Models for Free Nilpotent Lie Algebras

Matthew Grayson*
University of California, San Diego

Robert Grossman*
University of Illinois, Chicago

March, 1988

This is a draft of a paper which later appeared in the Journal Algebra, Volume 35, 1990, pp. 177-191.

If a Lie algebra g can be generated by M of its elements $E_1, \ldots, E_M$, and if any other Lie algebra generated by M other elements $F_1, \ldots, F_M$ is a homomorphic image of g under the map $E_i \to F_i$, we say that it is the *free* Lie algebra on M generators. The free nilpotent Lie algebra $\mathrm{g}_{M,r}$ on M generators of rank r is the quotient of the free Lie algebra by the ideal g^{r+1} generated as follows: $\mathrm{g}^1 = \mathrm{g}$, and $\mathrm{g}^k = [\mathrm{g}^{k-1}, \mathrm{g}]$. Let N denote dimension of $\mathrm{g}_{M,r}$. Free Lie algebras are those which have as few relations as possible: only those which are a conseqence of the anti-commutativity of the bracket and of the the Jacobi identity. Free nilpotent Lie algebras add the relations that any iterated Lie bracket of more than r elements vanishes. For further details, see [18] or [13].

In this paper we are interested in explicit computations of the Lie algebras $\mathrm{g}_{M,r}$. It is well known [19] that that there is a representation of $\mathrm{g}_{M,r}$ on upper triangular N by N matrices. The problem is that many computations are difficult using this representation. In this paper we present an algorithm that yields vector fields $E_1, \ldots, E_M$ defined in $\mathbf{R}^N$ with the property that they generate a Lie algebra isomorphic to $\mathrm{g}_{M,r}$. See [6] for another approach to this problem.

In Section 3, we restrict to two generators and give three consequences of this algorithm. First, the form of the vector fields is such that flows of the control system $\dot{x}(t) = (E_1 + u(t)E_2)(x(t))$ may be computed explicitly

*Research supported in part by Postdoctoral Research Fellowships from the National Science Foundation.

in closed form. Second, the coefficients of the Baker–Campbell–Hausdorff may easily be computed. Third, the Poincare–Birkhoff–Witt theorem, for the special case of a free, nilpotent Lie algebra, is a simple consequence of the algorithm. We conclude this introduction with a brief description of how nilpotent models arise in control theory and partial differential equations.

Nilpotent Lie algebras have been used in control theory beginning with the work of Krener [14], and continuing with the work of Hermes [9] and [10], Crouch [2], Bressan [1], and Hermes, Lundell and Sullivan [11]. Consider a control system evolving in $\mathbf{R}^N$

$$\dot{x}(t) = F_1(x(t)) + u(t)F_2(x(t)), \quad x(0) = 0 \in \mathbf{R}^N,$$

where F_1 and F_2 are two vector fields defined in a neighborhood of the origin of $\mathbf{R}^N$ and $t \rightarrow u(t)$ is a control. Suppose for a moment that at the origin all Lie brackets formed from r or fewer E's agree with the corresponding Lie bracket formed from the F's and that the E's generate a nilpotent Lie algebra of step r. Then there exists a diffeomorophism

$$\lambda : \mathbf{R}^N \longrightarrow \mathbf{R}^N$$

such that the trajectory $t \rightarrow x(t)$ of the control system defined by the F's is close for small time to the trajectory $t \rightarrow y(t)$ of the corresponding trajectory defined by the E's in the sense that the following estimate

$$\mid \lambda(y(t)) - x(t) \mid \leq \text{constant } t^r,$$

holds for small t. If the E's satisfy some mild technical conditions (they have polynomial coefficients and are homogeneous of weight 1), then it is easy to see that the trajectories of the E system may be explicitly computed by quadrature involving the controls u. While in general it is not possible to find vector fields E_1 and E_2 with prescribed Lie brackets at the origin, control systems whose defining vector fields generate a free, nilpotent Lie algebra form an interesting class of explicitly computable examples.

Beginning with the work of Hörmander [12] and continuing with the work of Folland and Stein [4], Rothschild and Stein [18], Rothschild [17], and Rockland [16] nilpotent Lie algebras have been used in analysis. See Helffer and Nourrigat [8] for more information. Consider the hypoellipticity of a partial differential operator

$$L = \sum_{j=1}^{M} F_j^2,$$

where the F_j are smooth real valued vector fields defined in a neighborhood of the origin of $\mathbf{R}^N$. Rothschild and Stein [18] showed how to add new variables in an appropriate fashion so that the vector fields $F_1, \ldots, F_M$ are replaced by vector fields $\tilde{F}_1, \ldots, \tilde{F}_M$ defined in a larger space $\mathbf{R}^{\tilde{N}}$ with the property that the latter vector fields are free (in an appropriate sense) at a given point. Finally the vector fields $\tilde{F}_1, \ldots, \tilde{F}_M$ are well approximated by the generators $E_1, \ldots, E_M$ of a free, nilpotent Lie algebra $\mathrm{g}_{M,r}$, for some r. The hypoellipticity of the well understood operator

$$L = \sum_{j=1}^{M} E_j^2$$

can then be used to determine the hypoellipticity of the original operator. In this area also there are few explicit examples of free, nilpotent Lie algebras in high dimensions.

