_{N \geqslant j > i} -[X_j, X_i]$$
$$= -\sum_i \left[X_i, \sum_{N \geqslant j > i} X_j \right] \to (\text{as } \Delta \to 0)$$
$$\to -\int_a^b \int_{u_1 = a}^{u_2} [K(u_1), K(u_2)]\, du_1\, du_2.$$

(c) If $k = 2$, then $T_N^{(2)}(0) = A + B$, where

$$A = 2 \sum_{N \geqslant p > q > j \geqslant 1} [X_p, [X_q, X_j]],$$
$$B = \sum_{N \geqslant m > j > 1} [X_m, [X_m, X_j]].$$

It is clear, again using the inverse ordering, that

$$\lim_{\Delta \to 0} A = 2 \int_{u_3 = a}^{b} \int_{u_2 = a}^{u_3} \int_{u_1 = a}^{u_2} [K(u_1), [K(u_2), K(u_3)]]\, du_1\, du_2\, du_3,$$

so that it only remains to show that $\lim_{\Delta \to 0} B = 0$. Also

$$|B| \leqslant \sum_{N \geqslant m \geqslant n \geqslant j \geqslant 1} \psi(m, n, j) |[X_m, [X_n, X_j]]| = c,$$

where $\psi(m, n, j) = 1$ if $m = n$ and zero otherwise. We have that

$$\lim_{\Delta \to 0} c = \iiint_{a \leqslant u_1 \leqslant u_2 \leqslant u_3 \leqslant b} \psi(u_1, u_2, u_3)$$
$$\times |[K(u_1), K(u_2)], K(u_3)]| \, du_1 \, du_2 \, du_3 = 0$$

and so the proof is completed for this special case, as $\psi \neq 0$ only if $u_1 = u_2$.

The proof of the general case is based on formula (65). Each term on the right side of (65)is a monomial of degree $(k+1)$ and it can be written in the form

$$\tilde{M}[X_{i_1}, [X_{i_2}, [\cdots [X_{i_k}, X_{i_{k+1}}] \cdots], \tag{67}$$

where $\tilde{M}$ depends only on $i_1, ..., i_{k+1}$ and where $N \geqslant i_1 \geqslant i_2 \geqslant \cdots \geqslant i_{k+1} \geqslant 1$. These sets of indices $(i_1, i_2, ..., i_{k+1})$ may be divided into groups,

$$\{(i_1, ..., i_{k+1}): N \geqslant i_1 \geqslant \cdots \geqslant i_{k+1} \geqslant 1\}$$
$$= \{(i_1, ..., i_{k+1}): N \geqslant i_1 > i_2 \geqslant \cdots \geqslant i_{k+1} \geqslant 1\}$$
$$\cup \{(i_1, ..., i_{k+1}): N \geqslant i_1 = i_2 \geqslant \cdots \geqslant i_{k+1} \geqslant 1\},$$
$$\{(i_1, ..., i_{k+1}): N \geqslant i_1 \geqslant i_2 \geqslant i_3 \geqslant \cdots \geqslant i_{k+1} \geqslant 1\}$$
$$= \{(i_1, ..., i_{k+1}): N \geqslant i_1 > i_2 > i_3 \geqslant \cdots \geqslant i_{k+1} \geqslant 1\}$$
$$\cup \{(i_1, ..., i_{k+1}): N \geqslant i_1 > i_2 = i_3 \geqslant \cdots \geqslant i_{k+1} \geqslant 1\},$$

etc., so that

$$(i_1, ..., i_{k+1}): N \geqslant i_1 \geqslant \cdots \geqslant i_{k+1} \geqslant 1\}$$
$$= \{(i_1, ..., i_{k+1}): N \geqslant i_1 > i_2 > \cdots > i_{i+1} \geqslant 1\}$$
$$\cup \bigcup_{j=1}^{k+2} \{(i_1, ..., i_{k+1}): N \geqslant i_1 > \cdots > i_j = i_{j+1} \geqslant \cdots \geqslant i_{k+1} \geqslant 1\}$$
$$= A + \bigcup_{j=1}^{k+2} B_j. \tag{68}$$

We now write (65) in the form

$$\sum_{(i_1,\dots,i_{k+1})\in A} \tilde{M}(i_1,\dots,i_{k+1})[X_{i_1},[X_{i_2},\dots,[X_{i_k},X_{i_{k+1}}]\cdots]$$
$$+\sum_{j=1}^{k+2}\sum_{(i_1,\dots,i_{k+1})\in B_j} \tilde{M}(i_1,\dots,i_{k+1})[X_{i_1},[X_{i_2},\dots,[X_{i_k},X_{i_{k+1}}]\cdots], \tag{69}$$

and note that $\tilde{M}\leqslant k!$. We have

$$\left|\sum_{(i_1,\dots,i_{k+1})\in B_j} \tilde{M}(i_1,\dots,i_{k+1})[X_{i_1},[X_{i_2},\dots,[X_{i_k},X_{i_{k+1}}]\cdots]\right|$$
$$\leqslant k!\sum_{(i_1,\dots,i_{k+1})\in B_j} |[X_{i_1},[X_{i_2},\dots,[X_{i_k},X_{i_{k+1}}]\cdots]|. \tag{70}$$

Recall that $X_i=K(s^*_{N-i+1})\,\Delta_{N-i+1}$, and writing for simplicity $K(s^*_{N-i+1})=\bar{K}_i$ and $\Delta_{N-i+1}=\bar{\Delta}_i$, we have

$$[X_{i_1},[X_{i_2},\dots,[X_{i_k},X_{i_{k+1}}]\cdots]$$
$$=[\bar{K}_{i_1},[\bar{K}_{i_2},\dots,[\bar{K}_{i_k},\bar{K}_{u_{k+1}}]\cdots]\,\bar{\Delta}_{i_1}\cdots\bar{\Delta}_{i_{k+1}}.$$

Let $C_j=\{(i_1,\dots,i_{k+1})\colon i_j=i_{j+1}\}$ so that $B_j\subset C_j$, and define

$$\psi_j(i_1,\dots,i_{k+1})=\begin{cases}1 & \text{if } \; i_j=i_{j+1}\\ 0 & \text{otherwise.}\end{cases}$$

