Free Lie Algebras

Christophe Reutenauer

Département de mathématiques, Université du Québec à Montréal, Case postale 8888, succursale Centre-Ville. Montréal (Québec), Canada H3C 3P8
E-mail: christo@math.uqam.ca

Contents

HANDBOOK OF ALGEBRA, VOL. 3
Edited by M. Hazewinkel

1. Introduction

Lie polynomials appeared at the end of the 19th century and the beginning of the 20th century in the work of Campbell, Baker and Hausdorff on exponential mapping in a Lie group, which has lead to the so-called Campbell–Baker–Hausdorff formula. Around 1930, Witt showed that the Lie algebra of Lie polynomials is actually the free Lie algebra, and that its enveloping algebra is the associative algebra of noncommutative polynomials. He proved what is now called the Poincaré–Birkhoff–Witt theorem, and showed how the free Lie algebra is related to the lower central series of a free group. About at the same time P. Hall and Magnus, with their commutator calculus, opened the way to bases of the free Lie algebra, by M. Hall.

In this chapter, our aim is to prove that the Lie algebra of Lie polynomials is indeed the free Lie algebra. We shall do this without using the Poincaré–Birkhoff–Witt theorem, but by constructing Hall bases of the free Lie algebras. The bases we consider are more general than the original ones; they include the Lyndon basis. We follow the method of Schützenberger (with improvements from Melançon). We shall also follow him to compute the dual basis.

In the Notes, we give statement of other results, and references for them, together with indications on recent work related to free Lie algebras.

In all that follows, **A** *is a commutative ring with unit.* In Section 14, we assume that it contains **Q**.

2. Free Lie algebra

Recall that a *Lie algebra* over **A** is an **A**-module L equipped with a bilinear mapping $L \times L \to L$, $(x, y) \mapsto [x, y]$, satisfying the two following properties, for any x, y, z in L:

$$[x, x] = 0,$$
$$[[x, y], z] + [[y, z], x] + [[z, x], y] = 0.$$

The latter identity is called the *Jacobi identity*. Note that the first one implies *antisymmetry*, that is,

$$[x, y] = -[y, x],$$

since, using bilinearity, we have $0 = [x + y, x + y] = [x, x] + [x, y] + [y, x] + [y, y] = [x, y] + [y, x]$. In view of antisymmetry, we may rewrite the Jacobi identity as

$$[[x, y], z] = [x, [y, z]] + [[x, z], y].$$

This identity, which will be useful later, means that $x \mapsto [x, z]$ is a derivation of the Lie algebra L.[1]

[1] A *derivation* of L is a linear mapping $L \to L, x \mapsto x'$, that satisfies the Leibnitz identity, that is $[x, y]' = [x, y'] + [x', y]$.

Homomorphisms, isomorphisms and Lie subalgebras of Lie algebras are defined as usual.

Given a set X, a *free Lie algebra* on X over $\mathbf{A}$ is a Lie algebra L over $\mathbf{A}$, together with a mapping $i : X \to L$, with the following *universal property*: for each Lie algebra K and each mapping $f : X \to K$, there exists a unique Lie algebra homomorphism $F : L \to K$ such that $f = F \circ i$.

A standard argument shows that a free Lie algebra on X is necessarily unique, up to Lie algebra isomorphism.

Its existence is shown as follows: a *tree* on X is a well-formed expression on X, recursively defined by: each x in X is a tree; if t_1, t_2 are trees, then so is $t = (t_1, t_2)$ (the terminology comes from the fact that each tree may be identified with a binary, complete, rooted and planary tree with leaves labelled by X). Let $M(X)$ denote the set of all trees on X. The mapping which sends the pair of trees t_1, t_2 to t as above is a binary law on $M(X)$. Let $D(X)$ denote the free $\mathbf{A}$-module with basis $M(X)$. Then the previous law defines a bilinear mapping $D(X) \times D(X) \to D(X)$, still denoted $(\ ,\ )$. An ideal I of $D(X)$ is a submodule such that for any $u \in I$, $v \in M(X)$, one has $(u, v), (v, u) \in I$. Let I be the ideal generated by all the elements of the form $[t, t]$, or $[[t_1, t_2], t_3] + [[t_2, t_3], t_1] + [[t_3, t_1], t_2]$, $t, t_1, t_2, t_3 \in M(X)$. Then the quotient $D(X)/I$ is clearly a Lie algebra. Furthermore, it is the free Lie algebra on X over $\mathbf{A}$. Indeed, define $i : X \to D(X)/I$ by $i(x) = x \bmod I$; now, if K is a Lie algebra with a mapping $f : X \to K$, then f clearly extends uniquely to a mapping $M(X) \to K$, by replacing each tree by the corresponding Lie bracketing, with $x \in X$ replaced by $f(x)$. By linear extension, this mapping extends to $D(X)$; the kernel of the latter clearly contains I, since K is a Lie algebra, hence we obtain a Lie homomorphism $F : D(X)/I \to K$, which satisfies $f = F \circ i$.

In other words, a Lie algebra over $\mathbf{A}$ is free on X if and only if it is generated by all possible Lie bracketings of elements of X (these bracketings are in bijection with trees), and if the only possible relations existing among these bracketings are consequence of the bilinearity of the bracketing, of the Jacobi identity and the identity $[u, u] = 0$.

3. Noncommutative polynomials

Given a set X of *noncommuting variables* (also called *letters*), denote by X^* the set of words on X, where by *word* we mean a finite sequence of elements of X, or equivalently, a *noncommutative monomial*; we include the *empty word*, denoted 1. With the product defined by concatenation of words, X^* becomes a monoid, which is the *free monoid* on X.

Indeed, for any mapping $f : X \to M$, where M is any monoid, there is a unique homomorphism of monoids $F : X^* \to M$ such that $f = F \circ i$, where i is the natural injection $X \to X^*$.

A *noncommutative polynomial* (we shall say polynomial) is a linear combination of words with coefficients in $\mathbf{A}$. The set they form is denoted $\mathbf{A}\langle X\rangle$. With the product inherited from the free monoid, it becomes an $\mathbf{A}$-algebra; note that it is the $\mathbf{A}$-algebra of the monoid X^*. Equivalently, it is the *tensor algebra* of the free $\mathbf{A}$-module with basis X.

This $\mathbf{A}$-algebra $\mathbf{A}\langle X\rangle$ is the *free associative* $\mathbf{A}$*-algebra* on X; indeed, for any mapping $f : X \to A$, where A is any $\mathbf{A}$-algebra, there is a unique homomorphism of $\mathbf{A}$-algebras $F : \mathbf{A}\langle X\rangle \to A$ such that $f = F \circ i$, where i is the natural injection $X \to \mathbf{A}\langle X\rangle$.