In §2 we review some needed facts about graded Lie algebras. In §3 we give an algorithm for determining M generators with polynomial coefficients which generate the free, nilpotent Lie algebra on M generators of of rank r, for all $M \geq 2$ and $r \geq 1$. In §3, we give the three applications of the algorithm mentioned above and an example.

The authors would like to thank Professor Rothschild for the many helpful comments she made on an earlier version of this paper.

1 Graded Lie algebras

In this section we review some facts about graded Lie algebras following Goodman [5]. We begin this section by making a number of definitions. Let V be a real vector space with a direct sum decomposition

$$V = V_1 \oplus \ldots \oplus V_r,$$

and let R be the ring of real polynomial functions on V. Define a one parameter group of *dilations* $\{\delta_t : t > 0\}$ by

$$\delta_t(\sum v_i) = \sum t^i v_i, \text{ where } v_i \in V_i.$$

and put

$$R_m = \{a \in R : a \circ \delta_t = t^m a\};$$

these are the polynomials *homogeneous of weight* m.

Fix a basis $e_1, \ldots, e_N$ of V and let $x_1, \ldots, x_N$ denote the dual basis of V^*. If $f = (f_1, \ldots, f_N)$, is a multi-index with non-negative entries, write

$$x^f = x_1^{f_1} \cdots x_N^{f_N}$$

and $f! = f_1! \cdots f_N!$. We will also need the difference of the multi-index f and the multi-index $g = (g_1, \ldots g_N)$

$$g - f = (g_1 - f_1, \ldots, g_N - f_N).$$

Notice that the weight of x_i is l in case $e_i \in V_l$ and the weight of x^f is $\sum_i f_i$ weight(x_i). Since $R_m R_n \subset R_{m+n}$, the decomposition

$$R = \bigoplus_{m \geq 0} R_m$$

is a *grading* of the algebra of polynomial functions R.

Since R is generated by 1 and the linear functions $x_1, \ldots, x_N$, any derivation E of the ring R is defined by its action on V^*. Therefore

$$E = \sum_{\mu=1}^{N} b^\mu \frac{\partial}{\partial x_\mu},$$

where $b^\mu = E(x_\mu)$. In other words the derivations of R are simply the vector fields on V with polynomial coefficients. We say that a derivation E is *homogenous* of weight m in case

$$E(a \circ \delta_t) = t^m (E(a)) \circ \delta_t, \text{ for all } t > 0 \text{ and } a \in R.$$

Given M derivations $E_1, \ldots, E_M$, we define the *Lie algebra generated by* $E_1, \ldots, E_M$ to be the smallest real subspace of derivations of R which is closed under the formation of Lie brackets

$$[E, F] = EF - FE;$$

this is denoted

$$\mathrm{g}(E_1, \ldots, E_M).$$

Elements in the linear span of

$$[E_{\alpha_1}, \ldots, [E_{\alpha_{k-1}}, E_{\alpha_k}] \ldots],$$

where $1 \leq \alpha_i \leq N$, are called *commutators* of length $\leq k$. Observe that, in $\mathrm{g}_{M,r}$, if E is a commutator of length $\leq k$ and F is a commutator of length

$\leq l$, then $[E, F]$ is a commutator of length $\leq k + l$. Also, observe that if $E_1, \ldots, E_M$ are homogenous of weight 1 with respect to the grading

$$V = V_1 \oplus \ldots \oplus V_r,$$

then the Lie algebra they generate is nilpotent of step r and any non-zero commutator of length l is homogenous of weight l. This is because if F_1 is a derivation homogenous of weight k and F_2 is a derivation homogenous of weight l, then $[F_1, F_2]$ either vanishes or is homogenous of weight $k + l$.

We turn now to free, nilpotent Lie algebras. In [7] Hall shows that the following is a basis for $\mathrm{g}_{M,r}$.