Equation (70) may be written

$$k!\sum_{(i_1,\dots,i_{k+1})\in B_j} |[X_{i_1},[X_{i_2},\dots,[X_{i_k},X_{i_{k+1}}]\cdots]|$$
$$\leqslant k!\sum_{N\geqslant i_1\geqslant\cdots\geqslant i_{k+1}\geqslant 1} \psi_j(i_1,\dots,i_{k+1})$$
$$\times|[\bar{K}_{i_1},[\bar{K}_{i_2},\dots,[\bar{K}_{i_k},\bar{K}_{i_{k+1}}]\cdots]|\,\bar{\Delta}_{i_1}\cdots\bar{\Delta}_{i_{k+1}},$$

and as $\Delta\to 0$, this expression converges to (recall the inverse ordering)

$$k!\int\cdots\int_{a\leqslant u_1\leqslant\cdots\leqslant u_{k+1}\leqslant b} |[K(u_1),[K(u_2),\dots,[K(u_k),K(u_{k+1})\cdots]|$$
$$\cdot\,\psi_j(u_1,\dots,u_{k+1})\,du_1\,du_2\cdots du_{k+1},$$

where

$$\psi_j(u_1,\dots,u_{k+1})=\begin{cases}1 & \text{if } \; u_{N-j+1}=u_{N-(j-1)+1}\\ 0 & \text{otherwise.}\end{cases}$$

As $\psi_j = 0$ almost everywhere with respect to $(k+1)$ dimensional measure, this implies that (70) converges to zero for all j, $j = 1, ..., k+2$. Thus the problem of evaluating the limit in (65) is reduced to the problem of considering the sum over A in (69).

For these terms the condition $N \geqslant i_1 > i_2 > \cdots > i_{k+1} \geqslant 1$ implies

$$\begin{aligned} &0 \leqslant k - \alpha_1 \leqslant 1, \\ &0 \leqslant \alpha_1 - \alpha_2 \leqslant 1, ..., \\ &0 \leqslant \alpha_{N-i-1} - \alpha_{N-i} \leqslant 1, \\ &0 \leqslant \alpha_{N-i} \leqslant 1 \end{aligned}$$

as an examination of (65) shows, as otherwise at least two of the indices would coincide. For example, if $k - \alpha_1 = 2$, then $\mathrm{ad}^{k-\alpha_1}_{-X_N} = [-X_N, [-X_N, \cdot]]$ and $i_1 = i_2$. It also follows from (65) that the factor $\tilde{M}$ for the terms of A is given by

$$\frac{k!}{(k-\alpha_1)!\,(\alpha_1-\alpha_2)!\,\cdots\,(\alpha_{N-i-1}-\alpha_{N-i})!\,\alpha_{N-i}!} = k!.$$

This implies that

$$\begin{aligned} &\sum_{(i_1, ..., i_{k+1}) \in A} \tilde{M}(i_1, ..., i_{k+1})[X_{i_1}, [X_{i_2}, ..., [X_{i_k}, X_{i_{k+1}}] \cdots]] \\ &= k! \sum_{i=2}^{N} \sum_{\substack{\varepsilon_N, \varepsilon_{N-1}, ..., \varepsilon_i \\ \varepsilon_N + \cdots + \varepsilon_i = k \\ \varepsilon_\alpha = 0 \text{ or } 1}} \mathrm{ad}^{\varepsilon_N}_{-X_N} \mathrm{ad}^{\varepsilon_{N-1}}_{-X_{N-1}} \cdots \mathrm{ad}^{\varepsilon_i}_{-X_i}(X_{i-1}) \\ &= k! \sum_{N \geqslant i_1 > i_2 > \cdots > i_{k+1} \geqslant 1} \mathrm{ad}_{-X_{i_1}} \mathrm{ad}_{-X_{i_2}} \cdots \mathrm{ad}_{-X_{i_k}}(X_{i_{k+1}}) \\ &= k!\,(-1)^k \sum_{N \geqslant i_1 > i_2 > \cdots > i_{k+1} \geqslant 1} [X_{i_1}, [X_{i_2}, ..., [X_{i_k}, X_{i_{k+1}}] \cdots] \\ &\to k!\,(-1)^k \int_{\Delta^{k+1}} \cdots \int [K(u_1), [K(u_2), \\ &\quad ..., [K(u_k), K(u_{k+1})] \cdots]\, du_1 \cdots du_{k+1}, \end{aligned}$$

and the theorem is proved.

Remark. Note that

$$\left| \int_{\Delta^{k+1}} \cdots \int [K(u_1), [K(u_2), ...[K(u_k), K(u_{k+1})] \cdots] \, du_1 \cdots du_{k+1} \right|$$

$$\leqslant \int_{I^{k+1}} \cdots \int | [K(u_1), [K(u_2), ..., [K(u_k), K(u_{k+1})] \cdots]| \, du_1 \cdots du_{k+1}$$

$$\leqslant M^k \int_{I^{k+1}} \cdots \int |K(u_1)| \cdots |K(u_{k+1})| \, du_1 \cdots du_{k+1}$$

$$\leqslant M^k \left[\int_a^b |K(s)| \, ds \right]^{k+1}$$

so that the integrals are finite. That the iterated integrals exist as Riemann integrals follows from the fact that the indefinite integrals are uniformly bounded continuous functions. This last fact is not needed, as the integrals could be considered as Lebesgue integrals.

The Main Theorem

THEOREM 8. *For each s, $a \leqslant s \leqslant b$, assume that $K(s) \in L$ and that $K(s)$ is a matrix. Assume that the Riemann integral of K exists, and that*

$$\int_a^b |K(s)| \, ds < \delta = \hat{b}/M. \tag{71}$$

Then the differential equation

$$\frac{dX}{dt} = K(t) \, X(t) \tag{72}$$

has a solution of the form

$$X(t) = e^{L \int_a^t K(s) \, ds} X(a) \tag{73}$$

which satisfies (72) *at each t for which* $\lim_{h \to 0} (1/h) \int_t^{t+h} K(s) \, ds = K(t)$, *where*

