Free systems of vector fields

Lars Hörmander and Anders Melin

In a recent paper Rothschild and Stein [1] have shown how systems of vector fields with commutators of maximal rank can be made free by introduction of auxiliary variables. In this note we shall give a short and elementary proof of this result (Theorem 4) and also of their theorem on approximation of the vector fields so obtained by left invariant vector fields on nilpotent Lie groups (Theorem 5).

Let $X_1, \dots, X_n$ be C^∞ vector fields near 0 in $\mathbf{R}^p$. By ad X we denote the linear operator sending Y to $[X, Y]$ when X and Y are vector fields. For a sequence $I=(i_1, \dots, i_k)$ of $k=|I|$ integers between 1 and n we shall write

$$X_I = X_{i_1} \dots X_{i_k}, \quad X_{[I]} = \operatorname{ad} X_{i1} \dots \operatorname{ad} X_{i_{k-1}} X_{i_k}.$$

Thus $X_{[I]}$ is a vector field and X_I is a differential operator of order $|I|$, $X_{[I]}=X_I=X_{i_1}$ if $|I|=1$. There are automatic relations between the vector fields $X_{[I]}$ such as $\operatorname{ad} X_{i_1} X_{i_2} + \operatorname{ad} X_{i_2} X_{i_1}=0$ and the Jacobi identity. Writing out $X_{[I]}$ explicitly gives for arbitrary vector fields

(1) $$X_{[I]} = \sum A_{IJ} X_J$$

where $A_{IJ}=0$ when $|I| \neq |J|$ and $A_{IJ}=\delta_{IJ}$ when $|I|=|J|=1$. It follows that for arbitrary vector fields X_j

$$\textstyle\sum_I a_I A_{IJ} = 0 \quad \text{for all} \quad J \Rightarrow \sum a_I X_{[I]} = 0.$$

Definition 1. $X_1, \dots, X_n$ are called *free or order s at* 0 if

(2) $$\textstyle\sum_{|I| \leqq s} a_I X_{[I]}(0) = 0 \Rightarrow \sum_{|I| \leqq s} a_I A_{IJ} = 0, \quad |J| \leqq s.$$

The following proposition is essentially contained in Witt's theorem [2] but we give a direct elementary proof.

Proposition 2. *$X_1, \dots, X_n$ are free of order s at* 0 *if and only if for arbitrary c_I, $|I| \leqq s$, it is possible to find $u \in C^\infty$ satisfying*

(3) $$X_I u(0) = c_I, \quad |I| \leqq s.$$

Proof. a) Assume that (3) can always be solved. If $\sum_{|I|\leqq s} a_I X_{[I]}(0)=0$ then (1) and (3) give $\sum_{|I|\leqq s} a_I A_{IJ} c_J=0$ for arbitrary c_J, hence $\sum a_I A_{IJ}=0$, $|J|\leqq s$.
(b) Assume that $X_1, \dots, X_n$ are free of order s. By induction with respect to j, $1\leqq j\leqq s$, we shall prove that one can find u such that

$$X_{[I_1]} \dots X_{[I_\nu]} u(0) = \sum A_{I_1 J_1} \dots A_{I_\nu J_\nu} c_{J_1 \dots J_\nu}, \quad \text{if} \quad \nu \leqq j, \quad |I_1|+\dots+|I_\nu| \leqq s. \tag{4}$$

When $\nu=s$ this is the same as (3). For $j=1$, thus $\nu=1$, the equations (4) mean that

$$\sum_{|I|\leqq s} a_I X_{[I]} u(0) = \sum a_I A_{IJ} c_J.$$

Since $X_1, \dots, X_n$ are free of order s, a linear form is uniquely defined by

$$\sum_{|I|\leqq s} a_I X_{[I]}(0) \to \sum a_I A_{IJ} c_J$$

on a subspace of the tangent space at 0. If we let $du(0)$ be an extension to all of $\mathbf{R}^p$, the assertion is proved when $j=1$, so we may assume $j>1$ and that there is a solution u_0 of (4) with j replaced by $j-1$. Set $u=u_0+v$ where v vanishes of order j at 0. Then the equations (4) with $\nu<j$ are fulfilled. With $p=v^{(j)}(0)$, which may be any symmetric j linear form, the remaining equations (4) are

$$\begin{aligned} p(X_{[I_1]}, \dots, X_{[I_j]}) &= \sum A_{I_1 J_1} \dots A_{I_j J_j} c_{J_1 \dots J_j} \\ - X_{[I_1]} \dots X_{[I_j]} u_0(0) &= d_{I_1 \dots I_j}; \quad |I_1|+\dots+|I_j| \leqq s. \end{aligned} \tag{4$'$}$$