Definition 1.1 *Each element of the Hall basis is a monomial in the generators and is defined recursively as follows. The generators $E_1, \ldots, E_M$ are elements of the basis and of length 1. If we have defined basis elements of lengths $1, \ldots, r-1$, they are simply ordered so that $E < F$ if length$(E) <$ length(F). Also if length$(E) = s$ and length$(F) = t$ and $r = s + t$, then $[E, F]$ is a basis element of length r if:*

1. *E and F are basis elements and $E > F$, and*
2. *if $E = [G, H]$, then $F \geq H$.*

We now give a definition, due to Stein and Rothschild [18]. Let $E_1, \ldots, E_M$ be derivations of a ring of polynomial functions. We say that $E_1, \ldots, E_M$ are *free to step* r at 0 in case

$$\dim \operatorname{span}\{\text{commutators of length} \leq r \text{ evaluated at } 0\} = \dim \mathrm{g}_{M,r}.$$

Note that if derivations $E_1, \ldots, E_M$ are free to step r and nilpotent of step r, then they generate the free, nilpotent Lie algebra $\mathrm{g}_{M,r}$. We are interested in the following problem.

Problem Given integers $r, M \geq 1$, find derivations $E_1, \ldots, E_M$ of a ring of polynomial functions R such that $\mathrm{g}(E_1, \ldots, E_M)$ is a free, nilpotent Lie algebra of step r.

2 Constructing the generators

Fix the number of generators M and the rank $r \geq 1$ of the free, nilpotent Lie algebra $\mathrm{g}_{M,r}$ and let N denote its dimension. We will use the grading

$$\mathrm{g}_{M,r} = V_1 \oplus \cdots \oplus V_r,$$

where V_i = the span of basis elements of length i. This induces a grading on the isomorphic vector space $\mathbf{R}^N$.

Number the basis elements for the Lie algebra by the ordering from Definition 1.1, *i.e.*, $E_{M+1} = [E_2, E_1]$, $E_{M+2} = [E_3, E_1]$, $E_{M+3} = [E_3, E_2]$, *etc.* Consider a basis element E_i as a bracket in the lower order basis elements, $[E_{j_1}, E_{k_1}]$, where $j_1 > k_1$. If we repeat this process with E_{j_1}, we get $E_i = [\,[E_{j_2}, E_{k_2}], E_{k_1}]$, where $k_2 \leq k_1$ by the Hall basis conditions. Continuing in this fashion, we get

$$E_i = [\,[\cdots[\,[E_{j_n}, E_{k_n}], E_{k_{n-1}}], \cdots, E_{k_2}], E_{k_1}],$$

where $k_n < j_n \leq M$, and $k_{l+1} \leq k_l$ for $1 < l < n-1$. This expansion involves n brackets, and we write $d(i) = n$ and define $d(1) = \cdots = d(M) = 0$. This process naturally associates a multi-index $I(i) = (a_1, \ldots, a_N)$, to each Hall basis element E_i defined by $a_s = \sharp\{t : k_t = s\}$. Note that $I(i) = (0, \ldots, 0)$ for $1 \leq i \leq M$. We say that E_i is a direct descendant of each E_{j_l}, and we indicate this by writing $j_l \prec i$. Note that $\prec$ is a partial ordering. If $j \prec i$, then each entry in $I(i)$ is at least as large as the corresponding entry in $I(j)$.

For every pair i and j with $j \prec i$, we define the monomial $P_{j,i}$ by

$$P_{j,i} = \frac{-1^{(d(i)-d(j))}}{(I(i)-I(j))!}\, x^{(I(i)-I(j))}.$$

By construction, every exponent in $P_{j,i}$ is non-negative.

For example, if $E_i = [[[[[[[E_2, E_1], E_1], E_1], E_2], E_4], E_4], E_7]$, and if $E_k = [[[E_2, E_1], E_1], E_1]$, then

$$I(i) = (3, 1, 0, 2, 0, 0, 1, 0, \ldots, 0), \quad P_{2,i} = -\frac{x_1^3 x_2 x_4^2 x_7}{3!2!}, \quad \text{and} \quad P_{k,i} = \frac{x_2 x_4^2 x_7}{2!}.$$

Theorem 2.1 *Fix $r \geq 1$ and $M \geq 2$ and let N denote the dimension of the free, nilpotent Lie algebra on M generators of rank r. Then the derivations*

$$\begin{aligned} E_1 &= \frac{\partial}{\partial x_1} \\ E_2 &= \frac{\partial}{\partial x_2} + \sum_{j \succ 2} P_{2,j} \frac{\partial}{\partial x_j} \\ &\vdots \\ E_M &= \frac{\partial}{\partial x_M} + \sum_{j \succ M} P_{M,j} \frac{\partial}{\partial x_j} \end{aligned}$$

have the following properties

1. *they are homogeneous of weight one with respect to the grading*

$$\mathbf{R}^N = V_1 \oplus \cdots \oplus V_r$$

   introduced at the beginning of the section;

2. *the Hall basis elements E_i they generate satisfy $E_i(0) = \partial/\partial x_i$; in other words, E_1 through E_M are free to step r at 0;*

