Solving nonlinear equations from higher order derivations in linear stages

Robert Grossman*
University of California, Berkeley

Richard G. Larson†
University of Illinois at Chicago

September 9, 1987

This is a draft of a paper which later appeared in Advances in Mathematics, Volume 82, 1990, pages 180-202.

This paper is concerned with the effective computation of approximations to higher order derivations. Let $\overline{R}$ be the ring of real smooth functions on some vector space V, with coordinates $x_1, \ldots, x_N$, and let R be the subring of real polynomial functions on V. Given integers $d, r > 0$ and several derivations $F_1, \ldots, F_M$ of $\overline{R}$, we want to find derivations $E_1, \ldots, E_M$ of R such that

$$E_\gamma(a)|_{x=0} = F_\gamma(a)|_{x=0}, \qquad (*)$$

for all polynomials $a \in R$ of degree less than or equal to d and all higher order derivations $E_\gamma = E_{\gamma_s} \cdots E_{\gamma_1}$ and $F_\gamma = F_{\gamma_s} \cdots F_{\gamma_1}$, where $\gamma = (\gamma_1, \ldots, \gamma_s)$, and $1 \le \gamma_i \le M$, of length s less or equal to r. The reason for wanting such derivations $E_1, \ldots, E_M$ is simple. Given several derivations $F_1, \ldots, F_M$, it is useful to have derivations $E_1, \ldots, E_M$ which are good local approximations to the F_i and which are easy to compute with. Notice that since the right hand sides of Equations $(*)$ are known, while the left hand sides involve the unspecified coefficients of the polynomial functions, the equations are equivalent to a system of nonlinear algebraic equations involving the coefficients. It is the purpose of this paper to give an algorithm which solves such a system, when a solution exists. In fact, as the title suggests, this algorithm will work by solving a sequence of r linear systems.

*This author is a National Science Foundation Postdoctoral Research Fellow.
†Research partially supported by National Science Foundation Grant DMS-8701805.

We next describe three situations in which the Equations $(*)$ arise. The first situation involves nilpotent approximations. Given a multi-index $\gamma = (\gamma_1, \ldots, \gamma_r)$, let

$$F_{[\gamma]} = [F_{\gamma_r}, \cdots [F_{\gamma_2}, F_{\gamma_1}] \cdots]$$

denote a Lie bracket of *length* $|\gamma| = r$. The Lie bracket $F_{[\gamma]}$ may be identified with a vector field on V. Therefore, for any $x \in V$, $F_{[\gamma]}(x)$ may be viewed as a vector; that is, an element of V. Assume that there is an integer $r > 0$ such that

$$\operatorname{span}\{F_{[\gamma]}(0) : |\gamma| \le r\} = V.$$

Equations $(*)$, with $d = 1$, become

$$E_\gamma(x_\mu)|_{x=0} = F_\gamma(x_\mu)|_{x=0},$$

for $\mu = 1, \ldots, N$. Since any Lie bracket $F_{[\gamma]}$ of length r may be expanded

$$F_{[\gamma]} = \sum_{|\alpha|=r} c_\alpha F_\alpha$$

where c_α are scalars, this implies

$$E_{[\gamma]}(x_\mu)|_{x=0} = F_{[\gamma]}(x_\mu)|_{x=0}$$

for $\mu = 1, \ldots, N$. Since $x_1, \ldots, x_N$ is the dual basis for V, this in turn implies that the corresponding Lie brackets agree at the origin. It is easy to arrange for the $E_1, \ldots, E_M$ to generate a nilpotent Lie algebra. Therefore, we have one means of generating nilpotent approximations to systems of vector fields.

Nilpotent approximations have become an important tool in control theory. Krener [12] was the was first to make explicit use of nilpotent Lie algebras in control theory, with other important contributions made by Hermes [8] and [9], Crouch [2], Bressan [1], and Hermes, Lundell and Sullivan [10]. The basic idea is simple to describe. Consider a control system evolving in $\mathbf{R}^N$

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

where F_1 and F_2 are two vector fields defined in neighborhood of the origin of $\mathbf{R}^N$ and $t \to u(t)$ is a control. Suppose for a moment that

1. 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

2. the E's generate a nilpotent Lie algebra of step r.

Then it is can be shown easily that the trajectory $t \to x(t)$ of the control system defined by the F's is close for small time to the trajectory $t \to y(t)$ of the corresponding trajectory defined by the E's in the sense that that the estimate

$$|y(t) - x(t)| \le Ct^r,$$

holds for small time t, for some constant C. If the E's satisfy some mild technical conditions (they have polynomial coefficients and are homogenous of weight 1), then it can be shown easily that the trajectories of the E system may be explicitly computed by quadrature involving the control u. In general, vector fields E_1 and E_2 with prescribed Lie brackets at the origin do not exist. Equations $(*)$ provide sufficient, but not necessary conditions for prescribing Lie brackets at a point.

Nilpotent approximations have also been used to study the hypoellipticity of partial differential equations. This was first done in Folland and Stein [4] and extended with contributions by Rothschild and Stein [15], Rothchild [14], and Rockland [13]. 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 [15] 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 $\widetilde{F}_1, \ldots, \widetilde{F}_M$ defined in a larger space $\mathbf{R}^{\widetilde{N}}$ with the property that the latter vector fields are free (in an appropriate sense) at a given point. It turns out that the vector fields $\widetilde{F}_1, \ldots, \widetilde{F}_M$ are well approximated by the generators $E_1, \ldots, E_M$ of a free, nilpotent Lie algebra $g_{M,r}$, 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.

Another application occurs in symbolic computation. The explicit computation of derivations F_i acting on general functions may be quite difficult. On the other hand, for many applications, it may be enough to make use of the E_i acting on polynomials. In the last section we give explicit formulas for such actions.

The arguments in this paper require a certain amount of combinatorics. We feel that this is made most palatable by encoding the information into finite rooted trees. In fact, the relation between finite rooted trees and higher order derivations is central to this paper. The necessary background material is covered in Section 1. Sections 2 and 3 contain the combinatorial lemmas which are the basis of the proof of the main theorem. The reader may want to skip the proofs in this section during the first reading. Recall that we are interested in higher order derivations with polynomial coefficients and certain nonlinear algebraic equations involving these coefficients. Section 4 presents the nonlinear algebraic equations which arise when higher order derivations act on linear monomials. The main theorem is stated and proved in Section 5. This expresses a higher order derivation acting on a linear monomial and evaluated at zero in terms of lower order expressions of the same type. Section 6 gives an analogous result for higher order derivations acting on arbitrary monomials.

1 Higher order derivations, trees, and Lie algebras

In this section we review some facts about graded Lie algebras, following Goodman [5]. We also describe an algebra structure on a family of rooted trees associated with higher derivations, and define some homomorphisms connecting this algebra to other algebraic structures we use.

We first make a number of definitions. Let V be a real vector space with a direct sum decomposition

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

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

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

these are the polynomials *homogenous of weight m.*

Fix a homogeneous basis $e_1, \ldots, e_N$ of V and let $x_1, \ldots, x_N$ denote the dual basis of V^*. Note that e_i is of weight l in case $e_i \in V_l$; we write $\mathrm{wt}(e_i) = l$. Let $f = (f_1, \ldots, f_N)$, be a multi-index, where

$$f_1 \geq 0, \ldots, f_N \geq 0$$

and denote

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

The *weight* of x^f is given by

$$\mathrm{wt}(x^f) = \sum_i f_i \mathrm{wt}(x_i).$$

For example, if $x_i \in R_l$, then $x_i^f \in R_{fl}$, since $fl = \mathrm{wt}(x_i^f)$. 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 $x_1, \dots, x_N$, any derivation E of the ring R is determined by its action on V^*. Therefore the derivation E may be written

$$E = \sum_{\mu=1}^{N} b^\mu D_\mu,$$

where

$$b_\mu = E(x_\mu), \text{ and } D_\mu = \frac{\partial}{\partial x_\mu}.$$

In other words the derivations of R are simply the vector fields on V with polynomial coefficients.

