Theory and Applications of Nonlinear Control Systems
C.I. Byrnes and A. Lindquist (editors)

A product expansion of the Chen series

H. J. Sussmann*

Department of Mathematics
Rutgers University
New Brunswick, NJ 08903
U.S.A.

We present an expansion for the Chen series as a product of exponentials of the members of a P. Hall basis multiplied by coefficients that have simple expressions as iterated integrals. We then discuss the convergence properties of the expansion.

The idea of associating a formal power series in noncommutative indeterminates to a control system has proven to be fruitful for many problems. The series (known as the "Chen series" in the work of M. Fliess and his coworkers) has a number of interesting properties, such as the fact that it is an exponential of a Lie series. However, no simple expression is known for the Lie series. The purpose of this note is to present a formula which expresses the Chen series as a product of exponentials of the elements of a Philip Hall basis, multiplied by coefficients that have very simple explicit expressions as iterated integrals of the controls.

Let $\underline{X} = \{X_i : i \in I\}$ be an indexed family of indeterminates. For a real number $T \geq 0$, let $U_T(I)$ denote the product of a family $\{L_i : i \in I\}$ of copies of the space $L^1([0,T],\mathbb{R})$. Finally, let $U(I) = \bigcup_{0\leq T} U_T(I)$. The elements of $U(I)$ are therefore indexed families $u = \{u_i(\cdot) : i \in I\}$ such that there is a $T \geq 0$ (called the duration of u) with the property that each $u_i(\cdot)$ is a Lebesgue integrable real-valued function on $[0,T]$. The set $U(I)$ is a semigroup with identity, with respect to the operation $\star : U(I) \times U(I) \to U(I)$ of concatenation. (The concatenation $u \star v$ of $u = \{u_i(\cdot) : i \in I\}$, with duration T, and $v = \{v_i(\cdot) : i \in I\}$, with duration T', is the element $w = \{w_j(\cdot) : i \in I\}$ of $U_{T+T'}(I)$ given by $w_i(t) = u_i(t)$ for $0 \leq t \leq T$ and $w_i(t) = v_i(t-T)$ for $T < t < T+T'$.)

Let $\mathcal{I}(I)$ be the set of all finite sequences $\sigma = (i_1,\ldots,i_k)$ of elements of I, of arbitrary length k. We use $|\sigma|$ to denote the length of σ. (The empty sequence $\emptyset$ is included.) Also, we use $\mathcal{I}_k(I)$ to denote the set of sequences of length k. For each $\sigma \in \mathcal{I}(I)$ we define the monomial

$$X_\sigma = X_{i_1}X_{i_2}\ldots X_{i_k} \tag{1}$$

*Partially supported by NSF Grant MCS78-02442

if $\sigma = (i_1, \ldots, i_k)$. (If $\sigma = \emptyset$, we let $X_\sigma = 1$.) The set of all formal linear combinations $\sum_{\sigma \in \mathcal{I}(I)} a_\sigma X_\sigma$ such that $a_\sigma \in R$ for all σ and $a_\sigma = 0$ for all but finitely many σ's is the *free associative algebra generated by* $\underline{X}$, and is denoted by $A(\underline{X})$. The algebra of all formal sums $\sum_{\sigma \in \mathcal{I}(I)} a_\sigma X_\sigma$, where the a_σ are real numbers, is denoted by $\hat{A}(\underline{X})$. If $\{S_j : j \in J\}$ is an indexed family of elements of $\hat{A}(\underline{X})$, and each S_j is given by $S_j = \sum_{\sigma \in \mathcal{I}(I)} S_{j,\sigma} X_\sigma$, then the sum $\sum_{j \in J} S_j$ will be said to be *convergent* if, for each σ, the set $\{j : S_{j,\sigma} \neq 0\}$ is finite. In that case, the sum of the series is a well defined element of $\hat{A}(\underline{X})$. For a $u \in U(I)$, given by $u = \{u_i(\cdot) : i \in I\}$, with duration T, and a multiindex $\sigma = (i_1, \ldots, i_k) \in \mathcal{I}(I)$, we define the *iterated integral*

$$\int_0^t u_\sigma = \int_0^t \int_0^{\tau_1} \int_0^{\tau_2} \ldots \int_0^{\tau_{k-1}} u_{i_1}(\tau_1)\, u_{i_2}(\tau_2) \ldots u_{i_k}(\tau_k)\, \mathrm{d}\tau_k \ldots \mathrm{d}\tau_1, \tag{2}$$

for $0 \leq t \leq T$. We then associate with u the $\hat{A}(\underline{X})$-valued function S_u on $[0, T]$ given by

$$S_u(t) = \sum_\sigma \left(\int_0^t u_\sigma \right) X_{\sigma^*} \tag{3}$$

(Here, for $\sigma = (i, \ldots, i_k)$, we use σ^* to denote the reversed sequence $(i_k, \ldots, i_1)$. Also, we let $\int_0^t u_\sigma = 1$ if $\sigma = \emptyset$.) The series $S_u(T)$ is the *formal power series* associated with u, and will be denoted by $\mathrm{Ser}(u)$. The map $\mathrm{Ser} : U(I) \to \hat{A}(\underline{X})$ is a semigroup homomorphism from $U(I)$ into $\hat{A}(X)$, with respect to the operations of concatenation and multiplication, respectively.

It is easy to see that the series-valued function S_u is the solution of the differential equation

$$\dot{S}(t) = S(t) \left(\sum_{i \in I} u_i(t) X_i \right) \tag{4}$$

with initial condition $S(0) = 1$. From this it follows easily that, if $u^\star$ denotes the element $\{u_i^\star(\cdot) : i \in I\}$ of $U_T(I)$ given by $u_i^\star(t) = -u_i(T - t)$, then $\mathrm{Ser}(u)\,\mathrm{Ser}(u^\star) = 1$. In particular, the image of $U(I)$ under Ser is a group under multiplication. This group will be denoted by $G_0(\underline{X})$. An important property of $G_0(\underline{X})$ is the inclusion $G_0(\underline{X}) \subseteq G(\underline{X})$, where $G(\underline{X})$ is the set of *exponential Lie series* in the indeterminates $\underline{X}$. Precisely, let $L(\underline{X})$ denote the Lie subalgebra of $A(\underline{X})$ generated by the X_i. Then let $\hat{L}(\underline{X})$ denote the set of all elements of $\hat{A}(\underline{X})$ that are equal to a convergent sum $\sum_{j \in J} S_j$ of elements of $L(\underline{X})$. The elements of $\hat{L}(\underline{X})$ are the *Lie series* in the indeterminates $\underline{X}$. If S is any series in $\hat{A}(\underline{X})$ whose constant term (i.e. the coefficient of 1) vanishes (in particular if $S \in \hat{L}(\underline{X})$), then the exponential $\exp(S)$ is well-defined by the series