By the Jacobi identity $\operatorname{ad} X_{[I]}=(\operatorname{ad} X)_{[I]}$ so a commutator $[X_{[I']}, X_{[I'']}]$ is a linear combination of commutators of length $|I'|+|I''|$. It is therefore clear that $d_{I_1 \dots I_j}$ is symmetric in the indices. Choose a minimal set B of sequences I with $|I|\leqq s$ such that $\{X_{[I]}(0)\}_{I\in B}$ span the same space at 0 as all $X_{[I]}$ with $|I|\leqq s$. When $|I_1|\leqq s$ we can write

$$X_{[I_1]}(0) = \sum_{I\in B} a_I X_{[I]}(0)$$

with $|I|=|I_1|$ in the sum, and this implies $A_{I_1 J}=\sum_B a_I A_{IJ}$, $|J|\leqq s$, since X_i are free of order s. Hence it suffices to satisfy (4)$'$ when $I_1\in B$, and similarly we may assume $I_2, \dots, I_j\in B$. But in a basis containing $X_{[I]}(0)$, $I\in B$, this means just that some coefficients of the multilinear form p are given in a symmetric way, so the existence of v is obvious.

Proposition 3. *Suppose that $X_1, \dots, X_n$ are free of order $s-1$ but not of order s at 0. Then one can find vector fields $\tilde{X}_j$ in $\mathbf{R}^{p+1}$ of the form*

$$\tilde{X}_j = X_j + u_j\, \partial/\partial t,$$

where $u_j\in C^\infty(\mathbf{R}^p)$, such that the $\tilde{X}_j$ remain free of order $s-1$ and for every $r\geqq s$ the number of linearly independent vectors $\tilde{X}_{[I]}(0)$ with $|I|\leqq r$ is one unit higher than the number of linearly independent $X_{[I]}(0)$, $|I|\leqq r$.

Proof. Induction with respect to $|I|$ gives for some $u_I \in C^\infty(\mathbf{R}^p)$

$$\tilde{X}_{[I]} = X_{[I]} + u_1\, \partial/\partial t.$$

It follows that the number of linearly independent $\tilde{X}_{[I]}$ with $|I| \leqq k$ is at least as large as the number of linearly independent $X_{[I]}$ with $|I| \leqq k$, and since this is maximal when $k = s-1$, it follows that the $\tilde{X}_i$ are free of order $s-1$. It remains to show that we can choose u_i so that $\partial/\partial t$ is a linear combination of $\tilde{X}_{[I]}(0)$, $|I| \leqq s$. This means that we must find a_I, $|I| \leqq s$, so that

$$\sum a_I X_{[I]}(0) = 0, \quad \sum a_I \tilde{X}_{[I]}(0) \neq 0. \tag{5}$$

By hypothesis one can find a_I with $\sum a_I A_{IJ} \neq 0$ for some J, $|J| \leqq s$, so that the first condition is fulfilled. Now we let

$$\sum a_I \tilde{X}_{[I]} = \sum a_I A_{IJ} \tilde{X}_J$$

operate on the function t, noting that $\tilde{X}_{Jj} t = X_J u_j$. By Proposition 2 we can choose u_j so that $X_J u_j(0)$ have arbitrary values for $|J| < s$. Hence

$$\sum a_I A_{I,Jj} X_J u_j(0) = \sum a_I \tilde{X}_{[I]} t(0)$$

is not 0 for every choice of u_j, which completes the proof.

Theorem 4. *Suppose that $X_1, \ldots, X_n$ are vector fields in $\mathbf{R}^p$ such that for some r the vectors $X_{[I]}(0)$ with $|I| \leqq r$ span $\mathbf{R}^p$. Then there exist an integer m and vector fields $\tilde{X}_k$ in $\mathbf{R}^{p+m}$ of the form*

$$\tilde{X}_k = X_k + \textstyle\sum_1^m u_{kj}(x, t)\, \partial/\partial t_j$$

which are free of order r, such that $\tilde{X}_{[I]}(0)$ span $\mathbf{R}^{p+m}$ when $|I| \leqq r$.