Fix M derivations $E_1, \dots, E_M$ of the ring R. We now define the higher order derivations generated by $E_1, \dots, E_M$. Let $\mathbf{R}{<}E_1, \dots, E_M{>}$ denote the polynomial algebra over $\mathbf{R}$ in the non-commuting variables $E_1, \dots, E_M$, that is, the free noncommutative algebra generated by $E_1, \dots, E_M$. If $p \in \mathbf{R}{<}E_1, \dots, E_M{>}$, then $p(E_1, \dots, E_M)$ induces a $\mathbf{R}$-linear map of R to itself, which we call a *higher order derivation,* and which we denote by $\chi(p)$. We denote the algebra of higher order derivations of R generated by $E_1, \dots, E_M$ by

$$\mathbf{Diff}(E_1, \dots, E_M; R).$$

The map

$$\chi : \mathbf{R}{<}E_1, \dots, E_M{>} \longrightarrow \mathbf{Diff}(E_1, \dots, E_M; R)$$

is an algebra homomorphism. We say that a higher order derivation is *homogeneous* of weight m in case

$$p(a \circ \delta_t) = t^m (p(a)) \circ \delta_t, \qquad \text{for all } t \in \mathbf{R} \text{ and all } a \in R.$$

At this point we introduce some notation which will be used in the examples throughout the remainder of the paper.

Example 1 *For $i = 1, \ldots, r$, let d_i denote the dimension of the subspace V_i and let $\{e_{i1}, \ldots, e_{id_i}\}$ and $\{x_{i1}, \ldots, x_{id_i}\}$ denote the basis and dual basis of V_i. Note that the subscripts indicate to which graded component the variable belongs. Throughout the examples, we assume that $E_1, \ldots, E_M$ are derivations of the ring R which are homogenous of weight 1. As an illustration, assume that V is the direct sum of two one dimensional spaces V_1 and V_2. Then $\{x_{11}\}$ and $\{x_{21}\}$ are bases of V_1^* and V_2^*; $E_1 = x_{11}\frac{\partial}{\partial x_{21}}$ and $E_2 = \frac{\partial}{\partial x_{11}}$ are two derivations homogenous of weight 1; and, E_2E_1 is a second order derivation homogeneous of weight 2.*

Let $\mathbf{R}\{\mathcal{LT}(E_1, \ldots, E_M)\}$ be the vector space which has as basis the set of all finite rooted trees, in which all nodes but the root are labeled using E_1, $\ldots$, E_M. (This labeling may have the same E_γ used as a label on more than one node.) We define a multiplication in $\mathbf{R}\{\mathcal{LT}(E_1, \ldots, E_M)\}$ as follows. Since the set of labeled rooted trees is a basis, it is sufficient to describe the product of two such trees. Suppose t_1 and t_2 are two labeled rooted trees. Let $s_1, \ldots, s_r$ be the children of the root of t_1. If t_2 has $n+1$ nodes (counting the root), there are $(n+1)^r$ ways to attach the r subtrees of t_1 which have $s_1, \ldots, s_r$ as roots to the labeled tree t_2 by making each s_i the child of some node of t_2, keeping the original labels from t_1. The product t_1t_2 is defined to be the sum of these $(n+1)^r$ labeled trees. It can be shown (see [7]) that this product is associative, and that the tree whose only node is the root is a multiplicative identity.

Since $\mathbf{R}{<}E_1, \ldots, E_M{>}$ is the free associative algebra generated by E_1, $\ldots$, E_M, there is a unique algebra homomorphism

$$\phi : \mathbf{R}{<}E_1, \ldots, E_M{>} \longrightarrow \mathbf{R}\{\mathcal{LT}(E_1, \ldots, E_M)\}$$

sending E_γ to the tree with two nodes: the root, and its one child, labeled with E_γ.

We define a map

$$\psi : \mathbf{R}\{\mathcal{LT}(E_1, \ldots, E_M)\} \longrightarrow \mathbf{Diff}(E_1, \ldots, E_M; R)$$

as follows:

1. Given a labeled rooted tree $t \in \mathcal{LT}(E_1, \ldots, E_M)$ with $m+1$ nodes, name the root 0, and name the other nodes $1, \ldots, m$. To each node k of t other than the root, associate the summation index μ_k.

2. Let k be a node of t, labeled with E_{γ_k} if k is not the root, and suppose that $l, \ldots, l'$ are the children of k. Let

$$R(k) = \begin{cases} D_{\mu_l} \cdots D_{\mu_{l'}} b_{\gamma_k}^{\mu_k} & \text{if } k \text{ is not the root;} \\ D_{\mu_l} \cdots D_{\mu_{l'}} & \text{if } k \text{ is the root.} \end{cases}$$

Note that $R(k)$ depends on the μ_i; note also that if $k > 0$, then $R(k) \in R$.

3. Define

$$\psi(t) = \sum_{\mu_1, \ldots, \mu_m = 1}^{N} R(m) \cdots R(1) R(0).$$

4. Extend ψ to all of $\mathbf{R}\{\mathcal{LT}(E_1, \ldots, E_M)\}$ by linearity.

Proposition 2 *The map*

$$\psi : \mathbf{R}\{\mathcal{LT}(E_1, \ldots, E_M)\} \longrightarrow \mathbf{Diff}(E_1, \ldots, E_M; R)$$

is an algebra homomorphism.

PROOF: Let $t_1, t_2 \in \mathcal{LT}(E_1, \ldots, E_M)$. We must show that $\psi(t_1)\psi(t_2) = \psi(t_1 t_2)$. Fix $a \in R$ and consider $\psi(t_1)\psi(t_2)(a)$. From the definition of ψ, if t_2 has $n + 1$ nodes,

$$\psi(t_2)(a) = \sum R(n) \cdots (R(0)a).$$

Now consider $\psi(t_1)\psi(t_2)(a)$. Each node of t_1, other than the root, adds a summation index to the expression for $\psi(t_1)\psi(t_2)(a)$. Call ν_i the summation index added by the node i. Note that both $\psi(t_1)\psi(t_2)(a)$ and $\psi(t_1 t_2)(a)$ involve sums over $m + n$ summation indices.

A node i of t_1 other than the root adds a factor $R(i)$ to each term in the expression for $\psi(t_1)\psi(t_2)(a)$. The same factor will also appear in each term of the expression for $\psi(t_1 t_2)(a)$, since the node i will appear as a non-root node of each tree in $t_1 t_2$. Suppose that the root of t_1 has r children. When $\psi(t_1)$ is applied to each term $R(n) \cdots (R(0)a)$ of $\psi(t_2)(a)$, the result will be $D_{\nu_1} \cdots D_{\nu_r}(R(n) \cdots (R(0)a))$, multiplied by the terms corresponding to the non-root nodes of t_1. The product rule for differentiation implies that $D_{\nu_1} \cdots D_{\nu_r}(R(n) \cdots (R(0)a))$ is the sum of terms, one corresponding to each possible way that the D_{ν_k} can be applied to the factors $R(n), \ldots, R(1)$, and $R(0)a$. But these correspond exactly to the possible ways in which the subtrees of t_1 whose roots are the children of the root of

t_1 can be attached to the nodes of t_2 in the definition of the product. In each term of $D_{\nu_1} \cdots D_{\nu_r}(R(n) \cdots (R(0)a))$, the factor $D_{\nu_{i_1}} \cdots D_{\nu_{i_k}} R(j)$ (or $D_{\nu_{i_1}} \cdots D_{\nu_{i_k}}(R(0)a))$ can be seen to be a factor which would arise in the computation of the term of $\psi(t_1 t_2)(a)$ which comes from the tree in $t_1 t_2$ in which the nodes $i_1, \ldots, i_k$ of t_1 are attached to node j (or to the root) of t_2. Thus both $\psi(t_1)\psi(t_2)(a)$ and $\psi(t_1 t_2)(a)$ equal the summation over equivalent sets of summation indices of the same sums of terms. This completes the proof of the proposition.