$$\exp(S) = \sum_{n=0}^{\infty} \frac{1}{n!} S^n. \tag{5}$$

The series of the form $\exp(S)$, $S \in \hat{L}(\underline{X})$, are the *exponential Lie series*. The Campbell-Hausdorff formula implies that the exponential Lie series form group under multiplication, which will be denoted by $G(\underline{X})$. There is a well defined *logarithm* map which assigns, to every series $S \in \hat{A}(\underline{X})$ whose constant term is equal to one, the series $\log S$ given by

$$\log S = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}(S-1)^n \tag{6}$$

The maps exp and log are inverses of each other. In particular, exp maps $\hat{L}(\underline{X})$ onto $G(\underline{X})$, and log maps $G(\underline{X})$ onto $\hat{L}(\underline{X})$.

Since our goal is to get an expression for $\mathrm{Ser}(u)$ that exhibits its significant properties, it is natural to write

$$\mathrm{Ser}(u) = \exp(Z(u)) \tag{7}$$

and seek an expression for $Z(u)$. Moreover, an explicit expression for $Z(u)$ can easily be derived, since $Z(u) = \log S(u)$, and there is an explicit formula for the logarithm. The resulting formula is

$$Z(u) = \sum_{n=1}^{\infty} \sum_{\sigma_1 \in \mathcal{I}'(I)} \cdots \sum_{\sigma_n \in \mathcal{I}'(I)} \frac{(-1)^{n+1}}{n} \left(\int_0^T u_{\sigma_1}\right) \cdots \left(\int_0^T u_{\sigma_n}\right) X_{\sigma_1^*} \cdots X_{\sigma_n^*}. \tag{8}$$

(Here $\mathcal{I}'(I) = \mathcal{I}(I) - \{\emptyset\}$.) However, this expression is not very useful. (For instance, the fact that $Z(u)$ is a Lie series is not at all evident from (8).)

A slight simplification can be achieved by using the fact that $Z(u) \in \hat{L}(\underline{X})$. For each monomial $X_\sigma = X_{i_1} \cdots X_{i_k}$, define $[X_\sigma]$ by

$$[X_\sigma] = \left[X_{i_1}, \left[X_{i_2}, \left[\ldots \left[X_{i_{k-1}}, X_{i_k}\right] \cdots\right]\right]\right]. \tag{9}$$

Then it can be proved that, if $A = \sum_\sigma a_\sigma X_\sigma$ happens to be a Lie series, then $A = \sum_\sigma \frac{a_\sigma}{|\sigma|}[X_\sigma]$. In particular, this yields the formula

$$Z(u) = \sum_{n=1}^{\infty} \sum_{\sigma_1 \in \mathcal{I}'(I)} \cdots \sum_{\sigma_n \in \mathcal{I}'(I)} \frac{(-1)^{n+1}}{n(|\sigma_1| + \cdots + |\sigma_n|)} \left(\int_0^T u_{\sigma_1}\right) \cdots \left(\int_0^T u_{\sigma_n}\right) [X_{\sigma_1^*} \cdots X_{\sigma_n^*}]. \tag{10}$$

This formula is clearly better than (8), but it is still far from solving our problem, since the Lie monomials that occur in (10) are not linearly independent, and there is no easy way to single out a basis and write down the corresponding coefficients.

As an alternative to the above formulas for $\mathrm{Ser}(u)$ and $Z(u)$, we propose to the write an expression for $\mathrm{Ser}(u)$ as an infinite product of exponentials. First, we have to give a precise definition of the infinite products. For any series $S \in \hat{A}(\underline{X})$, let us use $\langle S, \sigma\rangle$ to denote the coefficient of X_σ, so that $S = \sum \langle S, \sigma\rangle X_\sigma$. Let J be a set which is totally ordered by a relation $\leq$. Let $\left\{S^j : j \in J\right\}$ be a family of series in $\hat{A}(\underline{X})$, indexed by $j \in J$. We define the infinite product $\overleftarrow{\prod}_{j \in J} S^j$ by:

$$\overleftarrow{\prod_{j \in J}} S^j = \sum_{\sigma \in \mathcal{I}(I)} P_\sigma X_\sigma \tag{11}$$

where

$$P_\sigma = \sum_{n=0}^{\infty} \sum_{\substack{j_1\in J,\ldots,j_n\in J\\ j_1>j_2>\cdots>j_n}} \sum_{\substack{\sigma_1\in\mathcal{I}'(I),\ldots,\sigma_n\in\mathcal{I}'(I)\\ \sigma_1\#\sigma_2\#\cdots\#\sigma_n=\sigma}} \langle S^{j_1},\sigma_1\rangle\cdots\langle S^{j_n},\sigma_n\rangle. \tag{12}$$

Here, if $\sigma_1,\ldots,\sigma_n$ are in $\mathcal{I}(I)$, we use $\sigma_1\#\ldots\#\sigma_n$ to denote the concatenation of $\sigma_1,\sigma_2,\ldots,$ σ_n (so that $X_{\sigma_1\#\sigma_2\#\cdots\#\sigma_n} = X_{\sigma_1}X_{\sigma_2}\cdots X_{\sigma_n}$). If $\sigma=\emptyset$, then (12) implies that $P_\sigma = 1$. The product is defined provided that: (i) all the $S^j_\emptyset$ are equal to 1, and (ii) for each σ there are only finitely many choices of n, $j_1,\ldots,j_n$, $\sigma_1,\ldots,\sigma_n$, such that $j_1\in J,\ldots,j_n\in J$, $j_1>j_2>\cdots>j_n$, $\sigma_1\in\mathcal{I}'(I),\ldots,\sigma_n\in\mathcal{I}'(I)$, $\sigma_1\#\cdots\#\sigma_n=\sigma$, and $\langle S^{j_1},\sigma_1\rangle\langle S^{j_2},\sigma_2\rangle\cdots\langle S^{j_n},\sigma_n\rangle >\neq 0$.