$$L \int_a^t K(s) \, ds = H(K(s), a \leqslant s \leqslant t)$$

$$= H_1[t] + H_2[t] + \cdots = H[t]$$

converges, and $\{H_i[t]\}$ is uniquely defined by the recursion formulas

$$H_1[t] = \int_a^t K(s) \, ds,$$

and for $n \geqslant 1$, *with* $T_0 = H_1$,

$$(n+1) H_{n+1} = T_n + \sum_{r=1}^{n} \left\{ \frac{1}{2} [H_r, T_{n-r}] + \sum_{\substack{p \geqslant 1 \\ 2p \leqslant r}} k_{2p} \sum_{\substack{m_i > 0 \\ m_1 + \cdots + m_{2p} = r}} [H_{m_1}, [\ldots, [H_{m_{2p}}, T_{n-r}] \cdots]] \right\}, \tag{74}$$

where $k_{2p}(2p)!$ *are Bernoulli's numbers, and for* $k \geqslant 1$,

$$T_k[t] = \int_{u_1 = a}^{t} \left[\cdots \left[K(u_1), \int_{u_2 = a}^{u_1} K(u_2) \right], \cdots \int_{u_k = a}^{u_{k-1}} K(u_k) \right], \int_{u_{k+1} = a}^{u_k} K(u_{k+1}) \right] du_{k+1}\, du_k \cdots du_1, \tag{75}$$

or equivalently

$$T_k[t] = T_k = \int_{t \geqslant u_1 \geqslant \cdots \geqslant u_{k+1} \geqslant a} \cdots \int [\cdots [K(u_1), K(u_2)], \ldots, K(u_{k-1})], K(u_k)], K(u_{k+1})]\, du_1 \cdots du_{k+1}. \tag{75$'$}$$

Remark. In formula (75) the region of integration is the closed simplex $\Delta^{k+1} \subset \mathbb{R}^{k+1}$, where

$$\Delta^{k+1} = \{ u \in \mathbb{R}^{k+1},\ u = (u_1, \ldots, u_{k+1}) \text{ and } t \geqslant u_1 \geqslant \cdots \geqslant u_{k+1} \geqslant a \}.$$

(see Fig. 2). Formulas (75) and (75)′ may also be written using right brackets, yielding a *right bracket version.*

Proof of Theorem 8. The proof of Theorem 8 follows from Theorem 7 and formula (60). We have used the notation $T_k = T_k[t] = T_k([a, t]) = (1/k!)\, T_{\infty}^{(k)}(0)$ as the sequence $\{T_k[t]\}$ may be defined by (75). It is necessary, however, to check that

$$X(t) = e^{L \int_0^t K(s)\, ds} X(a) = e^{H[t]} X(a)$$

satisfies (71), or what is the same thing, that

$$\lim_{h \to \infty} \frac{1}{h} [e^{H[t+h]} - e^{H[t]}]\, X(a) = K(t)\, e^{H[t]} X(a).$$

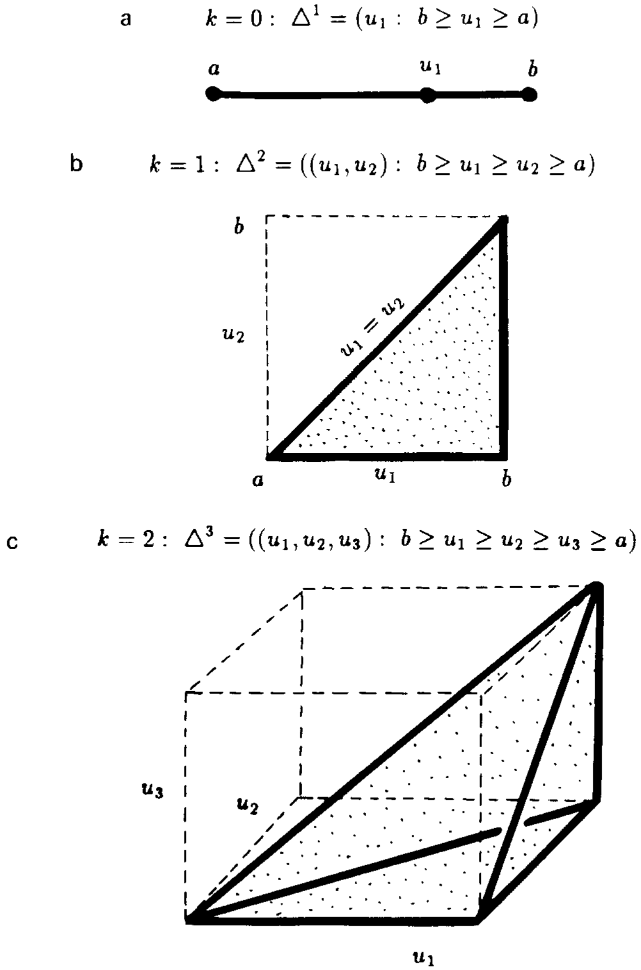

FIGURE 2

Since

$$\frac{1}{h}\left[e^{H[t+h]} - e^{H[t]}\right] = \frac{1}{h}\left[e^{H[t+h]}e^{-H[t]} - I\right] e^{H[t]},$$

the desired result follows from the fact that

$$\lim_{h \to 0} \frac{1}{h}\left[e^{H[t+h]}e^{-H[t]} - I\right]$$

$$= \lim_{h \to 0} \frac{1}{h}\left[e^{B_2(H[t+h], -H[t])} - I\right] = K(t)$$

as may be verified from the uniformity in h of convergence for $H[t+h]$ and thus of B_2.

Remark. It is well known in the theory of ordinary differential equations that if $K(t)$ commutes with $\int_a^t K(s)\,ds$ then the solution of (71) has the form

$$X(t)=e^{\int_a^t K(s)\,ds}X(a),$$

as may easily be checked by direct calculation. The general formulas (72)–(74) immediately yield this result since $[K(t), \int_a^t K(s)\,ds]\equiv 0$ imply $T_k\equiv 0$ for $k\geqslant 1$ (see (74)), so that $H=H_1=\int_0^t K(s)\,ds$.

It is useful to have an estimate for the norm of the solution (73) of Eq. (74). Theorem 9 gives a formal statement of such an estimate. See also Corollary 4 of Theorem 6.