4. Lie polynomials

For any polynomials P, Q in $\mathbf{A}\langle X\rangle$, define their *Lie bracket* by $[P, Q] = PQ - QP$. This defines a structure of Lie algebra on $\mathbf{A}\langle X\rangle$. Indeed, antisymmetry is immediate and a simple calculation shows that the Lie bracket satisfies the Jacobi identity.

Let $L_{\mathbf{A}}(X)$ (or simply $L(X)$) denote the smallest Lie subalgebra of $\mathbf{A}\langle X\rangle$ that contains each letter in X. It will be shown that this Lie algebra is free. A *Lie polynomial* is an element of $L(X)$.

5. Hall sets

Each tree t in the free magma $M(X)$ is either a letter $t \in X$, or is of the form $t = (t_1, t_2)$, for some trees t_1, t_2. We then write $t' = t_1$, $t'' = t_2$, and call t', t'' the *left, right immediate subtree* of t.

Consider a total order $>$ on $M(X)$ such that for any tree t, one has $t < t''$. Note that such orders surely exist. We define the *Hall set* H relative to this order recursively by: each letter is in H; if a tree t is not a letter, then it is in H if and only if t', t'' are in H, $t' < t''$ and: either t' is in X or $(t')'' \geqslant t''$. Elements of H are called *Hall trees*.

In the sequel, we fix a Hall set H.

6. Standard sequences

Given a Hall set H, a *standard sequence* is a sequence of Hall trees $(t_1, \ldots, t_n)$ with $n \geqslant 0$ and: for any i, either t_i is a letter, or $t_i'' \geqslant t_{i+1}, \ldots, t_n$.

Clearly, each sequence of letters is standard. Moreover, a *decreasing sequence* (that is, a sequence such that $t_1 \geqslant \cdots \geqslant t_n$) of Hall trees is always standard: indeed, if t_i is not a letter, then $t_i'' > t_i \geqslant t_{i+1}, \ldots, t_n$.

We call a *rise* of a sequence an index i such that $t_i < t_{i-1}$; in that case, we also say that (t_i, t_{i+1}) is a rise. An *inversion* of the sequence is a couple (i, j) such that $i < j$ and $t_i < t_j$ (note the opposite inequality to the classical one for inversions of permutations); here also, we say that (t_i, t_j) is an inversion. Observe that a sequence is decreasing if and only if it has no inversion, or equivalently, no rise. A *legal rise* is a rise i such that $t_{i-1} \geqslant t_{i+2}, \ldots, t_n$.

We now define a rewriting system on the set of standard sequences. Let $s = (t_1, \ldots, t_n)$ be a standard sequence and i some legal rise of s. We write $s \to s'$ if $s' = (t_1, \ldots, t_{i-1}, (t_i, t_{i+1}), t_{i+2}, \ldots, t_n)$. In other words, s' is obtained from s by multiplying in the free magma $M(X)$ the two trees that form the chosen legal rise. Note that s' is a sequence of Hall trees: indeed, either t_i is in X, or $t_i = (t_i', t_i'')$ and then $t_i'' \geqslant t_{i+1}$ since s is standard; this shows that (t_i, t_{i+1}) is in H since $t_i < t_{i+1}$, i being a rise. Furthermore, s' is standard: indeed, if $j = 1, \ldots, i-1$, then either t_j is a letter, or $t_j'' \geqslant t_{i+1}$ (since s is standard) $> (t_i, t_{i+1})$, by assumption on the order $>$. Moreover, $t_{i+1} \geqslant t_{i+2}, \ldots, t_n$, since i is a legal rise; thus s' is a standard sequence, since s is.

We denote by $\to^*$ the reflexive and transitive closure of the binary relation $\to$.

THEOREM 1.

(1) *For any standard sequences s, s_1, s_2 such that $s \to^* s_1$ and $s \to^* s_2$, there exists a standard sequence r such that $s_1 \to^* r$ and $s_2 \to^* r$.*

(2) *For each standard sequence s, there exists a sequence of letters r such that $r \to^* s$.*

(3) *For each standard sequence s, there exists a decreasing standard sequence t such that $s \to^* t$.*

PROOF. (1) Assume that $s \to s_1$ and $s \to s_2$. Write $s = (t_1, \ldots, t_n)$, $s_1 = (t_1, \ldots, t_{i-1}, (t_i, t_{i-1}), t_{i+2}, \ldots, t_n)$ and $s_2 = (t_1, \ldots, t_{j-1}, (t_j, t_{j+1}), t_{j+2}, \ldots, t_n)$, where i, j are legal rises of s. We may assume that $i < j$. Then $i + 1 < j$: indeed, otherwise, $i + 1 = j$; hence $t_{i+1} = t_j$, and $t_{i+1} < t_{j+1}$ since j is a rise; moreover, i is a legal rise, so that $t_{i+1} \geqslant t_{i+2} = t_{j+1}$, a contradiction.

Thus (t_i, t_{i+1}) is a rise of s_2 and (t_j, t_{j+1}) is a rise of s_1. We show that (t_j, t_{j+1}) is a legal rise of s_1, and that (t_i, t_{i+1}) is a legal rise of s_2. The first assertion is clear, since if t is at the right of t_{j+1} in s_1, then it is also in s; hence $t_{j+1} \geqslant t$, because j is a legal rise of s. For the second assertion, we have $t_{i+1} \geqslant t_{j+1}$ (since i is a legal rise of s) $> (t_j, t_{j+1})$ by the property of the order; hence t_{i+1} is greater or equal to each tree at its right in the sequence s_2, since this is true in s.

Define now $r = (t_1, \ldots, t_{i-1}, (t_i, t_{i+1}), t_{i+2}, \ldots, t_{j-1}, (t_j, t_{j+1}), t_{j+2}, \ldots, t_n)$. Then we have $s_1 \to r$ and $s_2 \to r$ by definition of $\to$.

This being done, the first assertion of the theorem follows by a straightforward induction on the lengths of the chains from s to s_1 and s to s_2.

(2) Let $s = (t_1, \ldots, t_n)$. If s is not a sequence of letters, then consider i such that t_i is not a letter, and for any $j = 1, \ldots, i-1$, t_j is a letter: such an i surely exists. Then let $r' = (t_1, \ldots, t_{i-1}, t_i', t_i'', t_{i+1}, \ldots, t_n)$. The sequence r' is standard: indeed, either t_i' is a letter, or $(t_i')'' \geqslant t_i''$ (since (t_i', t_i'') is a Hall tree) $\geqslant t_{i-1}, \ldots, t_n$, because s is standard; moreover, either t_i'' is a letter, or $(t_i'')'' > t_i''$ (by the property of the order) $\geqslant t_{i+1}, \ldots, t_n$; finally, for $j = 1, \ldots, i-1$, t_j is a letter. Thus r' is standard since s is. Moreover, (t_i', t_i'') is a rise of r', because $t_i' < t_i''$ (since (t_i', t_i'') is a Hall tree), which is legal, since $t_i'' \geqslant t_{i+1}, \ldots, t_n$. Thus $r' \to s$. We conclude by induction on the maximum degree of the trees in s.