Next we have to define the concept of a *P. Hall basis* of $L(\underline{X})$. First, we define the set $\mathrm{Br}(\underline{X})$ of *formal brackets* in the indeterminates $\underline{X}$. (This is essentially the same as the *magma* generated by $\underline{X}$.) Let A be the alphabet that consists of the X_i, the left and right square brackets, and the comma. Then $\mathrm{Br}(\underline{X})$ is the smallest set Ω of words in the alphabet A, such that (i) $X_i\in\Omega$ for all $i\in I$, and (ii) whenever α,β are two words that belong to Ω, then the word "$[\alpha,\beta]$" also belongs to Ω. Every formal bracket $B\in\mathrm{Br}(\underline{X})$ has a well-defined *degree* $\deg(B)$, which is an integer ≥ 1. If $\deg(B)\geq 2$, then B can be written in a unique way as $[B_1,B_2]$, with $B_1\in\mathrm{Br}(\underline{X}), B_2\in\mathrm{Br}(\underline{X})$ The formal brackets $B_1,_2 B_2$ are called the *left and right factors* of B, and are denoted by $\lambda(B)$ and $\rho(B)$, respectively. There is a mapping μ which associates with each $B\in\mathrm{Br}(\underline{X})$ an element of $L(\underline{X})$. The elements of $L(\underline{X})$ of the form $\mu(B)$, $B\in\mathrm{Br}(\underline{X})$, are the *Lie monomials* in the X_i. (But notice that: (a) the map μ is not one-to-one, and (b) the unique factorization is a property of the *formal brackets*, not of the Lie monomials. For instance, if X and Y are two different indeterminates, and we let $B_1=[X,[Y,[X,Y]]]$, $B_2=[Y,[X,[X,Y]]]$, then $\lambda(B_1)=X$ and $\lambda(B_2)=Y$. However, the Jacobi identity implies that $\mu(B_1)=\mu(B_2)$.)

A *Philip Hall basis* of $L(\underline{X})$ is a subset $\mathcal{B}$ of $\mathrm{Br}(\underline{X})$, totally ordered by a relation $\leq$, such that:

(I) every X_i is in $\mathcal{B}$,

(II) if $B_1\in\mathcal{B}, B_2\in\mathcal{B}$, and $B_1\leq B_2$, then $\deg(B_1)\leq\deg(B_2)$,

(III) if $B\in\mathrm{Br}(\underline{X})$ and $\deg(B)\geq 2$, so that $B=[B_1,B_2]$, $B_1\in\mathrm{Br}(\underline{X})$, $B_2\in\mathrm{Br}(\underline{X})$, then $B\in\mathcal{B}$ if and only if: (a) $B_1\in\mathcal{B}$ and $B_2\in\mathcal{B}$, (b) $B_1<B_2$, and (c) either $\deg(B_2)=1$ or $\lambda(B_2)\leq B_1$.

If $(\mathcal{B},\leq)$ is a P. Hall basis of $L(\underline{X})$, then it can be proved that μ is one to one on $\mathcal{B}$, so that $\mathcal{B}$ can also be regarded as a set of Lie monomials. When this is done, then B turns out to be a basis of $L(X)$ (cf. [1, 2, 3]).

We then have

Lemma. *Let $\mathcal{B}$ be a P. Hall basis of $L(\underline{X})$. Then:*

(i) for every familly $\alpha=\{\alpha_B : B\in\mathcal{B}\}$ of real numbers, the infinite product

$$\overleftarrow{P}_{\mathcal{B}}(\alpha) = \overleftarrow{\prod_{B\in\mathcal{B}}} \exp(\alpha_B B) \tag{13}$$

is defined,

(ii) $\alpha \to \overleftarrow{P}_{\mathcal{B}}(\alpha)$ *is a bijection from* $\mathbb{R}^{\mathcal{B}}$ *onto* $G(\underline{X})$.

Proof. Let $S^B = \exp(\alpha_B B)$. It is clear that $\langle S^B, \emptyset \rangle = 1$ for all B. Let $\sigma = (i_1, \ldots, i_k) \in \mathcal{I}(I)$. Then $\sigma \in \mathcal{I}(F)$ for some finite subset of I. If $B \in \mathrm{Br}(\underline{X})$, let $\mathrm{supp}(B)$ denote the set of those $i \in I$ that occur in B. (That is: $\mathrm{supp}\,(X_i) = \{i\}$, and $\mathrm{supp}\,[B_1, B_2] = \mathrm{supp}\,(B_1) \cup \mathrm{supp}\,(B_2)$.) Then it is easy to show that, if $B = \sum_{\tau \in \mathcal{I}(I)} B_\tau X_\tau$, then al1 the X_τ for which $B_\tau \neq 0$ are such that $\mathrm{supp}(\tau) = \mathrm{supp}(B)$. (If $\tau = (\tau_1, \ldots, \tau_m)$, we define $\mathrm{supp}(\tau) = \{\tau_1, \ldots, \tau_m\}$.) If $\alpha \in \mathbb{R}, B \in \mathrm{Br}(\underline{X})$, and $\alpha \neq 0$, the series $\exp(\alpha B)$ also has the property that the only monomials X_τ of positive degree that can occur are those for which $\mathrm{supp}(\tau) = \mathrm{supp}(B)$. If $\sigma_1, \ldots, \sigma_n$, $B_1, \ldots, B_n$ are such that $\sigma_i \in \mathcal{I}'(I)$, $B_i \in \mathcal{B}$, $\sigma_1 \# \cdots \# \sigma_n = \sigma$, $B_1 > B_2 > \cdots > B_n$, and $\langle B_i, \sigma_i \rangle \neq 0$ for $i = 1, \ldots, n$, then $\mathrm{supp}(B_i) = \mathrm{supp}\,(\sigma_i)$, and so $\mathrm{supp}\,(B_i) \subseteq F$. Moreover, $1 \leq \deg\,(B_i) \leq |\sigma_i| \leq |\sigma|$. Hence $n \leq |\sigma|$. Since there are only finitely many elements B of $\mathrm{Br}(\underline{X})$ such that $\mathrm{supp}(B) \subseteq F$ and $\deg(B) \leq |\sigma|$, we conclude that the number of choices of n, $\sigma_1, \ldots, \sigma_n$, $B_1, \ldots, B_n$ is finite. So $\overleftarrow{P}_{\mathcal{B}}(\alpha)$ is defined.