Proposition 3

$$\chi = \psi \circ \phi$$

Proof: It can be immediately checked that the two homomorphisms

$$\chi,\ \psi \circ \phi : \mathbf{R}{<}E_1, \ldots, E_M{>} \longrightarrow \mathbf{Diff}(E_1, \ldots, E_M; R)$$

agree on the generating set $E_1, \ldots, E_M$ of $\mathbf{R}{<}E_1, \ldots, E_M{>}$. The proposition follows.

2 A recurrence relation

In this section we give a recurrence relation which allows us to compute the action of a tree on a monomial in terms of the actions of smaller trees on monomials.

If $f = (f_1, \ldots, f_N)$, let x^f denote the monomial $x_1^{f_1} \cdots x_N^{f_N}$, and let $t \in \mathcal{LT}(E_1, \ldots, E_M)$ be a labeled tree. If the root of t has r children, removing the root of t and removing the labels from the roots of the resulting r trees gives r labeled trees which we will usually denote $t_1, \ldots, t_r$. We will call these trees the *trees associated with the children of the root of t.* If E_{γ_i} is the label removed from the root of the i^{th} tree to produce the labeled tree t_i, we will call E_{γ_i} the *label associated with the root of t_i in the tree t.*

We will prove

Theorem 4 *Let $t \in \mathcal{LT}(E_1, \ldots, E_M)$ be a labeled tree whose root has r children, which are labeled $E_{\gamma_1}, \ldots, E_{\gamma_r}$. Let the derivation*

$$E_{\gamma_i} = \sum_{\mu, f} c_f^{\mu}(\gamma_i) x^f D_\mu,$$

and let $t_1, \ldots, t_r$ be the labeled trees associated with the children of the root of t. Then

$$\psi(t)x^f = \sum c_{g_1}^{\mu_1}(\gamma_1)\cdots c_{g_r}^{\mu_r}(\gamma_r)(\psi(t_1)x^{g_1})\cdots(\psi(t_r)x^{g_r})D_{\mu_1}\cdots D_{\mu_r}x^f,$$

where the sum is taken over all $\mu_1, \ldots, \mu_r = 1, \ldots, N$, and over all vectors $g_1, \ldots, g_r$ of non-negative integers. Note that for each i and μ, the coefficient $c_g^\mu(\gamma_i)$ is non-zero for only finitely many g.

PROOF: From the definition of ψ in Section 1 we know that

$$\begin{aligned} \psi(t)x^f &= \sum_{\mu_1,\ldots,\mu_k} R(k)\cdots R(0)(x^f) \\ &= \sum_{\mu_1,\ldots,\mu_k} R(k)\cdots R(1)D_{\mu_1}\cdots D_{\mu_r}x^f. \end{aligned}$$

If $l_{i1}, \ldots, l_{is_i}$ are the numbers associated with the nodes in the sub-tree t_i of t, and if E_{γ_i} is the label that was associated with the root of t_i in the tree t, it is clear that

$$R(l_{i1})\cdots R(l_{is_i}) = \sum_{g_i} c_{g_i}^{\mu_i}(\gamma_i)\psi(t_i)x^{g_i}.$$

Grouping the factors in $R(k)\cdots R(1)$ so that the $R(l)$ corresponding to nodes in the same subtree are grouped together, we see that the theorem follows.

Note that the factor $D_{\mu_1}\cdots D_{\mu_r}x^f$ which appears in the theorem is simply a monomial whose coefficient is a product of binomial coefficients which depend on μ_i and f_j. We can give an explicit value for this in case $r = |f|$, where we denote

$$|f| = \sum_{i=1}^{N} f_i.$$

We say that $(\mu_1, \ldots, \mu_r) \rightharpoondown f$ if 1 occurs f_1 times among the $\mu_i, \ldots, N$ occurs f_N times among the μ_i. Note that if $(\mu_1, \ldots, \mu_r) \rightharpoondown f$, then $r = |f|$.

Definition 5 *Suppose that $(\mu_1, \ldots, \mu_r) \rightharpoondown f$. Then*

$$\epsilon(\mu_1, \ldots, \mu_r) = f_1!\cdots f_N!.$$

Lemma 6 *Suppose $r = |f|$. Then*

$$D_{\mu_1}\cdots D_{\mu_r}x^f = \begin{cases} \epsilon(\mu_1, \ldots, \mu_r) & \text{if } (\mu_1, \ldots, \mu_r) \rightharpoondown f, \\ 0 & \text{otherwise.} \end{cases}$$

PROOF: We prove this by induction on r. It is clear for $r = 1$. Let $k = \mu_r$. Then $D_{\mu_r} x^f = f_k x^{f'}$, where

$$f'_j = \begin{cases} f_j - 1 & \text{if } j = k; \\ f_j & \text{otherwise.} \end{cases}$$

Note that $(\mu_1, \ldots, \mu_r) \rightharpoondown f$ if and only if $(\mu_1, \ldots, \mu_{r-1}) \rightharpoondown f'$. By induction

$$D_{\mu_1} \cdots D_{\mu_{r-1}} x^{f'} = \begin{cases} f'_1! \cdots f'_N & \text{if } (\mu_1, \ldots, \mu_{r-1}) \rightharpoondown f', \\ 0 & \text{otherwise.} \end{cases}$$

It follows that

$$\begin{aligned} D_{\mu_1} \cdots D_{\mu_r} x^f &= f_k D_{\mu_1} \cdots D_{\mu_{r-1}} x^{f'} \\ &= f'_1! \cdots f_k \, f'_k! \cdots f'_N! \\ &= f_1! \cdots f_N! \\ &= \epsilon(\mu_1, \ldots, \mu_r), \end{aligned}$$

if $(\mu_1, \ldots, \mu_r) \rightharpoondown f$, and equals 0 otherwise. This completes the proof of the lemma.

We will need the following corollary to Lemma 6 in a later section.

Lemma 7

$$D_{\mu_1} \cdots D_{\mu_r} x^f|_{x=0} = \begin{cases} \epsilon(\mu_1, \ldots, \mu_r) & \text{if } (\mu_1, \ldots, \mu_r) \rightharpoondown f, \\ 0 & \text{otherwise.} \end{cases}$$

PROOF: If $|f| < r$, then differentiation of the monomial r times results in 0.

If $|f| > r$, then differentiation of the monomial r times results in a non-constant monomial, which evaluates to 0 when we set $x = 0$.

In the remaining case, differentiation results in a constant which we evaluate using Lemma 6. This completes the proof of the lemma.

3 Some combinatorial lemmas

In this section we prove lemmas which give recurrence relations on certain sums of trees. These lemmas will be used in the proofs of Theorems 20 and 22.

If t is a tree with $m + 1$ nodes, and $i_1 < \ldots < i_m$ are positive integers, by a *heap labeling* of t by $i_1, \ldots, i_m$ we mean an assignment of $i_1, \ldots, i_m$

to the nodes of t other than the root such that if j is assigned to a node of t, and k is assigned to a child of that node, then $j < k$. Ij j is assigned to a node, we will sometimes say that the node is *named* j, and sometimes simply refer to the node as j.

Definition 8 *Let $i_1 < \ldots < i_m$ be positive integers. Then*

$$\sigma_m(i_1, \ldots, i_m) = \sum t,$$

where the sum ranges over all heap labeled trees t with $m+1$ nodes, labeled by $i_1, \ldots, i_m$.