3. *the graded Lie algebra they generate is isomorphic to $\mathrm{g}_{M,r}$.*

We begin with the first assertion. Since $\frac{\partial}{\partial x_1}$ through $\frac{\partial}{\partial x_M}$ are homogeneous of weight 1, we see that $P_{l,j}\frac{\partial}{\partial x_j}$ is homogeneous of weight 1 for all $1 \leq l \leq M$, $j > M$. Therefore E_1 through E_M are homogeneous of weight one, which is assertion 1. It follows that each E_i is homogeneous with weight equal to its length. As a consequence, $E_i = 0$ whenever $i > N$, for there are no derivations in $\mathbf{R}^N$ with polynomial coefficients which are homogeneous of weight greater than r. In other words the Lie algebra generated by E_1 through E_M is nilpotent of step r. Combining this with assertion 2 immediately gives the third assertion. Therefore the proof is complete once we prove the second assertion. This will be done by induction and requires the following definition.

Definition 2.2 *Let A be any non-constant monomial in the variables x_1, $\ldots$, x_N. Define the minimum order of A by $m(A) = \min\{j : x_j|A\}$. If A is a polynomial with no constant term, then define $m(A)$ to be the maximum of the minimum orders of the monomials of A.*

Lemma 2.3 *If $E_i = [E_j, E_k]$, then*

$$E_i = \frac{\partial}{\partial x_i} + \sum_{c \succ i} P_{i,c}\frac{\partial}{\partial x_c} + \sum_d Q_{i,d}\frac{\partial}{\partial x_d},$$

where $P_{i,c}$ satisfies $k \leq m(P_{i,c}) < i$, and $Q_{i,d}$ is a polynomial with no constant term satisfying $m(Q_{i,d}) < k$.

Proof. We proceed by induction. The lemma is true for the generators, for the polynomials $Q_{i,d}$ are identically zero. Consider the parents of E_i:

$$E_j = \frac{\partial}{\partial x_j} + \sum_{e \succ j} P_{j,e}\frac{\partial}{\partial x_e} + \sum_f Q_{j,f}\frac{\partial}{\partial x_f},$$

and

$$E_k = \frac{\partial}{\partial x_k} + \sum_{g \succ k} P_{k,g} \frac{\partial}{\partial x_g} + \sum_h Q_{k,h} \frac{\partial}{\partial x_h},$$

by the inductive hypothesis.

Now $E_j = [E_r, E_s]$ and so $s \leq k$ by the basis condition. Therefore, we have the relations:

$$s \leq m(P_{j,e}) < j \quad \text{and} \quad m(Q_{j,f}) < s \leq k$$

And trivially,

$$m(P_{k,g}) < k \quad \text{and} \quad m(Q_{k,h}) < k.$$

Examine the terms in the bracket of E_j and E_k.

$$\begin{aligned}
[E_j, E_k] &= \sum_{g \succ k} \frac{\partial}{\partial x_j} P_{k,g} \frac{\partial}{\partial x_g} + \sum_h \frac{\partial}{\partial x_j} Q_{k,h} \frac{\partial}{\partial x_h} \\
&+ \sum_{e \succ j} \sum_{g \succ k} P_{j,e} \frac{\partial}{\partial x_e} P_{k,g} \frac{\partial}{\partial x_g} + \sum_{e \succ j} \sum_h P_{j,e} \frac{\partial}{\partial x_e} Q_{k,h} \frac{\partial}{\partial x_h} \\
&+ \sum_f \sum_{g \succ k} Q_{j,f} \frac{\partial}{\partial x_f} P_{k,g} \frac{\partial}{\partial x_g} + \sum_f \sum_h Q_{j,f} \frac{\partial}{\partial x_f} Q_{k,h} \frac{\partial}{\partial x_h} \\
&- \sum_{e \succ j} \frac{\partial}{\partial x_k} P_{j,e} \frac{\partial}{\partial x_e} \\
&- \sum_f \frac{\partial}{\partial x_k} Q_{j,f} \frac{\partial}{\partial x_f} \\
&- \sum_{g \succ k} \sum_{e \succ j} P_{k,g} \frac{\partial}{\partial x_g} P_{j,e} \frac{\partial}{\partial x_e} - \sum_{g \succ k} \sum_f P_{k,g} \frac{\partial}{\partial x_g} Q_{j,f} \frac{\partial}{\partial x_f} \\
&- \sum_h \sum_{e \succ j} Q_{k,h} \frac{\partial}{\partial x_h} P_{j,e} \frac{\partial}{\partial x_e} - \sum_h \sum_f Q_{k,h} \frac{\partial}{\partial x_h} Q_{j,f} \frac{\partial}{\partial x_f}.
\end{aligned}$$