Proof. The hypothesis implies that the dimension p is bounded by the rank of the matrix $A_{IJ}(|I|, |J| \leqq r)$. Ir also implies that the hypothesis of Proposition 3 is fulfilled with $s=1$ at least, unless $X_1, \ldots, X_n$ are already free of order r. It is then possible to lift the vector fields X_j according to Proposition 3 so that the hypotheses of the theorem are fulfilled by the new vector fields. After a finite number of steps we must therefore obtain vector fields which are free of order r.

We shall now examine the properties of the vector fields $\tilde{X}_1, \ldots, \tilde{X}_n$ obtained in Theorem 4. Changing the notations we assume that $X_1, \ldots, X_n$ are now C^∞ vector fields in a neighbourhood of $0 \in \mathbf{R}^p$ which are free of order r and whose commutators of order $\leqq r$ span $\mathbf{R}^p$. Let B be a subset of the set of sequences I of length $\leqq r$ such that the vectors $X_{[I]}(0)$ with $I \in B$ form a basis for $\mathbf{R}^p$. The map

$$\mathbf{R}^B \ni (u_I)_{I \in B} \to \left(\exp \textstyle\sum_B u_I X_{[I]}\right)(0) \in \mathbf{R}^p$$

gives a system of coordinates indexed by B such that

$$\textstyle\sum_B u_I X_{[I]} = \sum_B u_I e_I \tag{6}$$

where $e_I=\partial/\partial u_I$. We assign the weight $|I|$ to the coordinates u_I and $-|I|$ to e_I. Thus a C^∞ function is said to have weight $\geqq s$ at 0 if the Taylor expansion at 0 contains no term $au_{I_1}\dots u_{I_k}$ with $a\neq 0$ and $|I_1|+\dots+|I_k|<s$, and a vector field $Y=\sum_B f_I e_I$ is said to have weight $\geqq s$ is f_I has weight $\geqq s+|I|$ for every $I\in B$. (In [1, p. 272] Y is then said to have local degree $\leqq -s$.) By F^q_s and V^q_s we shall denote respectively the set of C^∞ functions and vector fields such that this is true for all terms in the Taylor expansion of degree $\leqq q$. The subsets of elements vanishing at 0 will be denoted $\mathring{F}^q_s$ and $\mathring{V}^q_s$.

The following theorem implies Theorem 5 of Rothschild—Stein [1] if one takes for Y_i left invariant vector fields from the appropriate nilpotent Lie group.

Theorem 5. *The vector fields X_i, $1\leqq i\leqq n$, have weight -1. If $Y_1,\dots,Y_n$ is another system of vector fields satisfying (6) in a neighbourhood of 0, then X_i-Y_i has weight $\geqq 0$.*

In the proof we need the following lemma.

Lemma 6. *The following inclusions are valid:*

(7) $$F^q_s F^q_t \subset F^q_{s+t}, \quad \mathring{F}^q_s F^{q-1}_t \subset \mathring{F}^q_{s+t},$$

(8) $$F^q_s V^q_t \subset V^q_{s+t}, \quad \mathring{F}^q_s V^{q-1}_t \subset \mathring{V}^q_{s+t}, \quad F^{q-1}_s \mathring{V}^q_t \subset \mathring{V}^q_{s+t},$$

(9) $$V^{q-1}_s(F^q_t) \subset F^{q-1}_{s+t}, \quad \mathring{V}^q_s(F^q_t) \subset \mathring{F}^q_{s+t},$$

(10) $$[V^q_s, V^q_t] \subset V^{q-1}_{s+t}, \quad [\mathring{V}^q_s, V^{q-1}_t] \subset V^{q-1}_{s+t}.$$

Proof. The terms of degree $\leqq q$ in the Taylor expansion of a product fg come from terms in the expansions of f and g of degree $\leqq q$, and if $f(0)=0$ then only terms in g of degree $<q$ contribute. This gives (7) which implies (8). Since $e_I(F^q_t)\subset F^{q-1}_{t-|I|}$ we also obtain (9) which implies (10) since a bracket $[X, Y]$ is formed by letting X operate on the coefficients of Y and Y on the coefficients of X.