Let $i_1 < \ldots < i_m$ be positive integers. Then

$$\sigma_{m,k}(i_1, \ldots, i_m) = \sum t,$$

where the sum ranges over all heap labeled trees t with $m+1$ nodes, labeled by $i_1, \ldots, i_m$, whose root has exactly k children.

Note that

$$\sigma_m(i_1, \ldots, i_m) = \sum_{k=1}^{m} \sigma_{m,k}(i_1, \ldots, i_m).$$

We will derive a recursive description of $\sigma_{m,1}(1, \ldots, m)$. Note that the root of each tree in this expression has exactly one one child, which is named 1. We will denote by e the tree with exactly one (unnamed) node. We will denote by e_k the tree with exactly one node, which is named k. If t is a tree with exactly one node (that is, if t is e, or e_k for some k), and if $X = \{t_1, \ldots, t_r\}$ is a set of trees, then $t \leftharpoondown X$ is the tree formed by making the roots of $t_1, \ldots, t_r$ children of the unique node of t. Note that it is possible to define the operation $\leftharpoondown$ even when the tree t has more than one node; see [7] for details.

If t_1 and t_2 are two heap labeled trees which are labeled with disjoint sets of positive integers, we will denote by $t_1 \odot t_2$ the heap labeled tree formed by identifying the roots of t_1 and t_2. Note that if t_1 has $m_1 + 1$ nodes and t_2 has $m_2 + 1$ nodes, then $t_1 \odot t_2$ has $m_1 + m_2 + 1$ nodes. We extend the definition of the operation $\odot$ to linear combinations of trees by requiring that it be bilinear.

Lemma 9

$$\sigma_{m,1}(1, \ldots, m) = \sum e \leftharpoondown \{e_1 \leftharpoondown \{\sigma_{l_1,1}(j_{11}, \ldots, j_{1l_1}), \ldots \sigma_{l_1,1}(j_{k1}, \ldots, j_{kl_k})\}\},$$

where the sum is taken over all partitions $\{\{j_{11}, \ldots, j_{1l_1}\}, \ldots, \{\{j_{k1}, \ldots, j_{kl_k}\}\}$ of the set $\{2, \ldots, m\}$.

PROOF: If t is a heap labeled tree with m nodes labeled by 2, ..., m, we have a corresponding partition $\{\{j_{11}, \ldots, j_{1l_1}\}, \ldots, \{\{j_{k1}, \ldots, j_{kl_k}\}\}$ of the set $\{2, \ldots, m\}$ gotten by decomposing t with respect to the operation $\odot$ into subtrees whose root has exactly one child, and taking as the subsets in the partition the sets of labels on the subtrees. Summing over all possible heap labeled trees, and partitioning the sum of the decomposed trees into groups of terms corresponding to partitions of the set $\{2, \ldots, m\}$, we get

$$\sigma_{m-1}(2, \ldots, m) = \sum \sigma_{l_1,1}(j_{11}, \ldots, j_{1l_1}) \odot \cdots \odot \sigma_{l_1,1}(j_{k1}, \ldots, j_{kl_k})\},$$

where the sum is taken over all partitions $\{\{j_{11}, \ldots, j_{1l_1}\}, \ldots, \{\{j_{k1}, \ldots, j_{kl_k}\}\}$ of the set $\{2, \ldots, m\}$. We now apply the linear map which sends each heap labeled tree into the heap labeled tree formed by naming the root 1 and making it the unique child of a new root, to get the equation in the lemma.

Lemma 10

$$\sigma_{m,k}(i_1, \ldots, i_m) = \sum \sigma_{l_1,1}(j_{11}, \ldots, j_{1l_1}) \odot \cdots \odot \sigma_{l_k,1}(j_{k1}, \ldots, j_{kl_k}),$$

where the sum ranges over all k-fold partitions $\{\{j_{11}, \ldots, j_{1l_1}\}, \ldots, \{j_{k1}, \ldots, j_{kl_k}\}\}$ of $\{i_1, \ldots, i_m\}$, with $j_{m1} < \cdots < j_{ml_m}$, and $j_{11} < \cdots < j_{k1}$.

PROOF: From Definition 8, we have that $\sigma_{m,k}(i_1, \ldots, i_m)$ is the sum of all heap labeled trees $t(i_1, \ldots, i_m)$ with $m+1$ nodes, named $i_1, \ldots, i_m$, whose root has exactly k children. For each k-fold partition $\{\{j_{11}, \ldots, j_{1l_1}\}, \ldots, \{j_{k1}, \ldots, j_{kl_k}\}\}$ of $\{i_1, \ldots, i_m\}$, group together those trees in $\sigma_{m,k}(i_1, \ldots, i_m)$ in which the elements of members of that partition are used to label the subtrees whose roots are the children of the root of the tree. It is easy to see that as the whole tree ranges over all heap labeled trees whose root has k children, and whose k subtrees are labeled with $\{j_{11}, \ldots, j_{1l_1}\}, \ldots, \{j_{k1}, \ldots, j_{kl_k}\}$, the k subtrees range independently over all trees labeled with $\{j_{11}, \ldots, j_{1l_1}\}, \ldots, \{j_{k1}, \ldots, j_{kl_k}\}$. It follows that the sum of these trees is

$$\sigma_{l_1,1}(j_{11}, \ldots, j_{1l_1}) \odot \cdots \odot \sigma_{l_k,1}(j_{k1}, \ldots, j_{kl_k}).$$

This completes the proof of the lemma.

If t is a heap labeled tree with $m+1$ nodes, named $j_1 < \cdots < j_m$, and if $(E_{\gamma_1}, \ldots, E_{\gamma_m})$ is an ordered sequence of labels, we will sometimes speak of of the labeled heap labeled tree, which we will denote $t(E_{\gamma_1}, \ldots, E_{\gamma_m})$, in which the node named j_k is labeled with E_{γ_k}. (Note that there is no requirement that $E_{\gamma_1}, \ldots, E_{\gamma_m}$ be distinct.) We define $\sigma_m(E_{\gamma_1}, \ldots, E_{\gamma_m})$ and $\sigma_{m,k}(E_{\gamma_1}, \ldots, E_{\gamma_m})$ in the obvious fashion. We will denote the tree with exactly one node, labeled with E_γ, by e_{E_γ}. Lemmas 9 and 10 generalize as follows.

Lemma 11

$$\begin{aligned}&\sigma_{m,1}(E_{\gamma_1}, \ldots, E_{\gamma_m}) = \\ &\qquad \sum e \leftharpoonup \{e_{E_{\gamma_1}} \leftharpoonup \{\sigma_{l_1,1}(E_{\gamma_{j_{11}}}, \ldots, E_{\gamma_{j_{1l_1}}}), \ldots \sigma_{l_k,1}(E_{\gamma_{j_{k1}}}, \ldots, E_{\gamma_{j_{kl_k}}})\}\},\end{aligned}$$

where the sum is taken over all partitions $\{\{j_{11}, \ldots, j_{1l_1}\}, \ldots, \{\{j_{k1}, \ldots, j_{kl_k}\}\}$ *of the set* $\{2, \ldots, m\}$.

Lemma 12

$$\begin{aligned}&\sigma_{m,k}(E_{\gamma_{i_1}}, \ldots, E_{\gamma_{i_m}}) = \\ &\qquad \sum \sigma_{l_1,1}(E_{\gamma_{j_{11}}}, \ldots, E_{\gamma_{j_{1l_1}}}) \odot \cdots \odot \sigma_{l_k,1}(E_{\gamma_{j_{k1}}}, \ldots, E_{\gamma_{j_{kl_k}}}),\end{aligned}$$

where the sum ranges over all k-fold partitions $\{\{j_{11}, \ldots, j_{1l_1}\}, \ldots, \{j_{k1}, \ldots, j_{kl_k}\}\}$ *of* $\{i_1, \ldots, i_m\}$, *with* $j_{m1} < \cdots < j_{ml_m}$, *and* $j_{11} < \cdots < j_{k1}$.