If A and B are any monomials with $m(A) < k$ then for any t, $A \frac{\partial}{\partial x_t} B$ either vanishes or has minimum order less than k. This implies that the non-zero terms in the third, sixth, and seventh lines have minimum orders $< k$. If A is any monomial satisfying $m(A) < k$ and if $u \geq k$, then $\frac{\partial}{\partial x_u} A$ either vanishes or has minimum order less than k. This implies that the non-zero terms in the first, second, and fifth lines have minimum orders $< k$. The remaining terms are

$$-\sum_{e \succ j} \frac{\partial}{\partial x_k} P_{j,e} \frac{\partial}{\partial x_e}.$$

If $m(P_{j,e}) = k$, then either $i = e$, in which case

$$-\frac{\partial}{\partial x_k} P_{j,e} \frac{\partial}{\partial x_e} = \frac{\partial}{\partial x_i},$$

or $e \succ i$, and

$$-\frac{\partial}{\partial x_k} P_{j,e} \frac{\partial}{\partial x_e} = P_{i,e} \frac{\partial}{\partial x_e}.$$

If $m(P_{j,e}) < k$, then $\frac{\partial}{\partial x_k} P_{j,e}$ is either zero, or it has minimum order less than k. If $m(P_{j,e}) > k$, then $\frac{\partial}{\partial x_k} P_{j,e} = 0$. We conclude that

$$-\sum_{e \succ j} \frac{\partial}{\partial x_k} P_{j,e} \frac{\partial}{\partial x_e} = \frac{\partial}{\partial x_i} + \sum_{e \succ i} P_{i,e} \frac{\partial}{\partial x_e} + \sum_q Q_{j,q} \frac{\partial}{\partial x_q},$$

where $m(Q_{j,q}) < k$. This proves the lemma. (2.3) ∎

Since $P_{i,j}(0) = 0$, for all $i \prec j$, assertion 2 follows immediately from the lemma, finishing the proof of the theorem. (2.1) ∎

3 Applications

In this section we restrict to the case of two generators and apply Theorem 2.1 to derive some results in control theory and Lie algebras. At the end of the section we write down the generators for the Lie algebra $\mathrm{g}_{2,6}$. First, we show that the control system $\dot{x}(t) = (E_1 + uE_2)(x(t))$ is explicitly integrable in closed form. We then show how the theorem can be used to compute the coefficients in the Baker–Campbell–Hausdorff formula with respect to a Hall basis, and how it can be used to compute the universal enveloping algebra of a free, Lie algebra.

We begin with the control theory application. For more details about control systems of the type considered here, see, for example Krener [14].

Theorem 3.1 *Let E_1 and E_2 be vector fields on $\mathbf{R}^N$ which generate the free, nilpotent Lie algebra $\mathrm{g}_{2,r}$ of dimension N. Then any trajectory of the control system*

$$\dot{x}(t) = E_1(x(t)) + u(t)E_2(x(t)), \qquad x(0) = 0 \in \mathbf{R}^N,$$

can be computed explicitly in terms of quadratures of the control $t \to u(t)$.

Proof. We see immediately that $x_1(t) = t$ and $x_2(t) = \int_0^t u(\tau)\ d\tau$. Consider the coefficient of $\frac{\partial}{\partial x_3}$ next. By homogeneity it must be a polynomial of weight 1 (in fact it is the monomial x_1). But all polynomials of weight 1 are already known as explicit functions involving quadratures of the control $t \to u(t)$. Therefore $x_3(t)$ can be computed. It is in fact $x_3(t) = -\frac{1}{2}t^2$. The theorem is easily proved by induction in this fashion as we now show.

More precisely, since E_1 and E_2 are both homogeneous of weight 1, $\dot{x_k}(t)$ is a homogenous polynomial function of weight strictly less than the weight of x_k. But by induction, any such polynomial can be written as an explicit function of t and of quadratures of the control $t \to u(t)$. Integrating, shows that the same can be said of $x_k(t)$, completing the induction and finishing the proof. (3.1) ∎

We turn now to the Baker–Campbell–Hausdorff formula. Let A denote the free, associative algebra in two variables E_1 and E_2, and let $\bar{A}$ denote the algebra of formal power series in the noncommuting variables E_1 and E_2. It turns out that there is an $F \in \bar{A}$ such that $\exp E_1 \exp E_2 = \exp F$ and that F turns out to be a sum of Lie elements; that is, a sum of elements which are in the Lie algebra generated by E_1 and E_2. The problem is to find a useful representation of F. This is a very basic problem and many such formulas exist. For example F may be written as a sum of words in the algebra A; or F may be written as a sum of Lie elements in the algebra A; or F may be written as a sum of Lie basis elements. We consider the third possibility here. Newman and Thompson [15] discuss several approaches that have been used to compute these expansions.