That $\overleftarrow{P}_{\mathcal{B}}(\alpha)$ is an exponential Lie series follows easily from the Friedrichs criterion (cf. Serre [2]). If S is an exponential Lie series, let $S = \exp(Z)$, $\Sigma \in \hat{L}(\underline{X})$. Write $Z = \tilde{Z}_2 + \sum_{B \in \mathcal{B}_1} \alpha_B B$, where $\mathcal{B}_1$ is the set of those $B \in \mathcal{B}$ such that: $\deg(B) = 1$, and $\tilde{Z}_2 \in \hat{L}(\underline{X})$ contains no monomial of degree one. Then the Campbell-Hausdorff formula implies that

$$\exp(Z) \cdot \left(\overleftarrow{\prod_{B \in \mathcal{B}_1}} \exp(\alpha_B B) \right)^{-1} = \exp(Z_2), \tag{14}$$

where Z_2 contains no monomials of degree one. Suppose that we have defined the α_B for $B \in \mathcal{B}_k$ (where $\mathcal{B}_k = \{B : B \in \mathcal{B}, \deg(B) \leq k\}$) such that

$$\exp(Z) \cdot \left(\overleftarrow{\prod_{B \in \mathcal{B}_k}} \exp\,(\alpha_B B) \right)^{-1} = \exp\,(Z_{k+1}) \tag{15}$$

where Z_{k+1} contains no monomial of degree $\leq k$. Then we let

$$Z_{k+1} = \tilde{Z}_{k+2} + \sum_{\substack{B \in \mathcal{B} \\ \deg(B) = k+1}} \alpha_B B,$$

where $\tilde{Z}_{k+2}$ contains no monomials of degree $\leq k+1$. Then (15) holds, with a suitable choice of Z_{k+2}, if k is replaced by $k+1$. So the α_B are defined for all $B \in \mathcal{B}$, and (15) holds for all k. From this it follows easily that $S = \overleftarrow{P}_{\mathcal{B}}(\alpha)$.

Finally, we must show that the map $\alpha \to \overleftarrow{P}_{\mathcal{B}}(\alpha)$ is injective. Let α, β be such that $\overleftarrow{P}_{\mathcal{B}}(\alpha) = \overleftarrow{P}_{\mathcal{B}}(\beta)$. Suppose that $\alpha \neq \beta$. Let B^* be the first element B of $\mathcal{B}$ such that $\alpha_B \neq \beta_B$. Then

$$\overleftarrow{\prod_{\substack{B \in \mathcal{B} \\ B \geq B^*}}} \exp\,(\alpha_B B) = \overleftarrow{\prod_{\substack{B \in \mathcal{B} \\ B \geq B^*}}} \exp\,(\beta_B B)\,. \tag{16}$$

The left side is equal to $1 + \Sigma^{\#} \alpha_B B + S_1$, where the sum $\Sigma^{\#}$ runs over those $B \in \mathcal{B}$ such that $B \geq B^*$ and $\deg(B) = \deg\,(B^*)$, and S_1 contains no monomials of degree $\leq 1 + \deg\,(B^*)$.

A similar expression holds for the right side of (16). So $\Sigma^{\#}\alpha_B B = \Sigma^{\#}\beta_B B$. Since $\mathcal{B}$ is a linearly independent set, we have $\alpha_{B^*} = \beta_{B^*}$, which is a contradiction. Hence $\alpha = \beta$, and our proof is complete. $\square$

The lemma implies, in particular, that

$$\text{(17)} \qquad \mathrm{Ser}(u) = \overleftarrow{\prod_{B\in\mathcal{B}}} \exp(\alpha_B(u)B),$$

where the α_B are real-valued functions on $U(I)$, uniquely determined by (17). Our goal is to exhibit an explicit formula for the $\alpha_B(\cdot)$ as iterated integrals. We shall associate with each formal bracket $B \in \mathrm{Br}(\underline{X})$, and each $u \in U(I)$, two functions $C_B(u)$, $c_B(u)$ defined on the interval $[0,T]$ (if T is the duration of u). The function $c_B(u)$ will be in $L^1([0,T],\mathbb{R})$, and then $C_B(u)$ will be given by

$$\text{(18)} \qquad C_B(u)(t) = \int_0^t C_B(u)(s)\,\mathrm{d}s.$$

The $c_B(u), C_B(u)$ are defined as follows. Let $u = \{u_i(\cdot) : i \in I\}$. For $B = X_i$, we let $c_B(u)(t) = u_i(t)$, $0 \le t \le T$, and then define $C_B(u)$ by (18). Assume that $c_B(u)$ and $C_B(u)$ have been defined for all $B \in \mathrm{Br}(\underline{X})$ such that $\deg(B) < k$, and let $B \in \mathrm{Br}(\underline{X})$ have degree $k+1$. Then B can be written in a unique way as $(\mathrm{ad}\, B_1)^m(B_2)$, with $\lambda(B_2) \ne B_1$, $m \ge 1$. (As usual, $\mathrm{ad}\, P$ is the mapping $Q \to [P,Q]$.) We then define

$$\text{(19)} \qquad c_B(u) = \frac{1}{m!} C_{B_1}(u)^m c_{B_2}(u)$$

and, again, we define $C_B(u)$ by (18). This completes the recursive definition of the $c_B(u)$, $C_B(u)$. We then have:

Theorem. *Let $u \in U(I)$ have duration T. Let $\mathcal{B}$ be a P. Hall basis of $L(X)$. Then*

$$\text{(20)} \qquad \mathrm{Ser}(u) = \overleftarrow{\prod_{B\in\mathcal{B}}} \exp\left(C_B(u)(T)B\right).$$

Formula (20) is the desired product expansion.

Proof. Let $\sigma \in \mathcal{I}(I)$ and let $F \subseteq I$ be finite and such that $\sigma \in \mathcal{I}(F)$. Let $u^F = \{u_i(\cdot) : i \in F\}$, Let $\underline{X}^F = \{X_i : i \in F\}$, and let $\pi^F : \hat{A}(\underline{X}) \to \hat{A}\left(\underline{X}^F\right)$ be the map that sends X_i to X_i if $i \in F$, and to zero if $i \notin F$. Then, if Q is any series in $\hat{A}(X)$ the coefficients $\langle Q, \sigma\rangle$ and $\langle \pi^F(Q), \sigma\rangle$ are equal. To prove that X_σ has the same coefficient in both sides of (20), it suffices to show that the results of applying F to both sides of (20) are the same. But $\pi^F(\mathrm{Ser}(u)) = \mathrm{Ser}\left(u^F\right)$, and $\pi^F\left(\overleftarrow{\prod}_{B\in\mathcal{B}} \exp\left(C_B(u)(T)B\right)\right) = \overleftarrow{\prod}_{B\in\mathcal{B}^F} \exp\left(C_B\left(u^F\right)(T)B\right)$ where $\mathcal{B}^F$ is the set of those $B \in \mathcal{B}$ that only contain X_i's that belong to F. Hence it suffices to prove (20) when I is a finite set. From now on, we assume that I is finite.