The following lemma gives a specific description of the map ϕ.

Lemma 13

$$\phi\left(E_{\gamma_m} \cdots E_{\gamma_1}\right) = \sigma_m\left(E_{\gamma_1}, \ldots, E_{\gamma_m}\right).$$

Proof: We prove this by induction on m. It is clear for $m = 1$. Assume $m > 1$. Since ϕ is an algebra homomorphism,

$$\begin{aligned}\phi\left(E_{\gamma_m} \cdots E_{\gamma_1}\right) &= \phi(E_{\gamma_m})\phi\left(E_{\gamma_{m-1}} \cdots E_{\gamma_1}\right) \\ &= \phi(E_{\gamma_m})\sigma_{m-1}\left(E_{\gamma_{m-1}}, \ldots, E_{\gamma_1}\right).\end{aligned}$$

Since $\phi(E_{\gamma_m})$ is the tree with two nodes, with the child of the root labeled with E_{γ_m}, the last product is formed by attaching a node labeled E_{γ_m} to every node in every tree in $\sigma_{m-1}\left(E_{\gamma_{m-1}}, \ldots, E_{\gamma_1}\right)$, and summing all the resulting trees. This sum is easily seen to be $\sigma_m\left(E_{\gamma_m}, \ldots, E_{\gamma_1}\right)$. This completes the proof of the lemma.

4 Higher order derivations acting on linear monomials

In this section we begin our study of higher order derivations acting on linear monomials. We are specifically interested in expressions such as

$$\begin{array}{l} E_{\gamma_1}(x_{1r})|_{x=0} \\ E_{\gamma_2}E_{\gamma_1}(x_{2r})|_{x=0} \\ E_{\gamma_3}E_{\gamma_2}E_{\gamma_1}(x_{3r})|_{x=0} \\ E_{\gamma_4}E_{\gamma_3}E_{\gamma_2}E_{\gamma_1}(x_{4r})|_{x=0}. \end{array}$$

Here we use the notation of Example 1. It is easy to see that these expressions may be written in terms of the coefficients that define the derivations $E_1, \ldots, E_M$: the purpose of this section is to derive these formulæ.

The derivation E_γ is of the form

$$E_\gamma = \sum_{\mu=1}^{N} b_\gamma^\mu D_\mu,$$

where $b_\gamma^\mu \in R$. We write

$$b_\gamma^\mu = \sum_{f \geq 0} c_f^\mu(\gamma) x^f,$$

where $c_f^\mu(\gamma) \in \mathbf{R}$. At the end of this section in Example 18 we will give explicit formulæ for the above expressions in terms of the $c_f^\mu(\gamma)$.

We begin by extending the definitions of Section 1 from $\mathbf{R}\{\mathcal{LT}(E_1, \ldots, E_M)\}$ to $\mathbf{R}\{\mathcal{LT}(E_1, \ldots, E_M)\} \otimes V^*$. Let $\overline{\chi}$ denote the evaluation map

$$\overline{\chi} : \mathbf{R}{<}E_1, \ldots, E_M{>} \otimes V^* \longrightarrow \mathbf{R},$$

defined by

$$\overline{\chi}(p \otimes a) = \chi(p) \cdot a|_{x=0},$$

for $p \in \mathbf{R}{<}E_1, \ldots, E_M{>}$, and $a \in V^*$. If W is a vector space over $\mathbf{R}$, we denote by $I_W : W \to W$ the identity map. The map $\overline{\chi}$ is simply the map $\chi \otimes I_{V^*}$, followed by the application map $p \otimes a \mapsto p \cdot a$, and the evaluation map $a \mapsto a|_{x=0}$.

In a similar manner, we define $\overline{\psi}$ to be the map $\psi \otimes I_{V^*}$, followed by the application map and the evaluation map. We can describe $\overline{\psi}$ directly as follows, giving its value on $t \otimes x_\mu$, where $t \in \mathcal{LT}(E_1, \ldots, E_M)$ has $m+1$ nodes, and $x_\mu \in V^*$, and extending it linearly. Suppose that the root of t has one child, named 1. Further suppose that the children of the node named j of the labeled tree t are named $l, \ldots, l'$. Note that if $(\mu_1, \ldots, \mu_r) \rightharpoondown f$, we will sometimes write $c^{\mu}_{\mu_1,\ldots,\mu_r}(\gamma)$ for $c^{\mu}_f(\gamma)$. To each node j of the labeled tree t, we define an element $C(j;t)$ of the field $\mathbf{R}$ as follows

$$C(j;t) = \begin{cases} \epsilon(\mu_l, \ldots, \mu_{l'}) c^{\mu}_{\mu_l,\ldots,\mu_{l'}}(\gamma_1) & \text{for } j = 1; \\ \epsilon(\mu_l, \ldots, \mu_{l'}) c^{\mu_j}_{\mu_l,\ldots,\mu_{l'}}(\gamma_j) & \text{for } j = 2, \ldots, m. \end{cases}$$

It is important to note that $C(j;t)$ depends both on the tree t and also on the indices $\mu_1, \ldots, \mu_m$. We will abbreviate $C(j;t)$ by $C(j)$.

Lemma 14

$$\overline{\psi}(t \otimes x_\mu) = \begin{cases} \sum_{\mu_2,\ldots,\mu_m} C(1) \cdots C(m) & \textit{if the root of } t \textit{ has exactly one child} \\ 0 & \textit{otherwise} \end{cases}$$

Proof: This formula is a simple consequence of the definitions. We compute

$$\begin{aligned} \overline{\psi}(t \otimes x_\mu) &= \psi(t)(x_\mu)|_{x=0} \\ &= \left(\sum_{\mu_1,\ldots,\mu_m=1}^{N} R(m) \cdots R(1) R(0) \right) (x_\mu)|_{x=0}. \end{aligned}$$

Observe first that if the root of t has more than one child then $\overline{\psi}(t \otimes x_\mu) = 0$, since $R(0)$ is a differential operator of degree 2 or more. Suppose then that the root of t has exactly one child. Then $R(0) = D_{\mu_1}$ and $R(0)(x_\mu)|_{x=0}$ is nonzero precisely when $\mu = \mu_1$. Consider next any of the other nodes of the tree t, say the one named j. Assume that j has children named $l, \ldots, l'$ and that j is labeled with E_{γ_j}. Then

$$\begin{aligned} R(j) &= D_{\mu_l} \cdots D_{\mu_{l'}} b^{\mu_j}_{\gamma_j} \\ &= D_{\mu_l} \cdots D_{\mu_{l'}} \sum_{f \geq 0} c^{\mu_j}_f(\gamma_j) x^f|_{x=0} \\ &= \sum_{f \geq 0} c^{\mu_j}_f(\gamma_j) \left(D_{\mu_l} \cdots D_{\mu_{l'}} x^f|_{x=0} \right) \\ &= \epsilon(\mu_l, \ldots, \mu_{l'}) c^{\mu_j}_{\mu_l,\ldots,\mu_{l'}}(\gamma_j), \end{aligned}$$

as required.