THEOREM 9. *Under the same assumptions as in Theorem* 8 *we have*

$$\text{(a)}\quad \left|L\int_a^t K(s)\,ds\right|\leqslant\frac{1}{M}h\left(M\int_a^t |K(s)|\,ds\right)\qquad \textit{and}$$

$$\text{(b)}\quad |X(t)|=|e^{L\int_0^t K(s)\,ds}|\leqslant e^{(1/M)h(M\int_a^t |K(s)|\,ds)},$$

where M is the constant such that $|[X, Y]|\leqslant M|X|\cdot|Y|$ for all $X, Y\in L$ and $h(z)$ is the solution of the scalar equation

$$\frac{dh}{dz}=e^{z}\left(q(h)+\frac{h}{2}\right),$$

with $h(0)=0$ and $q(h)=1+\sum_{p=1}^{\infty}|k_{2p}|\,h^{2p}$, where $k_{2p}(2p)!$ are Bernoulli's numbers.

LEMMA 11.

$$|T_n|\leqslant(1/M)\left(M\int_a^t |K(s)|\,ds\right)^{n+1}.$$

Proof of Lemma 11. An immediate consequence of the remark following Theorem 7.

Proof of Theorem 9 (see Corollary 4 of Theorem 6 and Step 3 at the end of Section 4). We have that

$$|H_n|\leqslant\left(\frac{1}{M}\right)\left(M\int_a^t |K(s)|\,ds\right)^n\rho_n,$$

using induction, as in the proof of Lemma 8, and we also have that

$$\left| L \int_a^t K(s)\, ds \right| \leqslant \sum_{n=1}^{\infty} |H_n| \leqslant \frac{1}{M} h\left(M \int_a^t |K(s)|\, ds \right),$$

as in the proof of Theorem 3, and (a) is proved. Part (b) follows immediately from (a), and the theorem is proved.

A Generalization to Homogeneous Manifolds

A homogeneous space V is a space of left cosets of a Lie group G modulo a closed subgroup P, i.e., $V = G/P$. Each Lie group is a special case of a homogeneous space, for example, by taking $P = \{I\}$, although there are other more interesting ways of considering Lie groups as homogeneous spaces in general. We consider the case in which G is finite dimensional, so that the theory developed applies without difficulty. Note that G is naturally represented as a transformation T_g, $g \in G$ with $T_g \colon V \to V$ defined by

$$T_g(hP) = ghP.$$

Consider an element $K \in L$, L the Lie algebra of G. We may regard K as defining an infinitely small transformation on V, $T_{I+\varepsilon K}$. Since each vector field on V can be interpreted as an ordinary differential equation,

$$\dot{v} = K(v),$$

where $K(v)$ is the value of the vector field of K at the point $v \in V$, so that $K(v) \in T_v V$. The solutions of the equation $\dot{v}(t) = K(v(t))$ are trajectories on the manifold V.

Assume that $K(s)$ is a smooth curve in L, the Lie algebra of G. Then for each s we obtain a vector field $K(s, v)$ on V, and the following differential equation:

$$\frac{dv(t)}{dt} = K(t, v(t)). \tag{76}$$

THEOREM 10. *Assume that* $\int_a^b |K(s)|\, ds < \delta$. *Then the solution of* (76) *has the form*

$$v(t) = e^{L \int_a^t K(s)\, ds}(v(a)),$$

where $e^{L \int_a^t K(s)\, ds}$ *is, for each* t, *an element of* G *and thus* $e^{L \int_a^t K(s)\, ds}(v(a))$ *denotes*

$$T_{e^{L \int_a^t K(s)\, ds}}(v(a)),$$

the action of the transformation representing $e^{L\int_a^t K(s)\,ds}$ *on* $v(a) \in V$. *The formulas for* $L \int_a^t K(s)\,ds$ *are* (74) *and* (75).

Proof of Theorem 10. Follows at once from Theorem 8, as it is just the representation of G which has changed.

8. The Connection between the Discrete and Continuous Cases of the Main Theorem. Geometrical Meaning of Theorem 7 for Step Functions

In this section we fix our attention on the following natural question: what form does the basic formulas (66), (66)′, (66)″ assume when the function $K(s)$ is a step function? It is intuitively clear that the formulas must reduce to (65), the formulas for the discrete case. It is our goal in this section to establish this. The plan is as follows.

(a) We prove that in the case of a step-function $K(s)$ the continuous formula (67) transforms into the discrete formula (65);

(b) As a consequence, we obtain an interesting geometrical interpretation of the discrete formula (65) and, in particular, of the geometrical meaning of the coefficients $k!/(k-\alpha_1)!\,(\alpha_1-\alpha_2)!\cdots(\alpha_{N-i})!$ in formula (65).

Consider the function $K(s)$, $a \leqslant s \leqslant b$, which has jumps at the points $a < s_1 < s_2 < \cdots < s_l < b$, where $l < \infty$ (Fig. 3) and represent the function $K(s)$ in slightly different drawings (Figs. 4, 5). Then we can consider the subdivision of the region of definition of the function F (namely, the simplex Δ^{k+1}) into the sum of polyhedra θ_α (Fig. 6), which are defined by

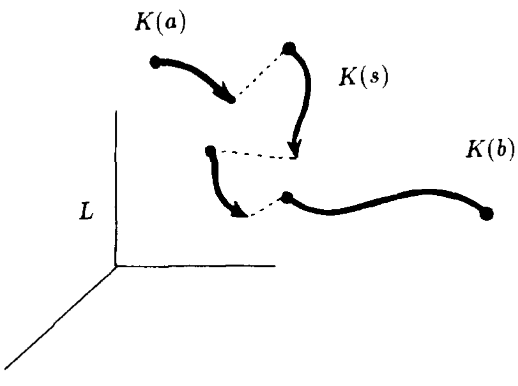

Figure 3

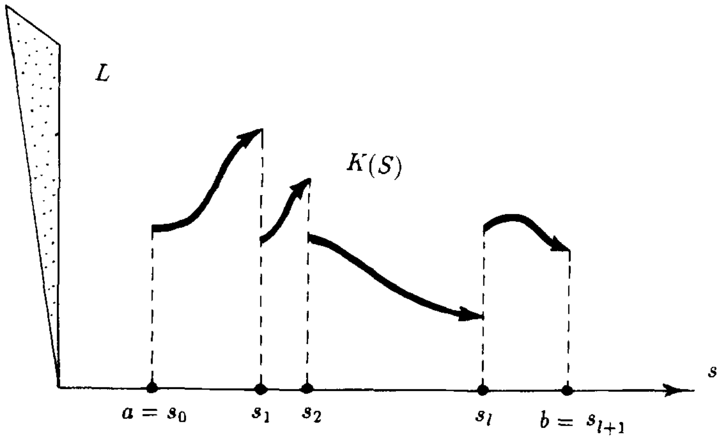

FIGURE 4

hyperplanes corresponding to the points $a < s_1 < \cdots < s_l < b$. The function F is defined as

$$F(u_1, ..., u_{k+1}) = [K(u_1), [K(u_2), ..., [K(u_k), K(u_{k+1})] \cdots]. \quad (77)$$

The union of the polyhedra θ_α is the $(k+1)$ dimensional cube, and the function F defined by (77) is smooth in the interior of θ_α and has jumps on its boundary, Fig. 6 is a picture for $k = 1$ (there $\Delta^{k+1} = \Delta^2$).