(3) If s is not decreasing, it has at least one rise. Choose i to be the rightmost one. Then $t_i < t_{i+1} \geqslant t_{i+2} \geqslant \cdots \geqslant t_n$. Hence this rise is legal and we obtain a shorter sequence s' such that $s \to s'$. We conclude by induction on the length of the sequence. □

7. Hall words

There is a canonical mapping from $M(X)$ onto X^*, denoted f; it is defined inductively by $f(x) = x$ if $x \in X$, and $f(t) = f(t')f(t'')$ if $t = (t', t'')$. For example, $f((y, ((x, y), z))) = yxyz$.

THEOREM 2. *Each word in X^* has a unique factorization $f(t_1) \ldots f(t_n)$, with $n \geqslant 0$, $t_i \in H$ and $t_1 \geqslant \cdots \geqslant t_n$.*

PROOF. Extend the mapping f to sequences by $f(t_1, \ldots, t_n) = f(t_1) \ldots f(t_n)$. Clearly, $s \to^* r$ implies $f(s) = f(r)$. For a word w, the sequence s of its letters is standard. By

Theorem 1, we have $s \to^* t$ for some decreasing sequence of Hall trees; hence w admits a factorization as in the statement.

Suppose it has another such factorization: $w = f(h_1) \dots f(h_p)$. Consider the standard sequences $s = (t_1, \dots, t_n)$ and $r = (h_1, \dots, h_p)$. By the theorem, we find sequences of letters s', r' such that $s' \to^* s$ and $r' \to^* r$. We have $f(s') = f(s) = w = f(r) = f(r')$. Thus $s' = r'$. Hence by Theorem 1 again, we have $s \to^* u$ and $r \to^* u$ for some standard sequence u. Since s, r are nondecreasing, hence have no rise, this is possible only if $s = r$. This proves unicity. □

Given a Hall set H, we call *Hall word* each word of the form $f(t)$, $t \in H$.

COROLLARY 1. *For each Hall word w, there exists exactly one Hall tree t such that $w = f(t)$.*

COROLLARY 2. *Each word has exactly one decreasing factorization into Hall words.*

8. Poincaré–Birkhoff–Witt basis

There is a canonical mapping g from $M(X)$ into $L(X)$. It is defined by $g(x) = x$ if $x \in X$, and $g(t) = [g(t'), g(t'')]$ if $t = (t', t'')$. In other words, $g(t)$ is obtained by replacing in t parentheses by Lie bracketing. For example, $g((y, ((x, y), z))) = [y, [[x, y], z]]$.

THEOREM 3. *The products $g(t_1) \dots g(t_n)$, $n \geqslant 0$, $t_1 \geqslant \dots \geqslant t_n$, form a basis of the* **A***-module* $\mathbf{A}\langle X \rangle$.

PROOF. It is enough to show this when X is finite.

For each standard sequence $s = (t_1, \dots, t_n)$, with a legal rise i, define

$$\lambda_i(s) = \big(t_1, \dots, t_{i-1}, (t_i, t_{i+1}), t_{i+2}, \dots, t_n\big)$$

and

$$\rho_i(s) = (t_1, \dots, t_{i-1}, t_{i+1}, t_i, t_{i+2}, \dots, t_n).$$

The first sequence is the same as for the definition of the rewriting system in Section 6, and is obtained by multiplying in $M(X)$ the two trees that form the legal rise; the second is obtained by interchanging them.

Extend naturally g to sequences by: $g(s) = g(t_1) \dots g(t_n)$ if $s = (t_1, \dots, t_n)$. Since clearly $ab = [a, b] + ba$, we obtain that $g(s) = g(\lambda_i(s)) + g(\rho_i(s))$.

Note that $\lambda_i(s)$ is shorter than s, and that $\rho_i(s)$ has one inversion less.

Thus we obtain by induction that $g(s)$ is equal to a sum of products as in the statement. This being true for each sequence of letters, we deduce that these elements span the module $\mathbf{A}\langle X \rangle$.

Note that these elements are homogeneous. Now Corollary 2 shows that for fixed degree d, the elements that have degree d are as numerous as the words of length d; but the latter form a basis of the submodule of homogeneous polynomials of degree d. Hence the polynomials of the statement which are of degree d form a basis of that submodule: indeed, in a free $\mathbf{A}$-module of rank N, if a generating set has cardinality N, then it is a basis (since a right invertible square matrix over $\mathbf{A}$ is invertible).

This show that the whole collection of polynomials as in the statement form a basis of $\mathbf{A}\langle X\rangle$. □

COROLLARY 3. *Each word, when written in the basis of the theorem, has coefficients in* $\mathbf{N}$.

9. Hall bases

We can now prove that $L(X)$ is a free $\mathbf{A}$-module with basis $g(H)$.

THEOREM 4. *The elements $g(t)$, $t \in H$, form a basis of $L(X)$.*

PROOF. By Theorem 3, the polynomials $g(t)$ are linearly independent. It is enough to show that they span $L(X)$. Again, we may suppose that X is finite. Since $L(X)$ is generated as a Lie algebra by X, and X is contained in $g(H)$, it is enough to show that: for any t_1, t_2 in H, the polynomial $g((t_1, t_2)) = [g(t_1), g(t_2)]$ is a linear combination over $\mathbf{Z}$ of polynomials $g(t)$, $t \in H$, with $t'' \leqslant \max(t_1, t_2)$. Let $|t|$ denote the *degree* of a tree t, which is defined by: $|t| = 1$ if t is in X; and $|t| = |t'| + |t''|$ if $t = (t', t'')$. We prove the previous statement by induction on the couple $(|t_1| + |t_2|, \max(t_1, t_2))$ where these couples are ordered by: $(d, u) < (e, v)$ if and only if either $d < e$ or $d = e$ and $u < v$; note that this is correct, since there are only finitely many Hall trees of a given degree.

By antisymmetry of the Lie bracket, we may assume that $t_1 < t_2$. Now, if t_1 is in X or if $t_1 = (t_1', t_1'')$ with $t_1'' \geqslant t_2$, then $t = (t_1, t_2)$ is a Hall tree, and $g(t)$ is in $g(H)$; moreover, $t'' = t_2 \leqslant \max(t_1, t_2)$.

So we may assume that $t_1 = (t_1', t_1'')$ and that $t_1'' < t_2$. By the property of the order, we have $t_1 < t_1''$, hence $t_1 < t_1'' < t_2$. Moreover, $t_1' < t_1''$, hence $t_1' < t_1'' < t_2$.