Example 15 *See Example 1 for unexplained notation.*

1. *Let t be the labeled tree consisting of the root, and a child of the root labeled E_{γ_1}. Then $t \otimes x_{1r}$ is sent by $\overline{\psi}$ to $c^{1r}_0(\gamma_1)$.*

2. *Let t denote the rooted tree consisting of three nodes such that the root has one child labeled E_{γ_1}, and the child of the root has one child labeled E_{γ_2}. Then $t \otimes x_{2r}$ is sent by $\overline{\psi}$ to*

$$\sum_{k=1}^{d_1} c_{1k}^{2r}(\gamma_1) c_0^{1k}(\gamma_2).$$

3. *Let t denote the rooted tree consisting of four nodes such that the root has one child labeled E_{γ_1}, the child of the root has one child labeled E_{γ_2}, and the grandchild of the root has one child labeled E_{γ_3}. Then $t \otimes x_{3r}$ is sent by $\overline{\psi}$ to*

$$\sum_{\substack{k=1,\ldots,d_2 \\ j=1,\ldots,d_1}} c_{2k}^{3r}(\gamma_1) c_{1j}^{2k}(\gamma_2) c_0^{1j}(\gamma_3).$$

4. *Let t denote the rooted tree consisting of four nodes such that the root has one child labeled E_{γ_1}, and the child of the root has two children labeled E_{γ_2} and E_{γ_3}. Then $t \otimes x_{3r}$ is sent by $\overline{\psi}$ to*

$$\sum_{j,k=1,\ldots,d_1} \epsilon(1j,1k) c_{1k,1j}^{3r}(\gamma_1) c_0^{1k}(\gamma_2) c_0^{1j}(\gamma_3).$$

Lemma 16

$$\overline{\chi} = \overline{\psi} \circ (\phi \otimes I_{V^*}).$$

PROOF: Since

$$\chi = \psi \circ \phi$$

it follows that

$$\chi \otimes I_{V^*} = (\psi \otimes I_{V^*}) \circ (\phi \otimes I_{V^*}).$$

Composing these maps with the map $p \otimes a \mapsto p \cdot a|_{x=0}$ proves the lemma.

We next prove a lemma which simplifies writing

$$E_{\gamma_m} \cdots E_{\gamma_1}(x_\mu)|_{x=0}$$

in terms of the coefficients of the polynomials defining the derivations E_1, $\ldots$, E_M.

Lemma 17 *Let* $E_{\gamma_m} \cdots E_{\gamma_1} \in \mathbf{R}{<}E_1, \ldots, E_M{>}$ *and let* $x_\mu \in V^*$. *Then*

$$E_{\gamma_m} \cdots E_{\gamma_1}(x_\mu)|_{x=0} = \sum \sum_{\mu_2,\ldots,\mu_m} C(1;t) \cdots C(m;t),$$

where the first sum is taken over all labeled heap labeled trees $t = t(E_{\gamma_1}, \ldots, E_{\gamma_m})$ *with* $m+1$ *nodes, whose root has one child.*

Proof: We have

$$\begin{aligned} E_{\gamma_m} \cdots E_{\gamma_1}(x_\mu)|_{x=0} &= \overline{\chi}\,(E_{\gamma_m} \cdots E_{\gamma_1} \otimes x_\mu) \\ &= \overline{\psi} \circ (\phi(E_{\gamma_m} \cdots E_{\gamma_1}) \otimes x_\mu) \\ &= \sum \sum_{\mu_2,\ldots,\mu_m} C(1;t) \cdots C(m;t), \end{aligned}$$

where the first sum in the last member of the equation is taken over all heap labeled trees $t = t(E_{\gamma_1}, \ldots, E_{\gamma_m})$ with $m+1$ nodes, whose root has one child. This completes the proof of the lemma.

Example 18 *See Example 1 for unexplained notation. Using Lemma 17, we can now write down formulæ for higher order derivations acting on linear monomials in terms of the coefficients of the defining derivations.*

$$\begin{aligned} E_{\gamma_1}(x_{1r})|_{x=0} &= c_0^{1r}(\gamma_1). \\ E_{\gamma_2}E_{\gamma_1}(x_{2r})|_{x=0} &= \sum_{k=1}^{d_1} c_{1k}^{2r}(\gamma_1) c_0^{1k}(\gamma_2). \\ E_{\gamma_3}E_{\gamma_2}E_{\gamma_1}(x_{3r})|_{x=0} &= \sum_{\substack{k=1,\ldots,d_2 \\ j=1,\ldots,d_1}} c_{2k}^{3r}(\gamma_1) c_{1j}^{2k}(\gamma_2) c_0^{1j}(\gamma_3) \\ &\quad + \sum_{j,k=1,\ldots,d_1} \epsilon(1k,1j) c_{1k,1j}^{3r}(\gamma_1) c_0^{1k}(\gamma_2) c_0^{1j}(\gamma_3). \end{aligned}$$

Also

$$\begin{aligned} &E_{\gamma_4}E_{\gamma_3}E_{\gamma_2}E_{\gamma_1}(x_{3r})|_{x=0} = \\ &\quad \sum_{\substack{l=1,\ldots,d_3 \\ k=1,\ldots,d_2 \\ j=1,\ldots,d_1}} c_{3l}^{4r}(\gamma_1) c_{2k}^{3l}(\gamma_2) c_{1j}^{2k}(\gamma_3) c_0^{1j}(\gamma_4) \\ &\qquad + \sum_{\substack{l=1,\ldots,d_3 \\ j,k=1,\ldots,d_1}} \epsilon(1j,1k) c_{3l}^{4r}(\gamma_1) c_{1j,1k}^{3l}(\gamma_2) c_0^{1j}(\gamma_3) c_0^{1k}(\gamma_4) \end{aligned}$$

$$+ \sum_{\substack{l=1,\ldots,d_2 \\ j,k=1,\ldots,d_1}} c^{4r}_{2l,1k}(\gamma_1) c^{2l}_{1j}(\gamma_2) c^{1j}_0(\gamma_3) c^{1k}_0(\gamma_4)$$

$$+ \sum_{\substack{l=1,\ldots,d_2 \\ j,k=1,\ldots,d_1}} c^{4r}_{2l,1k}(\gamma_1) c^{2l}_{1j}(\gamma_2) c^{1j}_0(\gamma_4) c^{1k}_0(\gamma_3)$$

$$+ \sum_{\substack{l=1,\ldots,d_2 \\ j,k=1,\ldots,d_1}} c^{4r}_{2l,1k}(\gamma_1) c^{2l}_{1j}(\gamma_3) c^{1j}_0(\gamma_4) c^{1k}_0(\gamma_2)$$

$$+ \sum_{j,k,l=1,\ldots,d_1} \epsilon(1j,1k,1l) c^{4r}_{1j,1k,1l}(\gamma_1) c^{1j}_0(\gamma_2) c^{1k}_0(\gamma_3) c^{1l}_0(\gamma_4)$$

These are examples of the systems of nonlinear algebraic equations that were mentioned in the introduction and at the beginning of this section. In the next section we will see how the solution of these systems can be reduced to solving a sequence of linear equations.

5 The main theorem

In this section we prove a theorem which allows us to compute $E_{\gamma_m} \cdots E_{\gamma_1} \cdot x_\mu|_{x=0}$ in terms of similar expressions involving smaller products of the E_{γ_i}. We first prove a lemma which describes $\overline{\psi}(t \otimes x_\mu)$ in terms of smaller trees. Note that if the E_j are homogeneous of degree 1 and if x_μ has weight m, only trees $t \in \mathcal{LT}(E_1, \ldots, E_M)$ with $m+1$ nodes, such that the root of t has one child, result in $\overline{\psi}(t, x_\mu)$ being nonzero. In fact

Lemma 19 *Assume $t \in \mathcal{LT}(E_1, \ldots, E_M)$ has $m+1$ nodes, and that the root of t has only one child. Moreover, assume that this child has l children, and let $s_1, \ldots, s_l$ denote the subtrees formed by attaching new roots to the l subtrees whose roots are the grandchildren of the root of t. Let $x_\mu \in V^*$. Then*

$$\overline{\psi}(t \otimes x_\mu) = \sum_{\mu_1,\ldots,\mu_l} \epsilon(\mu_1, \ldots, \mu_l) c^{\mu}_{\mu_1,\ldots,\mu_l}(\gamma_1) \overline{\psi}(s_1 \otimes x_{\mu_1}) \cdots \overline{\psi}(s_l \otimes x_{\mu_l}).$$

Proof: We name the nodes of the tree other than the root by $1, \ldots, m$, and to the node i we associate the summation index ν_i. In particular, we assume that the child of the root is 1, and that the grandchildren of the root are $2, \ldots, l+1$.