Proof of Theorem 5. We shall prove inductively for $q=0, 1, \dots$ that

(11) $$X_{[I]}\in V^q_{-|I|}, \quad X_{[I]}-Y_{[I]}\in V^q_{1-|I|}.$$

Here I is arbitrary, but (11) is obviously valid if $|I|>r$, since any vector field has weight $\geqq -r$. Moreover, the vectors $X_{[J]}(0)$ with $J\in B$ form a basis for $\mathbf{R}^p$ so we have for any I with $|I|\leqq r$

$$X_{[I]}(0) = \sum_{J\in B} c_{IJ} X_{[J]}(0).$$

Since $X_1, \dots, X_n$ are free of order r we may assume that $|J|=|I|$ in the sum and conclude that the same equation is valid everywhere for any vector field, in particular for X or Y. Thus (11) follows for all I if it is valid when $I\in B$.

If we multiply the identity (6) and the corresponding equation for $Y_{[I]}$ by ad e_I we obtain the equations

$$e_I = X_{[I]} + \sum_{K\in B} u_K \operatorname{ad} e_I X_{[K]} = Y_{[I]} + \sum_{K\in B} u_K \operatorname{ad} e_I Y_{[K]}, \quad I\in B. \tag{12}$$

In particular we have

$$X_{[I]}(0) = Y_{[I]}(0) = e_I, \quad I\in B,$$

which proves (11) when $q=0$. In what follows we assume that (11) is proved for a certain $q\geqq 0$ and want to prove (11) with q replaced by $q+1$. To do so it is in view of (12) and (8) sufficient to prove that for arbitrary $I, J\in B$

$$W = \operatorname{ad} X_{[J]} e_I \in V^q_{-(|I|+|J|)}, \quad Z = \operatorname{ad}(X_{[J]} - Y_{[J]}) e_I \in V^q_{1-(|I|+|J|)}. \tag{13}$$

If the first equation (12) is multiplied by ad $X_{[J]}$ we obtain

$$W = \operatorname{ad} X_{[J]} X_{[I]} + \sum_{K\in B} u_K \operatorname{ad} X_{[J]} \operatorname{ad} e_I X_{[K]} + \sum_{K\in B} X_{[J]}(u_K) \operatorname{ad} e_I X_{[K]}, \tag{14}$$

and (12) gives

$$X_{[J]}(u_K) - \delta_{JK} = -\sum_{L\in B} u_L \operatorname{ad} e_J X_{[L]}(u_K). \tag{15}$$

Now $u_L \operatorname{ad} e_J X_{[L]} \in \overset{0}{V}{}^q_{-|J|}$ by (10) and (8) so the right hand side of (15) is in $\overset{0}{F}{}^q_{|K|-|J|}$ by (9). Since $\operatorname{ad} e_I X_{[K]} \in V^{q-1}_{-(|I|+|K|)}$ it follows that the second sum in (14) is congruent to $\operatorname{ad} e_I X_{[J]} = -W$ modulo $V^q_{-(|I|+|J|)}$. By the Jacobi identity and the induction hypothesis we have $\operatorname{ad} X_{[J]} X_{[I]} \in V^q_{-(|I|+|J|)}$ which proves that

$$W \equiv -W + \sum_{K\in B} u_K \operatorname{ad} X_{[J]} \operatorname{ad} e_I X_{[K]} \quad \text{modulo } V^q_{-(|I|+|J|)}. \tag{16}$$

The term in the sum can be rewritten as follows

$$u_K \operatorname{ad} X_{[J]} \operatorname{ad} e_I X_{[K]} = -u_K \operatorname{ad} X_{[J]} \operatorname{ad} X_{[K]} e_I$$
$$= -u_K \operatorname{ad} X_{[K]} \operatorname{ad} X_{[J]} e_I - u_K \operatorname{ad}([X_{[J]}, X_{[K]}]) e_I.$$

The last term is in $V^q_{-(|I|+|J|)}$ by the Jacobi identity (10) and (8) so (16) gives

$$2W \equiv -\sum u_K \operatorname{ad} X_{[K]} W \quad \text{modulo } V^q_{-(|I|+|J|)}. \tag{17}$$