Let $\mathcal{B}(k) = \{B : B \in \mathcal{B}, \deg(B) = k\}$. Then each $\mathcal{B}(k)$ is a finite set. Since the total order $\le$ is such that $B_1 \le B_2$ implies $\deg(B_1) \le \deg(B_2)$, the elements of $\mathcal{B}$ can be arranged in a sequence $B_1, B_2, B_3, \ldots$ in such a way that $B_1 < B_2 < B_3 < \ldots$. We will construct, by induction on j (for $j = 0, 1, 2, \ldots$), a finite subset F_j of $\mathcal{B}$, such that:

(a) $\mathrm{card}\,(F_j) = j$,

(b) if we define, for a $u \in U(I)$ of duration t, the $\hat{A}(\underline{X})$-valued function S_u^j, by

$$S_u(t) = S_u^j(t) \cdot \overleftarrow{\prod_{B \in F_j}} \exp\left\{\left(\int_0^t c_B(u)(s)\,\mathrm{d}s\right) \cdot B\right\}, \quad 0 \le t \le T \tag{21}$$

then S_u^j satisfies the differential equation

$$\dot{S}_u^j(t) = S_u^j(t) \cdot \left(\sum_{B \in G_j} c_B(u)(t) B\right), \tag{22}$$

together with the initial condition $S_u^j(0) = 1$, where G_j is a subset of B such that: (i) $B_1 < B_2$ for all $B_1 \in F_j$, $B_2 \in G_j$, and (ii) if B_j^* is the $\le$-first element of G_j, then every $B \in G_j$ such that $B \ne B_j^*$ and $\deg(B) > 1$ satisfies $\lambda(B) \le B_j^*$

For $j = 0$, we let $F_j = \emptyset$, $G_j = \mathcal{B}(1)$, $S_u^j(t) = S_u(t)$. Then all the conditions are satisfied. Now assume F_j has been constructed for a particular j, in such a way that (a) and (b) hold, and let G_j, B_j^*, S_u^j be as specified in (b). We define $F_{j+1} = F_j \cup \left\{B_j^*\right\}$. Since $B_j^* \in G_j$, we have $B_j^* > B$ for every $B \in F_j$. In particular, $B_j^* \notin F_j$, and so $\mathrm{card}\,(F_{j+1}) = j + 1$. The series $S_u^{j+1}(t)$ then satisfies:

$$S_u^j(t) = S_u^{j+1}(t) \cdot \exp\left\{\left(\int_0^t c_{B_j^*}(u)(s)\,\mathrm{d}s\right) B_j^*\right\}. \tag{23}$$

Then

$$\dot{S}_u^j(t) = S_u^j(t) \cdot c_{B_j^*}(u)(t) B_j^* + \dot{S}_u^{j+1}(t) \exp\left\{\left(\int_0^t c_{B_j^*}(u)(s)\,\mathrm{d}s\right) B_j^*\right\}, \tag{24}$$

so that

$$S_u^j(t) \left(\sum_{\substack{B \in G_j \\ B \ne B_j^*}} c_B(u)(t) B\right) = \dot{S}_u^{j+1}(t) \exp\left(C_{B_j^*}(u)(t) B_j^*\right) \tag{25}$$

and then

$$\dot{S}_u^{j+1}(t) = S_u^{j+1}(t) \cdot D_u^j(t) \tag{26}$$

where

$$D_u^j(t) = \exp\left(C_{B_j^*}(u)(t) B_j^*\right) \cdot \left(\sum_{\substack{B \in G_j \\ B \ne B_j^*}} c_B(u)(t) B\right) \cdot \exp\left(-C_{B_j^*}(u)(t) B_j^*\right). \tag{27}$$

Then

$$D_u^j(t) = \sum_{m=0}^{\infty} \sum_{\substack{B\in G_j \\ B\neq B_j^*}} \frac{1}{m!} \left\{C_{B_j^*}(u)(t)\right\}^m c_B(u)(t) E_{m,B}, \tag{28}$$

where

$$E_{m,B} = \left(\mathrm{ad}\, B_j^*\right)^m (B). \tag{29}$$

The brackets $E_{m,B}$, for $B \in G_j$, $B \neq B_j^*$, and $m = \{0, 1, \ldots\}$, are in $\mathcal{B}$. This is clear for $m = 0$, because $G_j \subseteq \mathcal{B}$. For $m = 1$, $E_{m,B}$ is the bracket $\left[B_j^*, B\right]$. Both B_j^* and B are in $\mathcal{B}$. Moreover, $B_j^* < B$, because B_j^* was the $\leq$-first element of G_j. Finally, the inductive assumption (item (ii) of (b)) tells us that, if $\deg(B) > 1$, then the left factor of B satisfies $\lambda(B) \leq B_j^*$. Hence the definition of the P. Hall basis shows that $E_{m,B} \in \mathcal{B}$. Assume that $E_{m,B} \in \mathcal{B}$ for some $m \geq 1$. Then $E_{m+1,B} = \left[B_j^*, E_{m,B}\right]$, and both B_j^* and $E_{m,B}$ are in $\mathcal{B}$. Moreover, $\deg\left(B_j^*\right) < \deg\left(E_{m,B}\right)$, and so $B_j^* < E_{m,B}$. Finally, the left factor of $E_{m,B}$ is B_j^*, and so $\lambda\left(E_{m,B}\right) \leq B_j^*$. So $E_{m+1,B} \in \mathcal{B}$. This establishes that $E_{m,B} \in \mathcal{B}$ for all m and all $B \in G_j - \left\{B_j^*\right\}$.