Assume that the lemma holds for trees $t \in \mathcal{LT}(E_1, \ldots, E_M)$, with fewer than $m+1$ nodes. We compute

$$\overline{\psi}(t \otimes x_\mu) = \sum_{\nu_1,\ldots,\nu_m} C(m) \cdots C(1) D_{\nu_1} x_\mu|_{x=0}.$$

Now

$$\begin{aligned} C(1) &= \sum_f c_f^{\nu_1}(\gamma_1) D_{\nu_2} \cdots D_{\nu_{l+1}} x^f \\ &= \epsilon(\nu_2, \ldots, \nu_{l+1}) c^{\nu_1}_{\nu_2,\ldots,\nu_{l+1}}(\gamma_1) \end{aligned}$$

Let $k, \ldots, k'$ be the nodes in the subtree whose root is the grandchild $j+1$ of the root of t. Recall that s_j is formed from this subtree by attaching it to a new root. Now

$$\overline{\psi}(s_j \otimes x_{\nu_{j+1}}) = \sum_{\nu_k,\ldots,\nu_{k'}} C(k) \cdots C(k')$$

Decomposing $C(2) \cdots C(m)$ into factors corresponding to the subtrees whose roots are the grandchildren of the root of t, and replacing ν_{j+1} by μ_j, we see that the lemma follows.

We now prove our main result:

Theorem 20 *Let* $E_{\gamma_m} \cdots E_{\gamma_1} \in \mathbf{R}{<}E_1, \ldots, E_M{>}$ *be of the form*

$$E_{\gamma_i} = \sum_{\mu,f} c_f^{\mu}(\gamma_i) x^f D_\mu,$$

and let $x_\mu \in V^*$. *Then*

$$\begin{aligned} &E_{\gamma_m} \cdots E_{\gamma_1} \cdot x_\mu|_{x=0} = \\ &\quad \sum \sum_{\mu_1,\ldots,\mu_k} \epsilon(\mu_1, \ldots, \mu_k) c^{\mu}_{\mu_1,\ldots,\mu_k}(\gamma_1) \cdot \\ &\qquad (E_{\gamma_{j_{1l_1}}} \cdots E_{\gamma_{j_{11}}} \cdot x_{\mu_2}|_{x=0}) \cdots (E_{\gamma_{j_{kl_k}}} \cdots E_{\gamma_{j_{k1}}} \cdot x_{\mu_k}|_{x=0}) \end{aligned}$$

where the first sum is taken over all partitions $\{\{j_{11}, \ldots, j_{1l_1}\}, \ldots, \{\{j_{k1}, \ldots, j_{kl_k}\}\}$ *of the set* $\{2, \ldots, m\}$.

Proof: Apply Lemma 16 and Lemma 13 to get

$$\begin{aligned} E_{\gamma_m} \cdots E_{\gamma_1} \cdot x_\mu|_{x=0} &= \overline{\chi}(E_{\gamma_m} \cdots E_{\gamma_1} \otimes x_\mu) \\ &= \overline{\psi}(\phi(E_{\gamma_m} \cdots E_{\gamma_1}) \otimes x_\mu) \\ &= \overline{\psi}(\sigma_m(E_{\gamma_1}, \ldots, E_{\gamma_m}) \otimes x_\mu) \\ &= \overline{\psi}(\sigma_{m,1}(E_{\gamma_1}, \ldots, E_{\gamma_m}) \otimes x_\mu). \end{aligned}$$

Now Lemma 11 implies that this is equal to

$$\sum \overline{\psi}\left(e \leftharpoonup \{e_{E_{\gamma_1}} \leftharpoonup \{\sigma_{l_1,1}(E_{\gamma_{j_{11}}}, \ldots, E_{\gamma_{j_{1l_1}}}), \ldots, \sigma_{l_k,1}(E_{\gamma_{j_{k1}}}, \ldots, E_{\gamma_{j_{kl_k}}})\}\} \otimes x_\mu\right),$$

where the sum is taken over all partitions $\{\{j_{11},\, \ldots,\, j_{1l_1}\},\, \ldots,\, \{\{j_{k1},\, \ldots,\, j_{kl_k}\}\}$ of the set $\{2,\, \ldots,\, m\}$. Now Lemma 19 implies that this is equal to

$$\sum \sum_{\mu_1,\ldots,\mu_k} \epsilon(\mu_1, \ldots, \mu_k) c^{\mu}_{\mu_1,\ldots,\mu_k}(\gamma_1) \overline{\psi}\left(\sigma_{l_1,1}(E_{\gamma_{j_{11}}}, \ldots, E_{\gamma_{j_{1l_1}}}) \otimes x_{\mu_1}\right) \cdots \overline{\psi}\left(\sigma_{l_k,1}(E_{\gamma_{j_{k1}}}, \ldots, E_{\gamma_{j_{kl_k}}}) \otimes x_{\mu_k}\right),$$

which in turn is equal to

$$\sum \sum_{\mu_1,\ldots,\mu_k} \epsilon(\mu_1, \ldots, \mu_k) c^{\mu}_{\mu_1,\ldots,\mu_k}(\gamma_1) \overline{\psi}\left(\sigma_{l_1}(E_{\gamma_{j_{11}}}, \ldots, E_{\gamma_{j_{1l_1}}}) \otimes x_{\mu_1}\right) \cdots \overline{\psi}\left(\sigma_{l_k}(E_{\gamma_{j_{k1}}}, \ldots, E_{\gamma_{j_{kl_k}}}) \otimes x_{\mu_k}\right),$$

where the first sum is taken over all partitions $\{\{j_{11},\, \ldots,\, j_{1l_1}\},\, \ldots,\, \{\{j_{k1},\, \ldots,\, j_{kl_k}\}\}$ of the set $\{2,\, \ldots,\, m\}$. Now Lemma 13 implies that this is equal to

$$\sum \sum_{\mu_1,\ldots,\mu_k} \epsilon(\mu_1, \ldots, \mu_k) c^{\mu}_{\mu_1,\ldots,\mu_k}(\gamma_1) \cdot (E_{\gamma_{j_{1l_1}}} \cdots E_{\gamma_{j_{11}}} \cdot x_{\mu_2}|_{x=0}) \cdots (E_{\gamma_{j_{kl_k}}} \cdots E_{\gamma_{j_{k1}}} \cdot x_{\mu_k}|_{x=0}),$$

where the first sum is taken over all partitions $\{\{j_{11},\, \ldots,\, j_{1l_1}\},\, \ldots,\, \{\{j_{k1},\, \ldots,\, j_{kl_k}\}\}$ of the set $\{2,\, \ldots,\, m\}$. This completes the proof of the theorem.