Using (6) we can replace $X_{[K]}$ by e_K here, for

$$\sum u_K \operatorname{ad} X_{[K]} W = \sum \operatorname{ad}(u_K X_{[K]}) W + W(u_K) X_{[K]} \tag{18}$$
$$= \sum \operatorname{ad}(u_K e_K) W + W(u_K) X_{[K]} = \sum u_K \operatorname{ad} e_K W + \sum W(u_K)(X_{[K]} - e_K).$$

The last term is in $V^q_{-(|I|+|J|)}$ by (8), since $W(u_K) \in F^{q-1}_{|K|-|I|-|J|}$ and $X_{[K]} - e_K \in \overset{0}{V}{}^q_{-|K|}$. Hence

$$TW \in V^q_{-(|I|+|J|)} \quad \text{if} \quad TW = 2W + \sum u_K \operatorname{ad} e_K W. \tag{19}$$

But T just multiplies terms of degree μ in the Taylor expansion of W by $2+\mu$ so the first part of (13) follows.

To prove the second part of (13) we multiply the equations in (12) by $\operatorname{ad} X_{[J]}$ and $\operatorname{ad} Y_{[J]}$ and subtract. Since $\operatorname{ad} X_{[J]} X_{[I]} - \operatorname{ad} Y_{[J]} Y_{[I]} \in V^q_{1-(|I|+|J|)}$ by the Jacobi identity and the inductive hypothesis, we obtain

$$Z \equiv \sum_{K\in B} \operatorname{ad}(X_{[J]}-Y_{[J]})(u_K \operatorname{ad} e_I X_{[K]}) + \sum_{K\in B} \operatorname{ad} Y_{[J]}(u_K \operatorname{ad} e_I(X_{[K]}-Y_{[K]})). \tag{20}$$

This congruence and the following ones are modulo $V^q_{1-(|I|+|J|)}$. We have already proved that $X_{[K]} \in V^{q+1}_{-|K|}$, hence $\operatorname{ad} e_I X_{[K]} \in V^q_{-(|I|+|K|)}$ by (10) and $u_K \operatorname{ad} e_I X_{[K]} \in \overset{0}{V}{}^{q+1}_{-|I|}$ by (8) so the first sum in (20) is $\equiv 0$ by (10). Since $\operatorname{ad} e_I(X_{[K]}-Y_{[K]}) \in V^{q-1}_{1-|I|-|K|}$ by (10) we have by (8) that $u_K \operatorname{ad} e_I(X_{[K]}-Y_{[K]}) \in V^q_{1-|I|}$. Now $Y_{[J]}-e_J \in \overset{0}{V}{}^{q+1}_{-|J|}$ so (20) gives in view of (10)

$$Z \equiv \sum_{K\in B} \operatorname{ad} e_J(u_K \operatorname{ad} e_I(X_{[K]}-Y_{[K]})) = -\operatorname{ad} e_J(X_{[I]}-Y_{[I]}) \tag{21}$$

where the equality follows from the fact that $[\operatorname{ad} e_I, u_K] = \delta_{IK}$ and that $\sum u_K\,(X_{[K]}-Y_{[K]}) = 0$. Thus.

$$\operatorname{ad}(X_{[J]}-Y_{[J]})e_I \equiv -\operatorname{ad} e_J(X_{[I]}-Y_{[I]}) = \operatorname{ad}(X_{[I]}-Y_{[I]})e_J.$$

If we use this in the sum in (21) we obtain

$$Z \equiv \sum_{K\in B} \operatorname{ad} e_J(u_K \operatorname{ad} e_K(X_{[I]}-Y_{[I]}))$$
$$= \operatorname{ad} e_J(X_{[I]}-Y_{[I]}) + \sum u_K \operatorname{ad} e_K \operatorname{ad} e_J(X_{[I]}-Y_{[I]}),$$

for e_J and e_K commute. We can interchange I and J on the right hand side which gives $TZ \equiv 0$ with the notation in (19), hence $Z \equiv 0$, which completes the proof.

References

1. ROTHSCHILD, L. P. and STEIN, E. M., Hypoelliptic differential operators and nilpotent groups. *Acta Math.* 137 (1976), 247—320.
2. WITT, E., Treue Darstellung Liescher Ringe. *J. Reine Angew. Math.* 177 (1937), 152—160.

Received March 14, 1977; in revised form August 25, 1977.

Lars Hörmander and
Anders Melin
Dept. of Mathematics
Box 725
S-220 07 Lund
Sweden