By the Jacobi identity, we have

$$\begin{aligned} g\big((t_1, t_2)\big) &= \big[g(t_1), g(t_2)\big] = \big[\big[g(t_1'), g(t_1'')\big], g(t_2)\big] \\ &= \big[\big[g(t_1'), g(t_2)\big], g(t_1'')\big] + \big[g(t_1'), \big[g(t_1''), g(t_2)\big]\big]. \end{aligned}$$

Since $|t_1'| + |t_2|$ and $|t_1''| + |t_2|$ are both strictly smaller than $|t_1| + |t_2|$, the induction hypothesis shows that $[g(t_1'), g(t_2)] = \sum * g(u_i)$ and $[g(t_1''), g(t_2)] = \sum * g(v_j)$, where the $*$ indicate integers whose value is of no importance here, with the property that: $u_i'' \leqslant \max(t_1', t_2) = t_2$ and $v_j'' \leqslant \max(t_1'', t_2) = t_2$; moreover, by homogeneity, $|u_i| = |t_1'| + |t_2|$ and $|v_j| = |t_1''| + |t_2|$.

Thus we obtain

$$g\big((t_1, t_2)\big) = \sum * \big[g(u_i), g(t_1'')\big] + \sum * \big[g(t_1'), g(v_j)\big].$$

We have $|u_i| + |t_1''| = |t_1'| + |t_2| + |t_1''| = |t_1| + |t_2|$, and since $u_i < u_i'' \leqslant t_2$, $\max(u_i, t_1'') < t_2 = \max(t_1, t_2)$; thus by the induction hypothesis, we deduce that $[g(u_i), g(t_1'')]$ is a linear combination over $\mathbf{Z}$ of $g(t)$ with t in H and $t'' \leqslant \max(u_i, t_1'') < \max(t_1, t_2)$. Similarly, we have $|t_1'| + |v_j| = |t_1'| + |t_1''| + |t_2| = |t_1| + |t_2|$ and $\max(t_1', v_j) < t_2 = \max(t_1, t_2)$ since $v_j < v_j'' \leqslant t_2$. Thus by induction, $[g(t_1'), g(v_j)]$ is a linear combination of $g(t)$ with $t \in H$ and $t'' \leqslant \max(t_1', v_j) < \max(t_1, t_2)$. □

10. The algorithm

If we have a closer look to the previous proof, we see that it gives an algorithm which given any Lie bracketing of the letters, writes it as a linear combination over $\mathbf{Z}$ of Hall bracketings, that is, elements of the Hall basis $g(H)$.

More formally, the algorithm takes as input a linear combination of trees $S = \sum \alpha_t t$ and gives as output a linear combination of Hall trees $\sum \beta_h h$ such that $\sum \alpha_t g(t) = \sum \beta_h g(h)$. It works as follows: as a first step, look if each tree appearing in S is a Hall tree; then there is nothing to do and the algorithm stops. Otherwise, take some tree t appearing in S, which is not a Hall tree, and consider a subtree $s = (s', s'')$ of t which is not a Hall tree but such that s', s'' are Hall trees (s surely exists since letters are Hall trees). If $s' > s''$, replace s by (s'', s') in t, and replace α_t by $-\alpha_t$ in S. If $s' = s''$ then remove t from the linear combination. If $s' < s''$, note that since s is not in H, s' is not in X, and $s' = (a, b)$ with $b < s''$; now replace t in S by the sum of the two trees t_1, t_2 obtained as follows: t_1 is obtained by replacing in t the subtree s by $((a, s''), b)$ and t_2 is obtained by replacing s by $(a, (b, s''))$. Now go back to the first step of the algorithm.

Then this algorithm stops and does the desired job, as follows from the proof in the previous section.

Observe that if $S = \sum \alpha_t t$ and if $\sum \alpha_t g(t) = 0$, then, since the polynomials $g(h)$ are linearly independent, the algorithm outputs the 0 linear combination. As a consequence, we obtain the following result.

THEOREM 5. *$L(X)$ is the free Lie algebra on X.*

PROOF. There is a canonical surjective linear mapping $G : D(X) \to L(X)$, that sends each tree t onto the corresponding Lie bracketing $g(t)$ in $\mathbf{A}\langle X\rangle$. It suffices to show that $\operatorname{Ker} G = I$, with the notations of Section 2. Since $L(X)$ is a Lie algebra, $I \subset \operatorname{Ker} G$.

For the reverse inclusion, consider an element in $\operatorname{Ker} G$. It may be written as a linear combination $S = \sum \alpha_t t$ of trees. Then $0 = G(S) = \sum \alpha_t g(t)$. The previous algorithm shows the existence of a sequence of elements $S_0, \ldots, S_n$ of $D(X)$ such that $S_0 = S$, $S_n = 0$ and that $S_i \equiv S_{i+1} \bmod I$ for each i. Thus $S \in I$. □

11. Dimensions

We assume that X has cardinality q. Let α_n denote the dimension of the homogeneous part of degree n of $L(X)$. Equivalently, α_n is the number of Hall trees of degree n.

THEOREM 6. *One has* $\alpha_n = \frac{1}{n}\sum_{d|n}\mu(d)q^{n/d}$.

PROOF. This identity is equivalent to $n\alpha_n = \sum_{d|n}\mu(d)q^{n/d}$, hence by Möbius inversion to $q^n = \sum_{d|n} d\alpha_d$. By taking generating functions, this in turn is the same as

$$\sum_{n\geqslant 1} q^n s^n = \sum_{n\geqslant 1}\left(\sum_{d|n} d\alpha_d\right)s^n.$$

The latter is equal to $\sum_{k\geqslant 1} k\alpha_k \sum_{i\geqslant 1} s^{ki}$. Thus all we have to verify is $\frac{qs}{1-qs} = \sum_{k\geqslant 1}\frac{k\alpha_k s^k}{1-s^k}$. But this is a consequence of

$$\frac{1}{1-qs} = \prod_{k\geqslant 1}\frac{1}{(1-s^k)^{\alpha_k}}, \qquad (*)$$

by taking logarithmic derivatives with respect to s and multiplying by s. Note that by Corollary 2, one has the following equality of noncommutative formal power series: $(1-\sum_{x\in X} x)^{-1} = \prod_h (1-f(h))^{-1}$, where the product is over all Hall words in strictly decreasing order. The identity $(*)$ follows from the latter by applying the homomorphism sending every letter to the same variable s. □

12. The dual basis 1

Given any word w in X^*, it has a unique decreasing factorization $w = f(t_1)\dots f(t_n)$, where the t_i are Hall trees. We define $P_w = g(t_1)\dots g(t_n)$. Then the polynomials P_w form a basis of the $\mathbf{A}$-module $\mathbf{A}\langle X\rangle$ and the polynomials P_h, h a Hall word, form a basis of $L(X)$. In order to determine the dual basis, consider the scalar product $(\ ,\ )$ on $\mathbf{A}\langle X\rangle$ for which X^* is an orthonormal basis. Since the previous basis is homogeneous, and even finely homogeneous with respect to each partial degree, there exist polynomials S_w which represent the dual basis with respect to this scalar product, that is, such that $P = \sum_{w\in X^*}(S_w, P)P_w$, for any polynomial P.