Now consider the case when K is a step function, so that K is constant between jump points. In that case, (77) shows that F is constant in the interior of each θ_α and has a jump as the boundary is crossed from one θ_α to

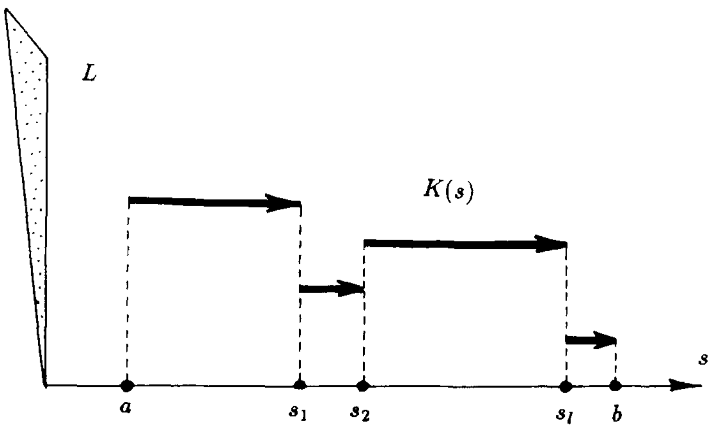

FIGURE 5

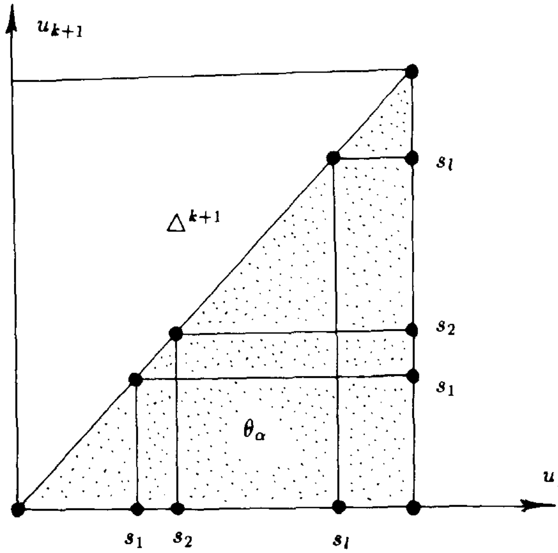

FIGURE 6

an adjacent $\theta_{\alpha'}$. Assume for simplicity that $a = s_0 < s_1 < \cdots < s_l < s_{l+1} = b$ and that $\Delta_i = s_i - s_{i-1}$, $i = 1, ..., l+1$ is independent of i, and in fact assume that $\Delta_i \equiv 1$. It will be clear that the calculations do not depend in an essential way on this assumption. In this case, Fig. 7 represents the subdivision of Δ^{k+1}. Each polyhedron θ_α which is completely contained in Δ^{k+1} is isomorphic to a unit cube, and each polyhedron θ_α which is not completely

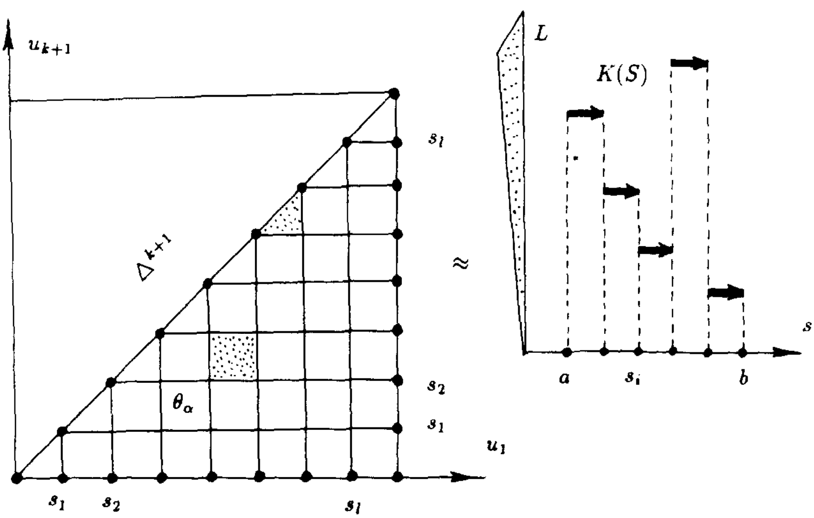

FIGURE 7

contained in Δ^{k+1} is isomorphic to a part of a unit cube, the part which remains after the removal of parts cut off by the hyperplanes which form the boundary of Δ^{k+1}.

THEOREM 11. *If $K(s)$ is a step function as described, then* (66) *transforms to* (65).

Proof of Theorem 11. The function F defined in (77) is constant on each θ_α.

(a) Consider first the case $k=1$, $\Delta^{k+1}=\Delta^2$ (see Fig. 7). Then, assuming K is left continuous,

$$T_\infty^{(1)}(0) = (\text{from } (66)) = \int_{\Delta^2=(b>u_1\geqslant u_2\geqslant a)} [K(u_1), K(u_2)]\, du_1\, du_2$$

$$= \sum_{N\geqslant i>j\geqslant 1} [K_i, K_j] \text{ area } \theta_{ij} + \sum_{N\geqslant i=j\geqslant 1} [K_i, K_j] \left(\frac{1}{2} \text{area } \theta_{ij}\right).$$