Next we show that the $E_{m,B}$ are all different, i.e. that $E_{m,B}$ cannot be equal to $E_{m',B'}$ unless $m = m'$ and $B = B'$. To see this, notice that each $B' \in \mathrm{Br}(\underline{X})$ can be written in a unique way as $\left(\mathrm{ad}\, B_1\right)^m (B_2)$, for some m, B_1, B_2 such that $B_1 \neq B_2$. (We let $B_1 = \lambda\left(B'\right)$; then we define H^i, K^i by $H^1 = B_1$, $K^1 = B_2, H^{i+1} = \lambda\left(K^i\right), K^{i+1} = \rho\left(K^i\right)$. The integer m is then the first i such that $H^i \neq H^{i+1}$, and B_2 can then be recovered, since $B_2 = K^m$.) If we apply this when $B' = E_{m,B}$, we see that m and B can be recovered from $E_{m,B}$. Since the $E_{m,B}$ are all different, and the coefficient of $E_{m,B}$ in (28) is obviously $c_{E_{m,B}}(u)(t)$, we can rewrite (28) as

$$D_u^j(t) = \sum_{B\in G_{j+1}} c_B(u)(t) B \tag{30}$$

where G_{j+1} is the set whose elements are the $E_{m,B}$. We must now show that F_{j+1}, G_{j+1} satisfy conditions (i), (ii) of (b). Let $B_1 \in F_{j+1}, B_2 \in G_{j+1}$. Then $B_2 = E_{m,B}$ for some m and some $B \in G_j, B \neq B_j^*$. If $m = 0$, then $B > B_j^*$ because B_j^* is the $\leq$-first element of G_j. So $B_2 > B_j^*$. If $m > 0$, then $\lambda\left(B_2\right) = B_j^*$, and so $\deg\left(B_2\right) > \deg\left(B_j^*\right)$, and therefore $B_2 > B_j^*$. In either case we have shown that $B_2 > B_j^*$. If $B_1 = B_j^*$, then $B_2 > B_1$. If $B_1 \in F_j$, then $B_2 > B_1$ as well, because $B_j^* \in G_j$ and so $B_1 < B_j^*$. So (i) holds. We now prove (ii). Let B_{j+1}^* be the $\leq$-first element of G_{j+1}. Let $B \in G_{j+1}, B \neq B_{j+1}^*$. Then $B = E_{m,B'}$ for some m and some $B' \in G_j$. Assume that $\deg(B) > 1$. If $m = 0$, then $B = B'$, and so $B \in G_j$ and $\deg(B) > 1$. Therefore $\lambda(B) \leq B_j^*$ by the inductive hypothesis. If $m > 0$, then $\lambda(B) = B_j^*$, and so $\lambda(B) \leq B_j^*$ as well. So $\lambda(B) \leq B_j^*$ in either case. Since $B_j^* < B_{j+1}^*$, we see that $\lambda(B) \leq B_{j+1}^*$, and the induction is complete.

So we have completed the inductive construction of the F_j. The construction shows that each F_{j+1} is obtained from F_j by adjoining an element B_j^* of $\mathcal{B}$ which does not belong to F_j and satisfies $B_j^* > B$ for all $B \in F_j$. If we let $B_j = B_{j-1}^*$, we see that the B_j are elements of $\mathcal{B}$ that satisfy $B_1 < B_2 < B_3 < \cdots$, and $F_j = \{B_1, B_2, \ldots, B_j\}$. Then (21) becomes

$$S_u(t) = S_u^j(t) \cdot P_u^j(t) \tag{31}$$

where

$$P_u^j(t) = \overleftarrow{\prod_{i \le i \le j}} \exp\left(C_{B_i}(u)(t) B_i\right). \tag{32}$$

If we let $\delta_j = \deg\left(B_j^*\right)$, it is clear that $S_u^j(t) = 1 + Q_u^j(t)$, where $Q_u^j(t)$ contains no monomials of degree $< \delta_j$. Let $R_u^0(t) = P_u^1(t)$, and $R_u^j(t) = P_u^{j+1}(t) - P_u^j(t)$ for $j \ge 1$. Then $R_u^j = Q_u^j P_u^j - Q_u^{j+1} P_u^{j+1}$, so that $R_u^j(t)$ only contains monomials of degree $\ge \delta_j$. Since $\delta_j \to \infty$ as $j \to \infty$, the sum $\sum_{j=0}^{\infty} R_u^j(t)$ converges, and so the infinite product $\overleftarrow{\prod}_{1 \le i < \infty} \exp\left(C_{B_i}(u)(t) B_i\right)$ converges as well. Since

$$S_u(t) = P_u^j(t) + Q_u^j(t) P_u^j(t), \tag{33}$$

we see that the coefficients of $S_u(t)$ and $P_u^j(t)$ agree up to degree $\delta_j - 1$. Therefore

$$S_u(t) = \overleftarrow{\prod_{1 \le i < \infty}} \exp\left(C_{B_i}(u)(t) B_i\right) \tag{34}$$

Equivalently, we have

$$S_u(t) = \overleftarrow{\prod_{B \in F}} \exp\left(C_B(u)(t) B\right) \tag{35}$$

where F is the set whose elements are $B_1, B_2, \ldots$, ordered by the restriction of $\le$ to F.

To complete our proof, we must show that $F = \mathcal{B}$. This can be established in a number of ways. We choose to use the Poincaré-Birkhoff-Witt theorem, according to which the elements of $A(\underline{X})$ that are of the form

$$B_1^{m_1} B_2^{m_2} \ldots B_k^{m_k} \tag{36}$$

with $B_1, \ldots, B_k$ in B, and $B_1 > B_2 > \ldots > B_k$, form a basis of $A(\underline{X})$. Let H denote the linear span of the elements of the form (36) that are such that $B_1, \ldots, B_k$ are in F. If $F \ne \mathcal{B}$, then H is a proper subspace of $A(\underline{X})$. We show that $H = A(\underline{X})$. Fix a multiindex $\sigma = (i_1, \ldots, i_r) \in \mathcal{I}(I)$. Let $A_r(\underline{X})$ be the quotient algebra obtained by setting all monomials of degree $> r$ equal to zero, and let $\pi_r : A(\underline{X}) \to A_r(\underline{X})$ be the canonical homomorphism. We can also regard $A_r(\underline{X})$ as a linear subspace of $A(\underline{X})$. Let $H_r = H \cap A_r(X)$. It is clear from (35) that $\pi_r\left(S_u(t)\right) \in H_r$ for every u, t. In particular, the product

$$\pi_r\left(\exp\left(t_1 X_{i_1}\right) \ldots \exp\left(t_r X_{i_r}\right)\right) \tag{37}$$

belongs to H_r for every $t_1, \ldots, t_r$. If we differentiate (37) with respect to $t_1, t_2, \ldots, t_r$ and then set $t_1 = \ldots = t_r = 0$, we find that $X_\sigma \in H_r$. So $X_\sigma \in H$. Since σ was arbitrary, we conclude that $H = A(\underline{X})$. So $F = \mathcal{B}$ and our proof is complete. □

We now turn to the question of the convergence of the product expansion.