Note that $S_1 = 1$.

THEOREM 7. *If $h = xv$ is a Hall word with $x \in X$, $v \in X^*$, then $S_h = xS_v$.*

We know that Hall words are in one-to-one correspondence with Hall trees; we use this bijection to transfer definitions on standard sequences of Hall trees to sequences of Hall words. In particular, we transfer the total order on Hall trees to Hall words; moreover, if $h = f(t)$ is a Hall word, the image under f of the Hall tree t, we put $g(h) = g(t)$, and for a standard sequence of Hall words $s = (h_1, \dots, h_n)$, we extend g by $g(s) = g(h_1)\dots g(h_n)$, which is also equal to $P_{h_1}\dots P_{h_n}$ with our previous notation.

Furthermore for each standard sequence of Hall words s, which has at least one rise, choose a fixed rise $i(s)$. Now, define a relation $\Rightarrow$ on standard sequences by: $s \Rightarrow s'$ if $i = i(s)$ exists and if $s' = \lambda_i(s) = (h_1, \dots, h_{i-1}, h_i h_{i+1}, h_{i+2}, \dots, h_n)$ or $s' = \rho_i(s) =$

$(h_1, \ldots, h_{i-1}, h_{i+1}, h_i, h_{i+2}, \ldots, h_n)$. Denote by $\Rightarrow^*$ the reflexive and transitive closure of $\Rightarrow$.

LEMMA 1. *Let s be a standard sequence $(h_1, \ldots, h_n)$ of Hall words, with $n \geqslant 2$ and $h_2 \geqslant \cdots \geqslant h_n$.*

- *If $h_1 \ldots h_n$ is a Hall word, there is exactly one chain $s \Rightarrow \cdots \Rightarrow (h_1 \ldots h_n)$.*
- *Given a standard sequence s' such that $s \Rightarrow^* s'$ and such that $s' \neq (h_1 \ldots h_n)$ (this inequality holds certainly if $h_1 \ldots h_n$ is not a Hall word), s' is of length at least 2.*

PROOF. We claim that if $u = (u_1, \ldots, u_m)$ is a standard sequence of Hall words with $m \geqslant 2$ and $u_1 \geqslant u_2, \ldots, u_m$, then for each standard sequence v with $u \Rightarrow^* v$, v is of length at least 2. Indeed, if i is a rise of u, then $i \geqslant 2$. Thus $\rho_i(s)$ satisfies the same hypothesis as u. Moreover, so does $\lambda_i(s)$, since $u_{i+1} > u_i u_{i+1}$, because $u_i u_{i-1}$ is a Hall word, image under f of the Hall tree $t = (t', t'')$ with $f(t') = u_i$ and $f(t'') = u_{i+1}$. Thus we conclude by induction on the length of the chain from u to v.

We now prove the lemma. If $h_1 \geqslant h_2$, then there is no nontrivial chain starting from s. Thus we may assume that $h_1 < h_2$. This will be the only rise, so that $i(s) = 1$; moreover, $\rho_1(s) = (h_2, h_1, h_3, \ldots, h_n)$ and $\lambda_1(s) = (h_1 h_2, h_3, \ldots, h_n)$. Since by assumption, the first element of $\rho_1(s)$ is its maximum, the claim implies that the sequences of each chain starting at $\rho_1(s)$ are all of length at least 2. Moreover, $\lambda_1(s)$ is shorter than s and either is of length 1, or satisfies the same hypothesis as s. Thus we conclude by induction. □

PROOF OF THE THEOREM. For any standard sequence $s = (h_1, \ldots, h_n)$ of Hall words, $P_{h_1} \ldots P_{h_n}$ is equal to the sum of all $g(s')$, for all possible decreasing sequences of Hall words s' such that $s \Rightarrow^* s'$. Indeed, this follows from Section 8.

The equality in the theorem is equivalent to $(S_h, yu) = \delta_{x,y}(S_v, u)$ $(x, y \in X,\ u, v \in X^*)$, since S_h has no constant term. We have $u = \sum_{w \in X^*} (S_w, u) P_w$, so that $yu = \sum_{w \in X^*} (S_w, u) y P_w$. Let w be a word with $w = h_1 \ldots h_n$, h_i Hall word, $h_1 \geqslant \cdots \geqslant h_n$. Then the sequence $s = (y, h_1, \ldots, h_n)$ is standard and

$$y P_w = \sum_{s'} \alpha_{s'} g(s'),$$

where the summation is over all decreasing sequences s' such that $s \Rightarrow^* s'$ and where $\alpha_{s'}$ is the number of chains $s \Rightarrow \cdots \Rightarrow s'$.

By Lemma 1, each such s' is of length at least 2, except when $yw = y h_1 \ldots h_n$ is a Hall word and $s' = (yw)$, in which case there is exactly one chain from s to s'. This implies that yu is equal to $\sum_{yw \text{ Hall word}} (S_w, u) P_{yw}$ + a sum of decreasing products $g(h'_1) g(h'_2) \ldots$ with at least two factors, $h'_i \in H$. Hence the coefficient of $P_h = P_{xv}$ in this sum is equal to 0 if $x \neq y$ and to (S_v, u) if $x = y$. In other words, $(S_h, yu) = \delta_{x,y}(S_v, u)$. □

13. Shuffle product

For each word $w = x_1 \ldots x_n$ of length n on X and each subset I of $\{1, \ldots, n\}$, denote by $w|I$ the word $x_{i_1} \ldots x_{i_k}$, with $I = \{i_1 < \cdots < i_k\}$. Given p words $u_1, \ldots, u_p$ whose

lengths $n_1, \ldots, n_p$ add up to n, their *shuffle product* is the polynomial $u_1 \times \cdots \times u_p = \sum w(I_1, \ldots, I_p)$, where the sum is over all p-tuples $(I_1, \ldots, I_p)$ of disjoint subsets of $\{1, \ldots, n\}$, whose union is $\{1, \ldots, n\}$, with $|I_j| = n_j$, and where the word $w(I_1, \ldots, I_p)$ is defined by $w(I_1, \ldots, I_p)|I_j = u_j$. This product extends linearly to polynomials, since words form a basis of $\mathbf{A}\langle X\rangle$. It is easy to see that the 2-ary shuffle product $u \times v$ is an associative product, with as neutral element the empty word.