Note that the assertion $\exp(E)(y^1) = y^2$ implies that y^1 and y^2 satisfy the following boundary value problem

$$\dot{x}(t) = E(x(t)), \qquad x(0) = y^1, \qquad x(1) = y^2.$$

Consider

$$\begin{aligned} \exp(E_2)\exp(E_1)(0) &= \exp(E_2)(y^1) \\ &= y^2, \end{aligned}$$

where y^1 is the vector $y^1 = (1, 0, \ldots, 0) \in \mathbf{R}^N$, and $y^2 \in \mathbf{R}^N$ is a point, which by theorem 3.1, is explicitly computable in closed form. On the other hand, the equation

$$\exp\left(\sum_{i=1}^{N} \alpha_i E_i\right)(0) = y^2$$

may be integrated in closed form, since Theorem 2.1 implies that $\dot{x_i}(t)$ is equal to a function of only those x_j which have weight less than or equal to the weight of x_i. Therefore the α's satisfy some linear equation in $\mathbf{R}^N$, which we write $T\alpha = y^2$. Observe that T is lower triangular; hence, the equation $T\alpha = y^2$ always has a solution. Since these formulas for the exponential of a vector field hold on all of $\mathbf{R}^N$, we could compute the coefficients α_i by flowing from an arbitrary point instead of the origin. The answers are, of course, the same. We have shown:

Theorem 3.2 *Let E_1 and E_2 be the generators of $\mathrm{g}_{2,r}$ and let $E_1, \dots, E_N$ be the Hall basis. Then*

1. *There exists constants $\alpha_1, \dots, \alpha_N$ such that*

$$\exp(E_2)\exp(E_1) = \exp(\sum_{i=1}^{N} \alpha_i E_i);$$

   that is, $\log(\exp E_1 \exp E_2)$ *is a Lie element;*

2. *the α's satisfy the linear equation $T\alpha = y$, where T is a lower triangular matrix, and $y \in \mathbf{R}^N$ is a point which may be computed in closed form.*

At the end of the section, we give the first twenty three coefficients.

Finally we prove the Poincare–Birkhoff-Witt theorem in the special case that the Lie algebra is equal to $\mathrm{g}_{2,r}$. In this case the universal enveloping algebra is equal to the free associative algebra generated by E_1 and E_2, and a basis is given by the next theorem. For the rest of this section, we write $D_i = \partial/\partial x_i$.

Theorem 3.3 *The elements $E_{\alpha_n} \cdots E_{\alpha_2} E_{\alpha_1}$, where $\alpha_{j+1} \geq \alpha_j$, form a basis for the free associative algebra generated by E_1 and E_2.*

Proof. The theorem is an immediate consequence of the following two lemmas. The next lemma shows that these elements are linearly independent, while the following one shows that they span. The proof that the elements span is standard (see Jacobson [13]), and is outlined here only for completeness.

Lemma 3.4

$$E_{\alpha_n} \cdots E_{\alpha_2} E_{\alpha_1} = D_{\alpha_n} \cdots D_{\alpha_2} D_{\alpha_1} + \sum_{|I| \leq n} P_I D^I,$$

where each P_I is a polynomial with no constant term satisfying $m(P_I) < \alpha_1$.

Proof. By induction on n. The case $n = 1$ is exactly Lemma 2.3. Consider the product $E^{\overline{\alpha}} = E_{\alpha_{n-1}} \cdots E_{\alpha_2} E_{\alpha_1}$, and write $D^{\overline{\alpha}}$ for the analogous products of the D's. By hypothesis, the minimum orders of the coefficients of the non-constant terms are all less than α_1. By Lemma 2.3, E_{α_n} is of the form

$$E_{\alpha_n} = D_{\alpha_n} + \sum_{\theta > \alpha_n} R_\theta D_\theta,$$

for some R_θ's satisfying $R_\theta(0) = 0$. Therefore,

$$\begin{aligned} E_{\alpha_n} E^{\overline{\alpha}} &= D_{\alpha_n} D^{\overline{\alpha}} + \sum_{|I| \leq n-1} D_{\alpha_n} P_I D^I + \sum_{|I| \leq n-1} P_I D_{\alpha_n} D^I \\ &+ \sum_{\theta > \alpha_n} \sum_{|I| \leq n-1} R_\theta D_\theta P_I D^I + \sum_{\theta > \alpha_n} \sum_{|I| \leq n-1} R_\theta P_I D_\theta D^I. \end{aligned}$$

Since $\alpha_n \geq \alpha_{n-1} \geq \alpha_1$ and $\theta > \alpha_n$, we see that each of $D_{\alpha_n} P_I$ and $D_\theta P_I$ either vanishes, or has minimum order less than α_1. (3.4) ∎

Lemma 3.5 *The non-increasing products of Hall basis elements,* $E_{\alpha_n} \cdots E_{\alpha_2} E_{\alpha_1}$*, where* $\alpha_{j+1} \geq \alpha_j$*, span the free associative algebra generated by* E_1 *and* E_2*.*