That the first term has the factor area θ_{ij} and the second, when $i=j$, has the factor $\frac{1}{2}\theta_{ii}$ may be seen from Figs. 7 and 8. It therefore follows that

$$T_\infty^{(1)}(0) = \sum_{N\geqslant i>j\geqslant 1} [K_i, K_j] + \frac{1}{2} \sum_{N\geqslant i\geqslant 1} [K_i, K_i]$$

$$= \sum_{N\geqslant i>j\geqslant 1} [K_i, K_j],$$

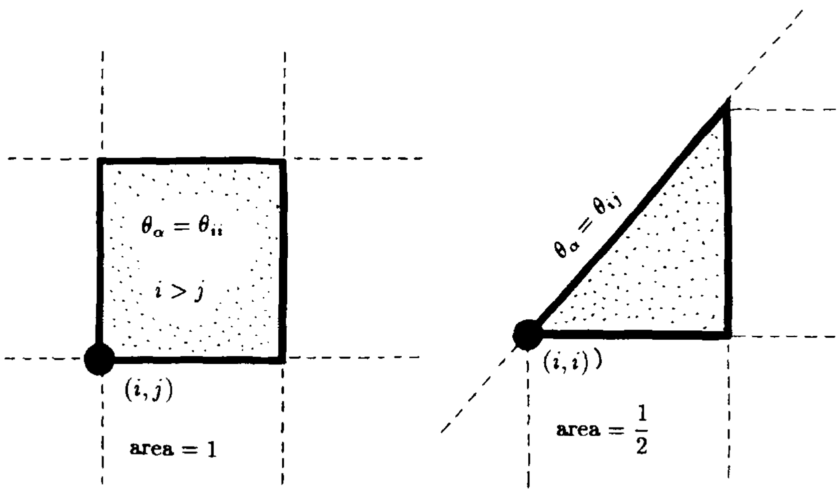

FIGURE 8

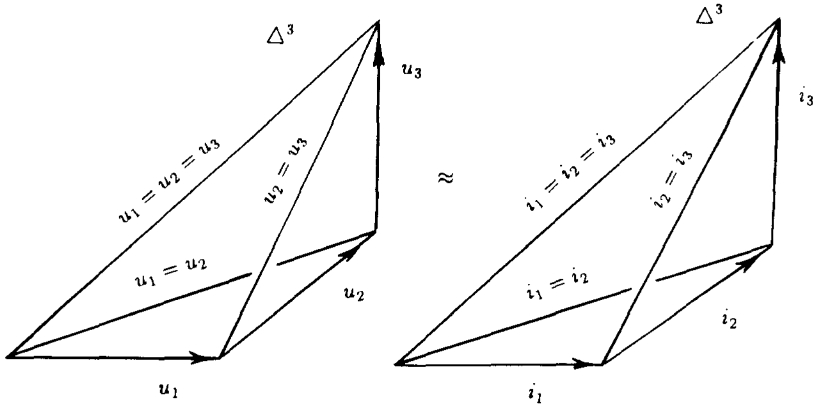

FIGURE 9

since $[K_i, K_i] = 0$. If we now use the fact that the inverse order is used in defining the integral, we obtain

$$T_{\mathscr{X}}^{(1)}(0) = \sum_{N \geq i > j \geq 1} [K_i, K_j] = - \sum_{N \geq i > j \geq 1} [X_i, X_j]$$

which agrees with (65).

(b) Consider next the case $k = 2$, so that $\Delta^{k+1} = \Delta^3$ (see Fig. 9), so that the two faces B and C of Δ^3 are defined by the equations $B = (u_1 = u_2)$ and $C = (u_2 = u_3)$ (see Fig. 10). Again assume that K has been chosen left continuous. Enumerate the polyhedra which intersect Δ^3 using the integer coordinates of their left lower front vertex so that we write

$$\theta_\alpha = \theta_{i_1, i_2, i_3},$$

(see Figs. 11, 12, and 13).

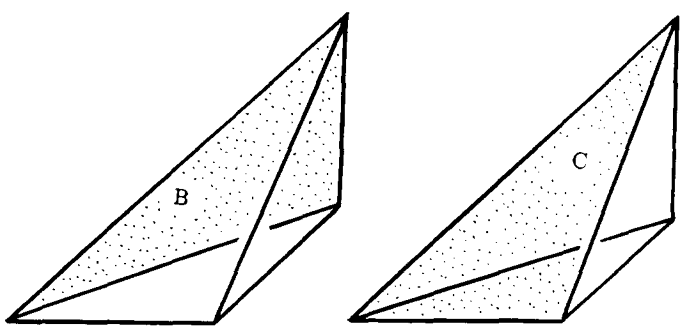

FIGURE 10

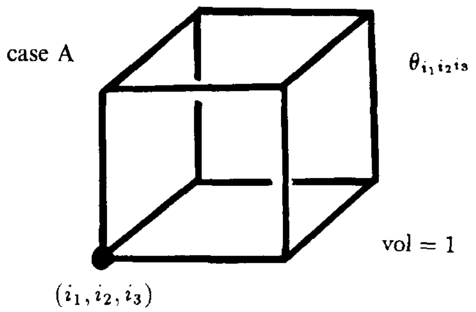

FIGURE 11

Each polyhedron $\theta_{i_1 i_2 i_3}$ which is totally embedded in Δ^3 is a standard unit cube of volume $= 1$ (Fig. 11). Here $i_1 > i_2 > i_3$.

Each polyhedron $\theta_{i_1 i_2 i_3}$, where $i_1 = i_2$, intersects the boundary face B (Fig. 12) and has volume $= \frac{1}{2}$. Each polyhedron $\theta_{i_1 i_2 i_2}$, where $i_1 = i_2$, intersects the boundary face B (Fig. 12) and has volume $= \frac{1}{2}$. Each polyhedron $\theta_{i_1 i_2 i_2}$, where $i_2 = i_3$, intersects the boundary face C (Fig. 13) and also has volume $= \frac{1}{2}$.