The shuffle product may be alternatively defined as follows. Let δ_p denote the homomorphism from the free associative algebra $\mathbf{A}\langle X\rangle$ into the p-fold tensor product $\mathbf{A}\langle X\rangle^{\otimes p}$, defined by $\delta_p(x) = x \otimes 1 \otimes \cdots \otimes 1 + 1 \otimes x \otimes \cdots \otimes 1 + \cdots + 1 \otimes 1 \otimes \cdots \otimes x$ for any letter x.

PROPOSITION 1. *For each polynomial P, one has*

$$\delta_p(P) = \sum_{u_1, \ldots, u_p \in X^*} (P, u_1 \times \cdots \times u_p) u_1 \otimes \cdots \otimes u_p.$$

Equivalently, $(\delta_p(P), P_1 \otimes \cdots \otimes P_p) = (P, P_1 \times \cdots \times P_p)$ for any polynomials $P, P_1, \ldots, P_p$, where the scalar product is naturally extended to the tensor product.

PROOF. It is enough to prove this when $P = w = x_1 \ldots x_n$ is a word in X^*. Then by definition $\delta_p(w)$ is equal to the product $\delta_p(x_1) \ldots \delta_p(x_n)$. Since $\delta_p(x_i) = x_i \otimes 1 \otimes \cdots \otimes 1 + 1 \otimes x_i \otimes \cdots \otimes 1 + \cdots + 1 \otimes 1 \otimes \cdots \otimes x_i$, the proposition follows by inspection. □

We also need the following result.

PROPOSITION 2. *If P is a Lie polynomial then*

$$\delta_p(P) = P \otimes 1 \otimes \cdots \otimes 1 + 1 \otimes P \otimes \cdots \otimes 1 + \cdots + 1 \otimes 1 \otimes \cdots \otimes P.$$

PROOF. This holds by definition when P is a letter. A simple computation shows that if it holds for two polynomials, then it holds also for their Lie bracket. Hence the proposition follows, since the set of polynomials for which the formula is true is a submodule closed under the Lie bracket. □

14. The dual basis 2

In Section 12, we have determined all S_h when h is a Hall word, knowing S_w for shorter words w. The next theorem completely solves the problem of recursively determining the dual basis (S_w) of the Poincaré–Birkhoff–Witt basis. We assume that the ring $\mathbf{A}$ contains $\mathbf{Q}$. We denote by $^{\times i}$ the shuffle exponentiation.

THEOREM 8. *For any word $w = h_1^{i_1} \ldots h_k^{i_k}$, where the h_j are Hall words with $h_1 > \cdots > h_k$ and i_j in $\mathbf{N}$, one has*

$$S_w = \frac{1}{i_1! \ldots i_k!} S_{h_1}^{\times i_1} \times \cdots \times S_{h_k}^{\times i_k}.$$

PROOF. Note that by definition of the dual basis, one has $(S_u, P_v) = \delta_{u,v}$. In particular, if u is a Hall word and v is not, $(S_u, P_v) = 0$.

Following Corollary 2, we may write w as a decreasing product of Hall words: $w = w_1 \dots w_i$, $w_1 \geqslant \dots \geqslant w_i$; hence $i = i_1 + \dots + i_k$. By the remark after Proposition 1, we have $(S_{w_1} \times \dots \times S_{w_i}, P_u) = (S_{w_1} \otimes \dots \otimes S_{w_i}, \delta_i(P_u))$.

Write $u = u_1 \dots u_n$ as decreasing product of Hall words. Then $P_u = P_{u_1} \dots P_{u_n}$ and the P_{u_j} are Lie polynomials. By Proposition 2, we have

$$\delta_i(P_{u_j}) = P_{u_j} \otimes 1 \otimes \dots \otimes 1 - 1 \otimes P_{u_j} \otimes \dots \otimes 1 + \dots + 1 \otimes 1 \otimes \dots \otimes P_{u_j}.$$

Moreover, $\delta_i(P_u) = \delta_i(P_{u_1}) \dots \delta_i(P_{u_n})$. Thus, by inspection, we find that $\delta_i(P_u)$ is a sum of terms $Q_1 \otimes \dots \otimes Q_i$ and, correspondingly, $(S_{w_1} \times \dots \times S_{w_i}, P_u)$ is a sum of products $(S_{w_1}, Q_1) \dots (S_{w_i}, Q_i)$: If $i > n$, then in each term at least one Q_j is equal to 1, hence since S_{w_j} has no constant term, we have $(S_{w_1} \times \dots \times S_{w_i}, P_u) = 0$. If $i < n$, then in each term, at least one Q_j is a decreasing product $P_{u'} = P_{u_{i_1}} \dots P_{u_{i_r}}$ with $r \geqslant 2$, so that $(S_{w_j}, P_{u'}) = 0$, since w_j is a Hall word and $u' = u_{i_1} \dots u_{i_r}$ is not; thus we also have $(S_{w_1} \times \dots \times S_{w_i}, P_u) = 0$. If finally $i = n$, then we obtain, again because S_{w_i} has no constant term,

$$\begin{aligned}(S_{w_1} \times \dots \times S_{w_i}, P_u) &= \sum_{\sigma \in S_n} (S_{w_1}, P_{u_{\sigma(1)}}) \dots (S_{w_n}, P_{u_{\sigma(n)}}) \\ &= \sum_{\sigma \in S_n} \delta_{w_1, u_{\sigma(1)}} \dots \delta_{w_n, u_{\sigma(n)}}.\end{aligned}$$

If $w \neq u$, then $(w_1, \dots, w_n) \neq (u_1, \dots, u_n)$ and since both sequences are decreasing, the right-hand side of the summation vanishes. If $w = u$, then both sequences are equal by Corollary 2; since $(w_1, \dots, w_n) = (h_1, \dots, h_1, \dots, h_k, \dots, h_k)$, each h_j repeated i_j times, the right-hand side is equal to the number of permutations fixing the previous sequence, that is $i_1! \dots i_k!$. Hence we obtain that

$$\left(\frac{1}{i_1! \dots i_k!} S_{h_1}^{\times i_1} \times \dots \times S_{h_k}^{\times i_k}, P_u \right) = \delta_{w,u},$$

which proves the theorem, by definition of the dual basis. □

COROLLARY 4. *The polynomials of the dual basis (S_w) have all coefficients in* **N**.

COROLLARY 5. *The shuffle algebra is a free commutative* **A***-algebra over the S_h (h = Hall word).*

15. Notes

Hall bases are due to Marshall Hall Jr. [16]. Similar constructions of "basic commutators" in a free group had been done previously by Philip Hall [18] and Wilhelm Magnus [25]. M. Hall's construction was generalized by Meier-Wunderli [27], Witt [51], Schützenberger

[37], Širšov [40], Michel [31], Viennot [47]. Unlike the original Hall basis, these generalizations include the Lyndon basis constructed by Viennot (loc. cit.; see also Lothaire [24]), following the lines of the commutator calculus of Chen, Fox, Lyndon [8], and the Širšov basis [39].