Formula (19) makes it possible to prove bounds for the coefficients $C_B(u)(t)$. Assume that I is a finite set. For a $u \in U(I)$ with duration T, let $u = \{u_i(\cdot) : i \in I\}$, and define $h_u(t) = \max\left\{|u_i(t)| : i \in I\right\}, H_u(t) = \int_0^t h_u(s)\,\mathrm{d}s$. Then an easy induction gives the bounds

$$|c_B(u)(t)| \le H_u(t)^{(\deg B) - 1} h(t) \tag{38}$$

and then

$$|C_B(u)(t)| \le \frac{H_u(t)^{\deg B}}{\deg B}. \tag{39}$$

These bounds are sufficient to prove that the product expansion (20) converges in a slightly stronger sense. Suppose that $A = \{A_i : i \in I\}$ are square $n \times n$ matrices or, more generally, bounded operators on some Banach space. Let M be an upper bound for the operator norms $\|A_i\|, i \in I$. For each bracket $B \in \mathrm{Br}(\underline{X})$, let B_A denote the operator obtained by substituting the A_i for the X_i in B. Then $\|B_A\| \le 2^{(\deg B)-1} M^{\deg B}$ for all B. For a given n, there are at most ν^n elements of $\mathcal{B}$ of degree n. Hence the series

$$\sum_B C_B(u)(t) B_A \tag{40}$$

converges absolutely and uniformly on every interval $0 \le t \le \bar{t}$ such that

$$\sum_{n=0}^{\infty} \frac{1}{n} \nu^n 2^{n-1} M^n (H_u(\bar{t}))^n < \infty. \tag{41}$$

In particular, the series (40) converges on $[0, \bar{t}]$ if $2M\nu H_u(\bar{t}) < 1$. Since the $u_i(\cdot)$ are integrable, (40) converges, for each u, on some interval $[0, \bar{t}]$, $\bar{t} > 0$.

On the other hand, it is easy to see that, whenever (40) converges absolutely, the infinite product $\prod_{B \in \mathcal{B}} \exp(C_B(u)(t) B_A)$ converges as well. So, for every choice of the A_i, and every u, the infinite product converges for sufficiently small times (or. for a fixed time interval $[0, \bar{t}]$, for sufficiently small controls u, the smallness being measured by the L^1 norm).

The presence of a factorial in (19) might be regarded as an indication that (39) is not the best possible bound for the $C_B(u)$, and that perhaps the $C_B(u)$ decrease faster than geometrically, and the series actually converges for all u and all times. We show that this is not so by means of an example.

In order to construct our example, we take a set I that has two elements, and let X, Y denote the two indeterminates. We let $\mathcal{B}$ be an arbitrary P. Hall basis of $L(\underline{X})$. We take $u \in U(I)$ such that both functions $u_i(\cdot)$ are identically equal to one. Then the functions c_x, c_y are equal to 1. An easy induction shows that, for each $B \in \mathrm{Br}(\underline{X})$, the coefficient $c_B(t)$ is given by

$$c_B(t) = \frac{t^{\deg B - 1}}{A_B} \tag{42}$$

where A_B is a positive constant. The constants A_B can be computed recursively by the formula

$$A_B = A_{B_1}^m n_1^m m! A_{B_2}, \tag{43}$$

valid if $B = (\mathrm{ad}\, B_1)^m B_2$, $\lambda(B_2) \ne B_1$, and $n_i = \deg B_i$ for $i = 1, 2$.

For $n \ge 1$, we let α_n denote the smallest of the A_B, for all $B \in \mathcal{B}$ such that $\deg B = n$. We will prove that the upper bound

$$\alpha_n \le K \gamma^n \tag{44}$$

holds for some $K > 0$, $\gamma > 0$. This implies that we can choose a sequence $B_1, B_2, B_3, \ldots$ of elements of $\mathcal{B}$, with deg $B_n = n$, such that the coefficients $c_{B_n}(t)$ satisfy $c_{B_n}(t) \geq \frac{\gamma^{-n} t^{n-1}}{K}$, so that $C_{B_n}(t) \geq \frac{1}{nK}\left(\frac{t}{\gamma}\right)^n$. Therefore, the series $\sum_n C_{B_n}(t)$ does not converge if $t > \gamma$.

To prove the bound (44) we begin by defining, for $k = 2, 3, 4, \ldots$, a set S_k of three elements of $\mathcal{B}$ of degree 2^k. We start by letting S_2 be the set whose elements are the three members of $\mathcal{B}$ that have degree 4 (i.e. $[X, [X, [X, Y]]]$, $[Y, [X, [X, Y]]]$ and $[Y, [Y, [X, Y]]]$). If S_k has been defined, we let S_{k+1} consist of the three brackets of the form $[B, B']$ with $B \in S_k$, $B' \in S_k$, $B < B'$. It is clear that $S_{k+1} \subseteq \mathcal{B}$. Next we let β_k denote the maximum of the constants A_B for $B \in S_k$. If $B \in S_{k+1}$, we can write

$$B = [B', B''], \quad B' \in S_k, B'' \in S_k. \tag{45}$$

Since $\deg B' = \deg B''$, the left factor of B'' cannot be equal to B'. Hence

$$A_B = A_{B'} A_{B''} 2^k. \tag{46}$$

Therefore,

$$\beta_{k+1} \leq 2^k \beta_k^2. \tag{47}$$

Then

$$\log \beta_{k+1} \leq 2 \log \beta_k + k \log 2. \tag{48}$$

Let $\theta_k = 2^{-k} \log \beta_k$. Then

$$\theta_{k+1} \leq \theta_k + 2^{-k-1} k \log 2 \tag{49}$$

Therefore

$$\theta_k \leq \theta_2 + \sum_{j=2}^{k-1} 2^{-j-1} j \log 2. \tag{50}$$

since the series $\sum_j 2^{-j-1} j \log 2$ converges, there is a constant η such that $\theta_k \leq \eta$ for all k. So $\log \beta_k \leq \eta 2^k$, and then $\beta_k \leq \exp\left(\eta 2^k\right)$ for all k. So, if we let $\gamma = e^n$, we see that

$$\beta_k \leq \gamma^{2^k} \quad \text{for } k = 2, 3, \ldots. \tag{51}$$

Since $\alpha_{2^k} \leq \beta_k$, we conclude that $\alpha_n \leq \gamma^n$ for all n that are of the form 2^k, $k = 2, 3, \ldots$. By taking a larger γ, if necessary, we can assume that the inequality also holds for $k = 0, 1$. So we have proved the desired bound, except for the fact that, so far, only values of n of the form 2^k have been considered.