Consider the integral (66)″

$$I = T_{\infty}^{(2)}(0) = 2 \iiint_{b \geqslant u_1 > u_2 > u_3 \geqslant a} [[K(u_1), K(u_2)], K(u_3)]\, du_1\, du_2\, du_3.$$

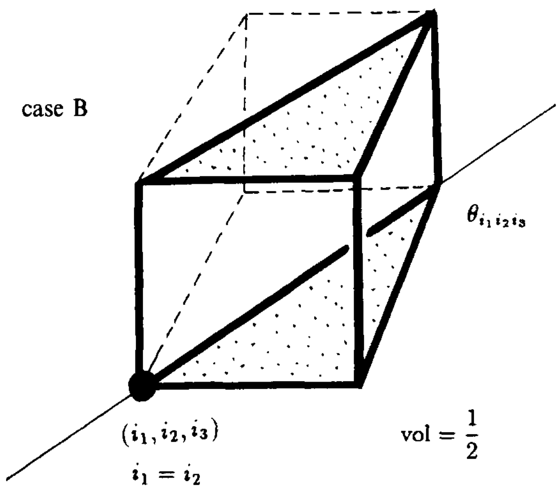

FIGURE 12

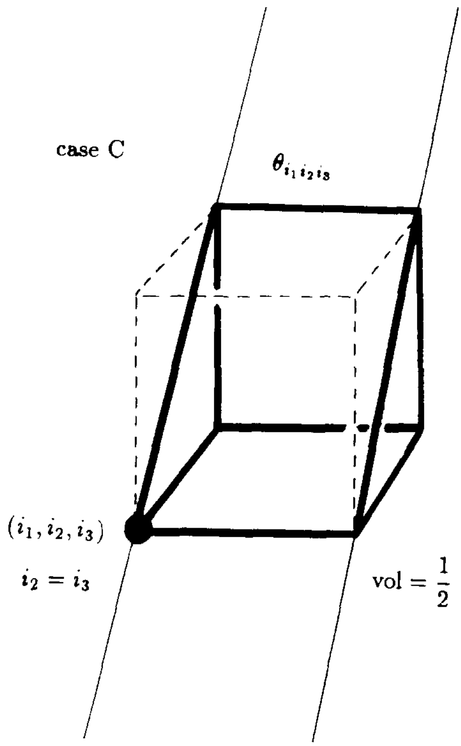

FIGURE 13

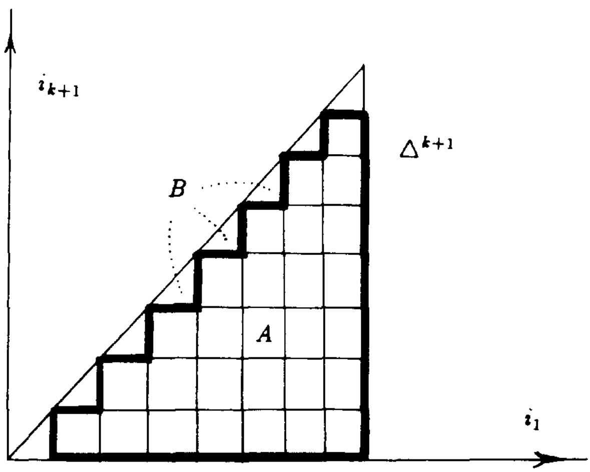

FIGURE 14

Because the function $F(u)$ is a step function (see above), the integral $T^{(2)}_{\infty}(0)$ transforms to the sum

$$I = 2 \sum_{\{\theta_\alpha\}} F(\theta_\alpha) \cdot \operatorname{vol} \theta_\alpha$$

$$= 2 \sum_{N \geqslant i_1 \geqslant i_2 \geqslant i_3 \geqslant 1} \operatorname{vol} \theta_{i_1 i_1 i_3} [[K_{i_1}, K_{i_2}], K_{i_3}],$$

where $[[K_{i_1}, K_{i_2}], K_{i_3}]$ is the value of the function $F(u)$ at the point (i_1, i_2, i_3), which is the indicated vertex of the polyhedron θ_α (see above).

The integral may be written in the form $I = A + B + C$, where

$$A = 2 \sum_{\substack{\theta_\alpha \text{ of type } A \\ \text{(Fig. 11)}}} F(\theta_\alpha) \operatorname{vol} \theta_\alpha,$$

$$B = 2 \sum_{\substack{\theta_\alpha \text{ of type } B \\ \text{(Fig. 12)}}} F(\theta_\alpha) \operatorname{vol} \theta_\alpha,$$

$$C = 2 \sum_{\substack{\theta_\alpha \text{ of type } C \\ \text{(Fig. 13)}}} F(\theta_\alpha) \operatorname{vol} \theta_\alpha.$$

This implies that

$$A = 2 \sum_{N \geqslant i_1 > i_2 > i_3 \geqslant 1} [[K_{i_1}, K_{i_2}], K_{i_3}] \cdot 1, \qquad \text{because } \operatorname{vol} \theta_{i_1 i_2 i_3} = 1;$$

$$B = 2 \sum_{N \geqslant i_1 > i_3 \geqslant 1} [[K_{i_1}, K_{i_1}], K_{i_3}] \cdot \frac{1}{2}, \qquad \text{because } \operatorname{vol} \theta_{i_1 i_1 i_3} = \frac{1}{2};$$

$$C = 2 \sum_{N \geqslant i_1 > i_2 \geqslant 1} [[K_{i_1}, K_{i_2}], K_{i_2}] \cdot \frac{1}{2}, \qquad \text{because } \operatorname{vol} \theta_{i_1 i_2 i_2} = \frac{1}{2}.$$

But $B \equiv 0$, because $[K_{i_1}, K_{i_1}] = 0$ and we have

$$I = A + C = 2 \sum_{N \geqslant i_1 > i_2 > i_3 \geqslant 1} [[K_{i_1}, K_{i_2}], K_{i_3}] + \sum_{N \geqslant i_1 > i_2 \geqslant 1} [[K_{i_1}, K_{i_2}], K_{i_2}].$$

This formula agrees wit (54)‴

$$\sum_{i=2}^{N} t^{(2)}_{N,i}(0) = 2 \sum_{N \geqslant p > q > j \geqslant 1} [X_p, [X_q, X_j]] + \sum_{N \geqslant m > j \geqslant 1} [X_m, [X_m, X_j]],$$

which is formula (65) for $k=2$. Thus Theorem 11 is proved in the case $k=2$.

Note that in writing $I=A+B+C$ we counted certain of the polyhedra twice, those along the line of intersection of B and C. That this apparently incorrect procedure gives the right answer follows because the function which is being summed has the value zero there (the reason that $C\equiv 0$).