Theorem 1 is due to Melançon [29], who extended a method of Schützenberger (loc. cit.), itself related to the "collecting process" of P. Hall (loc. cit.; see also M. Hall [17]). This theorem immediately implies Theorem 2, Corollaries 1 and 2, which constitute the combinatorial facts underlying Hall bases. The latter corollary also easily implies Theorem 6, which is the Witt formula (loc. cit.). Note that this formula gives also the number of *primitive circular words* on the alphabet X; an explanation of this fact is the following result: each primitive conjugation class of words contains exactly one Hall word (see, e.g., the book by Lothaire, loc. cit.).

The main result on Hall bases is Theorem 4; for the proof we have followed Schützenberger's proof (loc. cit.), which is algorithmic, and gives as a byproduct Theorem 5: the Lie algebra of Lie polynomials is the free Lie algebra; this theorem is due to Witt (loc. cit.). Note that this method does not use the theorem of Poincaré–Birkhoff–Witt (PBW). The counterpart for Hall bases of this theorem (Theorem 3) is established here directly. For the latter, Theorems 7 and 8, we have also followed Schützenberger (loc. cit.), with improvements from Melançon and Reutenauer [30] and Melançon [28].

Note that other bases, which are not Hall bases, are constructed by Kukin [23], Blessenohl and Laue [3], Garsia [13].

Proposition 2 characterizes Lie elements when $i = 2$: this is Friedrichs' criterion [12]. An equivalent version says that P is a Lie polynomial if and only if P is orthogonal to each shuffle product of two nonempty words for the scalar product of Section 12. In relation with Lie polynomials, the shuffle product was introduced by Ree [33], who simplified previous work of Chen on iterated integrals [7]. An alternative formulation of Friedrichs' criterion is due to Garsia (loc. cit.): say that a noncommutative polynomial $P(x, y, \ldots)$ on the alphabet $X = \{x, y, \ldots\}$ is *linear* if, taking a second alphabet $X' = \{x', y', \ldots\}$, in bijection with the previous one, such that each letter in X commutes with each letter in X', one has: $P(x + x', y + y', \ldots) = P(x, y, \ldots) + P(x', y', \ldots)$. Then Lie polynomials are exactly the linear ones.

Another well-known and earlier characterization of Lie polynomials is the following: denote by l the linear endomorphism of $\mathbf{A}\langle X\rangle$ "Lie bracketing from left to right" defined by $l(w) = [\ldots[[x_1, x_2], x_3], \ldots x_n]$ for any word $w = x_1 \ldots x_n$ of length n. Then a homogeneous polynomial P of degree n is a Lie polynomial if an only if $l(P) = nP$ (we assume that $\mathbf{A}$ is of characteristic 0). This characterization was found at the same time by three authors: Dynkin [11], Specht [44], Wever [48].

Actually, Lie polynomials appear implicitly in the work of Campbell [5,6], Baker [1] and Hausdorff [19], leading to their famous formula. It asserts that the noncommutative series $e^x e^y$ is the exponential of a Lie series (a Lie series is a series whose homogeneous components are Lie polynomials). Another famous Lie series is the Drinfeld *associator*, see [10].

Lie subalgebras of the free Lie algebra are again free, see Širšov [38], Witt [50,51]. Automorphism of a free Lie algebra are always tame (Cohn [9]) and are characterized by a Jacobian condition, see Shpilrain [41], Reutenauer [35], Umirbaev [46].

Representation-theoretic studies on the free Lie algebra began with the work of Thrall [45] and Brandt [4]. The full linear group acts on Lie polynomials, and the symmetric group acts on those which are multilinear. A result of Klyachko [21] characterizes the irreducible representations that occur, and another one of Kraskiewicz and Weyman [22] gives the exact multiplicities, in terms of the *major index* of Young tableaux. The Lie representation of the symmetric group is induced from any faithful representation of a subgroup generated by a full cycle (Klyachko, loc. cit.).

In relation with representation theory, many idempotents of the symmetric group algebra, called *Lie idempotents* may be found in the literature: the Dynkin–Specht–Wever idempotent, implicit in the work of these three authors (loc. cit.); the canonical idempotent, related to the PBW theorem, see Solomon [42], Mielnik, Plebanski [32], Reutenauer [34], Helmstetter [20]; the idempotent of Klyachko (loc. cit.), that has a fabulous definition using the major index of permutations, and its generalization by Bergeron, Bergeron, Garsia [2].

There are other idempotents related to the PBW decomposition of the tensor algebra; these idempotents constitute all primitive idempotents of the *descent algebra* of the symmetric group, see Garsia, Reutenauer [14]; the descent algebra was introduced by Solomon [43] for each finite Coxeter group. All previous Lie idempotents lie in the descent algebra. The latter is itself a Hopf algebra, isomorphic to a free associative algebra, and its primitive elements are exactly the quasi-Lie idempotents that lie in it, see Gelfand, Krob, Lascoux, Leclerc, Retakh, Thibon [15]. The descent algebra is dual to the ring of quasi-symmetric functions, see [26], which is therefore a free commutative algebra.

See the author's book on Free Lie Algebras [36] for more on the subject.

References

[1] H.F. Baker, *Alternants and continuous groups*, Proc. London Math. Soc. **3** (1905), 24–47.

[2] F. Bergeron, N. Bergeron and A. Garsia, *Idempotents for the free Lie algebra and q-enumeration*, Invariant Theory and Tableaux (Minneapolis, MN, 1988), IMA Vol. Math. Appl., Vol. 19, Springer, New York (1990), 166–190.

[3] D. Blessenohl and H. Laue, *A basis construction for free Lie algebras*, Exposition. Math. **11** (1993), 145–152.

[4] A. Brandt, *The free Lie ring and representations of the full linear group*, Trans. Amer. Math. Soc. **56** (1944), 528–536.

[5] J.E. Campbell, *On a law of combination of operators bearing on the theory of continuous transformation groups*, Proc. London Math. Soc. **28** (1897), 381–390.

[6] J.E. Campbell, *On a law of combination of operators bearing on the theory of continuous transformation groups*, Proc. London Math. Soc. **29** (1898), 14–32.

[7] K.T. Chen, *Integration of paths, geometric invariants and a generalized Baker–Hausdorff formula*, Ann. of Math. **65** (1957), 163–178.

[8] K.T. Chen, R.H. Fox and R.C. Lyndon, *Free differential calculus IV: The quotient groups of the lower central series*, Ann. of Math. **68** (1958), 81–95.