Example 21 *This is a continuation of the previous examples. Note that the equations given here are all linear in the c's.*

$$\begin{aligned}
E_{\gamma_1}(x_{1r})|_{x=0} &= c^{1r}_0(\gamma_1) \\
E_{\gamma_2}E_{\gamma_1}(x_{2r})|_{x=0} &= \sum_{j=1,\ldots,d_1} c^{2r}_{1j}(\gamma_1) E_{\gamma_2}(x_{1j})|_{x=0} \\
E_{\gamma_3}E_{\gamma_2}E_{\gamma_1}(x_{3r})|_{x=0} &= \sum_{j=1,\ldots,d_2} c^{3r}_{2j}(\gamma_1) E_{\gamma_3}E_{\gamma_2}(x_{2j})|_{x=0} \\
&\quad + \sum_{j,k=1,\ldots,d_1} \epsilon(1j, 1k) c^{3r}_{1j,1k}(\gamma_1) E_{\gamma_2}(x_{1j})|_{x=0} E_{\gamma_3}(x_{1k})|_{x=0}
\end{aligned}$$

Also

$$E_{\gamma_4}E_{\gamma_3}E_{\gamma_2}E_{\gamma_1}(x_{4r})|_{x=0} = \sum_{j=1,\ldots,d_1} c^{4r}_{3j}(\gamma_1) E_{\gamma_3}E_{\gamma_2}E_{\gamma_1}(x_{3j})|_{x=0}$$

$$+ \sum_{\substack{j=1,\ldots,d_1 \\ k=1,\ldots,d_2}} \epsilon(2j,1k) c^{4r}_{2j,1k}(\gamma_1) \, (E_{\gamma_3} E_{\gamma_2}(x_{2k})|_{x=0} E_{\gamma_4}(x_{1j})|_{x=0}$$

$$+ E_{\gamma_4} E_{\gamma_2}(x_{2k})|_{x=0} E_{\gamma_3}(x_{1j})|_{x=0} + E_{\gamma_4} E_{\gamma_3}(x_{2k})|_{x=0} E_{\gamma_2}(x_{1j})|_{x=0})$$

$$+ \sum_{j,k,l=1,\ldots,d_1} \epsilon(1j,1k,1l) c^{4r}_{1j,1k,1l}(\gamma_1) E_{\gamma_2}(x_{1j})|_{x=0} E_{\gamma_3}(x_{1k})|_{x=0} E_{\gamma_4}(x_{1l})|_{x=0}$$

(Note that $\epsilon(2j,1k) = 1$.)

Assume for the moment that the coefficients

$$E_{\gamma_1}(x_{1r})|_{x=0}, \quad \cdots, \quad E_{\gamma_d} \cdots E_{\gamma_1}(x_{dr})|_{x=0}$$

of the equations above are given for all higher order derivations of length less than or equal to d, and that our task is to determine the $c^{\mu_i}_{\mu_l,\ldots,\mu_{l'}}(\gamma)$ so that the equations above hold. The motivation for this was given in the introduction. We now present an algorithm for determining whether or not a solution exists, and for finding the solution if it does exist.

1. Solve the first set of equations for all $c^{1r}_0(\gamma)$. This is always possible.
2. For l equal to 2, …, d, determine whether or not the l^{th} system of linear equations has a solution. If it does, find the solution; if it does not, the original system of equations has no solution.

6 Actions on monomials

In this section we derive a formula which describes the action of higher order derivations on monomials in terms of derivations of lower order acting on linear terms, which we can compute using Theorem 20.

Theorem 22

$$E_{\gamma_m} \cdots E_{\gamma_1} \cdot x_{i_1} \cdots x_{i_k}|_{x=0} = \sum_{\pi} \sum (E_{\gamma_{j_{1l_1}}} \cdots E_{\gamma_{j_{11}}} \cdot x_{\pi(i_1)}|_{x=0}) \cdots (E_{\gamma_{j_{kl_k}}} \cdots E_{\gamma_{j_{k1}}} \cdot x_{\pi(i_k)}|_{x=0}),$$

where the first sum ranges over all permutations π of the multi-set $\{i_1, \ldots, i_k\}$, and the second sum ranges over all k-fold partitions $\{\{j_{11}, \ldots, j_{1l_1}\}, \ldots, \{j_{k1}, \ldots, j_{kl_k}\}\}$ of $\{1, \ldots, m\}$.

PROOF: From the definition of ψ and the fact that the only trees t for which $\psi(t)x_{i_1}\cdots x_{i_k}$ is non zero are those whose root has exactly k children, we have that

$$\begin{aligned} E_{\gamma_m}\cdots E_{\gamma_1}(x_{i_1}\cdots x_{i_k}) &= \sigma_m(E_{\gamma_1},\ldots,E_{\gamma_m})(x_{i_1}\cdots x_{i_k}) \\ &= \sigma_{m,k}(E_{\gamma_1},\ldots,E_{\gamma_m})(x_{i_1}\cdots x_{i_k}). \end{aligned}$$

From Lemma 12, it follows that

$$\sigma_{m,k}(E_{\gamma_1},\ldots,E_{\gamma_m}) = \sum \sigma_{l_1,1}(E_{\gamma_{j_{11}}},\ldots,E_{\gamma_{j_{1l_1}}})\odot\cdots\odot\sigma_{l_k,1}(E_{\gamma_{j_{k1}}},\ldots,E_{\gamma_{j_{kl_k}}}),$$

where the sum ranges over all k-fold partitions $\{\{j_{11},\ldots,j_{1l_1}\},\ldots,\{j_{k1},\ldots,j_{kl_k}\}\}$ of $\{1,\ldots,m\}$.

From the definition of ψ, we have that if $t = t_1\odot\cdots\odot t_k$, where the root of each t_r has exactly one child, then

$$\psi(t)(x_{i_1}\cdots x_{i_k}) = \sum R(m)\cdots R(1)D_{\mu_1}\cdots D_{\mu_k}x_{i_1}\cdots x_{i_k}.$$

Note that

$$D_{\mu_1}\cdots D_{\mu_k}x_{i_1}\cdots x_{i_k} = \sum_{\pi}\delta_{\mu_1,\pi(i_1)}\cdots\delta_{\mu_k,\pi(i_k)},$$

where π ranges over all permutations of the multi-set $\{i_1,\ldots,i_k\}$. Grouping each $\delta_{\mu_r,\pi(i_r)}$ with the $R(s)$ which correspond to nodes in the subtree t_r, we see that

$$\psi(t)x_{i_1}\cdots x_{i_k} = \sum_{\pi}(\psi(t_1)x_{\pi(i_1)})\cdots(\psi(t_k)x_{\pi(i_k)}),$$

where π ranges over all permutations of the multi-set $\{i_1,\ldots,i_k\}$. We sum over all labeled heap labeled trees $t(E_{\gamma_1},\ldots,E_{\gamma_m})$ whose root has exactly k children to get that

$$\begin{aligned} &\sigma_{m,k}(E_{\gamma_1},\ldots,E_{\gamma_m})(x_{i_1}\cdots x_{i_k})|_{x=0} \\ &\quad= \sum\sum_{\pi}(\psi(t_1)x_{\pi(i_1)}\cdots\psi(t_k)x_{\pi(i_k)})|_{x=0} \\ &\quad= \sum_{\pi}\sum(\psi(t_1)x_{\pi(i_1)}\cdots\psi(t_k)x_{\pi(i_k)})|_{x=0} \\ &\quad= \sum_{\pi}\sum(\psi(\sigma_{l_1,1}(E_{\gamma_{j_{11}}},\ldots,E_{\gamma_{j_{1l_1}}}))x_{\pi(i_1)}|_{x=0})\cdots \\ &\qquad\qquad(\psi(\sigma_{l_k,1}(E_{\gamma_{j_{k1}}},\ldots,E_{\gamma_{j_{kl_k}}}))x_{\pi(i_k)}|_{x=0}), \end{aligned}$$

where the second sum in the last member of the equation ranges over all k-fold partitions $\{\{j_{11},\ldots,j_{1l_1}\},\ldots,\{j_{k1},\ldots,j_{kl_k}\}\}$ of $\{1,\ldots,m\}$. The theorem now follows.

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, *The symbolic computation of higher-order derivations: symmetries of expressions and actions of group algebras,* in "Differential geometry: the interface between pure and applied mathematics," ed. M. Luksic, C. Martin, and W. Shadwick, American Mathematical Society, 1987.

[7] R. Grossman and R. G. Larson, *Hopf-algebraic structures of families of trees,* J. Algebra, to appear.

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

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

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

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

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

[13] C. Rockland, *Intrinsic nilpotent approximation,* Acta Applicandæ Math. **8** (1987), 23-270.

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

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