(c) Arbitrary k, so that the simplex is Δ^{k+1}. Denote the polyhedra θ_α which are completely contained in Δ^{k+1} by θ_β. Denote that part of the polyhedron which is contained in Δ^{k+1} among those θ_α which intersect Δ^{k+1} but are not totally contained in Δ^{k+1} by θ_α (Fig. 14). We let

$$A'=\sum_{(\beta)}\theta_\beta,\qquad B'=\sum_{(\gamma)}\theta_\gamma,$$

so that

$$\begin{aligned}I&=k!\int_{\Delta^{k+1}}F(u)\,du\\&=k!\int_{A'}F(u)\,du+k!\int_{B'}F(u)\,du=A+B\\&=k!\sum_{A'}F(\theta_\beta)\operatorname{vol}(\theta_\beta)+k!\sum_{B'}F(\theta_\gamma)\operatorname{vol}(\theta_\gamma).\end{aligned}$$

The sum

$$\begin{aligned}&k!\sum_{A'}F(\theta_\beta)\operatorname{vol}(\theta_\beta)\\&\quad=k!\sum_{N\geqslant i_1>\cdots>i_{k+1}\geqslant 1}[\bar{K}_{i_1},[\ldots,[\bar{K}_{i_k},\bar{K}_{i_{t+1}}]\cdots]\end{aligned}$$

which is the sum A from (69).

The sum

$$B=k!\sum_{B'}[\bar{K}_{i_1},[\cdots[\bar{K}_{i_k},\bar{K}_{i_{k+1}}]\cdots]\operatorname{vol}(\theta_{i_1i_2\cdots i_k})$$

corresponds to the boundary polyhedra and therefore may be decomposed into parts corresponding to the interior faces of Δ^{k+1}. The volumes of the part cut from these by the faces of Δ^{k+1} are equal to

$$\frac{1}{(k-\alpha_1)!\cdots(\alpha_{N-i-1}-\alpha_{N-i})!\,\alpha_{N-i}!}$$

in (65) depending on the face. We have therefore a geometrical interpretation of these coefficients.

COROLLARY 1. *In the discrete case when the function $K(s)$, $a \leqslant s \leqslant b$ is a step function (i.e., constant except for a finite number of jumps), the integral* (66) *transforms into the following sum*:

$$T_{\infty}^{(k)}(0) = A + B, \qquad \textit{where}$$

$$A = k! \sum_{N \geqslant i_1 > \cdots > i_{k+1} \geqslant 1} [\cdots [K(s_{i_1}), K(s_{i_2})], ..., K(s_{i_k})], K(s_{i_{k+1}})],$$

where $s_1, s_2, ..., s_{i_{k+1}}$ are the jump-points of the function $K(s)$ in the interval $a \leqslant s \leqslant b$;

$$B = \sum_{N \geqslant i_1 \geqslant \cdots \geqslant i_{k+1} \geqslant 1} \mathrm{vol}(\theta_{i_1 i_2 \cdots i_{k+1}}) \times [\cdots [K(s_{i_1}), K(s_{i_2})], ..., K(s_{i_k})], K(s_{i_{k+1}})],$$

where there exists at least one linear relation between the indices $i_1, ..., i_{k+1}$ (at least one pair of neighboring indices coincide). The $\mathrm{vol}(\theta_{i_1, ..., i_{k+1}})$ *can be expressed in terms of the coefficients $k!/(k-\alpha_1)!\,(\alpha_1 - \alpha_2)! \cdots (\alpha_{N-i})!$ from* (65).

REFERENCES

1. I. BIALYNICKI-BIRULA, B. MIELNIK, AND J. PLEBANSKI, Explicit solution of the continuous Baker–Campbell–Hausdorff problem and a new expression for the phase operator, *Ann. Phys.* **51** (1969), 187–200.
2. R. V. CHACON AND S. KOCHEN, to appear.
3. K.-T. CHEN, Integration of paths, geometric invariants and a generalized Baker–Hausdorff formula, *Ann. of Math.* **65** (1957), 163–178.
4. K.-T. CHEN, Formal differential equations, *Ann. of Math.* **73** (1961), 110–133.
5. K.-T. CHEN, An expansion formula for differential equations, *Bull. Amer. Math. Soc.* **68** (1962), 341–344.
6. K.-T. CHEN, Expansions of solutions of differential systems, *Arch. Rational Mech. Anal.* **13** (1963), 348–363.
7. K.-T. CHEN, On a generalization of Picard's approximation, *J. Differential Equations* **2** (1966), 438–448.
8. J. D. DOLLARD AND C. N. FRIEDMAN, Product integration, *in* "Encyclopedia of Mathematics and its Applications" (G.-C. Rota, Ed.), Vol. 10, Addison–Wesley, Reading, MA, 1979.
9. R. C. FEYNMAN, An operator calculus having applications in quantum electrodynamics, *Phys. Rev.* **84** (2) (1951), 108–128.
10. J. FROELICH AND N. SALINGAROS, The exponential mapping in Clifford algebras, *J. Math. Phys.* **25**(8) (1984), 2347–2350.

11. K. GOLDBERG, The formal power series for $\log e^x e^y$, *Duke Math. J.* **23** (1956), 13–21.
12. M. V. KARASEV AND M. V. MOSOLOVA, The infinite products and T-products of exponents, *J. Theoret. Math. Phys.* **28** (2) (1976), 189–200.
13. W. MAGNUS, On the exponential solution of differential equations for a linear operator, *Comm. Pure Appl. Math.* **7** (1954), 649–673.
14. V. P. MASLOV, "Operator Methods," Nauka, Moscow, 1973.
15. M. A. NAIMARK AND A. I. STERN, "Theory of Group Representations," Springer-Verlag, Berlin/New York, 1981.
16. R. S. STRICHARTZ, The Campbell–Baker–Hausdorff–Dynkin formula and solutions of differential equations, *J. Funct. Anal.* **72** (2) (1987), 320–345.
17. R. M. WILCOX, Exponential operators and parameter differentiation in quantum physics, *J. Math. Phys.* **8** (4) (1967), 962–982.
18. A. M. GARSIA, Combinatorics of the free Lie algebra and the symmetric group, preprint, La Jolla, pp. 1–62 (1988).
19. A. T. FOMENKO AND V. V. TROFIMOV, "Integrable Systems on Lie algebra and Symmetric Spaces," Gordon & Breach, New York, 1988.
20. M. NEWMAN AND R. C. THOMPSON, Numerical values of Goldberg's coefficients in the series for $\log(e^x e^y)$, *Math. Comp.* **48** (177) (1987), 265–271.