Proof. Consider an arbitrary word in the generators:

$$A = E_2^{b_1} E_1^{a_1} \cdots E_2^{b_n} E_1^{a_n},$$

where b_1 and a_n can be zero. In order to get A into non-increasing order, we must move all of the E_2's to the left. Start with the leftmost E_2 which is to the right of an E_1, that is, $A = E_2^{b_1} E_1^{a_2} E_2 W$. There is a non-increasing product to the left of the first element which we wish to move. By interchanging the E_1 and E_2, we obtain

$$A = E_2^{b_1} E_1^{a_1 - 1} E_2 E_1 W - E_2^{b_1} E_1^{a_1 - 1} E_3 W.$$

If we define the *badness* of a word to be the number of transpositions necessary to put it into non-increasing order, then the badness of the first word above is less than the badness of A, and the length of the second word less than the length of A. Note also that the sum of the weights of the Hall basis elements in each word is the same. If we continue to move the E_2 to the left, we will obtain

$$A = E_2^{b_1 + 1} E_1^{a_1} W - \sum_{i=1}^{a_1 - 1} E_2^{b_1} E_1^{a_1 - i} E_3 E_1^i W.$$

Define the *graded badness* B of a linear combination of words to be a vector whose ith component B_i is equal to the sum of the badnesses of the words of length i. We give the collection of B's the lexicographic ordering; a bad word of length l is worse than any word of any length $< l$. Note that the graded badness of A has decreased, and that the total badness of the remaining terms is bounded by the cube of the length of A. In general, the graded badness will decrease, and the total badness will be bounded, hence the process will terminate. (3.5) ∎

We have shown that the elements $E_{\alpha_n} \cdots E_{\alpha_1}$ span and are linearly independent, completing the proof of Theorem 3.3. (3.3) ∎

In the next lemma we show just how natural the Hall basis is for problems of this type. The words in the different expressions for A naturally form a binary tree. At each non-leaf node, there are two branches: one where the length is the same and the badness decreases, one where the out-of-order pair is replaced by its bracket. While it might be believable that the nodes of this tree can be written as linear combinations of words in the Hall basis elements, a much stronger and more surprising thing is true.

Lemma 3.6 *Every node in this tree* is *a word in the Hall basis elements.*

Proof. By induction, for A is a word in E_1 and E_2. The basis element which gets moved at a given node has a history of swaps and/or brackets. Since it is moving and/or bracketing through a non-increasing sequence product of Hall basis elements, the last term it bracketed with had a lower order than the one it faces now. Therefore, their bracket is another Hall basis element. The possibilities now are:

1. the out-of order elements can swap,
2. they can both be replaced with their bracket,
3. the moving element is now part of a non-increasing sequence, and some E_2 further to the right begins to move, or,
4. the word is in non-increasing order; it is a leaf node.

Each case is consistent with the conclusion of the lemma. (3.6) ∎

We conclude this section by giving an example of two vector fields defined in $\mathbf{R}^{23}$ which generate the free, nilpotent Lie algebra on two generators of rank 6. The basis elements are E_1, E_2, $E_3 = [E_2, E_1]$, $E_4 = [E_3, E_1]$, $E_5 = [E_3, E_2]$, $E_6 = [E_4, E_1]$, $E_7 = [E_4, E_2]$, $E_8 = [E_5, E_2]$, $E_9 =$

$[E_6, E_1]$, $E_{10} = [E_6, E_2]$, $E_{11} = [E_7, E_2]$, $E_{12} = [E_8, E_2]$, $E_{13} = [E_4, E_3]$, $E_{14} = [E_5, E_3]$, $E_{15} = [E_{10}, E_2]$, $E_{16} = [E_{11}, E_2]$, $E_{17} = [E_{12}, E_2]$, $E_{18} = [E_4, E_3]$, $E_{19} = [E_5, E_3]$, $E_{20} = [E_6, E_3]$, $E_{21} = [E_7, E_3]$, $E_{22} = [E_{18}, E_3]$, and $E_{23} = [E_5, E_4]$. The generators are $E_1 = D_1$ and

$$
\begin{aligned}
E_2 &= D_2 - x_1 D_3 + \frac{1}{2}\, x_1^2\, D_4 + x_1 x_2 D_5 - \frac{1}{6}\, x_1^3\, D_6 \\
&- \frac{1}{2}\, x_1^2\, x_2 D_7 - \frac{1}{2}\, x_1 x_2^2\, D_8 + \frac{1}{24}\, x_1^4\, D_9 + \frac{1}{6}\, x_1^3\, x_2 D_{10} \\
&+ \frac{1}{4}\, x_1^2\, x_2^2\, D_{11} + \frac{1}{6}\, x_1 x_2^3\, D_{12} - \frac{1}{120}\, x_1^5\, D_{13} \\
&- \frac{1}{24}\, x_1^4\, x_2 D_{14} - \frac{1}{12}\, x_1^3\, x_2^2\, D_{15} - \frac{1}{12}\, x_1^2\, x_2^3\, D_{16} \\
&- \frac{1}{24}\, x_1 x_2^4\, D_{17} - \frac{1}{2}\, x_1^2\, x_3 D_{18} - x_1 x_2 x_3 D_{19} + \frac{1}{6}\, x_1^3\, x_3 D_{20} \\
&+ \frac{1}{2}\, x_1^2\, x_2 x_3 D_{21} + \frac{1}{2}\, x_1 x_2^2\, x_3 D_{22} - x_1 x_2 x_4 D_{23}
\end{aligned}
$$