In order to extend the bound (44) to arbitrary values of n, we pick an n and express it as

$$n = 2^k + \varepsilon_1 2^{k-1} + \varepsilon_2 2^{k-2} + \ldots + \varepsilon_{k-1} 2 + \varepsilon_k \tag{52}$$

where each ε_j is either 0 or 1. We pick a $B \in S_k$ and modify it by suitable insertions to get an element of $\mathcal{B}$ that has degree n. Write $B = [B_1^\#, B_2]$, where $B_1^\# \in S_{k-1}$, $B_2 \in S_{k-1}$ and $B_1^\# < B_2$. If $\varepsilon_1 = 1$, let $B_1^* = [B_1^\#, [B_1^\#, B_2]]$. Then

$$A_{B_1^*} = 2A_{B_1^\#}^2 A_{B_2} 2^{k-1} \leq 2^k \beta_{k-1}^3 \tag{53}$$

and so

$$A_{B_1^*} \leq 2^k \gamma^{3 \cdot 2^{k-1}} = 2^k \gamma^{2^k + \varepsilon_1 2^{k-1}}. \tag{54}$$

If $\varepsilon_1 = 0$, define $B_1^* = B$. Then it is clear that (54) holds in this case as well.

We now have

$$B_1^* = \left(\mathrm{ad}\, B_1^\#\right)^{\varepsilon_1 + 1} (B_2)\,. \tag{55}$$

Write $B_2 = [B_2^\#, B_3]$, where $B_2^\#$ and B_3 are in S_{k-2}. Define

$$B_2^* = \left(\mathrm{ad}\, B_1^\#\right)^{\varepsilon_1 + 1} \left(\mathrm{ad}\, B_2^\#\right)^{\varepsilon_2 + 1} (B_3)\,. \tag{56}$$

If $\varepsilon_2 = 1$, then

$$A_{B_2^*} = 2A_{B_2^\#} 2^{k-2} A_{B_1^*}, \tag{57}$$

so that

$$A_{B_2^*} \leq 2^{k-1} \beta_{k-2} 2^k \gamma^{2^k + \varepsilon_1 2^{k-1}} \tag{58}$$

and then

$$A_{B_2^*} \leq 2^k 2^{k-1} \gamma^{2^k + \varepsilon_1 2^{k-1} + \varepsilon_2 2^{k-2}}. \tag{59}$$

If $\varepsilon_2 = 0$ then (59) holds as well.

This procedure can be continued, until we get an element B_{k-2}^* of degree $2^k + \varepsilon_1 2^{k-1} + \ldots + 4\varepsilon_{k-2}$ which satisfies

$$A_{B_{k-2}^*} \leq 2^k 2^{k-1} \cdots 2^3 \gamma^{2^k + \varepsilon_1 2^{k-1} + \ldots + 4\varepsilon_{k-2}} \tag{60}$$

and is of the form $(\mathrm{ad}\, B_1^\#)^{\varepsilon_1+1} \cdots (\mathrm{ad}\, B_{k-2}^\#)^{\varepsilon_{k-2}+1} (B_{k-1})$, where $B_{k-1} \in S_2$. Let $\nu = 2\varepsilon_{k-1} + \varepsilon_k$. Then $B_{k-1} = (\mathrm{ad}\, Z)^\lambda (\mathrm{ad}\, X)^\mu (Y)$, where either $\mu = 1, \lambda = 1$ and $Z = [X, Y]$, or $Z = Y$, $\mu \geq 1$, and $\mu + \lambda = 3$. Define $\tilde{B} = (\mathrm{ad}\, Z)^\lambda (\mathrm{ad}\, X)^{\mu+\nu} (Y)$, and then let $(\mathrm{ad}\, B_1^\#)^{\varepsilon_1+1} \cdots (\mathrm{ad}\, B_{k-2}^\#)^{\varepsilon_{k-2}+1} \left(\tilde{B}\right)$. Then $B^* \in \mathcal{B}$, $\deg B^* = n$, and

$$A_{B^*} = \frac{(\mu + \nu)!}{\mu!} A_{B_{k-2}^*}. \tag{61}$$

Since $1 \leq \mu \leq 3$ and $0 \leq \nu \leq 3$, we have $A_{B*} \leq 120 A_{B_{k-2}^*}$. Hence

$$A_B* \leq 15 \times 2^{1+2+\ldots+k} \times \gamma^n \tag{62}$$

so that

$$A_{B^*} \leq 15 \times 2^{\frac{1}{2}k(k+1)} \times \gamma^n \tag{63}$$

Since $2^{k+1} > n \geq 2^k$, we have

$$A_{B^*} \leq 15n^{\frac{1}{2}(k+1)}\gamma^n \tag{64}$$

and then

$$A_{B^*} \leq 15n^{\frac{1}{2}+\sigma \log n}\gamma^n, \tag{65}$$

where $\sigma = \frac{1}{2\log 2}\cdot$ Clearly,

$$15n^{\frac{1}{2}+\sigma \log n} \leq Ke^{\rho n} \tag{66}$$

for some K, ρ that do not depend on n. Hence, if we replace γ by $e^\rho\gamma$, we get

$$A_{B^*} \leq K\gamma^n.$$

So $\alpha_n \leq K\gamma^n$, and (44) is proved.

References

[1] Hall, M., *Theory of groups*. The MacMillan Company, 1959.

[2] Serre, J.P., *Lie groups and Lie algebras*. W.A. Benjamin, Inc, 1965.

[3] Viennot, G., *Algèbres de Lie libres et monoides libres*. Springer Verlag, 1978.