[9] P.M. Cohn, *Subalgebras of free associative algebras*, Proc. London Math. Soc. **14** (1964), 618–632.

[10] V.G. Drinfeld, *On quasitriangular quasi-Hopf algebras and on a group that is closely connected with* $\mathrm{Gal}(\overline{Q}/Q)$, Algebra i Analiz **2** (1990), 149–181; transl.: Leningrad Math. J. **2** (1991), 829–860.

[11] E.B. Dynkin, *Calculation of the coefficients in the Campbell–Hausdorff formula*, Dokl. Akad. Nauk SSSR **57** (1947), 323–326.

[12] K.O. Friedrichs, *Mathematical aspects of the quantum theory of fields. V: Fields modified by linear homogeneous forces*, Comm. Pure Appl. Math. **6** (1953), 1–72.
[13] A. Garsia, *Combinatorics of the free Lie algebra and the symmetric group*, Analysis, et Cetera, Academic Press, Boston, MA (1990), 309–382.
[14] A. Garsia and C. Reutenauer, *A decomposition of Solomon's descent algebra*, Adv. Math. **77** (1989), 189–262.
[15] I.M. Gelfand, D. Krob, A. Lascoux, B. Leclerc, V. Retakh and J.-Y. Thibon, *Noncommutative symmetric functions*, Adv. Math. **112** (1995), 218- 348.
[16] M. Hall, *A basis for free Lie rings and higher commutators of the free group*, Proc. Amer. Math. Soc. **1** (1950), 575–580.
[17] M. Hall, *The Theory of Groups*, Macmillan, New York (1959).
[18] P. Hall, *A contribution to the theory of groups of prime power order*, Proc. London Math. Soc. **36** (1933), 29–95.
[19] F. Hausdorff, *Die symbolische Exponentialformel in der Gruppentheorie*, Leipziger Ber. **58** (1906), 19–48.
[20] J. Helmstetter, *Série de Hausdorff d'une algèbre de Lie et projections canoniques dans l'algèbre enveloppante*, J. Algebra **120** (1989), 170–199.
[21] A.A. Klyachko, *Lie elements in the tensor algebra*, Sibirsk. Mat. Zh. (translation) **15** (1974), 1296–1304.
[22] W. Kraskiewicz and J. Weyman, *Algebra of coinvariants and the action of Coxeter elements*, Manuscript (1987).
[23] G.P. Kukin, *Bases of a free Lie algebra*, Mat. Zametki **24** (1978), 375–382.
[24] M. Lothaire, *Combinatorics on Words*, Addison-Wesley, Reading, MA (1983).
[25] W. Magnus, *Über Beziehungen zwischen höheren Kommutatoren*, J. Reine Angew. Math. **177** (1937), 105–115.
[26] C. Malvenuto and C. Reutenauer, *Duality between quasi-symmetric functions and the Solomon descent algebra*, J. Algebra **177** (1995), 967–982.
[27] H. Meier-Wunderli, *Note on a basis of P. Hall for the higher commutators in free groups*, Comment. Math. Helv. **26** (1951), 1–5.
[28] G. Melançon, *Récritures dans l'algèbre de Lie libre, le groupe libre et l'algèbre associative libre*, Thèse Math., Univ. du Québec à Montréal (1991).
[29] G. Melançon, *Combinatorics of Hall trees and Hall words*, J. Combin. Theory A **59** (1992), 285–308.
[30] G. Melançon and C. Reutenauer, *Lyndon words, free algebras and shuffles*, Canad. J. Math. **41** (1989), 577–591.
[31] J. Michel, *Bases des algèbres de Lie libres, études des coefficients de la formule de Campbell–Hausdorff*, Thèse de 3ème cycle, Université Paris XI (1974).
[32] B. Mielnik and J. Plebanski, *Combinatorial approach to Baker–Campbell -Hausdorff exponents*, Ann. Inst. H. Poincaré A **12** (1970), 215–254.
[33] R. Ree, *Lie elements and an algebra associated with shuffles*, Ann. of Math. **68** (1958), 210–220.
[34] C. Reutenauer, *Theorem of Poincaré–Birkhoff–Witt, logarithm and representations of the symmetric group whose orders are the Stirling numbers*, Combinatoire Énumérative, Proc., Montréal, P. Leroux and G. Labelle, eds, Lecture Notes in Math., Vol. 1234, Springer, Berlin (1986).
[35] C. Reutenauer, *Applications of a noncommutative Jacobian matrix*, J. Pure Appl. Algebra **77** (1992), 169–181.
[36] C. Reutenauer, *Free Lie Algebras*, Oxford University Press (1993).
[37] M.-P. Schützenberger, *Sur une propriété combinatoire des algèbres de Lie libres pouvant être utilisée dans un problème de mathématiques appliquées*, Séminaire P. Dubreil, Faculté des Sciences, Paris (1958).
[38] A.I. Širšov, *Subalgebras of free Lie algebras*, Mat. Sb. **33** (1953), 441–452.
[39] A.I. Širšov, *Free Lie rings*, Mat. Sb. **45** (1958), 113–122.
[40] A.I. Širšov, *On the bases of a free Lie algebra*, Algebra i Logika **1** (1962), 14–19.
[41] V. Shpilrain, *On generators of L/R^2 Lie algebras*, Proc. Amer. Math. Soc. **119** (1993), 1039–1043.
[42] L. Solomon, *On the Poincaré–Birkhoff–Witt theorem*, J. Combin. Theory A **4** (1968), 363–375.
[43] L. Solomon, *A Mackey formula in the group ring of a finite Coxeter group*, J. Algebra **41** (1976), 255–268.
[44] W. Specht, *Die linearen Beziehungen zwischen höheren Kommutatoren*, Mat. Z. **51** (1948), 367–376.
[45] R.M. Thrall, *On symmetrized Kronecker powers and the structure of the free Lie ring*, Amer. J. Math. **64** (1942), 371–388.

[46] U.U. Umirbaev, *Partial derivations and endomorphisms of some relatively free Lie algebras*, Sibirsk. Mat. Zh. **34** (1993), 179–188; transl.: Siberian Math. J. **34** (1993), 1161–1170.
[47] G.X. Viennot, *Algèbres de Lie libres et monoïdes libres*, Lecture Notes in Math., Vol. 691, Springer, Berlin (1978).
[48] F. Wever, *Über Invarianten in Lieschen Ringen*, Math. Ann. **120** (1949), 563–580.
[49] E. Witt, *Treue Darstellungen Liescher Ringe*, J. Reine Angew. Math. **177** (1937), 152–160.
[50] E. Witt, *Über freie Liesche Ringe und ihre Unterringe*, Math. Z. **58** (1953), 113–114.
[51] E. Witt, *Die Unterringe der freien Lieschen Ringe*, Math. Z. **64** (1956), 195–216.