Using these vector fields, it is easy to compute that the Baker–Campbell–Hausdorff coefficients $\alpha_1, \ldots, \alpha_{23}$ in the expansion

$$
\exp(E_2)\exp(E_1) = \exp(\sum_{i=1}^{N} \alpha_i E_i)
$$

are

$$
\begin{gathered}
\alpha_1 = 1, \alpha_2 = 1, \alpha_3 = -\frac{1}{2}, \alpha_4 = \frac{1}{12}\alpha_5 = -\frac{1}{12}, \alpha_6 = 0, \\
\alpha_7 = \frac{1}{24}, \alpha_8 = 0, \alpha_9 = -\frac{1}{720}, \alpha_{10} = -\frac{1}{180}, \alpha_{11} = \frac{1}{180}, \alpha_{12} = \frac{1}{720}, \\
\alpha_{13} = 0, \alpha_{14} = -\frac{1}{1440}, \alpha_{15} = -\frac{1}{360}, \alpha_{16} = -\frac{1}{1440}, \alpha_{17} = 0, \alpha_{18} = -\frac{1}{120}, \\
\alpha_{19} = -\frac{1}{360}, \alpha_{20} = 0, \alpha_{21} = -\frac{1}{240}, \alpha_{22} = -\frac{1}{720}, \text{ and } \alpha_{23} = \frac{1}{240}.
\end{gathered}
$$

References

[1] A. Bressan, *On the Taylor approximation of control systems,* to appear.

[2] P. E. Crouch, *Dynamical realizations of finite Volterra series,* SIAM J. Control Optim., **19** (1981), 177–202.

[3] P. E. Crouch *Graded vector spaces and applications to the approximations of nonlinear systems,* to appear.

[4] G. B. Folland and E. M. Stein, *Estimates for the $\bar{\partial}_b$ complex and analysis of the Heisenberg group,* Comm. Pure Appl. Math. **27** (1974), 429–522.

[5] R. W. Goodman, "Nilpotent Lie Groups: Structure and Applications to Analysis," Springer-Verlag (1976), Lecture Notes in Mathematics, No. 562.

[6] R. Grossman and R. G. Larson, *Solving nonlinear equations from higher order derivations in linear stages,* to appear in Adv. in Math.

[7] M. Hall, *A basis for free Lie rings and higher commutators in free groups,* Proc. Amer. Math. Soc., **1** (1950), 575–581.

[8] B. Helffer and J. Nourrigat, *Hypoellipticité maximale pour des operateurs polynomes de champs de vectuers,* Birkhaüser (1985) Boston.

[9] H. Hermes, *Control systems which generate decomposible Lie algebras,* J. Diff. Eqns., **44** (1982), 166–187.

[10] H. Hermes, *Nilpotent approximations of control systems and distributions,* SIAM J. Control Optim., **24** (1986), 731–736.

[11] H. Hermes, A. Lundell, and D. Sullivan, *Nilpotent bases for distributions and control systems,* J. Diff. Equations, **55** (1984), 385–400.

[12] L. Hörmander, *Hypoelliptic second order differential equations,* Acta. Math., **119** (1968), 147–171.

[13] N. Jacobson, "Lie algebras," John Wiley and Sons, New York, 1962.

[14] A. Krener, *Bilinear and nonlinear realizations of input-output maps,* SIAM J. Control Optim., **13** (1975), 827–834.

[15] M. Newman and R. Thompson, *Numerical values of Goldberg's coefficients in the series for* $\log(\exp x \exp y)$, Math. of Computation, **48** (1987), 265-271.

[16] C. Rockland, *Intrinsic nilpotent approximation,* Acta Applicandae Math. **8** (1987), 213-270.

[17] L. P. Rothschild, *A criterion for hypoellipticity of operators constructed from vector fields,* Comm. in PDE, **4** (1979), 645–699.

[18] L. P. Rothschild and E. M. Stein, *Hypoelliptic differential operators and nilpotent groups,* Acta. Math., **37** (1977), 248–315.

[19] V. S. Varadarajan, "Lie Groups, Lie Algebras, and their Representations," Prentice-Hall, Inc., 1974.